# Rural Windfalls and Urbanization: On Cocoa and Cities in Ivory Coast and Ghana

Rémi Jedwab*

Draft Version: September 15th 2010.

**Abstract**

While Africa was almost unurbanized at the turn of the 20th century, it has recently known spectacular urban growth. This is good news if cities are powerful engines of growth as emphasized by the economic geography literature. Yet, the agglomeration effects story was built on manufacturing and high value services, two sectors under-represented in African cities. We develop another story where rural windfalls feeds urban growth through consumption linkages, with a case study on cocoa production and cities in Ivory Coast and Ghana. We combine decadal data on cocoa production and cities at the district level from 1921 to 2000, and we show how cities have followed the cocoa front. Our identification strategy uses the fact that cocoa is produced by "eating" the virgin forest: (a) areas suitable to cocoa production are forested regions, basically the southern half of both countries, (b) for agronomic reasons, cocoa farmers move to a new forest every 25-50 years, this movement causing regional cycles, and (c) the cocoa front has started from the (South-)East of both countries. The cocoa front had to move westward, within the South. We can thus instrument cocoa production with a westward wave that we model. We find that cocoa production explains more than half of non-primate urbanization in both countries. We discuss and give evidence for the channels underlying this relationship, distinguishing what happens in new and old cocoa-producing regions.

JEL classification codes: N97, O12, O18, R11, R12.
Keywords: Urbanization, Cash Crops, Resource Curse, Africa.

---

*Paris School of Economics (e-mail: jedwab@pse.ens.fr, r.c.jedwab@lse.ac.uk). Web: http://www.pse.ens.fr/junior/jedwab/. I am extremely grateful to Denis Cogneau and Guy Michaels for their encouragement and support. I would like to thank Sylvie Lambert, Gareth Austin, Henry Overman, Thomas Piketty, Alexander Moradi, James Fenske, and seminar audiences at PSE and LSE for very helpful comments. I am also grateful to François Moriconi-Ebrard and Eric Denis from GEOPOLIS, Vincent Anchirinah, Francis Baah and Frederick Armon-Armah from CRIG, and Sandrine Mesple-Somps from DIAL-IRD for their help with data collection.

*"I had a marvellous dream [...]. Close to a castle, I have seen a man all dressed in white who told me: several years ago, this region was covered with forests. It was only missing hands to work. Compassionately, some men have come. [...] The forest has been gradually disappearing in front of labourers, tractors have replaced the daba [hoe] and beautiful cities, beautiful villages, beautiful roads have replaced the tracks only practicable during the dry season."*

Houphouët-Boigny's Presidential Address, 25 March 1974.

*"As the 1960s became the 1970s and then the 1980s, cocoa remained to Ivory Coast what oil was to Saudi Arabia or Nigeria: a geyser of cash. Brown gold. [...] After half a century of almost uninterrupted expansion, the Ivorian cocoa machine has begun to falter: cocoa yields are down and so is their quality. [...] It is early days, but the publicity-shy cocoa industry has started talking about a "chocolate crisis"."*

Financial Times, 28 May 2010.

# 1 Introduction

While Sub-Saharan Africa was unurbanized at the turn of the 20th century, it has registered dramatic urban growth in recent decades and it has now a larger urban population than Northern America or Western Europe (Satterthwaite 2007, WDR 2009).[1] This is potentially good news, as the growth literature has shown that development is highly correlated with urbanization (Acemoglu, Johnson and Robinson 2002, Henderson 2010). Development is indeed associated to the structural transformation, the economic transition from rural-based agriculture to city-based manufacturing and services (Caselli and Coleman II 2001, Michaels, Rauch and Redding 2008). Then, it can be argued that cities improve efficiency and promote growth in developing countries, making urbanization a potentially powerful agent of development (Duranton 2008, Venables 2010). Those works are based on previous studies showing there are strong agglomeration economies, both within sectors (localization economies) and across sectors (urbanization economies), in both developed countries (Rosenthal and Strange 2004, Henderson 2005, Combes, Duranton, Gobillon and Roux 2011) and developing countries (see Overman and

[1] While the urbanization rate of Sub-Saharan Africa is estimated at 5% around 1900 (Bairoch 1988), it increased from 11 to 37.2% between 1950 and 2010 and it is projected to be 60.1% by 2050 (WUP 2009). It has been estimated that post-1950 Africa has experienced amongst the highest rates of urban change ever registered in the history of mankind (Satterwhaite 2007).

Venables 2005, Henderson 2010 and Venables 2010 for references). As a result, a sound pro-urbanization public policy could consist in reducing the costs of urban congestion by providing public goods, implementing a titling policy for squatter settlements and deregulating the housing market, and removing barriers to internal migration and trade by ensuring peace, easing trade controls and investing in transport infrastructure (Duranton 2008, Venables 2010).

Yet, an optimistic view of urbanization in developing countries could be contradicted by empirical evidence on Africa. "Explosive urbanization", "overurbanization" or "urbanization without growth" are expressions frequently read in the literature on African cities (Bairoch 1988, Fay and Opal 2000). They imply that Africa has urbanized without it being fully explained by economic development, unlike developed countries. This excessive urbanization is often attributed to pull and push factors feeding rural exodus, using a simple Harris and Todaro (1970) model. First, cities are often associated to a parasital public sector, that feeds itself on the (over-)taxation of rural farmers (Bates 1981, Bairoch 1988). An extreme version of the urban bias story is primacy, when the largest city in a country is oversized compared to the rest of the urban population as it receives disproportionate public investments (Davis and Henderson 2003). Henderson (2003) and Duranton (2008) shows that primacy might be detrimental to growth. Second, increasing land scarcity and natural catastrophes can make rural living more and more difficult, this encouraging rural exodus (Barrios, Bertinelli and Strobl 2006). As poor rural migrants flock to the cities, they decrease urban welfare and governments of developing countries try to refrain them from doing so (Duranton 2008). They put barriers to formal residential development, this encouraging the formation of squalid slums (Satterwhaite 2007, Duranton 2008). Lastly, the agglomeration economies story is valid only for those sectors where localization or urbanization economies can take place, whether manufacturing or high value services. Manufacturing has been the main driver of development and urbanization in Europe and North America at the time of the industrial revolution (Bairoch 1988, Williamson 1990, Kim 2007). It is the main contributor to current growth in China (Bosworth and Collins 2008). But manufacturing is under-represented in African cities today, which asks two questions: (i) Where do African cities come from? What are the main sectors behind the African urbanization process? (ii) Given their economic composition, can we expect African cities to be powerful engines of growth? What is the future of those cities?

Our story of African urbanization differs from both views as it emphasizes the linkages between the rural-based cash crop sector and the urban sector. As some regions grow cash crops, the profits from this sector are spent and invested in localities that become cities or localities that are cities which will grow even further, this driving urban growth. Yet, as rural-urban linkages are mostly consumption linkages, most of the economic growth of cities takes place in the trade sector, with little impact on manufacturing. African cities are mostly "consumption cities", and one can thus have urbanization with growth, without it being a factor of long-term growth.[2] It builds on a case study of the role of cocoa production on urbanization in Ivory Coast and Ghana, two exemplary countries of the African "cash crop revolution" (Hill 1962, Tosh 1980, Austin 2007a). Cocoa has been the main motor of their economic development (Teal 2002, Cogneau and Mesplé-Somps 2002, Austin 2007b). Production boomed after the 1920s in Ghana and the 1960s in Ivory Coast (see figure 1). It has contributed to more than one third of their total exports and one tenth of their GDP during the 1948-2000 period. Yet, as cocoa can only be produced in forested areas (Ruf 1991, Ruf 1995, Balac 2002), its economic effects were confined to the Southern and forested part of both countries. Then, while Ghana and Ivory Coast were very little urbanized at the turn of the 20th century, their respective urbanization rate is 43.8% and 55.2% around 2000, making them two of the most urbanized countries in Africa.[3] As figure 2 shows, the total and urban populations of both countries have dramatically increased after World War II. Figure 3 displays the urbanization rate of both countries and their primacy rate, which we calculate as the size of the largest city (Accra for Ghana, Abidjan for Ivory Coast) over total population. First, Ghana started its urban transition earlier than Ivory Coast, but both experienced spectacular urbanization after 1948. Second, most of the post-1948 urbanization was not driven by the capital city. Understanding urbanization in Africa means understanding the growth of secondary cities. Nowadays, most of Ghanaian and Ivorian cities can be found in the former forested regions, those regions highly suitable to cocoa production (see figure 4). Assuming we are able to go beyond this strong but naive spatial correlation and show that cocoa production "causes" urbanization, this would strongly support our story.

We combine decadal data on cocoa production and cities at the district level

[2]Our study echoes the literature on rural-urban linkages, which relates urban economic activity to agricultural productivity shocks (Irz and Roe 2005, Tiffin and Irz 2006, Foster and Rosenzweig 2008, Henderson, Storeygard and Weil 2009).

[3]Ghana had nine cities of more than 5000 inhabitants in 1901 and its two largest cities were Cape Coast (28,948 inhabitants) and Accra (14,842). Ivory Coast did not have any such city, and Abidjan was then a small fishing village with less than 1000 inhabitants. The population of Accra and Abidjan were respectively estimated at 2,527,014 in 2000 and 2,955,578 in 1998.

from 1921 to 2000, and we show how the urban front has followed the cocoa front. Our identification strategy uses the fact that cocoa is produced by "eating" the virgin forest: (a) those areas suitable to cocoa production are forested regions, basically the southern half of both countries, (b) for agronomic reasons, cocoa farmers move to a new forest every 25-50 years, this movement causing regional cycles, and (c) the cocoa front has historically started from the (South-)East of both countries. This forced the cocoa front to move westward, within the South. We can thus instrument cocoa production with a westward wave that we model. First, our results suggest that local cocoa production has a strong impact on local urbanization, whether one considers new cities or already existing cities growing further. In total, we find that cocoa production explains more than half of non-primate urbanization in both countries. Second, we distinguishes what happens in new and old cocoa-producing regions.[4] We observe strong urban growth in both, which indicates that cities keep growing even when cocoa production decreases. Third, we discuss and give evidence for the various channels through which the cocoa sector impacts the urban sector, whether in the short or long run. Those long run effects are important because they explain why cities keep growing even in old cocoa-producing areas.[5] Yet, if their population increases, it is not clear yet what happens to per capita income. If labor productivity sufficiently increases to sustain a larger population, per capita income does not decrease and cash crops have a long-term development effect (optimistic scenario). If labor productivity does not increase enough, per capita income decreases and the development effect of cash crops is a short-term one (pessimistic scenario). Cities then pauperize.[6] We can use our framework to see what might happen in a few decades when both countries have entirely "eaten" their virgin forest, thus not being able to produce cocoa anymore. Beyond the microeconomic effects of resource exhaustion, this will have a negative impact on government revenue and spending, thus affecting urban

[4]Our study can then be compared to more general studies of the impact of local economic shocks on local labor markets, whether one considers booms (Carrington 1996) or busts (Hooker and Knetter 2001), or both (Black, McKinnish and Sanders 2005).

[5]Our paper is also related to a large body of work on the role of geographical endowments and agglomeration economies in long-term development. In new cocoa-producing regions, agglomeration economies have not arisen yet, and we can study the role of geographical endowments on development (Gallup, Mellinger and Sachs 1998, Engerman and Sokoloff 2000, Davis and Weinstein 2002, Nunn and Qian 2010). In old cocoa-producing regions, geographical endowments are "lost", but agglomeration economies are realized (Rosenthal and Strange 2004, Henderson 2005, Combes, Duranton, Gobillon and Roux 2011).

[6]This is in line with of Chen, Ravaillon and Sangraula (2007) who show that absolute poverty has urbanized in the developing world, as more and more poor people live in city.

poverty.

Finally, our research is also related to the study of cash crop windfalls, which have been highly relevant to the economic history of developing countries. Using FAO data, we could calculate that agricultural exports contributed to 53.4% of total exports for least developed countries in the 1960s, while it decreased to 21.4% in the 1990s as more and more countries specialized in the export of mineral products following discoveries of significant deposits. Then, many countries are still highly dependent upon one agricultural commodity. Amongst 125 developing countries in 2000, agricultural exports represent more than 50% of total exports for 20 countries, and more than 20% for 50 of them. Besides, it should be noted that most of those dependent countries are in Sub-Saharan Africa, with well-known stories such as cocoa in Ghana and Ivory Coast, tea and coffee in Kenya, Rwanda and Uganda, coffee in Burundi and Ethiopia, tobacco in Malawi, groundnut oil in Senegal, or cotton in Benin, Burkina-Faso or Mali. Like for mining products, the production of cash crops occurs through regional booms and busts, influenced by the cycle of international prices and the availability of suitable land. But while mining windfalls have been extensively studied by the resource curse literature (Sachs and Warner 1999, Sala-i-Martin and Subramanian 2003, Caselli and Michaels 2009, Vicente 2010, Michaels 2010), there are few studies on the economic effects on cash crop windfalls (Bevan, Collier and Gunning 1987, Maxwell and Fernando 1989, Angrist and Kugler 2008, Collier and Goderis 2009). Most mines are capital-intensive and employ few workers. Profits are usually transferred from the producing region to the capital city and abroad. On the contrary, the production of cash crops necessitates little technology, just suitable land and labor. Returns to scale being constant, the sector is dominated by a myriad of smallholders, whose bargaining power can be quite high. Although the cash crop sector can be taxed by the state (Bates 1981), we still expect a large share of sectoral profits to go to those producing regions and households, thus having large development effects. Considering urbanization as a valid development outcome (Acemoglu, Johnson and Robinson 2002), our study informs on the local benefits of cash crop production. But do those effects hold in the long run? When a developing country has accumulated massive wealth thanks to its cash crop sector, does it use it to fund its primitive accumulation of capital (physical, human and social) and move to a new steady state with a higher labor productivity? African countries have highly benefited from their primary exports till the early 1980s, but the subsequent period has been characterized by macroeconomic disequilibria, social and political unrest and sometimes general impoverishment. This implies that there are few long-term development effects of cash crops, which are then subject to a resource curse in the form of failed "intertemporal redistribution". A recent literature has emphasized that poverty reduction in Africa in the two latest decades could have been underestimated on the sole basis of national accounts (Miguel 2009, Young 2010, Pinkovskiy and Sala-i-Martin 2010). But we can wonder to what extent this result is specific to temporarily high terms of trade for African countries and if it does reflect a change in the structural capacity of those economies to produce growth and reduce poverty.

The remainder of this paper is organized as follows. Section 2 details a theoretical discussion of rural-urban linkages when the rural-based cash crop sector booms or busts. Section 3 presents the agronomic and historical background of cocoa production in Ghana and Ivory Coast, while section 4 introduces the data. Section 5 gives a graphic analysis of cocoa and urbanization, and discuss endogeneity issues. Section 6 explains our econometric framework and displays our main results. Section 7 addresses complementary issues. Section 8 discusses the potential future of African cities, while section 9 concludes.

# 2 Theoretical Discussion

The country can be divided into districts with district-specific locational fundamentals. Those districts suitable for cocoa production (the forested areas) experience urban growth when cocoa is produced (if the cocoa front has reached that district). Why would cocoa production lead to more urbanization? We need to develop a sequential model of urban settlement in a new forest.

In phase 1, which we label *no cocoa production yet*, a district with a virgin forest is relatively untouched and settlement is limited because it is difficult: land is not cleared yet, humidity and mortality are high.

In phase 2, which we label *new cocoa-producing area*, cocoa farmers settle there, the land is deforested and planted with cocoa trees, then cocoa production booms and the urbanization process is launched. The total population of the district increases, but we need this increase to be spatially concentrated for cities to appear and to grow. Indeed, only a few sublocations of the district will be driving the local urbanization process. Why would it be so? The first urbanizing effect of cocoa production is a pure settlement effect (effect A): when cocoa farmers move to new areas, they first settle in the few existing settlements or fund new settlements, to use them for the colonization of surrounding forested areas. The second urbanizing effect of cocoa production is a pure logistics effect (effect B):

cocoa beans need to be transported from cocoa-producing areas to the coast, so as to be exported abroad. Cities serve as collection points and transportation nodes for cocoa beans. The third urbanizing effect of cocoa production is a pure wealth effect (effect C): the income of cocoa farmers rise and they spend their extra income on non-essential consumption goods, in accordance with the Engel curve. If those goods are *produced in* or *distributed through* the cities, then more cocoa income means more local opportunities in the cities. As cocoa beans are not processed in producing countries, most of those inter-sectoral linkages are indeed consumption linkages. Migrants flow from non-producing regions to cities in producing areas. In other words, cocoa creates a large economic surplus, but this surplus being concentrated in the cities, those regions urbanize. The fourth urbanizing effect of cocoa production is an infrastructure effect (effect D): cocoa income allows those districts to pay the fixed costs of the primitive accumulation of physical capital (basic amenities such as roads, schools, health centres), and this has positive long-term effects on the size of those cities. Fifth, improved living standards and better infrastructure in those cities means reduced mortality and natural growth can quickly surpass rural-to-urban migration as the first source of urban growth (effect E). Effects D and E are long-run effects because cocoa does not directly impact urbanization as for effects A, B and C.

In phase 3, which we label *old cocoa-producing area*, cocoa leaves the region but cities do not collapse, on the contrary. The first two urbanizing effects of cocoa have disappeared, and we are left with the three other channels. First, inter-sectorial linkages and agglomeration economies make those cities thrive as they are able to "reinvent" themselves (effect C). Second, the many investments in basic amenities make those cities still very attractive (effect D). Lastly, the demographic transition is first an urban demographic transition, and the contribution of natural increase to urban growth increases (effect E). As regards living standards, per capita income is likely to fall given demographic growth and diminishing total income from cocoa production (pessimistic scenario). But we could expect another scenario whereby capital accumulation and agglomeration economies would raise labor productivity enough to increase or at least stabilize per capita income (optimistic scenario).

