Climate Migration and Urban Integration in Africa

Mélanie Gittard* and Eleanor Wiseman†

September 2026

Abstract

Does climate-induced migration contribute to the growth of low-quality housing in African cities? We construct bilateral yearly migration flows between districts of eight African countries from harmonized census microdata (1987-2019) and instrument migration with drought intensity at origin, selected by post-IV LASSO among candidate climate variables. Drought-induced migration raises the inflow into low-quality housing: of every hundred migrants that drought pushes into a city, about twenty-three settle in low-quality housing. These migrants are also more housing-vulnerable than others: the marginal drought-induced migrant is about 3 percentage points more likely to live in low-quality housing than the average arrival. Yet they are not economically excluded, but absorbed into employment everywhere. The cost of climate migration is thus residential rather than occupational, and it concentrates in urbanizing cities rather than the largest metropolises. Aggregating over our corridors, drought placed roughly 91,000 additional people in low-quality housing. These results shift the policy focus toward urban housing and toward the fast-urbanizing secondary cities where the pressure concentrates, through both greater absorption capacity and the targeting of the most vulnerable migrants.

Keywords: Climate Change, Migration, Low-Quality Housing, Urbanization, Labor Markets

JEL codes: Q54, R23, O15, O18, J61

*Stanford University, Center on Food Security and the Environment (FSE). mgittard@stanford.edu

†University of Wisconsin-Madison. wiseman3@wisc.edu.

We are grateful to David Lobell, Marshall Burke, Jennifer Burney, Rosamond Naylor, Pascal Polonik, Emmett Reynier, Sam Heft-Neal, Andrew Wilson, Aria Grossman, Lucie Boilet, and our many colleagues at Stanford University. We thank Marlon Seror, Valerie Mueller, Anja Benshaul-Tolonen, Victoire Girard, and Denis Cogneau for valuable feedback, as well as audiences at NOVAFRICA, the EEA, and the ICDE.

1 Introduction

African cities stand at the intersection of two major transitions: rapid urbanization and intensifying climate change and variability. By 2050, an estimated 68% of the continent's population will reside in urban areas, up from 41% in 2020 (Un-Habitat, 2020). This urban transition comes with major deficits in both housing and employment: 62.7% of Sub-Saharan Africa's urban population currently lives in slums (Un-Habitat, 2022), and a substantial share of urban residents remain employed in agriculture rather than in manufacturing or services (Henderson and Kriticos, 2018). Africa appears to be urbanizing without full structural transformation. At the same time, climate variability increasingly affects rural livelihoods: droughts deteriorate agricultural productivity, deplete household incomes, and push rural populations to migrate as a coping strategy (Henderson et al., 2017; Mueller et al., 2020). Yet where these climate-driven migrants settle, under what conditions, and whether they find work at destination remain poorly understood. The question matters because low-quality housing compounds residents' exposure to climate hazards such as flooding and extreme heat (Marx et al., 2019): populations fleeing climate-vulnerable rural areas may end up exposed anew in the city. Whether climate migration contributes to the growth of low-quality housing in African cities, and whether climate migrants are more likely to settle in such housing than other migrants, is therefore a question of first-order policy relevance.

This paper asks how drought-induced migrants are absorbed into African cities, in both housing and employment. For each outcome, we distinguish two channels. The *scale* channel asks whether drought-induced migration raises the inflow into low-quality housing and into employment by increasing the number of arrivals. The *propensity* channel asks whether the migrants that drought sends are more likely to end up in low-quality housing, and differ in their employment, from the migrants who would have moved anyway. Separating the two matters for policy. The scale channel points to a congestion effect from the number of additional migrants, calling for greater urban absorption capacity. The propensity channel asks whether climate migrants are a distinct, more vulnerable population, calling instead for targeted support.

We identify the causal effect of migration on low-quality housing by instrumenting bilateral migration flows with drought at origin, in a yearly panel built from census microdata across eight African countries. The first challenge is measurement: studying migrants' outcomes at destination requires knowing how many migrants move from each origin to each destination in each year - bilateral flows at annual frequency, which existing data for Africa seldom provide. We overcome this by building bilateral migration flows from IPUMS census microdata. We use retrospective questions recording each migrant's administrative unit of origin and year of arrival, letting us reconstruct yearly cohorts of migrants for every origin-destination pair at the admin-2 level. This yields a cohort-bilateral panel of migration flows between districts over 1987-2019, matched at the household level with a multi-criteria measure of housing quality based on the UN-Habitat slum definition, as well as each migrant's employment status. The second challenge is endogeneity: corridors with more migration differ systematically in ways correlated with destination housing. We instrument bilateral migration with drought intensity at origin, selected by a post-IV LASSO procedure among candidate climate variables to remove the discretionary choice of instrument. Drought strongly predicts migration, and the identifying assumption is that it affects destination housing only through migration along the corridor, conditional on corridor and arrival-year fixed effects.

A first result of the paper is the first stage, showing that drought moves people to cities. Where rural incomes depend on rain-fed agriculture, a failed growing season is an income shock, and moving to the city is one of the main margins of adaptation. Comparing the same origin-destination corridor across years, we find that in years when the origin's growing season is unusually dry, more people move to urban areas than in normal years. We show that a one-standard-deviation increase in drought at origin raises the corridor's urban migration rate by about 7% of its mean. The response is specific to the dry tail of the rainfall distribution - symmetric wet shocks do not move people - consistent with an agricultural push rather than generic weather variation. This first stage is the variation on which all subsequent results rest.

The main finding of the paper concerns where these additional migrants end up: the inflow of climate migrants contributes to the growth of low-quality housing in the cities that receive them. A 1 percentage-point (p.p) increase in the drought-induced migration rate raises the inflow into low-quality housing at destination by 0.23 p.p: of every hundred migrants whom drought pushes into a city - people who, absent the dry shock, would have stayed at origin - 23 settle in low-quality housing. In the counterfactual without drought, neither these migrants nor their housing demand would be in the city at all. We also find evidence of selection: drought-induced migrants are more likely to end up in low-quality housing than the migrants who would have moved anyway. Holding the size of the inflow fixed, the drought-induced migrant is about 3 p.p more likely to live in low-quality housing than the average migrant arriving on the same corridor - 49 % do, against 46 % of baseline arrivals.

Decomposing the total effect into these two margins, the scale channel - more migrants arriving, sorting into housing roughly like those already flowing along the corridor - accounts for about 90 % of the additional inflow into low-quality housing, and the propensity channel for the remaining 10 %. A back-of-the-envelope aggregation puts magnitudes on these estimates: over 1981-2018, drought generated roughly 400,000 of the 11.6 million urban moves recorded in our corridors (3.4%), and thereby placed about 91,000 additional people in low-quality housing - roughly 82,000 because drought sent more migrants to the cities, and 9,000 because it shifted the composition of arrivals toward low-quality housing.

The second main finding is that climate migration strains cities residentially rather than economically: drought migrants are absorbed into work, but into low-quality housing. Using the same design on employment outcomes, of every hundred marginal migrants about 34 arrive into employment, 6 into unemployment, and 27 into inactivity (the remainder having undetermined status). Drought-induced migrants are also positively selected into work: relative to the average arrival on the same corridor, the marginal migrant is about 3.6 p.p more likely to be employed - 54 % against 51 % - and correspondingly less likely to be inactive, with unemployment unchanged. Together with the housing results, this characterizes the drought-induced migrant precisely: employed more often than the typical arrival, yet housed worse. The margin of exclusion is residential, not occupational.

The effect operates where the agricultural-push mechanism predicts, and its geography is distinctive. At origin, it is strongest along rural-to-city corridors: drought displaces rural households whose livelihoods depend most directly on rainfall, and these are the migrants who most raise low-quality housing inflows. It also intensifies with the persistence of the shock: instrumenting with a cumulative three-season drought yields a larger effect (0.32 against 0.23), consistent with deeper shocks displacing more housing-vulnerable households, although the difference between the two estimates is imprecise. At destination, the two margins diverge, and this contrast is the key result: urban labor markets absorb drought migrants everywhere, but the sorting into low-quality housing concentrates in urbanizing cities - destinations in the middle of the urban transition - and fades in the largest metropolises. The working poor that drought sends to the cities - migrants who find work but not decent housing - are thus above all a phenomenon of intermediate cities. Further heterogeneity analyses, notably classifying destinations by city size with the Africapolis database, confirm this pattern.

We subject these results to an extensive robustness analysis. The main threat to identification is that drought at origin could affect destination housing through channels other than migration along the corridor; corridor and arrival-year fixed effects absorb the time-invariant and common components, the wet-side placebo rules out channels shared by all rainfall variation, and we discuss the remaining ones - notably food prices at destination and conflict-induced displacement - in Section 6. The estimates are stable across fixed-effects structures and across the alternative deprivation margins of the UN-Habitat definition, including a continuous count of deprived dimensions. Section 6 further reports checks on the instrument itself - alternative drought measures, out-of-sample validation of the LASSO selection - as well as alternative population denominators, an extension to fifteen countries using five-year migration brackets, and a leave-one-country-out exercise.

This paper makes four contributions. First, it provides the first causal evidence linking climate-induced migration to the growth of low-quality housing in Sub-Saharan Africa. A large literature establishes that climate shocks drive migration, from rural-rural moves (Gittard, 2024; McGuirk and Nunn, 2021; Gray and Mueller, 2012) to rural-urban (Henderson et al., 2017; Cattaneo and Peri, 2016; Mueller et al., 2020) and cross-border migration (Beine and Parsons, 2016; Missirian and Schlenker, 2017), while a separate literature debates whether informal settlements are transitional or persistent poverty traps and documents the vulnerability of their residents (Marx et al., 2013, 2019). This paper connects the two: it shifts the focus from out-migration at origin to the housing absorption of climate migrants at destination, and decomposes the effect into a scale and a propensity channel. Second, it characterizes the labor-market integration of these migrants. Whereas existing work emphasizes barriers that keep rural-urban migrants in low-productivity activities (Gollin et al., 2014; Bryan and Morten, 2019; Alvarez, 2020; Hamory et al., 2021) and negative labor-market spillovers of climate migration (Kleemans and Magruder, 2018; Gittard, 2024), we find that drought-induced migrants are absorbed into employment: the margin of exclusion is residential, not occupational, a distinction that matters for targeting policy. Third, it contributes to the literature on African urbanization and structural transformation (Henderson and Kriticos, 2018; McMillan et al., 2014; Jedwab et al., 2017; Gollin et al., 2016) by showing that the housing pressure of climate migration concentrates in urbanizing intermediate cities rather than the largest metropolises, shifting attention from primate cities to the secondary tier where urban growth is fastest and infrastructure thinnest. Fourth, methodologically, it illustrates a disciplined treatment of instrument choice in climate econometrics: a post-IV LASSO selection among candidate climate variables (Belloni et al., 2014), validated by a wet-side placebo, replaces the discretionary choice of climate instrument.

The remainder of the paper is organized as follows. Section 2 describes the data and the construction of the migration, weather, and housing-quality variables. Section 3 presents the identification strategy, including the post-IV LASSO selection of the instrument. Section 4 reports the main results on housing and employment, the scale-propensity decomposition, and the aggregate magnitude. Section 5 examines heterogeneity by destination urbanization. Section 6 presents robustness checks, and Section 8 concludes.

2 Data

This section describes the construction of the weather, migration, urbanization, and housing-quality variables. Appendix Table 7 reports descriptive statistics for all variables used in the analysis, and Appendix Figures 3–5 map the climate exposure of origins and the migration, urbanization, and migrant outcomes of destinations.

2.1 Construction of Weather Variables

Overview - The empirical model selects among a large set of candidate weather variables constructed from raw climate data, using a post-IV LASSO procedure (Belloni et al., 2014). The choice of candidate variables is grounded in recent climate-econ damage function literature (Hultgren et al., 2025; Auffhammer, 2018), which identifies the most influential weather terms for agricultural production. All weather variables are computed on a common historical reference period of 1983–2023. We distinguish temperature and precipitation terms. Weather variables are defined over the growing season (superscript g) for the main analysis, with yearly versions (superscript y) used in robustness checks. Growing seasons are defined at the national level using the main rainy season identified from historical monthly precipitation distributions for each country.

Raw data - Precipitation variables are derived from CHIRPS daily data (Funk et al., 2015). Temperature variables and evapotranspiration are derived from ERA5-Land reanalysis data (Muñoz-Sabater et al., 2021).

Pixel-level versus admin-unit-level transformations - We construct strongly nonlinear transformations (squared precipitation, squared temperature, GDD) at the pixel level before spatial aggregation, following Hsiang (2016), to preserve local nonlinear responses. Threshold-based drought variables (Drought^g_{pJ} and drought_int_z) are constructed using admin-unit-level historical precipitation distributions, consistent with the spatial scale of our migration outcomes.

Precipitation variables - We construct several precipitation candidate variables for the post-LASSO IV. The cumulative precipitation Precip^g_{cum} is the spatial mean of pixel-level monthly cumulative precipitation over the growing season. Its squared version $(\text{Precip}^g_{cum})^2$ is computed at the pixel level before spatial averaging, to capture non-linear responses. The number of dry days N^g_{dry} counts days with daily precipitation below 1 mm at the pixel level and is spatially averaged at the admin-unit level. Two threshold-based drought dummies Drought^g_{p10} and Drought^g_{p15} equal 1 if Precip^g_{cum} falls below the 10th and 15th percentile, respectively, of its historical distribution at the admin-unit level. Finally, the standardized drought intensity $\text{drought\_int\_z}_{o,t} = \max(0, -\tilde{z}_{o,t})$, where $\tilde{z}_{o,t} = (\text{Precip}^g_{o,t} - \overline{\text{Precip}}^g_o)/\sigma^g_{\text{Precip},o}$ is the seasonal precipitation anomaly, captures drought intensity asymmetrically: it activates only for below-average rainfall and measures the number of standard deviations below the historical mean.

For an extreme-precipitation analysis conducted separately, we construct two symmetric extreme-wet variables: ExtremeWet^g_{pJ} is a dummy equal to 1 if Precip^g_{cum} exceeds the 85th or 90th percentile of the historical distribution, and $N^g_{extreme}$ counts days with daily precipitation above 20 mm at the pixel level, spatially averaged at the admin-unit level.

Temperature variables - We construct three temperature candidate variables. The seasonal averages of daily minimum and mean 2m air temperatures, T^g_{min} and T^g_{mean}, are first computed at the pixel level and then spatially averaged. Their squared versions $(T^g_{min})^2$ and $(T^g_{mean})^2$ are computed at the pixel level (after conversion to Celsius) before spatial averaging, to capture non-linear thermal responses. Growing Degree Days GDD^g are computed daily at the pixel level as $\max\left(0, \frac{1}{2}(\tilde{T}_{min,d} + \tilde{T}_{max,d}) - 8\right)$, where $\tilde{T}_{min,d}$ and $\tilde{T}_{max,d}$ are daily minimum and maximum 2m temperatures capped at 30°C. The daily GDD are then summed over the growing season and spatially averaged at the admin-unit level. Extreme Degree Days EDD^g are in preparation for future versions of the candidate set.

Moisture variables - Moisture-related variables include potential evapotranspiration PET^g and actual evapotranspiration ET^g, both extracted from ERA5-Land at the daily level and aggregated over the growing season. PET captures atmospheric moisture demand (irrigation requirements), while ET captures realized evaporation. Vapor pressure deficit (VPD) is work in progress for future versions of the candidate set.

LASSO selection - The candidate set used in the post-IV LASSO selection comprises thirteen contemporaneous weather variables, all measured at the admin-unit level over the growing season - the drought dummies Drought^g_{p10} and Drought^g_{p15}, the standardized drought intensity drought_int_z, N^g_{dry}, the precipitation and temperature levels and polynomials, and the degree-day and evapotranspiration measures described above - augmented with the lags and three-season means of the drought-family variables, for 25 candidates in total (Section 3.3). The procedure selects drought_int_z as the unique relevant instrument for bilateral migration — a result we view as informative in itself, isolating an interpretable and asymmetric drought-intensity measure as the dominant climatic driver. We use this single-variable selection for the main analysis to maintain clarity in interpretation.

Robustness of the selection - The selection is probed in several directions (Section 3.3 and Appendix Figure 7). It is stable across LASSO penalty values, from the theoretical benchmark of Belloni et al. (2014) to more permissive thresholds; among lagged candidates, the three-season mean of drought intensity dominates, consistent with the cumulative instrument used in robustness checks; and a wet-side placebo - re-running the exercise on symmetric excess-rainfall candidates - finds essentially no predictive power, confirming that the instrument's strength is specific to the dry tail of the rainfall distribution. Extensions of the candidate set to the additional non-linear and moisture terms described above, a comparison with Double Machine Learning (Chernozhukov et al., 2018), and an out-of-sample validation across countries following Hultgren et al. (2025) are in progress.

2.2 Migration data

Sources - Both migration and housing quality data come from IPUMS census microdata (Minnesota Population Center, 2019) for fifteen African countries over the period 1987–2019. The sample includes exhaustive population censuses from Benin (2002, 2013), Burkina Faso (1996, 2006), Cameroon (1987, 2005), Egypt (1986, 1996, 2006), Ghana (2000, 2010), Kenya (1989, 1999, 2009, 2019), Malawi (1998, 2008), Mali (1998, 2009), Mauritius (1990, 2000, 2011), Mozambique (1997, 2007, 2017), Senegal (1988, 2002, 2013), Sierra Leone (2015), South Africa (1996, 2001, 2011), Togo (2010), Uganda (2002, 2014), and Zambia (1990, 2000, 2010). These censuses provide harmonized individual-level information on demographic characteristics, migration history (place of birth, year of arrival, previous residence), and housing conditions (dwelling type, access to basic services, urban/rural classification at the enumeration area level). One contribution of this paper is the construction of *bilateral yearly migration flows* at the IPUMS administrative level for 8 countries with available year-of-arrival information, and *five-year migration brackets* for the remaining 7 countries.[1]

Migrant identification - We define an individual i observed at destination d in census year t_c as a migrant from origin o in year t if (i) the IPUMS variable `migyrs1` reports a years-since-arrival value τ such that $t = t_c - \tau$, and (ii) the IPUMS variable for previous administrative unit of residence (or, when unavailable, place of birth) identifies origin o. The destination d is the administrative unit where the individual resides at the census date. For all censuses, we trace migration histories back twelve years prior to each census for the yearly migration analysis, capping the window in multi-census countries so that it does not overlap the preceding census. The individual migration indicator is:

$$\text{InMigrants}_{o,d,t} = \sum_i \mathbf{1}\{i \text{ at } d \text{ in } t_c, \text{ arrived from } o \text{ in } t\} \cdot \text{perwt}_i,$$

where perwt_i is the IPUMS person-weight and the sum runs over all individuals observed at destination d in census t_c.

We exclude from the main analysis origin–destination pairs that record zero migrants in every observed year. These always-zero corridors contribute no within-pair variation to identification under our origin–destination fixed effects specification, and their inclusion would mechanically inflate the sample size without adding identifying information.

[1] In a robustness exercise, we compare our IPUMS-derived migration measures for Kenya with the exhaustive population censuses for 1989, 1999, and 2009 obtained directly from the Kenya National Bureau of Statistics. Results are consistent across both sources.

Urban Migrant - We use the IPUMS variable `urban` to identify whether an individual resides in an urban or rural enumeration area within their destination administrative unit. This classification is finer than the admin unit level: a single administrative unit may contain both urban and rural enumeration areas. Urban migrants are individuals classified as residing in urban enumeration areas at destination (`urban` = 2 in IPUMS).

Population denominators - We construct two destination population measures. The *census-year population* $\text{Pop}^{\text{census}}_{d,T_c}$ is the sum of person-weights observed at destination d in census year T_c. The *urban census-year population* $\text{Pop}^{\text{census,urban}}_{d,T_c}$ restricts to individuals classified as urban residents (`urban` = 2). For robustness, we also reconstruct annual destination populations $\text{Pop}_{d,t}$ backward from the census year via the accounting identity:

$$\text{Pop}_{d,t} = \text{Pop}_{d,T_c} - \text{InMig}_{d,[t+1,T_c]} - \text{Births}_{d,[t+1,T_c]} + \text{OutMig}_{d,[t+1,T_c]} + \text{Deaths}_{d,[t+1,T_c]},$$

where births are estimated from age or birth-year information in the census, and deaths are derived from country-specific mortality rates in the corresponding period. Our main results use the census-year population, which is more stable; the annual reconstruction is used in robustness checks.

Migration rates. - The bilateral migration rate from origin o to destination d in year t is:

$$\text{MigRate}_{o,d,t} = \frac{\text{InMigrants}_{o,d,t}}{\text{Pop}^{\text{census}}_{d,T_c}} \times 100,$$

expressed in percentage points. The urban migration rate restricts the numerator to urban-resident migrants and uses the urban census-year population as denominator:

$$\text{MigRate}^{\text{urban}}_{o,d,t} = \frac{\text{InMigrants}^{\text{urban}}_{o,d,t}}{\text{Pop}^{\text{census,urban}}_{d,T_c}} \times 100.$$

For countries with only five-year migration brackets, we construct the equivalent five-year migration rate as the sum of migrants arriving over $[t-4, t]$, normalized by the same census-year denominator. In robustness checks, we re-estimate the main results using the annual reconstructed destination population $\text{Pop}_{d,t}$ as denominator in place of the census-year population, to verify that our results are not driven by the choice of a fixed denominator.

2.3 Urbanization data

Admin-unit urbanization. - For our heterogeneity and robustness analyses, we classify IPUMS administrative units by their urban population share at the most recent census:

$$\text{UrbanShare}_a = \frac{\sum_{i \in a, T_c} \text{perwt}_i \cdot \mathbf{1}\{\text{urban}_i = 2\}}{\sum_{i \in a, T_c} \text{perwt}_i}.$$

This admin-unit classification is distinct from the individual-level urban variable: a rural admin unit may contain some urban residents, and a city admin unit some rural enumeration areas. For the cities and rural sample restrictions, we label an admin unit as a *city* if $\text{UrbanShare}_a \geq 0.55$ and as *rural* otherwise.[2] For the urbanization heterogeneity analysis, we use three groups: rural-like ($\text{UrbanShare}_a < 0.15$), urbanizing cities ($[0.15, 0.55)$), and main cities ($\geq 0.55$).

