

14/12/2016

Breaking into Tradables: urban form and urban function in a developing city.*

Anthony J. Venables
University of Oxford,
CEPR and International Growth Centre

Abstract

Many cities in developing economies, particularly in Africa, are experiencing ‘urbanisation without industrialisation’. This paper conceptualises this in a framework in which a city can produce non-tradable goods and – if it is sufficiently competitive – also internationally tradable goods, potentially subject to increasing returns to scale. A city is unlikely to produce tradables if it faces high urban and hinterland demand for non-tradables, or high costs of urban infrastructure and construction. The paper shows that, if there are increasing returns in tradable production, there may be multiple equilibria. The same initial conditions can support dichotomous outcomes, with cities either in a low-level (non-tradable only) equilibrium, or diversified in both tradable and non-tradable production. We demonstrate the importance of history and of expectations in determining outcomes. Essentially, a city can be built in a manner which makes it difficult to attract tradable production. This might be a consequence of low (and self-fulfilling expectations) or of history. Predictions of the model are consistent with a number of observed features of African cities.

Keywords: city, urban, economic development, tradable goods, structural transformation.
JEL classification: O14, O18, R1, R3

* Forthcoming, Journal of Urban Economics

Acknowledgements: I gratefully acknowledge support from the Africa Research Program on Spatial Development of Cities at LSE and Oxford, funded by the Multi Donor Trust Fund on Sustainable Urbanization of the World Bank and supported by the UK Department for International Development. Thanks also to referees, an editor, Paul Collier, Avinash Dixit, Vernon Henderson, Patricia Jones, Somik Lall, and seminar participants for helpful comments, and to Sebastian Kriticos for research assistance. Opinions expressed in this paper are those of the author and do not represent those of the World Bank.

Author’s Address:

Tony Venables
Department of Economics
Manor Road
Oxford OX1 3UQ, UK
tony.venables@economics.ox.ac.uk

1. Introduction

Cities in the developing world face the challenge of accommodating a predicted 2.5bn more people by 2050, and the fastest urbanising region will be Sub-Saharan Africa where urban population is predicted to treble to more than 1bn. Urbanisation is occurring at lower per capita income levels than was historically the case, sometimes referred to as ‘urbanisation without growth’ (Fay and Opal, 2000; Jedwab and Vollrath 2015).¹ The performance of developing cities is heterogeneous, and in one area has been sharply dichotomous. Many Asian cities have been able to create jobs in tradable goods sectors and have become internationally competitive, producing large volumes of exports. African cities have failed to do this, and have instead grown on the basis of supplying local and perhaps regional markets. The phrase ‘urbanisation without industrialisation’ has gained currency, and Gollin et al. (2016) point to the prevalence of this phenomenon in resource rich developing countries.

The present paper analyses the factors that shape this aspect of performance and that determine the extent to which developing cities are able to succeed in attracting high productivity tradable goods (or service) sectors, or instead remain specialised in producing non-tradables for local markets. The paper is primarily theoretical, and is based on interactions between ‘urban function’ – the economic activity that takes place in the city – and ‘urban form’ – the way in which the city is constructed and the efficiency with which it operates. To capture this, the model that is developed in the paper has several key ingredients. On the production side we distinguish between non-tradable and tradable sectors of production. The former is likely to encounter diminishing returns because it is limited by the size of local markets, while the latter offers the prospect of increasing returns and agglomeration economies. On the residential side, urban form is captured by a standard urban model in which buildings are durable and density of construction is endogenous. The residential capital stock – and hence the size and density of the city – therefore depend on both past history and expectations of future returns.

Three sets of results are established. First, we establish conditions which are likely to lead to a city being specialised in non-tradables, as opposed to diversified into both non-tradables and tradables. Conditions include the presence of high urban and hinterland demand for non-tradables, and high costs of urban infrastructure and construction. Second, we show how, if there are increasing returns in tradable production, there may be multiple equilibria. The same initial conditions can support dichotomous outcomes, with cities either in a low-level (non-tradable only) equilibrium, or diversified in both tradable and non-tradable production. Third, we demonstrate the importance of history and of expectations in determining outcomes. Essentially, a city can be built in a manner which makes it difficult to attract

¹ Jedwab and Vollrath (2015) discuss alternative reasons for this, including urban technologies, push from agriculture, politics, and natural population increase.

tradable production. This might be a consequence of low (and self-fulfilling expectations) or of history.

While the paper is primarily theoretical, we start by outlining three features of African cities that are illuminated by the model. The first is African cities' failure to create jobs in internationally tradable goods or service sectors. Gollin et al. (2016) investigate this, principally at the national level, establishing the adverse effect of natural resource sectors on manufacturing employment. Natural resource dependence is only part of the story, and there are also significant regional differences. Focusing on Africa, Jones (2016) compares manufacturing shares of GDP in African and non-African economies at different stages of urbanisation. In non-African economies the manufacturing share of GDP rises from 10% to nearly 20% as the urban population share rises to 60%, above which it falls back. In Africa the manufacturing share remains flat (or somewhat falling) at around 10% of GDP through a cross-section of urbanisation rates ranging from 10% to 70%.² A more urban focus can be derived by using spatially disaggregated IPUMS data.³ This is sample data collected at the individual level, with self declared sector of employment. Table 1 reports the share of employment declaring to be in manufacturing in areas classified as urban. The data is presented by city (selected by size within country) and by country; (in India this is urban area by state, with separate city level data unavailable). While manufacturing is not synonymous with tradable goods, the data indicates clearly the extent to which Africa is different from other regions. Simple averages across countries suggest manufacturing shares of urban employment nearly three times greater in Asia than in Africa, and increasing through time, whereas Africa's have declined slightly. Accra appears as the African city with the highest share of employment in manufacturing, at 14.9%, whereas shares in Asian cities rise well in excess of 30%.⁴

The second feature is the problematic 'urban form' of African cities. The weakness of African urban form has numerous manifestations, showing up most obviously in low stocks of key capital assets, including housing and infrastructure. Whilst low stocks of these assets is partly a function of low income (urbanisation at relatively low levels of per capita income) it is also due to market and governance failures. Lack of clarity in land-tenure is widespread, disputes between multiple claimants on land are frequent, and invasions to seize urban land are a problem. Private investment is also discouraged by inappropriately high building and land use regulations, and mortgage finance is hard to obtain (Collier and Venables 2015). Provision of public capital is low, with estimates of the infrastructure gap suggesting that as

² Jones (2016). Manufacturing shares are predicted values from a quadratic function fitted to the cross section data.

³ Integrated Public Use Microdata Series (iPums) available for download at: <https://usa.ipums.org/usa-action/variables/group>

⁴ Ethiopian data dates from 1994; recent manufacturing investments suggest that it could now exceed this.

much as 20% of urban GVA needs to be spent over a period of decades to fill the gap (Foster et al. 2010). These features are captured in our modelling as high costs of building ('costs' including non-monetary obstacles) and of urban transport. Within the model, the consequences are low levels of residential investment and high costs of accessing jobs. This corresponds to the reality of widespread informal settlement, generally single storey and constructed of mud and sheet metal. Some 62% of Africa's urban population lives in slums (UN-Habitat, 2010) and there is often a hodgepodge of land use, with slum areas persisting next to modern developments near city centres. Inefficient land use and sub-optimal stocks of residential and infrastructure capital have costs for the functioning of the city as a whole. There are direct costs as, for example, data for Nairobi suggests that land near the city centre that is currently occupied by slums foregoes around 2/3s of its value compared to neighbouring formal developments (Henderson et al. 2016). There are wider costs of loss of connectivity between economic agents, imposing costs on firms and manifest in some of the longest commuting times in the world.

Table 1: Share of employment in manufacturing: urban areas.