In the very long run, the total stock of virgin forest is exhausted and cocoa can no longer be produced. All the districts that were suitable to cocoa production are in phase 3, where urban per capita income might be decreasing, stable or increasing, following which scenario is realized. Beyond those microeconomic effects, this could have a huge macroeconomic effect on government revenue and spending, if the cocoa sector is taxed by the state. We could then include to our model the redistributive effects of state taxation.

# 3 Agronomic and Historical Background

## 3.1 Agronomic Background

Cocoa is produced by "eating" the forest. Cocoa farmers go to a patch of virgin forest and replace forest trees with cocoa trees. Pod production starts after 5 years, peaks after 10 and continues up to 40 or even 50 years. When cocoa trees become too old, cocoa farmers have no choice but to move to another forest and start a new cycle. Indeed, removing forest trees alters the original environmental conditions and replanted cocoa trees are much less productive (Ruf 1991, Ruf 1995, Balac 2002).[7] That is why cocoa is characterized as a "migrant culture". Cocoa-producing countries have all experienced deforestation through regional cycles. When the forest rent is over, cocoa production moves to another country and even a new continent.[8] Thus, cocoa production has been and still is a significant contributor to deforestation in developing countries. The forested surface of Ivory Coast has decreased from 9 millions hectares in 1965 to 2.5 millions in 2000, while it has decreased from 8.2 millions in 1900 to 1.6 million in 2001 in Ghana. Recent studies have more generally emphasized the role of agricultural trade in worldwide deforestation (De Fries, Rudel, Uriarte and Hansen 2010).

Then, more economic and political factors can accelerate or decelerate those cycles. Both countries have more or less extracted the same quantity of cocoa throughout the 20th century (see figure 1): 24 millions of tons in Ghana vs. 22.1 millions in Ivory Coast. But, this amount has been extracted within a much shorter time period in Ivory Coast. In Ghana, three regional cycles did not overlap because

[7] By removing forest trees to plant cocoa trees, farmers change the environmental conditions that are nevertheless essential to the long-term profitability of their cocoa farms. Cocoa trees are affected by: (i) spreading heliophile weeds, (ii) diminishing pluviometry, (iii) a lower protection against winds, (iv) repeated attacks by new insects and diseases, (v) decreased soil fertility (the fertility of rainforests is contained in the trees and not the ground), and (vi) erosion. When cocoa trees are dying, cocoa farmers can plant new cocoa trees but the mortality rate of young cocoa trees is high while yields of those surviving trees is low. Discussions with agronomists have confirmed than replanting is twice more expensive than planting in a new forest. One can find those two interesting quotes in Ruf 1991: "Before, cocoa plantations were productive; it's difficult now, young cocoa trees die..." (p.105), and "An old plantation is like an old dying wife. Medicine would be too expensive to keep her alive. It's better to keep the money for a younger woman [a new plantation]" (p.107).

[8] Cocoa production was dominated by Caribbean and South American countries till the early 20th century, then moved to Africa and is now spreading in Asia.

the first two cycles were decelerated by extraordinary events and poor economic policy, which we are going to describe just thereafter. In Ivory Coast, those regional cycles have been perfectly imbricated as no decelerating factor showed up during the period.

## 3.2 Historical Background

Cocoa was first introduced to Ghana by missionaries in 1859, and then reintroduced in 1878 by a Ghanaian blacksmith coming back from Equatorial Guinea. But it took 30 years before seeing cocoa being widely grown in Ghana, making it the world's largest exporter as soon as 1911. Cocoa production spread out in the Eastern province from Aburi Botanical Gardens, where the British sold cocoa seedlings at a very low price. Figure 5 shows the various provinces of Ghana, the area suitable to cocoa production (basically, those regions with virgin forest one century ago) as well as Accra and Aburi (the historical starting point). Production peaked in the Eastern province in 1931, before plummeting as a result of both the Cocoa Swollen Shoot Disease and World War II which reduced international demand.[9] A second cycle started after the war in the Ashanti province. But low producer prices after 1958[10], restrictive migratory policies after 1969[11] and frequent droughts[12] precipitated the end of this cycle. Higher producer prices after 1983 then pushed cocoa farmers to launch a third cycle in the Western province, the last forested region of Ghana.

Cocoa was first planted in Ivory Coast in 1888 by two French farmers not far from Abidjan. But it was not till 1910-1912 that the French governor decided to seriously promote cocoa production, thus trying to replicate the Ghanaian success story. Ivorians were originally reluctant to grow cocoa instead of food crops, except in Indénié (Abengourou) where local farmers heard of the increasing wealth of Ghanaian cocoa farmers (see figure 5, which exhibits Ivorian provinces, the area

[9] The Cocoa Swollen Shoot Disease was first recorded in 1938 in the Eastern region. Because no attempt could be made to control the disease until after World War II, millions of trees were killed and more millions had to be removed to try to control it (Thresh and Owusu 1986).

[10] Ghana after 1948 and Ivory Coast after 1960 decided to fix the producer price, with the aim of protecting farmers against fluctuant international prices. The Ghana Cocoa Marketing Board (COCOBOD) was in charge with cocoa in Ghana while the *Caisse de stabilisation et de soutien des prix des productions agricoles* (CSSPPA, or "Caistab") was its Ivorian equivalent. However, since the producer price was fixed much below the international price, this policy served as a taxation mechanism of the cocoa sector (Bates 1981).

[11] In 1969, the government enacted the Aliens Compliance Order, which led to the massive exodus of laborers from neighboring countries and created labor shortages in the cocoa sector.

[12] The 1982-1983 drought was the worse in fifty years and many cocoa farms were burnt.

suitable to cocoa production and Abidjan and Abengourou). However, Ivorian production did not boom before the 1960s.[13] Cocoa moved from the Eastern forest to the Western Forest (see figure 5). Due to mounting deficits of the Caistab, the producer price was halved in 1989 and remained low thereafter, but this did not stop the colonization process as profits were still quite substantial.

Thus, in both countries, cocoa production was confined to the Southern (forested) areas and historically started in the South-East of the country, for rather exogenous reasons. Then, cocoa being a "migrant culture", it moved to the West and within the South of both countries (as it could not move anywhere else). It is like a *pacman* game, except that the number of players have been increasing with time. As population growth was high and cocoa remained more profitable than other crops, more and more individuals specialized in cocoa production and participated to the colonization of the forest, thus accelerating the westward movement. Yet, in Ghana, the colonization of the forest has not been as linear as in Ivory Coast, due to natural events and more economic and political factors. As the forest rent is going to disappear soon, so will cocoa production, unless technological innovations increase production yields in already deforested land.

# 4 Data

To study the effects of cocoa production on cities in Ghana and Ivory Coast, we combine data on cocoa production and urbanization at the district level over the period 1921-2000. We briefly describe the data here but the full methodology and the numerous sources used can be found in the data appendix. Cocoa production data mainly comes from reports published by the government agencies responsible for the organization of the cocoa production system in each country: the *Caisse de stabilisation et de soutien des prix des productions agricoles* (CSSPPA, or "Caistab") in Ivory Coast, and the Ghana Cocoa Marketing Board (COCOBOD) in Ghana. Our cocoa production data is available at the level of administrative districts in Ivory Coast, and at the level of *cocoa districts* in Ghana. Then, from census reports and administrative counts, we obtain the size of each locality of more than 5000 inhabitants for various years. We then geocode this

[13]Several factors could account for this Ivorian "lateness". First, cocoa production did not reach the Ghanaian border before the 1910s. Ivorian cocoa production was then increasing in the 1930s, but this first possible boom was short-lived, due to the Great Depression and World War II. Second, Ivorians had to provide the *corvée* (mandatory labour) for the colonial government, which forced them to grow more food crops and coffee.

data and we use GIS to extract urban population for any spatial decomposition we want. In Ivory Coast, we recreate urban and rural population data using the administrative districts. In Ghana, since cocoa districts significantly differ from administrative districts, we are only able to recreate urban population data and not rural population data. In the end, in our regression framework, we use a panel of 46 districts × 6 years (1948, 1955, 1965, 1975, 1988, 1998) = 276 observations in Ivory Coast, and 73 districts × 7 years (1921, 1931, 1948, 1960, 1970, 1984, 2000) = 511 observations in Ghana. Between each district-year observation, we know how many tons of cocoa beans have been extracted and how many more urban inhabitants live there. We can therefore relate more urban inhabitants and more cocoa production. Since we have a deflated series of national cocoa producer prices, we work on the effect of the value of cocoa production (in 2000$) on urbanization. Figure 6 shows the value of cocoa production going to cocoa farmers during the 1921-2000 period. We start with a more graphic analysis of the correlation between cocoa production and urbanization across time. We then study this relationship in an econometric framework. Lastly, we use various Ivorian and Ghanaian household surveys and census data sets to discuss and give evidence for the various channels underlying this relationship: the 1985-88 *Living Standards Measurement Study* (LSMS), and the 1998 and 2002 *Enquêtes sur le Niveau de Vie des Ménages* (ENV) for Ivory Coast, and the 1987-88 and 2005 *Ghana Living Standard Surveys* (GLSS) and the *2000 Population and Housing Census* IPUMS sample for Ghana.

# 5 Mapping and Econometric Framework

## 5.1 Cocoa and Cities: Mapping

Figures 7 to 11 show district density of cocoa production and cities every ten years or so (if two dates are reported, the first one is for Ghana while the second for Ivory Coast): 1948 (fig. 7), 1960-1965 (fig. 8), 1970-1975 (fig. 9), 1984-1988 (fig. 10) and 2000-1998 (fig. 11). We have created similar maps for 1921 and 1931 but they were not reproduced. In 1948 (fig. 7), Ghanaian cocoa production has already boomed in the Eastern province and is about to boom in the Ashanti province. Cocoa production is also spreading to Ivory Coast. Most of the Ghanaian cities at that time are coastal cities, administrative centers or localities in the cocoa-producing areas. The Ivorian urban structure is mostly the result of the colonial administrative system. We then see cocoa production moving westward in both countries. In Ghana, the main cocoa-producing area was Ashanti province in the 1960s and 1970s (fig. 8 and 9), and Western province in the 1990s (fig. 11). In Ivory Coast, production rapidly moved from the Eastern forest (fig. 7 and 8) to the Western forest (fig. 9, 10 and 11). This analysis shows that the correlation between cocoa production and urbanization is spatio-temporal. Pre-existing cities grow and new cities arise in both new and old cocoa-producing regions. This suggests why most of the cities of both countries can be found in the areas suitable to cocoa production.

## 5.2 Cocoa and Cities: Econometric Framework

The main hypothesis we wish to test is whether cocoa production drives urbanization. We focus on 1921-2000 Ghana and 1948-2000 Ivory Coast. We run panel data regressions for districts $d$ and years $t$ of the following form:

$$U_{d,t} = \alpha_d + \beta_t + \delta C_{d,t} + \gamma U_{d,t-1} + \phi_t X_d + u_{d,t} \tag{1}$$

where $\alpha_d$ and $\beta_t$ are district and year fixed effects, and our dependent variable is urban population (in inhabitants) of district $d$ at time $t$ ($U_{d,t}$), controlling for urban population at time $t-1$ ($U_{d,t-1}$). Since urban dependency varies across time, given agglomeration economies for instance, we might allow the effect of $U_{d,t-1}$ to be period-specific ($\gamma_t$). $C_{d,t}$ is our variable of interest and is equal to the value of cocoa beans (in millions of 2000$) produced between time $t-1$ and time $t$. $X_d$ is a vector of baseline demographic, economic and geographic controlling variables whose coefficients are also time-varying. Otherwise, they are included in the district fixed effect. In particular, we might wish to capture the time-specific effect of being suitable to cocoa production, which we define through a dummy equal to one if more than 50% of district area is suitable to cocoa production. As most of suitable districts are in the South of both countries, we control for the fact that trade and political economy factors could have a differential impact on the South and North of each country. $u_{d,t}$ are individual disturbances that are clustered at the district level.

Due to primacy, we expect the two main cities of both countries (Accra and Kumasi in Ghana and Abidjan and Bouaké in Ivory Coast) to receive a disproportionate share of public investments. Their city size is then explained by political economy factors (Davis and Henderson 1993). Since they bias downward our estimates, as those observations display a dramatic increase in urbanization without it being explained by cocoa, we drop the four districts that contain each city. We now have 71 districts and 7 time periods in Ghana, hence 497 observations. We have 44 districts and 6 time periods in Ivory Coast, hence 264 observations. Since we include the lag of urban population, we drop one round and obtain respectively 426 and 220 observations.

We assume that the OLS effect of cocoa production on urbanization is causal. Yet, urbanization could drive cocoa production. However, rainforests are dense forests where human settlement is difficult. That is why there are few cities in those forests before cocoa production booms. But more farmers are willing to overcome those settlement constraints when they obtain a high income, which is the case with cocoa and not with other less profitable crops. When cocoa production is sufficiently well-entrenched in the region and when economic mass rises, more migrants arrive to fill pre-existing and new cities. Besides, cities consume forested land and constrains potential land for cocoa production. Additionally, cities are not very useful to cocoa production, since cocoa production involves little technology. Lastly, being close to a city could mean being close to a local depot of the collection agency, but depots are usually created in new cocoa-producing regions after cocoa has boomed and not before.

Second, omitted factors could drive both cocoa production and urbanization, even when including district fixed effects. Obvious culprits are transportation networks[14], initial population (as it could provide cheap labor to both the urban and cocoa sectors) or rainfall (as it could support the production of food crops, thus lowering food prices for both the urban and cocoa sectors). Transportation networks were either pre-determined and/or can be controlled for, or resulting from cocoa production itself.[15] Regarding population, Ghana, Ivory Coast and their neighbors form a large labor market, and many of the workers of the cocoa and urban sectors were not directly originating from the producing regions. Regarding food production, the soil and climatic conditions that are suitable to cocoa production are also suitable to the growing of plantain, cassava or yam, which are the main crops consumed by people living in those regions. Lastly, random measurement errors on cocoa production could downward bias our coefficient. This cannot be solved unless we instrument cocoa production.

[14] Atack et al. 2009 look at the impact of railroad building on urbanization in the 19th century U.S. Other studies relating roads and development (but not urbanization) include Michaels 2008 in the U.S. and Banerjee, Duflo and Qian 2009 on China.

[15] Discussions with agronomists have confirmed that cocoa farmers first go to the rainforest to plant cocoa trees and when pod production starts, they lobby the state board to deliver proper evacuation routes in the form of roads.

Our instrumentation strategy relies on the fact that cocoa production is confined to suitable (forested) areas and is moving westward in both countries, as a result of historical factors. We first create a dummy equal to one if more than 50% of district area is suitable to cocoa production. One can also try 25 and 75% cut-offs. Figure 12 displays those districts that are suitable to cocoa production using those various cut-offs. We then arbitrarily divide the territory into longitude bands of one degree, using the centroid of each district. Figure 13 reproduces those various longitude bands. We assume that the cocoa front is moving one degree westward every $X$ time period. We take $X = 2$ for Ghana and $X = 1$ for Ivory Coast. Indeed, regional cycles were not imbricated in Ghana, contrary to Ivory Coast, due to natural events and more economic and political factors, as already explained in sections 3.2. The instrument is the dummy "being suitable to cocoa production" interacted with a dummy "being *on* the cocoa front". Then, cocoa production is high both at the cocoa frontier and in the two adjacent longitude bands. The right longitude band was the cocoa frontier at time $t-1$ and production does not immediately collapse. The left longitude band is the next cocoa frontier (at time $t+1$) and production is already increasing there. Thus, comparing the sole cocoa frontier to the other longitude bands might give a less powerful instrument than considering the cocoa frontier plus the two adjacent bands. The new instrument is then the dummy "being suitable to cocoa production" interacted with a dummy "being *close to* the cocoa front". We will test both. Here is the full IV model:

$$U_{d,t} = \alpha'_d + \beta'_t + \delta' C_{d,t} + \lambda F_{d,t} + \gamma' U_{d,t-1} + \phi'_t X_d + v_{d,t} \quad (2)$$

$$C_{d,t} = \alpha''_d + \beta''_t + \Pi S_d * F_{d,t} + \lambda' F_{d,t} + \gamma' U_{d,t-1} + \phi''_t X_d + w_{d,t} \quad (3)$$

with $S_d$ being a dummy equal to one if the district is suitable to cocoa production and $F_{d,t}$ equal to one if the district happens to be "*on* the cocoa front" or "*close to* the cocoa front", depending on which IV we use. The other variables are defined as above (see equation (1)). The coefficients of interest are $\delta$ from equation (1) and $\delta'$ from equation (2), the OLS and IV estimated impacts of the value of cocoa production on urbanization.

To conclude, we just instrument cocoa production by just saying "it has to move westward, within the South, as it cannot go anywhere else". Then, whether it is going North-Westward or South-Westward, as a result of transportation networks for instance, is irrelevant since the instrumentation permits us to get rid of those potentially contaminating factors.