We validate this classification against the Africapolis database (OECD/SWAC, 2024), which delineates all urban agglomerations above 10,000 inhabitants in Africa. Matching each destination to its largest agglomeration confirms that the classification orders city sizes correctly: the median largest agglomeration rises from about 47,000 inhabitants in rural-like destinations to 86,000 in urbanizing cities and 880,000 in main cities, and the share of destinations hosting a city of at least 100,000 inhabitants rises from 28% to 43% to 70%. The main-cities group contains the principal cities of each country (Nairobi, Kampala, Yaoundé, Bamako), the urbanizing-cities group classic secondary cities (Nakuru, Garoua, Fort Portal), and the rural-like group mostly small towns. Appendix 9.1.3 reports the full correspondence (Figure 6 and Table 9).

2.4 Housing Quality

Individual definition of housing quality. - The household-level low-quality housing classification follows the UN-Habitat multi-criteria slum definition (UN-Habitat, 2003).[3] The classification relies on IPUMS variables covering six housing dimensions grouped in two categories. Structural quality of the dwelling captures the materials of the floor, walls, and roof. Access to basic services and living conditions captures access to an improved water source, sanitation, and household crowding (persons per room). In the main definition, metal-sheet roofing is classified as durable, consistent with the UN-Habitat durability criterion that durable materials protect occupants from climatic conditions; the robustness definition instead counts metal-sheet roofing as deficient, reflecting its use as a low-cost material in the Sub-Saharan African urban context.[4]

A household is classified as residing in low-quality housing if at least one of these dimensions is deficient, following the UN-Habitat operational slum definition, under which a household lacking any one of the deprivation criteria is classified as slum (a logical OR across dimensions) (UN-Habitat, 2003). As a conservative robustness definition, we also use a stricter threshold requiring at least two deficient dimensions. The coding of each dimension - which materials, services, and crowding levels are classified as deficient, in harmonized

[2]The 55% threshold is calibrated on the median urban share across our sample of African administrative units; the rolling-window estimates in Figure 1 trace the estimates continuously along the urban-share distribution, showing that no single threshold drives the results.

[3]The census-based classification captures the material dimension of housing precarity - structural quality and access to basic services - but not the legal or tenure dimension that defines informal settlements in a stricter sense (unauthorized land occupation or unregulated construction). Throughout the paper, low-quality housing refers to the UN-Habitat statistical slum definition rather than to the legal sense of informal settlement.

[4]The UN-Habitat definition comprises five deprivations (durable housing, sufficient living area, improved water, improved sanitation, and secure tenure); we adapt it to the dimensions consistently available in harmonized census microdata, and cannot capture tenure security or, in some countries, sanitation. Our measure therefore reflects the material and service dimensions of housing precarity rather than the full UN-Habitat set.

IPUMS categories identical across countries, with dimension availability varying by census - is reported in Appendix Table 8.

From the household-level classification, two individual-level housing-quality measures are constructed: the binary indicator $\text{deprived}_i \in \{0,1\}$ equal to 1 if individual i resides in a household classified as low-quality, and the continuous count $\text{deprived\_count}_i \in \{0,1,...,6\}$ equal to the number of deficient housing dimensions for individual i's household.

Aggregate housing quality outcomes. - The main analysis focuses on origin-destination-year cohorts of migrants. From the individual-level housing-quality classification above, two aggregate outcomes are constructed at the corridor level. The first outcome, used in the main results to preserve the same estimation sample between first and second stages, is the *scale* outcome $S_{o,d,t}$, the low-quality housing inflow rate at destination:

$$S_{o,d,t} = \frac{\sum_i \text{perwt}_i \cdot \mathbf{1}\{i \text{ at } d \text{ in } t_c,\ \text{arrived from } o \text{ in } t, \text{deprived}_i = 1\}}{\text{Pop}^{\text{census}}_{d,T_c}} \times 100,$$

the share of destination population accounted for by migrants in low-quality housing arriving from origin o in year t. The second outcome, designed to isolate the propensity channel, is the *propensity* outcome $P_{o,d,t}$, the share of migrants residing in low-quality housing in each corridor:

$$P_{o,d,t} = \frac{\sum_i \text{perwt}_i \cdot \mathbf{1}\{i \text{ at } d \text{ in } t_c,\ \text{arrived from } o \text{ in } t, \text{deprived}_i = 1\}}{\sum_i \text{perwt}_i \cdot \mathbf{1}\{i \text{ at } d \text{ in } t_c,\ \text{arrived from } o \text{ in } t\}}.$$

This second outcome is defined only on cells with at least one migrant, which yields a smaller estimation sample than $S_{o,d,t}$. Analogous urban-restricted versions, $S^{\text{urb}}_{o,d,t}$ and $P^{\text{urb}}_{o,d,t}$, restrict the numerator to migrants residing in urban enumeration areas at destination and use the urban census-year population as denominator.

Decomposing scale and propensity. - The scale outcome embeds two distinct mechanisms by construction. The accounting identity

$$S_{o,d,t} = \frac{\text{MigRate}_{o,d,t}}{100} \times P_{o,d,t}$$

makes explicit that a causal effect of drought-induced migration on the low-quality housing inflow can operate through two channels. The *scale channel* reflects the mechanical contribution of drought-induced migrants arriving at the baseline propensity to reside in low-quality housing. The *propensity channel* reflects a shift in that propensity among migrants themselves: drought-induced migrants may be more or less likely to settle in low-quality housing than baseline migrants on the same corridor. Supplementary Appendix Figure S3 illustrates the distinction with a stylized example.

A first-order decomposition of the 2SLS coefficient on the scale outcome yields:

$$\widehat{\beta}^{2SLS}_S \approx \underbrace{\overline{P}}_{\beta^{\text{scale}}} + \underbrace{\frac{\overline{\text{MigRate}}}{100} \times \widehat{\beta}^{2SLS}_P}_{\beta^{\text{propensity}}},$$

where β^{scale} is the effect that would obtain if the propensity were fixed at its sample mean, and $\beta^{\text{propensity}}$ captures the additional contribution from a shift in migrants' propensity.

Estimating the propensity channel directly requires a separate 2SLS specification with $P_{o,d,t}$ as outcome. This approach faces three identification challenges. First, *ratio bias*: the endogenous migration rate enters the denominator of the propensity outcome, so variation in $P_{o,d,t}$ partly reflects mechanical denominator shifts rather than changes in migrant composition. Second, *measurement noise*: $P_{o,d,t}$ is noisily measured on cells with few migrants; this is addressed by weighting regressions by the time-mean of migrant count per corridor (Section 4). Third, *interpretive ambiguity*: a positive coefficient on $P_{o,d,t}$ cannot be cleanly separated between sorting of migrants into low-quality housing and the mechanical denominator effect just described.

2.5 Other Variables

Aridity index for heterogeneity analyses - For heterogeneity analyses by climate baseline, we use the CGIAR-CSI Global Aridity Index (Trabucco and Zomer, 2019), computed as the long-run mean ratio of precipitation to potential evapotranspiration. This index is extracted from the CGIAR-CSI global 30-arcsec dataset and spatially averaged at the IPUMS administrative level. Lower values indicate more arid conditions; we use within-country medians to classify origins as more arid or less arid, avoiding conflation between aridity and country fixed effects.

2.6 Descriptive Overview

Migration patterns - Appendix Figure 2 displays the average urban in-migration rate over time in the eight countries of the estimation sample, constructed from the yearly bilateral panel. Rates differ by an order of magnitude across countries, and within each census window they rise toward the census year - the mechanical recall and onward-migration attrition of earlier cohorts that the distance-to-census control absorbs in estimation (Section 3). The corresponding five-year-bracket rates for the full set of fifteen IPUMS countries are shown in Supplementary Appendix Figure S1.[5]

Spatial variation in climate and migration - Beyond temporal patterns, the data reveal substantial spatial heterogeneity within and across countries. Appendix Figure 3 maps mean growing-season precipitation and mean drought intensity by administrative unit, illustrating the diversity of agricultural climates spanned by the sample - from the arid regions of northern Mali and Kenya (below 500 mm per season) to the humid coastal zones of southern Cameroon (exceeding 2,500 mm). Appendix Figure 4 maps urban in-migration rates at the administrative unit level, revealing considerable within-country variation in destination patterns, with concentrated flows toward major coastal and capital cities. This spatial granularity allows us to construct origin-destination migration matrices at a fine geographic scale and to match each origin district with its specific climate exposure.

[5] A potential concern with retrospective migration data is recall bias: individuals may report their year of arrival with diminishing accuracy as the gap between migration year t and census year t_c widens. All specifications include a linear control for the survey-to-arrival gap $|t - t_c|$, absorbing systematic differences in measurement quality across migration cohorts.

Low-quality housing prevalence - Low-quality housing is widespread among the urban in-migrants we study: across the estimation panel, 46 percent of urban in-migrants reside in housing classified as low-quality under the main UN-Habitat definition (Appendix Table 7), a magnitude consistent with the prevalence UN-Habitat reports for Sub-Saharan Africa's urban population (Un-Habitat, 2022). Appendix Figure 5 maps this share by destination, showing pronounced variation both across and within countries. Cross-country levels partly reflect census vintage - Zambia's shares, the highest in the sample, are measured in its 1990 and 2000 censuses, the earliest we observe - and are absorbed by the corridor fixed effects in estimation; as a further check, Section 6 re-estimates the main result leaving out one country at a time, verifying that no single country drives the effect.

3 Identification Strategy

3.1 Empirical Model

This paper investigates whether weather-shock-induced migration increases migrant inflows into low-quality housing in African cities and identifies the channels through which this effect operates. Two related questions structure the analysis. First, does drought-induced migration increase the number of migrants arriving in deprived housing, with more migrants arriving and sorting into housing at the usual rate (the *scale* channel)? Second, are drought-induced migrants more likely than other migrants to settle in low-quality housing at destination (the *propensity* channel)? We then use the same design to characterize the labor-market integration of these marginal migrants (Section 4).

To address these questions, we implement an instrumental-variable strategy using the cohort-bilateral panel of migration flows (o, d, t). The first stage isolates variation in bilateral migration generated by exogenous weather shocks at origin:

$$\text{MigRate}_{o,d,t} = \alpha_1 Z_{o,t} + \gamma_1 |t - t_c| + \delta_{o,d} + \lambda_t + \varepsilon_{o,d,t}, \tag{1}$$

where $Z_{o,t}$ denotes the weather shock at origin o in year t, selected from a set of candidate climatic variables using post-IV LASSO (Section 3.3). The selected instrument is standardized drought intensity,

$$Z_{o,t} = \max\left(0, -\widetilde{z}_{o,t}\right),$$

as defined in Section 2.1. Origin-destination pair fixed effects, $\delta_{o,d}$, absorb all time-invariant corridor characteristics, so identification relies on within-corridor variation in the timing of droughts. Arrival-year fixed effects, λ_t, absorb shocks common to all corridors, while $|t - t_c|$ controls linearly for the distance between migration year t and the reference census year t_c. Standard errors are clustered at the origin administrative-unit level, where the instrument varies.

The second stage estimates the causal effect of bilateral migration on migrant inflows into low-quality housing at destination:

$$S_{o,d,t} = \beta_1 \widehat{\text{MigRate}}_{o,d,t} + \gamma_2 |t - t_c| + \mu_{o,d} + \lambda_t + \nu_{o,d,t}, \tag{2}$$

where $\widehat{\text{MigRate}}_{o,d,t}$ is the fitted migration rate from Equation 1, and the fixed effects and controls are defined as above. The scale outcome $S_{o,d,t}$, the inflow of migrants into low-quality housing per 100 destination residents, is defined as

$$S_{o,d,t} = \frac{\sum_{i \in \text{Migrants}(o,d,t)} \text{perwt}_i \mathbf{1}\{\text{deprived}_i = 1\}}{\text{Pop}^{\text{census}}_{d,t_c}} \times 100.$$

Because both the outcome and the migration rate are expressed per 100 residents at destination, β_1 measures the share of the marginal drought-induced migrant inflow that settles in low-quality housing. This specification identifies the *scale* channel.

We isolate the *propensity* channel using an analogous specification in which the outcome is the propensity outcome $P_{o,d,t}$, the share of arriving migrants living in deprived housing. A positive coefficient indicates that migrants induced to move by drought are more likely to live in low-quality housing than the average migrant (Section 4). We also estimate urban-restricted versions of both specifications using the urban migration rate, $\text{MigRate}^{\text{urb}}_{o,d,t}$, and the urban scale outcome, $S^{\text{urb}}_{o,d,t}$.

The distance-to-census control, $|t - t_c|$, addresses three sources of measurement error in retrospective migration data. First, recall error may increase with the time elapsed between migration and the census. Second, recent arrival cohorts may be more fully captured than earlier cohorts. Third, the census records migrants at their place of residence at enumeration, so members of earlier cohorts who subsequently moved to another destination or returned to their origin are no longer observed in the original corridor. Migrant counts are corrected for mortality attrition using crude death rates when constructing the panel (Section 2), while $|t - t_c|$ controls for remaining attrition that varies systematically with the distance to the census. Because these sources of measurement error affect both the migration rate and the outcome, the same control enters both stages.

3.2 Discussion

Interpretation of the IV estimate - The 2SLS estimate β_1 in equation (2) identifies a Local Average Treatment Effect: the causal effect of bilateral migration on the low-quality housing inflow for the subpopulation of drought-induced migrants - the marginal migrants whose decision to move from origin o in year t responds to drought at origin, the compliers. The estimate does not identify the average treatment effect across all migrants, but the effect on the climate-responsive margin.

Discussion of the compliers - In a standard application, these compliers are a technical by-product of the instrument, and the LATE is often viewed as a limitation relative to an average effect over all migrants. Here the interpretation is more direct: the population the instrument selects - households induced to migrate by drought - *is* the population of interest, since the paper asks how climate shocks reshape urban housing through migration. The LATE is therefore the estimand of interest rather than a nuisance parameter to be reconciled with a broader average. This alignment also gives the propensity results their meaning: the same drought variation that identifies the causal effect of migration on housing also characterizes the housing outcomes of climate migrants, because these are one and the same group.

Validity of the instrument and exclusion restriction - Identification rests on two key assumptions. The relevance condition requires that drought intensity predicts bilateral migration; the strong first-stage relationships reported in Section 3.4 ($F > 20$) confirm this. The exclusion restriction requires that drought at origin o affects the low-quality housing inflow at destination d only through its effect on migration from o to d, conditional on origin-destination pair and arrival-year fixed effects. This assumption is plausible in our setting: drought is a local agricultural shock at origin, and the origin-destination fixed effects absorb any time-invariant correlation between origin climate and destination housing conditions. Potential threats to this assumption - including direct effects of origin drought on destination outcomes via food prices or return migration -are discussed in Section 3.4.

3.3 Instrument Selection: Post-IV LASSO

The choice of the instrument $Z_{o,t}$ is validated by a data-driven selection procedure based on the post-IV LASSO of Belloni et al. (2014). The procedure runs two LASSO regressions on a candidate set of climate variables at origin - one for the endogenous migration rate and one for the outcome, both on Frisch-Waugh-Lovell residuals that partial out the origin-destination pair and arrival-year fixed effects and the distance-to-census control - and retains the union of the selected sets. This removes the discretionary choice of instrument among ex-ante plausible candidates and provides valid post-selection inference under the conditions of Belloni et al. (2014).

The candidate set comprises thirteen contemporaneous climate variables described in Section 2.1 - drought intensity, drought dummies at two thresholds, dry days, precipitation and temperature polynomials, degree-day and evapotranspiration measures - augmented with the lags and 3-season means of the drought-family variables (25 candidates in total). We implement the selection under three values of the penalty multiplier c, from the theoretical benchmark of Belloni et al. (2014) ($c = 1.1$) to more permissive thresholds ($c = 0.5$, 0.25) that let more instruments enter.

The selection confirms the paper's instrument. Drought intensity is the strongest single candidate on both samples ($F = 23.9$ on the scale sample, $F = 13.0$ on the propensity sample) and is selected at the default penalty in both; among lagged candidates its 3-season mean dominates ($F = 27.4$), consistent with the cumulative instrument used throughout. We use the single strongest candidate, keeping the model just-identified so that Anderson-Rubin inference remains exact. As a placebo, we repeat the exercise on wet-side mirror candidates: excess-rainfall intensity has essentially no predictive power ($F = 4.0$ and $F = 0.03$), and the flood dummies are similarly weak. The instrument's strength is thus specific to the dry tail of the rainfall distribution. Appendix 9.2.1 reports the full selection results (Figure 7).

3.4 First Stage

Table 1 reports the first stage under alternative sets of fixed effects and controls, for the sample of urban migrations to all destinations. The table shows the strength and the stability of the instrument for in-migration towards urban areas. Column (1) is the paper's main specification: origin-destination pair and arrival-year fixed effects with the distance-to-census control. Comparing columns (1) and (2) shows that this control is not driving the first

Table 1: First Stage: Drought Intensity and Urban Migration

	DV: Urban migration rate					
	(1) Main spec.	(2)	(3)	(4)	(5)	(6)
Drought intensity	0.00373***	0.00256***	0.00400***	0.00405***	0.00406***	0.00191**
	(0.00076)	(0.00078)	(0.00089)	(0.00083)	(0.00089)	(0.00089)
First-stage F	23.86	10.87	20.36	23.88	20.61	4.62
Effect of 1-SD drought (% mean)	+6.9%	+4.7%	+7.4%	+7.5%	+7.5%	+3.5%
Origin FE	No	No	No	No	No	No
Destination FE	No	No	No	No	No	No
Origin-Destination FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country × Year FE	No	No	No	No	No	No
Destination × Year FE	No	No	No	No	No	Yes
Dist. to census control	Yes	No	Yes	Yes	Yes	No[a]
Destination controls	No	No	No	Yes	Yes	No
Origin controls	No	No	Yes	No	Yes	No
R^2	0.359	0.355	0.359	0.359	0.359	0.384
Within R^2	0.006	0.000	0.006	0.006	0.006	0.000
Mean migration rate	0.027168	0.027168	0.027168	0.027168	0.027168	0.027168
Observations	245,292	245,292	245,292	245,292	245,292	245,292
OD pairs	16,888	16,888	16,888	16,888	16,888	16,888

Notes: Sample: urban migrants, all destinations; 8 countries, 12-year window. Instrument: contemporaneous drought intensity at origin, $\max(0, -\tilde{z})$. Column (1) is the paper's main specification; fixed effects and controls as indicated ([a]`dist_census` mechanically absorbed). First-stage F: cluster-robust Wald statistic on the instrument. Effect of 1-SD drought: coefficient × SD(instrument) ÷ mean migration rate × 100, on each column's own sample. Standard errors, clustered at `admin_origin`, in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

stage: the raw specification in column (2), without any control, already delivers a positive and highly significant coefficient, and adding the distance to the census mainly improves precision. Column (3) confirms that the instrument captures rainfall deficits rather than heat, and column (4) adds temperature at destination to absorb a potential climatic pull factor; column (5) includes all controls jointly. In all three cases the coefficient is essentially unchanged and the first stage remains strong. When accounting for destination-by-year fixed effects (column (6)), the coefficient remains positive and significant at the 5% level, but the first-stage F drops below conventional thresholds, as these fixed effects absorb a large share of the identifying variation, rainfall shocks being spatially correlated across origins and destinations within a country.

Quantitatively, the main specification in column (1) implies that a one standard deviation increase in the dryness of the growing season at origin raises in-migration to urban areas by 6.9% of the mean migration rate. Let's consider a destination of 10,000 inhabitants. At the mean migration rate, this corridor receives 2.72 migrants in a normal year. A one standard deviation drier growing season raises the migration rate by $0.00373 \times \text{SD(drought)}$, which scaled by the destination population adds about 0.19 migrants — bringing the inflow from 2.72 to 2.90, a 6.9% increase.

This magnitude is in line with existing estimates of drought-induced migration in sub-Saharan Africa (Henderson et al., 2017; Mueller et al., 2020; Gittard, 2024; Hoffmann et al., 2020). Relative to this literature, we focus on in-migration and identify the first stage within origin-destination corridors at annual frequency: a corridor receives more migrants precisely in the years its origin is hit by drought, net of fixed corridor characteristics and common year shocks.

Table 10 in Appendix 9.3 assesses the robustness of the first stage along two dimensions: the sample and the drought exposure window, reporting the paper's main specification throughout. Columns (3)-(4) retain only destinations classified as main cities.[6] Within this subsample, the first stage remains positive and statistically significant, although weaker: a one-standard-deviation increase in drought intensity raises in-migration by 6.4% of the sample mean, with a first-stage F-statistic of 10.5. Columns (5)-(6) further restrict the sample to migrants originating from rural areas and moving to these main cities, thereby isolating the rural-to-urban migration margin most relevant to our mechanism. In this subsample, the first stage is strongest among the restricted samples: a one-standard-deviation increase raises in-migration by 7.7% of the sample mean ($F = 21.7$). The even-numbered columns re-estimate each first stage with cumulative drought intensity, averaged over the migration year and the two preceding seasons. The cumulative measure is a stronger predictor in the full sample - a one-standard-deviation increase raises in-migration by 10% of the sample mean, against 6.9% for the contemporaneous measure - consistent with an asset-erosion mechanism, whereby households migrate after several consecutive poor seasons rather than in response to a single shock. We nonetheless retain the contemporaneous measure as our preferred instrument, for two reasons : the cumulative first stage collapses in the restricted samples (columns (4) and (6)), while the contemporaneous one remains strong; and cumula-

[6]Main cities are defined as destinations for which `share_urban_known_dest` ≥ 0.55, measured at the country's most recent census wave; rural origins are the complement, with an urban share strictly below the threshold or unmeasured.

tive exposure raises exclusion-restriction concerns - a multi-year accumulation is more likely to affect destination outcomes through channels other than the timing of migration - that are less salient for a single-season shock. Across both sample restrictions, the contemporaneous instrument remains a robust predictor of urban in-migration.

3.5 Reduced Form

Table 2 and Table 11 report the reduced-form relationship between drought at origin and the scale and propensity outcomes, respectively, estimated as

$$Y_{odt} = \beta\, D_{ot} + \gamma_1\, |t - t_c| + \gamma_{od} + \delta_t + \varepsilon_{odt}, \tag{3}$$

where Y_{odt} is either the low-quality housing inflow rate S_{odt} (the scale outcome) or the share of migrants in low-quality housing P_{odt} (the propensity outcome) in corridor (o, d) at time t; D_{ot} is drought intensity at origin; γ_{od} and δ_t are origin-destination pair and year fixed effects; and $|t-t_c|$ is the distance-to-census control. The reduced form regresses the outcome directly on the instrument, without instrumenting migration; it captures the total effect of a drought shock at origin on the outcome and rests only on the exogeneity of rainfall variation at origin.