Africa		India			Other Asia			
	1990s	2000s	States	1999	2004		1990s	2000s
Cameroon	-	10.2	Maharashtra	23.4	23.8	Thailand	10.1	12.2
Ethiopia	7.98	-	West Bengal	24.6	27.5	Bangkok	25.2	20.3
Addis Ababa	18.9	-	Delhi	23.1	25.5	Samut Prakan	24.9	48.9
Ghana	14.4	12.6	Tamil Nadu	27.0	30.7	Nonthaburi	18.8	15.0
Accra	17.5	14.9	Karnataka	23.8	21.9	Udon Thani	9.87	6.93
Liberia	-	1.6	Andhra Pradesh	17.7	20.1	Vietnam	16.8	16.9
Malawi	7.19	7.09	Gujarat	25.11	37.2	Ho Chi Minh	37.5	35.1
Blantyre	10.7	11.4	Uttar Pradesh	22.9	26.8	Ha Noi	21.6	18.4
Mali	6.26	6.55	Rajasthan	20.6	23.3	Da Nang	21.4	20.5
Mozambique	5.3	5.1	Bihar & Jharkhand	16.2	12.6	Hai Phong	24.8	24.3
Maputo	9.46	6.8	Punjab	22.7	27.2	Malaysia	20.1	22.4
Rwanda	-	2.58	Kerala	21.2	15.7	Kuala Lumpur	25.2	20.9
Kigali	-	4.92	Haryana	20.3	25.9	Seberang	40.3	44.4
Sierra Leone	-	0.74	Pondicherry	28.5	20.8	Indonesia	9.19	8.00
Sudan	-	6.12	Chhattisgarh	17.6	19.0	Jakarta	18.9	15.5
Tanzania	-	4.96	Orissa	16.8	14.3	Bandung	20.6	23.0
DaresSalaam	-	8.85	Chandigarh	15.7	17.4	Cambodia	5.01	6.92
Uganda	-	4.82	Daman, Diu & Goa	11.5	15.8	Phnom Penh	13.2	26.9
Kampala	-	6.92	Himachal Pradesh	8.86	14.3	Takeo	2.43	6.84
Zambia	6.11	-	Dadra & Nagar Haveli	52.5	27.9	Sihanoukville	8.02	17.5
Lusaka	10.4	-				Battambang	8.25	8.16

The third feature of African cities is the apparent paradox of relatively high nominal wages and prices in many cities, despite their low real income and lack of modern sector employment. This is a feature that, as we will see, can emerge as an outcome in the model. The evidence for it comes from several sources. Jones (2016) reports that firms in African cities pay wages (at official exchange rates) about 15% higher than in non-African cities, conditional on national real GDP pc. Labour costs are estimated at up to 50% higher. Corresponding to this, sales per worker are about 25% higher than in comparable non-African cities, but this is largely in the non-traded sectors and appears to simply reflect higher prices not physical productivity.⁵ High nominal wages are matched by high prices of goods and services. Nakamura et al. (2016) use ICP data to study the cost of living of urban households across countries and find that it is some 20- 30% higher in African countries than in other countries at similar income levels.⁶ Part of this is due to high rents and urban transport costs (respectively 55% and 42% higher than in comparable places elsewhere) although it extends to other commodities. Prices of food and other goods are also relatively high. Henderson and Nigmatulina (2016) show that, across developing country cities, high prices are negatively associated with a measure of connectivity between people in the city.

Our model captures these features, and is developed in a series of stages. Section 2 lays out the ingredients, characterises equilibrium and undertakes comparative statics to establish the determinants of city specialisation, as well as of city size, density and rent levels. While Section 2 maintains the assumption of constant returns to scale in tradable production, Section 3 moves to environments with increasing returns, establishing the possibility of multiple equilibria and the roles of history and expectations in shaping outcomes. Section 4 concludes and offers policy implications.

2. Production and urban form

2.1 The model

The model focuses on a single city which sells goods within and outside the city, and which is able to draw labour from the wider economy. Labour is the only input to production, and land is used for housing urban workers. Analysis is based around labour demand and labour supply. The former gives the relationship between the wage and the level of employment, depending on productive activity in the city ('urban function'). The latter is the relationship

⁵ These findings on labour costs and price levels are consistent with earlier work by Gelb et al. (2013, 2015), although this work does not have an urban focus.

⁶ Nakamura et al. also report findings from the Economist Intelligence Unit's Worldwide Cost of Living Survey. While this is produced principally for expatriates its findings are consistent, indicating African cities are 30% more costly for households than cities in low- and middle-income countries elsewhere.

between the wage and city population, this depending on migration and on the cost of living in the city, including costs of construction and of commuting ('urban form').

Production and labour demand: Labour is demanded by the production side of the city economy in which there are (potentially) two produced goods, tradables and non-tradables, with employment levels L_T and L_N giving total city employment $L \equiv L_T + L_N$. The wage is w , the same in both sectors, and since labour is the only input the value of output produced in the city is wL .⁷ To derive the city's labour demand function we look first at demand for labour in non-tradable production, and then in tradables.

Non-tradable goods meet demands from the local market and have price p_N , determined endogenously by supply and demand.⁸ They are produced under constant returns to scale so, choosing units such that one unit of labour produces one unit of output, $p_N = w$ and the value of supply is wL_N . The value of demand for non-tradables is $(1-\theta)wL + p_N h(p_N)$. In the first term, wL is the city wage bill and this is spent on a composite good which is a Cobb-Douglas aggregate of tradables with share θ , and non-tradables with share $(1-\theta)$. As we will see below, this spending takes different forms – final consumption, commuting and construction costs – but all demand the same composite, so the city's income generates demand for non-tradables $(1-\theta)wL$. The second term, $p_N h(p_N)$, is spending on non-tradables from income generated outside the city. Such spending consists of several elements. One is spending from transfer payments to the city, such as natural resource revenues, taxes, or foreign aid. Another is hinterland spending on non-tradables produced in the city. We refer to $h(p_N)$ as the hinterland demand function, and take it as exogenous and decreasing in price; it is shifted upwards by higher demand from these extra-city sources. Using $p_N = w$ and $L_N = L - L_T$, the equality of supply and demand for non-tradables is $(L - L_T)w = (1 - \theta)wL + wh(w)$. Rearranging,

$$w = h^{-1}(\theta L - L_T). \quad (1)$$

This is the wage at which net urban supply of non-tradables equals hinterland demand, where $h^{-1}(\cdot)$ is the inverse hinterland demand curve. Crucially, its downwards slope captures the fact that expanding urban employment in non-tradables reduces the wage paid, since increased supply reduces the price of non-tradables.

⁷ These assumptions imply full employment and an integrated labour market. It would be possible to add labour market imperfections (for example, part of non-tradable production being undertaken by an informal sector), but this is inessential for the arguments being made.

⁸ The definition of non-tradables is necessarily quite elastic. The crucial distinction is the extent to which the price depends on supply from the city, or is set on a wider regional or international market.

In contrast, tradable goods face perfectly elastic world demands, have fixed world price, and will be taken as numeraire. Labour productivity in the tradable sector is $a(L_T)$, and this is the wage offered in the sector. We assume that productivity is either constant or increasing in L_T , increasing returns arising because of agglomeration economies external to the firm but internal to the tradable sector. The sector operates if labour productivity is greater than or equal to the market wage, w , so the following relationships must hold,

$$w = a(L_T), L_T \geq 0; \quad w > a(L_T), L_T = 0. \quad (2)$$

Together, Eqns. (1) and (2) implicitly define the city's (inverse) labour demand schedule $w^D(L)$, i.e. they give the value of w at which urban employment (in non-tradables and tradables) fully employs a labour force of size L . This labour demand curve generally has a kink in it at the 'trigger wage', $a_0 \equiv a(0)$, at which tradable production commences. We summarise it as follows:

If the tradable sector is inactive, then $L_T = 0$ and, from (1),

$$w^D(L : L_T = 0) = h^{-1}(\theta L). \quad (3a)$$

If the tradable sector is active, then $L_T > 0$ and, from (1) and (2),

$$w^D(L : L_T > 0) = a(L_T), \text{ with } L_T \text{ solving } a(L_T) = h^{-1}(\theta L - L_T). \quad (3b)$$

The wage offered is the maximum,

$$w^D(L) = \max[w^D(L : L_T = 0), w^D(L : L_T > 0)]. \quad (3c)$$

This is illustrated by the bold curve in Fig. 1 which has city employment (= population) on the horizontal axis and the wage on the vertical. The figure is drawn with constant returns to scale in tradable production so labour productivity in tradables is a constant $a(L_T) = a_0$. On the downward sloping segment the entire labour force is employed in the non-tradable sector, $w^D(L : L_T = 0) = h^{-1}(\theta L)$, downward sloping as more employment increases supply of the non-traded good, reducing price and wages. If the ensuing wage is less than a_0 then tradable production is profitable, giving the horizontal segment, $w^D(L : L_T > 0) = a_0$. The dividing line is where the urban population is $L = h(a_0)/\theta$.

Urban costs and labour supply: Labour supply comes from the location decisions of mobile workers. It depends on utility outside the city, set as exogenous constant u_0 , and utility inside. This depends on the wage, the prices of goods consumed, and any further costs associated with urban living. These include direct utility costs (and benefits) from crowding,

congestion, and provision of public services. And financial costs, arising from land and house prices and from commuting costs.

