

Giants and midgets:
the effect of public goods' provision on urban
population concentration

Viacheslav Yakubenko

Department of Economics, University of Göttingen,
Platz der Göttinger Sieben 3, 37073 Göttingen, Germany

August 12, 2016

Abstract

Bigger cities increase income per capita and provide access to a larger pool of public goods, but also require higher living costs. Since individual utility is increasing both in private consumption and public goods (but at a decreasing rate), there exists some optimal city size that maximises the utility of its citizens and not only income. Moreover, all cities in the economy are interrelated and migrants affect economic outcomes both at the cities of origin and the destination. This paper presents a theoretical model (supported by empirical evidence) that explains migration decisions within a system of cities and provides an explanation for the existence of urban giants in developing countries. The model suggests that (1) differences in public goods' provision create incentives to migrate to the primate cities; (2) better national infrastructure development decreases these incentives and, hence, urban population concentration; (3) uneven distribution of public goods leads to the emergence of urban giants. These findings are especially relevant for developing countries, where rapid urbanisation is currently taking place.

Keywords: primacy, public goods, growth, urbanisation

JEL Codes: O1, R12, R53, H41, H71, E20

1 Introduction

Urbanisation has been ongoing since the very early stages of human development. In the twentieth century it has reached an extremely high pace. While in 1950 only one third of the population was living in urban areas, this number increased to half of the population by 2014. By 2050 two thirds of the world's population are expected to be urbanised (United Nations, Department of Economic and Social Affairs, Population Division 2015). These figures show how extremely rapid the urbanization process has been in the past decades and is expected to continue like this. However, countries are currently at different stages of urbanisation. For example, developed countries are already rather urbanised. In North America, 82 per cent of the population live in the urban areas and 73 per cent do so in Europe. On the contrary, Asia and Africa are still mostly rural with only 48 and 40 per cent of their population living in the urbanised areas, respectively. This connection between urbanisation rates and income levels is clear and not new to the literature.

Modern theory finds many reasons for the cities to emerge. For the purpose of this research we can roughly split them into two major (nevertheless, often interrelated) groups: economic and non-economic advantages of urban lifestyles. The latter group of reasons includes numerous factors and depends on individual preferences. The set of economic benefits is often described as urban agglomeration economies. Spatial concentration decreases transaction costs, provides easier and cheaper access to production infrastructure and gives the enterprises a larger pool of skilful employees. Duranton & Puga (2001, 2004) demonstrate micro-foundations of urban agglomeration economies in more detail. However, an increase in the size of urban agglomerations leads to both benefits and disadvantages. These drawbacks can be split into non-economic and economic, too. Large cities are associated with higher pollution and crime levels simply due to their increased size. Besides that, higher rents and commuting costs may harm economic development. Moreover, effects as crowding and decreasing returns to scale might have negative influence on the economy. There exists no universal size or number of citizens that maximize the difference between agglomeration benefits and drawbacks of the oversized cities. The optimal size is rather dependent on many factors that can substantially vary between cities. A number of studies explore the relation between population size and economic activity within the city (Fujita 1989, Fujita et al. 1999, Henderson 2005, Glaeser 2008). It is obvious that neither too small, nor too large cities are desired. If a city is

undersized, it provides urban benefits to fewer people than it potentially could holding their quality constant. The opposite holds for a city that is too big: disadvantages outweigh overall urban benefits.

The urbanization-growth link has been studied extensively in the economic literature. However, there are many aspects of urbanisation that remain understudied. One of them is the distribution of the population across the existing cities. Despite the lower overall urbanisation rates, the world's largest cities are located in developing countries. Only four out of twenty biggest cities are located in the developed economies. This raises a set of interesting questions. Do cities tend to be larger in poorer countries? Or do urban giants make the whole economy poorer? What are the underlying reasons for either of the processes? Modern literature has no single straightforward answer to these questions. A range of studies, e.g., by Ades & Glaeser (1995), Desmet & Rossi-Hansberg (2013), Jedwab & Vollrath (2015) provide different explanations of the emergence of urban giants. This paper attempts to provide another explanation for the existence of large cities in developing countries and illustrates the consequences of such population distribution pattern.