# 6 The Effect of Rural Windfalls on Urbanization

## 6.1 Local Cocoa Production and Local Urbanization

Tables 1 and 2 respectively present our first set of results, for Ivory Coast 1948-1998 and Ghana 1921-2000. For each country, we first show in panel A the OLS estimate (columns (1)), then the IV estimates without controls (column (2)), with a time-varying effect of lagged urbanization (column (3)), further adding a time-varying effect of being suitable to cocoa production (column (4)) and also including controls (column(5)). We report in panel B the first stage of our IV estimations. We privilege the IV estimates using the instrument ”suitable to cocoa production” * ”being *close to* the cocoa front” as the instrument is then powerful enough for both countries (see the Kleibergen-Paap rk Wald F stat which we report). The instrument ”suitable to cocoa production” * ”being *on* the cocoa front” is more powerful for Ivory Coast but much weaker for Ghana. Hopefully for us, the wave has no independent positive impact on cocoa production or urbanization (see the coefficients of ”Close to the cocoa front”). Only its interaction with cocoa suitability has a positive impact on the value of cocoa production (see the coefficients of ”Suitable to cocoa * Close to the cocoa front” in panel B). Then, our set of controls include: (a) dummies equal to 1 if the district is coastal, has a railway or a paved road going directly to the capital city, all interacted with a year trend, (b) year dummies interacted with the district distance to the coast, and the district average annual sum of rainfall for the period 1900-2000.

The OLS and IV estimates are not significantly different for Ivory Coast (see table 1). In Ghana (see table 2), the IV estimate is seven times higher than the OLS estimate. Since simultaneity and omissions biases are unlikely to vary across both countries, our guess is that our Ghanaian cocoa production data is subject to random measurement errors, and we know that such errors can generate a high bias of the coefficient of interest in panel data models. While our Ivorian cocoa production data mostly corresponds to real data obtained from the *Institute of Statistics of Côte d'Ivoire*, our Ghanaian cocoa production data was reconstructed using many different sources and assumptions had to be made to create a consistent data set across periods. If we believe our estimates from column (5) in tables 1 and 2, which is our full IV model, one million 2000$ of cocoa production increases urbanization by 107.4 inhabitants in Ivory Coast and 86.6 inhabitants in Ghana. The two effects are not significantly different, which is very comforting for our strategy. We then calculate the magnitude of each effect, that is to say how much of national urban growth between our first and last periods can be explained by this sole effect.[16] We find that this effect explains 58.2% of urban growth in Ivory Coast between 1948 and 1998 (excluding Abidjan and Bouaké) and 50.8% in Ghana between 1921 and 2000 (excluding Accra and Kumasi). If we do not drop those observations corresponding to those cities, the IV estimates are slightly reduced and their magnitudes decrease as our model cannot explain their growth. If we consider the sole districts suitable to cocoa production (basically, the southern districts), their magnitudes significantly increase, reaching around 80%.

## 6.2 Specification and Robustness Checks

In table 3, we show that those IV results (with controls) are robust to specification checks, with Ivory Coast in columns (1) to (4) and Ghana in columns (5) to (8). Columns (1) and (5) report those results for column (5) from tables 1 and 2. However, as our panel data model includes a lag of the dependent variable, our estimates are subject to a *dynamic panel bias* (Nickell 1981). We therefore estimate in columns (2) and (6) a model where we consider the change in urban population as the outcome without including any lag of the dependent variable. Coefficients are almost unaffected. In columns (3) and (7), we explain urban density (district urban population / district area in squared km.) by value density (district value of cocoa production / district area in squared km.) but this does not alter our message. Lastly, in columns (4) and (8), our variable of interest is cocoa production in volume (tons). The coefficient for Ghana is lower than the coefficient for Ivory Coast, but urbanization has also been lower in Ghana than in Ivory Coast.

We then check that the periodicity of our data does not impact our results. Instead of considering the change in urban population and the value of cocoa production between times $t-1$ and $t$, we run the same regression dividing them by the number of years between $t-1$ and $t$. Results are unchanged. We also verify that the selection of the suitability cut-off (25, 50 or 75%) for our instrumentation strategy does not affect our results. Lastly, we test that our results are not driven by a specific period.

[16]If $\delta$ is the impact of the value of cocoa production on urban population and if the total changes in urban population and cocoa production over our period are respectively $\phi$ and $\tau$, the total magnitude of this effect is $\frac{\tau \times \delta}{\phi} * 100$. This gives us how many percents of the total change in urban population can be attributed to this sole effect.

## 6.3 Decomposing the Population Effect of Cocoa

In table 4, we investigate whether this urbanization effect is part of a more general population effect where cocoa production would increase both urban density and rural density. As our Ghanaian cocoa production data use the spatial decomposition of *cocoa districts* and not *administrative districts*, we do not have total population data to include to our model. We decide to focus on Ivory Coast 1965-1998, for which we have both urban and rural population data. We investigate the impact of cocoa production value on total, urban and rural population (respectively columns (1), (2) and (3)), and on the urbanization rate (in %) which we define as urban population / total population * 100 (column (4)). We only report the OLS (see equation (1)) as the instrument is too weak for the rural population regression, making the estimated coefficient unreliable. We are confident in doing so as the OLS and IV estimates are not significantly different for Ivory Coast (see table 1). We find that one million 2000$ increases population by 83.7 urban inhabitants and that this population effect is concentrated in the cities as no impact is found for rural population: the coefficient is -1.7, which means the rural population has grown to an equal pace in cocoa producing and not-producing districts. This is also true if we consider rural density instead. This is again confirmed by the fact that cocoa production has dramatically increased the urbanization rates of those producing districts (column (4)).

We then decompose urban growth (column (2)) into the urban growth of cities already existing at time $t-1$ (column (5)) and the urban growth of new cities, those passing the 5000 threshold between $t-1$ and $t$ (column(6)). Both effects are not significantly different and account for half of the total urban growth effect. Cocoa production thus reinforces the power of pre-existing urban settlements. It also has a strong "city formation" power. Since the urban growth associated to each new city is small (from less than 5000 to more than 5000 between $t-1$ and $t$), such a strong urban growth effect of new cities must result from many new cities. This is confirmed by column (7), where 1 billion 2000$ is giving 9 new cities. Given the total value of cocoa production between 1965 and 1998, this gives 208 new cities, while there have been 313 new cities over the period. Thus, cocoa explains 66.4% of city formation. Results for Ghana are not shown but give very similar results.

## 6.4 Urban Growth in New vs. Old Cocoa-Producing Areas

We distinguish what happens in new and old-cocoa producing districts vs. the non cocoa-producing districts. We create a dummy equal to one if per capita cocoa production increases between $t-1$ and $t$ and 0 otherwise (those districts are located in new cocoa-producing areas). Then, we create a dummy equal to one if per capita cocoa production decreases between $t-1$ and $t$ and 0 otherwise (those districts are located in old cocoa-producing areas). We consider as an outcome total urban growth (col. (1) of table 5), urban growth in existing cities (col. (2)), urban growth in new cities (col. (3)) and the number of cities (col. (4)). As we do not have district population data for Ghana, we are unable to calculate per capita production. We nevertheless have regional population data for Ghana, so we use regional per capita production to create the same set of dummies. Results being very similar in both countries, we only show those results for Ivory Coast.

Results from column (1) indicate that old cocoa-producing regions experience higher urban growth: each old cocoa-producing district is receiving 78,188 more urban inhabitants, while it is 59,878 for each new cocoa-producing district. Yet, as there are fewer old cocoa-producing districts than new cocoa-producing districts, the latter are the main contributors to urban growth. Nevertheless, this indicates that cities in old cocoa-producing regions do not collapse, on the contrary. They grow even further, as if cocoa was just launching an urbanization process that was becoming self-reinforcing. Since no difference is noticeable as regards urban growth of existing cities (col. (2)), this difference between old and new cocoa-producing regions must be driven by urban growth of new cities, as confirmed by columns (3) and (4). This could be due to urban decentralization. As the existing cities become more congested, there are strong incentives for more secondary centers to appear. But this could also result from a cocoa front within the district. As land close to the already existing cities was fully exploited, the latest cocoa farmers entering the district have colonized the more remote forests where no settlement could be found. Those remote settlements become cities the next generation, when aggregate per capita production is already decreasing.

We replicate this analysis, this time using four dummies for highly/slightly decreasing/increasing per capita production at the district level (columns (5) to (8)). The goal of such an exercise is to show that our previous results are driven by those districts where production is either highly increasing or highly decreasing (for being highly decreasing, the production of those districts must have been highly increasing in the past). Thus, it is indeed cocoa production that is driving the urbanization process.

# 7 Rural-Urban Linkages and the Form of Urban Growth

We now discuss and give evidence for the various channels through which cocoa production drives urbanization, distinguishing those in new and old cocoa-producing areas. Unfortunately, we are unable to estimate the respective contribution of each of those effects to total urban growth in the forested areas. but we try to give some clue of the magnitude of each effect.

## 7.1 The Settlement of Cocoa Producers

If we follow our theoretical discussion in section 2, cocoa farmers move to new areas. They settle in the few existing settlements or fund new settlements, to use them for the colonization of surrounding forested areas. Results from the previous section indicates that urban growth in new cocoa-producing areas is driven equally by pre-existing cities and new cities. Then, new cities contribute relatively more than old cities to urban growth in old-cocoa producing regions, certainly due to a decentralization process. Then, some of those settlements where cocoa farmers settle are already urban or naturally evolve into cities. We thus expect a high share of urban inhabitants to be cocoa farmers. While they represent 52.3% of rural inhabitants in the Eastern and Western forests of Ivory Coast in 1985-88 (LSMS), they represent 20.5% of urban inhabitants. Similarly, we find that cocoa producers represent 45.2% of rural inhabitants and 14.8% of urban inhabitants in the forested regions of Ghana in 1987-88 (GLSS). This urban share has decreased to 9.7% in 2002 Ivory Coast (ENV) and 10.3% in 2005 Ghana (GLSS), as more and more cities diversified with time. We could also calculate that those cocoa farmers living in city are wealthier, both in terms of income and ownership of land and durable goods, than those living in countryside. Thus, a first impact of cocoa production on urbanization is purely demographic: there are many cocoa farmers and some of them live in city, that is why there are more cities in cocoa-producing areas (effect A of section 2).

## 7.2 The Logistics of Cocoa Beans Export

Cocoa beans must be transported from cocoa-producing areas to the ports for export. The logistics of cocoa beans export involves local and regional depots, transportation companies and port administration. Since their activity is mostly urban-based, we expect a significant share of the urban labor force to work for cocoa export. We use the 1985-88 LSMS survey to look at the industrial composition of the urban labor force. We focus on the urban individuals aged from 15 to 60 with a job in the last twelve months, and we estimate that respectively 18.1% of them work in the export of primary commodities considering the Eastern and Western Forests in Ivory Coast. Although cocoa beans are not the sole commodity exported abroad, it certainly involves most of the employees of this sector. Unfortunately, the sectoral decomposition offered in the other household surveys does not allow us to identify those workers related to the export sector.

## 7.3 Production and Consumption Linkages

The production of cocoa beans could have an impact on the birth and growth of urban sectors, through production and consumption linkages. Yet, as cocoa beans are directly exported abroad and not processed locally (given a lack of knowledge of processing processes), cocoa production has very few production linkages. In this regard, cocoa significantly differs from other cash crops whose processing can be a significant factor of development, such as cotton or sugar cane. Then, to study consumption linkages, assume a cocoa-producing household spends a share $u$ of its income $Y$ on non-essential goods. As non-essential goods are *produced in* or *distributed through* cities, they favor urbanization. If one region experiences a cocoa boom, the number of cocoa-producing households increases by $N$. The aggregate amount spent on *urbanizing* goods will then increase by $N \times uY$. A cocoa boom might increase the number of wealthy farmers in the district. Then, given the Engel curve, wealthy farmers are supposed to spend a higher share of their income on urbanizing goods.

First, using census (1988, 1998) and household survey (LSMS, EP) data, we could calculate that the total population living in cocoa-producing households in the Western Forest of Ivory Coast has increased by 775,000 people between 1988 and 1998. Given an average household size of 8.6, this corresponds to 90,210 additional households. Most of this increase was concentrated in the *Centre-Ouest* region, where the influx of cocoa-producing households accounted for 76.5% of total population change. We then replicate this exercise in Ghana (censuses 1984 and 2000, GLSS) where the total population living in cocoa-producing households in the Western province (the last regional cycle) has increased by 446,000 people between 1984 and 2000. Given an average household size of 6.6, this corresponds to 67,580 additional households. This influx then explains 61.5% of total population change in that province. We therefore understand that a cocoa boom significantly alters the population size and occupational composition of the affected region.

Second, we verify that cocoa farmers are much wealthier than the non-cocoa farmers of the same region. Using various household surveys, we regress household expenditure on a dummy equal to one if the household produces cocoa and we include village fixed effects so as to compare cocoa producers and non-cocoa farmers *within* the same village. We find that cocoa farmers are respectively 31.3% and 28.4% wealthier than their non-cocoa counterparts in the Eastern and Western forests of Ivory Coast in 1985-88 (LSMS) and 2002 (ENV). In the forest regions of Ghana, they are respectively 22.1% and 20.2% relatively wealthier in 1987-88 (GLSS) and 2002 (GLSS).

Third, we look at the structure of household expenditure for cocoa and non-cocoa farmers in the forest regions of each country. Table 6 shows this allocation for the sole cocoa farmers, in Ivory Coast (1985-88 and 2002) and Ghana (1987-88 and 2005). Total consumption is divided into three consumption aggregates: home production, food expenses and other (non-food) expenses (in % of total consumption). We then divide each consumption aggregate into six consumption subaggregates (in % of the consumption aggregate). The whole structure is rather stable through space and time. If we look at the structure of household expenditure in 1985-88 Ivory Coast, food represents 32.1% (home production) plus 25.8% (food expenses) = 57.9% of household expenditure. Home production is chiefly starchy roots that are intercropped with cocoa. Food expenses mainly concern seafood, cereals (in particular rice, which is considered as a treat in West Africa), sweets, alcohol and meat. Cocoa farmers then allocate a high share of their non-food expenditure to clothing, transfers and events, health and hygiene, housing and education. Although we cannot identify which good is *urbanizing*, we guess that food and non-food expenses imply the growth of the urban-based trade sector. Looking more specifically at other expenditure, cocoa income must have an impact on other sectors such as the textile industry, education and health, construction or public administration. Then, what is striking is that non-cocoa farmers have almost the same consumption structure (not reproduced here), although they are around 20-30% poorer. This invalidates the Engel curve. Nevertheless, as cocoa farmers are 20-30% wealthier than their non-cocoa counterparts, they still spend

20-30% more on urbanizing goods. We also show that cocoa farmers own more durable goods that non-cocoa farmers in 1985-88 (LSMS) and 1998 (EP) Ivory Coast. For each good, we regress a dummy equal to one if the household owns this good on a dummy equal to one if the household produces cocoa, including village fixed effects so as to compare cocoa producers and non-cocoa farmers within the same village. Results reported in table 7 indicate that cocoa producers more often owns a fan, a radio, a TV, a bicycle, a bike and a car. We find similar results using Ghanaian data. This indicates that cocoa farmers spend their relatively higher income to possess durable goods.

For the sake of concreteness, assume they spend 50% of their income on urbanizing goods. If the Western Forest of Ivory Coast has received 90,210 cocoa-producing households between 1988 and 1998, and if their income Y is 30% higher than the income of the non-cocoa farmers, it means that the total income spent on urbanizing goods has increased by $90,210 \times Y \times 0.5$, which is certainly a lot, and 30% more than if the region had been *counter-factually* colonized by non-cocoa farmers (the income gain would have then been $90,210 \times 0.7 \times Y \times 0.5$).

Lastly, amongst those urban individuals aged from 15 to 60 with a job in the last twelve months in the Eastern and Western forests of Ivory Coast in 1985-88 (LSMS), 35.3% of them work in the primary sector, 20.6% in the secondary sector and 44.1% in the tertiary sector. Amongst those who do not work in the primary sector, 28.6% work in the export of primary commodities, 11.8% in retail trade (mostly clothing and food trade), 11.6% in the leisure industry (hotels, restaurants, bars, hairdressers, etc.) and 6.7% in technical services (banking, insurance, professional services, etc.). Thus, except those specialized in the production and/or export of cocoa, a high share of urban inhabitants work in the distribution (and production) of goods and services that are consumed by cocoa farmers.

## 7.4 Cocoa Production and Infrastructure Investments

As aggregate income rises, infrastructure investments are realized. If infrastructure make individuals more productive or if people value infrastructure per se, those locations with better infrastructure are more attractive in the long run, causing population growth. We use various data sets to show that infrastructure today is spatially correlated with cocoa production in the past, for both rural and urban settlements. Cities of the cocoa-producing regions could have better infrastructure, which would make them grow relatively more in the long run than cities of the other regions. Or rural settlements of cocoa-producing regions could have better infrastructure, which could help their urbanization (passing the 5000 inhabitants threshold).

We first build an original GIS data set of paved roads for Ivory Coast for those years 1965, 1975, 1988 and 1998 (to be consistent with our population data). We estimate for each district-year the total length of paved roads (in kms). In a similar spirit to equation (1) (see subsection 5.2), we regress the length of paved roads at time $t$ on the value of cocoa production between $t-1$ and $t$, controlling for the length of paved roads at time $t-1$ and including district and time fixed effects. We find that cocoa production explains at least 50% of paved road building between 1965 and 1998 (results not reported but available upon request).

Second, using household surveys for Ivory Coast (EP + ENV), we estimate the share of rural and urban inhabitants with access to electricity, private tap water and toilet in 1998-2002. We drop those observations corresponding to Abidjan and Bouaké. We regress those shares on a dummy equal to 1 if per capita production is decreasing between 1965 and 1998 (the old cocoa-producing districts) and a dummy equal to one if it is increasing (the new cocoa-producing districts). Northern districts are taken as a control group. We expect residents of the old cocoa-producing regions to have a better access to infrastructure as they could realize such investments in the past. Results are reported in table 8. We do not notice any significant difference across cities of each group of districts (see columns (2), (4) and (6)). But villagers of the old cocoa-producing areas have a higher access to electricity (col. (1)), private tap water (col. (2)) and toilet (col. (3)). Considering the share of children attending school (col. (5) to (8)), this share is higher in the old cocoa-producing region than in the new cocoa-producing region where it is higher than in the Northern districts. We also have at our disposal administrative data on the number of primary and secondary schools in 1994 (Ministry of Education 1994) and the number of hospitals and health centers in 2003 (WHO 2003). This data does not distinguish rural and urban settlements, but it indicates that old cocoa-producing districts have more secondary schools and health centers per capita (results not shown but available upon request). No difference is noticed for primary schools and hospitals for which the spatial distribution is rather equal.