Residential status - Scale - Columns (1)-(4) use the contemporaneous drought instrument and columns (5)-(8) its cumulative (three-season mean) counterpart on the lag-restricted sample; within each group, the first column is our main specification and the remaining columns progressively saturate the fixed effects, up to destination-by-year effects that absorb all destination-year shocks. The table shows the stability of the estimate: under the contemporaneous instrument the coefficient is essentially identical across all four configurations, indicating that the relationship is driven by within-corridor variation in drought rather than by fixed differences across origins or destinations or by destination-level shocks. The main specification in column (1) implies that a one standard deviation increase in drought intensity at origin raises the low-quality housing inflow rate by 3.3% of its mean; under the cumulative instrument (column (5)), the same shock raises it by 6.7%, consistent with several consecutive poor seasons displacing more households than a single one.

Residential status - Propensity - Beyond raising the number of arrivals, drought may shift *who* arrives toward more housing-vulnerable migrants. Table 11 in Appendix 9.3.2 reports the reduced form for this propensity channel: the direct effect of drought at origin on the composition of the migrant inflow. Columns (1)–(2) use the contemporaneous drought instrument and columns (3)–(4) the cumulative (three-season mean) instrument on the lag-restricted sample. Odd columns report the extensive margin, based on the paper's main dependent variable: the share of migrants in corridor (o, d, t) living in low-quality housing,

$$P_{o,d,t} = \frac{\sum_{i \in \text{migrants}(o,d,t)} \text{perwt}_i \cdot \mathbf{1}\{\text{slum}_i = 1\}}{\sum_{i \in \text{migrants}(o,d,t)} \text{perwt}_i},$$

where slum_i equals one if migrant i's household lacks at least one housing dimension under the main definition. Even columns report the depth of deprivation: the mean number of deficient

Table 2: Reduced Form: Drought Intensity and Low-Quality Housing Inflows, Both Instruments

	Contemporaneous drought				*3-season mean drought (lagged)*			
	(1) Main spec.	(2)	(3)	(4)	(5) Main spec.	(6)	(7)	(8)
Drought intensity	0.000851***	0.000860***	0.000852**	0.000851**	0.00287***	0.00296***	0.00205**	0.00204**
	(0.000306)	(0.000307)	(0.000416)	(0.000420)	(0.000673)	(0.000680)	(0.000815)	(0.000825)
Effect of 1-SD drought (% mean)	+3.3%	+3.3%	+3.3%	+3.3%	+6.7%	+6.9%	+4.8%	+4.8%
Origin FE	No	Yes	No	No	No	Yes	No	No
Destination FE	No	Yes	No	No	No	Yes	No	No
Origin-Destination FE	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Year FE	Yes	Yes	No	Yes	Yes	Yes	No	Yes
Country × Year FE	No	No	Yes	No	No	No	Yes	No
Destination × Year FE	No	No	No	Yes	No	No	No	Yes
Dist. to census control	Yes	Yes	No[a]	No[a]	Yes	Yes	No[a]	No[a]
Destination controls	No	No	No	No	No	No	No	No
Origin controls	No	No	No	No	No	No	No	No
R^2	0.473	0.182	0.477	0.496	0.491	0.190	0.495	0.514
Within R^2	0.003	0.002	0.000	0.000	0.004	0.002	0.000	0.000
Mean outcome	0.01300	0.01300	0.01300	0.01300	0.01278	0.01278	0.01278	0.01278
Observations	245,292	245,292	245,292	245,292	239,568	239,568	239,568	239,568
OD pairs	16,888	16,888	16,888	16,888	16,888	16,888	16,888	16,888

Notes: Reduced form: the low-quality-housing inflow rate (main UN-Habitat definition, winsorized 1–99% by country, per 100 urban residents) regressed directly on drought intensity at origin. Columns (1)–(4) use the contemporaneous instrument; columns (5)–(8) its cumulative (three-season mean) counterpart on the lag-restricted sample. Within each group, the first column is the paper's main specification and the remaining columns saturate the fixed effects; the year-FE-only specification is omitted (its coefficient is several times larger and is reported in the replication package). [a]The distance-to-census control is mechanically absorbed under country × year or destination × year fixed effects. Effect of 1-SD drought: coefficient × SD(instrument) ÷ mean outcome × 100, on each column's own sample. Sample: urban migrants, all destinations; 8 countries, 12-year window; SEs clustered at *admin_origin*, in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

housing dimensions among the same migrants. With the contemporaneous instrument, a one standard deviation drought raises the share of arriving migrants in low-quality housing by 0.34 percentage points, off a 46% baseline, while the depth of deprivation moves little. The cumulative instrument delivers a stronger and broader pattern: the extensive-margin shift roughly doubles and the depth of deprivation also rises. Because the reduced form does not rely on the first stage, these lagged estimates remain interpretable despite the weak first stage of the cumulative instrument. These results indicate that persistent drought sends not only more migrants but also sends migrants that are more likely to live in low-quality housing.

Figure 9 in Appendix 9.3.3 plots binned migration and low-quality housing inflows against drought intensity at origin, after residualizing on the main specification's fixed effects and controls. Both the first-stage and reduced-form relationships are close to linear and hold across the full distribution of drought shocks, rather than being driven by extreme outliers.

Labor Statuts - Tables 12 and 13 report the reduced forms for the employment results, regressing the scale and propensity outcomes directly on origin drought for both instruments. On the scale margin, all six inflows rise significantly at the 1% level: a one-standard-deviation drought raises the employed, unemployed, and inactive inflows by 6.8%, 10.7%, and 4.9% of their means with the contemporaneous instrument, and by 8.2%, 21.3%, and 8.2% with the cumulative one. On the propensity margin, the same shock shifts the composition of arrivals by about half a percentage point from inactivity toward employment under both instruments (+0.43 and −0.51 pp contemporaneous; +0.42 and −0.56 pp lagged), while the unemployed share barely moves. Since the reduced form does not rely on the first stage, the lagged columns provide valid inference where the lagged first stage of Table 17 is weak, and confirm the same reallocation.

4 Main Results

4.1 Residential Integration

4.1.1 Scale

Table 3 reports the main estimates of the causal effect of drought-induced migration on the urban low-quality housing inflow (the scale outcome) at destination. Columns (1)-(3) report the OLS, first-stage and 2SLS estimates using the contemporaneous drought instrument - column (3) is the paper's main estimate - and columns (4)-(6) report the same estimates using the cumulative instrument, on the sample restricted to complete lag coverage. The sample is urban migrants to all destinations, with origin–destination pair and year fixed effects and the distance-to-census control throughout.

The main scale result shows that drought-induced migration causally raises the inflow into low-quality housing at destination. The main scale result is the 2SLS coefficient in column (3): a 1 p.p increase in the drought-induced urban migration rate raises the low-quality housing inflow rate at destination by 0.228 p.p. Since both rates are expressed per hundred urban residents, the coefficient reads directly as a ratio of flows: for every additional hundred

Table 3: Drought-Induced Migration and Low-Quality Housing Inflows: Main Results

	Contemporaneous drought			*3-season mean drought (lagged)*		
	OLS	FS	2SLS	OLS	FS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Urban migration rate	0.22886***		0.22798***	0.21911***		0.31502***
	(0.04561)		(0.06352)	(0.04462)		(0.05601)
Drought intensity		0.00373***			0.00913***	
		(0.00076)			(0.00174)	
First-stage F			23.86			27.43
AR 95% CI			[0.089, 0.358]			[0.206, 0.441]
Effect of 1-SD drought (% mean)	+3.3%	+6.9%	+3.3%	+4.7%	+10.0%	+6.7%
Origin-Destination FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Dist. to census control	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.648	0.359	0.648	0.654	0.366	0.623
Within R^2	0.333	0.006	0.333	0.322	0.006	0.261
Mean outcome		0.012997			0.012779	
Mean migration rate		0.027168			0.027170	
SD(drought intensity)		0.503			0.297	
Observations		245,292			239,568	
OD pairs		16,888			16,888	

Notes: Columns (1)–(3): contemporaneous instrument; columns (4)–(6): 3-season mean instrument on the lag-restricted sample, reported as robustness – column (3) is the paper's main estimate. OLS and 2SLS columns report the effect of the urban migration rate on the low-quality housing inflow rate (main UN-Habitat definition, winsorized 1–99% by country); FS columns report the first stage of the migration rate on drought intensity, reproduced from Table 1 and its lagged counterpart. First-stage F: cluster-robust Wald on the instrument. AR 95% CI: cluster-robust Anderson–Rubin confidence interval. Effect of 1-SD drought: for OLS/2SLS columns, coefficient × first-stage coefficient × SD(instrument) ÷ mean outcome × 100; for FS columns, coefficient × SD(instrument) ÷ mean migration rate × 100. Sample: urban migrants, all destinations; 8 countries (Egypt excluded), 12-year window, Mali 1998 included. SEs clustered at *admin_origin*, in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

migrants that drought pushes into a destination, about 23 settle in low-quality housing (see Table 14 for more details on the calculation).

To interpret the magnitude of the 2SLS estimate, consider a representative origin-destination corridor with 10,000 urban residents at destination. At the baseline urban migration rate of 0.0272 migrants per 100 residents (the sample mean reported in Table 3), this corridor receives $0.0272 \times 100 = 2.72$ migrants in a year without drought, of whom $0.0130 \times 100 = 1.30$ settle in low-quality housing, the mean inflow rate reported in the same table. Now let drought intensity at origin rise by one standard deviation ($SD = 0.503$, as reported in the table). The first stage translates this shock into migration: the migration rate rises by $\beta_{FS} \times SD = 0.00373 \times 0.503 = 0.0019$, or 6.9% of its mean, where β_{FS} is the first-stage coefficient from column (2) of Table 3 - 0.19 additional migrants for the corridor, bringing the inflow to 2.90. Combining the two stages translates the same shock into housing: the low-quality housing inflow rate rises by $\beta_{2SLS} \times \beta_{FS} \times SD = 0.228 \times 0.00373 \times 0.503 = 0.00043$ - equivalently, the reduced-form coefficient of Table 2 times SD, since with a single instrument the reduced form equals the product of the two stages - or 0.043 additional residents in low-quality housing. The ratio of the two changes is, by construction, the 2SLS coefficient itself: because the low-quality housing inflow and the migration rate share the same denominator - destination urban population - $\beta_{2SLS} = 0.043/0.19 = 0.23$ reads directly as the share of the marginal inflow that settles in low-quality housing. *Of every hundred migrants that drought pushes into a city - migrants who would not have moved otherwise - roughly twenty-three end up in low-quality housing.* This 23% is the central quantitative result of the paper; Appendix Table 14 reports the calculation step by step.

Column (6) reports the 2SLS estimate under the cumulative instrument: 0.315, significant at the 1% level, with a strong first stage ($F = 27.43$) and an Anderson-Rubin interval of $[0.21, 0.44]$. The interpretation is the same ratio reading as for the main estimate, applied to a different set of marginal migrants: of every hundred migrants displaced by *persistent* drought - three consecutive poor seasons rather than a single shock - about 32 settle in low-quality housing, against 23 for those displaced by a single-season shock. The estimate lies within the main specification's Anderson-Rubin interval, so the two instruments are statistically consistent; taken at face value, the larger point estimate suggests that households pushed out by repeated shocks arrive more depleted and are absorbed into the worst housing at a higher rate.

The OLS estimates (columns (1) and (4)) give the raw within-corridor correlation between migration and low-quality housing inflows, without the causal interpretation of the 2SLS estimates. They measure the effect of an additional percentage point of migration within the origin-destination corridor on the low-quality housing inflow rate at destination. Because corridors with more migration may differ systematically in ways correlated with housing quality, this reflects a correlation rather than the causal effect of drought-induced migration. Under the contemporaneous instrument, the OLS estimate in column (1), 0.229, is nearly identical to the 2SLS estimate: a 1 p.p increase in migration flows within corridors implies that about 23 of every 100 additional migrants settle in low-quality housing. Under the cumulative instrument, the 2SLS estimate in column (6) is larger than its OLS counterpart in column (4) (0.315 against 0.219), consistent with persistent drought selecting more housing-vulnerable migrants, though the difference is imprecisely estimated and we do not push this comparison. Neither comparison directly answers whether the migrants that drought adds are *individually* more likely to end up in low-quality housing than those who would have moved anyway - a question about the composition of the inflow, which the propensity analysis addresses next.

Table 15 in Appendix 9.4 restricts the sample first to destinations that are main cities, and then to rural-to-main-city corridors. The effect is robust across both restrictions and is strongest along the rural-to-urban margin, consistent with the agricultural-push mechanism at the core of the paper: drought displaces rural households whose livelihoods depend most directly on rainfall, and these are the migrants who most raise urban low-quality housing inflows.

4.1.2 Propensity

Table 4: Propensity for Low-Quality Housing: Main Results

	Contemporaneous drought				*3-season mean drought (lagged)*			
	Deprived Dummy - Main def.		Number of deprivations Continuous		Deprived Dummy - Main def.		Number of deprivations Continuous	
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Urban migration rate	0.04026***	1.08140**	0.04737**	0.46559	0.03948***	2.49567**	0.04834**	2.45690*
	(0.01012)	(0.54054)	(0.02063)	(0.74152)	(0.01040)	(1.11527)	(0.02064)	(1.32741)
First-stage F		12.97		12.97		6.73		6.73
AR 95% CI		[0.17, 2.80]		[-1.13, 2.31]		[1.06, 10.10]		[0.49, 10.27]
Effect of 1-SD drought	+0.01 pp	+0.34 pp	+0.0001	+0.0015	+0.01 pp	+0.68 pp	+0.0001	+0.0067
R^2	0.555	0.457	0.543	0.537	0.550	0.012	0.543	0.336
Within R^2	0.010	-0.207	0.006	-0.008	0.010	-1.173	0.006	-0.444
Mean outcome	0.464	0.464	0.605	0.605	0.456	0.456	0.593	0.593
Observations	72,821	72,821	72,821	72,821	71,304	71,304	71,304	71,304
OD pairs	10,561	10,561	10,561	10,561	10,536	10,536	10,536	10,536

Notes: Shares among urban in-migrants under the main UN-Habitat definition (metal roofs durable) and the mean number of deprived dimensions; 2SLS with migration instrumented by contemporaneous (cols (1)–(4)) or 3-season mean (cols (5)–(8), lag-restricted sample) drought. Weak lagged first stage ($F \approx 7$): lagged point magnitudes not interpretable – AR intervals only. Main specification (od-pair and year FE, `dist_census` control); SEs clustered at `admin_origin`. Effect of 1-SD drought in pp of the share / dimension units. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 4 examines the propensity channel: whether the migrants that drought adds at the margin - the compliers - are more likely to settle in low-quality housing than the average migrant already on the corridor. A positive coefficient indicates that the drought-induced compliers are more housing-vulnerable than the infra-marginal migrants who would have moved regardless. We estimate the same specification as for the scale outcome, replacing the inflow rate with the composition of the inflow: the share of in-migrants living in low-quality housing under the paper's main definition, $P_{o,d,t}$ (the extensive margin, columns (1)-(2) and (5-(6)), and the mean number of deprived housing dimensions among the same migrants (the depth of deprivation, columns (3)-(4) and (7)-(8)). Columns (1)-(4) use the contemporaneous drought instrument and columns (5)-(8) the cumulative instrument on the lag-restricted sample. Robustness across alternative deprivation definitions is reported in Appendix Table 20.

The main result of this section shows that drought-induced migrants are more likely to settle in low-quality housing than other migrants on the same corridor by 3 p.p, going from 46% to 49%. The main propensity result is the 2SLS estimate in column (2): a 1 p.p increase in drought-induced migration raises the share of migrants living in low-quality housing by 1.081 p.p (significant at the 5% level, with a first-stage F of 12.97 and an Anderson-Rubin interval of $[0.17, 2.80]$ excluding zero). For a realistic shock, a one-standard-deviation drought raises the share of arriving migrants living in low-quality housing by 0.34 p.p, off a baseline of 46.4%. Equivalently, scaling by the mean migration rate, the marginal drought-induced migrant is about $1.081 \times 0.0272 \approx 3$ percentage points more likely to live in low-quality housing than the average migrant in the same corridor- roughly 49% against 46%.[7] The depth of deprivation, by contrast, does not respond to a single-season shock (column (4)): contemporaneous drought shifts *whether* migrants end up in low-quality housing, not *how* deprived that housing is.

Under the cumulative instrument, the effect appears on both margins: persistent drought raises both the share of migrants in low-quality housing (column (6)) and the depth of their deprivation (column (8)), the latter now significant unlike under the single-season shock. Because the first stage is weak on this subsample ($F = 6.73$), inference rests on the Anderson-Rubin intervals, which remain valid under weak instruments and exclude zero on both margins ($[1.06, 10.10]$ and $[0.49, 10.27]$); the point magnitudes below are therefore indicative. The estimate in column (6) implies that a 1 p.p increase in persistent-drought-induced migration raises the share of migrants living in low-quality housing by 2.50 p.p; for a realistic shock, a one-standard-deviation cumulative drought raises the share of arriving migrants in low-quality housing by 0.68 p.p - roughly twice the contemporaneous shift. Equivalently, scaling by the mean migration rate, the marginal migrant displaced by persistent drought is about $2.50 \times 0.0272 \approx 7$ p.p more likely to live in low-quality housing than the average migrant - 53% against 46%. The depth of deprivation now responds as well (column (8)): a one-standard-deviation cumulative drought raises the mean number of deprived housing dimensions among arriving migrants by 0.007, off a baseline of 0.59, and, scaling by the mean migration rate, the marginal migrant lives in housing with about $2.46 \times 0.0272 \approx 0.07$ more deprived dimensions than the average migrant's. Migrants displaced by persistent drought are thus not only more often in low-quality housing, but in worse housing.

Reassuringly, these patterns are consistent with the reduced-form estimates (Table 11), which do not depend on first-stage strength, though they cannot net out the migration channel. Persistent drought thus displaces a more housing-vulnerable population across a broader range of margins, consistent with repeated poor seasons eroding household resources before departure.

The OLS coefficient in column (1) is small (0.040): across all migration, larger inflows barely change the deprived-housing share of arrivals - ordinary migration comes at roughly the average housing mix. The 2SLS estimate is more than twenty times larger. Unlike at the scale margin, where OLS and 2SLS coincide, drought-induced migration shifts the

[7]This conversion holds exactly under homogeneous corridor effects and is evaluated at sample means.

composition of the inflow far more than migration does on average. This shows the selection result: drought does not add migrants at the average housing mix, but migrants who are disproportionately housing-vulnerable.

For the restricted samples, the propensity point estimates are consistent with the full-sample result but too imprecise to establish it independently (Appendix Table 16).

4.2 Labor-Market Integration

The housing results show that drought-induced migration raises the inflow into low-quality housing. In this section, we address a complement: whether and how drought-induced migrants - the compliers of our IV estimate - are absorbed into the urban labor market, in terms of employment status: employed, unemployed, and inactive. For all three, we use the same identification strategy, counting the number of migrants within the corridor (o, d, t) that are employed, unemployed, and inactive per 100 urban residents, and we measure both the scale and the propensity channel.

4.2.1 Scale

Table 5: Drought-Induced Migration and the Labor-Force Status of the Inflow: Scale

	Employed		*Unemployed*		*Inactive*	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Urban migration rate	0.09679***	0.33946***	0.00880***	0.05931***	0.15981***	0.26934***
	(0.01925)	(0.04912)	(0.00203)	(0.01231)	(0.03144)	(0.03798)
First-stage F		23.86		23.86		23.86
AR 95% CI		[0.262, 0.478]		[0.040, 0.095]		[0.197, 0.359]
Effect of 1-SD drought (% mean)	+1.9%	+6.8%	+1.6%	+10.7%	+2.9%	+4.9%
R^2	0.612	-0.013	0.460	-0.270	0.616	0.533
Within R^2	0.210	-1.065	0.042	-1.254	0.322	0.176
Mean outcome	0.00934	0.00934	0.00104	0.00104	0.0103	0.0103
Observations	245,292	245,292	245,292	245,292	245,292	245,292

Notes: Outcomes: employed / unemployed / inactive in-migrants per 100 urban residents, winsorized 1–99% by country. 2SLS: migration instrumented by contemporaneous drought. Main specification (od-pair and year FE, `dist_census` control); SEs clustered at `admin_origin`. Undetermined labor-force status excluded from the known-status denominator. Sample: urban migrants, all destinations; 8 countries, 12-year window. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 5 addresses the scale margin, replacing the outcome of the main specification with the number of employed, unemployed, and inactive in-migrants per 100 urban residents; migrants whose labor-force status is undetermined in the census are excluded from the counts. The specification, instrument, and sample are otherwise identical to Table 3.

All three inflows increase significantly with migration, both in OLS and 2SLS. The 2SLS coefficient for the employed inflow is 0.339, significant at the 1% level: an increase in the migration rate of one in-migrant per 100 urban residents raises the number of employed in-migrants by 0.339 per 100 urban residents. The corresponding coefficients are 0.059 for the unemployed inflow and 0.269 for the inactive inflow, both also significant at the 1% level. The first stage is unchanged from the main table ($F = 23.86$), and the Anderson-Rubin intervals are bounded away from zero for all three outcomes. Because each outcome shares the denominator of the migration rate, these coefficients read directly as the labor-force status of the marginal migrants themselves: of every 100 additional drought-induced migrants, about 34 arrive into employment, 6 into unemployment, and 27 into inactivity, the remainder having undetermined status.

In relative terms, a one-standard-deviation drought raises the employed, unemployed, and inactive inflows by 6.8%, 10.7%, and 4.9% of their respective means. The larger relative rise in unemployment comes off a very small base - unemployed migrants are a negligible share of the inflow in levels - so in absolute terms employment dominates: the urban labor market absorbs a substantial part of the drought-induced inflow into work. These migrants find employment far more often than they join the ranks of the unemployed, even as they disproportionately settle in low-quality housing.