Our modelling of urban costs is in the tradition Alonso (1964) and others.⁹ Workers are employed in the central business district (CBD) and incur commuting costs. The city is linear, and x denotes distance from the CBD.¹⁰ Each worker occupies a unit of housing, and housing density (hence the amount of land occupied per unit housing) is chosen by profit maximising developers. The utility of a worker living at distance x from the CBD is $v(x)$,

$$v(x) \equiv [w - p(x) - xt w^{1-\theta}] / w^{1-\theta} = (w - p(x)) w^{\theta-1} - xt. \quad (4)$$

The term in square brackets is the wage net of housing and commuting costs. The cost of housing at distance x is $p(x)$ and commuting incurs t units of the composite good per unit distance, so a worker living at distance x from the CBD pays commuting costs $xt w^{1-\theta}$, where $w^{1-\theta}$ is the price index of the composite good.¹¹ Income net of housing and commuting costs (i.e. the term in square brackets) is spent on the composite good, so utility is derived by deflating by the price index.¹²

Workers are perfectly mobile, choosing between living outside the city with utility u_0 , or at locations x in the city. For all occupied urban locations, it must therefore be the case that $v(x) = u_0$, implying that house prices at each distance, $p(x)$, satisfy the indifference condition

$$p(x) = w - (u_0 + xt) w^{1-\theta}. \quad (5)$$

The supply of housing depends on residential construction decisions, taken to maximise land rent at each point. The number of housing units – or density – built at x is $N(x)$, and rent, $r(x)$, is revenue from housing minus construction costs, i.e.

$$r(x) = p(x)N(x) - cN(x)^\gamma w^{1-\theta}, \quad \gamma > 1. \quad (6)$$

Construction costs per unit land are $cN(x)^\gamma w^{1-\theta}$, an increasing and convex function of density (e.g. the cost of building taller).¹³ We assume that this relationship is iso-elastic with

⁹ See Duranton and Puga (2015) for full exposition and analysis of the Alonso-Muth-Mills urban model.

¹⁰ To minimise notation we work with a linear city with a single spoke from the CBD.

¹¹ As noted above, the composite is Cobb-Douglas with tradable share θ , non-tradable share $1 - \theta$.

Since tradables are the numeraire and the price of non-tradables is w , the price index is $w^{1-\theta}$.

¹² If there are transfer payments spent in the city (resource rents, taxes, foreign aid) we assume that, while they create demand for non-tradables, they do not provide utility for urban residents.

¹³ See Henderson, Tanner and Venables (2016) for a richer modelling of construction technologies that differentiates between formal and informal housing. Qualitative results of the present paper would be unchanged with this more general formulation.

parameter γ , and that costs are incurred in units of the composite good, i.e. have price $w^{1-\theta}$. Developers choose density to maximise rent, giving first order condition and maximised rent,

$$N(x) = \left(p(x)w^{\theta-1} / c\gamma \right)^{1/(\gamma-1)}, \quad (7)$$

$$r^*(x) = (1-1/\gamma)p(x)N(x) = (1-1/\gamma)p(x)^{\gamma/(\gamma-1)} \left(w^{\theta-1} / c\gamma \right)^{1/(\gamma-1)}. \quad (6')$$

The edge of the city is at distance \tilde{x} where rent equals the exogenous outside rent r_0 , i.e.

$$r^*(\tilde{x}) = r_0. \quad (8)$$

Housing capacity and total city population up to this edge are

$$L = \int_0^{\tilde{x}} N(x)dx. \quad (9)$$

Eqns. (5) – (9) give values of three variables that vary with distance, $N(x)$, $p(x)$, and $r^*(x)$ and two scalars, L and \tilde{x} , as functions of parameters and the wage. They implicitly define the residential structure of the city, its built density, house prices, land rents, and total population and city edge. They also define the labour supply function – the relationship between w and L . An explicit form of this relationship is useful for expositional purposes and can be derived with two simplifying assumptions.¹⁴ The first is that outside rent is zero, $r^*(\tilde{x}) = 0$; this implies that the edge of the city has $p(\tilde{x}) = N(\tilde{x}) = 0$ (from (6') and (7)), i.e. that the city converges to zero density at its edge. It follows from (5) that the city edge is given by:

$$r^*(\tilde{x}) = 0 \text{ implies } p(\tilde{x}) = 0 \text{ and hence } \tilde{x} = (w^\theta - u_0)/t. \quad (10)$$

Second, we assume that construction costs increase with the square of density, $\gamma = 2$, giving population

$$L = \int_0^{\tilde{x}} N(x)dx = \int_0^{\tilde{x}} \left(w^\theta - (u_0 + xt) \right) / 2c dx = \left(w^\theta - u_0 \right)^2 / 4tc. \quad (11)$$

This uses (7), (5) and (10) in (9); the appendix gives the case for general values of γ .

Inverting, the urban wage required to attract and accommodate population L is

$$w^S(L) = \left[u_0 + 2(tcL)^{1/2} \right]^{1/\theta}. \quad (12)$$

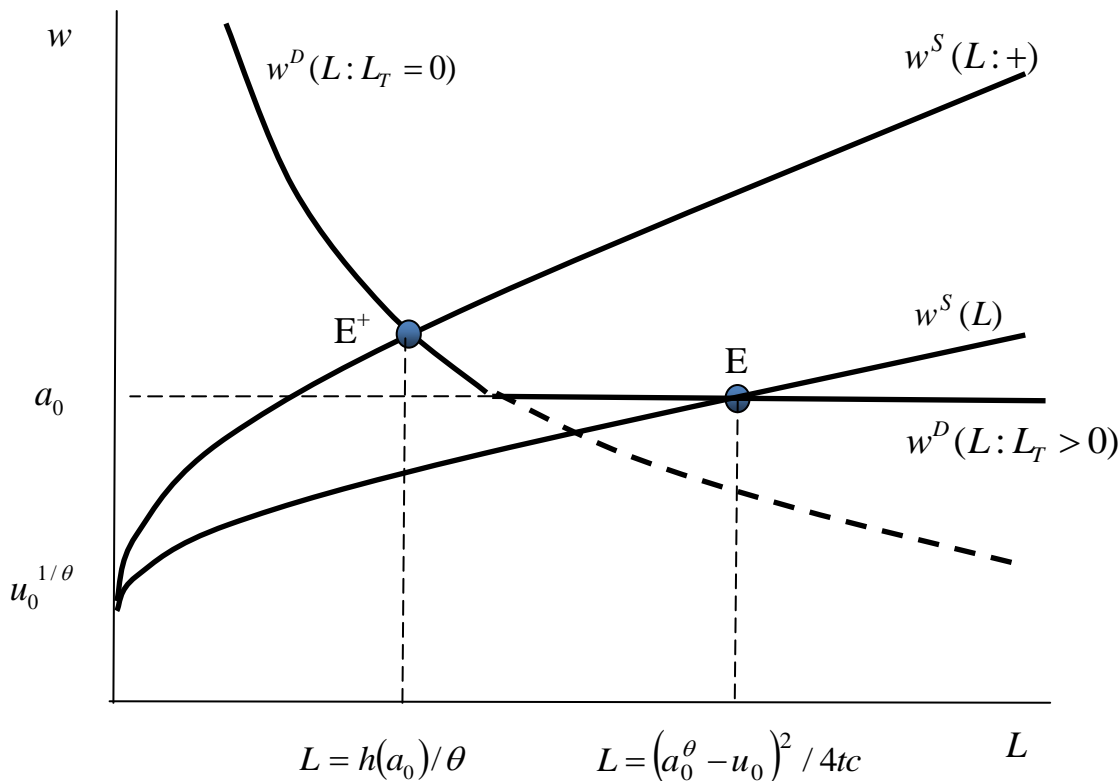
This is the inverse labour supply curve, $w^S(L)$ illustrated by the upwards sloping curves on Fig. 1, the upper one drawn for a city with higher urban costs.¹⁵

¹⁴ These assumptions are relaxed in simulations later in the paper.

¹⁵ This and subsequent Figs. are derived by numerical simulation, parameters given in the appendix.

Several remarks are in order. First, concavity is a general property of this supply function. Intuitively, if the wage (and hence land rent) is low the city is built at low density, so accommodates a small population.¹⁶ Higher wages increase population through two margins: the city becomes larger and is built at higher density, the interaction between the two giving concavity. Second, this supply curve is derived from the fundamentals of commuting costs, construction, and land prices, but these are only part of a more general urban cost relationship. If other factors – congestion, loss of amenity – increase with city size, this too will create the upward sloping relationship as a larger city requires higher wages as compensating differential to offset these costs (see e.g. Duranton’s 2008 discussion of the ‘cost of living curve’).

Figure 1: Urban equilibrium with constant returns to scale



2.2 Equilibrium with constant returns in tradable production

The urban labour demand and supply curves, Eqs. (3) and (12), give equilibrium values of the wage, city population and sectoral employment, as illustrated by points E and E⁺ on Fig. 1.

¹⁶ If $r_0 = 0$ the curve is vertical at $L = 0$, as illustrated.

Point E^+ gives the equilibrium if urban costs are high, $w^S(L: +)$, in which case the wage exceeds a_0 and city produces only non-tradables. At lower costs equilibrium is at E , with both sectors active, wage rate $w = a_0$, and hence city population and employment in tradables respectively, $L = (a_0^\theta - u_0)^2 / 4tc$ and $L_T = \theta L - h(a_0)$ (from Eqs. (11) and (3b)). The message is simple. Both cases have the same real wage, u_0 . However, the equilibrium without tradable production, E^+ , has higher nominal wages and higher non-tradable prices. High urban costs can be passed on to consumers of non-tradables, but cannot be passed on in tradable sector, meaning that the city is uncompetitive in the production of these goods.