Leaving aside the differences between rural and urban income, population distribution across the existing urban areas might have a significant effect on the economy. If economic activity concentrates only in one city, other areas are likely to be under-invested, due to the propensity to allocate firms in the most developed area. Relying on a neoclassical production function, while a city grows bigger the marginal income derived in it decreases (due to decreasing returns to scale), but negative externalities increase. However, people are not likely to migrate to smaller cities as they appear less attractive due to worse employment opportunities and a lower quality of public goods offered, as a result of low investments. Every country has its primate (largest) city, but degrees of spatial concentration can vary substantially. Typically, but not necessarily, the primate city is the capital. Existing literature provides us with a range of papers that consider how population distribution affects economic outcomes. Henderson (2003) demonstrates that at low levels urban concentration (measured by urban primacy) have a positive effect on productivity, but after a certain point it can harm the economy due to inefficient allocation of resources. Knowing the consequences of the urban concentration, we need to explain, which factors determine the degree of spatial concentration. This paper is therefore an attempt to introduce some of the new factors.

The research topic appears to be practically important for economic growth,

especially in the case of developing countries. Typically, countries at lower stages of economic development are more likely to have relatively smaller urban population due to concentration of workers in agriculture. Nevertheless, the ongoing industrialisation brings more people to the cities as a result of further economic development. In this case, governments can lack funds (or cut them down intentionally) for maintenance of other cities rather than capitals (Ades & Glaeser 1995). However, such a process might be rather harmful for the country as a whole. The alternative way to distribute funds might significantly improve the aggregate well-being and provide basis for more equal and faster growth.

In this paper I will present a theoretical model that considers both individual income and public goods offered in the city. The model is able to explain the migration decisions resulting into the emergence of urban giants in the poorer countries. To highlight the novelty of this paper, it is important to note that the previous studies of urban primacy were considering only monetary effects of population concentration. For example, Henderson (2003) in his model explains migration decisions with differences in income obtained at different locations. However, there might be other important factors affecting urban concentration. Using the existing findings on the topic, I develop a new model assuming that people decide to reside in the areas, where they will be able to maximize their utility (in other words, be generally happier). This approach looks more realistic and allows us to better explain the existence of urban giants in the developing countries. Moreover, utility in this model is derived not only from income and private consumption, but also from public goods. Given this, citizens migrate to the areas where they will not only earn more, but also have the opportunity to live more comfortable lifestyles due to a wider range of amenities offered to them.

This paper is organised as follows. After the introduction I present a brief literature overview on the topic. Then I describe the theoretical model that includes private and public goods to explain migration and population concentration. The fourth section presents the results of empirical testing of the theoretical model. The conclusion section summarises the results of the paper.

2 Theoretical background

To start the paper, first of all, I need to define what is a city. The question might sound too simple, but it is very important to decide on what exactly we mean

by a “*city*”. Modern literature usually refers to the “*metro areas*” consisting of many municipalities. By defining metro areas we can cover the “entire labour market of the area, service and residential activities radiating from the core city, until activity peters out into farm land or very low density development” (Henderson 2005, p.1548). Employing the metro area definition, rather than the formal bureaucratic interpretation, we concentrate on functions that a city fulfils. However, sometimes, especially in the case of highly urbanised countries, metro areas can grow extremely big and even absorb areas that are specialised on non-urban economic activities (agriculture) or satellite cities. The intermediate solution would be the concept of urban areas that have a common infrastructure system, but do not contain the rural land and close-located settlements. Further in this paper when I refer to a “city” one should generally consider the “urban area”.

