Do rent-seeking and interregional transfers contribute to urban primacy in sub-Saharan Africa?

Kristian Behrens* Alain Pholo Bala†

August 12, 2011

— REVISED VERSION —

Abstract

We develop an economic geography model where mobile skilled workers choose to either work in a production sector or to become part of an unproductive elite. The elite sets income tax rates to maximize its own welfare by extracting rents, thereby influencing the spatial structure of the economy and changing the available range of consumption goods. We show that either unskilled labor mobility, or rent-seeking behavior, or both, are likely to favor the occurence of agglomeration and of urban primacy. In equilibrium, the elite may tax the unskilled workers but does not tax the skilled workers, and there are rural-urban transfers towards the agglomeration. The size of the elite and the magnitude of the tax burden that falls on the unskilled decrease with product differentiation and with the expenditure share for manufacturing goods. All these results are broadly in line with observed patterns of urban primacy and economic development in sub-Saharan African countries.

Keywords: economic geography; rent-seeking; interregional transfers; urban primacy; Sub-Saharan Africa.

JEL Classification: D72; F12; R12

*Canada Research Chair, Département des Sciences Économiques, Université du Québec à Montréal (UQAM), Canada; CIRPÉE; and CEPR. E-mail address: {behrens.kristian@uqam.ca, kristian.behrens@gmail.com}

†*Corresponding author:* Department of Economics and Econometrics, University of Johannesburg, APK Campus, D Ring 213, PO Box 524, Auckland Park 2006, Johannesburg, South Africa. E-mail address: {apholo@uj.ac.za, alainpholo@gmail.com}

"*It seems to be the thinking that Africans obtain an education in order to enter the government bureaucracy so as to be able to share the wealth of the nation rather than create wealth for the nation.*" (Bassey, 1999, p.106)

1 Introduction

Several explanations have been put forward to explain the urban growth in SSA. On the one hand, there are *push factors* linked to rural population growth, environmental degradation, climate change (Barrios *et al.*, 2006), and armed civil conflicts (Lozano-Gracia *et al.*, 2010). On the other hand, there are *pull factors* which are essentially linked to the attraction exerted by higher incomes and productivity, as well as a larger range of available goods and services, in the urban centers on rural migrants. Productivity improvements and higher incomes may stem from agglomeration economies arising from the diversity of intermediate goods, the matching process in the local labor market, from knowledge spillovers across firms, or any combination of these. There is an extensive literature on the mechanisms that generate aggregate increasing returns to scale at the urban level (see Duranton and Puga, 2004; Fujita and Thisse, 2002). While evidence for agglomeration economies in developed countries and in Asia is fairly well documented, this is not the case for SSA. Yet, agglomeration economies may play an important role in African urban development. According to McGranahan *et al.* (2009), careful analyses of the relationship between economic change and urbanization show that in Africa, as elsewhere, urbanization has been associated with economic growth. The potential role that African cities might play in the development of their continent is even clearly praised in the 2009 World Development report which claims that "urbanization, done right, can help development more in Africa than elsewhere" (World Bank, 2008, p.285) .

However, agglomeration economies in SSA seem less important than those in Asia and in OECD countries (Collier, 2006). Indeed, because countries in that region are too small and not integrated enough, many African cities tend to be too small compared to the optimum. As shown by Au and Henderson (2006) for the case of Asia, this may have serious impacts in terms of foregone growth. Research on agglomeration economies and international competitiveness further suggests that late-comers to industrialization, such as Africa's natural resource exporters, face a disadvantage linked to the spatial distribution of the global industry (Page, 2008). As the overall productivity of sub-Saharan Africa is fairly low compared to international standards, the economic drive does not seem that important in explaining city growth in that region. Indeed, the economic content of African cities "lacks the dynamism, specialisation, diversity and economies of scale normally associated with urban life" (Bryceson and Potts, 2006).

Hence, complementary explanations are required to understand the urban growth spurt occuring in that region. One cause for Africa's rapid urban growth may be the political role of its primate cities. The location of the central administrations in African primate cities has favored their development well beyond what is economically feasible and expected. The extent of the bureaucracy is often the most characteristic aspect of city development in SSA. Employment is predominantly generated in the government sector and parastatals (Bryceson and Potts, 2006). 21 out of the 27 African cities belonging to the 100 largest cities in the world that had the fastest population growth between 1950 and 2000 are capitals. They are the principal location of state administration and public entreprises. The preponderant share of capitals among large and fast growing cities is an African specificity. Indeed, this ratio is much lower in other regions of the world — 10 out of 34 in Asia, 4 out of 21 in Latin America & Carribean, and 0 out of 4 in North America (Satterthwaite, 2005).

The dominant role of capitals within the national urban hierarchy is further strengthened by the fact that a disproportionate part of the national budget is spent within them – which explains why they are usually the subject of massive in-migration by those seeking economic and political opportunities not available elsewhere. This is where African political economy kicks in: political leaders allocate a potentially economically indefensible share of resources to primate cities to satisfy the urban mob because they fear the pressure of city elites (Lipton, 1983; Bairoch, 1985). Spatial proximity is likely to increase political influence in non-democratic countries since the leadership is more sensitive to the claims of the urban elites than to those of people living in the country-side. Hence, as empirically shown by Ades and Glaeser (1995), politics directly affects urban concentration as rent-seeking agents have to be spatially close to the political power. The rents they reap, together with the government's net transfer of resources from the country-side to the cities to run the state, further raise city population by attracting economic activity to the main centers of purchasing power.[1] The net result is that still more population is attracted by these transfers, which draw ressources from the hinterland and make cities in the developing world absorb a disproportionate fraction of overall economic activity. Eventually, the process becomes self-reinforcing and leads to the formation of large urbanized core regions, as emphasized by the new economic geography (see Fujita *et al.*, 1999; Fujita and Thisse, 2002; Baldwin *et al.*, 2003).

The agglomeration process described in the foregoing highlights that economical, political, and spatial factors should be jointly taken into consideration when analyzing the formation of large urban agglomerations in the developing world.[2] Stated differently, *political rent-seeking, rural-urban*

[1]Similar mechanisms, leading to the formation of 'parasite cities' due to the transfers of resources, have already been vividly depicted by Cantillon (1730, ch. I.V, in paragraph I.V.2) and Bairoch (1985). As shown by Dascher (2000), income transfers which violate Ohlson's principle of 'fiscal equivalence' also seem to play a role in explaining the growth of regional capitals in Germany. The key idea is, as always, the breaking of the tax-benefits link.

[2]Kanbur and Venables (2003, p.473) clearly emphasize this aspect by asking the question: "To what extent are inequalities due to lumpiness in infrastructure investments, to industrial agglomeration forces, or to political economy factors that bias policy to selected regions?"

transfers, migration, increasing returns, and geography should be important ingredients of the whole story. The main objective of this paper is to present a new economic geography (henceforth, NEG) framework that combines these ingredients to shed some light on agglomeration, regional imbalances, and urban primacy in developing countries. Although the NEG literature has been rapidly growing these last years there are, to the best of our knowledge, only few contributions dealing with political factors. Robert-Nicoud and Sbergami (2004) present a model that analyzes the impacts of political factors, increasing returns, and economic integration on agglomeration. Their main objective is to explain how and why the regional integration process of the European Union leads to important transfers of resources from the 'core regions' to the 'peripheral regions'. Using a probabilistic voting model, the authors endogenize regional policy and transfers that are determined by the swing voters of the ideologically most homogeneous group. As the latter is predominantly located in the country-side, the peripheral regions can obtain transfers because of their relative political homogeneity. By contrast, the large region will keep the 'core' only if its relative economic size overcomes its political weakness due to ideological fragmentation. Robert-Nicoud and Sbergami's (2004) main result, namely that the political process leads to a more even distribution of economic activity than the market mechanism, is quite opposite to what we seem to observe in many developing countries, namely transfers of resources from the hinterland to the cities, which increases agglomeration. Furthermore, a democratic process, as embodied in the probabilistic voting model, does not adequately characterize the political environment of many developing countries, especially in regions like SSA. We thus propose an alternative model in which the 'political process' consists in agents deciding on whether or not to enter a political elite in order to extract rents to maximize their own welfare. In such a setting, rent-seeking leads to rural-urban transfers that exacerbate regional imbalances, thus showing that the nature of the political process matters for the direction of net transfers and the degree of agglomeration. This illustrates that "when farmers form a majority of the population, they tend to subsidize the urban minority. When farmers form a minority, the urban majority subsidizes them" (Friedman, as quoted by Mbeki, 2005, p.5).

The remainder of the paper is organized as follows. Section 2 presents some facts about urban primacy and its links with corruption in sub-Saharan Africa. In Section 3, we develop the model and discuss the market outcome. Section 4 then investigates the spatial equilibrium and shows that rural-urban transfers make the emergence of agglomeration more likely. Section 5 deals with the issues of elite formation and tax setting, and derives the equilibrium taxes and the equilibrium size of the elite. Section 6 concludes.

2 Some facts about urbanization in SSA

The urbanization process in SSA has led to a major population redistribution between rural and urban areas. Even among the urban centers, this redistribution has been very skewed towards the very large cities and metropolitan regions (Mabogunje, 1994). SSA countries have, as a result, become increasingly affected by regional imbalances and urban primacy.[3]

Insert Table 1 about here.

Table 1 illustrates the prevalence of urban primacy in SSA. As one can see, there are only few countries in that region with primacy levels below 15% – South Africa and Nigeria being among those few. While South Africa, with its comparatively high level of economic development, is characterized by a more balanced urban system, Nigeria departs from the typical regional pattern because its large population is hardly compatible with the existence of a single primate city. We can further see from Table 1 that while primacy is typical of small countries like Gambia, Burundi, Rwanda and Togo, where the capitals dominate the economic landscape, it is also characteristic of larger countries like Gabon, Angola, Congo, Ivory Coast, and Senegal.

Another noteworthy feature is that the biggest cities in Africa experienced the fastest growth rates. The share of SSA urban population living in cities of more than 500,000 inhabitants rose from 6% to 41% between 1960 and 1980, and the number of such centers increased from 3 to 28.[4] Moreover, the number of agglomerations with at least one million inhabitants jumped from 2 in 1950 to 21 in 1990 (Max Miller and Singh, 1994). Smaller cities grew at a much slower pace. According to the United Nations (2002), while the average population growth rate of cities with fewer than 500,000 inhabitants was of 3.8%, those of cities with between one and five million inhabitants and of cities with more than five million inhabitants were of 6.67% and 5.35%, respectively.

Figure 1 reveal that urban primacy is significantly correlated with central government expenditures and with corruption.[5] Put differently, corrupt countries with large central government expenditures on goods, services, and compensation of employees, generally have larger primate cities than the remaining countries. As can be further seen from Figure 1, SSA countries (depicted in red) are more strongly affected by these two factors. Hence, more corruption and larger central government expenditures are more strongly associated with urban primacy in SSA countries than in other parts

[3]Urban primacy can be measured in a variety of ways, e.g., as the proportion of urban population living in the largest city or as the ratio of the largest city population to the second-largest city population. The largest city is usually considered as being a 'primate city' if it has a population that exceeds twice the population of the second-largest city, i.e., if it deviates from the rank-size rule.

[4]World Bank (1989) *Sub-Saharan Africa: From Crisis to Sustainable Growth - A Long Term Perspective Study.* Washington D.C.: World Bank (as quoted by Mabogunje, 1994).

[5]The source and the description of the data used in figure 1 are given in Appendix B.

of the world.[6]

To further assess the link between urban primacy and corruption in a multivariate regression setting with more controls, we also make use of panel data. Our econometric model that follows is essentially based on Davis and Henderson (2003), who derived their specification from a theoretical framework modeling rural-urban allocation and primacy.

We estimate the following specification:

$$\begin{aligned} primacy_{it} &= \beta_0 \ \ln\left(urban\ population\right)_{it} + \beta_1 \left[\ln\left(urban\ population\right)_{it}\right]^2 \\ &\quad + \beta_3 \ \ln\left(\mathrm{GDP}pc\right)_{it} + \beta_4 \left[\ln\left(\mathrm{GDP}pc\right)_{it}\right]^2 + \beta_5 \ X_{it} + \beta_6 \ Z_i + \delta_t + \xi_{it} \end{aligned} \tag{1}$$

Observations are in five-year intervals from 1980 to 2005 for a panel of countries (see the Appendix for details). As explanatory variables, we firstly include national urban population and GDP per capita (adjusted for purchasing power parity, PPP). Those variables are inserted in quadratic form to account for non-linearities.[7] The X_{it} covariate is an institutional variable on political regimes. Indeed, some specific political institutions may mitigate overconcentration. For instance, democracy may allow the hinterland to have more influence in the political game (Davis and Henderson, 2003). Z_i denotes time invariant covariates.[8] In this category, we have data on transportation infrastructure: waterways per km^2, railways per km^2, roadways per km^2. We also include an index of quality of the road network: the share of kms of roadways that are paved. The vector of time invariant variables also includes data on the land area of each country, on religion (percentage of catholics, protestants and muslims), on ethnolinguistic fractionalization, and dummies indicating whether the primate city of a country is a capital, a port or whether a country is landlocked. Eventually, our most interesting time invariant variable is the non-corruption index. We add to this corruption measure an interaction term of the non-corruption index with a dummy for SSA countries. This interaction variable will allow us to assess whether corruption affects primacy differently in SSA countries than in other countries.[9] Finally, the δ_t control for time shocks/trends that are common across countries.

[6]The relationship between 'central governement expenditures on goods, services and compensation of employees' and urban primacy appears only slightly stronger for the SSA subsample. This is strongly driven by Uganda (point UGA in Figure 1), which is a leverage point of the regression. Without Uganda, the correlation becomes much stronger (from 0.43 to 0.68) and the slope larger (from 0.52 to 1.31) for the SSA subsample.

[7]For the specific case of GDP per capita, the principal economic rationale behind the non-linearity assumption is the so-called 'Williamson effect'. According to Williamson (1965), urban primacy first rises with economic growth and then declines. At the outset of economic progess, a high degree of urban concentration may be viewed as necessary since it enables the economy to make an efficient use of infrastructure and managerial resources. Nevertheless, as the economy keeps on growing, deconcentration finally occurs since the economy may now spread economic infrastructure and knowledge resources to the country-side.

[8]We also include in this category time varying covariates for which we do not have enough information over the considered time span. Such covariates are generally averaged over the period in which their data are available.

[9]Countries endowed with a substantial high-quality transportation infrastructure are less likely to concentrate their

Two comments are in order. First, Equation (1) does not explicitly include fixed effects because their presence would entail an identification issue for the time invariant variables. However, since country fixed effects allow to control for unobserved cross-country time invariant factors that may impact urban primacy, it is important to include them (Davis and Henderson, 2003). In Appendix C, we present a method that allows to estimate consistently time invariant variables in a panel setting with fixed effects. Second, specification (1) is prone to endogeneity issues as discussed at length in Davis and Henderson (2003). We also describe in Appendix C the two-step procedure that allows to obtain consistent 3SLS estimates of time varying variables.