Third, we use the 2000 Ghanaian Facility Census to test whether past cocoa production has permitted infrastructure investments. For each administrative district and each type of settlement (rural/urban), we estimate the share inhabitants that are less than 10 kms away from various facilities: primary school, junior secondary school (JSS), senior secondary school (SSS), health center, hospital, post office, telephone. We then use 2000 Population and Housing Census to calculate for each district and type of settlement the share of inhabitants with access to electricity, private tap water and toilet. We then create a dummy equal to one if this district belongs to a region where cocoa production boomed in the 1930s (the very old cocoa-producing districts), one dummy equal to one if it belongs to a region where cocoa production boomed in 1960s (the old cocoa-producing districts) and a dummy equal to one if it belongs to a region where cocoa production boomed in the 1990s (the new cocoa-producing districts). Northern districts are taken as a control group. We expect very old cocoa-producing districts to be better endowed than old cocoa-producing districts, themselves better endowed than new cocoa-producing districts and non-producing districts. We drop those observations corresponding to Accra and Kumasi. Results reported in table 8 show that both cities and villages of the very old and old cocoa-producing areas tend to have a much better infrastructure than the other districts.

To conclude, although we cannot definitively prove that cocoa production causes infrastructure improvements, the previous analysis indicates that the old cocoa-producing regions are relatively more well-endowed in infrastructure than the other regions, and this seems to be true along several dimensions: road, electricity, water, hygiene, education, health and communications. This is certainly important for explaining why cities keep growing and rural settlements urbanize in old cocoa-producing regions.

## 7.5 Natural Increase and Urban Growth

In countries of the First Industrial Revolution, mortality was much higher in city than in countryside (Bairoch 1988, Clark and Cummins 2009). As a result, cities could not grow without massive influx of rural migrants. As both the rural-urban mortality gap and the urban-rural income differential were closing, natural increase became the main contributor to urban growth (Williamson 1990, Voigtländer and Voth 2010). In the Third World, mortality has always been much lower in city, making natural increase a strong factor of urbanization. To study this issue, we look at the urban and rural dimensions of the demographic transition in Ivory Coast and Ghana. Following Williamson (1990), we know that:

$$U_t - U_{t-1} = (UCRB_{t-1} - UCRD_{t-1}) * U_{t-1} + IM_{t-1} + UEM_{t-1} \quad (4)$$

$$R_t - R_{t-1} = (RCRB_{t-1} - RCRD_{t-1}) * R_{t-1} - IM_{t-1} + REM_{t-1} \quad (5)$$

where $U_t$ and $R_t$ are urban and rural population at time $t$, $U_t - U_{t-1}$ is urban population change between $t-1$ and $t$, and $R_t - R_{t-1}$ rural population change between $t-1$ and $t$. $CRB$ and $CRD$ are crude rate of birth and crude rate of death in city (U) and countryside (R). $IM$ is the number of internal migrants, i.e. rural-to-urban migrants. $UEM$ and $REM$ are the number of external (foreign) migrants going to the cities and the countryside. For the model to be valid, the internal migration estimates ($M_{t-1}$) in the urban and rural equations must be consistent. For each inter-census subperiod in Ivory Coast (1965-1975, 1975-1988 and 1988-1998) and Ghana (1960-1970, 1970-1984 and 1984-2000), we know urban and rural growth, as well as the urban and rural crude rates of birth and death (the difference between the two being the crude rate of natural increase $CRNI$). Since our urban data can be decomposed between Abidjan/Accra and the other cities, our model has one rural equation, one equation for non-capital cities and one equation for the main city.

We first start our analysis by looking at the evolution of urban/rural crude rates of birth, death and natural increase between the 1960s and the 1980s (see table 10). At the time of independence, there is no major natality differential between cities and the countryside. Yet, we observe a strong reduction in urban natality after 1960 in Ghana and 1975 in Ivory Coast. Regarding mortality in 1965, it was lower in Abidjan/Accra than in the other cities, where it was lower than in the countryside. Throughout the period, it has been decreasing across all places of residence, but this evolution was more impressive in countryside. In Ivory Coast, natural increase in Abidjan and the other cities peaked in 1975, while rural natural increase peaked in 1988. In Ghana, natural increase had already peaked in Accra and the other cities in 1960, while it remained high in countryside. This confirms that the demographic transition is first "urban" then "rural".

We then use equations (4) and (5) to gauge the contribution of natural increase to urban/rural growth. For each intercensal subperiod, we estimate the urban/rural population change that can be explained by natural increase. We then compare it with the observed population change. The difference between both population changes is necessarily explained by either internal or external migration. Results are reported in table 11. In Ivory Coast, the contribution of natural increase has risen from 31% in 1965-75 to 80% 1988-98 in Abidjan and from 20% in 1965-75 to 45% in 1988-98 in the other cities. In Ghana, the contribution of natural increase to urban growth has peaked during the 1970-84 period. For instance, in the other cities, it increased from 56% in 1960-70 to 90% in 1970-84. Most of urban growth at that time was fed by newborns and not rural migrants. Then, with the end of the political and economic crisis and a new cocoa boom in the Western region, migration has become again the main contributor to urban growth. If one considers the last period for both countries, natural increase explains around 45% of urban growth in non-capital cities.

To conclude, natural increase has become a determining factor of non-primate urban growth, thus making urbanization a self-reinforcing process. Cocoa production has had a strong impact on the birth and growth of cities. By permitting household and community investments in physical and human capital (better housing conditions, education, health), it certainly contributed to long-term urban growth.

# 8 Discussion

## 8.1 The Potential Effects of Resource Exhaustion

The previous subsection has shown that cities were likely to keep growing even when cocoa production is decreasing. Yet, predictions are not straightforward about per capita income in the cities of the old cocoa-producing areas. Per capita income is likely to fall given demographic growth and diminishing total income from cocoa production (pessimistic scenario). But we could imagine another scenario whereby capital accumulation and agglomeration economies would raise labor productivity enough to increase or at least maintain constant per capita income (optimistic scenario). In other words, what happens to all the newborns in the cities of the old cocoa-producing regions? Do they live better or worse than their parents? This is an essential issue as cocoa production is doomed to vanish in both Ivory Coast and Ghana in a few decades. Cocoa is produced by "eating" the forest, and both countries are eating their last available forests (in their south-west). Their entire southern territory will then be in phase 3, with old and unproductive cocoa plantations.

Beyond those microeconomic effects, resource exhaustion is likely to have a huge macroeconomic impact in both countries, one channel being a collapse of government revenue and spending. As already explained before, the governments of both countries have always fixed the producer price much below the international price, this working as an implicit taxation system of the cocoa sector. But the taxation rate of the cash crop export sector has always been quite high in African countries (Bates 1981). For the period 1961-2006, the average taxation rate is 43.8% in Ivory Coast and 49.5% in Ghana.[17] Then, although this cocoa tax was

[17]Part of this discrepancy is explained by the costs of transportation of cocoa beans from cocoa-producing regions to the ports. Yet, those costs represent very little compared to the total

supposed to be saved for stabilization purposes, it was captured by the state who used it to fund its own current consumption and investments. Figure 11 shows the cocoa tax and government consumption and investment in Ivory Coast. A similar figure is available for Ghana but not reproduced here. The loss of the cocoa tax would have a strong impact on government revenue and spending.

As we do not have precise data on the spatial distribution of government spending, we use the 2002 ENV household survey in Ivory Coast to guess who might be affected by a fall in government spending. First, as most state employees are concentrated in the cities, this fall would harm the urban sector. Indeed, we calculate that 94% of civil servants live in city. This share decreases to 80.4% if we include employees of the education and health systems. Those civil servants then represent 4.8% of the labor force of the cities. This share increases to 10.8% if we use our second definition of the public sector. Second, as already shown by the urban primacy literature (Davis and Henderson 2003), we expect the capital city to be disproportionately favored by the central government. Then, since governments also adopt redistributive regional policies, we expect the state to be more represented per capita in the poorest regions of the country, basically the Northern regions. Using both the 2002 ENV household survey and our population data, we estimate that the number of civil servants per thousand inhabitants is 23.8 in Abidjan, 4.3 in the North, 3.1 in the Eastern Forest and 2.2 in the Western Forest. Abidjan and the North will be hit harder by a fall in government spending. We now look at the 1988 Urban Infrastructure Census and we find that cities of the North receives 2.61 millions of 1988 CFAF per thousand inhabitants of "national aid" against 1.87 for the Eastern Forest, 1.30 for the Western Forest and 0.9 for Abidjan. The North is therefore significantly subsidized per capita by the central government. Unfortunately, we could not find spatial expenditure data for the more recent period.

## 8.2 The Future of African Cities?

[To be completed] This part will discuss the potential long-term effects of rural-driven urbanization given a lack of production linkages. Cities are engines of growth if urbanization is associated to the rise of sectors where agglomeration economies can be realized. The structural transformation that historically took

tax of the cocoa sector. Besides, our taxation rate is a lower-bound estimate if we consider that those governments have maintained an overvalued exchange rate, which happened in Ghana but is less true in Ivory Coast (Bates 1981, Teal 2002).

place in developed countries and that is taking place right now in China has permitted the growth of the manufacturing sector, and we know how this sector can promote long-term development when trade is liberalized. I want to show in that section that urban non-agricultural employment in my two African countries was based on the tertiary sector, especially the trade of goods that were either imported or locally produced (but without favoring intersectoral linkages). As the tertiary sector _ excluding high tech services _ cannot be a sustainable factor of development, we can wonder to what extent African cities are doomed to remain poor. High urban demographic growth might be a constraint if the young generations cannot find employment, whether in the formal or informal sectors.

# 9 Conclusion

We look at the effect of one cash crop, cocoa, on urbanization in two African countries, Ghana and Ivory Coast, during the 20th century. Our results suggest that it explains more than half of non-primate urbanization in both countries. Thus, rural windfalls can drive urban growth through mainly consumption linkages. Cities then keep growing and arising in old cocoa-producing areas, as the urbanization process becomes self-reinforcing. While not being able yet to infer what will happen to per capita income in old cocoa-producing regions, we wonder whether cities might be powerful engines of growth in Africa. A *missing* manufacturing sector and an *overgrown* trade sector in African cities could prevent them from driving national development.

# References

[1] Acemoglu, Daron, Simon Johson, and James Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117(4): 1231-1294.

[2] Angrist, Joshua, and Adriana Kugler. 2008. "Rural Windfall or a New Resource Curse? Coca, Income, and Civil Conflict in Colombia." *Review of Economics and Statistics* 90(2): 191-215.

[3] Atack, Jeremy, Fred Bateman, Michael Haines, and Robert Margo. 2009. "Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850-60." *NBER Working Paper* no 14640.

[4] Austin, Gareth. 2007a. "Resources, Techniques and Strategies South of the Sahara: Revising the Factor Endowments Perspective on African Economic Development, 1500-20001." *Economic History Review* 61(3): 587-624.

[5] Austin, Gareth. 2007b. "A Shifting Ratio and an Institutional Revolution." *Australian Economic History Review* 47(1): 95-120.

[6] Bairoch, Paul. 1988. *Cities and Economic Development: Frow the Dawn of History to the Present*. The University of Chicago Press.

[7] Balac, Ronan. 2002. "Dynamiques migratoires et économie de plantation." In *La Côte d'Ivoire à l'aube du 21ème siècle : défis démographiques et développement durable*, ed. by Editions Karthala, Paris.

[8] Banerjee, Ahbijit, Esther Duflo and Nancy Qian. 2009. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." Unpublished manuscript, Massachusetts Institute of Technology, Department of Economics.

[9] Barrios, Salvador, Luisito Bertinelli, and Eric Strobl. 2006. "Climatic Change and Rural-Urban Migration: The Case of Sub-Saharan Africa." *Journal of Urban Economics* 60(3): 360-371.

[10] Bates, Robert. 1981. *Markets and States in Tropical Africa: The Political Basis of Agricultural Policies*. Berkeley: University of California Press.

[11] Bevan, David, Paul Collier, and Jan Willem Gunning. 1987. "Consequences of a Commodity Boom in a Controlled Economy: Accumulation and Redistribution in Kenya 1975-83." *World Bank Economic Review* 1(3): 489-513.

[12] Black, Dan, Terra McKinnish, and Seth Sanders. 2005. "The Economic Impact of the Coal Boom and Bust." *Economic Journal* 115(503): 449-476.

[13] Bosworth, Barry, and Susan M. Collins. 2008. "Accounting for Growth: Comparing China and India." *Journal of Economic Perspectives* 22(1): 45-66.

[14] Carrington, William. 1996. "The Alaskan Labor Market during the Pipeline Era." *Journal of Political Economy* 104(1): 186-218.

[15] Caselli, Francisco and Wilbor John Coleman II. 2001 "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation." *Journal of Political Economy* 109(3): 584-616.

[16] Caselli, Francesco, and Guy Michaels. 2009. "Do Oil Windfalls Improve Living Standards? Evidence from Brazil." *NBER Working Paper* no 15550.

[17] Chen, Shaohua, Martin Ravaillon and Prem Sangraula. 2007. "New Evidence on the Urbanization of Global Poverty." *Population and Development Review* 33(4): 667-701.

[18] Clark, Gregory, and Neil Cummins. 2009. "Urbanization, Mortality and Fertility in Malthusian England." *American Eonomic Review* 99(2): 242-247.

[19] Cogneau, Denis et Sandrine Mesplé-Somps. 2002. "Côte d'Ivoire : la fin du mirage ?" *DIAL DT* no 2002/18.

[20] Collier, Paul, and Benedikt Goderis. 2009. "Commodity Prices, Growth, and the Natural Resource Curse: Reconciling a Conundrum." *CSAE Working Paper* no 2007-15.

[21] Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2011. "Estimating Agglomeration Economies with History, Geology, and Worker Effects." in *The Economics of Agglomeration*, edited by E. L. Glaeser, NBER and University of Chicago Press.

[22] Davis, Donald and David Weinstein. 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity ." *American Economic Review* 92(5): 1269-1289.

[23] Davis, James and J. Vernon Henderson. 2003. "Evidence on the Political Economy of the Urbanization Process." *Journal of Urban Economics* 53: 98-125.

[24] De Fries, Ruth, Thomas Rudel, Maria Uriarte, and Matthew Hansen. 2010. "Deforestation Driven by Urban Population Growth and Agricultural Trade in the Twenty-First Century." *Nature Geoscience* 3: 178-181.

[25] Duranton, Gilles. 2007. "Urban Evolutions: The Fast, the Slow, and the Still." *American Economic Review* 97(1): 197-221.

[26] Engerman, Stanley, and Kenneth Sokoloff. 2000. "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives* 14(3): 217-232.

[27] Fay, Marianne, and Charlotte Opal. 2000. "Urbanization Without Growth: A Not-So-Uncommon Phenomenon." *World Bank Policy Research Working Paper* 2412.

[28] Foster, Andrew, and Mark Rosenzweig. 2008. "Economic Development and the Decline of Agricultural Employment." Chapter 47 in (eds) T. Schultz and John Strauss, *Handbook of Development Economics*, Volume 4(5).

[29] Gallup, John Luke, Jeffrey Sachs and Andrew Mellinger. 1998. "Geography and Economic Development." *NBER Working Paper* no W6849.

[30] Harris, John, and Michael Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *American Economic Review* 60(1): 126-42.

[31] Henderson, J. Vernon. 2005. "Urbanization and Growth." in (eds) Philippe Aghion and Steven Durlauf, Handbook of Economic Growth, North Holland.

[32] Henderson, J. Vernon, Adam Storeygard, and David Weil. 2009. "Measuring Economic Growth from Outer Space." *NBER Working Paper* no 15199.

[33] Henderson, J. Vernon. 2010. "Cities and Development." *Journal of Regional Science* 50(1): 515-540.

[34] Hill, Polly. 1962. *The Migrant Cocoa Farmers of Southern Ghana: A Study in Rural Capitalism*. Cambridge University Press.

[35] Hooker, Mark, and Michael Knetter. 2001. "Measuring the Economic Effects of Military Base Closures." *Economic Inquiry* 39(4): 583-598.

[36] Irz, Xavier, and Terry Roe. 2005. "Seeds of Growth? Agricultural Productivity and the Transitional Dynamics of the Ramsey Model." *European Review of Agricultural Economics* 32(2):143-65.

[37] Kim, Sukkoo. 2007. "Immigration, Industrial Revolution and Urban Growth in the United States, 1820-1920: Factor Endowments, Technology and Geography." *NBER Working Paper* 12900.

[38] Maxwell, Simon, and Adrian Fernando. 1989. "Cash Crops in Developing Countries: The Issues, the Facts, the Policies." *World Development* 17(11): 1677-1708.

[39] Michaels, Guy. 2010. "The Long Term Consequences of Resource Based Specialization." Forthocming in *Economic Journal*.

[40] Michaels, Guy. 2008. "The Effect of Trade on the Demand for Skill - Evidence from the Interstate Highway System." *Review of Economics and Statistics* 90(4): 683-701.

[41] Michaels, Guy, Ferdinand Rauch and Stephen Redding. 2008. "Urbanization and Structural Transformation." *CEPR DP* 7016.

[42] Miguel, Edward. 2009. "Africa's Turn?" Cambridge: the MIT Press.