It is important to demonstrate the benefits of urban agglomerations for the economy. Modern literature has both theoretical and empirical evidence that cities can spur economic development (Sveikauskas et al. 1988, Fujita 1989, Duranton & Puga 2001, Henderson 2003, Glaeser 2008, Desmet & Rossi-Hansberg 2013). Duranton & Puga (2004) describe the benefits of concentration of economic activities in the cities (also called *urban economies*). These advantages can be split into three mechanisms: sharing, matching and learning. Moving to the city gives individuals and firms access to the common infrastructure. Building power plants or water supply systems might be extremely expensive, so these costs are shared by the citizens. Moreover, matching of economic agents reduces transaction costs and increases the workers efficiency. In a bigger city firms have larger pool of potential workers, so better chances to hire qualifications they particularly need. Individuals are more likely to find jobs matching their profile without requiring additional training. Furthermore, workers learn faster in the larger cities. More diversified environments facilitate information exchange and foster innovations (Duranton & Puga 2001). This makes individuals more likely to learn new skills and become more productive. These three mechanisms are enough to justify why cities emerge. However, along with urban economies, larger urban areas are associated with congestion. Higher crime incidence, pollution, traffic are far not the only drawbacks of the large agglomerations. As both positive and negative effects increase with city population, there might be some trade-off between the two.

Given that both urban economies and diseconomies grow with size, one might expect some optimal city size to exist. Henderson (2005) considers that

the “*optimal*” size maximises the real income of the residents. This approach seems totally reasonable, however the model does not explain the existence of mega-cities in developing countries. To see it explicitly, we need to look at the efficient population size n^* , derived from the model:

$$n^* = (\delta 2b^{-1}D)^{2/(1-2\delta)}h^{2\varepsilon}, \quad (1)$$

where δ is elasticity of output with respect to population¹, b is a linear function of transport costs (increasing in per unit transport costs), D is a productivity parameter and h is a human capital. Given that $\delta \in (0, \frac{1}{2})$ and $\varepsilon > 0$, the model implies that the optimal size is increasing in technological innovations: δ , D and h are expected to increase with technological progress and lead to a higher n^* . At the same, time improvements in transport technology enable cheaper and faster commuting, decreasing b .

Given the above mentioned outcomes of the model, why do we have mega-cities in poor countries? Technology is expected to be more advanced in the developed countries, but cities rarely reach such a giant size there. This does not necessarily mean that the model is wrong. First of all, cities, especially those in poorer countries, might be of not efficient size. Moreover, some assumptions of the Henderson’s model might be not likely to hold in the real world. Particularly, the equation for n^* is derived as the maximisation problem of the city authorities. Basically, n^* allows to maximise the local budget revenue. However, control over migration flows requires a set of restrictive policies. Unlike China, many countries do not have severe migration restrictions. As a result, local authorities do not really determine the city size, but rather face population size as an exogenous factor. Further in this paper I will present the model that rests on a set off less weaker assumptions.

As local authorities have no direct control of the city size, there should be other factors affecting the city size. Jedwab & Vollrath (2015) explain the existence of urban giants as a result of post-war mortality transition. Countries that had been industrialised by the end of WWII already had lower birth rates. As a result, substantial drop in death rates did not entail exploding population growth. On the other hand, countries that were not industrialised at the moment kept high birth rates even after a reduction in number of deaths. This lead to a fast increase in population that was much higher than wage growth.

¹It is important to note that population is not similar to labour. Here the population literally stands for the number of people residing in the city. It does not consider the amount of working hours and qualification of workers.

These cities ended up being large, but poor, as their population grew too fast for a city to be able to maintain income growth high enough to compensate increasing congestion. This hypothesis is perfectly reasonable and sheds some light on the problem. However, there might be other factors that affect urban concentration. Moreover, fertility transition alone does not fully explain the population distribution patterns. According to Jedwab & Vollrath (2015), all cities should be larger in developing countries at the moment. However, if we compare Canada and Uzbekistan, two countries with comparable population, but different stages of economic development, we will see some contradiction to this hypothesis. Uzbekistan's largest city, Tashkent, has around 2 million inhabitants (The State Committee of Republic Uzbekistan on statistics 2015), while in Canada three cities have larger population and three more have over 1 million citizens (Statistics Canada 2016). This is a very rough comparison, but it contradicts the mortality transition argument. Another fact worth noting is the difference in urban structures of two countries. Toronto, Canada's primate city, is less than 1.5 times larger than Montreal, the second-largest urban agglomeration nationwide (Statistics Canada 2016). In Uzbekistan the contrast is much sharper: Tashkent is 4.6 times bigger than the second-largest Samarqand (The State Committee of Republic Uzbekistan on statistics 2015). We clearly see that Canada has more even urban population distribution compared to Uzbekistan. This is just a special case not leading to any general conclusions, but it leads us back to the question: does *relative* size of a city matter for the economic outcomes? This question is not entirely new to the literature. Henderson (2003) demonstrated that countries with higher per capita income tend to have lower primacy rates. This was an empirical study which clearly indicated the pattern, but did not explain the underlying mechanisms in a bigger detail. Moreover, Henderson (2003) leaves aside possible consequences of a particular urban structure for the national economy.

