



Urban Agglomeration and Income Inequality: Is Kuznets Hypothesis Valid for Sub-Saharan Africa?

Isaiah Maket¹ · Izabella Szakálné Kanó¹ · Zsófia Vas¹

Accepted: 15 September 2023 / Published online: 15 October 2023
© The Author(s) 2023

Abstract

This paper contributes to the ongoing debate on the relationship between urban agglomeration and income inequality. The World Bank and the United Nations place Sub-Saharan Africa among the leading urbanizing regions with sizable urban agglomeration inequality challenges. Therefore, the main research question of this study was whether there is a significant relationship between urban agglomeration and income inequality. This study also aimed to determine whether the relationship is nonlinear, estimated using a dynamic panel model, an inverted U-shaped Kuznets hypothesis, and balanced panel data from 2000 to 2020 for 22 countries in Sub-Saharan Africa. The findings revealed a nonlinear relationship between urban agglomeration and income inequality in Sub-Saharan Africa. The findings showed that income inequality increases with urban agglomeration in the first stage and decreases in the later stages of urbanization. Based on the findings, our recommendations are to enhance governance capacity in providing urban infrastructural investment, improve industrialization capacity, and open up the peri-urban connecting rural regions through public–private development partnerships to shorten the urbanization-driven income inequality inverted U-shaped Kuznets’ turning point in Sub-Saharan Africa.

Keywords Urban agglomeration · Dynamic panel model · Gini index · Sub-Saharan Africa

1 Introduction

Urbanization is a valuable outcome of economic development. As nations develop, the urban share of the population increases due to population drift from underdeveloped rural to developed urban areas with economic opportunities such as employment and increases in income (Castells-Quintana, 2018; Kuznets, 1955; Moreno, 2017). Metropolitan regions are a critical potential factor reflecting human prosperity, development, and sustained economic growth because they contribute to consumption, innovation, and investment in developed and developing economies (Moreno, 2017). As such, the population in urban regions has been increasing because the migration to urban areas is increasing as individuals

✉ Isaiah Maket
maket.isaiah.juma@o365.u-szeged.hu

¹ Faculty of Economics and Business Administration, University of Szeged, Szeged, Hungary

endeavor to improve their income, employment, education, trade, and commerce opportunities and avail excellent communication and transportation services (Ahrend et al., 2017; Ikwuyatum, 2016).

A contributing factor to burgeoning urbanization and subsequent urban agglomeration (unregulated, continuous concentration of individuals in urban regions) worldwide is the uncontrollable widening variation in wealth and economic resources across rural and urban regions (Hardoon et al., 2016; Liddle, 2017; Tripathi & Kaur, 2017). Data from the World Bank has demonstrated that the population in Africa is remaining in urban regions, which supports the aforementioned finding. For instance, the World Bank reported that by 2030, over 50% of Africa's population will reside in metropolitan areas (World Bank, 2015a, 2015b; World Bank, 2019). Specifically, in Sub-Saharan Africa (SSA), the urban agglomeration has increased gradually, from 13.09 million people in 2000 to 14.35 million people in 2010 and to greater than 17.97 million people in 2020, with a positive propensity for further increases (United Nations Population Division, 2018; World Bank, 2022a, 2022b).

The sudden, unregulated upsurge in urban agglomeration presents conflicting outcomes in developed and developing economies. On the one hand, it leads to positive economic outcomes such as increased regional economic performance resulting from an increased labor supply pool, specialization, and proximity to urban industries (Ahrend et al., 2017; Maket, 2021). On the other hand, it leads to deleterious outcomes such as increases in income inequality, urban poverty, and the share of the urban population residing in slums and dilapidated urban settlements with inadequate access to public infrastructural services and employment opportunities, negatively affecting overall well-being and urban livability (Li, Chiu & Lin, 2019a, 2019b; Liddle & Messinis, 2015; Maket et al., 2022; Pereira, 2016; UN-Habitat, 2017).

SSA is among the world's leading regions experiencing rapid urbanization and the subsequent urban agglomeration coupled with insufficient government policy measures to provide safe public services such as water, electricity, and sanitation within the major metropolitan regions (Liddle, 2013; World Bank, 2020). This is attributed to the relative economic significance of this region's urban areas. Most of the nations in this region have relatively low incomes that vary based on the gross domestic product (GDP) produced (Liddle, 2013; United Nations, 2015). Therefore, no economy can realize sustainable economic growth without spontaneous urbanization and urban agglomeration, and this argument has been empirically tested. For instance, Castells-Quintana and Royuela (2015) observed that economic growth strongly correlates with urbanization and income inequality. Moreover, most of the developing countries in SSA are experiencing difficulty in overcoming the socioeconomic challenges due to urban agglomeration and continuous urbanization in terms of providing adequate housing, health care, water, energy, schooling, and employment (Manteaw, 2020; Tuholske et al., 2020; UN-DESA, 2018).

Parallel to SSA's rapid urbanization and urban agglomeration is the widening income inequality (Manteaw, 2020). Although fair post-reform economic development was supposed to mitigate—to some extent—the rural and urban poverty of millions of people, income inequality has widened due to the constant ineffective policy changes and ravaging disparity in the distribution systems and preferences (Bloch et al., 2015). For instance, income inequality measured in SSA by the Gini index averaged between 0.68 and 0.70 from 2000 to 2020, depicting widening income inequality (Standard World Income Inequality Database, 2022; World Bank, 2022a, 2022b).

Because of the co-occurrence of the variables, the research question posed in this study was whether increasing income inequality is nonlinearly related to urban agglomeration. Classical development theory, popularized by Kuznets (1955), considers urban

agglomeration crucial in rearranging the developing economies dichotomized by rural subsistence and an industrializing urban sector. The increasing rural–urban population drift is a significant dimension of economic structural processes (Kuznets, 1955). As an increasing number of individuals are migrating from the perceived lower-income rural agricultural areas to the perceived higher-income urban industrial sector, income inequality increases in the first urbanization stages and declines in later urbanization stages beyond the turning point (Kuznets, 1955). Most economies in SSA have not passed the turning point—income inequality is still increasing with urban agglomeration and is unlikely to begin to decline in the short term (Kanbur & Zhuang, 2013). This evaluation is vital for SSA because its income inequality is primarily attributed to widening rural–urban and urban–urban income gaps; hence, urbanization cannot reduce the impact of the disparity due to inadequate fiscal capitation and government technical ability during first stages, but the impact can be decreased after passing certain turning point (Bloch et al., 2015).

Against this backdrop, the innovation of this paper is its contribution to the literature of new knowledge regarding the Sub-Saharan African context. First, this study attempted to determine whether there is a nonlinear relationship between income inequality and urban agglomeration, following an inverted U-shaped Kuznets hypothesis. Second, this study used various estimation methods, such as pooled OLS, fixed effects (FE), random effects (RE), difference, and system GMM, in estimating nonlinear relationships by using a sample of the latest data, from 2000 to 2020, from 22 countries in SSA. The study is unique because of its primary focus: determining the validity of the inverted U-shaped Kuznets hypothesis in the case of SSA, which requires further research. Moreover, this study aimed to fill the gap in the literature due to the limitations of studies that have focused on SSA (Adams & Klobodu, 2019; Nkalu et al., 2020). Thus, rather than determining how urban agglomeration influences income inequality, we used current panel data to determine the turning point of urban agglomeration, from which the income inequality curve starts to decline.

The remainder of this paper is structured as follows: Sect. 2 presents the literature that departs from the theoretical evaluation of the inverted U-shaped Kuznets hypothesis and presents empirical evidence of either a linear or nonlinear relationship between urban agglomeration and income inequality; Sect. 3 presents the data, data sources, and estimation strategy; Sect. 4 presents the statistical results and discussions; and Sect. 5 summarizes the findings, draws deductive conclusions, and proposes policy recommendations.