What factors shape the likelihood of each of these outcomes occurring? The dividing line between cases is where wages and employment levels satisfy $a_0 = w^S(L) = w^D(L: L_T = 0)$, so the city is specialised in non-tradables if $a_0 < [u_0 + 2(cth(a_0) / \theta)^{1/2}]^{1/\theta}$ (using Eqs. 12 and 3a). This is more likely if, from the left-hand side of the inequality, a_0 is small, i.e. the productivity of workers in the tradable sector is low. This could arise because of low productivity in the sector, or transport barriers imposing additional costs on exporting. On the right hand side of the inequality four parameters enter multiplicatively, $c, t, H, 1/\theta$, where H represents a multiplicative shift parameter in the hinterland demand function.¹⁷

Before discussing the impact of these parameters it is useful to establish their impact on other aspects of the equilibrium, in particular the area of the city, its average density, and urban land rents. Area is captured by \tilde{x} , Eqn. (10), and average density is L / \tilde{x} . Total rents, R , are derived by integrating over rents at each point in the city (Eqn. 6'),

$$R = (1 - 1/\gamma) \int_0^{\tilde{x}} p(x)N(x)dx = w^{1-\theta} (w^\theta - u_0)^3 / 12tc \quad (13)$$

where the last equation uses $\gamma = 2$ (see appendix for general form). The effect of parameters on endogenous variables is found by log-linearising the equilibrium conditions, and full expressions are given in the appendix. Here we simply note the comparative static signs, with Table 1 giving the sign of an increase in each of the parameters on specialisation and other endogenous variables.

The city is more likely to produce non-tradables only, $L_T = 0$, the higher are c, t, H , and $1/\theta$. The last two of these capture high demand for non-tradables, either from the hinterland or from a high share of city income being spent on non-tradables. The first two capture high urban costs, of construction and commuting respectively. While demand and costs have the same qualitative impact on city specialisation and on nominal wages, they have opposite

¹⁷ I.e., the demand function becomes $h(p) = H\bar{h}(p)$

effects on city population. If $L_T = 0$, higher demand is associated with larger population, and higher costs with a smaller population. Higher values of each of these parameters increase nominal wages and the ratio of rent to wage bill. Notice however that the two cost parameters have opposite implications for city density and geographic size. High construction costs reduce density, and city size is greater despite lower population. High transport cost reduce size, raising rents and hence city density

If the city is diversified, $L_T > 0$, then nominal wages are set by productivity in tradables, a_0 . Demand for non-tradables has no effect on city population – high demand is accommodated by lower tradable production. High costs do however reduce city population, because of their direct effect on density of building (for construction costs) or the extent of the city (for commuting costs).

Table 1: Comparative statics.

	Likelihood $L_T = 0$	Population* \hat{L}	Wage* \hat{w}	Size* \hat{x}	Density* $\hat{L} - \hat{x}$	Rent* $\hat{R} - \hat{L},$ $\hat{R} - \hat{L} - \hat{w}$
		$L_T = 0: L_T > 0$	$L_T = 0: L_T > 0$	$L_T = 0: L_T > 0$	$L_T = 0: L_T > 0$	$L_T = 0: L_T > 0$
$\Delta H > 0$	+	+ 0	+ 0	+ 0	+ 0	+ 0
$\Delta c > 0$	+	- -	+ 0	+ 0	- -	+ 0
$\Delta t > 0$	+	- -	+ 0	- -	+ 0	+ 0
$\Delta a_0 > 0$	-					

* ^ denotes a proportionate change in a variable. The first column gives comparative statics when equilibrium has $L_T = 0$ and the second when $L_T > 0$.

In summary, characteristics of the city equilibrium are determined by multiple factors, and the model gives clear predictions about the mapping between parameters and outcome variables. Of course, the model parameters are themselves just summaries of complex realities, some elements of which can be changed by policy, others not. Thus, building costs include the ad valorem equivalent of the multiple obstacles to private construction that were outlined in the introduction. Commuting costs are to do with transport investment, and also reflect income levels – cities in which most people have to walk to work. Similarly, demand for non-tradables may be driven by the distribution of tax revenues, foreign aid or natural resource revenues which may themselves be the outcome of a political economy of urban bias (Bates 1981, Lipton 1977, Ades and Glaeser 1991). Hinterland demand for non-tradables is also a function of the income and economic geography of the region in which the city is located; for example, demand will be high if there are no nearby cities offering alternative sources of supply.

3. Increasing returns, expectations and coordination failure

We now open up two central features of the model. The first is increasing returns to scale – agglomeration economies – in tradable production, this creating the possibility of multiple equilibria.¹⁸ We show that it is possible that there is a low equilibrium in which the city produces tradables only, and also a high equilibrium in which the city is active in both sectors. The possibility of being trapped in the low-equilibrium arises because of coordination failure.¹⁹ In the simplest case (section 3.1) this is an inter-firm coordination failure; potential producers of tradables do not coordinate to internalise the external economies of scale associated with agglomeration. However, given endogenous choice of the way the city is built – its size and density – the coordination failure goes deeper. Section 3.2 turns to the role of sunk costs and expectations in shaping construction decisions, and shows how (self-fulfilling) expectations may be such that the city is constructed in a way that locks it into the low equilibrium. It is possible that – even if the inter-firm coordination failure between producers of tradables were somehow resolved – the built urban form is incompatible with expanding into tradable goods production.

3.1 Increasing returns and multiple equilibria.

Suppose that tradable production is subject to agglomeration economies so that productivity in tradable production is increasing with the size of the sector. The left-hand segment of the labour demand curve, $w^D(L: L_T = 0)$ is unchanged, and the segment with tradable sector active, $w^D(L: L_T > 0)$, becomes upwards sloping. Fig. 2 illustrates for the case in which productivity is linear in tradable employment between lower and upper bounds, a_0, a_m , so $a(L_T) = \min[a_0 + \alpha L_T, a_m]$, $\alpha > 0$, $a_0 < a_m$.

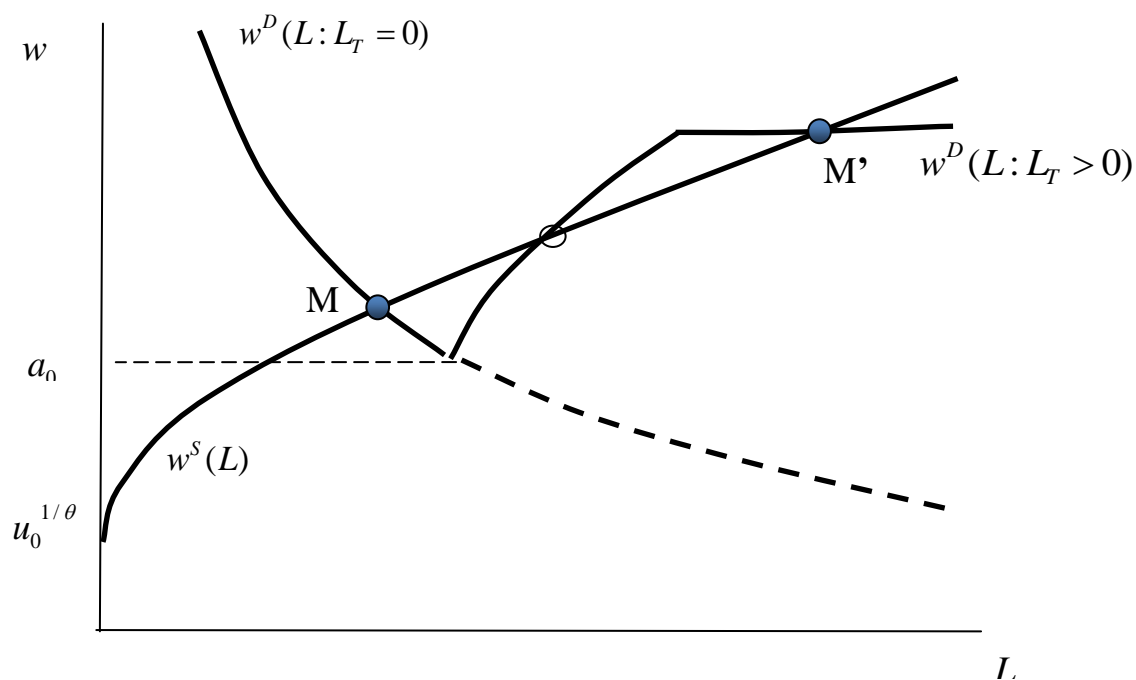
As illustrated, returns to scale are strong enough for labour demand to have three intersections with labour supply. Points M and M' are stable equilibria, while the intermediate intersection is unstable (under a dynamic in which the tradable sector expands or contracts according to whether profits are positive or negative). At the lower point, M, the city is specialised in non-tradable production and wages are above the trigger point, a_0 , at which tradable production commences. Entry of a small mass of tradable produces is not profitable; this point is an equilibrium if coordinated entry of a sufficiently large mass of tradable producers (who would reap the benefits of agglomeration economies) is not possible. At the upper equilibrium, M', the tradable sector is active and is large enough for

¹⁸ For discussion of agglomeration economies in the context of developing economies see World Bank (2009) and for evidence see Chauvin et al. (2016).