Ades & Glaeser (1995) had an attempt to explain the existence of megacities in developing economies from a political point of view. According to their hypothesis, countries with a higher degree of political instability are more prone to have relatively larger primate cities. Besides this, dictatorships lead to more centralised urban populations. Ades & Glaeser (1995) demonstrate the exact mechanisms for these effects to take place. In their model government is assumed to set different tax rates in the capital and the hinterland to maximise the expected tax revenue. If the tax is too high, this leads to the population disapproval, which can be expressed in two ways: legal and illegal (violent).

The latter are assumed to be successful only in the capitals, as rebellions can be suppressed in the hinterland. The opposite holds for the legal protest (elections): as more than a half of the population is assumed to live outside the capital, the median voter resides in the hinterland. As a result, the survival probability of the government depends on the tax rates. Too high tax in the capital leads to the loss of power through violent revolts. At the same time, excessive taxes in the hinterland result into the failure to get re-elected. However, existing political systems might significantly affect these two mechanisms controlling the government tax-setting power. Countries with a higher degree of political instability are expected to have lower taxes in the capital, that will attract more migrants from other areas. The same will hold for dictatorships: without a threat (or with a lower one) of loosing the elections the government can set higher taxes in the hinterland causing migration flow to the capital.

Ades & Glaeser (1995) demonstrate how political systems might affect urban concentration through the taxes. However, tax rates are not the only mechanism for such an influence. The government might decrease the revolt probability in the capital by other means, rather than lower taxes. One of the obvious ways to “please” the citizens would be the increased provision of public goods. Better amenities increase utility of individuals and make them happier and less prone to the violent revolts. It is quite natural to expect that more public goods and higher potential utility will attract more migrants to the city. As a result, population of the primate city (hence, urban concentration) will increase. Henderson & Wang (2007) demonstrate that better democracies have lower urban population concentration rates. Their explanation is based on the fact that smaller cities are much more likely to be represented in the government, if the democratic institutions are better developed. Davis & Henderson (2003) come up with the same conclusion that more fiscally decentralised countries have lower primacy rates. However, their way of measuring fiscal decentralisation rather represents the level of political independence of regions from the centre, what does not make the results less valid.

The topic of this paper is not entirely new to the literature. However, the theoretical studies tend to leave aside the effect the population size of the primate city on the rest of the country. On the other hand, empirical articles demonstrate this link, but do not explain the underlying mechanisms. The model presented in the next section considers a system of cities and explains how population distribution can affect different outcomes in these cities.

3 Urban system model with public goods

3.1 The model set-up

It is natural to believe that people decide to reside in the areas, where they will have opportunities to live happier lives. In the economic theory world this can be represented through the different utility levels they could achieve in particular locations. I assume that individuals derive utility from consumption, as commonly done in the literature. All goods can be split into two categories: private and public. For simplicity let all the disposable income to be spent on individual consumption. In other words, one component of the utility function is net income. Moreover, individuals enjoy public goods in the cities. Better roads, parks or hospitals also increase the utility of the citizens. To put it formally, we can assume the standard log-linear utility function:

$$U(c, \theta) = a_1 \ln(I) + a_2 \ln(\theta), \quad (2)$$

where I is a disposable income of the individuals (as was said above, entirely spent on privately consumed goods) and θ is a sum of public goods. For simplicity I neglect prices and relative preferences towards separate goods in the subgroups aggregating all goods into two major categories (privately consumed and public ones). As both types of goods are arguments of a log function, they have to be positive. This is quite a realistic assumption, as an individual needs at least some little amount of each type of goods to survive (e.g., food and healthcare). a_i are the parameters indicating relative weights of each goods category in the utility function ($a_i > 0$).