In comparison, the OLS estimates are smaller (0.097 for the employed, 0.009 for the unemployed, and 0.160 for the inactive inflow), and their ordering differs: in OLS the inactive inflow is the largest component of migration, whereas in 2SLS it is employment. The OLS coefficients pool all variation in migration - including flows unrelated to drought and measurement error in census-recalled moves, which attenuates the estimates - while the 2SLS isolates the drought-driven component. The comparison therefore suggests that the drought-induced inflow is more work-oriented than the typical migration flow, though the two estimates answer different questions and the gap cannot be read as a measure of bias. Whether drought-induced migrants indeed differ from the average arrival in their labor-force status is precisely what the propensity margin tests directly, holding the size of the inflow fixed; we turn to it next.

4.2.2 Propensity

Table 6 turns to the propensity margin: the outcomes are now the shares of employed, unemployed, and inactive migrants among in-migrants with known labor-force status. As in the housing propensity analysis, the denominator is the number of migrants rather than the urban population, so the estimates hold the size of the inflow fixed and isolate changes in its composition. Because the three shares sum to one, the 2SLS coefficients sum to zero by construction: drought can only reallocate migrants across statuses. The specification and instrument are identical to the main table; on this smaller sample of corridors with observed labor-force status, the first stage remains strong ($F = 13.49$).

Drought tilts the inflow toward employment. The 2SLS coefficient on the employed share is 1.319, significant at the 5% level, and its mirror image is a decline in the inactive share of -1.586, also significant at the 5% level; the unemployed share is unchanged (0.267, not significant, with an Anderson-Rubin interval that includes zero). Multiplying by the mean migration rate of 0.0272 (Table 3) converts these coefficients into the gap between marginal and average migrants: drought-induced migrants are about 3.6 percentage points more likely to be employed than the average in-migrant (roughly 54% against 51%) and

Table 6: Drought-Induced Migration and the Labor-Force Status of the Inflow: Propensity

	Employed		*Unemployed*		*Inactive*	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Urban migration rate	-0.09706***	1.31911**	0.00161	0.26703	0.09545***	-1.58614**
	(0.01522)	(0.57790)	(0.00612)	(0.22926)	(0.01868)	(0.62105)
First-stage F		13.49		13.49		13.49
AR 95% CI		[0.404, 3.217]		[-0.149, 0.936]		[-3.749, -0.644]
Effect of 1-SD drought	−0.03 pp	+0.43 pp	+0.00 pp	+0.09 pp	+0.03 pp	−0.51 pp
R^2	0.331	0.098	0.243	0.198	0.326	-0.012
Within R^2	0.005	-0.342	0.000	-0.059	0.006	-0.491
Mean outcome	0.5053	0.5053	0.0578	0.0578	0.4368	0.4368
Observations	71,587	71,587	71,587	71,587	71,587	71,587

Notes: Outcomes: employed / unemployed / inactive shares among known-status in-migrants; the three shares sum to one, so the 2SLS coefficients sum to zero by construction. 2SLS: migration instrumented by contemporaneous drought. Main specification (od-pair and year FE, `dist_census` control); SEs clustered at `admin_origin`. Undetermined labor-force status excluded from the known-status denominator. Sample: urban migrants, all destinations; 8 countries, 12-year window. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

about 4.3 percentage points less likely to be inactive (roughly 39% against 44%). A one-standard-deviation drought shifts the composition of arrivals by +0.43 percentage points toward employment and −0.51 percentage points away from inactivity.

The OLS estimates again differ from the 2SLS in a telling way: they are small and of the opposite sign (−0.097 for the employed share, +0.095 for the inactive share, both significant at the 1% level), so typical variation in migration is associated with slightly less employed inflows, while the drought-driven component is more employed. This is the direct confirmation of the pattern suggested by the scale margin: the marginal, drought-induced migrant differs from the average arrival, and differs in the direction of work. Together with the housing results, this points to a precise characterization of drought-induced migrants: they integrate the urban labor market more than the typical arrival, yet disproportionately settle in low-quality housing.

Table 17 in Appendix 9.4.3 re-estimates both margins with the cumulative drought instrument: the scale results are unchanged, and the shift of arrivals from inactivity toward employment persists.

4.3 Interpreting the Effect

4.3.1 Decomposing the Effect: Scale versus Propensity

Residential Integration - To assess the relative importance of the two channels, the additional low-quality housing inflow from a one-standard-deviation drought decomposes as

$$\Delta S \;\approx\; \underbrace{\Delta \text{MigRate} \times \overline{P}}_{\text{scale}} \;+\; \underbrace{\overline{\text{MigRate}} \times \Delta P}_{\text{propensity}},$$

where the scale term captures additional migrants arriving with the average propensity for low-quality housing, and the propensity term the baseline flow of migrants arriving with a higher propensity. The scale component is $0.00188 \times 0.464 = 0.00087$ additional migrants in low-quality housing per 100 urban residents, and the propensity component $0.02717 \times 0.0034 = 0.00009$: about 90 % of the additional inflow operates through the number of migrants and 10% through their composition. The reason is transparent in the two sets of estimates - drought raises the number of migrants by 6.9 % while raising their propensity for low-quality housing by a third of a p.p. (from 46.4 to 46.8 percent of arrivals). Drought therefore fills low-quality housing primarily by sending more migrants, and secondarily by sending more vulnerable ones. Appendix Table 18 details the calculation. These shares are approximate: the two channels are estimated on different samples - the propensity margin requires migrants with observed housing characteristics - and the identity is evaluated at the average migrant's propensity rather than the marginal migrant's, so it measures the relative importance of the channels rather than the level of the effect.

This decomposition also explains why OLS and 2SLS coincide at the scale margin (Table 3) while differing sharply at the propensity margin. A gap between OLS and 2SLS on the scale outcome would require a large propensity channel - drought-induced migrants entering low-quality housing at a markedly different rate than ordinary migrants. But since this channel accounts for only about 10% of the total effect, the selection it reflects is too small to move the scale estimate detectably, and OLS and 2SLS are nearly identical there. The selection is not absent - it is directly visible on the propensity outcome, where the 2SLS estimate is more than twenty times the OLS one - but it is dominated by the scale channel. The two findings are consistent: drought-induced migrants are more housing-vulnerable than average, yet this composition shift is second-order relative to the increase in migrant numbers.

Labor Integration - The employment results decompose along the same lines as the housing results. Applying the identity of Section 4.3.1 to the employed inflow, the scale component is $0.00188 \times 0.505 = 0.00095$ additional employed migrants per 100 urban residents and the propensity component $0.02717 \times 0.0043 = 0.00012$: about 89% of the additional employed inflow comes from the larger number of migrants and 11% from their higher propensity to be employed, with the same caveats as before - the two margins are estimated on different samples, here migrants with known labor-force status, and the identity is evaluated at the average arrival's employment share. Drought therefore expands urban employment among migrants mainly by scale, exactly as it fills low-quality housing: the two outcomes share not only the same shock but the same anatomy.

4.3.2 The Aggregate Picture

Residential Integration - The corridor-level estimates can be aggregated to gauge the size of the phenomenon. Over 1981-2018, the censuses record about 11.6 million urban in-migrants in the corridors we observe across the eight countries. Drought accounts for roughly 400,000 of these moves - about 3.4% of observed urban in-migration.[8] Applying the 2SLS estimate, under which 23 of every 100 marginal migrants settle in low-quality housing (Table 3), drought placed roughly 91,000 additional people in low-quality housing over the period. The decomposition of Section 4.3.1 attributes about 90% of this inflow to the scale channel and 10% to the propensity channel: of the roughly 91,000 additional residents of low-quality housing, about 82,000 arrive because drought sends more migrants to the cities, and about 9,000 because it shifts the distribution of arrivals toward low-quality housing. Drought degraded urban housing conditions mostly by adding people, and as well by re-sorting a given flow of migrants into worse housing.

These magnitudes are conservative. They cover only the corridors observed in the censuses, within the twelve-year windows preceding each census; they capture permanent, inter-regional moves and miss temporary or short-distance migration; and they hold destination conditions fixed, abstracting from any equilibrium response of housing supply.

Labor Integration - The same aggregation applies to the labor-market results: of the roughly 400,000 drought-induced migrants, about 136,000 arrived into employment, 24,000 into unemployment, and 108,000 into inactivity, the remainder having undetermined status (Table 5). Drought thus added workers to the cities on a similar scale as it added residents of low-quality housing.

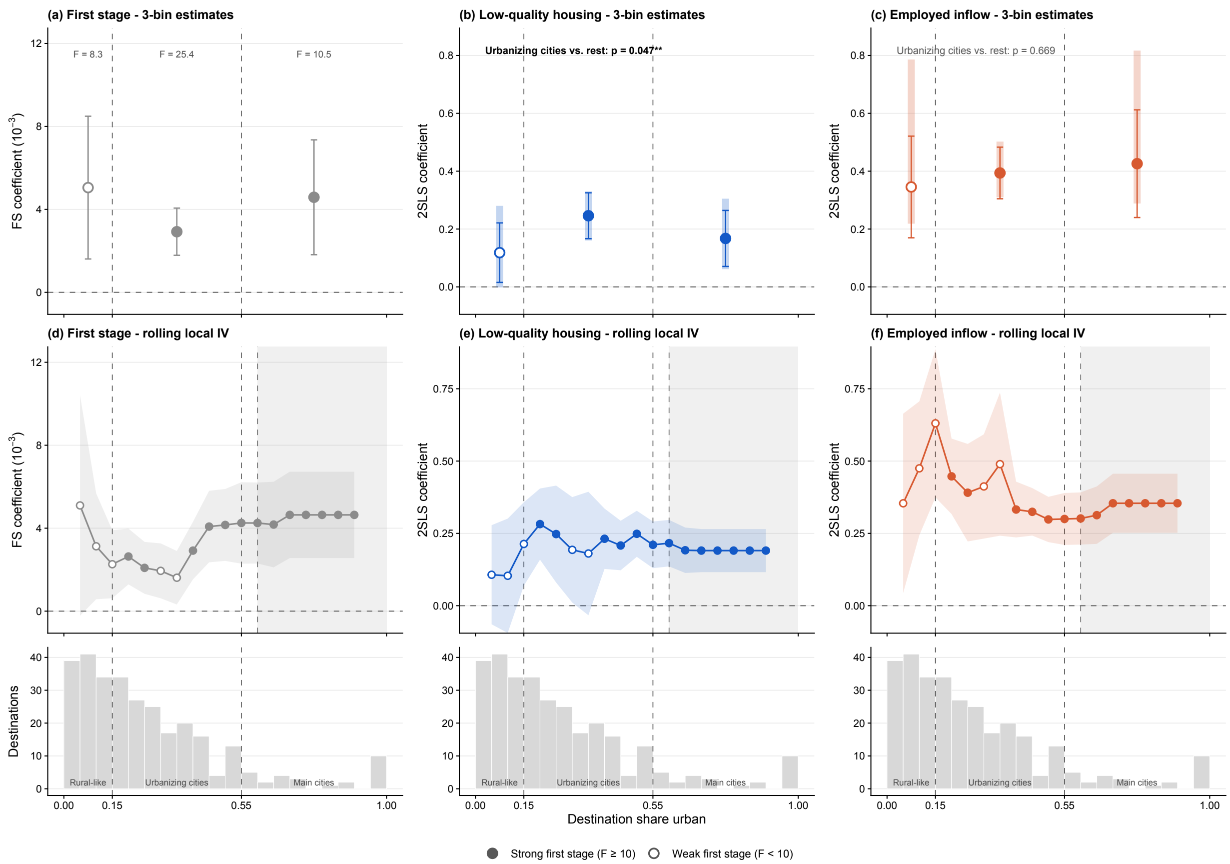
Working Poor - Table 19 in Appendix 9.4.4 cross-tabulates the two dimensions on the same individuals, splitting the inflow into four employment $\times$ housing cells. All four inflows rise with drought-induced migration, but unevenly: the two employed cells and the not-employed cell outside low-quality housing increase significantly at the 1% level, while the most deprived cell - not employed and in low-quality housing - shows the weakest and least precise response. The key ratio concerns the composition of the low-quality housing inflow: about two-thirds of the drought-induced migrants entering low-quality housing arrive employed (0.124/0.187), against 42% of the baseline inflow (0.004362/0.010490, from the mean outcomes in columns (1) and (3)). The marginal residents of low-quality housing are thus predominantly working poor rather than idle poor - drought-induced migrants integrate the labor market yet still settle in deprived housing, now verified on the same individuals rather than inferred from the separate housing and employment results.

[8] The number of drought-induced migrants is obtained by multiplying each corridor-year's realized drought intensity by the first-stage coefficient and the destination's urban population, and summing across corridors and years, which yields $\approx$ 400,250 migrants, or 3.4% of the 11.6 million observed. Applying the 2SLS estimate to this flow gives the additional residents of low-quality housing: $400{,}250 \times 0.228 \approx 91{,}250$. Both coefficients are taken from Table 3.

5 Heterogeneity

5.1 Heterogeneity by Destination Urbanization

Figure 1: Heterogeneity by Destination Urbanization: Low-Quality Housing and Employment



Notes: Heterogeneity by destination urbanization (urban population share at the most recent census; zones: rural-like < 0.15, urbanizing cities $[0.15, 0.55)$, main cities ≥ 0.55). Panels (a)-(c): first-stage and 2SLS estimates from a single specification saturated by urbanization bin -the migration rate interacted with each bin indicator, instrumented by drought intensity interacted with the same indicators, with origin–destination and arrival-year × bin fixed effects and bin-specific distance-to-census controls; standard errors clustered at the origin district. Estimates are plotted at bin midpoints; whiskers are 95% Wald intervals and shaded bars Anderson-Rubin 95% intervals; outcomes in panels (b)–(c) are the low-quality-housing and employed inflows per 100 urban residents. For housing, the urbanizing-cities bin differs from the other two pooled ($p = 0.047$); pairwise: urbanizing − rural-like $p = 0.031$, urbanizing − main cities $p = 0.113$, rural-like − main cities $p = 0.463$. All employment contrasts have $p > 0.55$. Panels (d)–(f): rolling-window local IV – the main specification re-estimated on corridors into the ∼60 destinations nearest each urban-share grid point; pointwise 95% Wald bands; indicative only, as windows overlap. The shaded area marks where the window no longer slides (too few destinations above 0.55); the binned estimates in panels (a)–(c) carry the inference there. Hollow points flag a weak first stage (Sanderson–Windmeijer conditional $F < 10$; bin values 8.3 / 25.4 / 10.5). Bottom row: distribution of destination urban shares with the three zones labeled. N = 193,290 corridor-years (64,720 / 109,456 / 19,114 across bins).

We examine whether the effect varies with the type of destination, splitting destinations into three groups by their urban population share at the most recent census: rural-like (< 0.15), urbanizing cities ($[0.15, 0.55)$), and main cities (≥ 0.55). We estimate a single specification in which the migration rate is fully interacted with the destination-type indicators:

$$S_{o,d,t} = \sum_{k=1}^{3} \beta_k \big(\text{MigRate}_{o,d,t} \times \mathbf{1}\{d \in k\}\big) + \sum_{k=1}^{3} \delta_k \big(\text{dist}_{o,d,t} \times \mathbf{1}\{d \in k\}\big) + \alpha_{od} + \mu_{t\times k} + \varepsilon_{o,d,t}, \tag{4}$$

where each interacted migration term is instrumented by the corresponding interaction of drought intensity at origin, $Z_{o,t} \times \mathbf{1}\{d \in k\}$ (three endogenous variables, three instruments: just-identified), α_{od} are corridor fixed effects, and $\mu_{t\times k}$ arrival-year $\times$ bin fixed effects. No uninteracted migration term is included, so each β_k measures the absolute effect within destinations of type k rather than a difference from a reference category. The system is numerically equivalent to estimating the main specification separately by destination type, while the pooling delivers the covariance across the β_k used to test for differences.[9] Figure 1 reports the binned estimates with the first stage by group (panels (a)–(c)), alongside a continuous rolling-window counterpart that re-estimates the main specification along the urban-share distribution (panels (d)–(f), indicative only) and the distribution of destinations; full estimates and contrasts are in Supplementary Appendix Table S1.

Drought-induced migrants sort into low-quality housing primarily in urbanizing cities. The effect follows an inverted U (panel (b)): of every 100 marginal migrants, about 25 enter low-quality housing in an urbanizing city (0.246), against 12 in rural-like destinations (0.118, weakly identified) and 17 in main cities (0.167). Urbanizing cities differ significantly from the other types pooled ($p = 0.047$) and from rural-like destinations ($p = 0.031$); the gap with main cities is of the same sign but imprecise ($p = 0.113$), as few destinations exceed the 0.55 threshold (bottom row). The rolling estimates in panel (e) trace the same pattern continuously: the effect climbs as destinations cross into the urbanizing range and flattens beyond it. By contrast, absorption into employment is uniform across destination types (panels (c) and (f)), with all contrasts far from significance ($p > 0.55$). The heterogeneity is thus specific to the housing margin: urban labor markets absorb drought migrants everywhere, but in urbanizing cities this absorption runs disproportionately through low-quality housing.

The pattern is unchanged when destinations are instead classified by city size using the Africapolis database: the housing effect again vanishes in the largest cities while employment absorption remains uniform (Supplementary Appendix Figure S4).

Supplementary Appendix Figure S5 shows a consistent (though noisier) pattern for the propensity margin: the sorting of climate migrants into low-quality housing is detectable only in urbanizing cities (where the weak-instrument-robust interval excludes zero).

6 Robustness

This section presents a series of robustness checks and alternative specifications designed to assess the stability of the main results and the validity of the identification strategy. Several of these checks are reported in the appendix - the fixed-effects saturation in Table 2, the

[9] Saturation by bin is required for identification: with pooled year fixed effects or a continuous interaction, the interacted instruments do not separately identify the interacted endogenous variables (Sanderson-Windmeijer conditional F near zero).

definition sweep, and the wet-side placebo of Figure 7 - and the remaining estimates are in progress.

Alternative specifications. - We consider four alternative specifications. First, we estimate a manually specified two-instrument design combining $\text{Drought}^g * p10$ and $N^g * dry$ to examine whether the results are robust to multi-instrument identification despite a weaker joint first stage. Second, we replace drought intensity with extreme precipitation measures based on high-percentile rainfall shocks to test whether migration responses are symmetric across the rainfall distribution or primarily driven by adverse drought events. Third, following Hultgren et al. (2025), we perform an out-of-sample validation to evaluate the stability of the post-LASSO instrument selection across held-out countries. Fourth, we replace the census-based destination population denominator with the annually reconstructed population series, $\text{Pop}_{d,t}$, to assess the sensitivity of the estimated migration rates to the choice of normalization.

Alternative samples. - To assess the external validity of the main findings, we extend the analysis to all fifteen countries in the dataset by exploiting five-year migration intervals rather than restricting attention to the eight countries with annual migration information. This exercise evaluates whether the estimated scale effect generalizes beyond the baseline sample.

Sensitivity analyses. - We conduct two additional sensitivity analyses. First, we replace the binary low-quality housing indicator with a continuous measure of housing deprivation defined as the number of deficient UN-Habitat dimensions (ranging from 0 to 6), allowing us to test whether drought-induced migration affects the entire distribution of housing quality rather than only the extensive margin of low-quality housing residence. Second, we perform a leave-one-country-out exercise to verify that the estimated effects are not driven by any single country in the sample.

Exclusion restriction. - The identifying assumption may be violated if drought at origin affects destination outcomes through channels other than migration, such as food price transmission, remittance flows, or government infrastructure responses. Origin–destination pair fixed effects absorb time-invariant correlations between origin characteristics and destination outcomes, while arrival-year fixed effects capture aggregate shocks common to all migration corridors. To further assess the plausibility of the exclusion restriction, we are investigating whether origin-specific drought shocks predict destination food prices after conditioning on the baseline specification. Evidence from this exercise will be reported in future versions of the paper.

Comparison with other migration drivers. - Two additional exercises currently in progress will compare drought-induced migration with migration generated by alternative shocks. The first instruments migration using conflict events at origin (ACLED data); the second uses exogenous variation in mine openings at the admin-unit level. These comparisons sharpen the interpretation of the LATE estimated in the main analysis by contrasting push driven (drought, conflict) and pull-driven (mining) migration mechanisms and their respective urban absorption patterns.

7 A Model of Post-Migration Housing Integration

7.1 Motivation

The empirical analysis establishes three main results. First, drought at origin causally increases rural-to-urban migration. Second, this drought-induced migration raises urban low-quality housing inflows: of every hundred additional climate migrants, about twenty-three settle in low-quality housing, mainly through the scale of arrivals but also through selection - drought-induced migrants are about three percentage points more likely to end up in low-quality housing than the migrants who would have moved anyway. Third, these same migrants are absorbed into urban employment: the housing and employment margins decouple, and the integration gap of climate migrants is residential rather than occupational.

While these reduced-form results document a systematic residential integration gap, they do not quantify the extent to which this gap reflects binding integration barriers - deposits and guarantor requirements in the formal rental market, documentation, lack of local networks - as opposed to selection on income or preferences, nor do they allow for counterfactual policy analysis. The framework below addresses the propensity margin only. The scale margin - how many migrants arrive - is taken as given, since it is governed by the migration decision the empirical strategy instruments rather than by post-arrival sorting within the city. To make progress on these questions, we sketch a simple conceptual framework of post-migration housing integration within cities.

The framework is static and partial equilibrium. Migrants have already arrived in the city, and it abstracts from migration decisions across locations, focusing on the housing integration choice after arrival. The decoupling documented above disciplines its key ingredient: because climate migrants find work at least as often as other arrivals, the barrier they face is specific to the housing margin.

7.2 A Two-State Example

Consider migrants who have already arrived in a destination city, and two housing integration states: a low-integration state L (low-quality housing) and a high-integration state H (formal, adequate housing). Each migrant draws a pair of state-specific capacities (s_{iL}, s_{iH}), indexing how well the migrant fares in each state. Draws are margin-specific: (s^E_{iL}, s^E_{iH}) for employment and (s^R_{iL}, s^R_{iH}) for housing. Formal housing is valued for the bundle of amenities and services it delivers rather than for a differential return to labor, which keeps the two margins separable and consistent with the decoupling documented above. Let w_H and w_L denote the value per unit of capacity in each state - on the housing margin, inclusive of the amenity bundle. Absent barriers, migrants choose the state that maximizes value: those with $w_H\, s_{iH} > w_L\, s_{iL}$ sort into formal housing. Graphically (Appendix Figure 11), these migrants lie above the ray OC with slope w_L/w_H.