¹⁹ As in, for example Murphy et al. (1989), Henderson and Venables (2009).

agglomeration economies to have cut in, raising productivity. This is associated with larger city size and employment and with higher nominal wages (although real wages are held to u_0 by migration and labour supply).

Figure 2: Increasing returns and multiple equilibria.



What are the conditions that support this configuration? The appendix gives conditions on parameters, and here we just make three remarks. First, point M exists if

$a_0 < w^S(L) = w^D(L: L_T = 0)$, exactly as discussed in section 2.2. Second, city size must (we assume) be bounded above, so it must be the case that as L becomes very large so $w^S(L) > w^D(L: L_T > 0)$, as illustrated. Third, a necessary condition for multiple

intersections is that economies of scale in tradables are large enough that, in some range, the slope of the labour demand curve exceeds that of labour supply. Recall that, along $w^D(L: L_T > 0)$, L_T solves $a(L_T) = h^{-1}(\theta L - L_T)$, Eqn. (3b). The slope is therefore

$$\frac{dw^D}{dL} = a'(L_T) \frac{dL_T}{dL} = \frac{a' \theta}{1 + h' a'}$$

The slope is steeper the greater are marginal returns to scale, $a'(L_T)$, adjusted for the fact that as L increases some extra jobs are created in the non-tradable sector so $dL_T/dL < 1$. This is larger the fewer non-tradable jobs are created by demand from additional urban population (high θ), and the more that the increasing wage

chokes off hinterland demand for non-tradables, (more elastic hinterland demand for non-tradables).

There are two further comments. First it can be shown that aggregate welfare is higher in the equilibrium M' than M . Comparing the two, workers' utility is, by construction, unchanged. Total city rent is higher, since along the labour supply curve it is increasing in wage and city size. There is a third element of welfare change, which is that of 'hinterland' consumers of the non-tradable good, whose welfare depends on the price of non-tradables. We show in the appendix that the sum of these elements is increasing in the wage and hence, comparing points on the labour supply curve, welfare is higher at M' than M . However, the distributional impact, given constant real wages in the city, is no change for workers, gain for urban landowners, and loss for outsiders consuming non-traded goods produced in the city.

Second, we suggested in the introduction that cities' experience seemed to be dichotomous, some trapped in non-tradable production, while others have been successful in attracting tradables and have grown large clusters of tradable activities. The possibility of multiple equilibria captures this dichotomous response, showing how small initial differences can lead to quite different outcomes. However, the low equilibrium is supported by a simple inter-firm coordination failure between tradable sector producers. We now enrich this, adding interaction between the tradable sector and the way in which the city is constructed.

3.2 Expectations, sunk costs and construction

Buildings are long lived and most of the costs incurred in constructing them are sunk. This has two implications. One is that construction decisions involve forward looking expectations, and the other is that a historical legacy of urban form shapes the present urban equilibrium. It is therefore possible that cities are 'locked-in' to outcomes not simply because of coordination failure between firms, as we saw in the previous section, but because of a coordination failure spanning time periods and the actions of builders and firms. To capture this a time dimension must be added to the model, and this is done simply by adding a second time period. We will show how low expectations mean that the city may be constructed at a scale and density that precludes attracting tradable sector production. High expectations can be self-fulfilling as they support construction at greater scale and density, an urban form that lowers costs, attracts tradables, and generates an outcome with higher welfare.

The extension to two periods is straightforward. Production and labour demand are as above, but residential building is durable. Buildings constructed in the first period last into the second, meaning that first period construction decisions depend on prices in period 1 and on expected prices in period 2. Time periods are indicated by subscripts 1, 2 and expectations of future values by superscript E , so period 1 expectations of period 2 prices of houses and

labour are, $p_2^E(x), w_2^E$. Worker mobility means that house prices are expected to satisfy indifference condition (5) in both periods, so

$$p_1(x) = w_1 - (u_0 + xt)w_1^{1-\theta}, \quad p_2^E(x) = w_2^E - (u_0 + xt)(w_2^E)^{1-\theta}. \quad (5a)$$

The expected present value of a house built at x in period 1 is $\delta p_1(x) + (1-\delta)p_2^E(x)$, where δ and $1-\delta$ are weights attached to each period, and it is this that guides period 1 construction decisions.²⁰ The expected present value of construction at density $N_1(x)$ on land at x in period 1, and the optimal choice of density are

$$r_1^E(x) = \{\delta p_1(x) + (1-\delta)p_2^E(x)\}N_1(x) - cN_1(x)^\gamma w_1^{1-\theta}, \quad (6a)$$

$$N_1(x) = \left[\{\delta p_1(x) + (1-\delta)p_2^E(x)\} w_1^{\theta-1} / c\gamma \right]^{1/(\gamma-1)}, \quad (7a)$$

The edge of the city is at distance \tilde{x}_1 where rent equals the exogenous outside rent r_0 , and city population follows, so

$$r_1^*(\tilde{x}_1) = r_0, \quad L_1 = \int_0^{\tilde{x}_1} N_1(x) dx. \quad (8a), (9a)$$

These equations are analogous to Eqs (5) – (9), except that future expected prices influence construction decisions.

In the second and final period house prices satisfy the indifference condition

$$p_2(x) = w_2 - (u_0 + xt)w_2^{1-\theta}. \quad (5b)$$

Building takes place on land that is not previously developed so, for $x > \tilde{x}_1$, rent, construction decisions and population are given by

$$r_2(x) = p_2(x)N_2(x) - cN_2(x)^\gamma w_2^{1-\theta}, \quad (6b)$$

$$N_2(x) = \left[p_2(x)w_2^{\theta-1} / c\gamma \right]^{1/(\gamma-1)}, \quad (7b)$$

$$\text{and } r_2^*(\tilde{x}_2) = r_0, \quad L_2 = L_1 + \int_{\tilde{x}_1}^{\tilde{x}_2} N(x) dx. \quad (8b), (9b)$$

Once again, these are analogous to Eqs. (5)-(9), except that they now apply only to land that is not previously developed, $x > \tilde{x}_1$, and the housing stock (and hence population) includes that inherited from period 1.

²⁰ We continue to use the word house ‘price’ rather than rent, although payments are made in both periods.

The full equilibrium comes from the labour demand functions (as previously) and from noting that Eqs. (5a) – (9a) and (5b) – (9b) respectively define first and second period inverse labour supply functions, $w_1^S(L_1 : w_2^E)$ and $w_2^S(L_2 : L_1)$. Period 1 labour supply now depends also on expectations about period 2: and period 2 labour supply depends also on inherited residential stock and hence population L_1 . It is not generally possible to derive explicit expressions for these supply functions, so we proceed by demonstrating some special cases, and then going to numerical simulation.

First, notice that, by construction, the two equilibria of sub-section 3.1 are stationary perfect foresight equilibria of the two period model. Thus, if $w_1 = w_2^E = w_2$ so that

$p_1(x) = p_2^E(x) = p_2(x)$ then $L_2 = L_1$ and the model gives stationary outcomes identical to those in section 3.1.²¹ This is illustrated on Fig. 3, with points M, M' and labour supply curve $w^S(L)$ exactly as in Fig. 2. We refer to these equilibria as $\{M, M\}$ and $\{M', M'\}$, the two elements in brackets representing the outcome in the two time periods.

Second, suppose that the first period equilibrium is at M (so the city was built with stationary expectations, $w_2^E = w_1$). What effect does this have on the set of opportunities attainable in period 2 – or more formally, what does $w_2^S(L_2 : L_1)$ look like? To answer this we assume that $r^*(\tilde{x}) = 0$ and $\gamma = 2$, in which case the edge of the city in period 1 is $\tilde{x}_1 = (w_1^\theta - u_0)/t$ (Eqn. 10). If the same assumptions hold in period 2 then (5b) – (9b) give $\tilde{x}_2 = (w_2^\theta - u_0)/t$. This enables the integral giving city population in period 2 to be evaluated as

$$L_2 = L_1 + \int_{\tilde{x}_1}^{\tilde{x}_2} N_2(x)dx = L_1 + \int_{\tilde{x}_1}^{\tilde{x}_2} [w_2^\theta - (u_0 + xt)]/2c \cdot dx = L_1 + (w_2^\theta - w_1^\theta)^2 / 4tc$$

Rearranging, the inverse labour supply curve is

$$w_2^S(L_2 : L_1) = [w_1^\theta + 2(tc\{L_2 - L_1\})^{1/2}]^{1/\theta}. \quad (14)$$

This is the dashed line $w_2^S(L_2 : L_1)$ through point M on Fig 3. The steepness of this curve close to M is apparent from Eq. (14) and reflects that fact that, if w_2 is only slightly greater than w_1 , then the only new building undertaken will be close to the edge of the period 1 city, and low price, rent, and density, therefore accommodating few workers.²² Quite generally, $w_2^S(L_2 : L_1)$ must lie on or above $w^S(L)$. The latter optimises construction for each value of