Insert Table 2 about here.

Table 2 presents estimation results for equation (1). Column (1) gives pooled OLS results with time invariant covariates. The only time-varying covariate coefficient that is significant is the coefficient of the log of urban population. It is negative as expected: a greater urban scale entails that the economy can support more cities, reducing the population share of each of them. Column (1) yields other interesting results: primacy increases with the capital city dummy, which suggests that the extent of bureaucracy allows capital cities to achieve higher size than other primate cities. Urban primacy also decreases with the landlocked dummy. This supports the claim that long distances, deep divisions, and poor integration of transportation systems badly affect the size of African primate cities (Collier, 2006; World Bank, 2008). Results regarding the impact of transportation infrastructure are informative as well. While primacy is unaffected by the railways and waterways densities, it is positively affected by the density of roadways. However, it decreases with the share of paved roads. In words, a larger stock of roads raises urban primacy, but the higher the quality of the roadways network, the lower the increase in urban primacy.

One may also notice that, in the pooled OLS regression, the non-corruption index has a positive and significant impact on urban primacy. Its interaction term with the SSA dummy is negative, however, thus suggesting that there is a SSA-specific effect. Since the balance of the two terms is negative (-0.0016), more corrupt SSA countries have higher urban primacy.

Our pooled OLS estimates are prone to two kinds of problems: the first one is the neglect of unobserved heterogeneity, and the second is endogeneity. To mitigate the first issue, we present fixed effects estimates in column (2). Time invariant regressors, estimated in a second step, are displayed in blue. Concerning time-varying covariates, the estimated coefficients of urban population and its

ressources in just one city. Large countries are also more likely to spread their urban population across several cities. Moreover, we may conjecture that if the primate city in a country is a port, its size will be larger due to the transport hub effect (Fujita and Mori, 1996; Davis and Henderson, 2003). We previously mentioned that political economy in SSA favors the development of capital cities well beyond economic lines. This suggests that in case the biggest city of a country is a capital, its size will be larger as well. Finally, many authors attribute regional effects to cultural differences. We control these effects with measures of ethnic and cultural heterogeneity and religious affiliations.

square suggest that primacy varies non-monotonically with urban population. However, primacy declines with urban population throughout the sample range of the urban population variable.[10] Observe that most of the fixed effects results for time invariant covariates are in line with pooled OLS estimates: primacy increases with the capital dummy, and decreases with the share of paved roads. Moreover, urban primacy increases significantly with the non-corruption index but declines with the interaction term. As with pooled OLS, the balance of the two terms is negative (-0.0014). Conversely to pooled OLS results, the fixed effects estimate of the landlocked dummy is no longer significant. Moreover, urban primacy now declines significantly with the logarithm of land area.

Fixed effects estimation implies a substantial loss of degrees of freedom due to the inclusion of country dummies. Multivariate regression estimates from the first difference of (1) allows to mitigate that issue and to obtain more efficient estimates. Column (3) gives results for this regression procedure. It shows that primacy declines with urban population throughout the sample range of the urban population variable. It also indicates that primacy increases as income rises, peaks, and then declines. The peak is at an income of about 2120 US$, in line with the Williamson hypothesis. Concerning the road infrastructures covariates, we find the same kind of results in column (3) than in the pooled OLS specification (1): primacy decreases with the share of paved roads while it increases with roadways density. Findings concerning the capital city dummy are similar to previous ones, confirming that being the seat of central administration and of political institutions increases the size of primate cities. Results pertaining to the political institutions variables are also similar to those of columns (1) and (2): the non-corruption index and the interaction term have opposite signs, with the same negative balance as in column (2).

Our 3SLS estimates in colums (4) and (5) allow to handle endogeneity. They yield coefficients of the ln(urban population) and square of ln(urban population) that have higher absolute values than in previous findings. Concerning time invariant covariates, most of the results in column (4) are in line with previous results. Yet, small differences can be noticed: the interaction of primacy with the SSA dummy is still negative, as in previous results, but the coefficient of the non-corruption index is no longer significant. Hence, corruption favors urban primacy in SSA, whereas it has apparently no significant impact in other regions of the world. In column (5), we finally add the current value of the democracy index as a time-varying covariate. We obtain a negative coefficient that is significant at the 10% level. This suggests that the less democratic a country is, the higher its primacy. As indicated in Table 3, SSA countries are more autocratic than countries in other regions or groupings, except those of Middle East & North Africa. This further drives up average urban primacy in SSA.

To sum up, our panel data analysis suggest that corruption, autocracy, landlockedness and central

[10]In column (2), the minimum of the curve is attained at $ln\,(urban\ population) = 0.1731/\,(2 * 0.0040) = 21.64$. This implies a minimum value of urban population of more than two billions. Since such a value is clearly far beyond our sample range, primacy is monotonically decreasing with urban population.

administration in capital cities raise urban primacy in SSA countries. Therefore, urban primacy in SSA may stem from other factors than in the rest of the developing world: high corruption which leads to a redistribution of purchasing power from the country-side to the cities, where it is spent by an essentially unproductive elite. SSA may be the developing region where politics has the strongest impact on regional imbalances and primacy. In what follows, we propose a 'simple model of economic geography' that captures these features.

3 The model

We extend the analytically solvable NEG model of Forslid and Ottaviano (2003) to formalize the interactions between an elite ('corruption'), the redistribution of purchasing power across regions ('interregional income transfers'), and agglomeration ('urban primacy'). We consider a country with two regions, labeled 1 and 2. Variables associated with each region will be subscripted accordingly. The population consists of exogenously given masses $\overline{U}$ of unskilled and $\overline{S}$ of skilled workers. *All workers are geographically mobile.* The skilled may be either production workers or part of an *unproductive elite* — on top of being geographically mobile they are thus also socially mobile. The unskilled are socially immobile and are always production workers. Yet, contrary to the skilled, they can choose between working in the formal or in the informal sector of the economy (e.g., because they are harder to monitor and to tax than the skilled).[11] All agents spend their incomes locally and work in the region they live in. In what follows, we denote by S the mass of skilled in the production sector and by E the mass of unproductive skilled constituting the political elite. Both S and E are endogenously determined, with $S + E = \overline{S}$. We denote by $0 \leq \lambda^S \leq 1$ and $0 \leq \lambda^U \leq 1$ the share of productive skilled S and unskilled workers $\overline{U}$ living in region 1. We assume that the political elite is clustered into a historically determined center of power, which we henceforth refer to as the *capital* of the country (e.g., the historical capital). We do not attempt to endogenously determine where this center is located. Although this is an interesting question, it is secondary to the aspects we are interested in. While skilled workers are thus a priori mobile across regions, they become de facto immobile if they want to be part of the elite. In other words, social mobility of the skilled can only

[11] We assume that the unskilled cannot become skilled. Adding this possibility is formally equivalent to endogenizing the wage elasticity of labor supply from agriculture to manufacturing (Puga, 1998). This in turn is equivalent to allowing the geographically mobile population to increase, which would only reinforce our agglomeration results. We also assume, for simplicity, that the total population is fixed. Yet, one should keep in mind that strong population growth and the associated pressure on scarce resources in the countryside are important factors to explain why large streams of migrants are 'pushed' from rural areas to cities. Adding such considerations would, again, reinforce our results.

occur in the capital and once they are part of the elite they become regionally immobile.[12] Without loss of generality, we assume that region 1 is the capital. Our model may be viewed as a game with four stages: (i) the elite sets the tax rates t^U and t^S for the unskilled and the skilled; (ii) skilled workers decide whether or not to enter the elite; (iii) skilled production workers choose the region they live and work in; (iv) firms maximize profits, and production and consumption takes place. We solve that game by backward induction.[13]

3.1 Preferences

A representative consumer in region $i = 1, 2$ (whether he belongs to the skilled, to the unskilled, or to the elite) has Cobb-Douglas upper-tier preferences over agricultural and manufactured goods, with a CES sub-utility over a continuum of horizontally differentiated varieties. Formally, he solves the following consumption problem:[14]

$$\max_{A_i,\, q_{ji}(\omega)} \quad \frac{1}{\mu^\mu (1-\mu)^{1-\mu}} A_i^{1-\mu} \left[\int_{\Omega_i} q_{ii}(\omega)^{\frac{\sigma-1}{\sigma}} \mathrm{d}\omega + \int_{\Omega_j} q_{ji}(\omega)^{\frac{\sigma-1}{\sigma}} \mathrm{d}\omega \right]^{\frac{\mu\sigma}{\sigma-1}}$$

$$\text{s.t.} \quad p^A A_i + \int_{\Omega_i} p_{ii}(\omega) q_{ii}(\omega) \mathrm{d}\omega + \int_{\Omega_j} p_{ji}(\omega) q_{ji}(\omega) \mathrm{d}\omega = y_i$$

where A_i is the consumption of agricultural good; $q_{ji}(\omega)$ and $p_{ji}(\omega)$ stand for the consumption and the price of variety ω in region i when it is produced in region j; Ω_i stands for the set of varieties produced in region i, with measure n_i; y_i stands for the agent's income, which depends on the social group he belongs to (skilled, unskilled, or elite); and $0 < \mu < 1$, and $\sigma > 1$ are parameters. Since preferences are homothetic, we obtain the following aggregate demand for firm ω in region i when it is located in region j:

$$D_{ji}(\omega) = \frac{p_{ji}(\omega)^{-\sigma}}{\mathbb{P}_i^{1-\sigma}} \mu Y_i \tag{2}$$

where Y_i is the total income of agents in region $i = 1, 2$ and $\mathbb{P}_i$ is the CES price index. We assume that all varieties produced in the same region are symmetric, which allows us to alleviate notation

[12]In many countries, the political elite is strongly clustered and its location is very immobile spatially (China is an example). Observe, however, that this does not preclude that the people constituting the elite are initially spatially mobile and come from faraway place to the centers of power.

[13]As usual, the timing of the game is potentially important for the equilibrium. We could, e.g., consider the case in which tax setting and the formation of the elite take place simultaneously. Yet, we believe that our timing is natural since for agents to decide on tax rates they must already have chosen to be part of the elite. Furthermore, since our game is one-shot, we disregard the fact that the elite wants to remain the elite in the long run.

[14]As correctly pointed out by a referee, the assumption that the skilled, the unskilled and the elite have identical preferences is a strong one. We could relax it by assuming that the expenditure share on differentiated goods is higher for the skilled and the elite. Doing so would reinforce the agglomeration tendencies in our model, however at the expense of much heavier notation and algebra.

by dropping the variety index ω. The price index $\mathbb{P}_i$ then reduces to

$$\mathbb{P}_i^{1-\sigma} = \left(n_i p_{ii}^{1-\sigma} + n_j p_{ji}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} . \tag{3}$$

3.2 Technology, taxes, and transportation

There are two production factors, skilled and unskilled labor, and two sectors, manufacturing and agriculture. The agricultural sector produces a homogeneous good using unskilled labor only. We assume that this good is costlessly tradable across regions. Without loss of generality, we normalize the unit input coefficient in that sector to one. Perfect competition and costless trade then imply that the unskilled wages w^U are equalized across regions: $w_1^U = w_2^U = p^U = 1$, where the last equality comes from our choice of numéraire.[15] All unskilled workers are a priori free to work in agricultural or in manufacturing. Hence, the wages for unskilled in the two sectors will be equalized.

Skilled workers can choose to remain in the production sector to earn a wage w_i in region i, or to get involved in politics and become part of the elite in the capital. The benefit of belonging to the elite is to participate in running the country and to extract rents for personal consumption. The elite determines the tax rates for the different groups of agents. To keep things simple, we suppose that the elite levies *proportional income tax rates* t^S and t^U on the incomes of skilled and unskilled workers in the manufacturing sector.[16] The agricultural sector is assumed to be informal and untaxed because it is hard to monitor. The unskilled hence face the choice of working either in the formal manufacturing sector and to pay taxes, or to work in the agricultural shadow economy, hence evading taxation. Put differently, the unskilled tax base is not perfectly inelastic but generally shrinks with the level t^U of taxation. This observation fits well with the empirical fact that all African countries have large shadow economies (fiscal evasion seems to be fairly easy due to the lack of enforcement). Recent estimates using a sample of 37 African countries indeed reveal that the average share of the shadow economy is about 43.2% of GDP (Schneider, 2004, Figure 3.1.1). The skilled do not work in the shadow economy and can thus evade taxation only by becoming part of the elite.

Manufacturing firms produce varieties of a horizontally differentiated consumption good using both skilled and unskilled labor. More precisely, each firm requires F units of skilled labor as a fixed input requirement and m units of unskilled labor per unit of output as a variable input requirement. The unskilled are willing to work in the formal sector if and only if their net income is equal to the one they can secure in the informal sector, which implies that firms have to pay $1/(1-t^U)$ per unit

[15] Strictly speaking, such factor price equalization only holds when the mass of unskilled workers is large enough for some agricultural production to take place in both regions at equilibrium. In what follows, we assume that this condition holds (formally, μ must be small compared to σ; see Forslid and Ottaviano, 2003). More precisely, we consider that $\mu < \sigma - 1$.

[16] Since tax revenue is spent entirely unproductively by the elite, we may view $1 - t^S$ and $1 - t^U$ as proxies for the 'degree of property rights enforcement' for skilled and unskilled workers, respectively (Ades, 1995).

of unskilled labor. Hence taxing the unskilled raises firms' production costs because workers have the outside option of the shadow economy. Total production costs for producing a quantity X_i in region $i = 1, 2$ are then given by

$$\mathrm{TC}_i(X_i) = \frac{m}{1 - t^U} X_i + F w_i,$$

where w_i stands for the skilled wage in region i and where we have used $w^U \equiv 1$. Given the fixed labor requirement, skilled labor market clearing then requires that the masses of firms in the two regions are as follows:

$$n_1 = \frac{\lambda^S S}{F} \qquad \text{and} \qquad n_2 = \frac{(1 - \lambda^S) S}{F}. \tag{4}$$

To ship one unit of any variety between the two regions entails an iceberg trade cost of $\tau > 1$. Taking into account this resource waste, the profit of a representative firm in region i is given by

$$\Pi_i = \left(p_{ii} - \frac{m}{1 - t^U} \right) D_{ii} + \left(p_{ij} - \frac{m\tau}{1 - t^U} \right) D_{ij} - F w_i,$$

where the demands are evaluated at (2). Since those demands are isoelastic, the profit-maximizing prices display a constant mark-up over marginal cost:

$$p^*_{ii} = \frac{\sigma m}{(\sigma - 1)(1 - t^U)} \qquad \text{and} \qquad p^*_{ij} = \frac{\sigma m \tau}{(\sigma - 1)(1 - t^U)}.$$

Substituting these prices into (3), letting $\phi \equiv \tau^{1-\sigma}$ stand for the freeness of trade between the two regions, and using the skilled labor market clearing conditions (4), the price indices can be expressed as follows:

$$\begin{aligned} \mathbb{P}_1 &= \frac{\sigma m}{(\sigma - 1)(1 - t^U)} \left(\frac{S}{F} \right)^{\frac{1}{1-\sigma}} \left[\lambda^S + (1 - \lambda^S)\phi \right]^{\frac{1}{1-\sigma}} \\ \mathbb{P}_2 &= \frac{\sigma m}{(\sigma - 1)(1 - t^U)} \left(\frac{S}{F} \right)^{\frac{1}{1-\sigma}} \left[\lambda^S \phi + (1 - \lambda^S) \right]^{\frac{1}{1-\sigma}} \end{aligned} \tag{5}$$

Note that, contrary to other NEG models (Krugman, 1991; Ottaviano *et al.*, 2002) where the total mass of varieties is proportional to the exogenously fixed population and the price index only changes with the spatial distribution of that population, there are two 'price index effects' in our model:

- for any given value of S, the price index in a region decreases with the share of firms located in that region ('regional market crowding effect'); and
- for any given distribution of firms, *the price indices in both regions decrease with the mass S of productive skilled workers* ('global market crowding effect').