Residential integration barriers - deposits for formal rentals, guarantor and documentation requirements, lack of local networks - reduce the effective value of the high-integration state. If climate migrants face a proportional integration cost τ, their effective return in formal housing becomes $(1-\tau)\,w_H$, and only those above the steeper ray OE, with slope $w_L/\big[(1-\tau)\,w_H\big]$, choose it. The rays partition migrants into four groups. In Region I, above OE, migrants integrate into formal housing even in the presence of barriers. Regions II and III, between OE and OC, contain the migrants the barrier misallocates: they would choose formal housing absent barriers - whether their comparative advantage lies in the high-integration state (Region II, above the 45° line) or not (Region III) - yet remain in low-quality housing. Region IV, below OC, comprises migrants who optimally remain in low-quality housing regardless of barriers.

Two qualifications on how the rays are read. First, baseline migrants face their own barriers $\tau^0 \geq 0$, so OC is their cutoff rather than a frictionless benchmark: the propensity gap identifies the differential barrier $\tau - \tau^0$, and recovering its level - and hence the welfare gain from a given reduction in τ - requires an external restriction on τ^0. Second, the trapped fan is defined in the space of housing draws, so labelling it the working poor presumes that migrants in that fan also clear the employment margin. This holds if capacities are positively correlated across margins, which the cross-tabulation supports but does not impose.

This diagram organizes the interpretation of our propensity result: a higher τ for drought-induced migrants shifts part of their distribution from Region I into the trapped Regions II–III, raising their share in low-quality housing relative to baseline migrants on the same corridor - precisely the three-percentage-point gap we estimate - without affecting their employment. The same logic applies margin by margin, and this is where the framework connects to the paper's distinctive finding. Consider two integration margins - employment and housing - each with its own two-state choice and its own proportional barrier, τ_E and τ_R (Appendix Figure 11). Our employment results discipline the first wedge: drought-induced migrants are absorbed into work at least as often as baseline arrivals, so $\tau_E \approx 0$ and the employment margin has no trapped region (panel (a)). The housing results discipline the second: a residential gap alongside absorption into work requires $\tau_R > 0$, opening the trapped fan of panel (b). Jointly, a residential-only barrier predicts that climate migrants concentrate in the employed-but-badly-housed cell - the working poor - which is precisely what the cross-tabulation of housing and employment outcomes shows (Appendix Table 19).

The diagram is an interpretive device rather than an identification result. Two alternatives generate the same propensity gap without any barrier. Migrants may sort on preferences or prices, choosing low-quality housing to economise on rent. More importantly, the gap may reflect the composition of draws rather than a rotation of the cutoff: drought-induced migrants originate from agricultural areas and may draw a lower s_H/s_L ratio, shifting the distribution across a fixed OC. The diagram does not separate the two. A proportional barrier rotates the cutoff while a fixed cost - a rental deposit - shifts it, so the two are distinguishable along the income distribution with price and income data. Doing so, estimating the size of the trapped region for climate migrants relative to other migrants, and quantifying the welfare gains from policies that reduce τ (housing vouchers, formal rental access programs), is the object of the structural extension of this paper.

8 Conclusion

This paper provides causal evidence on the contribution of drought-induced migration to the growth of low-quality housing in Sub-Saharan Africa. Using harmonized census microdata for eight countries spanning 1987–2019, it combines a cohort-bilateral panel of migration flows with high-resolution rainfall data to construct exogenous variation in drought intensity at origin. A post-IV LASSO procedure selects drought intensity as the instrument for bilateral migration, and the 2SLS estimates identify a Local Average Treatment Effect on drought-induced migrants, precisely the population climate policy concerns.

Drought-induced migration raises urban low-quality housing inflows at destination: of every hundred migrants that drought pushes into a city, about twenty-three settle in low-quality housing, and over 1981–2018 drought placed roughly 91,000 additional people in low-quality housing across our corridors. This effect operates mainly through the scale of arrivals, but also through selection: drought-induced migrants are about three percentage points more likely to settle in low-quality housing than the migrants who would have moved anyway. Yet these migrants are absorbed into urban employment across destinations; the margin of exclusion is residential, not occupational. The housing effect is concentrated along rural-origin corridors, consistent with an agricultural-push mechanism, and in urbanizing cities rather than the largest metropolises, a pattern confirmed when destinations are classified by agglomeration size using Africapolis data.

These findings have direct implications for urban and climate policy in Sub-Saharan Africa. Because the effect runs mainly through the scale of arrivals, the immediate priority is expanding urban absorption capacity; the selection channel calls, in addition, for targeting the most vulnerable climate migrants. The concentration in urbanizing secondary cities suggests that housing policies focused on primary metropolitan areas may overlook the destinations most affected by climate-induced migration. If low-quality housing itself exposes residents to climate hazards (Marx et al., 2019), climate-induced growth in low-quality housing may amplify rather than mitigate the costs of climate variability.

Several avenues remain open for future research. A structural model of migrant integration, jointly modeling housing and labor-market choices, would quantify the welfare cost borne by climate migrants and allow the evaluation of targeted policies such as housing support in urbanizing cities. Subnational crop data, such as HarvestStat Africa, would allow a direct test of the agricultural mechanism - drought reducing harvests, and harvests driving migration. As climate change intensifies and urbanization accelerates, understanding the interaction between climate shocks, migration, and urban integration will be increasingly important for policies promoting both adaptation and inclusive urban development.

References

Alvarez, J. A. (2020). The agricultural wage gap: Evidence from brazilian micro-data. *American Economic Journal: Macroeconomics*, 12(1):153–173.

Auffhammer, M. (2018). Quantifying economic damages from climate change. *Journal of Economic Perspectives*, 32(4):33–52.

Beine, M. and Parsons, C. R. (2016). Climatic factors as determinants of international migration: Redux. CREA Discussion Paper Series 16-11, Center for Research in Economic Analysis, University of Luxembourg.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.

Bryan, G. and Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Cattaneo, C. and Peri, G. (2016). The migration response to increasing temperatures. *Journal of Development Economics*, 122(Supplement C):127 – 146.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2(1):1–21.

Gittard, M. (2024). Droughts, migration and population in kenya.

Gollin, D., Jedwab, R., and Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth*, 21(1):35–70.

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

Gray, C. L. and Mueller, V. (2012). Natural disasters and population mobility in bangladesh. *TProceedings of the National Academy of Sciences*, 109(16):6000–6005.

Hamory, J., Kleemans, M., Li, N. Y., and Miguel, E. (2021). Reevaluating agricultural productivity gaps with longitudinal microdata. *Journal of the European Economic Association*, 19(3):1522–1555.

Henderson, J. V. and Kriticos, S. (2018). The development of the african system of cities. *Annual Review of Economics*, 10(1):287–314.

Henderson, J. V., Storeygard, A., and Deichmann, U. (2017). Has climate change driven urbanization in africa? *Journal of development economics*, 124:60–82.

Hoffmann, R., Dimitrova, A., Muttarak, R., Crespo Cuaresma, J., and Peisker, J. (2020). A meta-analysis of country-level studies on environmental change and migration. *Nature Climate Change*, 10(10):904–912.

Hsiang, S. (2016). Climate econometrics. *Annual Review of Resource Economics*, 8:43–75.

Hultgren, A., Carleton, T., Delgado, M., Gergel, D. R., Greenstone, M., Houser, T., Hsiang, S., Jina, A., Kopp, R. E., Malevich, S. B., McCusker, K. E., Mayer, T., Nath, I., Rising, J., Rode, A., and Yuan, J. (2025). Estimating global impacts to agriculture from climate change accounting for adaptation. *Nature*, 641:644–652.

Jedwab, R., Christiaensen, L., and Gindelsky, M. (2017). Demography, urbanization and development: Rural push, urban pull and… urban push? *Journal of Urban Economics*, 98:6–16.

Kleemans, M. and Magruder, J. (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal*, 128(613):2032–2065.

Marx, B., Stoker, T., and Suri, T. (2013). The economics of slums in the developing world. *Journal of Economic perspectives*, 27(4):187–210.

Marx, B., Stoker, T. M., and Suri, T. (2019). There is no free house: Ethnic patronage in a kenyan slum. *American Economic Journal: Applied Economics*, 11(4):36–70.

McGuirk, E. and Nunn, N. (2021). Transhumant pastoralism, climate change, and conflict in africa. *Working Paper*.

McMillan, M., Rodrik, D., and Verduzco-Gallo, Í. (2014). Globalization, structural change, and productivity growth, with an update on africa. *World development*, 63:11–32.

Minnesota Population Center (2019). Integrated public use microdata series, international: Version 7.2 [dataset]. Minneapolis, MN: IPUMS.

Missirian, A. and Schlenker, W. (2017). Asylum applications respond to temperature fluctuations. *Science*, 358(6370).

Mueller, V., Gray, C., and Hopping, D. (2020). Climate-induced migration and unemployment in middle-income africa. *Global Environmental Change*, 65:102183.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N. (2021). Era5-land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9):4349–4383.

OECD/SWAC (2024). Africapolis (database). https://africapolis.org.

Trabucco, A. and Zomer, R. (2019). Global aridity index and potential evapotranspiration (ET0) climate database v2. CGIAR Consortium for Spatial Information (CGIAR-CSI).

UN-Habitat (2003). *The Challenge of Slums: Global Report on Human Settlements 2003*. Earthscan, London.

Un-Habitat (2020). Strengthen un-habitat regional representation in sub-saharan africa. *Un-Habitat report*.

Un-Habitat (2022). Envisaging the future of cities. *World Cities Report*.

9 Appendix

9.1 Descriptive Statistics

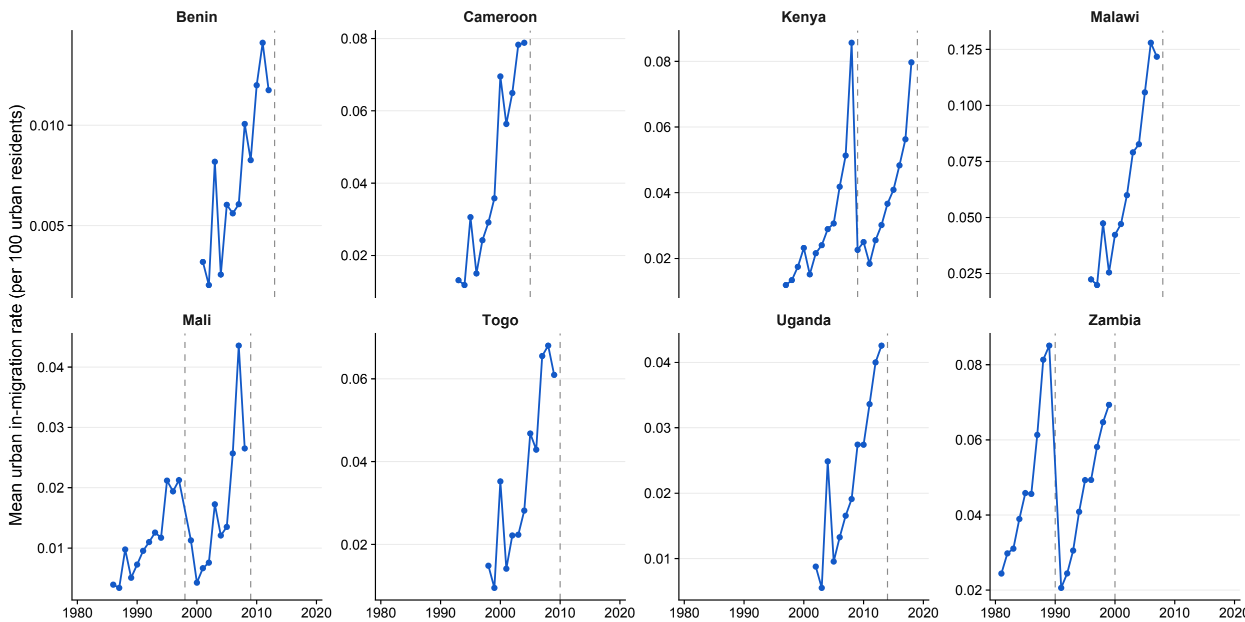
Table 7 reports descriptive statistics for all variables used in the analysis: the corridor-year estimation panel (Panel A), the migrant propensity outcomes (Panel B), and destination characteristics (Panel C).

Table 7: Descriptive Statistics

Variable	Mean	SD	P25	Median	P75	N
Panel A: Corridor-year panel (main estimation sample)						
Urban migration rate (per 100 urban residents)	0.0272	0.1192	0.0000	0.0000	0.0092	245,292
Low-quality housing inflow rate (per 100)	0.0130	0.0524	0.0000	0.0000	0.0000	245,292
Employed inflow rate (per 100)	0.0093	0.0293	0.0000	0.0000	0.0000	245,292
Unemployed inflow rate (per 100)	0.0010	0.0056	0.0000	0.0000	0.0000	245,292
Inactive inflow rate (per 100)	0.0103	0.0363	0.0000	0.0000	0.0000	245,292
Drought intensity at origin	0.352	0.503	0.000	0.000	0.628	245,292
Cumulative 3-season drought at origin	0.367	0.297	0.107	0.348	0.553	239,568
Distance to census $\|t - t_c\|$ (years)	6.07	3.29	3.00	6.00	9.00	245,292
Urban population at destination ('000)	121.9	313.5	15.7	40.0	94.7	245,292
Panel B: Migrant propensity outcomes (corridors with positive denominators)						
Share of migrants in low-quality housing	0.464	0.414	0.000	0.420	1.000	72,821
Mean number of deprived dimensions (0–6)	0.605	0.657	0.000	0.500	1.000	72,821
Share of migrants employed	0.505	0.363	0.200	0.500	0.843	71,587
Share of migrants unemployed	0.058	0.155	0.000	0.000	0.035	71,587
Share of migrants inactive	0.437	0.358	0.000	0.429	0.667	71,587
Panel C: Destination characteristics (one row per destination, most recent census)						
Urban population share of the destination	0.256	0.226	0.092	0.198	0.358	299

Notes: Descriptive statistics for the variables used in the analysis. Panel A: the main corridor-year estimation panel (8 countries, Egypt excluded; Mali 1998 included; 12-year capped windows with the census year excluded). Corridors with at least one migrant over the window are retained; 70% of corridor-years record zero urban migrant arrivals, which accounts for the zero lower quartiles of the rate variables. Rates are counts per 100 urban residents at destination; outcome rates are winsorized at the 1st and 99th percentiles by country, as in estimation; the migration rate and raw counts are not winsorized. Drought intensity is the standardized negative seasonal rainfall anomaly at origin, $\max(0, -\tilde{z})$, equal to zero by construction in wetter-than-average seasons; the cumulative version averages the current and two preceding seasons (defined where lags exist). The inactive inflow counts migrants out of the labor force with determined status. Panel B: shares among migrants (housing: denominator = migrants; employment: denominator = migrants with known labor-force status); defined only for corridor-years with a positive denominator. Panel C: census urban population share, one observation per destination at its most recent census.

Figure 2: Average Urban In-Migration Rates Over Time, by Country

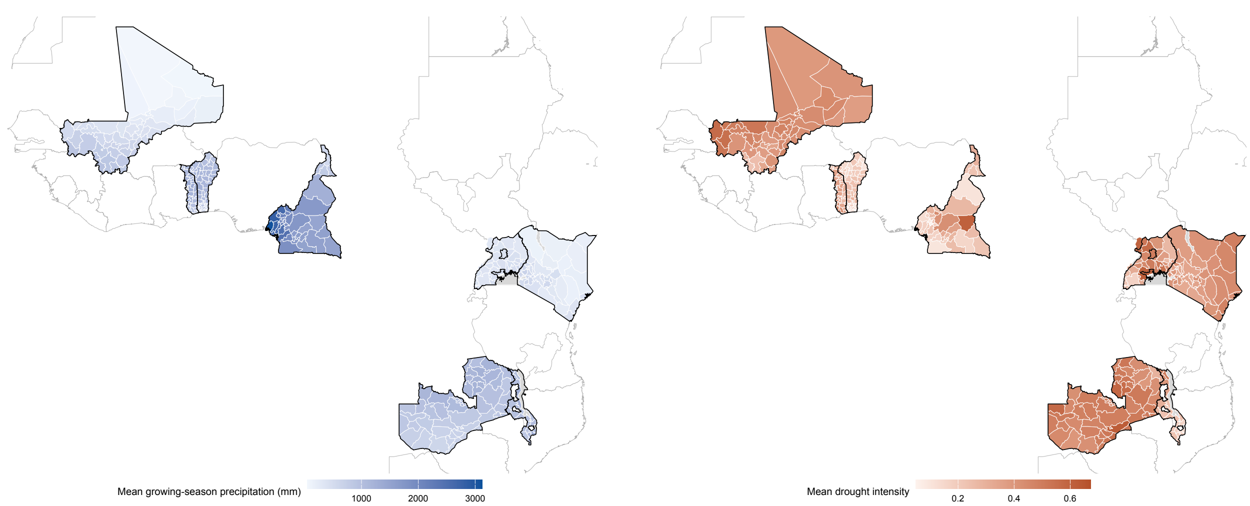


Notes: Mean urban in-migration rate (in-migrant arrivals per 100 urban residents at destination) across corridor-year observations of the yearly estimation panel, by calendar year and country. Dashed vertical lines mark census years; each census contributes a capped retrospective window of yearly migration cohorts. Within each window, rates rise mechanically toward the census year, reflecting recall and onward-migration attrition of earlier cohorts; the distance-to-census control absorbs this gradient in estimation. Y-axis scales differ across panels.

Source: Authors' computations from IPUMS census microdata.

Figure 3: Descriptive Maps: Climate at Origin

(a) Mean growing-season precipitation (b) Mean drought intensity

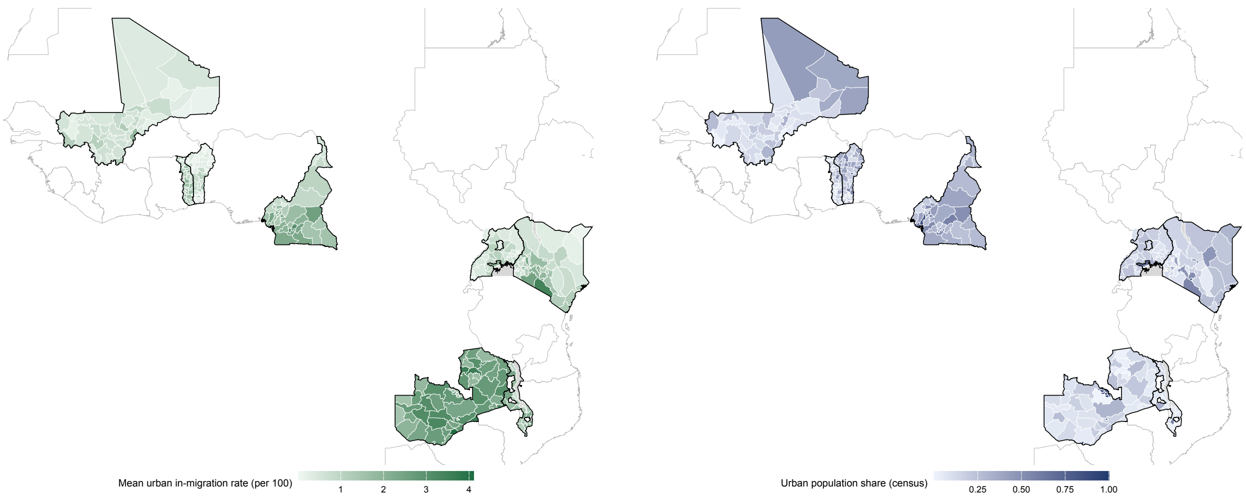


Notes: The eight countries of the estimation sample are outlined in black, with their administrative units (second level; first level for Malawi and Uganda); other African countries covered by IPUMS are outlined in grey. Panel (a): mean cumulative precipitation over the growing season (mm), by origin unit – the Precip^{g}_{cum} measure underlying the instrument. Panel (b): mean of the standardized drought intensity $\max(0, -\tilde{z})$ over the panel years, by origin unit.

Source: Authors' computations from CHIRPS data and IPUMS harmonized shapefiles.

Figure 4: Descriptive Maps: Migration and Urbanization at Destination

(a) Mean urban in-migration rate (b) Urban population share

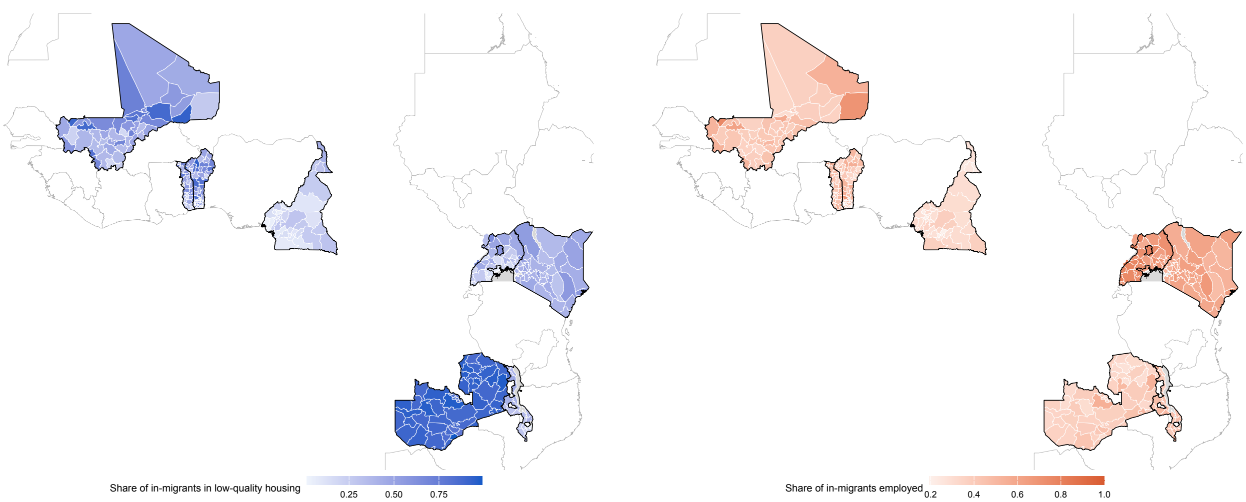


Notes: Panel (a): mean yearly urban in-migrant arrivals per 100 urban residents, by destination unit, aggregated over all origins. Panel (b): urban population share of the destination (urban over total census population, most recent census). Grey units: not observed in the estimation panel.