2 Literature Review

Urban agglomeration refers to the concentration of individuals in urban regions orchestrated by rapid urbanization and rural–urban migration (Duranton, 2015). Thus, increased urban agglomeration (urban share of population) creates income inequality (economic differences) in terms of access to social amenities, employment opportunities, and the general well-being of the urban population (Brulhart & Sbergami, 2009; Harris & Todaro, 1970). Therefore, to theorize the relationship between urban agglomeration and income inequality, we first evaluated the inverted U-shaped Kuznets curve (Fig. 1), fronted by Kuznets (1955), in a specific region. Next, we explored income inequality in the regional context by considering urban agglomeration driven by the urbanization rate.

In Fig. 1, income inequality is plotted on the y-axis. The urban share of the population is displayed on the x-axis to show the nonlinear or quadratic relationship between income

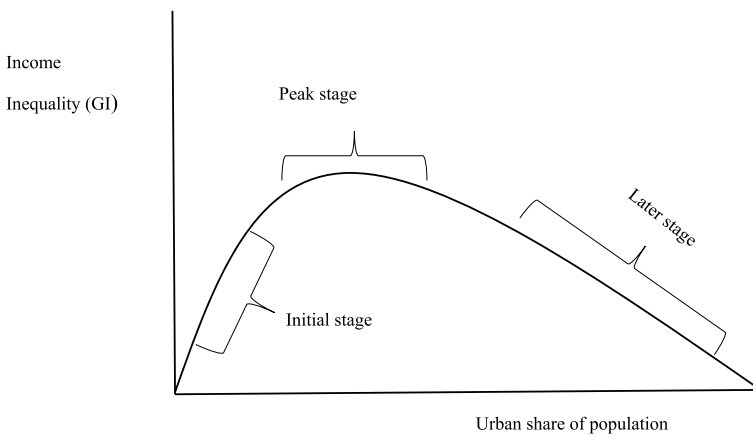


Fig. 1 Inverted U-shaped Kuznets Curve: Urban Agglomeration and Income Inequality. *Source* Anand and Kanbur (1993)

inequality and urban agglomeration (urban share of the population) (Anand & Kanbur, 1993). Illustratively, as urban agglomeration increases, income inequality increases in the initial stages, peaks at some level, and then deepens under the prevailing increasing urban share of the population throughout the urbanization and economic processes. Therefore, according to Kuznets (1955), urban agglomeration pursued by rural–urban migration processes fundamentally increases income inequality during the early stages of urban agglomeration. Additionally, as urban agglomeration increases, urban economic performance and industrialization widen the per capita income gap among the urban population (Anand & Kanbur, 1993).

Kuznets (1955) further outlines urbanization’s contribution to the nonlinear relationship between urban agglomeration and income inequality. This study considered urbanization a significant factor in increasing the speed of urban agglomeration and subsequent general income inequality in a region (Ha et al., 2019; Krugman, 1991). The rate of urbanization propels the urban agglomeration and economic performance of the urban areas, at least during the first stages of progress, signifying that a balance exists between urban agglomeration and income distribution beyond a specific turning point (Brulhart & Sbergami, 2009). Income inequality results from urban agglomeration and regional economic performance (Harris & Todaro, 1970; Lewis, 1954).

In summary, the Kuznets model explains the variation in income distribution in the first stages of urban agglomeration, and it decreases in later stages of urban development and population growth. With the illustration above, Kuznets depicts an inverse U-shaped (non-linear or quadratic) relationship between urban agglomeration and income inequality. Thus, the continuous migration of individuals from rural to urban regions is inherently linked with increased income inequality by increasing the share of the urban population in the early stages of urbanization (Ha et al., 2019).

In agreement with the theorized nonlinear relationship, Liddle (2017) empirically observed that increasing economic growth reduces poverty and narrows the rural–urban income gap. For instance, if a more significant share of the rural population migrates to urban regions with disproportionate urban economic opportunities, they become unemployed or engage in casual jobs which cannot fulfill their basic needs, worsening their

income gap (Arouri et al., 2017; Tuholske et al., 2020; UN-Habitat, 2017). Nonetheless, if the size of the rural–urban population matches the urban regions’ number of available economic opportunities, urbanization could be linked to reduced income inequality at later stages (Khan et al., 2016). Similarly, Wu and Rao (2017) found supportive evidence of a nonlinear or inverted U-shaped relationship between urban agglomeration and income inequality in provinces in China. In a subsequent study in Mozambique, Mahumane and Mulder (2022) observed an inverted U-shaped relationship between urban agglomeration driven by urbanization and income inequality associated with energy expenditure. Moreover, Christiansen and Weerdt (2017) used Tanzanian data between 1991 and 2010 and observed no significant relationship between urban agglomeration and income inequality. Focusing on SSA from a regional perspective, Adams and Klobodu (2019) and Sulemana et al. (2019) have observed an inverted U-shaped relationship between income inequality and the urban share of the population.

We divided the empirical literature into two broad categories. One category comprises studies on the relationship between urban agglomeration from an in-country analysis perspective (Castells-Quintana et al., 2015; Chen et al., 2017; Cottineau et al., 2019; Christiansen & Weerdt, 2017; Mahumane & Mulder, 2022; Martinez Posada & Garcia, 2017; Wu & Rao, 2017). The other category comprises studies that used cross-country or regional analysis, a growing literature strand that is shifting the research focus on world sample (Castells-Quintana et al., 2015; Li & Liu, 2018; Li et al., 2019a, 2019b; Naguib, 2017). Nevertheless, a growing number of studies is shifting the latest strands of the literature (cross-country or regional) from developing regions, such as the Sub-Saharan African perspective (Adams & Klobodu, 2019; Castells-Quintana, 2018; Sulemana et al., 2019).

The reviewed studies confirm the existence of a significant relationship between urban agglomeration and income inequality. However, further research is necessary to understand the exact relationship in the context of SSA and whether it follows the inverted U-shaped hypothesis. Due to the increasing policy and research focus on SSA, our paper contributes to the literature by presenting granular evidence on whether the relationship between urban agglomeration and income inequality is nonlinear and follows an inverted U-shaped hypothesis by using a sample of 22 countries in SSA; the current dataset from 2000 to 2020; and the panel dynamic data model estimated by pooled OLS, FE, RE, difference, and system GMM techniques.

3 Data and Methodology

We used a balanced panel dataset from 22 countries in SSA from 2000 to 2020. The data availability informed the inclusion of the countries and period of urban agglomeration data; thus, countries with data breaks were dropped from the initial list of all 48 countries in SSA, and only 22 fulfilled this threshold. Regarding variable inclusion, we included income inequality as the dependent variable, measured by the Gini index (Gini, 1909). Specifically, this study considered the Gini index computed from mean income differences in the country’s population while excluding location, age, and employment status (Solt, 2016). The panel data for income inequality was sourced from the Standardized World Income Inequality Database, a source derived from World Income Inequality Database, owing to its ability to include imputation or fill data gaps (Jenkins, 2015). Moreover, the Standardized World Income Inequality Database was used as the primary source of income inequality owing to its current data compilation, which was the main focus of this study

(Clark, 2013; Jenkins, 2015). Urban agglomeration measured by the urban share of the population was included as the independent variable. Urban agglomeration was measured using the urban share (%) of the population determined by dividing the total urban population by the country's population (Frick & Pose, 2018). Additionally, the urban share of the population from agglomerations above a 1 million thresholds was included as an additional measure of urban agglomeration for a robustness check (Asogwa et al., 2020; Frick & Pose, 2018). The summary of the study variables, measures, and data sources is shown in Appendix I.