This second effect, which implies that a smaller mass of productive skilled workers decreases welfare by reducing product diversity and by increasing consumer prices, will be important in the subsequent analysis of the elite's behaviour.

Product market clearing for each variety requires that a firm located in region i produces the total quantity

$$X_i = D_{ii} + \tau D_{ij} = \frac{\mu(\sigma-1)(1-t^U)}{m\sigma}\left(\frac{Y_i}{n_i+\phi n_j} + \frac{\phi Y_j}{\phi n_i + n_j}\right). \tag{6}$$

Since firms price above marginal cost, there exist pure operating profits which are competed away by firms' bidding for skilled labor. Therefore, in equilibrium the skilled wages absorb all operating profits:

$$w_i = \frac{mX_i}{F(\sigma-1)(1-t^U)} = \frac{\mu}{\sigma F}\left(\frac{Y_i}{n_i+\phi n_j} + \frac{\phi Y_j}{\phi n_i + n_j}\right), \tag{7}$$

which using (4), can finally be expressed as follows:

$$w_1 = \frac{\mu}{\sigma S}\left[\frac{Y_1}{\lambda^S + \phi(1-\lambda^S)} + \frac{\phi Y_2}{\phi\lambda^S + (1-\lambda^S)}\right] \tag{8}$$

$$w_2 = \frac{\mu}{\sigma S}\left[\frac{\phi Y_1}{\lambda^S + \phi(1-\lambda^S)} + \frac{Y_2}{\phi\lambda^S + (1-\lambda^S)}\right]. \tag{9}$$

3.3 Rent-seeking and elite formation

As stated before, skilled workers can only evade taxation by becoming part of the elite. When they do so, they spend their time in an 'unproductive' way which de facto reduces the amount of skilled labor available for the production of differentiated varieties. Although this is a strong assumption it partly captures the fact that, in many African countries, political participation essentially stems from a rent-seeking motive, which is time intensive (e.g., because of lobbying) and therefore reduces the productive labor supply of agents involved in this kind of activity. A skilled worker will become a member of the elite if and only if the rent he can secure from doing so exceeds the skilled wage he can secure in the production sector.[17] Formally, a skilled worker will enter the elite if his rent r exceeds $(1-t^S)w_i$, where the rent satisfies

$$\begin{aligned} rE &\equiv m[n_1X_1 + n_2X_2]\frac{t^U}{1-t^U} + S\left[\lambda^S w_1 + (1-\lambda^S)w_2\right]t^S \\ &= S\left[(\sigma-1)t^U + t^S\right][\lambda^S w_1 + (1-\lambda^S)w_2], \end{aligned} \tag{10}$$

the right-hand side being the elite's total tax revenue. Note that we assume, for simplicity, that the elite *does spend all the tax revenue unproductively.* This captures the idea that "in those African countries in which corruption has become quite pervasive, the cost of public goods and services is highly inflated, usually to provide additional income for the individuals whose job it is to serve the public" (Mbaku, 2003, p.317). Adding some partial tax spending on the provision of public goods and services is not likely to weaken our results if this spending is largely allocated to cities and financed from general taxes instead of specific urban ones.

[17] There is a long-standing tradition which assumes that there are some fixed entry costs for becoming a member of the elite (see, e.g., Ades, 1995, for further references). For simplicity, we disregard these costs in this paper. Adding them would reinforce our results but make the algebra more involved.

3.4 Market outcome

We first analyze the market outcome for any *given* allocation of the skilled between the production sector and the elite, and for any given spatial distribution (λ^S, λ^U) of skilled and unskilled workers across regions. Furthermore, t^U and t^S are considered fixed at this stage.

Since the elite is, by assumption, immobile and tied to region 1, the net aggregate incomes accruing to the agents in both regions are given by

$$\begin{aligned} Y_1 &= \lambda^U \overline{U} + \lambda^S S w_1 (1 - t^S) + Er \\ &= \lambda^U \overline{U} + \lambda^S S w_1 (1 - t^S) + S\left[(\sigma - 1)t^U + t^S\right]\left[\lambda^S w_1 + (1 - \lambda^S) w_2\right] \quad (11) \\ Y_2 &= (1 - \lambda^U)\overline{U} + (1 - \lambda^S) S w_2 (1 - t^S). \quad (12) \end{aligned}$$

The market outcome is a solution to the four equations (8), (9), (11) and (12) in the four unknowns w_1, w_2, Y_1 and Y_2. The unique solution in wages w_1^* and w_2^* to this linear system is relegated to Appendix A.1.[18] It can readily be verified that w_1^* and w_2^* reduce to the corresponding expressions in Forslid and Ottaviano (2003) in the no-tax case with a symmetric distribution of the unskilled ($t^U = t^S = 0$ and $\lambda^U = 1/2$).

Some comments are in order. First, both w_1^* and w_2^* are *increasing* in the unskilled-to-skilled ratio $\overline{U}/S$, which will itself be endogenously determined later in our analysis. The reason for this is that, on top of standard endowment effects, a larger supply of skilled production workers increases the mass of competing firms, which leads to global product market crowding and, therefore, lower equilibrium wages. Such an effect does not arise in standard NEG models where the mass of firms is usually proportional to the exogenously fixed mass of skilled workers (Krugman, 1991; Ottaviano *et al.*, 2002). Second, since the denominator of w_i^* does not depend on λ^U, we clearly have $\partial w_1^*/\partial \lambda^U > 0$ and $\partial w_2^*/\partial \lambda^U < 0$. In words, there is 'complementarity in agglomeration' as the clustering of the unskilled in one region raises skilled workers' wages there and reduces skilled workers' wages in the other region. Hence, agglomeration forces in our setup are stronger than in traditional models where the unskilled are immobile and assumed to be evenly spread across regions. Finally, standard but longer calculations show that w_2^* is decreasing in t^S, whereas w_1^* is increasing in t^S. Stated differently, increasing taxation of the skilled shifts nominal wages in favor of the capital region and away from the periphery. The reason is that as t^S increases, the elite spends proportionally more tax revenues on varieties produced in the capital region, thereby raising demand and wages there. As we show later, the widening interregional wage gap induced by taxation increases the tendency for agglomeration of the mobile sector.

[18]Because Y_1^* and Y_2^* are complicated expressions that are not required for the subsequent analysis, we do not provide their analytical expressions.

4 Spatial equilibrium

Let ΔU^{S*} stand for the indirect utility differential between regions 1 and 2 of a skilled worker, which is a function of tax rates and given as follows:

$$\Delta U^{S*}(t^S, t^U) \equiv U_1^{S*} - U_2^{S*} = (1 - t^S)\left(\frac{w_1^*}{\mathbb{P}_1^{\mu}} - \frac{w_2^*}{\mathbb{P}_2^{\mu}}\right).$$

Recalling that unskilled wages are normalized to unity and equal across locations, the indirect utility differential ΔU^{U*} between regions 1 and 2 of an unskilled worker is given by:

$$\Delta U^{U*}(t^S, t^U) \equiv U_1^{U*} - U_2^{U*} = \frac{1}{\mathbb{P}_1^{\mu}} - \frac{1}{\mathbb{P}_2^{\mu}}.$$

Both skilled and unskilled workers are spatially mobile. We assume that the unskilled have heterogeneous regional tastes, whereas the skilled do not display any particular regional attachment.[19] Following Tabuchi and Thisse (2002) and Murata (2003), we assume that the unskilled have idiosyncratic preferences for unobserved regional characteristics. More precisely, an unskilled worker chooses a location to maximize

$$V_r^{U*} \equiv U_r^{U*} + \xi_r^U$$

where ξ_r^U is the realization of a random variable that is double exponentially distributed with zero mean and variance equal to $\pi^2\beta^2/6$. The parameter β is referred to as the unskilled workers' degree of regional taste heterogeneity. When β is large, regional tastes are very heterogeneous and the unskilled distribute themselves more equally across regions. In the limit where $\beta \to \infty$, the unskilled have extremely heterogeneous tastes and a so strong regional attachment that they become immobile and split themselves symmetrically across regions. On the other hand, the smaller is β the more the unskilled choose regions as a function of economic opportunities only. In the limit where $\beta \to 0$, the unskilled have the same tastes and all react in the same way to differences in regional economic incentives.

As is well known, the probability of an unskilled to choose to live in region $r = 1, 2$ is given by

[19] A few comments are in order. First, what matters for our analysis here is that the unskilled are not perfectly mobile and thus generate a dispersion force in the model. Yet, we want to make them somewhat mobile to account for the observed unskilled migration to cities in many developing countries. Second, we could readily introduce another parameter that governs the imperfect interregional mobility of the skilled, but doing so would make the model very heavy. It is well known that the skilled are more mobile than the unskilled, and we capture this differential mobility in a parsimonious way. Furthermore, if we think of the unskilled as being primarily employed in agriculture, land ownership puts additional constraints on their possible migration behavior. Last, as pointed out by a referee, the elite is usually immobile in many countries and this seems to conflict with our assumed migration behavior of the skilled. Yet, recall that for skilled to be in the elite, they have to be de facto immobile and located in the capital. Ex ante, the skilled who want to be part of the elite are mobile, yet ex post they are not.

$\mathcal{P}_r = \Pr\left(U_r^{U*} \geq U_s^{U*},\ s \neq r\right)$ which can be expressed as

$$\mathcal{P}_r = \frac{\mathrm{e}^{U_r^{U*}/\beta}}{\mathrm{e}^{U_r^{U*}/\beta} + \mathrm{e}^{U_s^{U*}/\beta}}, \quad r \neq s.$$

The law of motion of the unskilled is then given by $\mathrm{d}\lambda^U/\mathrm{d}t = (1-\lambda^U)\mathcal{P}_r - \lambda^U\mathcal{P}_s$, which is the inflow minus the outflow of agents given the intial distribution and their choice probabilities. A spatial equilibrium is such that $\mathrm{d}\lambda^U/\mathrm{d}t = 0$, i.e., $\mathcal{P}_r/\mathcal{P}_s = \lambda^U/(1-\lambda^U)$. It then follows that

$$\frac{\lambda^U}{1-\lambda^U} = \frac{\mathcal{P}_r}{\mathcal{P}_s} = \mathrm{e}^{(U_r^{U*} - U_s^{U*})/\beta}.$$

Taking logarithms of the foregoing expression, a spatial equilibrium for the unskilled is such that

$$\Delta V^{U*} \equiv \frac{1}{\mathbb{P}_1^{\mu}} - \frac{1}{\mathbb{P}_2^{\mu}} - \beta \ln\left(\frac{\lambda^U}{1-\lambda^U}\right) = 0, \tag{13}$$

where

$$\frac{1}{\mathbb{P}_1^{\mu}} - \frac{1}{\mathbb{P}_2^{\mu}} = \left(\frac{S}{F}\right)^{\frac{\mu}{\sigma-1}} \left[\frac{(1-t^U)(\sigma-1)}{m\sigma}\right]^{\mu} \left[(\lambda^S + \phi(1-\lambda^S))^{\frac{\mu}{\sigma-1}} - ((1-\lambda^S) + \phi\lambda^S)^{\frac{\mu}{\sigma-1}}\right] \tag{14}$$

is the regional cost-of-living difference. Since the regional price indices $\mathbb{P}_1$ and $\mathbb{P}_2$ do not directly depend on λ^U, there exists a unique solution $\lambda^U(\lambda^S)$ to equation (13). The expression being cumbersome, we relegate it to Appendix A.2. For given values of S and t^U,

$$\partial\left(\frac{1}{\mathbb{P}_1^{\mu}} - \frac{1}{\mathbb{P}_2^{\mu}}\right)/\partial\lambda^S > 0 \quad \text{so that} \quad \partial\lambda^U(\lambda^S)/\partial\lambda^S > 0. \tag{15}$$

This reflects again the 'complementarity in agglomeration' between the skilled and the unskilled. By symmetry, it is readily verified that $\lambda^U(1/2) = 1/2$. Furthermore, it is easy to check that $\lambda^U(\lambda^S) < 1$ for all parameter values of the model. Finally, observe that *for any given distribution λ^S of skilled workers*, $\partial U^{U*}/\partial\lambda^U = -\beta/[\lambda^U(1-\lambda^U)] < 0$. In words, for any spatial allocation where skilled workers have no incentives to move, unskilled workers have no incentives to move either. This property will ease the subsequent stability analysis of the spatial equilibrium.

Having a unique and well-behaved solution $\lambda^U(\lambda^S)$ – as well as complementarity between λ^U and λ^S – allows us to reduce an a priori complex problem with mobility of both the skilled and the unskilled to a simpler problem where we can only focus on the mobility of the skilled, but take into account the implied distribution of the unskilled. In words, the essentially two-dimensional problem boils down to a one-dimensional one, which will make the analysis of the equilibrium (and its stability) easier.

A *spatial equilibrium* is such that no worker (neither skilled nor unskilled) has an incentive to change location, conditional upon the fact that product markets clear at the equilibrium prices and factor markets clear at the equilibrium wages. Formally, a spatial equilibrium arises at $\lambda^{S*} \in (0,1)$

and $\lambda^{U*} = \lambda^U(\lambda^{S*})$ when $\Delta U^{S*}(\lambda^{S*}) = 0$; or at $\lambda^{S*} = 0$ and $\lambda^{U*} = \lambda^U(0)$ if $\Delta U^{S*}(0) \leq 0$; or at $\lambda^{S*} = 1$ and $\lambda^{U*} = \lambda^U(1)$ if $\Delta U^{S*}(1) \geq 0$. Following Fujita *et al.* (1999), an interior equilibrium is said to be stable if and only if the slope of the indirect utility differential ΔU^{S*} is negative in a neighborhood of the equilibrium, whereas the two agglomerated equilibria are always stable whenever they exist (recall that ΔV^{U*} is always downward-sloping in λ^U for any given value of λ^S, i.e., there is never locational instability because of the migration incentives of the unskilled).