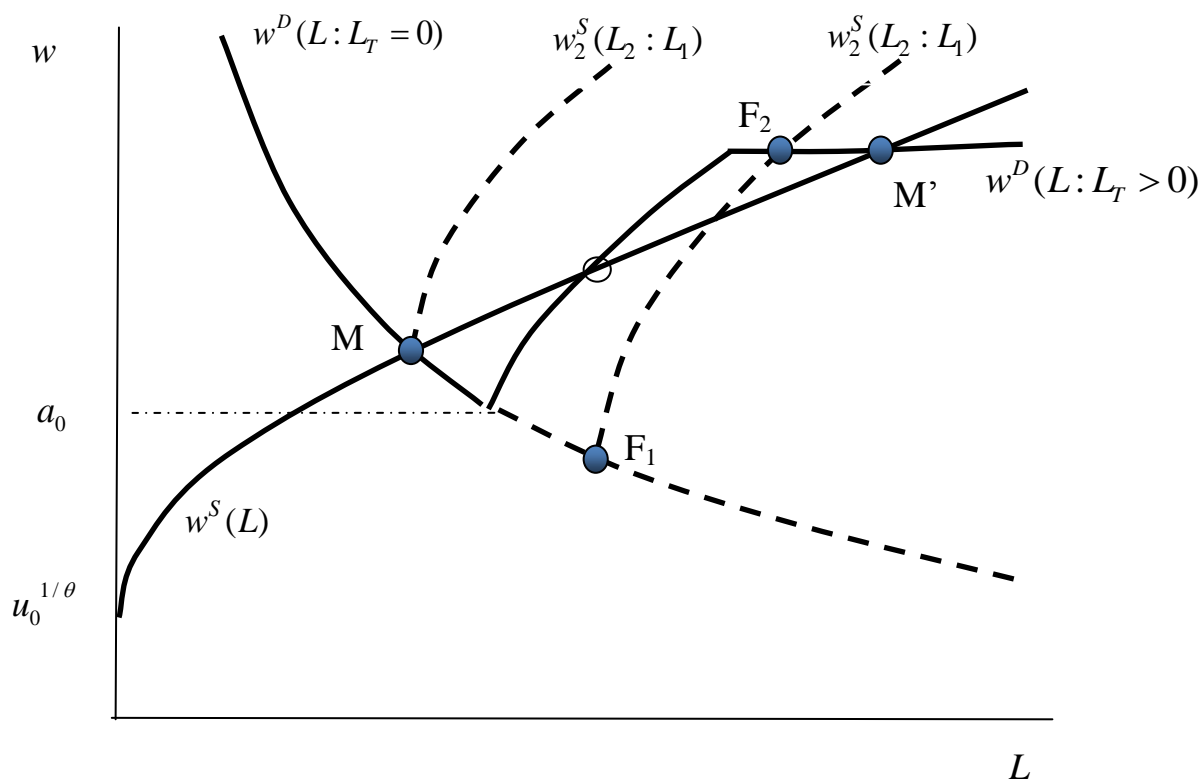
²¹ With $w_1 = w_2^E = w_2$ and $p_1(x) = p_2^E(x) = p_2(x)$ Eqns. 5a-8a, 5b-8b and 5-8 are identical. It follows that $\tilde{x}_1 = \tilde{x}_2$ and $L_1 = L_2$.

²² See Eqn. 12 and following discussion for comparison.

L ; the former inherits a housing stock optimised for first period value of L_1 , and then builds beyond. Some of the city building stock is therefore optimised for the relatively small population at M , a building stock that is not optimal for the larger population.

The consequences are clear, from the fact that – as illustrated – labour supply $w_2^S(L_2 : L_1)$ does not intersect with labour demand except at M . Thus, even if inter-firm coordination failure were to be resolved, the high equilibrium is unattainable. Starting from period 1 equilibrium at M , the only equilibrium is the stationary one, remaining at M .

Figure 3: Multiple equilibria with sunk costs



The previous case worked with period 1 equilibrium at M – as would be the case with stationary (and self-fulfilling) expectations. What if expectations are more positive, expecting tradable production in period 2? We have already seen that $\{M', M'\}$ is an equilibrium, but it is more interesting to study the transition from non-tradable specialisation to tradables. We therefore *assume* that tradable production is technologically impossible in the first period (e.g. city tradable productivity is very low), but becomes possible in the second period. The first period equilibrium therefore lies on $w^D(L : L_T = 0)$ but, if

expectations are more positive, will not be at point M. If tradable production is expected in period 2 then $w_2^E = w_2 = a_M$. This raises the returns to building in period 1 (Eqns. 5a-7a), so there is more and denser period 1 building, accommodating a larger city population and moving the first period equilibrium along $w^D(L: L_T = 0)$ to point F_1 . The second period labour supply curve through this point, $w_2^S(L_2: L_1)$, does not have a closed form solution, but is qualitatively similar to (14), as illustrated on Fig. 3, and gives the perfect foresight equilibrium pair $\{F_1, F_2\}$.

The essential difference between the two equilibria, $\{M, M\}$ and $\{F_1, F_2\}$ lies in expectations. Both start with $L_T = 0$, but in $\{F_1, F_2\}$ expectations are ‘optimistic’, and perfectly foresee the second period outcome with tradable production. This gives higher present value rents, more first period construction, and hence greater first period housing capacity which reduces house prices and first period nominal wages. Importantly if point F_1 lies below a_0 , then the inter-firm coordination problem is resolved. City population and the supply of non-tradables is large enough at F_1 for the wage (absent tradable production) to be less than the trigger wage, ensuring that tradable production will take place in the second period. Once again, the built structure of the city overcomes a potential inter-firm coordination failure,

3.3 The urban profile

Finally, we illustrate results by developing a numerical example. This puts numbers on the arguments above and demonstrates the existence of the equilibria discussed. It also enables us to relax assumptions (in particular letting outside land rent be positive), establish welfare effects and undertake comparative statics. Perhaps most importantly, it generates illustrations of the city profile – the way in which land rent, house prices and building density varies across the city and between equilibria. The simulations use the same parameter values as underlie earlier figures, except that $r_0 > 0$. Values are given in the appendix; to scale results we note that outside utility is set at unity and the two time-periods have equal weight, $\delta = 1/2$.

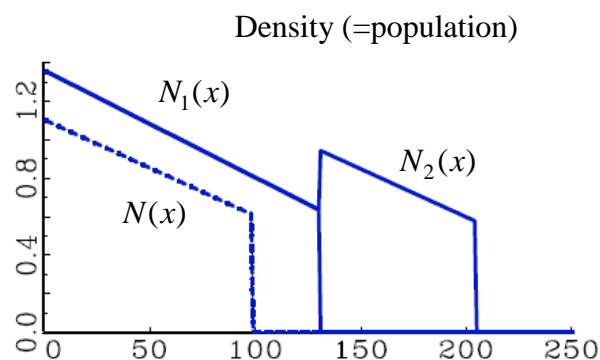
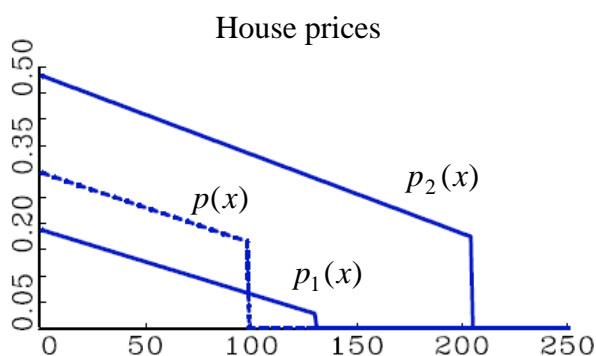
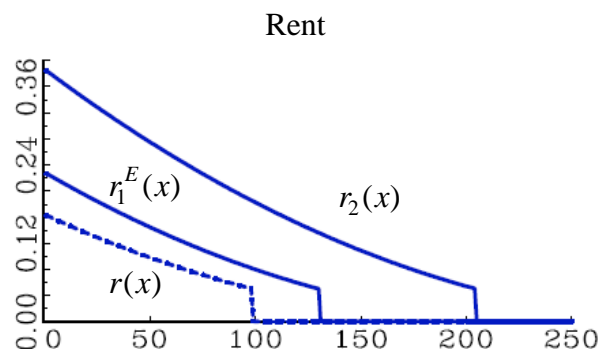
Outcomes are illustrated in the panels of Fig. 4. Values of city-wide variables are given in the table and intra-city profiles of rent, house prices, and density are given in the three plots which have distance from the CBD on the x -axis. The table reports outcomes for three equilibria: the stationary low and high equilibria, $\{M, M\}$ and $\{M', M'\}$ respectively, and the equilibrium $\{F_1, F_2\}$. For clarity, the plots just give $\{M, M\}$ (dashed lines) and $\{F_1, F_2\}$, for which the two time periods are subscripted.

Equilibrium $\{M, M\}$ has wage in each period $w = 1.65$, with a 65% premium over outside utility going to meet the urban cost of living, including commuting and housing costs. This wage is consistent with no tradable production occurring if it is greater than the trigger wage, i.e. $a_0 < 1.65$. The intra-city profiles are given by the dashed lines on the figure. The city

edge is the kink in the rent function (up to which $r(x) \geq r_0$ and beyond which land is undeveloped). House prices fall linearly, since commuting costs are assumed linear in distance. Density and rents decline with distance, as expected.