Before proceeding, we must present some stylized facts about urban agglomeration and income inequality trends in SSA from 2000 to 2020. Income inequality growth averaged 62%, and the mean share of the urban population growth of the 22 selected countries in SSA was 37.4% over the period. Similarly, the average growth of the urban share of the population in agglomerations with more than 1 million people was 16.5% (see descriptive statistics in Appendix II). More notable than the aforementioned data is the increasing urban agglomeration, where the urban share of the population increased from 31% in 2000 to 36% in 2010 and 41% in 2020 (World Bank, 2023). In the same breath, income inequality in SSA averaged approximately 69% between 2000 to 2007 before slightly declining to an average of 68% between 2008 and 2016. Further, the recent income inequality growth averaged 67% from 2017 onward (Standardized World Income Inequality Database, 2023; World Bank, 2023; Solt, 2016).

Moreover, to establish the empirical link between urban agglomeration and income inequality, we included control variables, such as urbanization rate, regional economic performance, industrialization, education level, and governance policy preferences, identified in the literature as structural factors influencing both urban agglomeration and income inequality (Bloom et al., 2010). The datasets of these variables were obtained from the World Bank Development Indicators, World Governance Index, and World Penn World Tables (Heston et al., 2012). Regional economic performance due to high labor productivity and a large pool of skilled individuals migrating from rural to urban regions was included to capture the structural implication of the increasing urbanization rate, compounded by rural–urban migration (de Bruin & Liu, 2020; Lengyel & Szakálné, 2012, 2018). For instance, the wage income per capita decreases with income per capita at the early stages of development before declining at the later stages of the increasing urban share of the population (Behrens et al., 2014; Kuznets, 1955).

We included industrialization as a control variable measured by the number of employed individuals in urban industries (Altunbas & Thornton, 2019). It increases the rate of urbanization; urban agglomeration; and, subsequently, income inequality. Notably, the continuous migration of individuals from rural areas creates competition for the limited employment and opportunities in urban areas (Ali et al., 2021; Ike et al., 2020). Similarly, we incorporated the urbanization rate as part of the control variables measured by the ratio of the urban share of the population and the country's share of the rural population. We used this method because the literature has found that rural–urban migration is the key driver of the pace of urban agglomeration growth; moreover, income inequality in that the rural–urban income gap measured by the difference of agricultural rural sector employment and industrial urban sector employment income has a most considerable marginal impact on the income inequality (Khan et al., 2020). Ideally, as individuals migrate from rural to urban regions, the income gap between the rural and urban populations widens the gap in income inequality (Guo et al., 2019).

The study also included education level computed as the human capital index per person from schooling years and returns to education as part of the structural measures

preceding rapid urbanization and urban agglomeration, where individuals migrate to cities searching for education and career opportunities (United Nations, 2019; UN-Habitat, 2017). The continuous concentration of skilled youths from diverse backgrounds with their various innovative ideas creates a labor force pool around urban industries. Income increases at the first phase of education before decreasing later as returns to education decrease, as indicated by the high unemployment rate among youths in SSA (Castells-Quintana, 2017; Tripathi, 2021). Last, the study considered governments' policy preferences in containing the socioeconomic challenges posed by rapid urbanization in terms of their effectiveness in implementing and monitoring public provisioning service policies such as urban infrastructural development (Fossaceca, 2019; Satterthwaite et al., 2015; Thacker et al., 2019). Notably, SSA, similar to most of the urbanizing regions, is attempting to manage the increasing numbers of slums and unplanned settlements, in which over 50% of the population reside under the acute proliferation of inadequate access to sanitation, water, and energy services (Castells-Quintana, 2017; UNDP, 2016; Shi, 2019).

This study followed a dynamic panel data model of income inequality and urban agglomeration in line with pastulations of Elhorst (2014) in a quadratic form as follows:

Beginning with the panel dynamic data model in a reduced linear format:

$$G_{it} = \alpha_0 + \alpha_1 G_{it-1} + \beta_1 UA_{it} + \gamma UrbF_{it} + \eta_i + \mu_t + \epsilon_{it}; i = 1, \dots, N, t = 1, \dots, T \quad (1)$$

By stating Eq. (1) in a nonlinear dynamic panel model, we expressed income inequality as equivalent to urban agglomeration and its squared term together with the remaining control variables as follows:

$$G_{it} = \alpha_0 + \alpha_1 G_{it-1} + \beta_1 UA_{it} + \beta_2 UA_{it}^2 + \gamma UrbF_{it} + \eta_i + \mu_t + \epsilon_{it} \quad (2)$$

where G_{it} is income inequality (Gini index), G_{it-1} is lagged income inequality, UA and UA^2 are the matrix of the independent variable (urban agglomeration and its squared term) measured by the urban share of the population (ratio of total urban population to country's population) and the urban share of the population from agglomerations more than 1 million people, and $UrbF_{it}$ is the vector of the unobserved urbanization factor covariates of income inequality taken care of by the aforementioned control (urbanization, industrialization, regional economic performance, education level, and government policy preferences). The intercept is defined by α_0 , β_1 is the slope parameter for urban agglomeration, γ is the coefficient's vector for the control variables, η_i is the FE of the region i , μ_t is the RE at a particular time t , and the random error term is ϵ_{it} (Baltagi, 2008; Bond, 2002; Dang et al., 2015; Hsiao et al., 2002). The subscript indexes i refer to country 1, ..., 22, and t refers to time in years from 2000 to 2020.

Dynamic panel data models (1) and (2) were estimated using different estimation methods—pooled OLS, FE, RE, difference GMM, and systems GMM—because the use of pooled OLS, for instance, by default, assumes that at least a portion of the regression estimators is similar across the entire panel (pooling assumption) (Alvarez & Arellano, 2022). For this reason, Arellano and Bond (1991) and Arellano and Bover (1995) have proposed that GMM estimators are capable of producing unbiased estimates when the panel data has sufficiently large cross-sections (N) and short time series components (T), as is the case in this study where $N = 22$ and $T = 21$. Particularly, GMM techniques present the ability to use the difference operator Δ (difference GMM), a system of equations consisting of lag levels and lagged first differences as instrumental variables (system GMM), ideal for taking care of probable FE and Nickel bias (Nickell, 1981; Roodman, 2009). Expressly, the

system GMM technique assumes that additional first difference instruments are uncorrelated with the FE, dramatically enhancing estimation efficiency (Hansen, 1982; Roodman, 2009).

Summarily, all the variables whose measurements were in percentages—urbanization rate (ratio of urban population to rural population), the urban share of the population, the share of the urban population from agglomeration above 1 million (measures of urban agglomeration), the share of the population employed in urban industries (industrialization measure), and GDP per capita growth (a measure of regional economic performance)—were converted into an index or ratio form to maintain uniformity with the other variables measured in ratios (i.e., income inequality [Gini index], human capital index, and policy preference [governance effectiveness index]) and ensure a uniform distribution of approximately zero (skewness). First, in our estimation process, we assessed the order of integration of the study variables to determine whether they were stationary. We employed several panel unit root tests, such as those in Levin et al. (2002), known as LLC, and Im et al. (2003), known as the IPS test, to check for stationarity traits. LLC and IPS perform well in cases of panels with a small T. Our rationale for using panel unit root tests (i.e., IPS and LLC) instead of first-generation unit root tests (ADF and PP) was to increase the robustness of the test by using the available information provided by the cross-sections under consideration. Moreover, we employed the augmented cross-sectional IPS (CIPS) by Pesaran (2007) to account for the possibility of cross-sectional dependence in our panel data. Any dataset found not stationary was differenced to the first or second difference to make it stationary.

4 Empirical Results and Discussion

Figure 2 presents the relationship between urban agglomeration and income inequality following an inverted U-shaped Kuznets curve. In the first panel, the scatter plot shows an increasing relationship in the initial stages of urban agglomeration, peaking at some point and decreasing at later stages. The second panel shows the two-way plot with a turning point indicated by the vertical red line. Using Stata software, the turning point was 0.6206, calculated using regression coefficients (Fig. 2).

Note: The turning point shown by the vertical red line and the value of the urban share of the population at the turning point was generated using Stata software using command codes attached in Appendix III.