4.1 No taxes but mobile unskilled labor

Appendix A.3 presents the benchmark case without taxes and with symetrically distributed immobile unskilled labor. This case has been previously analyzed by Forslid and Ottaviano (2003) and we report their results as a benchmark only. Note that our model boils down to theirs when $(t^U, t^S) = 0$ and $\beta \to \infty$, in which case $\lambda^U = 1/2$ irrespective of the spatial distribution of the skilled.

In what follows, we investigate the spatial equilibrium with either mobile unskilled workers, or positive taxes, or both. Let us first consider the case without taxes ($t^U = 0$ and $t^S = 0$) but where the unskilled are interregionally mobile. In this case, some longer calculations show that that the break-point is a solution to the following stability condition for the symmetric equilibrium:[20]

$$\frac{\partial(\Delta U^{S*})}{\partial \lambda^S}\Big|_{\lambda^S=1/2,\lambda^U(\lambda^S)} < 0 \quad \Longleftrightarrow \quad \frac{\left(\frac{S}{F}\right)^{\frac{\mu}{\sigma-1}} \mu \left(\frac{\sigma-1}{\sigma}\right)^{\mu} (\sigma-\mu)(1-\phi)\left(\frac{\phi}{2}+\frac{1}{2}\right)^{\frac{\mu}{\sigma-1}}}{\beta m^{\mu}} \\ +\mu^2(\phi-1)+(\sigma-1)\sigma(\phi-1)-\mu(1-2\sigma)(\phi+1) < 0, \tag{17}$$

where the term on the second line corresponds to the one that determines the break-point ϕ^b in Forslid and Ottaviano (2003). Since the term on the first line is always positive at $\phi = \phi^b$ and decreasing in ϕ, whereas the term on the second line is increasing in ϕ and by definition nil at ϕ^b, it is clear that the break-point with mobile unskilled labor must be *smaller that the break-point with immobile unskilled labor.* In words, symmetry must be broken and agglomeration will arise for higher values of transport costs than when unskilled labor is immobile. Furthermore, it is easy to see that symmetry-breaking must occur sooner the smaller the value of β. Of course, when $\beta = 0$, only full agglomeration can be sustained as an equilibrium since the only dispersion force vanishes (i.e., all

[20]Observe that the Jacobian

$$\mathbf{J} = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} \quad \text{of the system} \quad \begin{pmatrix} \dot\lambda^S \\ \dot\lambda^U \end{pmatrix} = \begin{pmatrix} U_1^S(\lambda) - U_2^S(\lambda) \\ U_1^U(\lambda) - U_2^U(\lambda) \end{pmatrix} \tag{16}$$

is negative definite whenever (17) holds. Both of its eigenvalues are negative if and only if $J_{11} < 0$, $J_{22} < 0$ and $J_{11}J_{22} > |J_{12}J_{21}|$. The first condition ($J_{11} < 0$) holds when (17) is satisfied, and the other two ($J_{22} < 0$ and $J_{11}J_{22} > |J_{12}J_{21}|$) are then trivially satisfied since $J_{21} = 0$ and $J_{22} < 0$ in our model. Since the foregoing conditions on J_{21} and J_{22} do not depend on the value of λ^U, some condition on J_{11} is sufficient to make sure that an interior equilibrium is locally stable.

unskilled react in the same way to regional economic differences). Last, as shown in Appendix A.4, full agglomeration can be quite naturally sustained for a larger range of transport costs when the unskilled are mobile than when they are not.

The foregoing results can be summarized as follows:

Proposition 1 (mobility of the unskilled and agglomeration) *When the unskilled are mobile across regions: (i) the capital hosts a larger share of the unskilled population; (ii) full agglomeration in the capital is more likely and can occur for higher values of transport costs than in the case where the unskilled are immobile.*

Proof. See Appendix A.4. ■

4.2 Taxes and immobile unskilled labor

Assume next that the elite levies taxes ($t^U > 0$ and/or $t^S > 0$) but that the unskilled are interregionally immobile ($\beta \to \infty$ so that $\lambda^U = 1/2$). For now, we consider given tax rates. In this case, some longer calculations show that

$$\begin{aligned}\frac{\Delta U^{S*}(t^S, t^U)}{K(t^S, t^U)} &= \frac{\Delta U^{S*}(0,0)}{K(0,0)} \\ &+\mu\left(1-\phi^2\right)\left[\lambda^S \frac{t^U(\sigma-1)}{[\lambda^S\phi + (1-\lambda^S)]^{\frac{\mu}{1-\sigma}}} + (1-\lambda^S)\frac{2t^S + t^U(\sigma-1)}{[\lambda^S + \phi(1-\lambda^S)]^{\frac{\mu}{1-\sigma}}}\right],\end{aligned} \tag{18}$$

where the first term on the right-hand side corresponds to Forslid and Ottaviano's (2003) indirect utility differential, as given by (A.4) in Appendix A.3. Since the second term on the right-hand side of (18) is unambiguously positive for all $t^U > 0$ and/or $t^S > 0$, and since what matters for the spatial equilibrium is the sign of ΔV^*, we have the following result.

Proposition 2 (taxation and agglomeration) *When compared with the no-tax case, (i) in case of an interior equilibrium, the capital hosts a larger share of the mobile skilled population under positive taxes, i.e.,* $\lambda^{S*}(t^S, t^U) \geq \lambda^{S*}(0,0)$ *when* $(t^S, t^U) > 0$*; and (ii) full agglomeration in the capital is more likely under positive taxes.*

Proof. See Appendix A.5. ■

Proposition 2 shows that taxation increases agglomeration in the capital. This is because of the redistribution of purchasing power from the periphery to the capital ('urban bias'), which entices firms and agents to locate there. In the words of Ades and Glaeser (1995, p.199), "*the political power [induces] the government to transfer resources to the capital, and these transfers will attract migrants*".[21] Before proceeding, a few remarks are in order.

[21] The conditions for full agglomeration in the other region are no longer symmetric, due to asymmetries in regional spending induced by the presence of the elite (see also Forslid and Ottaviano, 2003, pp.237–239).

First, although taxation by the elite increases agglomeration, we cannot compute the (partially agglomerated) stable interior equilibrium analytically. Indeed, the indirect utility differential is a transcendental function, which does not allow for explicit solutions in λ^S in the general case. By consequence, it is also impossible to explicitly characterize the break point for $\lambda^S \neq 1/2$, since this requires to evaluate the sign of the derivative of ΔV^{S*} at the interior equilibrium that we do not know. In this respect, our analysis faces the same problems than those encountered by Forslid and Ottaviano (2003) and Baldwin and Krugman (2004). Yet, these authors have shown that exogenous size differences or differential taxation favor the degree of agglomeration in the region having the larger market (in terms of consumption expenditure), and this result continues to hold true here.

Second, contrary to Forslid and Ottaviano (2003) and Baldwin and Krugman (2004), the value of S will be endogenously determined by the process of elite formation in our model. Observe that what usually matters in the CES model are the expenditure share μ and the trade costs τ, so that the size of the elite does not appear a priori to be of fundamental importance for the spatial equilibrium. Although this is true when the unskilled are immobile (see (18)), it no longer holds true in the case where the unskilled are mobile (see Appendix A.2). The reason is that more skilled workers generally impact on the price indices, and those drive location decisions of the unskilled. Hence, when the unskilled are mobile, the size of the elite has an impact on their location pattern. Furthermore, one should keep in mind that the size of the elite will also influence tax-setting in all cases we consider, which has itself an impact on the spatial equilibrium and which is the focus of our subsequent analysis.[22]

Last, observe that the degree of agglomeration monotonically increases with the taxation of the unskilled:

$$\frac{\partial\left(\Delta U^{S*}/K\right)}{\partial t^U} = \mu(\sigma-1)(1-\phi^2)\left\{\frac{1-\lambda^S}{[(1-\lambda^S)\phi+\lambda^S]^{\frac{\mu}{1-\sigma}}} + \frac{\lambda^S}{[\lambda^S\phi+(1-\lambda^S)]^{\frac{\mu}{1-\sigma}}}\right\} > 0.$$

The reason is, as in the foregoing, that higher elite expenditure in the capital attracts mobile firms and agents (both skilled and unskilled).

4.3 Taxes and mobile unskilled labor

Last, let us consider the spatial equilibrium in the most general case with mobile unskilled labor *and* taxation. As should be clear from the foregoing results, summarized in Propositions 1 and 2, both unskilled labor mobility and taxation make the occurence of agglomeration of the skilled and the

[22]The spatial equilibrium would depend even more strongly on S if skilled workers and members of the elite had different expenditure shares for differentiated goods. An additional channel through which the size of the elite, and thus the number of firms, is likely to influence the spatial equilibrium is through more product-market competition in frameworks that allow for variable demand elasticities and pro-competitive effects (e.g., Ottaviano *et al.*, 2002; Behrens and Murata, 2007).

unskilled in the capital more likely. In the former case, the symmetric equilibrium becomes unstable for a larger range of trade costs; in the latter case, the indirect utility differential shifts upwards, which increases the value of λ^{S*} at any stable interior equilibrium and makes full agglomeration more likely. Unfortunately, the analytical expressions become very cumbersome and no longer informative when unskilled labor mobility and taxation are jointly taken into consideration. Since both the slope and the locus of the indirect utility differential are affected by unskilled mobility and taxation, the analysis no longer yields clear results. Yet, we can easily depict the different cases numerically as in Figure 2.

Insert Figure 2 about here.

Figure 2 illustrates the four possible cases of: (a) immobile unskilled labor and no taxes; (b) mobile unskilled labor and no taxes; (c) immobile unskilled labor and taxes; and (d) mobile unskilled labor and taxes. Comparing panels (a) and (b), we can see that unskilled mobility does not fundamentally shift the locus ΔU^{S*} yet does make the symmetric equilibrium less stable. Comparing panels (a) and (c), one can see that taxation, and the associated expenditure shift towards the capital, shifts the locus ΔU^{S*} upwards. As a result, more skilled workers agglomerate in the capital as the market size expands there. Last, putting both taxes and mobility together, we can see by comparing (a) and (d) that both reinforce each other and lead to more (eventually full as shown in the figure) agglomeration of the mobile skilled in the capital. All these results mirror those established analytically in Propositions 1 and 2.

5 Elite formation and taxation

In this section, we focus on the dual issues of *elite formation* and *tax setting* by the elite.

5.1 Elite formation

Only the skilled have the opportunity of entering the political elite of the country. Every skilled worker therefore faces the binary choice of: (i) staying in the production sector, earning an after-tax wage of $w_i(1 - t^S)$ when he is located in region i; or (ii) leaving the production sector to become part of the political elite, thereby securing a claim to the revenues generated by taxation. In the latter case, the worker has to move to the capital since the benefits of the elite can only be reaped by being close to the center of power. Recalling that E denotes the mass of elite agents, the mass of productive skilled workers is given by $S = \overline{S} - E$. As pointed out in Section 3.3, we assume that skilled workers' only cost of entering the elite is to forfeit the wage they could have earned in the production sector.[23] The benefit of doing so is to secure a claim to an equal share of tax rents levied by the elite. Formally, skilled workers will choose to enter the elite if doing so yields a higher indirect utility than staying in the productive sector.

In what follows we focus, for analytical tractability, on the case with full agglomeration of all skilled into the capital region 1 (i.e., $\lambda^{S*} = 1$).[24] We know from Proposition 2 that the presence of an elite reinforces agglomeration, which implies that we can focus on the case in which there is dispersion in the absence of an elite, whereas there is full agglomeration in its presence (numerical examples are easy to construct). Although this scenario is a special case, our results extend to the cases with partial agglomeration, as shown at the end of this section with the help of a numerical example.

We start with the question of how the unproductive political elite is formed, and how it sets taxes to maximize its well-being. Let

$$\Delta U^{E*} = \frac{r^*}{\mathbb{P}_1^{\mu}} - \frac{w_1^*(1-t^S)}{\mathbb{P}_1^{\mu}}, \tag{19}$$

stand for the indirect utility differential between a member of the elite and a productive skilled in the capital. Elite formation will take place until indirect utilities are equalized. Using the elite members' income r^*, given by (10), the value of E will be determined such that

$$S\frac{(\sigma-1)t^U + t^S}{E}[\lambda^S w_1^* + (1-\lambda^S)w_2^*] - w_1^*(1-t^S) = 0.$$

Letting $\lambda^{S*} = 1$ and using $S \equiv \bar{S} - E$, the mass E of the elite will be determined such that

$$(\bar{S} - E)\frac{(\sigma-1)t^U + t^S}{E} - (1-t^S) = 0. \tag{20}$$

The unique solution E^* to (20) is such that:

$$E^* = \bar{S}\left[1 - \frac{1-t^S}{1+t^U(\sigma-1)}\right], \quad \text{and hence} \quad S^* = \bar{S}\frac{1-t^S}{1+t^U(\sigma-1)}. \tag{21}$$

Expression (21) shows that there is no political elite when there is no taxation ($t^S = t^U = 0$), whereas when rent-extraction from the skilled is complete ($t^S = 1$) all skilled workers want to be part of the elite. Furthermore, taxation reduces the mass of productive skilled by inflating the unproductive

[23]At an aggregate level, there is an additional cost of workers moving into the elite: the range of consumption varieties decreases. Yet, since each worker is negligible in the aggregate economy, this is not a cost stemming from an individual decision to enter the elite and is, therefore, disregarded when making that choice.

[24]This assumption is justified by the fact that the elite would always choose to implement full agglomeration in its region of residence because this maximizes its access to product variety (see also Charlot *et al.*, 2005, Prop. 3). A convenient by-product of letting $\lambda^{S*} = 1$ is that we can draw clearer analytical conclusions since the optimal tax-setting behavior of the elite does not depend on the spatial distribution of the unskilled (which is not true at an interior equilibrium, as discussed at the end of this section).

elite, *thereby decreasing product variety and consumption benefits.* This latter aspect may be quite important in practice but has, to the best of our knowledge, not really been highlighted in models of agglomeration and rent-seeking until now. Some straightforward calculations further show that $(\partial E^*)/(\partial \sigma) > 0$. These results may be summarized as follows:

Proposition 3 (size of the elite) *Under full agglomeration in the capital, the size of the unproductive elite decreases with the degree of product differentiation (smaller σ) and increases with both tax rates t^S and t^U.*

Note that a smaller degree of product differentiation (higher σ) tends to inflate the size of the unproductive elite. The intuition underlying this result is that when firms in the manufacturing sector have little market power they pay lower wages, which raises the payoff to political rent-seeking when compared to production. Everything else equal, this fosters entry into the elite.