[43] Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49(6): 1417-1426.

[44] Nunn, Nathan, and Nancy Qian. 2010. "The Potato's Contribution to Population and Urbanization: Evidence from an Historical Experiment." Unpublished manuscript, Department of Economics, Harvard University.

[45] Overman, Henry, and Anthony Venables. 2005. "Cities in the Developing World." *CEP Discussion Paper* dp0695.

[46] Pinkovskiy, Maxim, and Xavier Sala-i-Martin. 2010. "Africa Poverty is Falling... Much Faster than You Think!" Unpublished manuscript, Massachussets Institute of Technology, Departement of Economics.

[47] Rosenthal, Stuart S. and William C. Strange. 2004 "Evidence on the Nature and Sources of Agglomeration Economies," Handbook of Urban and Regional Economics, (eds) J. Vernon Henderson and Jacques-Francois Thisse.

[48] Ruf, François. 1991. "Les crises cacaoyères. La malédiction des âges d'or ?" *Cahiers d'études africaines* 31(121): 83-134.

[49] Ruf, François. 1995. *Booms et crises du cacao: Les vertiges de l'or brun.* Ministère de la Coopération, CIRAD-SAR and Karthala.

[50] Sachs, Jeffrey, and Andrew Warner. 2001. "The Curse of Natural Resources." *European Economic Review* 45: 827-838.

[51] Satterthwaite, David. 2007. "The Transition to a Predominantly Urban World and Its Underpinnings." Paper presented at the UNU-WIDER International Workshop on Beyond the Tipping Point: Development in an Urban World, London, October 19.

[52] Teal, Francis. 2002. "Export Growth and Trade Policy in Ghana in the Twentieth Century." *The World Economy* 25: 1319-1337.

[53] Thresh, John M., and Maxwell Owusu. 1986. "The Control of Cocoa Swollen Shoot Disease in Ghana: An Evaluation of Eradication Procedures." *Crop Protection* 5: 41-52.

[54] Tiffin, Richard, and Xavier Irz. 2006. "Is Agriculture the Engine of Growth?" *Agricultural Economics* 35(1): 79-89.

[55] Tosh, John. 1980. "The Cash Crop Revolution in Tropical Africa: An Agricultural Reappraisal." *African Affairs* 79, 314: 79-94.

[56] Venables, Anthony J. 2010. "Economic Geography and African Development." *Journal of Regional Science* 89(3): 469-483.

[57] Vicente, Pedro. 2010. "Does Oil Corrupt? Evidence from a Natural Experiment in West Africa." *Journal of Development Economics* 92(1): 28-38.

[58] Voigtländer, Nico, and Hans-Joachim Voth. 2010. "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe." Unpublished manuscript, UCLA, Department of Economics.

[59] Williamson, Jeffrey. 1990.*Coping with City Growth During the British Industrial Revolution.* Cambridge University Press.

[60] WDR. 2009. *World Development Report: Reshaping Economic Geography.* Washington, DC, The World Bank.

[61] WUP. 2009. *World Urbanization Prospects: The 2009 Revision Population Database.* New York, The United Nations.

[62] Young, Alwyn. 2010. "The African Growth Miracle." Unpublished manuscript, London School of Economics, Department of Economics.

# Data Appendix

### Urban Data

Population data on Ivory Coast comes from the following documents: (i) *Annuaire Statistique de l'A.O.F.* 1949-1951 & 1950-1954, (ii) *Rapports périodiques des gouverneurs et chefs des services* 1895-1940 and *Rapports Statistiques* 1818-1920, collections of the French colonial archives, (iii) *Population de l'A.O.F. par canton et groupe ethnique 1950-1951*, Haut-Commissariat de l'A.O.F., Service de la statistique générale, (iv) *Répertoire des villages de la Côte d'Ivoire 1955*, Service de la statistique générale et de la mécanographie, Territoire de Côte d'Ivoire, (v) *Inventaire Economique de la Côte d'Ivoire 1947-1958*, (vi) *Côte d'Ivoire 1965: Population, Etudes régionales 1962-1965*, Synthèse, Ministère du Plan de Côte d'Ivoire, (vii) *Recensement général de la population 1975*, (viii) *Population de la Côte d'Ivoire, Analyse des données démographiques disponibles* 1984, Ministère de l'Economie et des Finances de Côte d'Ivoire, Direction de la statistique, (ix) *Recensement général de la population et de l'habitat 1988*, (x) *Recensement général de la population et de l'habitation 1998.* As regards population data, administrative boundaries have changed with time and we are able to get a consistent sample for the 1965-1998 period only. As regards urban data, we know the size and the geographical coordinates of any locality with more than 5000 inhabitants for the period 1901-1998. Using GIS, we are then able to recalculate district urban population for any spatial decomposition of the territory. Since we have cocoa production for 46 districts, we use those district boundaries to estimate total and urban populations.

Population data on Ghana comes from the reports of the following *Population and Housing Censuses*: 1901, 1911, 1921, 1931, 1948, 1960, 1970, 1984 and 2000. As regards total population data, administrative boundaries have changed with time and we are not able to get a consistent sample. As regards urban data, we know the size and the geographical coordinates of any locality with more than 5000 inhabitants for the whole period. Using GIS, we are then able to recalculate district urban population for any spatial decomposition of the territory. Since we have cocoa production for 73 *cocoa districts*, we use those district boundaries to estimate urban population. For the time being, as cocoa districts significantly differ from administrative districts in Ghana, we are not able to include total population to our sample.

### Cocoa Production Data

Cocoa production data on Ivory Coast is obtained by crossing the information contained in many different sources. For the pre-independence period, our two major sources are: (i) *Annuaire Statistique de l'A.O.F.* 1949-1951, and (ii) *Inventaire Economique de la Côte d'Ivoire 1947-1958.* They list cocoa production at the colonial district level for the 1945-1958 period. We then use more minor sources to obtain data for the pre-1945 period as well as more refined spatial data for the post-1944 period[18]: (i) *Documentary Material on Cacao for the Use of the Special Committee on Cacao of the Inter-American Social and Economic Council*, 1947, Pan American Union, (ii) *Félix Houphouët-Boigny: Biographie*, Frédéric Grah Mel (2003), Editions du CERAP, Maisonneuve & Larose, (iii) *Problèmes de l'économie rurale en A.O.F.*, Ch. Robequain (1937), *Annales de Géographie* 46 (260): 137-163, (iv) "Immigration, Land Transfer and Tenure Changes in Divo, Ivory Coast, 1940-80", Robert Hecht (1985), *Africa: Journal of the International African Institute* 55(3): 319-336, and (v) "Immigration et économie de plantation dans la région de Vavoua (Centre-Ouest Ivoirien)", P. Brady (1983), unpublished thesis University of Paris 10. For the post-independence period, our major sources are: (i) *Annuaire rétrospectif de statistiques agricoles et forestières 1900-1983*, Ministère de l'agriculture et des eaux et des forêts de Côte d'Ivoire, 1983, and (ii) *Caisse de stabilisation et de soutien des prix des productions agricoles (CSSPPA)*, the agricultural marketing board of Ivory Coast till its dismantling in 1999. They list cocoa production at the district level for the 1959-1997 period. Those major sources were then complemented with more refined spatial data from: (i) *La boucle du cacao en Côte d'Ivoire, Etude régionale des circuits de transport* ,P. Benvéniste (1974), Travaux et Documents de l'ORSTOM, and (ii) *Atlas de Côte d'Ivoire, 1971-1979*, Ministère du Plan de Côte d'Ivoire. In the end, we obtain cocoa production in tons for 46 districts for the following years: 1924, 1930, 1936, 1945-1997. We use linear interpolation to recalculate cocoa production for the missing years: 1925-1929, 1931-1935 and 1937-1944. Lastly, we calculate how much tons of cocoa production were produced between each census year for each district. We then use the following sources to obtain the producer price (in CFA Francs) for the period 1948-2006: (i) *Annuaire Statistique de l'A.O.F. 1949-1951 & 1950-1954*, and (ii) FAOSTAT. We obtain CFAF/$ exchange rate data and $ deflator data from: (i) UN 2010, (ii) IFS 2010, World Bank, (iii) Teal 2002. By multiplying cocoa production and the deflated producer price (in 2000$), we get the deflated total value (also in 2000$) of cocoa production going to cocoa farmers. Likewise, we calculate how much 2000$ of cocoa production were earned between each census year for each district.

Cocoa production data on Ghana is obtained similarly, although only for the main crop[19]. For the pre-independence period, we use the following documents: (i) *1927 Yearbook of the Gold Coast Department of Agriculture*, Government of Ghana, (ii) *A Historical Geography of Ghana*, Dickson (1968), (iii) *Report on the Cocoa Industry in Sierra Leone, and Notes on the Cocoa Industry of the Gold Coast*, Cadbury (1955). Those documents respectively display a very detailed map of cocoa production for the years 1926, 1936 and 1950. We can then use GIS to recalculate cocoa production using any

[18]There are 19 colonial districts ("*cercles coloniaux*") at the time of independence. By obtaining additional data at the subdivision level (the spatial unit below the colonial district), we are able to reconstruct data using the post-independence district boundaries.

[19]The main crop extends from October to June, while the light crop is between July and September. We could only obtain data for the main crop, but this is not an issue as we could calculate that the main crop amounts to 94.7% of the total crop on average during the period 1947-2000.

district boundaries. We obtain national and regional data from the following documents: (i) *The Gold Coast Cocoa Farmer*, by Polly Hill (1956), Oxford University Press, (ii) *The Gold Coast Cocoa Industry: Prices, Production and Structural Change*, Christer Gunnarson (1978), Economic History Association, Lund, Sweden, (iii) *Annual Reports and Accounts of the Ghana Marketing Board* 1957-1962, 1965, 1970, (iii) *Enquiry into the Gold Coast Cocoa Industry, 1918-1919*, W.S. Tudhope (1919), (iv) *Reports of the Department of Botanical and Agricultural Department* 1904-1955, Government of the Gold Coast, and (v) *The Economics of Cocoa Production and Marketing*, Proceedings of Cocoa Economics Research Conference in Legon, April 1973, University of Ghana.

For the post-independence period, all the following documents list cocoa production at the cocoa district level for their respective year (it should be noted that cocoa districts differ significantly from administrative districts): (i) *Analysis of Cocoa Purchases by Societies, Districts and Regions* are repots edited by the Produce Department of the Ghana Cocoa Marketing Board and available for the following years: 1961-1975, 1989 and 1994-1999, (ii) *Ghana Cocoa Marketing Board Newsletter* 1966-1974, (iii) *Ghana Cocoa Marketing Board Monthly Progress Reports* 1972-1985, and (iv) a summary of 2001-2008 district cocoa purchases which was obtained from the Ghana Cocoa Marketing Board. Since district boundaries change from year to year, we use GIS to reaggregate our data so as to get a consistent sample of district cocoa production. In the end, we obtain cocoa production for districts for the following years: 1926, 1936, 1950, 1961-1982, 1989, 2001. We use linear interpolation to recalculate cocoa production for the missing years: 1911-1925, 1927-1935, 1937-1949, 1951-1960, 1983-1988 and 1990-2000. Lastly, we calculate how much tons of cocoa production were produced between each census year for each district. We then use the following sources to obtain the producer price (in 2000 Ghanaian 2nd Cedi) for the period 1900-2006: (i) *Cocoa in the Ghanaian Economy*, Merryl Bateman (1965), unpublished thesis, MIT, (ii) FAOSTAT, (iii) "Export Growth and Trade Policy in Ghana in the Twentieth Century", Teal (2002), The World Economy 25: 1919-1937. We obtain Cedi/$ official and parallel exchange rate data and $ deflator data from: (i) Dordunoo, Cletus. 1994. "The Foreign Exchange Market and the Dutch Auction System in Ghana." *AERC Research Paper* no 24, (ii) Lawrence H. Officer. 2009. "Exchange Rates Between the United States Dollar and Forty-one Currencies." *Measuring Worth*, (iii) Teal 2002, and (iv) UN 2010. We use parallel exchange rate data when the black market premium is significantly different from 0. By multiplying cocoa production and the deflated producer price (in 2000$), we get the deflated total value (also in 2000$) of cocoa production going to cocoa farmers. Likewise, we calculate how much 2000$ of cocoa production were earned between each census year for each district.

**Other Data**

In addition to urban and cocoa production data, we collect data from various sources on Ivory Coast and Ghana. First, forest data is derived from land cover GIS data compiled by Globcover 2009. The data indicates at a very fine spatial level those areas with virgin forest or mixed virgin forest/croplands, which were areas with virgin forest before it was cleared for crop production. We are then able to know the location of the virgin forest one century ago, before cocoa production even started.

For Ivory Coast, we first use three household surveys to calculate a range of statistics that we use for our empirical analysis: (i) the 1985-1988 Living Standards and Measurement Study (LSMS), and (ii) the 1998 and 2002 *Enquêtes sur le niveau de vie des ménages* (ENV). Third, in addition to the household surveys, we use the following infrastructure datasets: (i) a GIS data set on paved roads in 1965, 1975, 1988 and 1998 using information from Michelin road maps and the book *Elephants d'Afrique 1995-2000*, (ii) the 1988 urban infrastructure census (*Recensement des Infrastructures des Communes Urbaines*), (iii) allocation maps of primary and secondary schools in 1992 from the book *Elephants d'Afrique 1995-2000*, (iv) allocation map of health facilities in 2003 from the WHO website (http://gamapserver.who.int/mapLibrary/default.aspx). Fourth, demographic transition data is compiled crossing information from the following documents: (i) reports of *Recensement général de la population et de l'habitation 1998*, (ii) *Temps des villes, temps des vivres : L'Essor du vivrier marchand en Côte-d'Ivoire*, by Jean-Louis Chaléard (2000), Karthala, (iii) *La Côte d'Ivoire à l'aube du XXIe siècle : Défis démographiques et développement durable*, by Georges Tapinos, Philippe Hugon and Patrice Vimard (2003), Karthala, (iv) *Données démographiques sur la croissance des villes en Côte d'Ivoire*, by Jean-Paul Duchemin et Jean-Pierre Trouchaud (1969), Cahiers de l'ORSTOM, Série Sciences Humaines, 1-1969. Fifth, the cocoa tax is estimated using annual FAO data on the international price in dollars of cocoa beans and exchange rate UN data. Sixth, data on government total spending, consumption and investment comes from the African Development Indicators dataset of the World Bank.

For Ghana, we first use three household surveys and two censuses to calculate a range of statistics that we use for our empirical analysis: (i) the 1987-88 and 2005 *Ghana Living Standard Survey*, (ii) the 2000 *Population and Housing Census* IPUMS sample, and (iii) the 2000 *Facility Census*. Second, demographic transition data is compiled crossing information from the following documents: (i) Patterson, David. 1979. "Health in Urban Ghana: the Case of Accra 1900-1940." *Social Science and Medicine* 13B: 251-268, (ii) Caldwell, J.C. 1967. "Fertility Differentials as Evidence of Incipient Fertility Decline in a Developing Country: The Case of Ghana." population Studies 21(1): 5-21, (iii) *The Population of Ghana 1974*, CICRED Report, (iv) *Demographic and Household Survey* 1988, (v) Agyei-Mensah, Samuel. 2005. "The Fertility Transition in Ghana Revisited." Paper prepared for the 25th IUSSP International Population Conference, Tours, France, and (vi) *Ghana's Development Agenda and Population Growth: The Unmet Need for Family Planning*, National Population Council 2006.

Table 1: Cocoa Production and Urbanization, Ivory Coast, 1948-1998.

| Dependent Variable: | District Urban Population (Pop. in ≥5000 Localities, Excluding Abidjan and Bouaké) | | | | |
|---|---|---|---|---|---|
| | *OLS* | *IV* | *IV* | *IV* | *IV+Ctrls* |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: Main Equation** | | | | | |
| District value of cocoa production | 86.9*** | 117.5*** | 120.1*** | 129.1*** | 107.4** |
| (between $t-1$ and $t$, millions of 2000$) | [16.1] | [34.3] | [37.4] | [46.9] | [41.1] |
| Close to the cocoa front | | 4,711.1 | 3,877.8 | 2,513.7 | 3,811.3 |
| | | [3,204.7] | [3,369.3] | [3,480.5] | [3,542.2] |
| **Panel B: 1st Stage** | | | | | |
| Suitable to cocoa * Close to the cocoa front | | 212.6*** | 193.3*** | 145.5*** | 129.0*** |
| | | [59.3] | [55.8] | [49.1] | [45.3] |
| Close to the cocoa front | | -46.9* | -31.8 | -1.8 | -4.3 |
| | | [25.2] | [23.6] | [15.1] | [19.4] |
| Kleibergen-Paap rk Wald F stat | | 12.8 | 12 | 8.8 | 8.1 |
| Observations | 220 | 220 | 220 | 220 | 220 |
| R-squared | 0.9 | 0.89 | 0.9 | 0.9 | 0.92 |
| Year Fixed Effects | Y | Y | Y | Y | Y |
| District Fixed Effects | Y | Y | Y | Y | Y |
| Lag of Dependent Variable | Y | Y | Y | Y | Y |
| Year Dummies * Lag of Dependent Variable | N | N | Y | Y | Y |
| Year Dummies * Suitable to Cocoa | N | N | N | Y | Y |
| Controls | N | N | N | N | Y |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the district level. All the regressions include the lag of the dependent variable and year and district fixed effects. In column (3), we allow for the effect of the lag of the dependent variable to vary with time. In column (4), we include year dummies interacted with a dummy equal to 1 if the district is suitable to cocoa production. The set of controls we use in column (5) is: (a) dummies equal to 1 if the district is coastal, has a railway or a paved road going directly to Abidjan, all interacted with a year trend, (b) year dummies interacted with the district distance to the coast, and the district average annual sum of rainfall for the period 1900-2000.