Turning point calculation follows a regression expressing average income inequality as a function of the average urban share of the population and its squared value for the selected countries over the period stated as follows:

$$\text{Income inequality} = \beta_0 + \beta_1 * \text{Urban share pop} + \beta_2 * \text{Urban share pop}^2 \quad (3)$$

By differentiating Eq. 3 with respect to the urban share of the population and equating the outcome to zero, we obtained

$$\beta_1 + \beta_2 * 2 * \text{Urban share pop} = 0 \quad (4)$$

Letting the urban share of the population be the subject of the formula in Eq. 4, we obtained the turning point value of the urban share of the population when the curve starts tilting:

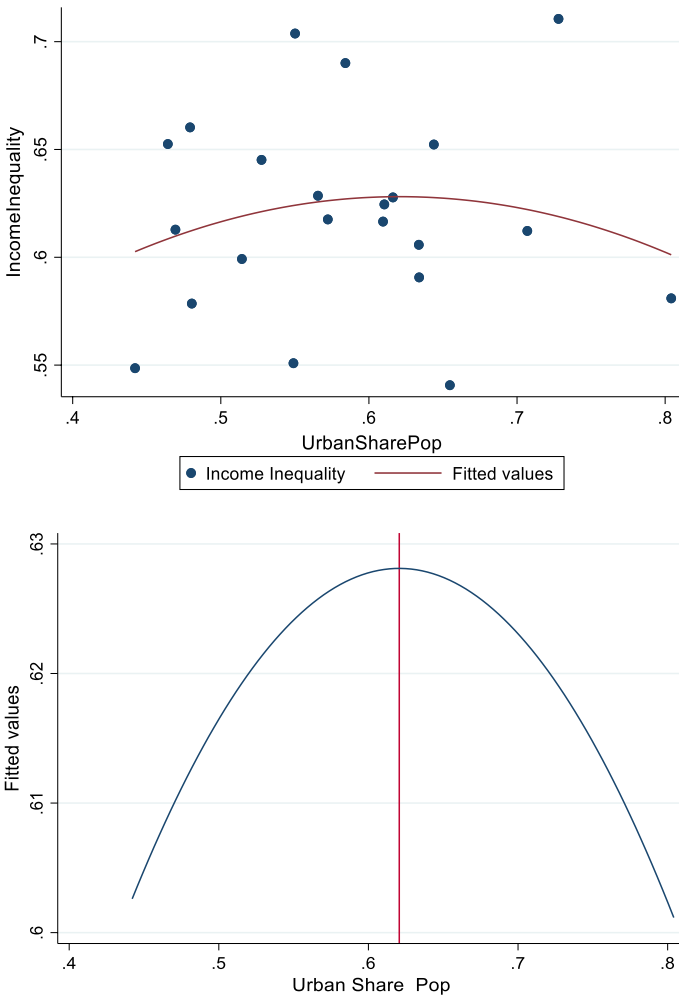


Fig. 2 Quadratic and Turning Point of Income Inequality and Urban Agglomeration. *Source* Author’s Computations from Stata software (2023).

$$\text{Urban share pop at turning point} = \frac{-\beta_1}{2\beta_2} = -0.5 \frac{\beta_1}{\beta_2} = -0.5 \left(\frac{0.99356}{-0.80046} \right) = 0.6206 \tag{5}$$

Thus, the turning point of the inverted U-shaped Kuznets curve occurred when the urban share of the population is at a 62.06% mark.

The cross-sectional dependence (CD) and panel unit root tests’ findings using the second generation unit root test are presented in Table 1 in level form, first difference, and second difference. Beginning with CD, the null hypothesis of no presence of cross-sectional dependence failed to be rejected for income inequality and government policy preference, implying the absence of cross-sectional dependence. However, the null hypothesis of no presence of cross-section dependence was rejected for the urban share

Table 1 Panel stationarity and cross-sectional dependence. *Source* Author's Computation from Eviews 12 (2023)

Variable	LLC	IPS	CIPS	CD	Status
Income inequality	-4.364**	-3.568**	-3.326**	-1.811	I (0)
Urban share of population	-28.114**	-14.235	-2.883**	-2.161**	I (2)
Urban share above 1M	-3.064**	-7.383**	-2.095**	31.067**	I (2)
Urbanization rate	-42.305**	-18.222**	-2.392**	55.452**	I (2)
Industrialization	-3.478**	-3.366**	-2.293**	6.362**	I (1)
GDP per capita growth	-8.314**	-12.751**	-3.043**	13.762**	I (1)
Human capital index	-4.273**	3.063	-2.635**	67.947**	I (1)
Policy preference	-2.487**	-2.158**	-2.815**	-1.438	I (0)

** indicates significant probabilities of the unit root and cross-sectional dependence (CD) tests at a 5% confidence interval. Urban Share Above 1M is the share of urban from agglomeration with more than 1 million people.

of the population, urban share above 1 million, urbanization rate, industrialization, GDP per capita growth, and human capital index, implying the presence of CD. Further, the null hypothesis of the presence of the unit root is rejected for income inequality and policy preference in level form when using LLC and IPS, implying integration to order 0 (I (0)). Additionally, LLC and IPS statistic values are significant at 5% for industrialization, GDP per capita growth, and human capital index at the first difference, implying the rejection of the null hypothesis of the presence of the unit root and that the variables are integrated to order 1 (I (1)).

Moreover, LLC and IPS statistical values are significant at 5%, signifying the rejection of the null hypothesis and the conclusion that the urban share of the population, urban share above 1 million, and urbanization rate are integrated to order 2 (I (2)). Last, we assessed whether the variables were stationary regardless of the presence of CD. CIPS confirmed significance for all variables, implying that all variables are stationary in CD. Therefore, because the variables are stationary in different forms, namely, a level form (I (0)), the first difference (I (1)), and the second difference (I (2)), considering testing for the relationship between variables using pooled OLS, FE, RE, difference and system GMM models is necessary.

After confirming the variables' panel unit root and CD, we empirically modeled the nonlinear relationship between urban agglomeration and income inequality. Pooled OLS, FE, and RE were used as benchmark techniques. The conclusion and observation are from the difference and system GMM regression results. Four principles informed this decision: (i) $N = 22$ is considerably greater than $T = 21$, although this T may produce unreliable estimates if the conclusion is based entirely on pooled OLS, FE, and RE (Maket et al., 2023; Dorn & Schinke, 2018); (ii) the income inequality measure (Gini index) is persistent over time and has a weak correlation with its first lag (-0.3023), lower than the threshold of establishing a good relationship between a variable and itself (Asongu & Aca-Anyi, 2019); (iii) GMM methods help preserve cross-economy variations, given the presence of CD and potential endogeneity among regressors in our panel data; and (iv) GMM methods mitigate all time-invariant and unobserved heterogeneity country-specific effects and account for simultaneity in the independent variable (Bond & Windmeier, 2002; Tchmyou et al., 2019).

Further, in choosing between the difference and system GMM, we used the estimated findings of pooled OLS, FE, RE, and difference GMM and compared the corresponding values of α_0 in the dynamic panel model in Eqs. 1 and 2 (Baum-Snow & Pavan, 2013). In this case, pooled OLS and FE were regarded, correspondingly, as upper-bound and lower-bound estimates of α_0 . Because a *priori* expectation is that α_0 is correlated positively with the error term (ϵ_{it}), the pooled OLS value will be biased upward, and the FE value will be biased downward; thus, the estimated GMM value for the valid parameter should be within this range (Bond, 2002; Roodman, 2009). Additionally, the Hausman test was conducted to determine whether relying on the FE or RE estimate of α_0 (Hausman, 1978). RE was relied upon because the null hypothesis that FE is more appropriate than RE was rejected (Tables 2 and 3).