Insert Figure 3 about here.

Turning to the size of the elite in the absence of full agglomeration, we have to resort to numerical illustrations to gauge the qualitative behavior of the model. As can be seen from Figure 3, increasing taxation of the unskilled leads to (a) an increase in the equilibrium size of the elite, (b) an associated decrease in the size of the productive skilled workforce, (c) more agglomeration of the skilled in the capital, and (d) increasing agglomeration of the unskilled in the capital. All of these results are in line with the qualitative behavior of our model established in the previous sections and in expression (21). It is thus fair to say that, even though the case with $\lambda^{S*} = 1$ is a special one, it's qualitative behavior is not radically different from the one we observe at an interior equilibrium.

5.2 Tax setting

The final stage involves tax setting by the elite, taking into account how its decisions influence subsequent entry into the elite by skilled workers. In what follows, we assume that the elite sets tax rates such as to maximize its indirect utility and not necessarily its tax revenue. Stated differently, the elite takes into account the fact that higher taxes and the subsequent entry into the elite reduces product variety and welfare by diverting productive resources.[25]

[25]In public economic terms, the elite is not a leviathan but a welfare maximizer (for itself). Note that the leviathan case is likely to lead to an even larger elite as it disregards the negative variety effect. Note also that the 'variety constraint' faced by the elite, which limits its ability to extract rents, may be less stringent nowadays because imports can supplement the needs of the elite. Stated differently, international trade may be a driver for agglomeration in a setting where an unproductive elites maximizes its welfare since it increases their extractive capacity. Though interesting, we disregard this possibility in what follows as modeling these issues is beyond the scope of this paper.

The indirect utility of a member of the political elite is given by $U^{E*} = r^*/P_1^\mu$. Letting $S^* = \overline{S} - E^*$, $\lambda^{S*} = 1$, and using r^* as given by (10), we then obtain:

$$U^{E*}\mid_{\lambda^S=1} = \left[\frac{(\sigma-1)(1-t^U)}{\sigma m}\right]^\mu \left(\frac{\overline{S}-E^*}{F}\right)^{\frac{\mu}{\sigma-1}} \frac{(\overline{S}-E^*)[(\sigma-1)t^U + t^S]}{E^*} w_1^*.$$

Substituting the expression of E^* as given by (21), and using w_1^*, we finally get

$$\Delta U^{E*}\big|_{\lambda^S=1} \;=\; \kappa(1-t^S)^{\frac{\mu}{\sigma-1}}\xi(t^U), \tag{22}$$

where

$$\xi(t^U) \equiv \frac{\left(1-t^U\right)^\mu \left[(\sigma-1)t^U+1\right]^{1-\frac{\mu}{\sigma-1}}}{\sigma-\mu-(\sigma-1)t^U\mu} \tag{23}$$

and where $\kappa > 0$ is a bundle of parameters. The elite maximizes (22) with respect to the taxes t^U and t^S. Since $\mu/(\sigma-1) > 0$, we see that ΔU^{E*} is always decreasing in t^S. Stated differently, the elite sets $t^{S*} = 0$. This suffices to establish the following result:

Proposition 4 (optimal skilled taxation) *At any spatial equilibrium with full agglomeration, the elite will set a zero tax rate on the skilled ($t^{S*} = 0$).*

The intuition underlying Proposition 4 is that $V^{E*} = V^{S*}$ by arbitrage, so that the objective of maximizing the welfare of the elite is congruent with the objective of maximizing the welfare of the skilled. Consequently, the skilled will not be taxed. Results with respect to the unskilled tax rate t^U are more complicated to establish. Yet, we can show the following:

Proposition 5 (optimal unskilled taxation) *At any spatial equilibrium with full agglomeration, there exists a threshold*

$$\sigma^t \equiv \frac{1}{2} + \mu + \frac{1}{2}\sqrt{1+4(1-\mu)\mu} < 2$$

such that (i) $t^{U*} = 0$ and there is no entry into the elite ($E^* = 0$) when $\sigma \leq \sigma^t$; and (ii) there is a unique $0 < t^{U*} < 1$ with entry into the elite ($E^* > 0$) when $\sigma > \sigma^t$.

Proof. See Appendix A.6. ■

Proposition 5 shows that when products are sufficiently differentiated and firms have enough market power, there will be no elite formation in equilibrium, even if it is a priori possible. The intuition underlying this result is that, since it is profitable enough to work in the production sector, agents cannot win by pursuing unproductive rent-seeking. This finding suggests that a strong reliance on relatively homogeneous products, as often observed in developing countries, may be a strong driver for elite formation as the relative payoff of unproductive to productive activity is large. Entry of skilled into the elite both erodes per capita rents in a standard way and, by taxing agents and thus raising production costs, increases wages in the modern sector. Both effects are penalizing and may be strong enough to prevent elite formation in equilibrium.

Insert Figure 4 about here.

It can be shown that t^{U*} is strictly decreasing in μ when $\sigma^t < \sigma \leq 2$, whereas it is first decreasing and then increasing in μ when $\sigma > 2$ (see panel (a) in Figure 4 for the latter case). The intuition underlying the second effect is that when expenditures on the modern sector account only for a small part of their budget, consumers are willing to buy even when the price is high, which allows for more taxation of the unskilled without reducing tax revenues and product variety. Clear-cut comparative static results with respect to σ are not easily derived since the various thresholds are parametrized by σ. Yet, numerical examples suggest that, as expected, t^{U*} is increasing in σ (see panel (b) in Figure 4 for an illustration). The less differentiated the varieties are, the higher the unskilled tax rate. The reason is that higher unskilled taxes raise production costs in the differentiated sector, thereby eroding profits, triggering entry into the elite and reducing product diversity – the latter effect is, however, not too penalizing when diversity matters only little for welfare.

To summarize our main findings, both a large expenditure share on the agricultural good and little product differentiation in the manufacturing sector are drivers for high tax rates on the unskilled and a large elite, as they erode the returns to production activity. Both of these features seem to be largely prevalent in the developing world and may, therefore, be linked to the presence of a sizeable unproductive elite.[26]

6 Conclusions

The model we presented suggests that both unskilled labor mobility and rent-seeking behavior may fuel the formation of large urban centers in developing countries via interregional income shifts and transfers. Even when increasing returns and the freeness of trade are too low for agglomeration to occur in equilibrium in the 'standard' NEG model, spatial concentration may nevertheless arise because of the additional purchasing power generated in the capital by either unskilled labor mobility or tax-based transfers from the country-side.[27] We have shown that rent-seeking will always lead to the formation of an elite provided that goods are not too differentiated. Furthermore, the elite sets

[26] As pointed out in the foregoing, the case where the elite is not constrained to be fully agglomerated in the capital is more complicated to analyze as analytical solutions for λ^{S*} are not available. Knowing the spatial distribution of the unskilled λ^{U*} for any given triple (λ^S, t^S, t^U), we could solve the system of equations $\Delta U^S = 0$, $\partial U^S / \partial t^S = 0$ and $\partial U^S / \partial t^U = 0$ for those three variables to simulate the model. We conjecture that results would be similar.

[27] A particularily illuminating example of high trade costs in developing countries is detailed in the article "The road to hell is unpaved" of the December 19, 2002, print edition of The Economist (available online at `http://www.economist.com/ PrinterFriendly.cfm?Story_ID=1487583`). Due to administrative hassle, 47 roadblocks and poor infrastructure, a 500 kilometer trip by truck from Douala to Bertoua took four days, with only two-thirds of the original load of Guinness arriving at its destination (a real 'iceberg' trade cost). Bad infrastructure and high transaction costs therefore add substantially to total shipping costs.

zero tax rates for skilled workers when maximizing its own welfare. This result is consistent with the observation that the main bulk of taxation falls heavily on the poor in developing countries. The smaller the elite, the higher the incentives for skilled to engage in rent-seeking behavior because the payoff in doing so rises. If we proxy the share of skilled by that of people with higher education, we see that SSA (with only 2.43% of enrollment in higher education; Barro and Lee, 2001) has one of the worst performances in this respect. When combined with the low returns to private investment, the payoff to skilled of entering the political elite may be quite large in this region of the world.

Other factors also contribute to increase the size of the elite in SSA. Indeed, we have shown that in equilibrium the size of the elite is inversely related to the share of manufacturing expenses in national income (via the equilibrium tax rates) and to the degree of product differentiation. As a matter of fact, low product differentiation is likely to be a feature of SSA, because economic development is positively correlated with product diversity (Falkinger and Zweimüller, 1996). Moreover, most SSA countries are characterized by small manufacturing sectors and small expenditure shares on manufactured goods. SSA was in 2002 the region exhibiting the lowest ratio of manufacturing value added to GDP (13.6% versus 15.8%, 18.2%, 20.1% and 33.1% for Latin America & Caribbean, Europe & Central Asia, the European Union, and East Asia & Pacific, respectively). Looking only at SSA is even more informative: of the 43 SSA countries, 23 have ratios of less than 10% and 9 have ratios of even less than 5%.[28] This suggests that the payoff to productive work is quite low, enticing skilled workers to engage in rent-seeking activities, which might drive urban primacy.

Observe finally that the presence of a rent-seeking elite also reduces the range of available goods because it diverts productive resources. In an endogenous growth setting where varieties also serve as intermediate inputs (Grossman and Helpman, 1991), less variety due to the presence of the elite may damage long-run growth, which might lead to a relatively bad growth performance as observed in SSA. Indeed, that region has witnessed an urban growth rate significantly outpacing its economic growth rate over the last two decades.

Acknowledgements. We thank two anonymous referees, the editor Takatoshi Tabuchi, Luisito Bertinelli, Yasusada Murata, Dominique Peeters, Pierre M. Picard, Debraj Ray, Frédéric Robert-Nicoud, Yasuhiro Sato, and Jacques Thisse for helpful comments and suggestions. Kristian Behrens gratefully acknowledges financial support from the European Commission under the Marie Curie Fellowship MEIF-CT-2005-024266. Kristian Behrens is holder of the *Canada Research Chair in Regional Impacts of Globalization.* Financial support from the CRC Program of the Social Sciences and Humanities Research Council (SSHRC) of Canada is gratefully acknowledged. Kristian Behrens also gratefully acknowledges financial support from FQRSC Québec (Grant NP-127178). Alain Pholo Bala

[28] All figures are taken from the World Bank Development Indicators, 2005, available online at the following address: `http://devdata.worldbank.org/data-query`.

would like to thankfully recognize the financial support from Economic Research Southern Africa (ERSA), South Africa. The views expressed in this paper and any remaining errors are ours.

References

[1] Ades, A. (1995) Economic development with endogenous political participation, *Economics and Politics* 7, 93–117.

[2] Ades, A. and E.L. Glaeser (1995) Trade and circuses: Explaining urban giants, *Quarterly Journal of Economics* 110, 195–227.

[3] Aryeetey-Attoh, S. (1997) *The Geography of Sub-Saharan Africa*. Upper Saddle River: Prentice-Hall.

[4] Au, C.-C. and J.V. Henderson (2006) How migration restrictions limit agglomeration and productivity in China, *Journal of Development Economics* 80, 350–388.

[5] Bairoch, P. (1985) *De Jericho à Mexico: Villes et économie dans l'histoire*. Paris: Gallimard. English translation *Cities and Economic Development. From the Dawn of History to the Present* (1988), London: Mansell Publishing Ltd.

[6] Baldwin, R.E., R. Forslid, Ph. Martin, G.I.P. Ottaviano and F. Robert-Nicoud (2003) *Economic Geography and Public Policy*. Princeton: Princeton Univ. Press.

[7] Baldwin, R.E. and P.R. Krugman (2004) Agglomeration, integration and tax harmonization, *European Economic Review* 48, 1–23.

[8] Barrios, S., L. Bertinelli and E. Strobl (2006) Climatic change and rural-urban migration: The case of sub-Saharan Africa, *Journal of Urban Economics* 60, 357-371.

[9] Barro, R. J. and J.-W. Lee (2001) International data on educational attainment: updates and implications, *Oxford Economic Papers* 3, 541–563.

[10] Bassey, M.O. (1999) *Western Education and Political Domination in Africa. A Study in Critical and Dialogical Pedagogy*. Westport: Bergin & Garvey.

[11] Behrens, K. and Y. Murata (2007) General equilibrium models of monopolistic competition: a new approach, *Journal of Economic Theory* 136, 776–787.

[12] Bryceson, D.F. and D. Potts (2006) *African Urban Economies: Viability, vitality, or vitiation*. London: Palgrave Macmillan.

[13] Cantillon, R. (1730) *Essai sur la Nature du Commerce en Général*. Full english version available online at `http://www.econlib.org/library/NPDBooks/Cantillon /cntNTContents.html`

[14] Charlot, S., C. Gaigné, F. Robert-Nicoud and J.-F. Thisse (2006) Agglomeration and welfare: the core-periphery model in the light of Bentham, Kaldor, and Rawls, *Journal of Public Economics* 90, 325–347.

[15] Collier, P. (2006) *Africa: Geography and Growth*. Mimeo.

[16] Dascher, K. (2000) Are politics and geography related? Evidence from a cross-section of capital cities, *Public Choice* 105, 373–392.

[17] Davis, J.C. and J.V. Henderson (2003) Evidence on the political economy of the urbanization process, *Journal of Urban Economics* 53, 98–125.

[18] Duranton, G. and D. Puga (2004) Micro-foundations of urban agglomerations economies. In: Henderson, J.V. and J.F. Thisse (eds.) *Handbook of Regional and Urban Economics, vol. 4*, North-Holland: Elsevier Science Publishers, pp. 2063–2117.

[19] Falkinger, J. and J. Zweimüller (1996) The cross-country Engel-curve for product diversification, *Structural Change and Economic Dynamics* 7, 79–97.

[20] Forslid, R. and G.I.P. Ottaviano (2003) An analytically solvable core-periphery model, *Journal of Economic Geography* 3, 229–240.

[21] Fujita, M., P.R. Krugman and A.J. Venables (1999) *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

[22] Fujita, M. and T. Mori (1996), The role of ports in the making of major cities: Self-agglomeration and hub-effect, *Journal of Development Economics* 49, 93-120

[23] Fujita, M. and J.-F. Thisse (2002) *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge: Cambridge Univ. Press.

[24] Grossman, G.M., and E. Helpman (1991) *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press.

[25] Kanbur, R. and A.J. Venables (2003) Editor's introduction, *Journal of African Economies* 12, 473–475.