Figure 4: City profiles

{M,M}	{M',M'}	{F ₁ , F ₂ }	
$w_1 = w_2 = 1.65$	$w_1 = w_2 = 2$	$w_1 = 1.42$	$w_2 = 2$
$L_1 = L_2 = 84$	$L_1 = L_2 = 221$	$L_1 = 129$	$L_2 = 186$
$R_1 = R_2 = 5.1$	$R_1 = R_2 = 29$	$R_1 = 10.1$	$R_2 = 29$
<i>Hinterland</i> $CS_1 = CS_2 = 28$	<i>Hinterland</i> $CS_1 = CS_2 = 19$	<i>H-land</i> $CS_1 = 37$	<i>H-land</i> $CS_2 = 19$
-	$CV_1 = CV_2 = 13$	$CV_1 = 16$	$CV_2 = 13$



The solid lines, subscripted by period, are equilibrium $\{F_1, F_2\}$ in which there is tradable production in the second period, but not the first. The example has $a_m = 2$, and this is the second period wage. This generates high second period house prices and land rents, and also therefore relatively high expected returns to period 1 building, $r_1^E(x)$. The city is therefore built both larger and denser in the first period than is the case in equilibrium $\{M, M\}$, as is clear from the density profile $N_1(x)$. Period 2 land rents are higher again, so that there is further construction and this is at high density, curve $N_2(x)$. The relatively large amount of building in period 1 accommodates a large population, as a consequence of which period 1 wage, $w_1 = 1.42$, and house prices, $p_1(x)$, are lower than that in equilibrium $\{M, M\}$. If the

wage is less than a_0 enough (in this example, requiring $a_0 > 1.42$) tradable production is triggered. By assumption, it does not occur in period 1, but will surely do so in period 2.

The table in Fig. 4 also gives population, real incomes and welfare measures. Equilibrium $\{F_1, F_2\}$ supports nearly twice the first period population of equilibrium $\{M, M\}$, and more than three times the second period population. Real income gains accrue in the form of total land rents, R , and the present value of these are nearly four times larger in equilibrium $\{F_1, F_2\}$ than $\{M, M\}$ (and are also larger relative to the wage bill); deflating by the price index, real rents are around 3.7 times higher. The growth of the city and expansion of tradable production is associated with varying supply of non-tradables to the hinterland, this increasing hinterland consumer surplus in period 1 but reducing it in period 2. The bottom row of the table gives the combined welfare change from $\{M, M\}$ to other cases as a compensating variation. The gain ranges from 9% to 12% of the wage bill in $\{M, M\}$.²³

The effects of varying parameters of the model are as would be expected. Higher demand for non-tradables, H , shifts labour demand $w^D(L : L_T = 0)$ to the right, raising the period 1 wage and increasing city population. Higher construction and commuting costs shift points $\{F_1, F_2\}$ to the left, raising the first period wage and reducing population in both periods. The importance of future expectations can be seen by changing δ , the weight on the first period relative to the second. A higher δ has no effect on equilibrium $\{M, M\}$ and again moves points $\{F_1, F_2\}$ to the left, raising w_1 and reducing population. Each of these parameter changes increases the likelihood that w_1 exceeds a_0 , in which case the city fails to attract tradable production and equilibrium of type $\{F_1, F_2\}$ does not exist.

4. Concluding comments

The introduction to this paper sets out stylised facts about developing country cities, particularly those in Africa, indicating the presence of relatively high urban costs, a high cost of living, high nominal wages (alongside a low real wage), and failure to attract investment in tradable sectors of activity. Each of these is captured in the model developed in this paper. The analysis shows how the combination of increasing returns to scale, durable capital and sunk costs make for multiple equilibria, and the possibility – depending on parameters – that the city is built in a way that precludes establishment of the tradable goods sector.

Several broad policy messages follow. The first is the need to see the city as a whole. Policy has often been siloed, while our analysis highlights the interaction between all aspects of a city's urban form – residential construction, infrastructure, transport – and its economic

²³ The compensating variation is a discrete version of the marginal welfare change given in the appendix.

performance. The second is the high cost of policy failure. Small differences in initial conditions can – with increasing returns, sunk costs, and expectations – set cities on quite different development paths.

In terms of specific policy instruments, the paper points to the importance of efficient land use and infrastructure provision. The consequences of high building and commuting costs go far beyond their direct effects as they shape the sort of economic activity that takes place in the city. The many obstacles to residential construction that were noted in the introduction are all sources of inefficiency, which necessarily raise urban costs and thereby make the city a less attractive place for tradable goods production. Reducing these obstacles has direct benefits, and also increases the likelihood that the city will develop new sectors of activity. Expectations also matter, as the form and extent of investment in durable structures depends on the expected future prosperity of the city. Expectations need to be coordinated in some way, so that investors have confidence that the city – or a particular area within it – is likely to grow. Setting these expectations in a credible way is difficult and may require commitment in the form of investment in public infrastructure.

Stepping outside the confines of the model, two further points can be made. While the model focuses on investment in physical capital, investment in human capital is at least as important. Acquisition of the specialist skills needed to run modern production – and to run the city – will take place only if the costs of the investment are not too high and there are expectations of positive returns. Thus, the arguments that the paper has made with respect to physical capital apply with at least equal force to human capital. The model also draws out the possibility of vicious or virtuous circles leading to multiple equilibria, and here too further forces might be at play. In particular a fiscal feedback – not present in the model – may be important. A weak tax base is likely to lead to poor infrastructure and public services. This will raise urban costs, directly in the form of high transport and congestion costs and limited availability of power and other utilities, and indirectly via reducing the well-being of workers who require compensating wage payments. As we have seen, higher urban costs will undermine the city's economic performance and hence its tax base (e.g. lowering rents and land values). This in turn reduces the city's ability to provide such services, completing the vicious circle.

Finally – and for future research – this urban model needs to be placed in a wider model of urban hierarchy, within which different cities perform different functions. Some cities will specialise in non-tradables but, in all but the most natural resource abundant countries, some cities that are able to compete in non-resource tradable activities will surely be needed.

Appendix:

Parameters used in simulation:

$u_0 = 1$; $\theta = 0.4$; $t = 0.001$; $\delta = 0.5$; $c = 0.1$; $\gamma = 2$; $h(w) = Hw^{-\varepsilon}$, $H = 150$, $\varepsilon = 3$.

In Fig. 1 $a_0 = 1.45$, $a_m = 1.45$, $\alpha = 0$. In figures 2-4, $a_0 = 1.45$, $a_m = 2$; $\alpha = 0.008$.

In Fig. 1- 3 $r_0 = 0$. In Fig. 4 $r_0 = 0.05$.

Section 2.1:

1) Derivation of city population, (11), with general γ .

$$\begin{aligned} L &= \int_0^{\tilde{x}} N(x) dx = \left[\frac{w^{\theta-1}}{c\gamma} \right]^{1/(\gamma-1)} \int_0^{\tilde{x}} p(x)^{1/(\gamma-1)} dx = \left[\frac{w^{\theta-1}}{c\gamma} \right]^{1/(\gamma-1)} \int_0^{\tilde{x}} (w - (u_0 + xt)w^{1-\theta})^{1/(\gamma-1)} dx \\ &= \frac{\gamma-1}{t\gamma} \left[\frac{1}{c\gamma} \right]^{1/(\gamma-1)} (w^\theta - u_0)^{\gamma/(\gamma-1)}. \end{aligned}$$

2) Derivation of total land rent, (13), with general γ .

$$\begin{aligned} R &= (1 - 1/\gamma) \int_0^{\tilde{x}} p(x) N(x) dx = \left(1 - \frac{1}{\gamma} \right) \left[\frac{w^{\theta-1}}{c\gamma} \right]^{1/(\gamma-1)} \int_0^{\tilde{x}} p(x)^{\gamma/(\gamma-1)} dx \\ &= \frac{(\gamma-1)^2}{t\gamma(2\gamma-1)} \left[\frac{1}{c\gamma} \right]^{1/(\gamma-1)} w^{1-\theta} (w^\theta - u_0)^{(2\gamma-1)/(\gamma-1)}. \end{aligned}$$

$$\frac{R}{wL} = \frac{\gamma-1}{2\gamma-1} \left[\frac{w^\theta - u_0}{w^\theta} \right].$$

Section 2.2:

3) Comparative statics:

Equilibrium conditions:

Labour supply (11): $L = (w^\theta - u_0)^2 / 4tc$.

Labour demand (3): If $L_T = 0$, $L = h(w)H / \theta$, elasticity of demand $\eta \equiv -wh'(w)/h(w)$.

If $L_T > 0$, $a(L_T) = a_0$, elasticity of demand $\eta = \infty$.

City area (10): $\tilde{x} = (w^\theta - u_0) / t$.

Total rent (13): $R = w^{1-\theta} (w^\theta - u_0)^3 / 12tc$.