We clearly see that in this model utility is increasing in disposable income and public goods, but at decreasing rates. If we assume that individuals migrate within the country in search of higher utility, the migration decisions can be motivated by (1) higher disposable incomes and (2) better (or simply more) public goods at the new place of residence. However, it might be the case that higher level of public goods comes at a cost of increased tax payments. This lowers disposable income, hence, private consumption. As a result, we can expect some trade-off between net income (spent on private consumption) and public goods that yield a higher utility of an individual. As it is seen from the equation 2, individuals would consider not only their private consumption level (hence, disposable income), but a level of public goods provision, as well. On the other hand, there might be some factors that can prevent migration. To put

it formally, a person will migrate from A to B if the utility increase associated with the move is high enough to compensate the resistance term ρ :

$$U_B(c_B, \theta_B) - \rho > U_A(c_A, \theta_A). \quad (3)$$

It is also worth noting that ρ can be expressed both in monetary terms (e.g., cost of moving) and utility terms (e.g., joy of living at the birthplace). A practical outcome of the equation 3 is the fact that intercity migration must be driven by substantial differences in utility levels of two cities. This explains the absence of constant mass-migration within the country. Individuals are not constantly chasing slightly higher utility, but rather reside in the areas with the acceptable utility level. This kind of behaviour was already described by Simon (1955) as *satisfising*: instead of accepting only the utility-maximising option, individuals are indifferent between those that yield utility above some threshold level.

To consume private goods individuals earn some income by producing non-differentiated output. I assume the same production function as in Henderson (2005). It already incorporates urban economies implying that individuals become more productive in bigger agglomerations. All firms located in a city are consisting of one worker. Each worker-firm uses its individual human capital, h_i , to produce the non-differentiated good x . The individual production function can be written as:

$$x_i = D(n^\delta h^\psi)h_i^\phi, \quad 0 < \delta < \frac{1}{2}, \quad (4)$$

where h is the average amount of human capital per worker in the city and n is the population of the city. These two terms drive the urban economies effect. The individual worker is more productive, if he is working in a creative environment. Besides lower transaction costs and better infrastructure urban concentration fosters generation of new ideas. The innovative environment has two possible sources: the quantitative (number of people living in the city, n) and the qualitative (the average level of “intellectual quality”, h) components. An increase in both variables makes the worker more productive with the degree of substitution depending only on the elasticity of output with respect to the factor of production: δ and ψ . Due to the fact that there is only one production factor in the model, workers retain their whole output, so the wage equals the production x_i .

The restriction $\delta < 1/2$ is added to prevent an endlessly increasing utility, as will be seen later. Besides the technical reasons, limiting δ also appears

plausible if we apply the model to the real world. Sveikauskas et al. (1988) have analysed the effect of city and industry size on urban productivity in the food processing industry. They have found strong evidence for the economies of urban agglomeration to exist, indicating that larger size of a city is associated with higher productivity. The scale effect was estimated to be no bigger than 1.3 percent, meaning that in our formulation δ would not exceed 0.013, far below the 0.5. This fact still implies an exponential growth of the aggregate product in population of urban agglomeration, but justifies the restriction of δ . Moreover, this value of a parameter seems more realistic, as it requires big values of n to result into major productivity synergies. Even though, the study by Sveikauskas et al. (1988) was conducted for a particular industry, they have demonstrated that productivity might increase simply due to the larger number of workers involved in production.