Table 2: Pooled OLS, FE, RE, difference, and system GMM model linearity results. *Source* Estimations from Eviews 12 and Gretl (2023)

Variable	Pooled OLS (1)	FE (2)	RE (3)	GMM-Diff (4)	GMM-Sys (5)
Const.	-0.001 (0.007)	0.066*** (0.023)	0.009 (0.009)	0.0002 (0.001)	0.011 (0.008)
Income inequality (-1)	0.998*** (0.014)	0.903*** (0.035)	0.981*** (0.018)	0.792*** (0.019)	0.971*** (0.014)
DD_Urban share	-1.731 (1.695)	-2.077 (2.037)	-1.727 (1.624)	-2.917* (1.516)	-2.968*** (1.133)
DD_Urban share >1M	-0.003 (0.023)	-0.0005 (0.023)	1.131E-5 (0.023)	-0.031 (0.059)	-0.004 (0.051)
DD_Urbanization rate	0.771 (0.494)	0.595 (0.372)	0.680* (0.409)	0.614 (0.509)	0.834** (0.386)
D_GDP per capita growth	-0.015 (0.041)	-0.017 (0.036)	-0.021 (0.040)	-0.001 (0.006)	-0.004 (0.005)
D_Industrialization	0.235*** (0.076)	0.263*** (0.086)	0.252*** (0.0772)	0.303*** (0.0373)	0.272*** (0.043)
Governance policy preference	0.001 (0.001)	0.003 (0.004)	0.002 (0.001)	0.004** (0.002)	0.002** (0.001)
Human capital index	0.003*** (0.002)	-0.001 (0.008)	0.003 (0.002)	0.013*** (0.005)	0.004 (0.001)
Year FE	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES
F-statistic	19.689**	85.237**	2.765**		
Hausman test			32.32***		
Wald Chi ²				1980.0**	14370.3**
Hansen's J-stat				433.84	688.522
AR(1)				-2.306	-2.450
AR(2)				-2.692	-3.743
N	418	418	418	396	396

All columns included the lagged income inequality as part of the exogenous. Also, all estimations were done with the inclusion of asymptotic standard errors and time dummies. The standard errors are in parentheses. For Difference and System GMM estimations, model adjustment from the 1-st step to the 2-step was considered, out of which the model with reliable estimates was reported. D and DD denote the first and second differenced variables.

** $p < 0.05$, *** $p < 0.01$, * $p < 0.1$.

Table 3 Non-linear estimates of pooled OLS, FE, RE, difference and system-GMM. *Source* Estimations from Eviews 12 and Gretl (2022)

Variable	Pooled OLS (1)	FE (2)	RE (3)	GMM-Diff (4)	GMM-Sys (5)
Const.	-0.001 (0.008)	0.078*** (0.028)	0.006 (0.008)	0.001 (0.004)	0.002 (0.001)
Income inequality (-1)	0.998*** (0.015)	0.894*** (0.040)	0.985*** (0.018)	0.812*** (0.018)	0.795*** (0.060)
DD_Urban share	-1.771 (2.047)	-1.824*** (3.050)	-1.192 (2.075)	0.913*** (2.273)	0.860*** (1.352)
DD_Urban Share.Sq	-2.708 (2.704)	0.636*** (2.264)	-2.045 (2.315)	1.084*** (1.883)	0.705*** (2.065)
DD_Urban share >1M	-0.029 (0.063)	-0.094* (0.053)	-0.680 (0.060)	-0.154 (0.156)	0.006 (0.022)
DD_Urbanisatation rate	0.778 (0.926)	-0.532*** (1.727)	0.321 (0.944)	-0.80*** (1.603)	-0.569** (1.916)
DD_Urbanisatation rate.Sq	1.514 (2.223)	-0.726*** (0.766)	-0.863 (0.934)	-1.34*** (0.756)	-0.302*** (0.832)
D_GDP per capita Growth	-0.007 (0.007)	-0.006 (0.005)	-0.007 (0.006)	-0.003 (0.005)	-0.006 (0.005)
D_GDP per capita Growth.Sq	-0.010 (0.043)	-0.022 (0.037)	-0.016 (0.043)	0.024 (0.037)	-0.020 (0.036)
D_Industrialization	0.239*** (0.078)	0.270*** (0.083)	0.249*** (0.078)	0.260*** (0.055)	0.230** (0.083)
Governance policy preference	0.001 (0.001)	0.002 (0.004)	0.002 (0.001)	0.005** (0.002)	0.001 (0.005)
Human capital index	0.003 (0.002)	-0.003 (0.007)	0.003 (0.002)	0.009*** (0.002)	0.002 (0.012)
Year FE	YES	YES	YES	YES	YES
Time dummies	YES	YES	YES	YES	YES
F-statistic	17.154**	4.882**	2.765**		
Hausman test			1.987**		
Wald Chi ²				1980.0**	14370.3**
Hansen's J-stat				435.1	171.210
AR(1)				-2.759	-3.567
AR(2)				-2.436	-4.122
N	418	418	418	396	396

All variables are measured at the beginning of the period. The time spans from 2000 to 2020. All columns included the lagged income inequality as part of the exogenous. Also, all estimations were done with the inclusion of asymptotic standard errors and time dummies. The standard errors are in parentheses. D and DD denote the first and second differenced variables.

** $p < 0.05$, *** $p < 0.01$, * $p < 0.1$.

Table 2 presents the pooled OLS, FE, RE, difference, and system GMM regression results. The pooled OLS, FE, and RE estimates of income inequality as determined by urban agglomeration and selected control variables are shown in columns 1, 2, and 3. Columns 4 and 5 report the complete dynamic panel model estimates that used the difference and system GMM techniques, with all control variables included. However, as indicated, the estimate ($\alpha_0 = 0.011$) for system GMM was within the threshold required between the

pooled OLS estimate ($\alpha_0 = -0.001$) and RE estimate ($\alpha_0 = 0.009$). Therefore, we followed the system GMM findings, which showed a significant negative relationship between the urban share of the population (urban agglomeration measure) and income inequality at a 1% confidence interval (column 5). Further, the system GMM findings showed that control variables (urbanization rate, industrialization, and governance policy preference) were significant and positively related to income inequality at 1%, except for the human capital index and D_GDP per capita growth (column 5). The results also showed insignificant Hansen's J-statistics for system GMM, depicting no evidence of model misspecification. In addition, the AR (1) and AR (2) serial correlation statistic values were statistically insignificant, demonstrating the absence of serial correlation in the error terms (Arellano & Bond, 1991). The Wald and F-statistic values were significant at 5%, indicating the overall significance of the parameter estimates of the models under consideration.

We estimated the dynamic panel model in Eq. 2 to test for the nonlinearity (confirmation of inverted U-shaped Kuznets hypothesis) between urban agglomeration and income inequality.

$$G_{it} = \alpha_0 + \alpha_1 G_{it-1} + \beta_1 UA_{it} + \beta_2 UA_{it}^2 + \gamma UrbF_{it} + \eta_i + \mu_t + \epsilon_{it} \quad (2)$$

We effected this by including the urban share of the population and its squared term. Moreover, we included the urbanization rate and GDP per capita growth together with their squared terms and the remaining control variables in their unit form. Urbanization and GDP per capita growth were included in the nonlinearity analysis because, as aforementioned, income inequality tends to increase in early states of urbanization when economic performance is developing and decrease in later stages of urbanization when economic systems improve, and the majority of urban residents can access improved state provision and economic opportunities (Wu & Rao, 2016; Zhou & Qin, 2012). As shown in Table 3, including the quadratic term of the urban agglomeration, urbanization rate, and GDP per capita growth in all the models produced results different from those in Table 1. First, the results showed a significant positive relationship between the urban share of the population and income inequality at a 1% significance level (column 5). Second, imposing a quadratic term on urban agglomeration (urban share of population), the results revealed a significant positive relationship between the squared value of the urban share of the population and the income inequality at a 1% significance level, implying a nonlinear relationship between urban agglomeration and income inequality in SSA (column 5).