Define $\Omega \equiv \theta w^\theta / (w^\theta - u_0) > 0$, and note that $\Omega - \theta = \theta u_0 / (w^\theta - u_0) > 0$.

Totally differentiating and solving the log-linearised system gives comparative static responses:

$$\hat{w} = [\hat{H} + \hat{t} + \hat{c}] / (2\Omega + \eta)$$

$$\begin{aligned}\hat{L} &= -\eta\hat{w} + \hat{H} = [2\Omega\hat{H} - \eta(\hat{t} + \hat{c})]/(2\Omega + \eta) \\ \hat{x} &= \Omega\hat{w} - \hat{t} = [\Omega(\hat{H} + \hat{c}) - (\Omega + \eta)\hat{t}]/(2\Omega + \eta) \\ \hat{R} &= (1 - \theta + 3\Omega)\hat{w} - \hat{t} - \hat{c} = [(1 - \theta + 3\Omega)\hat{H} + (1 - \theta - \eta + \Omega)(\hat{t} + \hat{c})]/(2\Omega + \eta) .\end{aligned}$$

It follows that:

$$\begin{aligned}\text{Average density:} \quad \hat{L} - \hat{x} &= [\Omega\hat{H} - (\Omega + \eta)\hat{c} + \Omega\hat{t}]/(2\Omega + \eta) \\ \text{Average rent:} \quad \hat{R} - \hat{x} &= (1 - \theta + 2\Omega)\hat{w} - \hat{c} = [(1 - \theta + 2\Omega)(\hat{H} + \hat{t}) + (1 - \theta - \eta)\hat{c}]/(2\Omega + \eta) \\ \text{Rent per person:} \quad \hat{R} - \hat{L} &= (\Omega + 1 - \theta)(\hat{H} + \hat{t} + \hat{c})/(2\Omega + \eta) \\ \text{Rent/ wage bill:} \quad \hat{R} - \hat{L} - \hat{w} &= (\Omega - \theta)(\hat{H} + \hat{t} + \hat{c})/(2\Omega + \eta)\end{aligned}$$

Section 3.1: Multiple equilibria.

Existence of equilibrium M (Fig. 2). Non-tradable production commences at $L = h(a_0)/\theta$ so

existence of point M requires $w^s(L) = \{u_0 + 2(tc h(a_0)/\theta)^{1/2}\}^{1/\theta} > a_0$.

Existence of equilibrium M': (Fig. 2). If the productivity relationship is linear in tradable employment between lower and upper bounds, a_0, a_m , then $a(L_T) = \min[a_0 + \alpha L_T, a_m]$. The maximum level of productivity is first attained at $L_T(a_m - a_0)/\alpha$. From Eqn. (3b), at point M', $h(a_m) = \theta L - L_T$, hence $L = [(a_m - a_0)/\alpha + h(a_m)]/\theta$. Point M' exists if

$$w^s(L) = \{u_0 + 2(tc [(a_m - a_0)/\alpha + h(a_m)]/\theta)^{1/2}\}^{1/\theta} < a_m.$$

Section 3.1. Welfare analysis

We measure the change in welfare between situations 1 and 0 as the compensating variation $CV = (R_1 + CS_1 - CS_0)/w_1^{1-\theta} - R_0/w_0^{1-\theta} \approx d(R/w^{1-\theta}) - dCS/w^{1-\theta}$ where CS is hinterland consumer surplus. The interpretation is the change in the real value of rents net of 'compensation' of hinterland consumers for a change in consumer surplus. Total rents are

$$R = w^{1-\theta} (w^\theta - u_0)^3 / 12tc .$$

$$L = (w^\theta - u_0)^2 / 4tc , \text{ gives } d(R/w^{1-\theta}) = L\theta w^{\theta-1} dw .$$

The change in hinterland consumer surplus is simply quantity times price change, so $dCS = \{L_T - \theta L\}dw$. Hence,

$$CV = L_T w^{\theta-1} dw , \text{ i.e. the real value of the productivity increase in tradables.}$$

Section 3.2. Second period population

Using (5b) and (7b),

$$\int_{\tilde{x}_1}^{\tilde{x}_2} N_2(x)dx = \int_{\tilde{x}_1}^{\tilde{x}_2} [w_2^\theta - (u_0 + xt)]/2c \cdot dx = [w_2^\theta - u_0 - t(\tilde{x}_1 + \tilde{x}_2)/2](\tilde{x}_2 - \tilde{x}_1)/2c$$
 which, with $\tilde{x}_1 = (w_1^\theta - u_0)/t$ and $\tilde{x}_2 = (w_2^\theta - u_0)/t$ gives the expression in the text.

References:

- Ades, A. and E. Glaeser (1995). ‘Trade and Circuses: Explaining Urban Giants.’ *Quarterly Journal of Economics* 110 (1): 195–227
- Bates, R. (1981) *Markets and States in Tropical Africa*. Berkeley: University of California Press
- Bernard, L., O. D’Aoust, and P. Jones (2016). *The Urban Premium in Africa: Is it Real?* Mimeo, University of Oxford.
- Chauvin, J.P., E. Glaeser, Y. Ma and K. Tobio (2016). “What is different about urbanisation in rich and poor countries? Cities in Brazil, China, India and the United States”, NBER working paper 22002.
- Collier, P. and A.J. Venables ‘Housing and urbanization in Africa: unleashing a formal market process’, in E. Glaeser and A. Joshi-Ghani (eds) (2015), *The urban imperative; towards competitive cities*, OUP, Oxford.
- Duranton, G. (2008) Viewpoint: From cities to productivity and growth in developing countries. *Canadian Journal of Economics* 41(3):689–736.
- Duranton G. and D. Puga (2005) ‘From sectoral to functional urban specialisation’, *Journal of Urban Economics*, 57: 343–370.
- Duranton, G. & D. Puga, (2015). ‘Urban Land Use’, in: G. Duranton, J. V. Henderson and W. Strange (eds.), *Handbook of Regional and Urban Economics*, volume 5, chapter 8, pages 467-560, Elsevier.
- Fay, M. and C. Opal, (2000). “Urbanization Without Growth : a not-so uncommon phenomenon.” The World Bank Policy Research Working Paper Series 2412.
- Foster, V. and C. Briceno-Garmendia (2010) ‘*Africa’s infrastructure; a time for transformation*’, World Bank, Washington DC.
- Gelb, A., Meyer, C., and Ramachandran, V. (2013). “Does Poor Mean Cheap? A Comparative Look at Africa’s Industrial Labor Costs”. Center for Global Development Working Paper 325, Washington, DC: Center for Global Development.
- Gelb, A., and A. Diofasi, (2015). What determines purchasing power parity exchange rates? Center for Global Development Working Paper 416.
- Gollin, D., R. Jedwab, and D. Vollrath, (2016) “Urbanization with and without industrialization.” *Journal of Economic Growth*, 21(1): 35–70.

- Jedwab, Remi & Vollrath, Dietrich, 2015. 'Urbanization without growth in historical perspective,' *Explorations in Economic History*, 58, 1-21.
- Nakamura, S., Harati, R., Lall, S.V., Dikhanov, Y.M., Hamadeh, N., Oliver, W.V., Rissanen, M.O., and Yamanaka, M. (2016). *Is Living in African Cities Expensive?* World Bank Policy Research Working Paper, No. 7641
- Henderson, J.V. and A .J. Venables (2009) 'The dynamics of city formation', *Review of Economic Dynamics*, 12, 233-254.
- Henderson, J.V., T. Regan and A .J. Venables (2016) 'Building the city; sunk capital and sequencing', *in progress*
- Henderson, J.V., A. Storeygard and U. Deichman, (2015) 'Is climate change driving urbanisation in Africa' processed, LSE
- Henderson, J.V. and D. Nigmatulina (2016) The fabric of African cities: How to think about density and land use. Draft. The London School of Economics.
- Jones, P. (2016) 'Closed for Business; high urban costs', processed, Dept of Economics, Oxford.
- Lipton M. (1977). 'Why poor people stay poor: urban bias in world development.' Cambridge University Press, Cambridge
- Murphy, K., A. Shleifer and R. Vishny (1989) "Industrialization and the Big Push" *Journal of Political Economy*, 1003-1026.
- Nakamura, S., Harati, R., Lall, S.V., Dikhanov, Y.M., Hamadeh, N., Oliver, W.V., Rissanen, M.O., & Yamanaka, M. (2016). *Is Living in African Cities Expensive?* Mimeo, World Bank.
- Toulmin, C. (2009), "Securing land and property rights in sub-Saharan Africa: the role of local institutions", *Land Use Policy* (26.1): 10-19
- UN-Habitat (2010) 'The challenge of slums: global report on human settlements 2003' (updated 2010) UN, New York
- Williamson, J.G. (1988) 'Migration and Urbanisation', in *Handbook of Development Economics*, vol 1, pp425-465, eds. H. Chenery and T.N. Srinivasan, North Holland, Amsterdam.
- World Bank (2009) '*Reshaping Economic Geography; World Development Report 2009*', World Bank, Washington DC