Source: Authors' computations from IPUMS census microdata.

Figure 5: Descriptive Maps: Migrant Outcomes at Destination

(a) Share of in-migrants in low-quality housing (b) Share of in-migrants employed



Notes: The propensity outcomes of the paper, aggregated by destination over the panel: among urban in-migrants, the share residing in low-quality housing (panel (a), main UN-Habitat definition; denominator: migrants) and the share employed (panel (b); denominator: migrants with known labor-force status). Grey units: no in-migrants observed.

Source: Authors' computations from IPUMS census microdata.

9.1.1 Descriptive Maps

9.1.2 Housing-Quality Definition

Table 8: Coding of the Housing-Quality Dimensions

Dimension	Classified as deficient
Panel A: Structural quality of the dwelling	
Floor	Unfinished natural floor – earth, sand, or dung – or salvaged wood (codes 100–120, 235)
Walls	No walls, or rudimentary, organic, or salvaged materials – cane, palm, trunks, mud, adobe, scrap (codes 100–299)
Roof	Wood, thatch, or palm; mud or earth; cardboard, plastic, or salvaged materials; other (codes 40–53, 60–61, 70–73, 90)
Panel B: Access to basic services and living conditions	
Water	Unimproved water source – unprotected wells or springs, surface water, vendors (codes 23, 25, 27–28, 29 and above)
Sanitation	No toilet (code 10)
Crowding	More than three household members per room

Notes: Coding of the six housing-quality dimensions, in harmonized IPUMS categories (codes in parentheses). The coding is identical across countries; the availability of each dimension varies by census (Egypt records no floor, wall, or roof variable and is excluded from the housing analysis; sanitation is unavailable in some censuses). Sheet-metal roofing (codes 30–39) is classified as durable in the main definition and as deficient in the robustness definition. A household resides in low-quality housing if at least one available dimension is deficient (main definition); the robustness definition requires at least two. The continuous deprivation count sums the deficient dimensions (0–6).

9.1.3 Urbanization Data

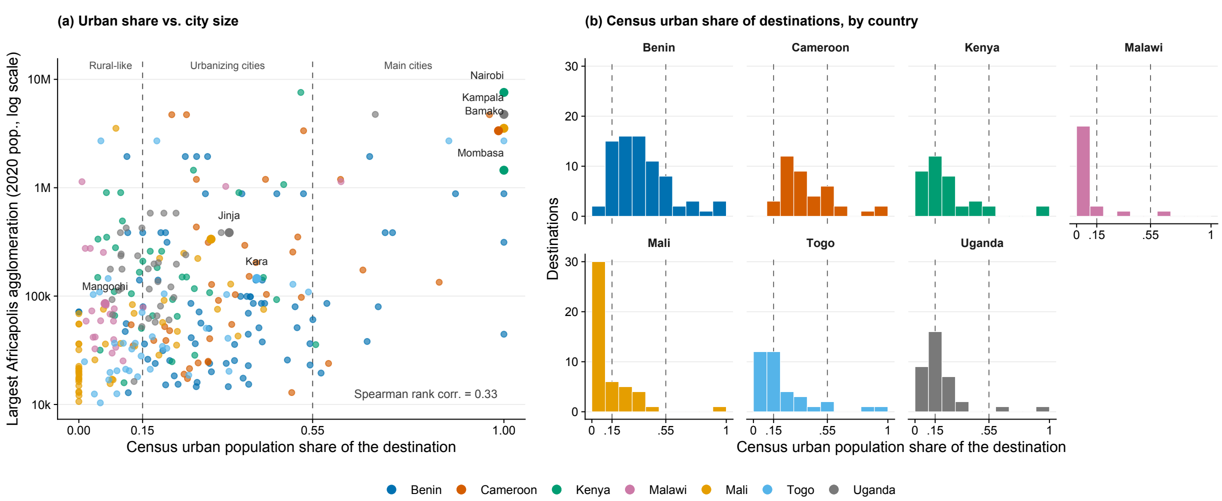
Table 9 and Figure 6 validate the census-based urbanization classification against an external measure of city size, the population of the largest Africapolis agglomeration at destination. The two measures line up where it matters: across the three bins, the median agglomeration size rises from 47,000 to 86,000 to 880,000, the share of destinations hosting a city of at least 100,000 inhabitants rises from 28 to 43 to 70 percent, and the main-cities bin contains the region's actual metropolises (Nairobi, Kampala, Yaoundé, Bamako). The destination-level rank correlation is positive but moderate (Spearman 0.33), as expected: the census share captures the urban composition of the destination's population, while Africapolis measures the physical size of the nearest agglomeration, and coarser administrative units in Malawi and Uganda mechanically dilute urban shares (panel (b)). Appendix Figure S2 reports the country-specific cumulative distributions of both measures.

Table 9: Validation of the Urbanization Classification: Census Urban Share versus Africapolis City Size

Bin (urban share)	Destinations	Corridor-years	Largest agglomeration Median ('000)	Largest agglomeration ≥ 100k (%)	Example cities
Rural-like (< 0.15)	113	64,720	47	28%	Naivasha, Arua, Molo, Kitui
Urbanizing cities ([0.15, 0.55))	155	109,456	86	43%	Nakuru, Lilongwe, Kisumu, Fort Portal
Main cities (≥ 0.55)	27	19,114	880	70%	Nairobi, Kampala, Yaounde, Bamako

Notes: Correspondence between the census urbanization bins and Africapolis city sizes. Bins: census urban population share of the destination. Agglomeration populations and names: largest Africapolis (2024) agglomeration in or within 10 km of the destination unit (2020 population; recovered boundary file, Nairobi and 15 further Kenya units included). Example cities are assigned to the bin of their own unit (highest urban share among the units they map to). Four rows carrying conurbation-cluster totals (Kisumu, Embu, Mbale, Kikima) are corrected using each agglomeration's pre-merge decadal growth (e.g., Kisumu 15.5M → 901k). Zambia lacks the census urban variable and is absent. Corridor-years from the estimation sample ($N = 193{,}290$).

Figure 6: Census Urban Share versus Africapolis City Size at Destination



Notes: Panel (a) plots one observation per destination, colored by country. The horizontal axis reports the destination's urban population share in the most recent census. The vertical axis reports the 2020 population of the largest agglomeration in the Africapolis 2024 database located within, or within 10 km of, the destination unit, on a logarithmic scale. Panel (b) shows the distribution of census urban shares across destinations, separately by country. Dashed vertical lines indicate the classification thresholds of 0.15 and 0.55. Destinations without a matched Africapolis agglomeration of at least 10,000 inhabitants are excluded from panel (a). Four Africapolis agglomeration records reporting conurbation-cluster totals are corrected as described in the notes to Table 9. Zambia is excluded because its census data do not contain the urban-status variable. The Spearman rank correlation between the two measures is 0.33. Appendix Figure S2 reports their country-specific cumulative distributions.

Source: Authors' computations using IPUMS census data and the Africapolis 2024 database.

9.2 Identification Strategy

9.2.1 Post-IV Lasso Selection

Figure 7 reports a data-driven validation of the instrument choice. We build a pool of candidate instruments - drought intensity, drought dummies at alternative thresholds, precipitation and temperature polynomials, degree-day and evapotranspiration measures, and their lags and 3-season means - and apply the post-IV-LASSO selection of Belloni et al. (2014), alongside the single-candidate first-stage F. We split the exercise between drought shocks (panels (a)-(b)) and their extreme-wet counterparts (panels (c)-(d)).

The drought-intensity instrument is the strongest single candidate on both samples ($F = 23.9$ on the scale sample, $F = 13.0$ on the propensity sample) and is selected by the LASSO in both; adding lagged candidates, its 3-season mean dominates ($F = 27.4$), consistent with the cumulative instrument used throughout. At the default penalty the LASSO retains a small set of dry-side candidates rather than a single winner (drought intensity, dry days, and extreme degree days on the scale sample); we nevertheless use the single strongest candidate, both for transparency and to keep the model just-identified, so that Anderson-Rubin inference remains exact and the complier population is not mixed across instruments. he robustness section reports the main results under the LASSO-selected instrument sets, showing that our findings are not tied to this choice.

By contrast, Figure 8 shows the results for the excess-rainfall intensity - the exact positive mirror of the drought instrument - has essentially no predictive power ($F = 4.0$ and $F = 0.03$), and the flood dummies are similarly weak. The instrument is thus data-selected, and its power comes specifically from the dry tail of the rainfall distribution.

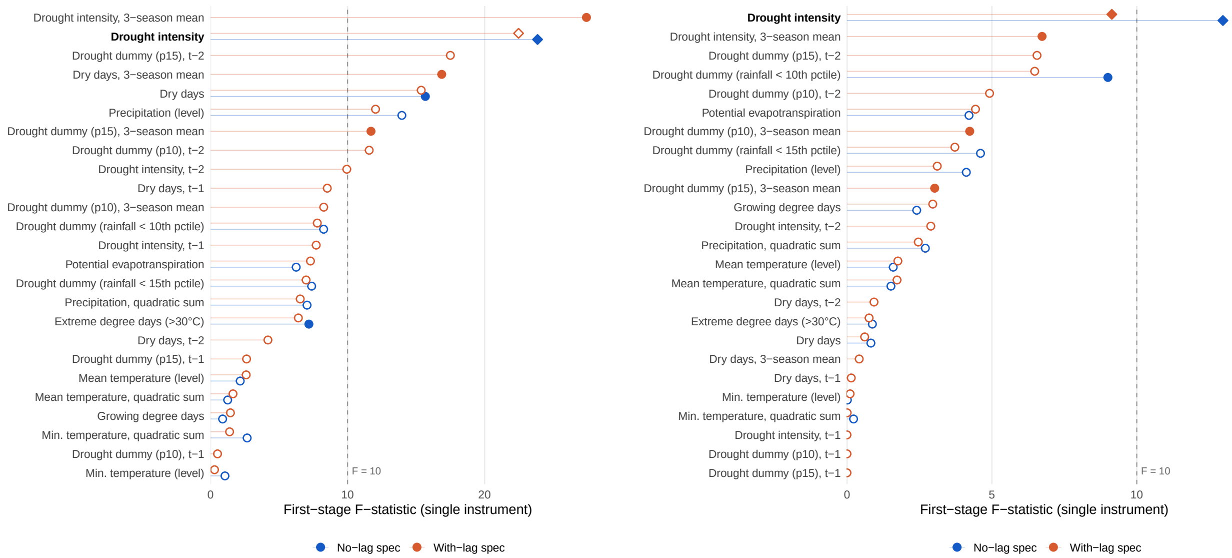
9.3 First Stage and Reduced Form

9.3.1 First Stage

Table 10 assesses the robustness of the first stage to sample restrictions and to the drought exposure window, reporting the paper's main specification for each sample and instrument; specification robustness itself is established on the full sample in Table 1. Columns (1)-(2) cover the main sample of urban migrations to all destinations, columns (3)-(4) restrict destinations to main cities,
footnoteMain cities are defined as destinations with an urban population share weakly above the threshold, $\text{share}_d^{\text{urb}} \geq 0.55$, measured at the country's most recent census wave and held fixed across waves. Rural origins are defined as the complement: locations with $\text{share}_o^{\text{urb}} < 0.55$ or with no measured urban share. Robustness checks vary this threshold and use alternative definitions and datasets. and columns (5)-(6) further restrict to moves from rural origins to main cities. Within each pair, the first column uses the contemporaneous drought instrument and the second its cumulative (three-season mean) counterpart on the lag-restricted sample.

The contemporaneous instrument is a strong and significant predictor of migration in every sample, including moves from rural origins to major cities. On the full sample, the cumulative instrument delivers an even stronger first stage ($F = 27.4$ against 23.9), but its strength collapses once the sample is restricted to city destinations (F below conventional

Figure 7: Post-IV-LASSO Instrument Selection: Drought Candidates

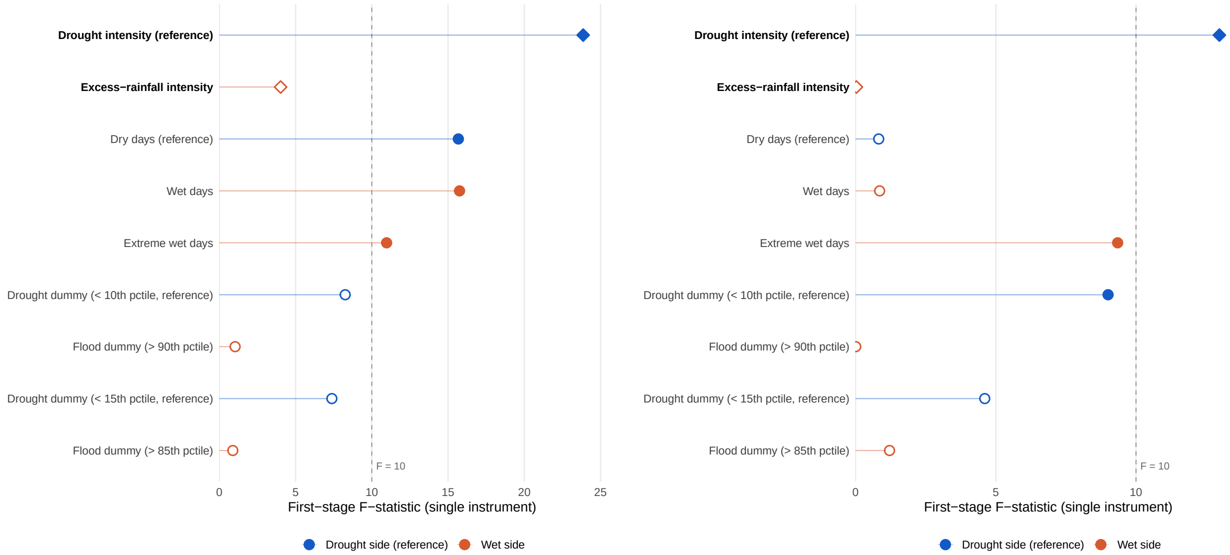


(a) Scale sample (b) Propensity sample

Notes: Each point reports the first-stage F-statistic of a single candidate instrument for the urban migration rate under the main specification, including origin–destination and year fixed effects and the distance-to-census control. Standard errors are clustered at the origin-district level. The panels compare candidate sets without and with lagged instruments for the scale and propensity samples, respectively. Filled markers denote candidates selected by post-IV-LASSO.

Source: Authors' computations using IPUMS census and CHIRPS climate data.

Figure 8: Placebo Instrument Selection: Wet-Side Mirror Candidates



(a) Scale sample (b) Propensity sample

Notes: This figure replaces each drought candidate with its wet-side mirror. Excess-rainfall intensity is defined as the positive mirror of drought intensity. Each point reports the single-instrument first-stage F-statistic under the main specification. The weaker first stages indicate that instrument strength is specific to the dry tail of the rainfall distribution.

Source: Authors' computations using IPUMS census and CHIRPS climate data.

thresholds in columns (4) and (6)). This asymmetry, together with exclusion-restriction considerations-a multi-year accumulation of drought is more likely to affect destination outcomes through channels other than the timing of migration, such as persistent income losses at origin-leads us to retain the contemporaneous instrument as the paper's main one. The cumulative instrument is nonetheless reported alongside the main results throughout, and delivers consistent conclusions wherever its first stage is strong.

Table 10: First Stage: Restricted Samples and Lagged Instrument (Main Specification)

	All destinations		*Main cities*		*Rural → main cities*	
	Contemp.	Lagged	Contemp.	Lagged	Contemp.	Lagged
	(1)	(2)	(3)	(4)	(5)	(6)
Drought intensity	0.00373***	0.00913***	0.00458***	0.00051	0.00434***	0.00415
	(0.00076)	(0.00174)	(0.00141)	(0.00481)	(0.00093)	(0.00270)
First-stage F	23.86	27.43	10.53	0.01	21.65	2.36
Effect of 1-SD drought (% mean)	+6.9%	+10.0%	+6.4%	+0.4%	+7.7%	+4.3%
R^2	0.359	0.366	0.776	0.776	0.767	0.766
Within R^2	0.006	0.006	0.031	0.030	0.053	0.051
Mean migration rate	0.027168	0.027170	0.032214	0.032214	0.025474	0.025474
Observations	245,292	239,568	19,114	19,114	16,990	16,990
OD pairs	16,888	16,888	1,462	1,462	1,290	1,290

Notes: First stage (DV: urban migration rate) at the paper's main specification (origin–destination pair and arrival-year fixed effects, distance-to-census control) in each sample and for each instrument – specification robustness is established on the full sample in Table 1. 'Contemp.': $\max(0, -\tilde{z})$ at origin; 'Lagged': its 3-season mean over $\{t, t-1, t-2\}$ (lag-restricted sample). Main cities: destinations with *share_urban_known_dest* ≥ 0.55; rural origins additionally restrict to origins below the threshold. The lagged instrument has essentially no first stage in the restricted samples – the reason the city-level analyses use the contemporaneous instrument. First-stage F: cluster-robust Wald on the instrument. Effect of 1-SD drought: coefficient $\times$ SD(instrument) $\div$ mean migration rate $\times$ 100. 8 countries, 12-year window; SEs clustered at *admin_origin*, in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

9.3.2 Reduced Form

Residential Status - This section reports the reduced-form counterparts of the paper's main estimates: each outcome is regressed directly on the drought instrument at origin, under the main specification, so that no result depends on the strength of the first stage. Table 11 reports the reduced form for the propensity margin. Columns (1)-(2) use the contemporaneous instrument and columns (3)-(4) the cumulative instrument on the lag-restricted sample; within each pair, the first column reports the share of in-migrants living in low-quality housing under the main definition (dummy) and the second the mean number of deprived housing dimensions among the same migrants (continuous). Its significance pattern mirrors the Anderson-Rubin intervals of Table 4, as the two tests coincide.

Table 11: Reduced Form: Drought Intensity and the Propensity for Low-Quality Housing

	Contemporaneous drought		*3-season mean (lagged)*	
	Deprived Dummy - Main def. (1)	Number of deprivations Continuous (2)	Deprived Dummy - Main def. (3)	Number of deprivations Continuous (4)
Drought intensity	0.00679**	0.00292	0.02360***	0.02323**
	(0.00297)	(0.00463)	(0.00622)	(0.00968)
Effect of 1-SD drought	+0.34 pp	+0.0015	+0.68 pp	+0.0067
R^2	0.555	0.543	0.550	0.543
Within R^2	0.010	0.006	0.010	0.006
Mean outcome	0.464	0.605	0.456	0.593
Observations	72,821	72,821	71,304	71,304
OD pairs	10,561	10,561	10,536	10,536

Notes: Reduced form: propensity outcomes (main UN-Habitat share; mean nb. of deprived dimensions) regressed directly on origin drought – no first-stage reliance; RF t-test = AR test of zero in the 2SLS table. Cols (3)–(4): lagged instrument, lag-restricted sample, directly interpretable. Main specification; SEs clustered at `admin_origin`. 1-SD effects in pp / dimension units. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Labor-Force Status - Tables 12 and 13 report the reduced forms for the employment results, regressing the scale and propensity outcomes directly on origin drought for both instruments. On the scale margin, all six inflows rise significantly at the 1% level. On the propensity margin, drought shifts the composition of arrivals from inactivity toward employment under both instruments; since the reduced form does not rely on the first stage, the lagged columns provide valid inference where the lagged first stage of Table 17 is weak, and confirm the same reallocation.

9.3.3 Identifying Variation

Figures 9 and 10 provide a graphical representation of the relationships underlying the identification strategy. Each panel plots mean outcomes within forty equal-sized bins of drought intensity at origin, after residualizing all variables on origin-destination pair and arrival-year fixed effects and the distance-to-census control; the fitted line is estimated on the underlying corridor-year data, so its slope equals the corresponding regression coefficient exactly, and the shaded area is the 95% confidence band based on standard errors clustered at the origin level. All relationships are positive, close to linear, and supported across the entire distribution of drought shocks rather than driven by a few extreme bins. Corridors experiencing unusually dry growing seasons send more migrants to urban areas, more migrants into low-quality housing, and more employed migrants (Figure 9), the visual counterparts of the first-stage and reduced-form estimates of Tables 1 and 2. The same holds for the propensity outcomes (Figure 10): among migrants, the shares in low-quality housing and in employment both rise with origin drought, the reduced-form counterpart of the sorting results.