[26] Krugman, P. (1991) Increasing Returns and Economic Geography, *Journal of Political Economy* 99, 483–499.

[27] Lipton, M. (1983) *Why Poor People Stay Poor: A Study of Urban Bias in World Development*. London: Temple Smith.

[28] Lozano-Garcia, N., G. Piras, A.M. Ibanes, and G.J.D. Hewings (2010) The journey to safety: Conflict-driven migration flows in Colombia, *International Regional Science Review* 33, 157–180.

[29] Mabogunje, A.L. (1994) Introduction: Cities and Africa's Economic Recovery. In: Tarver, J.D. (ed.) *Urbanization in Africa: A Handbook*. Westport: Greenwood Press, pp. xxi–xxxii.

[30] Max Miller, H. and R.N. Singh (1994) Urbanization during the Postcolonial Days. In: Tarver, J.D. (ed.) *Urbanization in Africa: A Handbook*. Westport: Greenwood Press, pp. 408–422.

[31] Mbaku, J.M. (2003) Corruption and economic development. In: Nnadozie, E. (ed.) *African Economic Development*, Academic Press, San Diego, pp. 313–349.

[32] Mbeki, M. (2005) Underdevelopment in sub-Saharan Africa: the role of the private sector and political elites, CATO Institute Foreign Policy Briefing #85.

[33] McGranahan, G., D. Mitlin, D. Satterthwaite, C. Tacoli and I. Turok (2009) Africa's urban transition and the role of regional collaboration, Human Settlements Working Paper Series Theme: Urban Change #5.

[34] Murata, Y. (2003) Product diversity, taste heterogeneity, and geographic distribution of economic activities: market vs. non-market interactions, *Journal of Urban Economics* 53, 126–144.

[35] Oaxaca, R.L. and I. Geisler (2003) Fixed effects models with time invariant variables: a theoretical note, *Economics Letters* 80, 373-377.

[36] Ottaviano, G.I.P., T. Tabuchi and J.-F. Thisse (2002) Agglomeration and trade revisited, *International Economic Review* 43, 409–436.

[37] Page, J. (2008) Rowing against the current. The diversification challenge in Africa's resource-rich economies, Brookings Global Economy and Development Working Paper #29.

[38] Puga, D. (1998) Urbanization patterns: European vs less developed countries, *Journal of Regional Science* 38, 231–252.

[39] Robert-Nicoud, F. and F. Sbergami (2004) Home-market vs. vote-market effect: Location equilibrium in a probabilistic voting model, *European Economic Review* 48, 155–179.

[40] Robert-Nicoud, F. (2005) The structure of simple 'New Economic Geography' models (or, On identical twins), *Journal of Economic Geography* 5, 201–234.

[41] Satterthwaite, D. (2005) The scale of urban change worldwide 1950-2000 and its underpinnings. Mimeographed.

[42] Schneider, F. (2004) The size of the shadow economies of 145 countries all over the world: First results over the period 1999-2003, IZA Discussion Paper #1431.

[43] Tabuchi, T. and J.-F. Thisse (2002) Taste heterogeneity, labor mobility and economic geography, *Journal of Development Economics* 69, 155–177.

[44] United Nations (2002) *World Urbanization Prospects: The 2001 Revision. Data Tables and Highlights*. New York: United Nations Secretariat, Population Division.

[45] Williamson, J. (1965) Regional inequality and the process of national development, *Economic Development and Cultural Change* 13, 3–45.

[46] World Bank (2008) *World Development Report 2009: Reshaping Economic Geography*. Washington D.C.: The World Bank.

Appendix A: Proofs and technicalities

A.1. Skilled wages

Some standard but longer computations yield:

$$\begin{aligned} w_1^* &= \frac{\sigma\phi[\lambda^S(1-\phi)+\phi]+(1-\lambda^S)(1-\phi^2)[t^S\mu+\lambda^U(\sigma-\mu)+t^U\mu(\sigma-1)(1-\lambda^U)]}{D(t^S,t^U,\lambda^S)}\mu\overline{U} \\ w_2^* &= \frac{\sigma\phi+\lambda^S(1-\phi)\left[\sigma-(1+\phi)(\mu(1+t^U(\sigma-1))(1-\lambda^U)+\lambda^U(\sigma-\mu))\right]}{D(t^S,t^U,\lambda^S)}\mu\overline{U}, \end{aligned}$$

where

$$\begin{aligned} D(t^S,t^U,\lambda^S) \;\equiv\; & S[\sigma-\mu(1+t^U(\sigma-1))] \\ & \times\left\{\sigma\phi+\lambda^S(1-\lambda^S)(1-\phi)[\sigma(1-\phi)-(1-t^S)\mu(1+\phi)]\right\}>0 \qquad \text{(A.1)} \end{aligned}$$

for all t^S, t^U and λ^S given the restrictions $0\le\lambda^S\le1$, $0\le\lambda^U\le1$, $\sigma>1$ and $0<\mu<1$.

A.2. Indirect utility differential

The indirect utility differential of the skilled can be expressed as follows:

$$\begin{aligned}\frac{\Delta U^{S*}(t^S,t^U)}{K(t^S,t^U)} &= (\lambda^S+\phi(1-\lambda^S))^{\frac{\mu}{\sigma-1}}\Big[\sigma\phi(\lambda^S+(1-\lambda^S)\phi) \\ &+ (1-\lambda^S)\left(1-\phi^2\right)\left(t^S\mu+t^U(1-\lambda^U)(\sigma-1)\mu+\lambda^U(\sigma-\mu)\right)\Big] \\ &- ((1-\lambda^S)+\phi\lambda^S)^{\frac{\mu}{\sigma-1}}\Big[\lambda^S(1-\phi)(\sigma+((\lambda^U-1)\mu(t^U(\sigma-1)+1) \\ &- \lambda^U\sigma)(\phi+1))+\sigma\phi\Big],\end{aligned} \tag{A.2}$$

where $\lambda^U=\lambda^U(\lambda^S)$ is given by

$$\lambda^U(\lambda^S)=\left\{1+\exp\left(\frac{\left(\frac{S}{F}\right)^{\frac{\mu}{\sigma-1}}\left[\frac{(1-t^U)(\sigma-1)}{m\sigma}\right]^{\mu}\left[(\phi\lambda^S-\lambda^S+1)^{\frac{\mu}{\sigma-1}}-(\lambda^S+(1-\lambda^S)\phi)^{\frac{\mu}{\sigma-1}}\right]}{\beta}\right)\right\}^{-1},$$

where

$$K(t^S,t^U)\equiv\frac{(1-t^S)\mu\overline{U}}{D(t^S,t^U,\lambda^S)}\left[\frac{m\left(\frac{S}{F}\right)^{\frac{1}{1-\sigma}}\sigma}{(1-t^U)(\sigma-1)}\right]^{-\mu}>0 \tag{A.3}$$

and where $D(t^S,t^U,\lambda^S)$ is given by (A.1) in Appendix A.1. Observe that since K is strictly positive, it does not affect the sign of the right-hand side of expression (A.2) and may thus be neglected in the determination of the spatial equilibrium (even though that term does depend on λ^S). Yet, we need to take it into account when evaluating the derivative of $\Delta U^{S*}(t^S,t^U)$.

A.3. Benchmark case

Let us start with the case where $t^S=t^U=0$ and $\beta\to\infty$. It is readily verified that the latter condition implies that $\lambda^U=1/2$ for all values of λ^S, which corresponds to the benchmark case of Forslid and Ottaviano (2003). In that case, ΔU^{S*} is given by

$$\frac{\Delta U^{S*}(0,0)}{K(0,0)}=\frac{2\lambda^S\sigma\phi+(1-\lambda^S)\left[\sigma\left(1+\phi^2\right)-\mu\left(1-\phi^2\right)\right]}{\left[(1-\lambda^S)\phi+\lambda^S\right]^{\frac{\mu}{1-\sigma}}}-\frac{2(1-\lambda^S)\sigma\phi+\lambda^S\left[(\mu+\sigma)\phi^2+(\sigma-\mu)\right]}{\left[\lambda^S\phi+(1-\lambda^S)\right]^{\frac{\mu}{1-\sigma}}}.$$

As shown by Forslid and Ottaviano (2003), full agglomeration may be sustained as an equilibrium if and only if

$$\frac{\Delta U^{S*}(0,0)}{K(0,0)}\Big|_{\lambda^S=1}=-\frac{\Delta U^{S*}(0,0)}{K(0,0)}\Big|_{\lambda^S=0}=2\sigma\phi-\frac{(\mu+\sigma)\phi^2+(\sigma-\mu)}{\phi^{\frac{\mu}{1-\sigma}}}>0, \tag{A.4}$$

which implicitly defines the sustain point ϕ^s as the value of ϕ that equates the above expression to zero. Full agglomeration can be sustained for all $\phi\geq\phi^s$. Additionally, there are at most three interior equilibria in the no-tax case (Robert-Nicoud, 2005), of which the symmetric one ($\lambda^{S*} = 1/2$) always exists. The stability of the equilibrium $\lambda^{S*} = 1/2$ depends on the sign of the derivative of the indirect utility differential, whereas the other two interior equilibria are always unstable. Computing $\partial(\Delta U^{S*})/\partial\lambda^S$ and evaluating it at $\lambda^S = 1/2$, the break-point is such that

$$\phi^b \equiv \frac{\sigma - \mu}{\sigma + \mu} \frac{\mu - \sigma + 1}{1 - \mu - \sigma}. \tag{A.5}$$

Hence, $\lambda^* = 1/2$ is a stable spatial equilibrium for all $\phi \leq \phi^b$. Note, finally, that both types of equilibria occur for values $\phi^s \leq \phi \leq \phi^b$, in which case both full agglomeration and full dispersion are stable spatial equilibria.

A.4. Proof of Proposition 1

Proof. We have

$$\begin{aligned} &\Delta U^{S*} \Big|_{\lambda^S=1, \lambda^U=\lambda^U(\lambda^S)} - \Delta U^{S*} \Big|_{\lambda^S=1, \lambda^U=1/2} \\ &= - \frac{U\mu \left[\frac{m\left(\frac{S}{F}\right)^{\frac{1}{1-\sigma}} \sigma \phi^{\frac{1}{1-\sigma}}}{\sigma - 1} \right]^{-\mu} (1 - \phi^2) \tanh \left[\frac{m^{-\mu} \left(\frac{S}{F}\right)^{\frac{\mu}{\sigma-1}} \left(\frac{\sigma-1}{\sigma}\right)^{\mu} \left(\phi^{\frac{\mu}{\sigma-1}} - 1\right)}{2\beta} \right]}{2S\sigma\phi} > 0 \end{aligned}$$

since $\tanh(x) < 0$ for all $x < 0$. Hence, full agglomeration can be sustained for a larger range of parameter values when the unskilled are mobile than when they are not. ■

A.5. Proof of Proposition 2

Proof. In both cases, let λ_0^{S*} denote the value of λ^S such that $\Delta U^{S*}(0,0) \big|_{\lambda^S=\lambda_0^{S*}} = 0$. Evaluating (18) at this value yields

$$\frac{\Delta U^{S*}(t^S, t^U)}{K(t^S, t^U)} \Big|_{\lambda^S=\lambda_0^{S*}} = \mu \left(1 - \phi^2\right) \left\{ \lambda_0^{S*} \frac{t^U(\sigma - 1)}{[\lambda_0^{S*}\phi + (1 - \lambda_0^{S*})]^{\frac{\mu}{1-\sigma}}} + (1 - \lambda_0^{S*}) \frac{2t^S + t^U(\sigma - 1)}{[\lambda_0^{S*} + \phi(1 - \lambda_0^{S*})]^{\frac{\mu}{1-\sigma}}} \right\}$$

which is strictly positive. This shows that: (i) when there is a dispersed equilibrium without taxes, the equilibrium with taxes must involve more agglomeration since the utility differential at that no-tax spatial equilibrium is strictly positive; and (ii) when there is full agglomeration without taxes, there is also full agglomeration with taxes since

$$\frac{\Delta U^{S*}(t^S, t^U)}{K(t^S, t^U)} \Big|_{\lambda^S=1} = \frac{\Delta U^{S*}(0,0)}{K(0,0)} \Big|_{\lambda^S=1} + \mu(1 - \phi^2) t^U (\sigma - 1) \phi^{\frac{\mu}{\sigma-1}} \tag{A.6}$$

is strictly positive because $\Delta U^{S*}(0,0)/K(0,0) \big|_{\lambda^S=1} > 0$ (since full agglomeration is sustainable without taxes). Finally, since the second term in (A.6) is strictly positive we see that: (iii) there are parameter values for which full agglomeration is sustainable with taxation, whereas it is not in the absence of taxation. When taken together, conditions (i), (ii), and (iii) show that taxation cannot make the economy less agglomerated in the capital, which establishes Proposition 2. ■

A.6. Proof of Proposition 5

Proof. To prove our claim, we need to analyze the function ξ, as given by (23). First, note that ξ is continuous in t^U and non-negative on the interval $[0,1]$. Furthermore, since $\mu < \sigma - 1$ by assumption, we have

$$\lim_{t^U \to 1} \xi(t^U) = 0 \qquad \text{and} \qquad \lim_{t^U \to 0} \xi(t^U) = \frac{1}{\sigma - \mu} > 0. \tag{A.7}$$

Stated differently, prohibitive taxation leads to zero tax revenue and zero utility for the elite, whereas sufficiently low taxation yields positive utility. Second, one can check that $\partial\xi/\partial t^U = 0$ if and only if

$$t^U = \frac{(4-3\sigma)\mu^2 + (\sigma-2)\sigma\mu + \sigma\left\{\sigma \pm \sqrt{(1-\mu)(\sigma-\mu-1)[\sigma+\mu(\mu+3\sigma-4)-1]} - 1\right\}}{2\mu^2(\sigma-2)(\sigma-1)}.$$

It is readily verified that the larger of the two roots exceeds 1 and, therefore, does not belong to the range of admissible values. Consequently, ξ is either monotonously decreasing on $[0,1]$, or admits a unique maximum (the existence of an interior minimum is ruled out by (A.7) and because $\xi \geq 0$). Some standard computations show that

$$\lim_{t^U \to 0} \frac{\partial\xi(t^U)}{\partial t^U} = \frac{2\mu^2 - 2\sigma\mu + (\sigma-1)\sigma}{(\mu-\sigma)^2}.$$

Equating this expression to zero, the two roots with respect to σ are such that the smaller one is less than 1 and must be ruled out, whereas the larger one is given by:

$$\sigma^t = \frac{1}{2} + \mu + \frac{1}{2}\sqrt{1 + 4(1-\mu)\mu} < 2.$$

Hence, when $\sigma < \sigma^t$ the function ξ is strictly decreasing for all tax rates and the elite's best choice is to not tax the unskilled. Obviously, there will be no elite in that case. When $\sigma > \sigma^t$, the elite chooses $0 < t^{U*} < 1$ and there will be elite formation. The optimal tax rate is therefore given as follows:

$$t^{U*} = \begin{cases} 0 & \text{if} \quad 1 < \sigma < \sigma^t \\ \frac{(4-3\sigma)\mu^2+(\sigma-2)\sigma\mu+\sigma\left\{\sigma-\sqrt{(1-\mu)(\sigma-\mu-1)[\sigma+\mu(\mu+3\sigma-4)-1]}-1\right\}}{2\mu^2(\sigma-2)(\sigma-1)} & \text{if} \quad \sigma^t < \sigma,\ \sigma \neq 2 \\ \frac{1-\mu}{1+\mu} & \text{if} \quad \sigma = 2 \end{cases}$$

which establishes Proposition 5. ■

Appendix B: Data and descriptives

The panel dataset covers the 1980–2005 period in five years interval. The "city states" of Hong Kong and Singapore and Macau were excluded from estimations and descriptive statistics. It draws on publicly available data only. The dataset, as well as the list of countries, is available upon request (Table B.1 below provides summary statistics). The following variables are available:

1. *Urban primacy* is defined as the percentage of urban population living in the country's largest city. The panel data are obtained from the World Bank online data `http://data.worldbank.org/`. Retrieved on 2011-05-10.