Table 2: Cocoa Production and Urbanization, Ghana, 1921-2000.

| Dependent Variable: | District Urban Population (Pop. in ≥5000 Localities, Excluding Accra and Kumasi) | | | | |
|---|---|---|---|---|---|
| | *OLS* | *IV* | *IV* | *IV* | *IV+Ctrls* |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: Main Equation** | | | | | |
| District value of cocoa production | 12.1 | 30.4 | 44.8* | 87.0*** | 86.6*** |
| (between $t-1$ and $t$, millions of 2000$) | [16.2] | [23.3] | [22.5] | [31.1] | [31.2] |
| Close to the cocoa front | | 1,483.3 | 1,211.9 | -1,060.4 | -1,329.9 |
| | | [1,170.9] | [1,267.4] | [834.0] | [1,725.0] |
| **Panel B: 1st Stage** | | | | | |
| Suitable to cocoa * Close to the cocoa front | | 87.6*** | 86.7*** | 78.3*** | 94.4*** |
| | | [16.6] | [16.1] | [17.2] | [20.8] |
| Close to the cocoa front | | -14.7* | -14.1* | -4 | -20.1** |
| | | [8.3] | [8.1] | [3.2] | [9.5] |
| Kleibergen-Paap rk Wald F stat | | 27.9 | 29 | 20.8 | 20.6 |
| Observations | 426 | 426 | 426 | 426 | 426 |
| R-squared | 0.9 | 0.9 | 0.92 | 0.91 | 0.91 |
| Year Fixed Effects | Y | Y | Y | Y | Y |
| District Fixed Effects | Y | Y | Y | Y | Y |
| Lag of Dependent Variable | Y | Y | Y | Y | Y |
| Year Dummies * Lag of Dependent Variable | N | N | Y | Y | Y |
| Year Dummies * Suitable to Cocoa | N | N | N | Y | Y |
| Controls | N | N | N | N | Y |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the district level. All the regressions include the lag of the dependent variable and year and district fixed effects. In column (3), we allow for the effect of the lag of the dependent variable to vary with time. In column (4), we include year dummies interacted with a dummy equal to 1 if the district is suitable to cocoa production. The set of controls we use in column (5) is: (a) dummies equal to 1 if the district is coastal, has a railway or a paved road going directly to Accra, all interacted with a year trend, (b) year dummies interacted with the district distance to the coast, and the district average annual sum of rainfall for the period 1900-2000.

Table 3: Cocoa Production and Urbanization, IV, Specification Checks.

| Dependent Variable: | District Urban Population (Pop. in ≥5000 Localities, Excluding the Two Main Cities) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Ivory Coast, 1948-1998** | | | | **Ghana, 1921-2000** | | | |
| | Level (1) | △ (2) | Density (3) | Level (4) | Level (5) | △ (6) | Density (7) | Level (8) |
| **Main Equation** | | | | | | | | |
| District value of cocoa production | 107.4** | 113.4*** | | | 86.6*** | 99.8** | | |
| (between $t-1$ and $t$, millions of 2000$) | [41.1] | [31.5] | | | [31.2] | [47.1] | | |
| District density of value of cocoa production | | | 78.5** | | | | 54.2* | |
| (between $t-1$ and $t$, millions of 2000$/sq.km.) | | | [32.4] | | | | [32.0] | |
| District cocoa production | | | | 0.15** | | | | 0.08** |
| (between $t-1$ and $t$, tons) | | | | [0.06] | | | | [0.03] |
| Close to the cocoa front | 3,811 | 4,062 | 1.0** | 2,992 | -1,330 | -2,806 | -1.5** | -1,392 |
| | [3,542] | [3,534] | [0.5] | [4,052] | [1,725] | [2,701] | [0.7] | [1,650] |
| Kleibergen-Paap rk Wald F stat | 8.1 | 7.1 | 14.8 | 10.2 | 20.6 | 19.8 | 29.1 | 16.7 |
| Observations | 220 | 220 | 220 | 220 | 426 | 426 | 426 | 426 |
| R-squared | 0.92 | 0.73 | 0.93 | 0.93 | 0.91 | 0.35 | 0.99 | 0.91 |
| Year Fixed Effects | Y | Y | Y | Y | Y | Y | Y | Y |
| District Fixed Effects | Y | Y | Y | Y | Y | Y | Y | Y |
| Year Dummies * Lag of Dependent Variable | Y | N | Y | Y | Y | N | Y | Y |
| Year Dummies * Suitable to Cocoa | Y | Y | Y | Y | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y | Y | Y | Y | Y |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the district level. All the regressions include year and district fixed effects, and year dummies interacted with a dummy equal to 1 if the district is suitable to cocoa production. Except in columns (2) and (5) where the outcome is the change variable, we also include year dummies interacted with the lag of the dependent variable. The set of controls we use include: (a) dummies equal to 1 if the district is coastal, has a railway or a paved road going directly to the capital city, all interacted with a year trend, (b) year dummies interacted with the district distance to the coast, and the district average annual sum of rainfall for the period 1900-2000. Those results are robust to the exclusion of controls.

Table 4: Cocoa Production and Total, Urban and Rural Populations, Ivory Coast, 1965-1998.

| Dependent Variable: | **Population** | | | | **Urban Population** | | |
|---|---|---|---|---|---|---|---|
| | Total | Urban | Rural | Urban. Rate | In Old Cities | In New Cities | Number of Cities |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| District value of cocoa production | 83.7* | 91.3*** | -1.7 | 0.01** | 54.6*** | 56.8** | 0.009*** |
| (between $t-1$ and $t$, millions of 2000$) | [49.4] | [22.9] | [42.8] | [0.004] | [10.8] | [21.6] | [0.003] |
| Observations | 132 | 132 | 132 | 132 | 132 | 132 | 132 |
| R-squared | 0.87 | 0.87 | 0.75 | 0.74 | 0.88 | 0.43 | 0.75 |
| Year Fixed Effects | Y | Y | Y | Y | Y | Y | Y |
| District Fixed Effects | Y | Y | Y | Y | Y | Y | Y |
| Year Dummies * Lag of Dependent Variable | Y | Y | Y | Y | Y | Y | Y |
| Year Dummies * Suitable to Cocoa | Y | Y | Y | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y | Y | Y | Y |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the district level. All the regressions include year and district fixed effects, and year dummies interacted with: (i) the lag of the dependent variable, and (ii) a dummy equal to 1 if the district is suitable to cocoa production. Those results are robust to the inclusion of controls.

Table 5: Population Change in Old vs. New Cocoa-Producing Areas, Ivory Coast 1965-1998, OLS.

| Dependent Variable: | **District Urban Population** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | All Cities (1) | In Old Cities (2) | In New Cities (3) | Number of Cities (4) | All Cities (5) | In Old Cities (6) | In New Cities (7) | Number of Cities (8) |
| Decreasing pc Production | 78,188*** | 59,004*** | 55,847*** | 7.3*** | | | | |
| | [13,791] | [8,068] | [12,485] | [1.7] | | | | |
| Increasing pc Production | 59,878*** | 60,998*** | 39,137*** | 5.2*** | | | | |
| | [20,128] | [5,362] | [7,084] | [0.8] | | | | |
| Highly Decreasing pc Prod. | | | | | 126,539*** | 76,296*** | 90,645*** | 13.0** |
| | | | | | [32,552] | [21,017] | [31,963] | [5.0] |
| Slightly Decreasing pc Prod. | | | | | 64,904*** | 51,477*** | 48,739*** | 6.5*** |
| | | | | | [13,147] | [8,331] | [9,967] | [1.2] |
| Highly Increasing pc Prod. | | | | | 55,207*** | 57,402*** | 36,721*** | 5.0*** |
| | | | | | [14,942] | [5,701] | [5,690] | [0.7] |
| Slightly Increasing pc Prod. | | | | | 90,444*** | 66,337*** | 59,154*** | 7.8*** |
| | | | | | [19,326] | [8,500] | [14,203] | [1.4] |
| Observations | 132 | 132 | 132 | 132 | 132 | 132 | 132 | 132 |
| R-squared | 0.84 | 0.86 | 0.33 | 0.72 | 0.88 | 0.87 | 0.44 | 0.77 |
| Year Fixed Effects | Y | Y | Y | Y | Y | Y | Y | Y |
| District Fixed Effects | Y | Y | Y | Y | Y | Y | Y | Y |
| Year Dummies * Lag of Dep. Var. | Y | Y | Y | Y | Y | Y | Y | Y |
| Year Dummies * Suitable | Y | Y | Y | Y | Y | Y | Y | Y |
| Controls | N | N | N | N | N | N | N | N |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the district level. All the regressions include year and district fixed effects, and year dummies interacted with: (i) the lag of the dependent variable, and (ii) a dummy if more than 50% of district area is suitable to cocoa production. Those results are robust to the inclusion of controls.

Table 6: Allocation of Household Expenditures for Cocoa-Producing Households (in %).

| **Ivory Coast, Forest, 1985-1988 and 2002** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1988 | 2002 | | 1988 | 2002 | | 1988 | 2002 |
| Home Production | 32.1 | 28.9 | Food Expenses | 25.8 | 27.9 | Other Expenses | 42 | 43.2 |
| *Of which (%):* | | | *Of which (%):* | | | *Of which (%):* | | |
| Starchy roots | 51.7 | 40.1 | Seafood | 30.0 | 29.0 | Clothing | 22.9 | 21.7 |
| Cereals | 25.2 | 24.5 | Cereals | 16.1 | 22.3 | Transfers & Events | 15.0 | 14.3 |
| Vegetables | 10.1 | 11.9 | Sweets | 13.9 | 8.6 | Health & Hygiene | 14.0 | 23.6 |
| Meat | 5.8 | 6.3 | Alcohol | 9.3 | 3.1 | Housing | 14.0 | 17.4 |
| Oils | 5.2 | 9.1 | Meat | 8.6 | 8.5 | Education | 12.2 | 6.6 |
| Fruits | 1.3 | 5.8 | Oils | 8.0 | 8.0 | Bills & Fuel | 9.1 | 11.3 |

| **Ghana, Forest, 1987-1988 and 2005** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1988 | 2005 | | 1988 | 2005 | | 1988 | 2005 |
| Home Production | 36.3 | 18.5 | Food Expenses | 33.3 | 37.2 | Other Expenses | 30.1 | 44.3 |
| *Of which (%):* | | | *Of which (%):* | | | *Of which (%):* | | |
| Starchy roots | 59.5 | 67.5 | Seafood | 35.8 | 35.1 | Clothing | 26.6 | 19.4 |
| Vegetables | 10.9 | 8.8 | Starchy roots | 17.5 | 6.0 | Health & Hygiene | 20.5 | 16.3 |
| Cereals | 9.9 | 6.7 | Vegetables | 11.0 | 12.1 | Housing | 9.4 | 10.6 |
| Fruits | 7.8 | 4.2 | Cereals | 9.1 | 17.5 | Transfers & Events | 11.9 | 10.9 |
| Meat | 6.2 | 4.0 | Meat | 8.0 | 7.4 | Education | 6.9 | 15.2 |
| Oils | 5 | 6 | Oils | 6 | 5 | Bills & Fuel | 8.9 | 11.3 |

Note: We use the 1985-88 LSMS and 2002 ENV household surveys for Ivory Coast and the 1987-88 and 2005 GLSS (1, 2 & 5) household surveys to estimate the allocation (in %) of total household expenditure for cocoa-producing households in the Forest regions of both countries. We first show the allocation across three consumption aggregates: home production, food expenses and other expenses. Second, for each of those consumption aggregates, we show the six main consumption subaggregates and their contribution (in %) to the value of the consumption aggregate.

Table 7: Cocoa Production and Ownership of Durable Goods, Ivory Coast, 1985-1988 and 1998.

| **Forest, Ivory Coast, 1985-1988** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Owns: | Gas Cooker | Fridge | Fan | Radio | TV | Bicycle | Bike | Car |
| Cocoa HH | -0.01*** | -0.01 | 0.01** | 0.11*** | 0.02*** | 0.13*** | 0.08*** | 0.03*** |
| | [0.00] | [0.01] | [0.01] | [0.01] | [0.01] | [0.01] | [0.01] | [0.00] |
| Observations | 20443 | 20443 | 20443 | 20443 | 20443 | 20443 | 20443 | 20443 |
| R-squared | 0.19 | 0.34 | 0.4 | 0.19 | 0.37 | 0.32 | 0.22 | 0.17 |
| Village-Time FE | Y | Y | Y | Y | Y | Y | Y | Y |

| **Forest, Ivory Coast, 1998** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Owns: | Gas Cooker | Fridge | Fan | Radio | TV | Bicycle | Bike | Car |
| Cocoa HH | 0.01* | 0.01 | 0.03*** | 0.14*** | 0.08*** | 0.22*** | 0.02*** | 0.02*** |
| | [0.01] | [0.01] | [0.01] | [0.01] | [0.01] | [0.01] | [0.01] | [0.00] |
| Observations | 9206 | 9206 | 9206 | 9206 | 9206 | 9206 | 9206 | 9206 |
| R-squared | 0.23 | 0.4 | 0.35 | 0.27 | 0.31 | 0.32 | 0.19 | 0.24 |
| Village-Time FE | Y | Y | Y | Y | Y | Y | Y | Y |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. We use the 1985-88 LSMS and 1998 ENV household surveys and we create a set of dummies equal to 1 if the household owns the following goods: gas cooker, fridge, fan, radio, TV, bicycle, bike, car. We then regress each dummy on a dummy equal to 1 if the household produces cocoa. We also regress the number of spouses of the household head on the cocoa household dummy. As we drop those households that are not involved in agricultural activity, we mechanically compare cocoa producers and non-cocoa farmers (which we take as a benchmark). Since, we include a set of village-time fixed effects, we control for villages characteristics and we compare cocoa producers and non-cocoa farmers *within* their village.

Table 8: Cocoa Production and Infrastructure Investments, Ivory Coast, 1998-2002.

| Dependent Variable: | **Share of Households with Access to** | | | | | | **Share of Children Attending School** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Electricity | | Private Tap Water | | Toilet | | 6-11 yo | | 12-15 yo | |
| | Rural | Urban | Rural | Urban | Rural | Urban | Rural | Urban | Rural | Urban |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Old Cocoa Area | 0.28*** | -0.03 | 0.09*** | -0.04 | 0.46*** | 0.05 | 0.18** | -0.05 | 0.20** | -0.02 |
| | [0.09] | [0.11] | [0.03] | [0.08] | [0.15] | [0.07] | [0.08] | [0.06] | [0.08] | [0.08] |
| New Cocoa Area | 0.12** | -0.03 | -0.02 | -0.07 | 0.16** | -0.01 | 0.13*** | -0.03 | 0.1 | 0.01 |
| | [0.05] | [0.05] | [0.02] | [0.06] | [0.07] | [0.06] | [0.05] | [0.05] | [0.06] | [0.05] |
| Observations | 44 | 45 | 44 | 45 | 44 | 45 | 44 | 45 | 44 | 45 |
| R-squared | 0.22 | 0.01 | 0.2 | 0.03 | 0.27 | 0.01 | 0.19 | 0.02 | 0.11 | 0 |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. The 1998 EP and 2002 ENV surveys are used to estimate the share of inhabitants with access to private tap water, electricity, and toilets, as well as the share of children attending school for two age groups: 6-11 corresponding to primary schooling, and 12-15 corresponding to secondary schooling. Old cocoa areas correspond to districts whose per capita cocoa production has been decreasing between 1965 and 1998 (mostly in the Eastern Forest). New cocoa areas correspond to districts whose per capita cocoa production has been increasing between 1965 and 1998 (mostly in the Western Forest). Northern districts are taken as a control group.

Table 9: Cocoa Production and Infrastructure Investments, Ghana, 2000.

| Dependent Variable: | **Share of Inhabitants ≤ 10 Kms From** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Primary School | | JSS | | SSS | | Health Centre | | Hospital | |
| | Rural | Urban | Rural | Urban | Rural | Urban | Rural | Urban | Rural | Urban |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Very Old Cocoa Area | 0.04* | 0.00 | 0.13* | 0.00 | 0.16* | 0.07* | 0.14* | 0.02** | 0.12* | 0.10 |
| | [0.01] | [0.00] | [0.05] | [0.00] | [0.07] | [0.03] | [0.05] | [0.01] | [0.05] | [0.06] |
| Old Cocoa Area | 0.04* | 0.00 | 0.12 | 0.00 | 0.18 | 0.02 | 0.14* | 0.01 | 0.17* | 0.09 |
| | [0.01] | [0.00] | [0.05] | [0.00] | [0.08] | [0.03] | [0.06] | [0.01] | [0.06] | [0.07] |
| New Cocoa Area | 0.03 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.09 | -0.02 |
| | [0.01] | [0.00] | [0.05] | [0.00] | [0.07] | [0.03] | [0.05] | [0.01] | [0.05] | [0.06] |
| Observations | 104 | 100 | 104 | 100 | 104 | 100 | 104 | 100 | 104 | 100 |
| R-squared | 0.27 | 0.00 | 0.30 | 0.00 | 0.13 | 0.04 | 0.19 | 0.02 | 0.13 | 0.03 |

| Dependent Variable: | **Share of Inhabitants ≤ 10 Kms From** | | | | **Share of Inhabitants with Access to** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Post Office | | Telephone | | Electricity | | Private Tap | | Toilet | |
| | Rural | Urban | Rural | Urban | Rural | Urban | Rural | Urban | Rural | Urban |
| | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) | (20) |
| Very Old Cocoa Area | 0.36** | 0.11* | 0.33** | 0.17** | 0.11** | 0.12** | 0.02** | 0.07** | 0.60** | 0.27* |
| | [0.06] | [0.04] | [0.04] | [0.04] | [0.02] | [0.02] | [0.00] | [0.02] | [0.14] | [0.08] |
| Old Cocoa Area | 0.30* | 0.10* | 0.25* | 0.12* | 0.15** | 0.14** | 0.01 | 0.01 | 0.55** | 0.25* |
| | [0.09] | [0.04] | [0.08] | [0.04] | [0.03] | [0.02] | [0.01] | [0.03] | [0.14] | [0.09] |
| New Cocoa Area | 0.12 | 0.01 | 0.13* | 0.05 | 0.12** | 0.18** | 0 | 0 | 0.42* | 0.21* |
| | [0.06] | [0.04] | [0.04] | [0.04] | [0.02] | [0.02] | [0.00] | [0.02] | [0.14] | [0.08] |
| Observations | 104 | 100 | 104 | 100 | 104 | 100 | 104 | 100 | 104 | 100 |
| R-squared | 0.33 | 0.15 | 0.24 | 0.08 | 0.29 | 0.2 | 0.07 | 0.05 | 0.61 | 0.39 |

Note: Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors are clustered at the region level. The 2000 Facility Census is used to calculate the average distance (in km) to various types of facility. JSS and SSS are junior and senior secondary schools. The 2000 Population and Housing Census is used to estimate the share of inhabitants with access to electricity, private tap water and toilets. Observations corresponding to Accra and Kumasi are dropped. Very old cocoa areas correspond to districts whose maximum cocoa production was reached during the 1930s. Old cocoa areas correspond to districts whose maximum cocoa production was reached in the 1960s. New cocoa areas correspond to districts whose maximum cocoa production was reached in the 1990s. Northern districts are taken as a control group.