The results also showed a significant negative relationship between the urbanization rate and income inequality. Similarly, imposing a quadratic term on the urbanization rate measure, the relationship remained significantly negative at a 1% significant interval, implying that urbanization exhibits a nonlinear relationship with income inequality (column 5). However, GDP per capita growth and its square value did not significantly influence income inequality. The GMM-system results demonstrated a significant positive relationship between industrialization and income inequality. The Wald Test was significant in all the dynamic models, signifying the reliability of the findings in making deductions. Furthermore, Hansen's J-statistics were insignificant, depicting the correct specification of the difference and system GMM models. In addition, the AR (1) and AR (2) serial correlation statistic values were statistically insignificant, demonstrating the absence of serial correlation in the error terms (Arellano & Bond, 1991).

The reviewed literature provided significant evidence regarding the nonlinear relationship between urban agglomeration and income inequality. Although some studies have confirmed a linear relationship, and others have established a nonlinear relationship

between urban agglomeration and income inequality, our case focusing on SSA showed a significant nonlinear relationship. Mainly, the conflict in the literature can be attributed to the difference in a particular region's urban agglomeration level. In this study, we focused on determining whether there is a significant relationship between urban agglomeration and income inequality and whether the relationship is nonlinear in SSA. Our findings align with the theoretical fact in the inverted U-shaped Kuznets hypothesis: income inequality increases with urban agglomeration at the first stage, peaks in the middle, and decreases at later stages of urban agglomeration. The study modeled the nonlinear relationship between urban agglomeration and income inequality by using balanced panel data from 2000 to 2020 for 22 countries in SSA by using a dynamic panel data model estimated using multiple methods: pooled OLS, FE, RE, difference GMM, and system GMM estimation. However, all deductions were based on system GMM results.

This study used the urban share of the population and its quadratic term as measures of urban agglomeration in estimating a nonlinear relationship between urban agglomeration and income inequality. The results showed a significant positive relationship between urban agglomeration and income inequality. Further, imposing a quadratic term on the urban share of the population produced similar findings. Collaboratively, these findings align with the theoretical assertions of Kuznets (1955). Two arguments justify this observation. First, increasing agglomeration at the first stages occurs due to increased rural–urban migration, resulting in high-income differences as individuals take time to settle and find livelihoods owing to unmatching skills and education transitions in their early years in the urban regions (World Bank, 2019). Second, at the peak of urbanization, few individuals may be forced to migrate to greener cities, leaving a sizable number of economic opportunities to the remaining individuals, causing income inequality to decrease gradually as additional individuals access improved incomes, government services, and returns to education (Demont, 2013).

5 Conclusion and Policy Recommendations

We aimed to evaluate the nonlinear relationship between urban agglomeration and income inequality in SSA. We used a balanced panel dataset from 2000 to 2020 for 22 countries in SSA to empirically test the theoretical underpinnings of the inverted U-shaped Kuznets hypothesis (Kuznets, 1955). The system GMM results revealed a significant nonlinear relationship between urban agglomeration and income inequality in SSA, aligning with the inverted U-shaped Kuznets hypothesis and empirical studies in the literature, which confirmed a nonlinear relationship.

Based on our findings, our conclusion is that income inequality increases with urban agglomeration through increased rural–urban migration, which shifts the skilled labor factors of production to formal urban regions, leaving informal and rural sectors with limited economic productivity (Liddle, 2017). Moreover, the disproportionate productivity and the inappropriate state policy preferences that focus on engaging hopeful rural–urban migrants in economic production contribute to competition for limited resources in the urban region in their early years in cities (Liddle, 2013). Regarding the second part of the inverted U-shaped Kuznets curve, an argument could be that income inequality starts to decline when the majority of the urban population is accessing improved public provisions and returns to education, resulting from the distributive power of government through increased urban investment and development (Kanbur & Zhuang, 2013). Moreover, scaled

industrialization through public–private partnerships and increased government investment capacity regarding social amenities, such as water, sanitation, and energy access, results in a gradual decline in income inequality at later urbanization stages (Castells-Quintana, 2017). However, the critical policy puzzling question is what can be done to ensure that the inflection or turning point of the inverted U-shaped Kuznets curve occurs sooner than what has been observed.

In line with this deduction, we propose several policy recommendations for the rapidly urbanizing region of SSA. First, economies in SSA should tap into the increasing active population migrating to cities by implementing urban policies that favor employment creation, improved accessibility to social amenities supported by scientific and technological innovations, and a favorable business environment. These improvements can be achieved using collaborative inward-looking industrialization frameworks such as public–private partnerships. Moreover, business and scientific innovation opportunities can be increased, as Nkulu et al. (2020) suggested.

Additionally, the leadership and development stakeholders in SSA should review policies from fiscal allocation to social development, such as those involved in providing quality education and improving access to health care (Lee, 2005). These policy measures will ensure the distributional effect of the productive urban and rural populace. Last, government agencies charged with the responsibility of urban development should dissuade their focus on decentralizing services and development projects because this will be the means to opening up the peri-urban and connecting rural regions in SSA through industrialization and agro-processing (Sow, 2015; Sulemana et al., 2019). To achieve this objective, development agencies in the agricultural sector should enact policy incentives to make agriculture attractive to educated and uneducated individuals, reducing massively the rural–urban migration pushed by the perception that migrating to urban results in better jobs and economic prosperity than if they remained in their rural communities (Nkulu et al., 2020).

Overall, our findings reveal the need to chart a new development path that is pegged on enhancing a favorable governance environment in terms of quality government policy preferences and effectiveness in policy implementation. The net impact is that working institutions in SSA are pivotal in realizing rural-urban migration to ensure that the income inequality turning point comes early enough. As pointed out by Adams and Klobodu (2019) and Sy (2016), sound governance policy preferences and effectiveness, driven by accountability, and strict oversight, SSA can overcome the widening income inequality, retarding economic performance, and biting urban poverty.

Appendix I: Variable description, measurements, and data sources

Variable name	Measure	Data source	Source URL link
Income Inequality	Gini index	SWIID	https://fsolt.org/swiid/swiid_source/ .
Urban Agglomeration <i>Measure 1</i>	Urban share (%) of population determined by dividing total urban population by total country's population.	World Bank	https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=ZG .

Variable name	Measure	Data source	Source URL link
Urban Agglomeration <i>Measure 2</i>	Share (%) of urban population in agglomeration of more than 1 million people.	World Bank	https://data.worldbank.org/indicator/EN.URB.MCTY.TL.ZS?locations=ZG .
Urbanisation rate	The ratio of the urban population to the rural population	World Bank/UNDP	https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=ZG . and https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?locations=ZG .
Industrialization	Share (%) of the population employed in urban industries	World Bank	https://data.worldbank.org/indicator/SL.IND.EMPL.ZS?locations=ZG .
Economic Performance	GDP Per Capita Growth (annual % change)	World Bank	https://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG?locations=ZG .
Policy Preferences	Governance Effectiveness Index proxy measure	WGI/World Bank	https://info.worldbank.org/governance/wgi/ .
Education Level	Human Capital Index per person calculated based on years of schooling and returns to education	World Penn Tables	https://febpuwt.webhosting.rug.nl/Dmn/AggregateXs/PivotShow .

Source: Author's construction (2023).

Appendix II: Descriptive statistics

Variable name	N	Mean	Maximum	Minimum	Std. dev.
Income inequality	462	0.6205	0.7615	0.5231	0.0504
Urban share	462	0.3739	0.6783	0.1461	0.1474
Share above 1m	462	0.1647	0.6318	0.0335	0.1163
Urbanisation rate	462	0.7036	2.1084	0.1711	0.4699
Industrialization	462	0.1061	0.2879	0.0284	0.0523
GDP per capita growth	462	0.0155	0.1994	–	0.0437
Human capital index	462	1.7478	2.9133	1.0695	0.4242
Policy preferences	462	-0.8289	0.6459	– 1.8414	0.4397

Source Author's Computations from Eviews 12 (2023).