2. *GDP per capita, PPP* $US is taken from the World Bank online data, available at `http://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD?page=1`. Retrieved on 2011-05-10.

3. *Urban population* available online at `http://data.worldbank.org/indicator/SP.URB.TOTL`. Retrieved on 2011-07-06.

4. *Democracy index* used is the *POLITY2* Revised Combined Polity Score proposed by the Polity IV project (Political Regime Characteristics and Transitions, 1800–2009). This score is obtained by subtracting the Autocracy score ($AUTOC$) from the Democracy score ($DEMOC$); the resulting unified polity scale ranges from -10 (strongly autocratic) to +10 (strongly democratic). Data and full documentation on variables are available online at `http://www.systemicpeace.org/polity/polity4.htm`. Retrieved on 2011-06-06.

5. *Non corruption index* (5 previous years) used is the *Freedom from corruption* available online from 1995 to 2011 at `http://www.heritage.org/index/explore?view=by-region-country-year`. Last accessed on 2011-07-06.

6. *Ethnolinguistic fractionalization index* used is the 1985 index proposed by Philip G. Roeder, Department of Political Science, University of California, San Diego available online at the address `http://weber.ucsd.edu/~proeder/elf.htm`. Last accessed on 2011-07-06.

7. *Data on religious affiliation* has been gathered from several sources: national census (through `http://www.wikipedia.org`), CIA World Factbook (from `https://www.cia.gov/library/publications/the-world-factbook/`), Religiously Remapped - Mapping Religious Trends In Africa (`http://www.religiouslyremapped.info/Religiously_Remapped/Homepage.html`). Last accessed on 2011-06.

8. *Data on infrastructure: roadways, railways, waterways densities & percentage of paved roads* obtained from the CIA World Factbook (`https://www.cia.gov/library/publications/the-world-factbook/`). Last accessed on 2011-06.

The source and the description of the cross-sectional data used in figure 1 are the following:

- *Urban primacy*: those data are computed from the UN Populations Division "World Urbanization Prospects Population Database: The 2009 Revision" for the year 2005, available online at `http://esa.un.org/unpd/wup/index.htm`. Retrieved on 2011-05-10.

- *Non-corruption index* is the "2005 Transparency International Corruption Perceptions Index" (CPI). It ranges from 1 (most corrupt) to 10 (least corrupt). It "[...] focuses on corruption in the public sector and defines corruption as the abuse of public office for private gain. The surveys used in compiling the CPI ask questions that relate to the misuse of public power for private benefit, with a focus, for example, on bribe-taking by public officials in public procurement." Available online at `http://www.transparency.org/policy_and_research/surveys_indices/cpi/2005`. Retrieved on 2011-05-10.

- *Central government expenditures on goods, services, and compensation of employees* in % of central government expenditures in 2005 is computed from the World Bank online data, available at `http://data.worldbank.org`. Retrieved on 2011-05-10.

Table B.1: Summary statistics (full panel sample)

Variable	Obs.	Min	Max	Mean	Std. dev.
urban primacy	1,191	0.025	1.221	0.413	0.246
ln(urban population)	1,213	8.437	20.082	14.249	2.297
ln(GDP per capita, PPP)	948	5.016	11.466	8.434	1.247
Democracy index (prev. 5 years)	727	-10	10	0.996	7.360
Non-corruption index	894	8.75	95.75	40.190	21.979
Non-corruption*SSA dummy	894	0	58.833	6.486	12.789
ln(land area)	1,260	0.693	16.611	10.878	2.976
Ethnolinguistic fractionalization index	1,008	0	0.984	0.459	0.274
Percentage of paved roads	972	0.799	100	50.971	33.289
Roadways/km2	1,200	0.005	25	0.903	2.133
Railways/km2	1,260	0	8.940	0.076	0.665
Waterways/km2	1,182	0.00	0.184	0.007	0.020
Percent catholics	1,116	0	98	31.786	32.704
Percent protestants	972	0	98.4	25.747	27.558
Percent muslims	954	0	100	30.788	38.193

Appendix C: Panel data econometric procedure

C.1. Fixed effects estimation with time invariant variables

With two-stage procedures, it is possible to identify and consistently estimate the effects of time invariant regressors in panel data settings with fixed effects. We use the two-stage GLS estimator proposed by Oaxaca and Geisler (2003). They demonstrate the equivalence between that estimator and the pooled OLS estimator with time invariant covariates. Yet, the estimated standard errors differ between these two procedures. The first stage of the GLS procedure implies the fixed effects within regression. This allows to obtain the vector $\hat{\omega}^{FE}$ of time varying covariates coefficients estimates, with $\omega = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)'$. The second step implies the following regression

$$primacy_{i.} - W_{i.}\hat{\omega}^{FE} = \alpha + Z_i\gamma + \eta_{i.} \tag{A.8}$$

where the subscript $i.$ denotes country time averages, where $\eta_{i.} = \xi_{i.} + W_{i.}\left(\omega - \hat{\omega}^{FE}\right)$, and where

$$W_{i.} = \left(ln\left(urban\ population\right)_{i.}, \left[ln\left(urban\ population\right)_{i.}\right]^2, ln\left(GDPpc\right)_{i.}, \left[ln\left(GDPpc\right)_{i.}\right]^2, X_{i.}\right)'.$$

Equation (A.8) is estimated by GLS to correct the heteroscedasticity inherent in the diagonal elements of the variance-covariance matrix of $\eta_{i.}$.

C.2. Instrumental variables estimation with time invariant variables

Some components of ξ_{it} are likely to be correlated with regressors like income and the democracy index. Furthermore, contemporaneous ξ_{it} shocks impacting migration to primate cities will also affect covariates either in the current year or in the future. To deal with those issues, we follow the methodology used by Davis and Henderson (2003). Hence, we first difference (1), yielding a set of equation-years (2005–2000, 2000–1995, 1995–1990, 1990–1985, 1985–1980). This eliminates the fixed effects as well as time invariant variables Z_i. Then, we estimate those equation-years jointly by constraining coefficients to be equal across years, considering all the time-varying regressors as potentially endogeneous. Thus, for a given equation-year t minus $t-1$ we instrument for all time varying regressors with level values of covariates in $t-2$ and $t-3$.[29]

To recover the effects of time invariant variables, we proceed as in the preceding subsection. We consider the coefficients estimated in the previous step and we compute the difference between the primacy country average and the sum of the products of estimated coefficients and country averages of time varying variables. By this procedure, we obtain country average residuals that we regress by GLS on Z_i.

[29]Following Davis and Henderson (2003), we supplement our predetermined list of variables by including predictors of income and population growth, like gross capital formation and fertility rates, at $t-2$ and $t-3$.

Table 1: Urban primacy in sub-Saharan Africa

	% Urban Population in the largest city					
Country	1980	1990	2000	2001	2005	2009
West Africa						
Benin		12.0	31.5	24.69	22.87	21.93
Burkina Faso	41	50.6	38.3	48.55	52.77	56.44
Cote d'Ivoire	34	44.5	44.3	40.09	39.57	38.48
Gambia		100.0	100.0	50.43	46.32	44.70
Ghana	35	27.1	19.4	19.34	18.95	18.75
Guinea	80	76.1	46.6	46.82	46.37	45.43
Liberia		56.6	43.2	57.30	62.05	36.69
Mali	24	37.4	31.0	37.80	37.92	38.21
Mauritania	39	80.3	39.3	52.99	52.68	52.28
Niger	31	38.5	34.0	38.46	39.73	39.52
Nigeria	17	20.4	17.1	13.59	13.47	13.43
Senegal	65	55.9	44.2	50.74	51.86	51.95
Sierra Leone	47	51.7	49.5	45.23	41.75	40.36
Togo	60	56.1	52.6	53.38	54.79	56.35
Central Africa						
Cameroon	21	21.9	22.4	18.12	18.25	18.27
Central African Rep	36	50.8	45.6	41.05	41.21	40.95
Chad	39	43.6	42.6	32.02	28.89	26.58
Congo, Rep.	56	68.6	54.4	56.06	57.00	56.85
Congo, Dem. Rep.	28	33.1	32.2	37.11	37.47	36.80
Gabon	56		59.7	49.52	49.24	49.08
East and NE Africa						
Burundi		82.5	67.4	60.02	55.25	51.20
Ethiopia	37	29.2	25.5	23.81	21.91	19.98
Kenya	57	27.0	20.3	36.38	37.96	38.72
Rwanda	54		62.5	45.07	49.23	48.81
Somalia	13	12.4	36.5	47.52	48.12	40.08
Tanzania	50	25.4	18.8	27.93	28.39	28.24
Uganda	52	35.2	39.3	37.04	36.74	35.71
Southern Africa						
Angola	64	62.6	56.6	37.50	39.38	42.34
Bostwana		36.2	23.0	19.61	18.08	16.65
Madagascar	36	23.6	36.0	32.59	31.67	30.98
Malawi	19		34.2	27.52	28.02	27.86
Mozambique	83	40.7	19.1	19.29	18.65	18.45
Namibia		30.3	40.5	19.29	41.05	42.06
South Africa	13	12.4	11.1	11.03	11.73	11.95
Swaziland		19.2	28.9	26.06	26.30	24.70
Zambia	35	23.5	35.7	29.75	30.79	30.71
Zimbabwe	50	31.8	32.6	32.85	33.78	33.92

Sources: For 1980 and 1990 data – United Nations (1991), *United Nations Statistical Yearbook*, Washington, DC: United Nations; World Bank (1990), *World Development Report*, New York: Oxford University Press; Rondenelli (1983), *Secondary Cities in Developing Countries, Beverly Hills: Sage* as cited by Aryeetey-Attoh (1997). For 2000 data – Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat (2004). *World Urbanization Prospects: The 2003 Revision.* Data Set Name (POP/DB/WUP/Rev.2003/Data set number/File number), data set in digital form. For 2001, 2005, and 2009 data – World Bank online data available at `http://data.worldbank.org`. Retrieved on 2011-05-10.

Table 2: Dependent variable = primacy.

	(1) OLS	(2) FE	(3) MVREG	(4) 3SLS	(5) 3SLS
ln(urban pop)	-0.0988**	-0.1731**	-0.2076**	-0.2531**	-0.3834**
	(0.0395)	(0.0339)	(0.0606)	(0.0719)	(0.0803)
Square of ln(urban pop)	0.0014	0.0040**	0.0049**	0.0052**	0.0082**
	(0.0013)	(0.0012)	(0.0021)	(0.0025)	(0.0028)
ln(GDPpc)	-0.0003		0.1149**		0.0105
	(0.0660)		(0.0393)		(0.0283)
Square of ln(GDPpc)	0.0001		-0.0075**		-0.0011
	(0.0041)		(0.0025)		(0.0017)
Democracy (prev. 5 years)					-0.0003*
					(0.0002)
Protestant	-0.0018**	-0.0017**	-0.0018**	-0.0018**	
	(0.0003)	(0.0006)	(0.0006)	(0.0006)	
Muslim	-0.0006**	-0.0004	-0.0005	-0.0004	
	(0.0002)	(0.0005)	(0.0005)	(0.0006)	
Catholic	-0.0008**	-0.0007	-0.0007	-0.0003	
	(0.0002)	(0.0005)	(0.0005)	(0.0005)	
Port city indicator	0.0085	0.0077	0.0053	0.0188	
	(0.0134)	(0.0264)	(0.0297)	(0.0280)	
Capital city indicator	0.0716**	0.0811**	0.0730**	0.0582*	
	(0.0127)	(0.0325)	(0.0330)	(0.0347)	
ln(land area)	-0.0077	-0.0162*	-0.0080	0.0222**	
	(0.0069)	(0.0082)	(0.0088)	(0.0091)	
Waterways/km2	0.9992	0.7031	0.5719	0.9966	
	(0.6152)	(1.2998)	(1.3954)	(1.4187)	
Railways/km2	-0.3322	-0.4163	-0.0902	0.1035	
	(0.3365)	(0.5542)	(0.7248)	(0.6270)	
Roadways/km2	0.0544**	0.0476*	0.0540*	0.0800**	
	(0.0138)	(0.0278)	(0.0301)	(0.0321)	
% paved roads	-0.0028**	-0.0029**	-0.0028**	-0.0022**	
	(0.0003)	(0.0005)	(0.0005)	(0.0005)	
Ethnolinguistic fractionalization	0.0125	0.0145	0.0138	0.0058	
	(0.0263)	(0.0691)	(0.0749)	(0.0769)	
Corruption index	0.0010**	0.0013**	0.0016**	0.0003	
	(0.0004)	(0.0006)	(0.0007)	(0.0007)	
Corrup. Index*SSA dummy	-0.0026**	-0.0027**	-0.0030**	-0.0025**	
	(0.0004)	(0.0008)	(0.0009)	(0.0009)	
Landlock dummy	-0.0610**	-0.0486	-0.0608**	-0.0878**	
	(0.0181)	(0.0294)	(0.0293)	(0.0285)	
French colonisation	0.0654**	0.0733*	0.0775*	0.0673	
	(0.0151)	(0.0391)	(0.0415)	(0.0421)	
English colonisation	0.0320**	0.0231	0.0256	0.0161	
	(0.0140)	(0.0341)	(0.0346)	(0.0372)	
Time dummies	NO	NO	YES	YES	YES
N [countries]	403[70]	1197[205]	620[124]	585[117]	450[90]
R^2	0.7188	0.5367			

Notes: Standard errors in parenthesis are Huber-White. Coefficients with ** are significant at the 5% confidence level, those with * are significant at the 10% confidence level. Results in blue are those of a GLS regression of first step country average residuals on time invariant variables.