Table 12: Reduced Form: Drought Intensity and the Labor-Force Status of the Inflow: Scale

	Contemporaneous drought			*3-season mean (lagged)*		
	Employed (1)	Unemployed (2)	Inactive (3)	Employed (4)	Unemployed (5)	Inactive (6)
Drought intensity	0.00127***	0.00022***	0.00101***	0.00257***	0.00076***	0.00283***
	(0.00019)	(0.00003)	(0.00022)	(0.00043)	(0.00010)	(0.00055)
Effect of 1-SD drought	+6.8%	+10.7%	+4.9%	+8.2%	+21.3%	+8.2%
R^2	0.513	0.439	0.440	0.523	0.442	0.450
Within R^2	0.008	0.003	0.010	0.008	0.004	0.011
Mean outcome	0.00934	0.00104	0.0103	0.00937	0.00106	0.0102
Observations	245,292	245,292	245,292	239,568	239,568	239,568
OD pairs	16,888	16,888	16,888	16,888	16,888	16,888

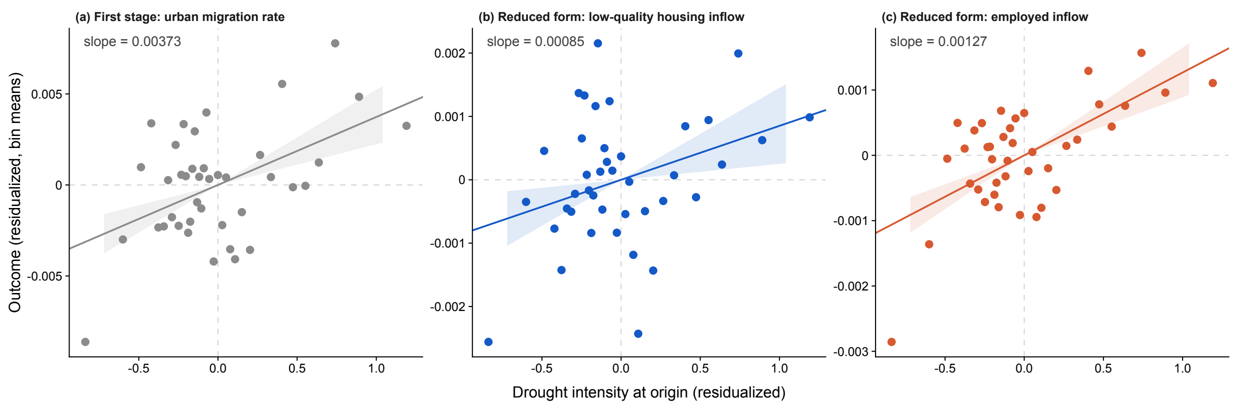
Notes: Reduced form: employed / unemployed / inactive in-migrants per 100 urban residents (winsorized 1–99% by country) regressed directly on origin drought – no first-stage reliance; the RF t-test = AR test of zero in the 2SLS table. Cols (4)–(6): lagged instrument, lag-restricted sample. Effect of 1-SD drought in % of the mean outcome. Main specification (od-pair and year FE, `dist_census` control); SEs clustered at `admin_origin`. Undetermined labor-force status excluded. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 13: Reduced Form: Drought Intensity and the Labor-Force Status of the Inflow: Propensity

	Contemporaneous drought			*3-season mean (lagged)*		
	Employed (1)	Unemployed (2)	Inactive (3)	Employed (4)	Unemployed (5)	Inactive (6)
Drought intensity	0.00850***	0.00172	-0.01022***	0.01454**	0.00480*	-0.01934***
	(0.00303)	(0.00137)	(0.00293)	(0.00603)	(0.00270)	(0.00578)
Effect of 1-SD drought	+0.43 pp	+0.09 pp	−0.51 pp	+0.42 pp	+0.14 pp	−0.56 pp
R^2	0.330	0.243	0.325	0.335	0.245	0.330
Within R^2	0.003	0.000	0.005	0.003	0.000	0.005
Mean outcome	0.505	0.0578	0.437	0.506	0.0586	0.435
Observations	71,587	71,587	71,587	70,119	70,119	70,119
OD pairs	10,476	10,476	10,476	10,445	10,445	10,445

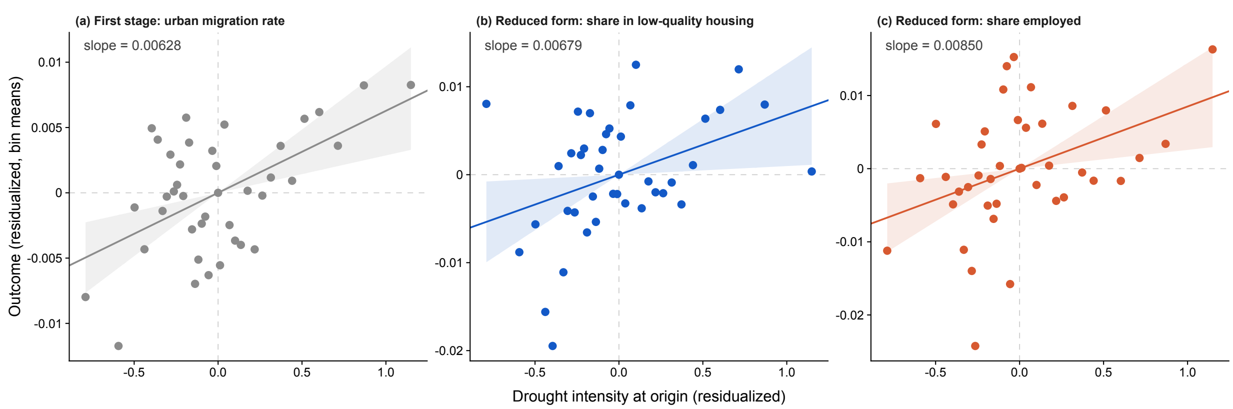
Notes: Reduced form: employed / unemployed / inactive shares among known-status in-migrants regressed directly on origin drought – no first-stage reliance; the RF t-test = AR test of zero in the 2SLS table, the honest inference for the weak lagged first stage there. The three shares sum to one, so coefficients sum to zero by construction within each instrument. Cols (4)–(6): lagged instrument, lag-restricted sample. Effect of 1-SD drought in percentage points. Main specification (od-pair and year FE, `dist_census` control); SEs clustered at `admin_origin`. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure 9: First Stage and Reduced Form: Binned Scatterplots under the Main Specification



Notes: Binned scatterplots of the first-stage (panel a) and reduced-form (panel b) relationships under the main specification. All variables are residualized on origin–destination pair and arrival-year fixed effects and the distance-to-census control (Frisch–Waugh); each dot is the mean within one of forty equal-sized bins of residualized drought intensity at origin. Panel (a) plots the urban migration rate; panel (b) the low-quality-housing inflow rate (main slum definition, winsorized at the 1st and 99th percentiles by country). The fitted line is estimated on the underlying corridor-year data and passes through the origin by construction; its slope equals the corresponding regression coefficient exactly (0.00373 in panel a, 0.00063 in panel b). The shaded area is the 95% confidence band based on standard errors clustered at the origin administrative-unit level. Sample: urban migrants, all destinations; 8 countries; 245,292 corridor-years.
Source: Authors' computations from IPUMS census and CHIRPS climate data.

Figure 10: First Stage and Reduced Form for the Propensity Outcomes: Binned Scatterplots



Notes: Binned scatterplots of the first-stage (panel a) and propensity reduced-form (panels b–c) relationships under the main specification. All variables are residualized on origin–destination pair and arrival-year fixed effects and the distance-to-census control (Frisch–Waugh); each dot is the mean within one of forty equal-sized bins of residualized drought intensity at origin. Panel (a) plots the urban migration rate; panel (b) the share of migrants in low-quality housing (main definition, denominator: migrants); panel (c) the share of migrants employed (denominator: known-status migrants). Shares are not winsorized. The fitted line is estimated on the underlying corridor-year data and passes through the origin by construction; its slope equals the corresponding regression coefficient exactly (0.00628 in panel a, 0.00679 in panel b, 0.00850 in panel c). The shaded area is the 95% confidence band based on standard errors clustered at the origin administrative-unit level. Sample: urban migrants, all destinations; 8 countries; 72,821 corridor-years in panels (a)–(b) (corridors with at least one migrant) and 71,587 in panel (c) (corridors with at least one known-status migrant).

Source: Authors' computations from IPUMS census and CHIRPS climate data.

9.4 Main Results

9.4.1 Residential Integration

Scale- Table 14 details the calculation behind the interpretation of the main estimate: for a representative corridor with 10,000 urban residents, it translates a one-standard-deviation drought at origin into additional migrants and additional residents in low-quality housing, whose ratio equals, by construction, the 2SLS coefficient of Table 3. Table 15 replicates the main table under the two sample restrictions used in the paper - destinations limited to main cities, and rural origins moving to main cities - with the contemporaneous instrument, whose first stage remains strong in both samples.

Table 14: Interpreting the Main Estimate: A Representative Corridor

Panel A. Step-by-step calculation		
	Formula	Value
Baseline (no drought)		
Migrants	$\overline{\text{MigRate}}/100 \times 10{,}000$	$0.02717 \times 100 = 2.72$
Migrants in low-quality housing	$\overline{S}/100 \times 10{,}000$	$0.01300 \times 100 = 1.30$
Effect of a 1-SD drought shock		
Δ Migration rate	$\beta_{FS} \times SD(\text{drought})$	$0.00373 \times 0.503 = 0.00188$
Δ Migrants	$\Delta\text{MigRate}/100 \times 10{,}000$	+0.19
Δ Low-quality housing inflow (S)	$\beta_{2SLS} \times \beta_{FS} \times SD(\text{drought})$	$0.228 \times 0.00373 \times 0.503 = 0.00043$
Δ Migrants in low-quality housing	$\Delta S/100 \times 10{,}000$	+0.043
Share of new migrants in low-quality housing		
Δ Low-quality / Δ Migrants	β_{2SLS}	0.043/0.19 = 23%

Panel B. Summary				
	Without drought	With 1-SD drought	Change	Share of new migrants in low-quality housing
Number of migrants	2.72	2.90	+0.19 (+6.9%)	–
Migrants in low-quality housing	1.30	1.34	+0.043 (+3.3%)	23%

Notes: Illustration for a single origin–destination corridor with 10,000 urban residents at destination (all-destinations sample, main definition of low-quality housing). Notation as in the text: MigRate is the urban migration rate and S the low-quality housing inflow rate, both per 100 urban residents; bars denote sample means. Because S and MigRate share the same denominator (destination urban population), β_{2SLS} reads directly as the share of marginal, drought-induced migrants who settle in low-quality housing – about 23 of every 100. $\beta_{FS} = 0.00373$, $\beta_{2SLS} = 0.22798$, $SD(\text{drought}) = 0.503$; all values from Table 3.

9.4.2 Propensity

Propensity - Table 16 re-estimates the propensity outcomes - the share of in-migrants in low-quality housing under the main definition and the mean number of deprived dimensions - in the two restricted samples, with the contemporaneous instrument. The propensity

Table 15: Drought-Induced Migration and Low-Quality Housing Inflows: Main Cities and Rural-Origin Samples

	Main cities			*Rural origins → main cities*		
	OLS	FS	2SLS	OLS	FS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Urban migration rate	0.10960***		0.16744***	0.18824***		0.14702***
	(0.02335)		(0.04944)	(0.02419)		(0.04476)
Drought intensity		0.00458***			0.00434***	
		(0.00141)			(0.00093)	
First-stage F			10.53			21.65
AR 95% CI			[0.061, 0.305]			[0.046, 0.237]
Effect of 1-SD drought (% mean)	+2.9%	+6.4%	+4.4%	+5.2%	+7.7%	+4.0%
Origin-Destination FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Dist. to census control	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.839	0.776	0.822	0.869	0.767	0.864
Within R^2	0.270	0.031	0.195	0.400	0.053	0.381
Mean outcome		0.007863			0.007081	
Mean migration rate		0.032214			0.025474	
SD(drought intensity)		0.449			0.449	
Observations		19,114			16,990	
OD pairs		1,462			1,290	

Notes: Contemporaneous instrument throughout; the cumulative instrument is omitted, as its first stage collapses in these samples (see the first-stage appendix table). Columns (1)–(3): destinations restricted to main cities (*share_urban_known_dest* $\geq$ 0.55 at the country's most recent census). Columns (4)–(6): additionally, rural origins only. OLS and 2SLS columns report the effect of the urban migration rate on the low-quality housing inflow rate (main UN-Habitat definition, winsorized 1–99% by country); FS columns report the sample-specific first stage of the migration rate on drought intensity. First-stage F: cluster-robust Wald on the instrument. AR 95% CI: cluster-robust Anderson–Rubin confidence interval. Effect of 1-SD drought: for OLS/2SLS columns, coefficient $\times$ first-stage coefficient $\times$ SD(instrument) $\div$ mean outcome $\times$ 100; for FS columns, coefficient $\times$ SD(instrument) $\div$ mean migration rate $\times$ 100. 8 countries (Egypt excluded), 12-year window, Mali 1998 included. SEs clustered at *admin_origin*, in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

subsamples are considerably smaller than for the scale outcome, since they require corridors with observed housing characteristics: in the main-cities sample the first stage is weak ($F = 3.4$), so only the Anderson-Rubin intervals are informative there, while in the rural-to-cities sample the first stage is strong ($F = 13.8$). The 2SLS point estimates on the extensive margin (1.13 and 1.29) are close to the full-sample estimate of 1.08, but the confidence intervals are wide and include zero: the subsample results are consistent with the main propensity result, though too imprecise to establish it independently.

Table 16: Propensity for Low-Quality Housing: Main Cities and Rural-Origin Samples

	Main cities				*Rural origins → main cities*			
	Deprived Dummy - Main def.		Number of deprivations Continuous		Deprived Dummy - Main def.		Number of deprivations Continuous	
	OLS (1)	2SLS (2)	OLS (3)	2SLS (4)	OLS (5)	2SLS (6)	OLS (7)	2SLS (8)
Urban migration rate	0.05626**	1.13021	0.06344*	0.10967	0.12736**	1.29137	0.15541**	0.48288
	(0.02254)	(1.41542)	(0.03391)	(1.72605)	(0.05550)	(1.30591)	(0.06955)	(1.59140)
First-stage F		3.42		3.42		13.75		13.75
AR 95% CI		[-15.80, 18.06]		[-20.54, 20.76]		[-1.31, 4.70]		[-3.06, 4.18]
Effect of 1-SD drought	+0.01 pp	+0.20 pp	+0.0001	+0.0002	+0.02 pp	+0.25 pp	+0.0003	+0.0009
R^2	0.458	0.408	0.486	0.486	0.459	0.435	0.489	0.488
Within R^2	0.026	-0.063	0.019	0.019	0.026	-0.018	0.021	0.019
Mean outcome	0.276	0.276	0.327	0.327	0.276	0.276	0.323	0.323
Observations	11,017	11,017	11,017	11,017	9,692	9,692	9,692	9,692
OD pairs	1,163	1,163	1,163	1,163	1,016	1,016	1,016	1,016

Notes: Propensity (composition) outcomes in the restricted samples, contemporaneous instrument only (the cumulative instrument has no first stage in these samples; see the first-stage appendix table). Columns (1)–(4): destinations restricted to main cities (*share_urban_known_dest* ≥ 0.55); columns (5)–(8): additionally, rural origins only. Within each sample, the first pair of columns reports the share of in-migrants living in low-quality housing under the main definition (dummy) and the second pair the mean number of deprived housing dimensions (continuous); denominator: migrants, unweighted, not winsorized. Main specification (origin–destination pair and arrival-year FE, distance-to-census control); SEs clustered at *admin_origin*, in parentheses. First-stage F: cluster-robust Wald on the instrument, on each column's estimation sample. AR 95% CI: cluster-robust Anderson–Rubin confidence interval. Effect of 1-SD drought: coefficient × first-stage coefficient × SD(instrument), in pp of the share (dummy) or dimension units (continuous). 8 countries, 12-year window. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

9.4.3 Labor Market Integration

Table 17 re-estimates the employment results of Tables 5 and 6 with the 3-season cumulative drought instrument. On the scale margin the first stage is strong ($F = 27.4$) and the employed, unemployed, and inactive inflows all rise significantly. On the propensity margin the first stage is weak ($F = 7.1$), so inference relies on the Anderson-Rubin intervals, which confirm the contemporaneous pattern: the employed share of arrivals rises, the inactive share falls, and the unemployed share is unchanged.

Table 17: Drought-Induced Migration and the Labor-Force Status of the Inflow: Lagged Instrument

	Scale Inflows per 100 urban residents			*Propensity* Shares of known-status in-migrants		
	Employed	Unemployed	Inactive	Employed	Unemployed	Inactive
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Urban migration rate	0.28188***	0.08347***	0.31042***	1.49204*	0.49241	-1.98445**
	(0.03191)	(0.01390)	(0.03130)	(0.81025)	(0.32886)	(0.91797)
First-stage F	27.43	27.43	27.43	7.08	7.08	7.08
AR 95% CI	[0.227, 0.363]	[0.062, 0.123]	[0.250, 0.383]	[0.290, 5.989]	[-0.050, 2.142]	[-7.630, -0.745]
Effect of 1-SD drought	+8.2%	+21.3%	+8.2%	+0.42 pp	+0.14 pp	−0.56 pp
R^2	0.252	-1.091	0.454	0.049	0.098	-0.174
Within R^2	-0.556	-2.731	0.018	-0.425	-0.195	-0.743
Mean outcome	0.00937	0.00106	0.0102	0.506	0.0586	0.435
Observations	239,568	239,568	239,568	70,119	70,119	70,119
OD pairs	16,888	16,888	16,888	10,445	10,445	10,445

Notes: 2SLS estimates with the 3-season cumulative drought instrument on the lag-restricted sample. Columns (1)–(3): employed / unemployed / inactive in-migrants per 100 urban residents, winsorized 1–99% by country; the effect of a 1-SD drought is in % of the mean. Columns (4)–(6): shares among known-status in-migrants (summing to one, so coefficients sum to zero by construction); the effect of a 1-SD drought is in percentage points. The propensity-side first stage is weak ($F \approx 7$): the lagged point magnitudes in columns (4)–(6) are not interpretable and inference should rely on the Anderson-Rubin intervals. Main specification (od-pair and year FE, `dist_census` control); SEs clustered at `admin_origin`. Undetermined labor-force status excluded. Sample: urban migrants, all destinations; 8 countries, 12-year window. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

9.4.4 Interpreting the effect

Decomposing the Effect for housing - Table 18 details the decomposition of the additional low-quality housing inflow from a one-standard-deviation drought into its scale and propensity components, reporting the formula, inputs, and implied contribution of each channel - roughly 90% through the number of migrants and 10% through their propensity to live in low-quality housing.

Table 18: Decomposition of the Drought Effect on Low-Quality Housing Inflows: Scale versus Propensity

Channel	Formula	Calculation	Per 10,000 residents	Share
Scale *(more migrants, same propensity)*	$\Delta\text{MigRate} \times \overline{P}$	$0.00188 \times 0.464 = 0.00087$	+0.087	$\sim 90\%$
Propensity *(same migrants, higher propensity)*	$\overline{\text{MigRate}} \times \Delta P$	$0.02717 \times 0.0034 = 0.00009$	+0.009	$\sim 10\%$
Total	Scale + Propensity	0.00096	+0.096	100%

Notes: Decomposition of the additional low-quality housing inflow from a one-standard-deviation drought, $\Delta S \approx \Delta\text{MigRate} \times \overline{P} + \overline{\text{MigRate}} \times \Delta P$, expressed per 100 urban residents at destination (the "Calculation" column) and per 10,000 residents in the representative corridor. Inputs: $\Delta\text{MigRate} = \beta_{FS} \times SD(\text{drought}) = 0.00373 \times 0.503 = 0.00188$ and $\overline{\text{MigRate}} = 0.02717$ from Table 3; $\overline{P} = 46.4\%$, the baseline share of in-migrants in low-quality housing, and $\Delta P = +0.34$ percentage points, the effect of a one-standard-deviation drought on that share, from Table 4. The implied total (+0.096 migrants per 10,000 residents) exceeds the direct estimate of +0.043 in Table 14 because the identity evaluates the scale channel at the average migrant's propensity for low-quality housing (46.4%) rather than the marginal migrant's (23%), and the two channels are estimated on different samples – the propensity margin requires migrants with observed housing characteristics. The decomposition therefore measures the relative importance of the two channels, not the level of the effect.

Working Poor - Table 19 splits the drought-induced inflow into four employment × housing cells, measured on the same individuals. About two-thirds of the migrants entering low-quality housing arrive employed (0.124/0.187), against 42% of the baseline inflow (mean outcomes, columns (1) and (3)): the marginal residents of low-quality housing are predominantly working poor rather than idle poor.

9.5 Robustness

9.5.1 Housing Quality Gradient

9.6 Model

Figure 11 illustrates the two-state housing integration framework of Section 7.

Table 19: Drought-Induced Migration across Employment and Housing Cells: The Working Poor

	Employed		*Not employed*	
	Low-quality housing 2SLS (1)	Not low-quality housing 2SLS (2)	Low-quality housing 2SLS (3)	Not low-quality housing 2SLS (4)
Urban migration rate	0.12370***	0.20234***	0.06332*	0.26368***
	(0.02326)	(0.03666)	(0.03475)	(0.04737)
First-stage F	23.86	23.86	23.86	23.86
AR 95% CI	[0.084, 0.185]	[0.142, 0.301]	[-0.021, 0.128]	[0.189, 0.397]
Mean outcome	0.004362	0.004631	0.006128	0.004924
Observations	245,292	245,292	245,292	245,292

Notes: Outcomes: in-migrants per 100 urban residents in the four employment $\times$ low-quality housing cells (main definition), measured on the same individuals – migrants with both labor-force status and housing observed (94.6% of urban migrants) – and winsorized 1–99% by country. Column (1) is the "working poor" cell. 2SLS: migration instrumented by contemporaneous drought. Main specification (od-pair and year FE, `dist_census` control); SEs clustered at `admin_origin`. The estimates describe the corridor-level composition of the marginal inflow, not an individual-level causal link between employment and housing. Column (3) is imprecise: its AR interval includes zero. Sample: urban migrants, all destinations; 8 countries, 12-year window. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 20: Robustness: Scale and Propensity across Definitions of Low-Quality Housing

	Scale (2SLS, inflow per 100 residents)		*Propensity (2SLS, share of migrants)*	
Definition of low-quality housing	Coef.	AR 95% CI	Coef.	AR 95% CI
≥1 dimension, metal roofs deprived	0.73014***	[0.59, 0.96]	0.57441*	[-0.02, 1.61]
	(0.08370)		(0.33726)	
≥2 dimensions, metal roofs deprived (conservative)	0.16875***	[0.08, 0.26]	0.49407	[-0.44, 1.84]
	(0.04124)		(0.48526)	
≥1 dimension, metal roofs durable (main)	0.22798***	[0.09, 0.36]	1.08140**	[0.17, 2.80]
	(0.06352)		(0.54054)	
≥2 dimensions, metal roofs durable	0.01100	[-0.04, 0.05]	-0.27568	[-1.13, 0.38]
	(0.02088)		(0.32278)	
Mean nb. of deprived dimensions (metal durable)	0.000757	[-0.01, 0.01]	0.46559	[-1.13, 2.31]
	(0.00392)		(0.74152)	

Notes: Both channels re-estimated under each definition (contemporaneous instrument, main specification; SEs clustered at `admin_origin`). Scale: inflow per 100 urban residents, $F = 23.86$, N = 245,292. Propensity: share among migrants, $F = 12.97$, N = 72,821. Main definition: UN-Habitat criterion (≥1 deprivation, metal roofing durable). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure 11: The Two-Margin Integration Model: Employment versus Housing

Notes: Each panel is a two-state Roy diagram for one integration margin; each migrant draws state-specific capacities (S_L, S_H). The two panels show distinct, margin-specific simulated draws, positively correlated across margins. Absent barriers, migrants choose the high-integration state above the ray OC with slope w_L/w_H (blue dashed); OC is the cutoff of baseline migrants, who face their own barriers τ^0, so the hatched fan represents the differential misallocation $\tau - \tau^0$. Panel (a), employment margin: the low state is valued at the shadow value of non-market time; the employment results imply no differential barrier for drought-induced migrants ($\tau_E \approx 0$), so the cutoff is unchanged and there is no trapped region. Panel (b), housing margin: a residential barrier $\tau_R > 0$ steepens the cutoff to the ray OE with slope $w_L/[(1 - \tau_R)w_H]$ (red dotted); the hatched fan between OE and OC contains the migrants whom the barrier misallocates - they clear the employment margin but remain in low-quality housing, the barrier-induced working poor of Appendix Table 19. This reading presumes positive correlation of capacities across margins, and the position of Region II above the 45° line presumes $(1 - \tau_R)w_H < w_L < w_H$. The hatched region captures the barrier-induced working poor only: migrants below OC may also be employed and badly housed, but optimally so. The 45° line indicates equal capacities across states.