Appendix III: Turning point regressions and plots

```

tway qfit IncomeInequality UrbanSharePop
gen UrbanSharePopSQ = UrbanSharePop^2
reg IncomeInequality UrbanSharePop UrbanSharePopSQ
di -_b[UrbanSharePop]/(2 * _b[ UrbanSharePopSQ ])
local tp = -_b[UrbanSharePop]/(2 * _b[ UrbanSharePopSQ ])
tway qfit IncomeInequality UrbanSharePop , xli(`tp')

```

Variable	Coefficients
Constant (β_0)	3198 (0.3632)
Urban share of population (β_1)	0.99356** (1.2145)
Urban share of population SQ (β_2)	- 0.80046** (0.9997)

Source Author's Computations from Stata 17 (2023).

Acknowledgment We would like to express gratitude for the detailed reviews, suggestions, and modifications from two anonymous reviewers.

Funding Open access funding provided by University of Szeged. We acknowledge the funding from the African Economic Research Consortium (AERC), which facilitated the completion of this paper.

Declarations

Competing interests Authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adams, S., & Klobodu, E. K. M. (2019). Urbanisation, economic structure, political regime, and income inequality. *Social Indicators Research*, 142(3), 971–995.
- Ahrend, R., Lembecke, A. C., & Schumann, A. (2017). The role of urban agglomerations for economic and productivity growth. *International Productivity Monitor*, 32(3), 161–179.
- Ali, I. M. A., Attiaoui, I., Khalfaoui, R., & Tiwari, A. K. (2021). The effect of urbanization and industrialization on income inequality: An analysis based on the method of moments quantile regression. *Social Indicators Research*, 161(1), 29–50.
- Altunbaş, Y., & Thornton, J. (2019). The impact of financial development on income inequality: A quantile regression approach. *Economics Letters*, 175(3), 51–56.
- Alvarez, J., & Arellano, M. (2022). Robust likelihood estimation of dynamic panel data models. *Journal of Econometrics*, 226(1), 21–61.

- Anand, S., & Kanbur, S. R. (1993). The Kuznets process and the inequality—development relationship. *Journal of Development Economics*, 40(1), 25–52.
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297.
- Arouri, M., Youssef, A. B., & Nguyen, C. (2017). Does urbanisation reduce rural poverty? Evidence from Vietnam. *Economic Modelling*, 60(1), 253–270.
- Asogwa, F. O., Amuka, J. I., Igwe, A. A., & Nkalu, C. N. (2020). Dynamics of population, an urban agglomeration, and economic growths in Sub-Saharan Africa: Evidence from panel data. *Journal of Public Affairs*, 20(4), 178–190.
- Asongu, S. A., & Acha-Anyi, P. N. (2019). The murder epidemic: A global comparative study. *International Criminal Justice Review*, 29(2), 105–120.
- Baltagi, B. H. (2008). Forecasting with panel data. *Journal of Forecasting*, 27(2), 153–173.
- Baum-Snow, N., & Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics*, 95(5), 1535–1548.
- Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and aggregation. *Journal of Political Economy*, 122(3), 507–553.
- Bloch, R., Monroy, J., Fox, S., & Ojo, A. (2015). Urbanisation and urban expansion in Nigeria (URN) Research Report. London: ICF International.
- Bloom, D. E., Canning, D., & Fink, G. (2008). Urbanisation and the wealth of nations. *Science*, 319(5864), 772–775.
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115–143.
- Bond, S. R. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, 1(2), 141–162.
- Bond, S., & Windmeijer, F. (2002). Finite sample inference for GMM estimators in linear panel data models. *Cenmap Working Paper Series No. CWP04/02*. Institute of Fiscal Studies, London.
- Brühlhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics*, 65(1), 48–63.
- Castells-Quintana, D. (2017). Malthus living in a slum: Urban concentration, infrastructure, and economic growth. *Journal of Urban Economics*, 98, 158–173.
- Castells-Quintana, D. (2018). Beyond Kuznets: Inequality and the size and distribution of cities. *Journal of Regional Science*, 58(3), 564–580.
- Castells-Quintana, D., Ramos, R., & Royuela, V. (2015). Inequality in European regions: Recent trends and determinants. *Review of Regional Research*, 35(2), 123–146.
- Chen, B., Liu, D., & Lu, M. (2017). *City Size, Migration, and Urban Inequality in the People's Republic of China (No. 723)*. ADBI Working Paper Series.
- Christiaensen, L., & Weerdt, J. D. (2017). Urbanisation, growth, and poverty reduction: The role of secondary towns. Final Report. International Growth Center.
- Clark, R. (2013). Convergence in national income distributions. *Social Forces*, 92(2), 413–436.
- Cottineau, C., Finance, O., Hatna, E., Arcaute, E., & Batty, M. (2019). Defining urban clusters to detect agglomeration economies. *Environment and Planning B: Urban Analytics and City Science*, 46(9), 1611–1626.
- Dagum, C. (1997). Decomposition and interpretation of Gini and the generalised entropy inequality measures. *Statistica*, 57(3), 295–308.
- Dang, V. A., Kim, M., & Shin, Y. (2015). Searching robust methods for dynamic panel data models in empirical corporate finance. *Journal of Banking & Finance*, 53(1), 84–98.
- De Bruin, A., & Liu, N. (2020). The urbanization-household gender inequality nexus: Evidence from time allocation in China. *China Economic Review*, 60(3), 101–131.
- Demont, M. (2013). Reversing urban bias in African rice markets: A review of 19 National
- Dorn, F., & Schinke, C. (2018). Top income shares in OECD countries: The role of government ideology and globalisation. *The World Economy*, 41(9), 2491–2527.
- Duranton, G. (2015). Growing through cities in developing countries. *World Bank Research Observer*, 30(1), 39–73.
- Elhorst, J. P. (2014). Dynamic spatial panels: Models, methods, and inferences. In *Spatial Econometrics* (pp. 95–119). Springer, Berlin, Heidelberg.
- Fossaceca, A. (2019). Assessing the determinants of the human development index in oil-dependent nations. *Undergraduate Economic Review*, 16(1), 1–14.
- Gini, C. (1909). Concentration and dependency ratios (in Italian). *Rivista di Politica Economica*, 87(1), 769–789.