Table 10: Crude Rates of Birth, Death and Natural Increase per Place of Residence.

| **Ivory Coast 1965-1998** | | | | **Ghana 1960-2000** | | | |
|---|---|---|---|---|---|---|---|
| Crude Rate of (%o) | | | | Crude Rate of (%o) | | | |
| | Birth | Death | Natural increase | | Birth | Death | Natural increase |
| **Rural** | | | | **Rural** | | | |
| 1965 | 50 | 30 | 20 | 1960 | 52 | 23 | 29 |
| 1975 | 48 | 20 | 28 | 1970 | 51 | 21 | 30 |
| 1988 | 52 | 15 | 37 | 1984 | 48 | 17 | 31 |
| **Urban** | | | | **Urban** | | | |
| 1965 | 46 | 26 | 20 | 1960 | 49 | 20 | 29 |
| 1975 | 51 | 14 | 37 | 1970 | 45 | 14 | 31 |
| 1988 | 42 | 13 | 29 | 1984 | 37 | 14 | 23 |
| **Abidjan** | | | | **Accra** | | | |
| 1965 | 47 | 14 | 33 | 1960 | 43 | 14 | 30 |
| 1975 | 50 | 9 | 41 | 1970 | 36 | 7 | 28 |
| 1988 | 42 | 9 | 33 | 1984 | 34 | 11 | 23 |

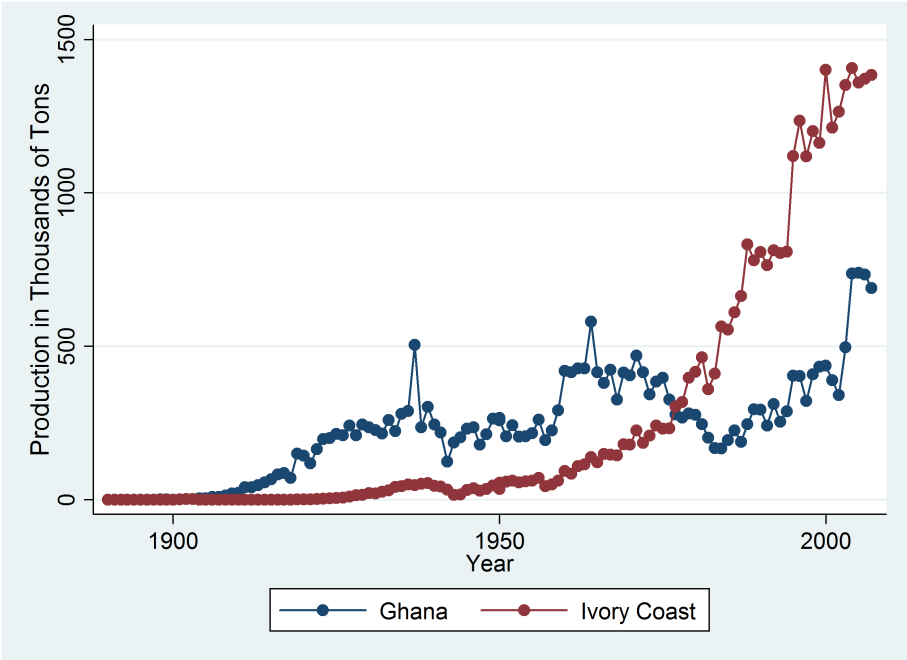
Note: see data appendix for details.

Table 11: Contribution (%) of Natural Increase and Migration to Population Change.

| | Ivory Coast 1965-1998 | | | Ghana 1960-2000 | |
|---|---|---|---|---|---|
| | Natural increase (%) | Migration (%) | | Natural increase (%) | Migration (%) |
| **Rural** | | | **Rural** | | |
| 1965-1975 | 84 | 16 | 1960-1970 | 227 | -127 |
| 1975-1988 | 131 | -31 | 1970-1984 | 139 | -39 |
| 1988-1998 | 162 | -62 | 1984-2000 | 264 | -164 |
| **Urban** | | | **Urban** | | |
| 1965-1975 | 20 | 80 | 1960-1970 | 56 | 44 |
| 1975-1988 | 46 | 54 | 1970-1984 | 90 | 10 |
| 1988-1998 | 45 | 55 | 1984-2000 | 46 | 54 |
| **Abidjan** | | | **Accra** | | |
| 1965-1975 | 31 | 69 | 1960-1970 | 33 | 67 |
| 1975-1988 | 62 | 38 | 1970-1984 | 76 | 24 |
| 1988-1998 | 80 | 20 | 1984-2000 | 39 | 61 |

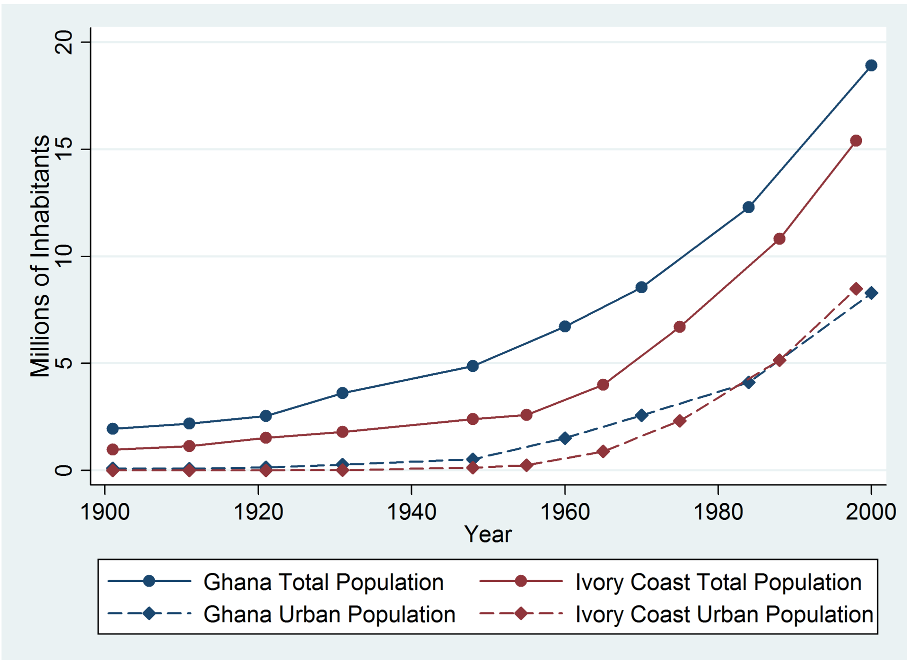
Note: We use historical data on population and crude rates of birth and death by place of residence (rural / urban / main city) to estimate the contribution (in %) of natural increase and migration to population change for each subperiod between two census dates. For each subperiod - place of residence, the contribution of natural increase (in %) is calculated as initial population times the rate of natural increase between year t-1 and year t over population change times 100. The contribution of migration (in %) is calculated as 100 minus the contribution of natural increase. Unfortunately, we cannot distinguish internal and external migration.

Figure 1: Cocoa Production (in thousands tons), 1890-2007.



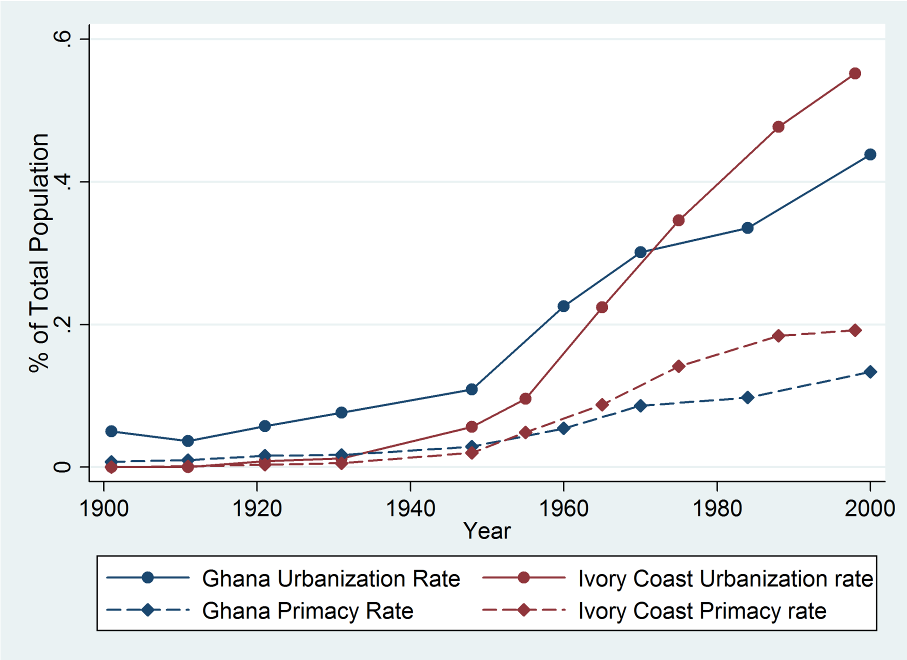

Sources: see data appendix for more details.

Figure 2: Total and Urban Populations, 1900-2000.



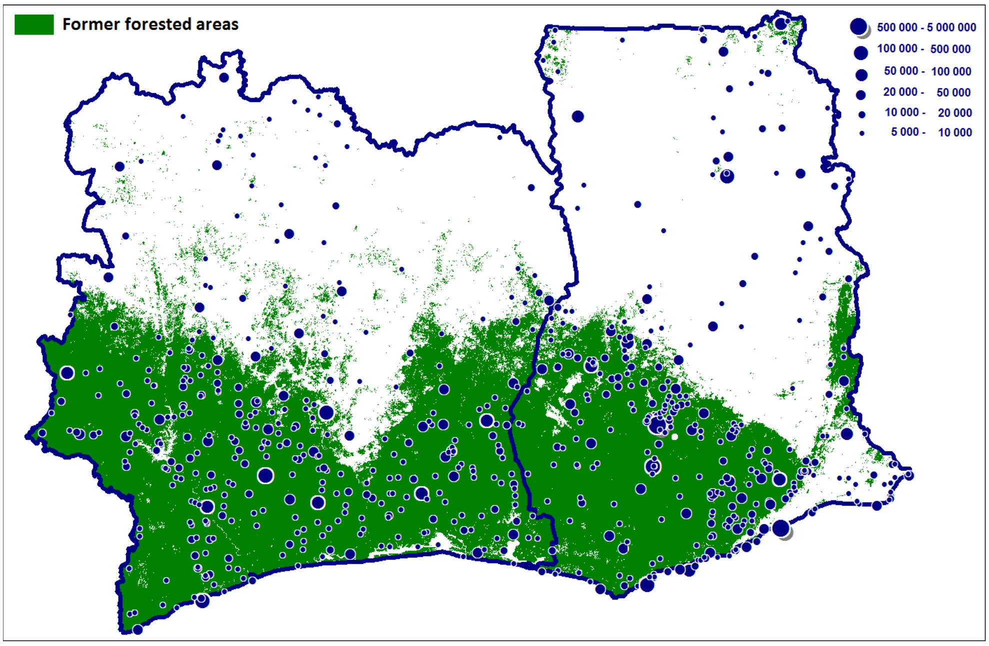

Sources: see data appendix for more details.

Figure 3: Urbanization and Primacy Rates, 1900-2000.



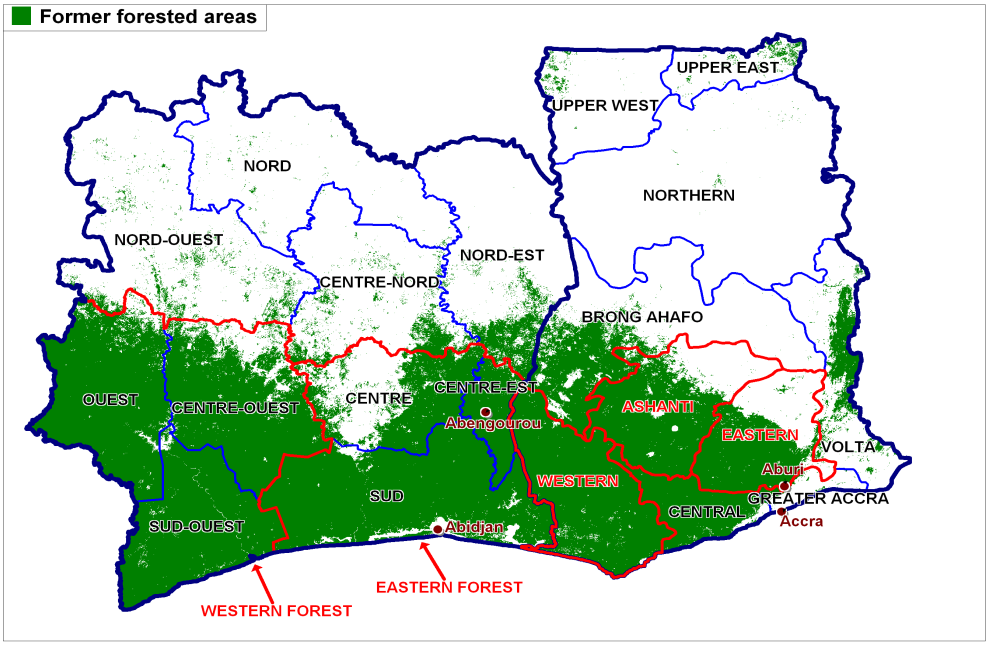

Sources: see data appendix for more details. The urbanization rate is defined as urban population over total population * 100, and the primacy rate is the size of the capital city over total population * 100.

Figure 4: Former Forested Areas and Cities in 1998 (Ivory Coast) / 2000 (Ghana).



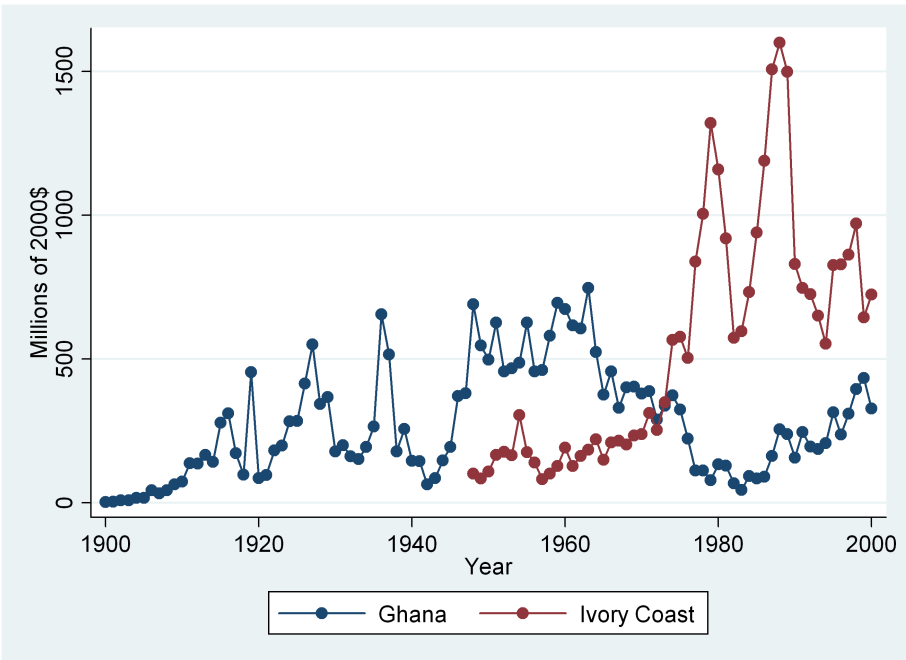

Sources: see data appendix for more details. Former forested areas were a virgin forest one century ago, before cocoa production started.

Figure 5: Former Forested Areas, Regions and Historical Starting Points.




Sources: see data appendix for more details. Former forested areas were virgin forests one century ago, before cocoa production started. Aburi is the historical starting point of Ghanaian cocoa production. Abengourou is the historical starting point of Ivorian cocoa production.

Figure 6: Value of Cocoa Production Going to Cocoa Farmers, 1900-2000.