Source: Authors' illustration.

S1 Supplementary Appendix

S1.1 Descriptive Statistics

S1.1.1 Migration Data: Fifteen-Country Sample

Figure S1 reports average in-migration rates over time for the full set of fifteen IPUMS countries, using five-year migration brackets for cross-country comparability; the yearly rates for the eight-country estimation sample are in Appendix Figure 2.

S1.1.2 Urbanization Data

Figure S2 reports the cumulative distributions of the census urban share and of Africapolis city size at destination, separately by country. The height of a country's curve at the 0.15 and 0.55 thresholds gives the share of its destinations in each urbanization bin; the curves confirm that the lower bins are populated by all countries rather than driven by any single one, with Malawi and Mali most concentrated at low urban shares.

S1.2 Scale versus Propensity: A Stylized Example

Figure S3 illustrates the two channels through which drought-induced migration can raise the low-quality housing inflow, using the language of the identification strategy. Migrants who move regardless of the shock (always-takers) settle in low-quality housing at some baseline rate. A drought shock adds compliers - the drought-induced migrants of our 2SLS estimates. Under the *scale* channel alone (middle row), compliers sort into housing exactly like always-takers: the share of migrants in low-quality housing is unchanged, and only the number of arrivals rises - a congestion problem, calling for housing supply where flows land. Under the *propensity* channel (bottom row), compliers are additionally more likely to end up in low-quality housing, shifting the composition of arrivals - a vulnerability problem, calling for support targeted at the marginal migrants themselves. The numbers in the figure are illustrative; in our estimates, the baseline share is 46 percent, the marginal drought-induced migrant is about 3 percentage points more likely to be in low-quality housing, and the propensity channel accounts for roughly 10 percent of the total effect (Section 4.3.1).

S1.3 Heterogeneity

Table S1 reports the estimates underlying Figure 1: the 3-bin interacted IV by destination urbanization, for the low-quality housing and employed inflows. For housing, the effect peaks in secondary cities - about 25 of every 100 marginal migrants enter low-quality housing there, against 12 in rural-like and 17 in megacity destinations - and the secondary-bin coefficient differs significantly from the other two bins pooled ($p = 0.047$) and from rural-like destinations ($p = 0.031$). For employment, the coefficients are statistically indistinguishable across destination types (all contrasts $p > 0.55$). The rural-like bin is weakly identified (Sanderson–Windmeijer conditional $F = 8.3$), so its Anderson–Rubin interval carries the inference; the first stage, common to both outcomes, is otherwise strong.

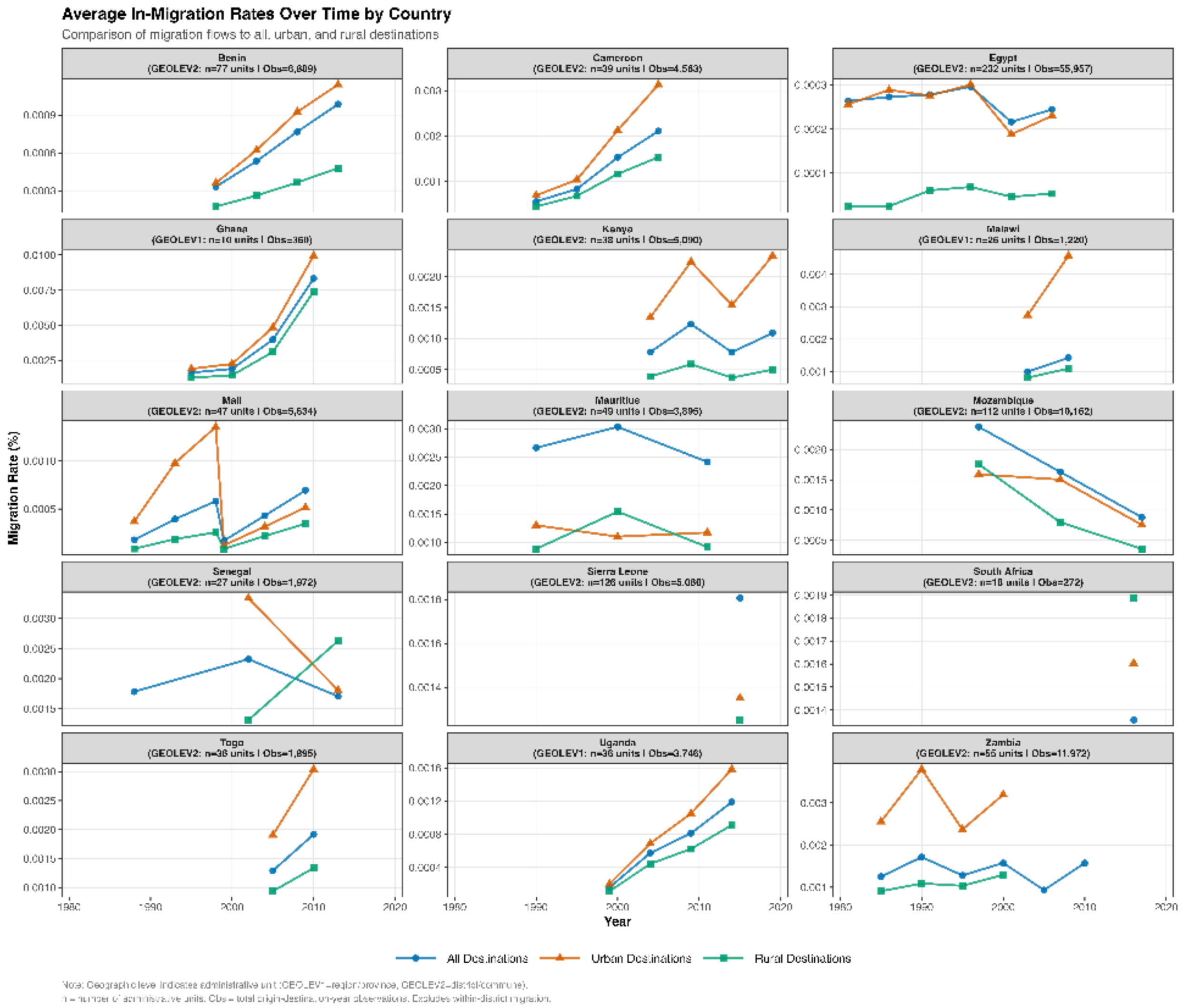
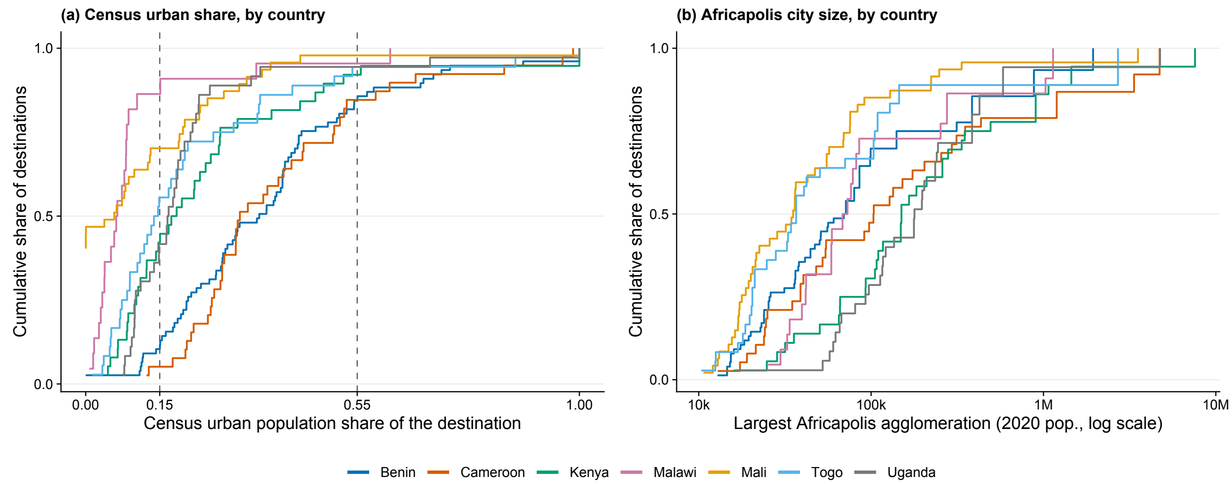


Figure S1: Average In-Migration Rates Over Time by Country.

Notes: This figure displays mean in-migration rates across administrative units for 15 African countries, disaggregated by destination type. Migration flows are calculated from IPUMS census microdata, measuring the number of in-migrants (weighted by `perwt`) divided by the destination population, aggregated into 5-year brackets. The panel titles show the geographic level (GEOLEV1 = region/province, GEOLEV2 = district/commune), the number of administrative units, and total observations. Note that Y-axis scales vary across countries to facilitate within-country comparisons. Countries with a single census year (Sierra Leone, South Africa) provide limited temporal variation. The figure reveals substantial heterogeneity in migration patterns, with urban destinations consistently attracting higher rates than rural areas, particularly in rapidly urbanizing countries like Ghana, Togo, and Uganda.

Source: Authors' elaboration based on IPUMS census data.

Figure S2: Distributions of Census Urban Share and Africapolis City Size, by Country



Notes: The figure reports empirical cumulative distribution functions across destinations, separately by country. Panel (a) shows the destination's urban population share in the most recent census. At each dashed threshold, 0.15 and 0.55, the height of a country's curve gives the proportion of its destinations below that threshold. Panel (b) shows the 2020 population of the largest agglomeration in the Africapolis 2024 database located within, or within 10 km of, each destination unit, on a logarithmic scale. Destinations without a matched Africapolis agglomeration of at least 10,000 inhabitants are excluded from panel (b). Four Africapolis agglomeration records reporting conurbation-cluster totals are corrected as described in the notes to Table 9. Zambia is excluded because its census data do not contain the urban-status variable.

Source: Authors' computations using IPUMS census data and the Africapolis 2024 database.

Table S1: Heterogeneity by Destination Urbanization: 3-Bin Interacted IV Estimates

	Low-quality housing inflow			*Employed inflow*		
	Rural-like†	Urbanizing cities	Main cities	Rural-like†	Urbanizing cities	Main cities
	(1)	(2)	(3)	(4)	(5)	(6)
Urban migration rate	0.1183**	0.2460***	0.1674***	0.3454***	0.3939***	0.4260***
	(0.0525)	(0.0405)	(0.0494)	(0.0897)	(0.0455)	(0.0950)
SW conditional F	8.27	25.36	10.55	8.27	25.36	10.55
AR 95% CI	[-0.002, 0.280]	[0.160, 0.330]	[0.061, 0.305]	[0.218, 0.786]	[0.310, 0.502]	[0.288, 0.817]
Mean outcome	0.0077	0.0057	0.0079	0.0099	0.0074	0.0130
Observations	64,720	109,456	19,114	64,720	109,456	19,114
Urbanizing − rural-like		$p = 0.031$			$p = 0.611$	
Urbanizing − main cities		$p = 0.113$			$p = 0.748$	
Urbanizing − rest		$p = 0.047$			$p = 0.669$	

Notes: 3-bin interacted IV: the urban migration rate interacted with each destination-urbanization bin (share urban at the most recent census; zones: rural-like < 0.15, urbanizing cities $[0.15, 0.55)$, main cities ≥ 0.55), instrumented by contemporaneous drought intensity interacted with the same indicators; origin–destination and arrival-year × bin fixed effects, bin-specific `dist_census` control, SEs clustered at `admin_origin`. Outcomes per 100 urban residents, winsorized 1–99% by country. SW F: Sanderson–Windmeijer conditional first-stage F; † weak first stage ($F < 10$) – the AR interval carries the inference there. The first stage is common to both outcomes (identical sample); per-bin coefficients 5.05 / 2.92 / 4.58 ($\times 10^{-3}$). Contrast p-values from the same system's clustered covariance; "rest" pools the rural-like and main-cities bins. Sample: corridors with non-missing destination urbanization (N = 193,290). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure S3: Scale versus Propensity: More Migrants, or Different Migrants?

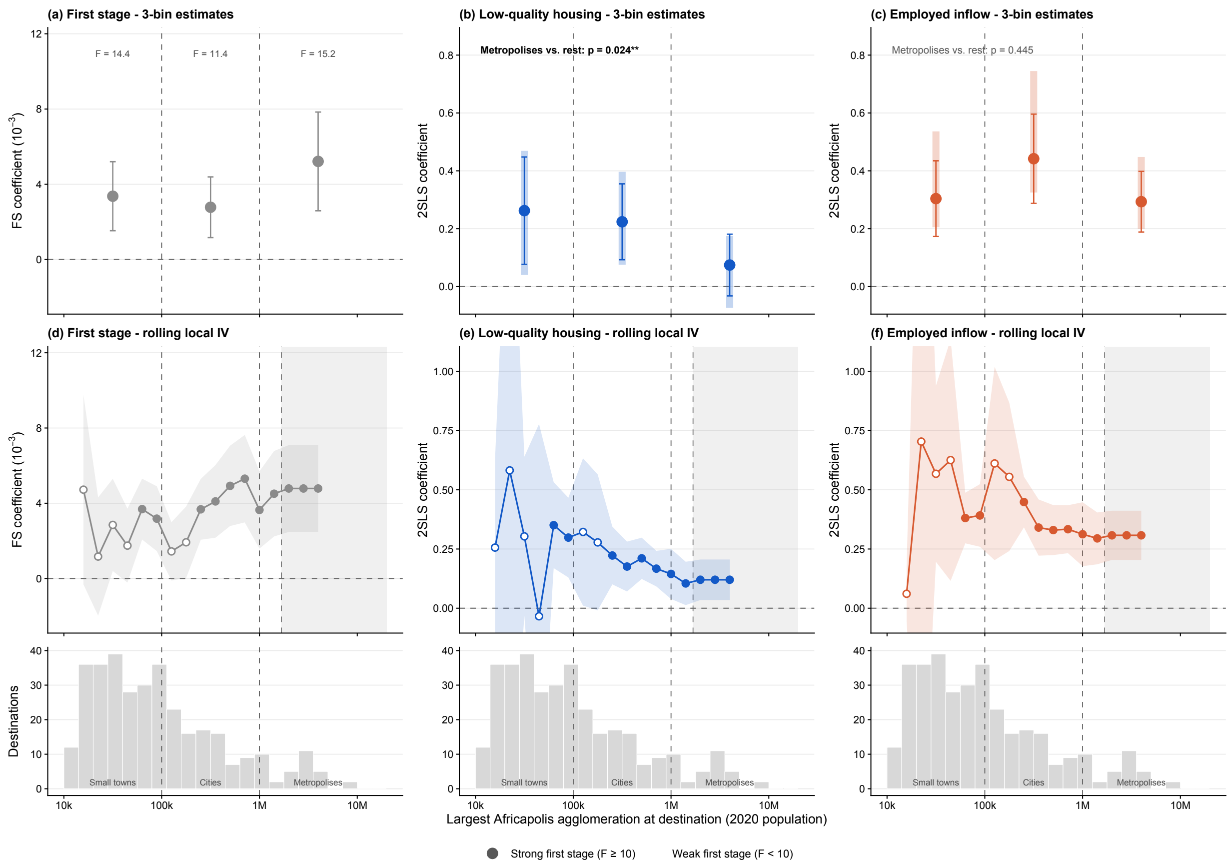
Notes: Stylized example. Grey dots: always-takers (migrants who move regardless of the drought shock); yellow dots: compliers (drought-induced migrants); green outline: higher-quality housing; orange outline: low-quality housing. Top row: no shock - always-takers settle in low-quality housing at the baseline rate. Middle row: a drought shock adds compliers who sort like always-takers - the scale channel; the share in low-quality housing is unchanged. Bottom row: compliers are additionally more likely to end up in low-quality housing - the propensity channel; the share among all migrants rises. Shares are illustrative; see Section 4.3.1 for the estimated counterparts.

Source: Authors' illustration.

Figure S4 re-examines the destination heterogeneity using a measure independent of the census urban share: destinations are classified by the population of the largest Africapolis agglomeration located in or within 10 km of the destination unit - small towns (< 100k), cities (100k–1M), and metropolises ($\geq$ 1M). This classification also extends the sample: it does not require the census urban variable, so Zambia (Lusaka and the Copperbelt) re-enters, and Nairobi and 15 further Kenya units missing from the harmonized Africa shapefile are recovered ($N = 244{,}092$ corridor-years; all three groups have strong first stages, F between 11 and 15). The pattern confirms the main classification: drought-induced migrants sort into low-quality housing in small-town and city destinations - about 26 and 22 of every 100 marginal migrants - but not in the metropolises, where the effect is close to zero (0.075) and significantly below the rest ($p = 0.024$). The rolling estimates trace this as a continuous decline in the housing effect as the destination's principal city grows, while the employed inflow is again flat across all destination sizes (all contrasts $p > 0.11$). The null effect in the largest cities is thus not an artifact of the urban-share measure, and it is robust to the size threshold defining the top group.

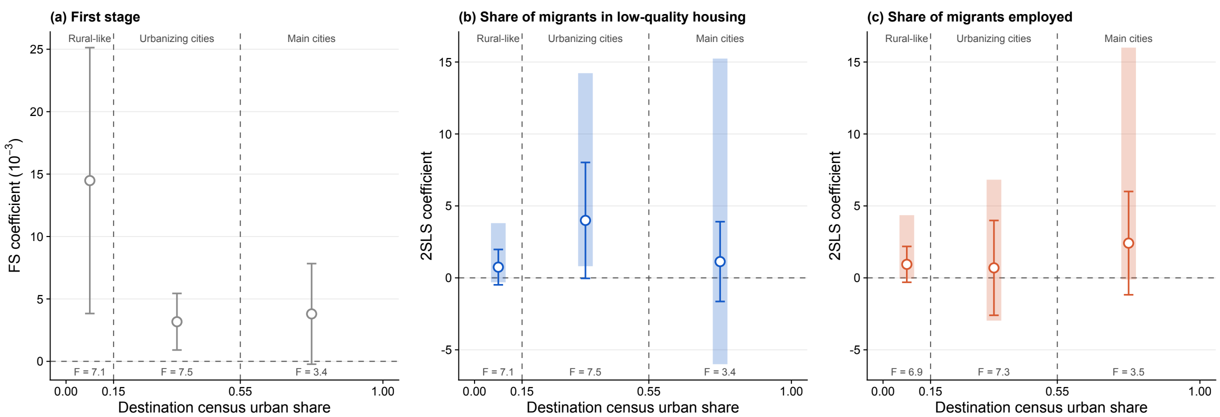
Figure S5 reports the urbanization heterogeneity for the propensity (sorting) margin.

Figure S4: Heterogeneity by Destination City Size (Africapolis): Low-Quality Housing and Employment



Notes: Heterogeneity by destination city size: destinations are classified by the population of the largest Africapolis (OECD/SWAC, 2024) agglomeration located in or within 10 km of the destination unit (2020 population; zones: small towns < 100k, cities 100k–1M, metropolises ≥ 1M). Unlike the census urban-share classification, this sample includes Zambia (no census urban variable) and recovers Nairobi and 15 further Kenya units via the harmonized boundary file used for the climate extraction ($N = 244{,}092$ corridor-years). Panels (a)–(c): 3-bin saturated interacted-IV estimates (migration $\times$ bin instrumented by drought $\times$ bin, origin–destination and arrival-year $\times$ bin fixed effects, bin-specific distance-to-census control, SEs clustered at the origin district), plotted at zone midpoints; whiskers: 95% Wald intervals, shaded bars: Anderson–Rubin 95% intervals. Pairwise contrasts, housing: cities $-$ metropolises $p = 0.060$, small towns $-$ metropolises $p = 0.061$, small towns $-$ cities $p = 0.76$; all employment contrasts insignificant. Panels (d)–(f): rolling-window local IV along the logarithm of destination city size (main specification on corridors into the $\sim$60 destinations nearest each grid point; pointwise 95% Wald bands; indicative only). The shaded zone marks where the window no longer slides; the binned estimates carry the inference there. Hollow points: weak first stage (Sanderson–Windmeijer conditional $F < 10$). Housing and employment panels share a common coefficient axis; Anderson–Rubin intervals and rolling bands extending beyond it are truncated. Bottom row: distribution of destination city sizes.
Source: Authors' computations using IPUMS census, CHIRPS climate, and Africapolis 2024 data.

Figure S5: Heterogeneity by Destination Urbanization: Migrant Propensity



Notes: The figure reports estimates from a three-bin, fully interacted IV specification examining heterogeneity in migrant propensity by destination urbanization. Propensity outcomes are defined only for corridor-years with positive migrant flows; the samples contain $N = 55{,}698$ observations for housing and $N = 55{,}323$ for employment. Panel (a) reports the first-stage estimates and the corresponding Sanderson–Windmeijer conditional F-statistics. The conditional F-statistic is below 10 in each bin, as indicated by hollow markers; estimates in panels (b) and (c) should therefore be interpreted cautiously. Whiskers report conventional 95% Wald confidence intervals, while shaded bars report weak-instrument-robust 95% Anderson–Rubin confidence intervals. Intervals extending beyond the displayed axis range are truncated. The Anderson–Rubin interval excludes zero only for the share of migrants living in low-quality housing in urbanizing cities, with a confidence interval of $[0.81, 14.2]$. This result is consistent with the scale heterogeneity reported in Figure 1, although none of the pairwise contrasts is statistically significant. The specification otherwise matches that used in Figure 1.

Source: Authors' computations using IPUMS census and CHIRPS climate data.