- Guo, C., Buchmann, C. M., & Schwarz, N. (2019). Linking urban sprawl and income segregation—Findings from a stylized agent-based model. *Environment and Planning B: Urban Analytics and City Science*, 46(3), 469–489.
- Ha, N. M., Le, N. D., & Trung-Kien, P. (2019). The impact of urbanisation on income inequality: A study in Vietnam. *Journal of Risk and Financial Management*, 12(3), 146–160.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 50(1), 1029–1054.
- Hardoon, D., Ayele, S., & Fuentes-Nieva, R. (2016). *An economy for the 1%*. Nairobi: Oxfam International.
- Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment, and development: A two-sector analysis. *The American Economic Review*, 60(1), 126–142.
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251–1271.
- Heston, A., Summers, R., & Aten, B. (2012). Center for international comparisons of production, income, and prices at the University of Pennsylvania. *Penn World Table Version*, 7, 1.
- Hsiao, C., Pesaran, M. H., & Tahmiscioglu, A. K. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short periods. *Journal of Econometrics*, 109(1), 107–150.
- Ike, G. N., Usman, O., & Sarkodie, S. A. (2020). Testing the role of oil production in the environmental Kuznets curve of oil-producing countries: New insights from the method of moments quantile regression. *Science of the Total Environment*, 711(2), 135–208.
- Ikwuyatun, G. (2016). Migration and urbanisation: Exploring the factors of the Nexus in Nigeria. *International Journal of Humanities and Social Science*, 6(8), 161–175.
- Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1), 53–74.
- Jenkins, S. P. (2015). World income inequality databases: An assessment of WIID and SWIID. *The Journal of Economic Inequality*, 13, 629–671.
- Kanbur, R., & Zhuang, J. (2013). Urbanisation and inequality in Asia. *Asian Development Review*, 30(1), 131–147.
- Khan, A. U., Saboor, A., Ali, I., Malik, W. S., & Mahmood, K. (2016). The urbanisation of multidimensional poverty: Empirical evidence from Pakistan. *Quality and Quantity*, 50(1), 439–469.
- Khan, H., Khan, I., & Binh, T. T. (2020). The heterogeneity of renewable energy consumption, carbon emission, and financial development globally: A panel quantile regression approach. *Energy Reports*, 6(2), 859–867.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3), 483–499.
- Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, 45(1), 1–28.
- Lee, C. (2005). Income inequality, democracy, and public sector size. *American Sociological Review*, 70(1), 158–180.
- Lengyel, I., & SzakálnéKanó, I. (2012). Competitiveness of Hungarian urban micro-regions: Localization agglomeration economies and regional competitiveness function. *Regional Statistics*, 2(1), 27–44.
- Levin, A., Lin, C. F., & Chu, C. S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1), 1–24.
- Lewis, W. A. (1954). Economic development with unlimited supplies of labor. *The Manchester School*, 22(5), 139–191.
- Li, J., Cheong, T. S., Shen, J., & Fu, D. (2019). Urbanisation and rural-urban consumption disparity: Evidence from China. *The Singapore Economic Review*, 64(4), 983–996.
- Li, Y., & Liu, X. (2018). How did urban polycentricity and dispersion affect economic productivity? A case study of 306 Chinese cities. *Landscape and Urban Planning*, 173(2), 51–59.
- Li, Y., Chiu, Y.-H., & Lin, T.-Y. (2019). Economic growth and air pollution impact public health in 31 Chinese cities. *International Journal of Environmental Research and Public Health*, 16(3), 1–26.
- Liddle, B. (2013). Urban density and climate change: A STIRPAT analysis using city-level data. *Journal of Transport Geography*, 28(3), 22–29.
- Liddle, B. (2017). Urbanisation and inequality/poverty. *Urban. Science*, 1(35), 1–7.
- Liddle, B., & Messinis, G. (2015). Which comes first—urbanisation or economic growth? Evidence from heterogeneous panel causality tests. *Applied Economics Letters*, 22(5), 349–355.
- Mahumane, G., & Mulder, P. (2022). Urbanization of energy poverty? The case of Mozambique. *Renewable and Sustainable Energy Reviews*, 159, 112089.
- Maket, I. J. (2021). Population dynamics and economic growth in Kenya. *Hungarian Statistical Review*, 4(2), 18–33.

- Maket, I. J., Kano, I. S., & Vas, Z. B. (2022). Reverse causal nexus between pro-poor policies and income inequality in Kenya. *Regional Science Policy & Practice*. <https://doi.org/10.1111/rsp3.12540>
- Maket, I., Kanó, I. S., & Vas, Z. (2023). Estimations of the pooled dynamic panel data model with time-space dependence of selected Sub-Saharan African urban agglomerations, 2000–2020. *Regional Statistics*, 13(4), 1–23.
- Manteaw, B. O. (2020). Sanitation dilemmas and Africa's urban futures: Foregrounding environmental public health in contemporary urban planning. *Academic Journal of Interdisciplinary Studies*, 9(5), 177–177.
- Martínez Posada, M. A., & Vélez García, H. M. (2017). *How economic agglomeration could impact the economic development at the region level and spatial analysis in Colombia* (Doctoral dissertation, Universidad EAFIT).
- Moreno, E. L. (2017). Concepts, definitions and data sources for urbanisation study: The 2030 Agenda for Sustainable Development. United Nations Expert Group Meeting On Sustainable Cities, Human Mobility and International Migration, UN/POP/EGM/2017/8.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 49(6), 1417–1426.
- Pereira, M. (2016). Urban population estimation based on residential building volume using IKONOS-2 images and LIDAR data. *International Journal of Remote Sensing*, 37(1), 1–28.
- Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2), 265–312.
- Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), 86–136.
- Satterthwaite, D., Mitlin, D., & Bartlett, S. (2015). Is it possible to reach low-income urban dwellers with good-quality sanitation? *Environment and Urbanization*, 27(1), 3–18.
- Shi, L. (2019). Promise and paradox of metropolitan regional climate adaptation. *Environmental Science & Policy*, 92, 262–274.
- Solt, F. (2016). The standardized world income inequality database. *Social Science Quarterly*, 97(5), 1267–1281.
- Sow, M. (2015). Foresight Africa 2016: Urbanization in the African context [Blog post]. Retrieved December 4, 2016.
- Stern, D. I., & Common, M. S. (2001). Is there an environmental Kuznets curve for Sulfur? *Journal of Environmental Economics and Management*, 41(2), 162–178.
- Sulemana, I., Nketiah-Amponsah, E., Codjoe, E. A., & Andoh, J. A. N. (2019). Urbanisation and income inequality in Sub-Saharan Africa. *Sustainable Cities and Society*, 48(2), 101–154.
- Sy, A. (2016). *Foresight Africa: Top priorities for the continent in 2016*. Brookings.
- Tchamyou, V. S. (2019). The role of information sharing in modulating the effect of financial access on inequality. *Journal of African Business*, 20(3), 317–338.
- Thacker, S., Adshear, D., Fay, M., Hallegatte, S., Harvey, M., Meller, H., & Hall, J. W. (2019). Infrastructure for sustainable development. *Nature Sustainability*, 2(4), 324–331.
- Tripathi, S., & Kaur, S. (2017). Do negative externalities have any impact on population agglomerations? Evidence from urban India. *Theoretical Empirical Research Urban Management*, 13 (3), 5–24.
- Tuholske, C., Andam, K., Blekking, J., Evans, T., & Caylor, K. (2020). Comparing measures of urban food security in Accra, Ghana. *Food Security*, 12(2), 417–431.
- UNDESA, (2018). Revision of world urbanization prospects. United Nations Department of Economic and Social Affairs. Retrieved October 9, 2022. <https://www.un.org/development/desa/publications/2018-revision-of-world-urbanization-prospects.html>
- UN-Habitat. (2017). World Cities Report 2016. Urbanisation and development emerging futures. United Nations-Habitat, Nairobi.
- United Nations. (2019). World Urbanization Prospects: The 2018 Revision. United Nations, United Nations Department of Economic and Social Affairs (UN-DESA). (2018). *World Urbanization Prospects: The 2018 Revision*. United Nations.
- United Nations Development Programme (2016). *Africa Inequality Study, UNDP Regional Bureau for Africa Working Paper Series*. Volume 1, Numbers 1-4.
- United Nations Human Settlements Programme (UN-Habitat, 2017). *Global Urban Indicators Database*. Nairobi.
- World Bank (2015). *Urbanisation in Africa: Trends, Promises, and Challenges*. Retrieved October 7, 2017, from <http://www.worldbank.org/en/events/2015/06/01/urbanization-in-africa-trends-promises-and-challenges>.

- World Bank. (2015). *Stocktaking of the Housing Sector in Sub-Saharan Africa: Challenges and Opportunities: Summary Report*. Washington, DC.
- World Bank (2019). *World Development Indicators*. Retrieved February 27, 2019, from <https://databank.worldbank.org/data/source/world-development-indicators#>.
- World Bank (2022). *World Development Indicators*. Retrieved April 6, 2022, from <https://data.worldbank.org/indicator/SP.URB.TOTL?locations=ZG>.
- World Bank. (2022). *Income Inequality in Sub-Saharan Africa*. Retrieved April 6, 2022, <https://data.worldbank.org/indicator/SI.POV.GINI?locations=ZG>.
- World Bank (2023). Urban Population Growth in Sub-Saharan Africa. Retrieved February 22, 2023, from <https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=ZG>.
- Wu, D., & Rao, P. (2017). Urbanisation and income inequality in China: An empirical investigation at the provincial level. *Social Indicators Research*, 131(1), 189–214.
- Zhou, Y., & Qin, Y. (2012). *Empirical analysis on income inequality of Chinese residents* (Vol. 1). Springer Science & Business Media.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.