As all workers are assumed to be identical, we can consider that $h_i = h$. Now the aggregate output of the city can be represented the following way:

$$X = x_i n = Dh^{\phi+\psi} n^{1+\delta}. \quad (5)$$

It follows from the equation 4 that larger cities are associated with higher gross incomes. However, disposable income I_i does not increase infinitely, as larger cities also imply higher living costs. Mohring (1961) has demonstrated that these costs increase in a number of citizens, but at a decreasing rate. I slightly simplify his findings and apply them to the model without a loss of generality. I assume that spatially a city can be represented as a circle filled with n number of land lots, each one of them hosting one citizen. Then $S_{city} = n$ is the area of the city measured in these lots of land. Individuals are assumed to travel between the edge of the circle and the centre, so the distance they cover is equal to the radius r of the circle.

$$\begin{aligned} S_{city} &= \pi r^2 = n \\ r &= \left(\frac{n}{\pi}\right)^{1/2} \end{aligned} \quad (6)$$

Transport costs per unit of distance, v , are assumed to be exogenous and independent from the distance. Having travel distance determined, we can use the equation 6 to determine commuting costs:

$$C = vr = v \left(\frac{n}{\pi}\right)^{1/2}. \quad (7)$$

The main difference of the equation 7 with the Mohring (1961) formulation of total urban costs is the absence of land rents. Including them would not alter the results, so I just leave them aside for the simplicity reasons. Henderson (2005) has described the city costs as a sum of commuting and rents:

$$\text{total costs} = \frac{3}{2}bn^{1/2}, \quad (8)$$

where parameter b is a linear function of per unit travel costs.

As one can see, the difference between the equations 7 and 8 is not dramatic. I have decided to exclude housing rents on purpose. This does not change the mechanics of the model, but allows to avoid a set of strong assumptions about the rents and city budgets. Henderson (2005) assumes that rents are collected by the local authorities to finance the governmental functions, what might be not fully realistic. To exclude the rents from the analysis I assume that all citizens possess their own housing and do not pay for it. Alternatively, I could keep rents in the model and assume that they go to the real estate owners that are very few and can be neglected from total population calculations. Including Mohring-type rents would not significantly alter the model.

Having both gross income and costs determined, we can determine the disposable income I . Individuals are assumed to earn the whole output x_i . However, besides productivity increases cities provide access to the range of public goods. This employs the sharing mechanism of urban economies. To be able to provide a range of public goods city authorities need to collect taxes. I assume a simple income tax at a rate $t \in (0, 1)$. Another major expenditure for an individual is the cost of commuting within the city. Given that, the disposable income can be presented the following way:

$$I = (1 - t)Dn^\delta h^{\psi+\phi} - v \left(\frac{n}{\pi}\right)^{1/2}. \quad (9)$$

As is was assumed before that $\delta < 1/2$, the function determined by the equation 9 has only one maximum point and is not increasing infinitely in population. This ensures that individual disposable income (unlike the nominal one) first grows due to higher productivity associated with larger agglomerations, but once it has reached its maximum, every additional inhabitant enhances congestion more than productivity.

The “income” of the city government is a tax revenue collected:

$$Revenue = \theta = x_i t n = t D h^{\phi+\psi} n^{1+\delta}. \quad (10)$$

Legislators are assumed not to make any profit, hence, they do not keep any part of the revenue. The whole tax income is paid back to the citizens in a form of public goods, so I assume that θ equals the whole revenue. One can easily see that amenities offered in the city are proportional to the city output and increase in population size, so the provision level is higher in bigger cities. Despite the tax payments being returned in amenities, the city population is not indifferent between the tax rates: individuals that value private consumption higher than the public will be in favour of lower taxes, and vice versa. Moreover, the utility in the equation 2 is a non-linear function of disposable income. Hence, the tax rate affects not just a level, but also the marginal utility of private income: high tax rates will result into a relatively larger marginal utility compared to the low ones. Having both components of the function determined we rewrite the equation 2 plugging in the values for I and θ :

$$U(I, \theta) = a_1 \ln[(1 - t) D h^{\phi+\psi} n^\delta - v n^{1/2}] + a_2 \ln[t D h^{\phi+\psi} n^{1+\delta}]. \quad (11)$$

The equation 11 can be also referred as an indirect utility function of a population n and a tax rate t . Keeping the tax rate constant, the function has an inverted U-shape in population. The intuition behind it is similar to the situation with disposable income. Given the decreasing marginal utility of both private and public consumption, both types of goods significantly increase utility, when they are scarce, but add little additional benefit, when they are abundant. At the same time, disposable income is approaching 0, if n is above the optimal point, but public goods are only increasing. Figure 1 illustrates these effects.