Table 3: Mean comparison tests between SSA and other groups of countries

SSA versus	Mean comparison test	p-value
	Democracy index (Polity2)	
Industrial countries	$\mu_{ssa} < \mu_{IND}$	0.0000
Latin America & Caribbean	$\mu_{ssa} < \mu_{LA}$	0.0000
East Asia & Pacific	$\mu_{ssa} < \mu_{AS}$	0.0001
South Asia	$\mu_{ssa} < \mu_{AS}$	0.0241
Middle East & North Africa	$\mu_{ssa} > \mu_{NA}$	0.0000

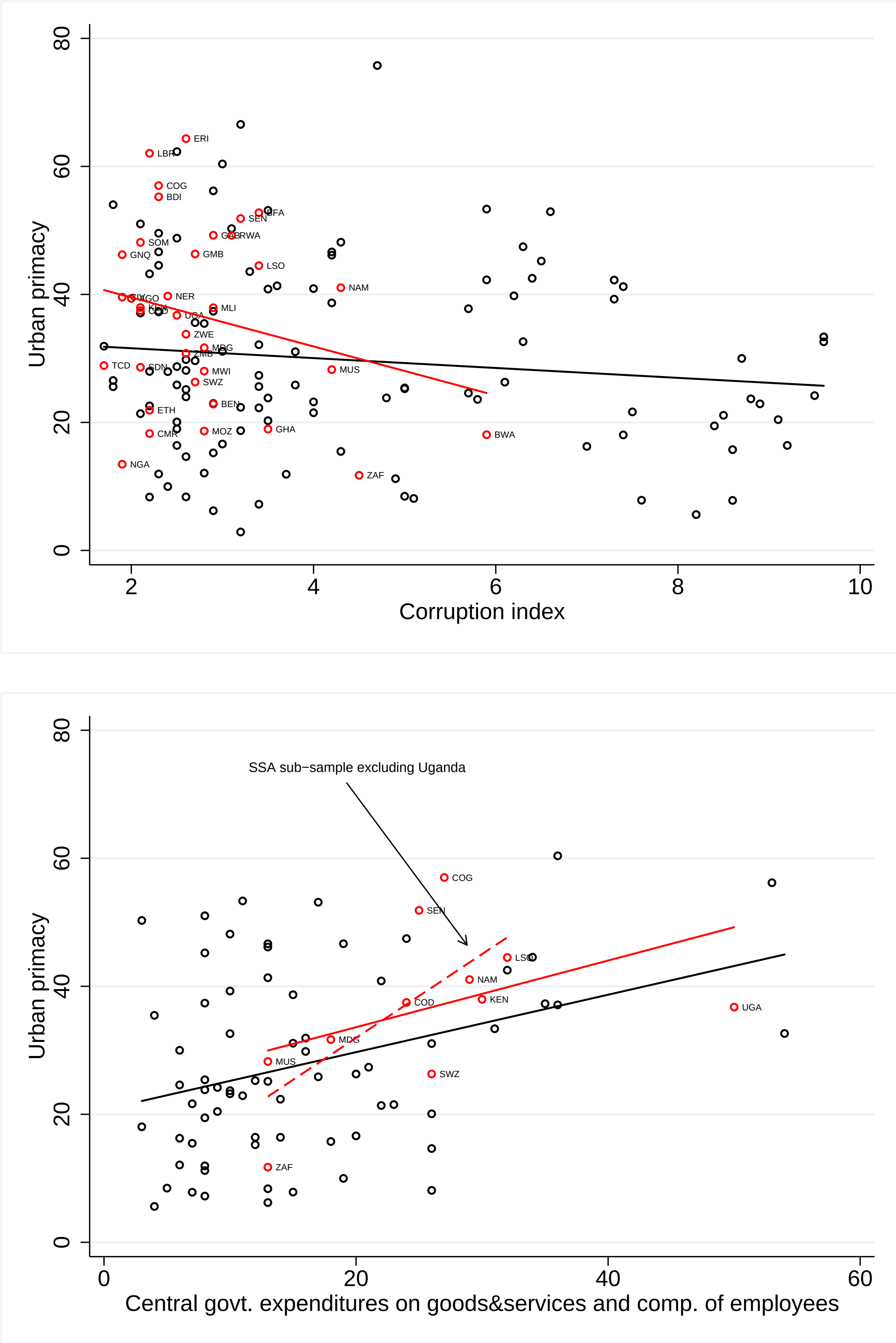
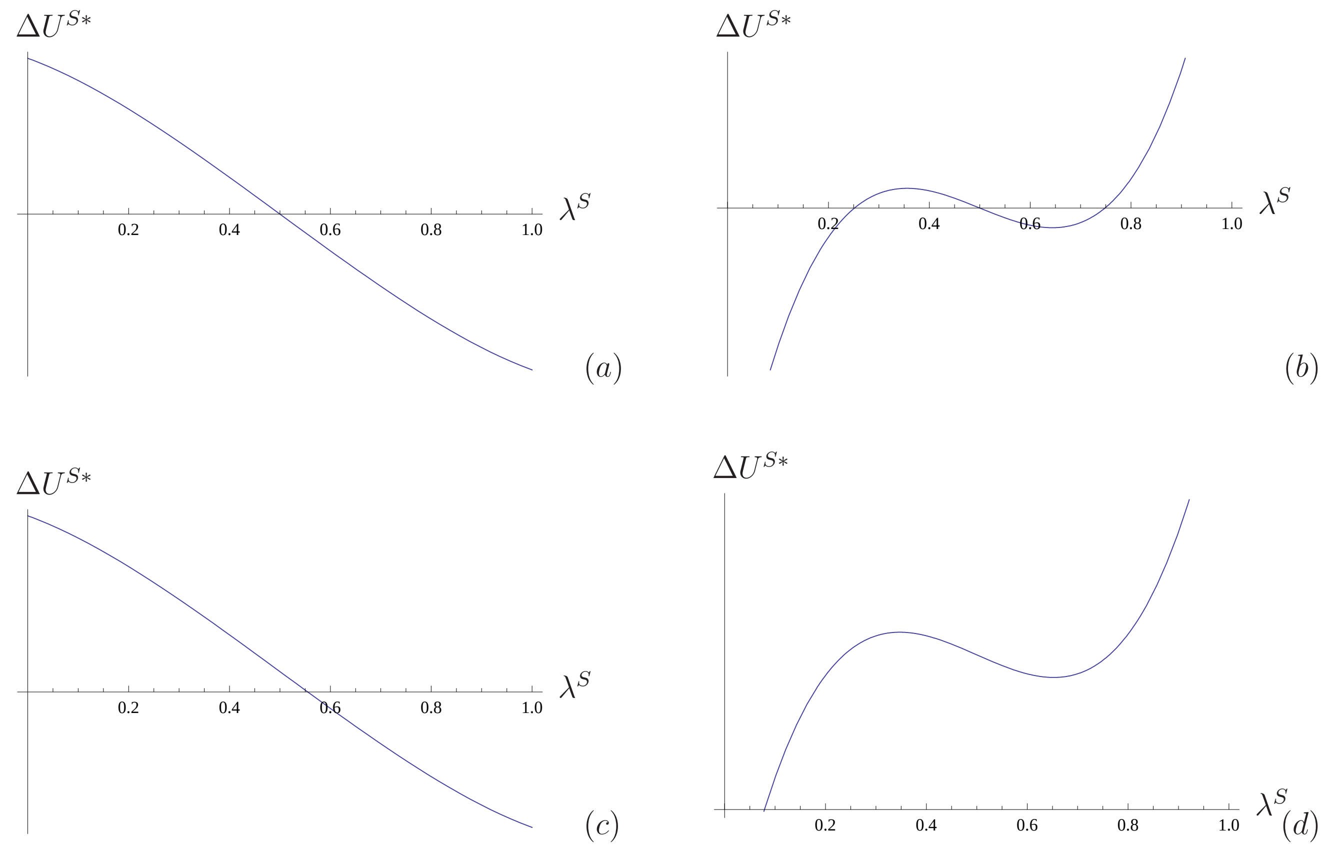
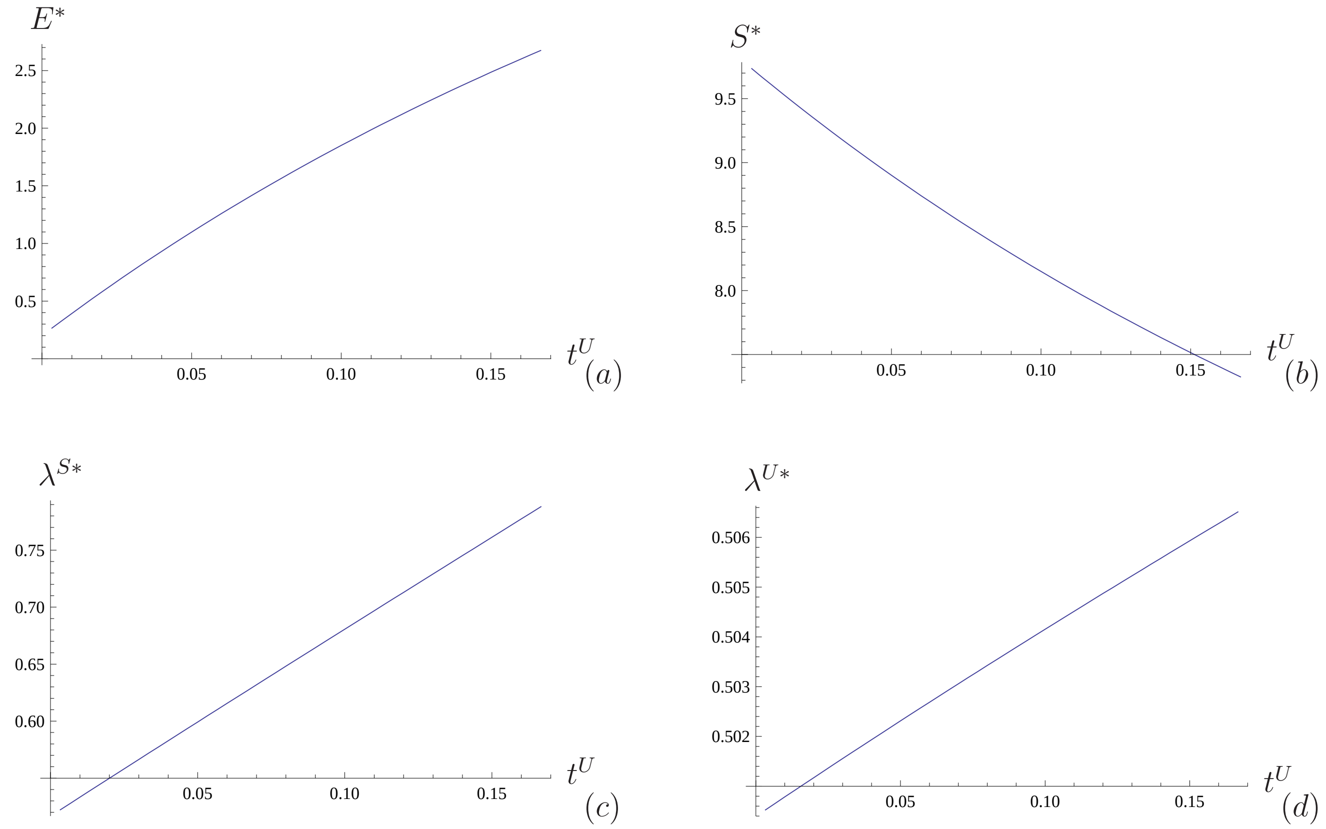


Figure 1: Corruption, central government expenditures, and urban primacy (SSA countries in red, the other countries in black; simple OLS regression lines)



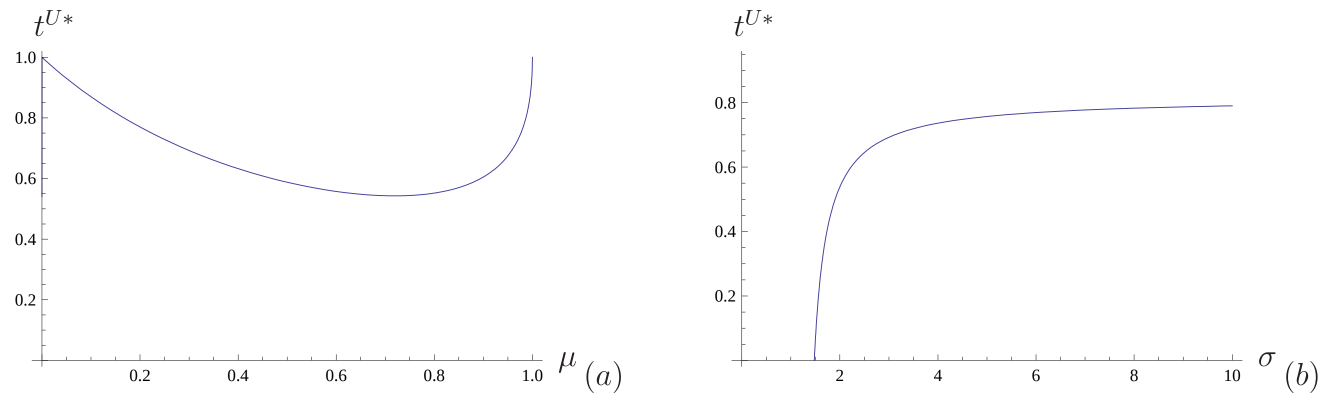
All panels are depicted with $\sigma = 3$, $\mu = 0.3$, $\phi = 0.582$, $\overline{S} = \overline{U} = 10$, $\beta = 4$, $F = 1$, $m = 0.2$. In the cases with positive taxes, we let $t^U = 0.01$ and $t^S = 0$.

Figure 2: ΔU^{S*} with (a) immobile unskilled labor and no taxes; (b) mobile unskilled labor and no taxes; (c) immobile unskilled labor and taxes; and (d) mobile unskilled labor and taxes



All panels are depicted with $\sigma = 3$, $\mu = 0.3$, $\phi = 0.51$, $\overline{S} = \overline{U} = 10$, $\beta = 4$, $F = 1$, $m = 0.2$. We also set the tax rate for skilled to $t^S = 0.02$.

Figure 3: Impacts of increasing unskilled taxation at the interior spatial equilibrium ($0 < \lambda^{S*}, \lambda^{U*} < 1$)



All panels are depicted with $\phi = 0.51$, $\overline{S} = \overline{U} = 10$, $\beta = 4$, $F = 1$, $m = 0.2$. In panel (a) $\sigma = 3$, and in panel (b) $\mu = 0.3$.

Figure 4: The equilibrium unskilled tax rate t^{U*} as a function of the parameters σ and μ