Sources: see data appendix for more details. The value of cocoa production in year $t$ is calculated as the quantity (in tons) produced that year multiplied by the price per ton (in 2000$) that year.

Figure 7: District Density of Cocoa Production and Cities around 1948.

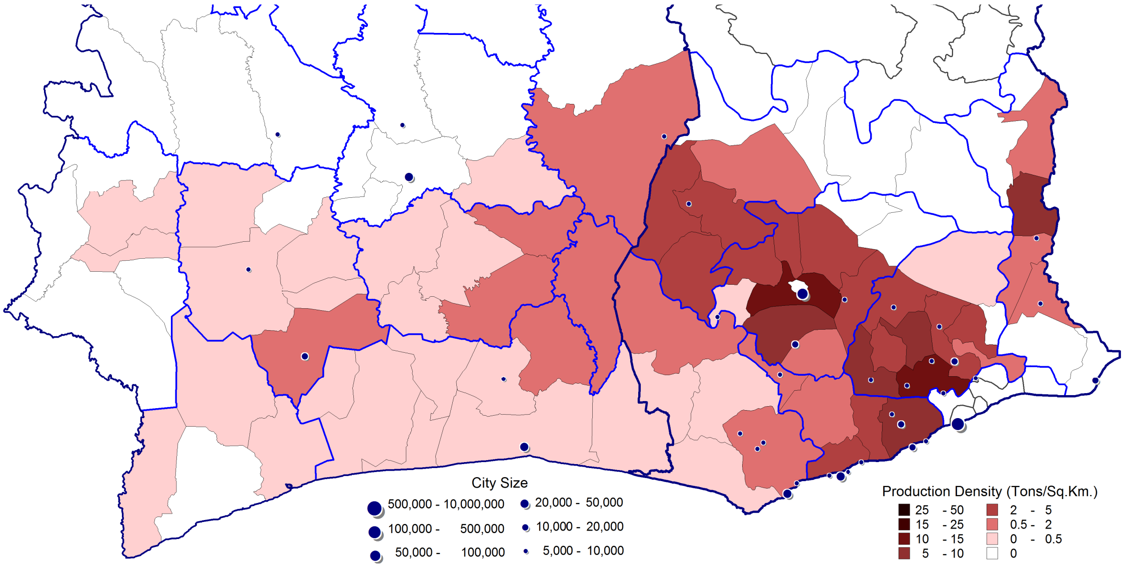



Figure 8: District Density of Cocoa Production and Cities around 1960 (Ghana) / 1965 (Ivory Coast).

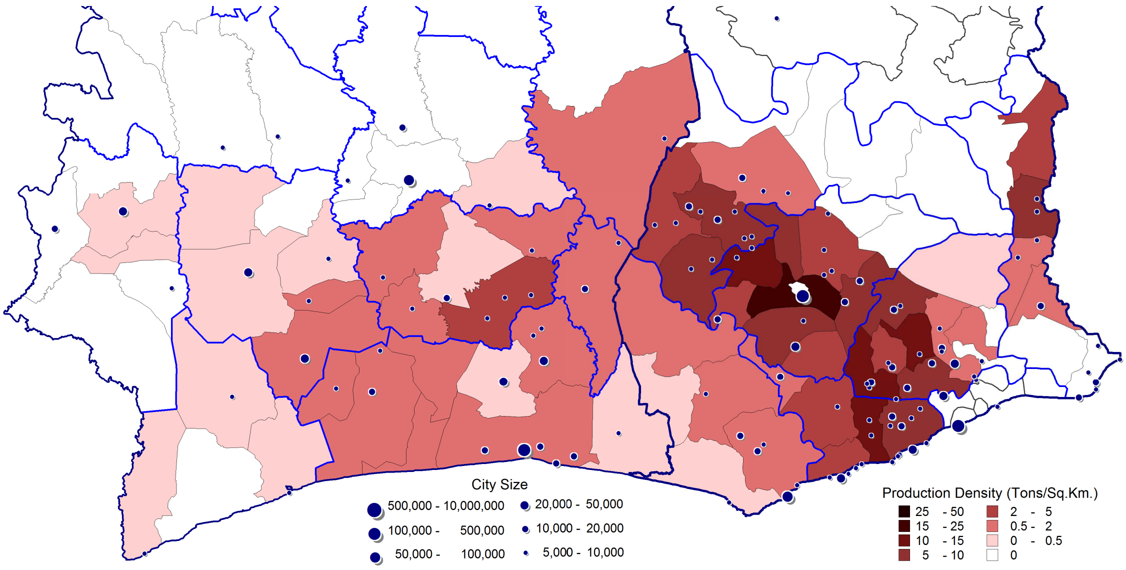


Figure 9: District Density of Cocoa Production and Cities around 1970 (Ghana) / 1975 (Ivory Coast).

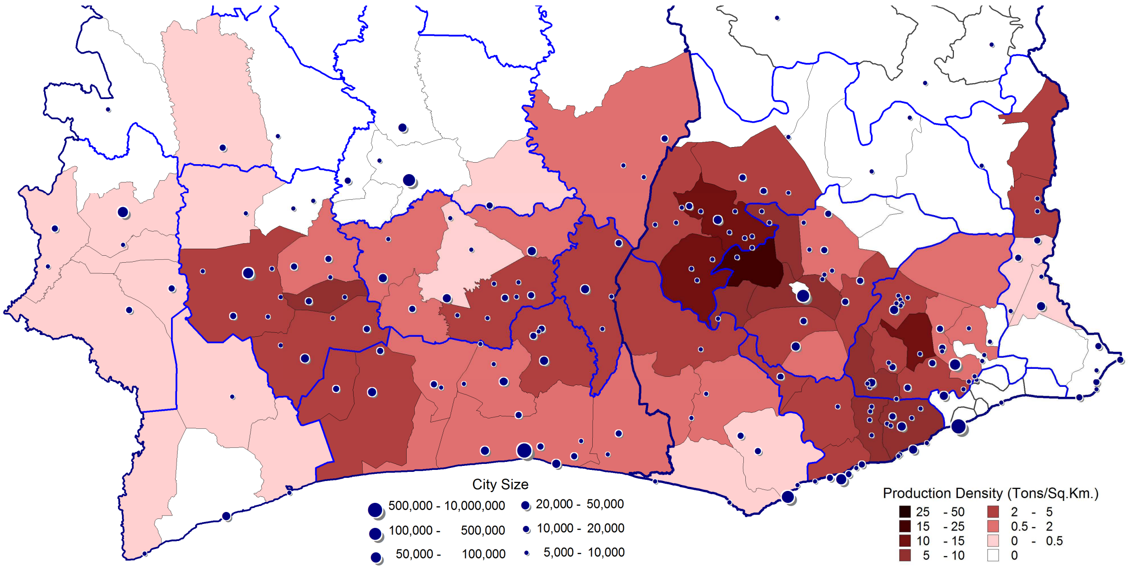



Figure 10: District Density of Cocoa Production and Cities around 1984 (Ghana) / 1988 (Ivory Coast).

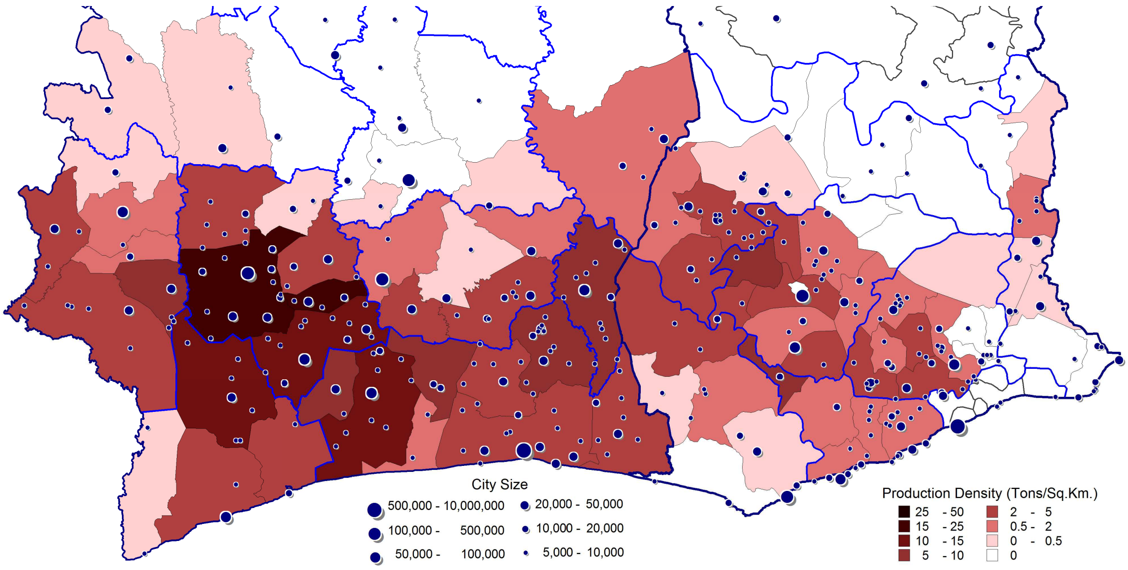


Figure 11: District Density of Cocoa Production and Cities around 2000 (Ghana) / 1998 (Ivory Coast).

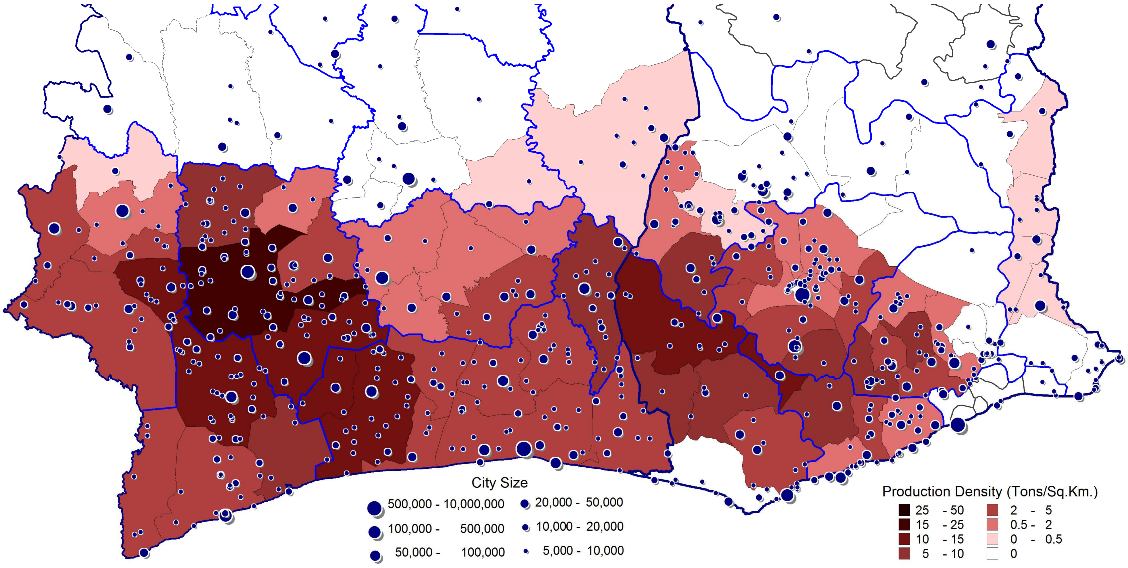
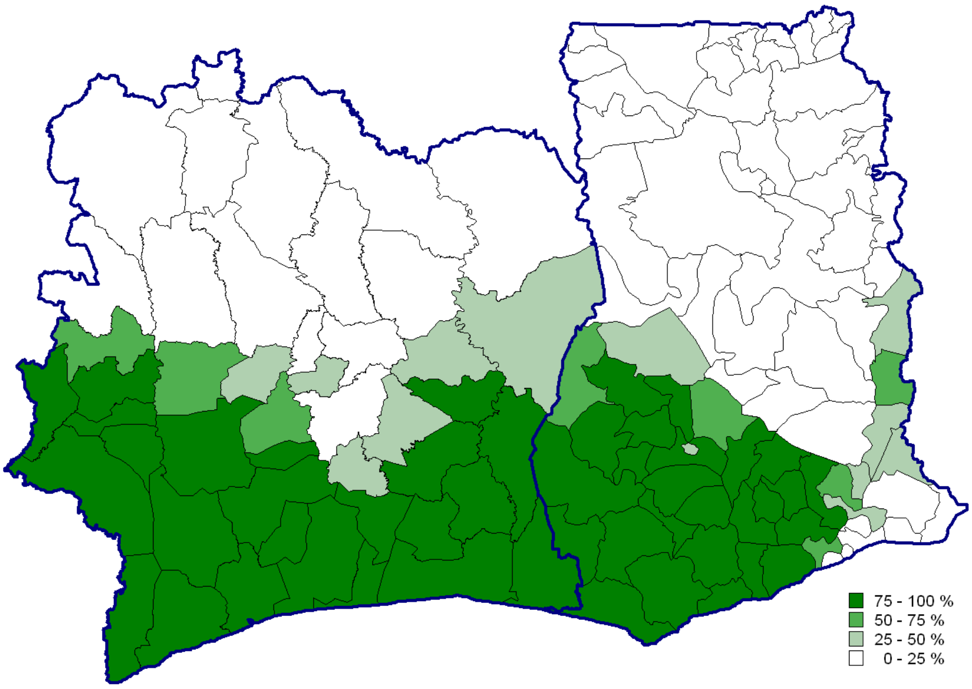



Figure 12: Share of District Area Suitable to Cocoa Production.




Sources: for each district, we calculate the share of district area being suitable to cocoa production (with a virgin forest one century ago). We create four categories of land suitability: when the share is inferior to 25%, when it is between 25 and 50%, 50 and 75% or more than 75%.

Figure 13: Centroids and Longitude Bands of One Degree.

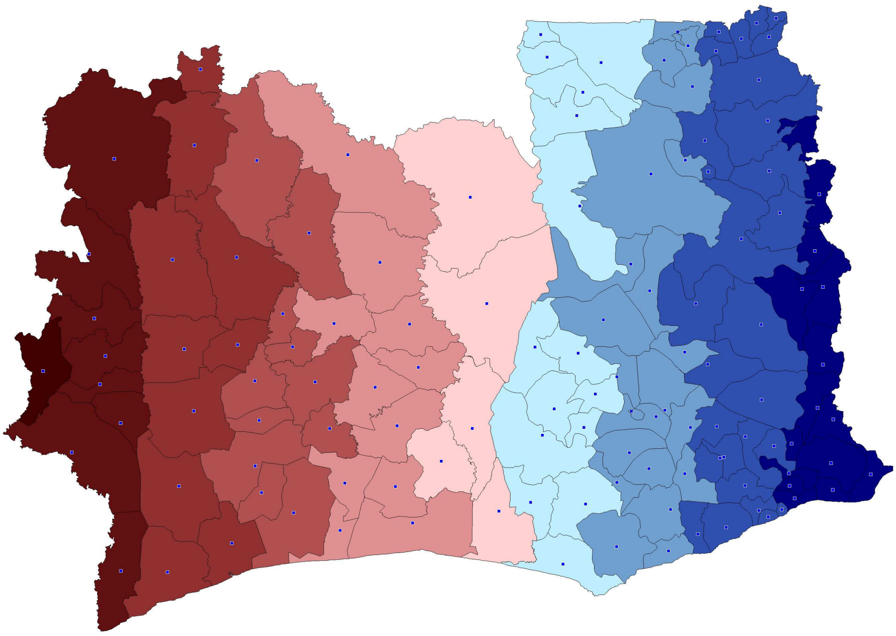
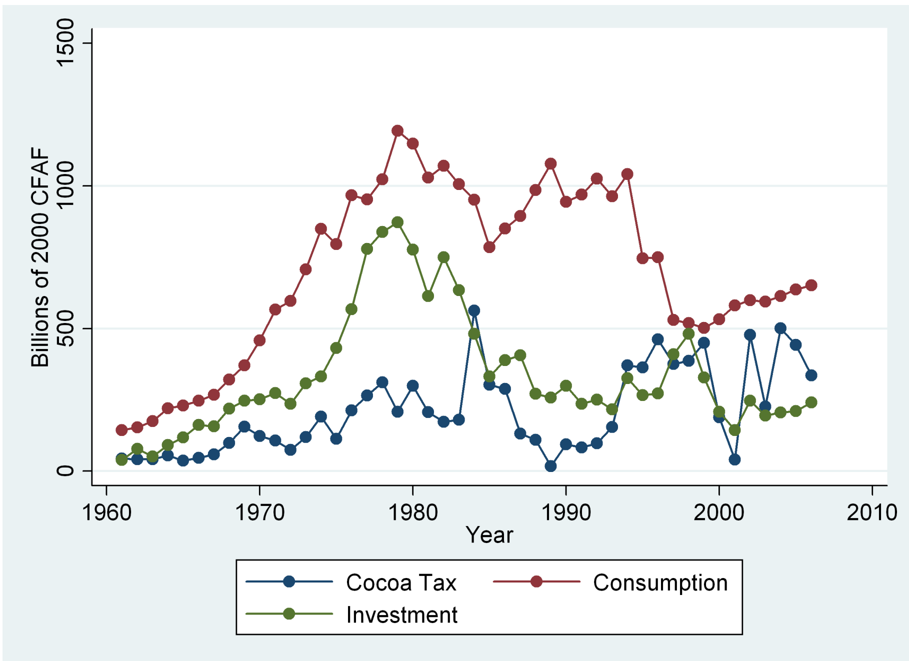


Figure 14: Cocoa Tax and Central Government Spending, Ivory Coast 1961-2006.




Sources: see data appendix for more details. Cocoa tax is the total revenue going to the state, as a result of the difference between the international price and the producer price of cocoa beans. Government spending includes government consumption and investment.