The above mentioned explains the shape of the aggregate utility curve with two components included: (1) at the very low values of n both gross wages and public goods are low; (2) until a certain point growing population increases both private consumption and provision of public goods; (3) further increase in n lowers the disposable income, but individuals still can get an overall utility increase due to a higher level of amenities; (4) after the maximum point overall utility decreases in number of citizens, as increasing public goods are no longer able to compensate the losses in private consumption. These phases are presented in a more illustrative form in the figure 2.

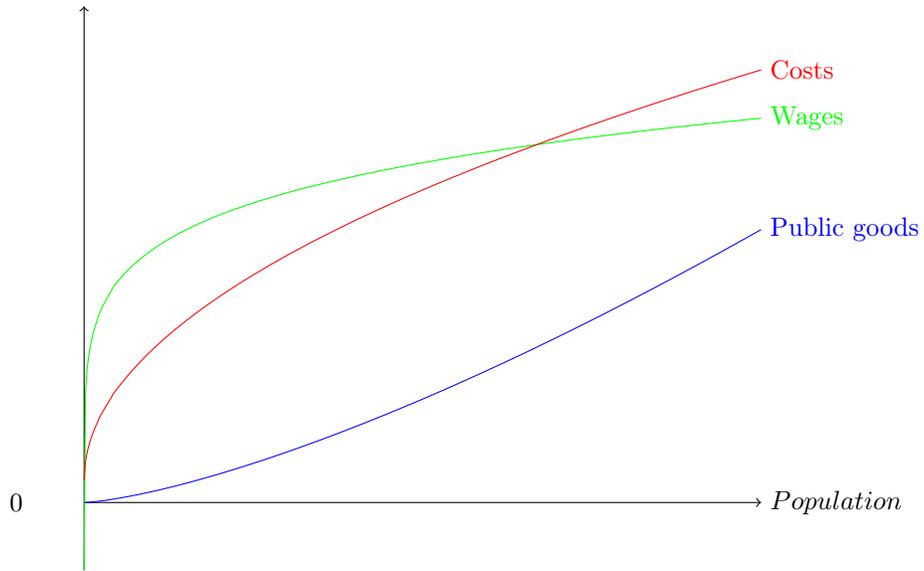


Figure 1: Wages, costs, public goods and city size

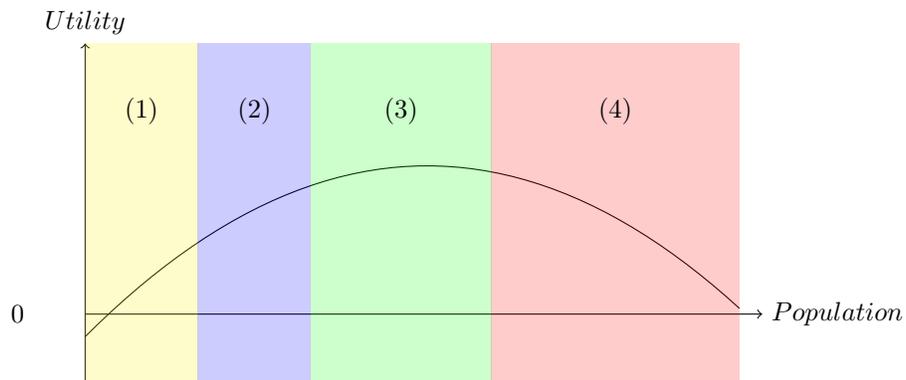


Figure 2: City size and utility

3.2 Model implications for a system of cities

As in Henderson (2005), the figure 2 implies that all cities have some optimal size that maximises the utility of citizens. However, I do not consider single cities, but rather the whole urban system. Now I will demonstrate how the utility level of one city can affect the well-being in other parts of the country within the model framework. For this purpose we can consider a simple case of two

cities, A and B, in an economy with free internal migration. Hence, the decision whether to migrate or not is solely described by the equation 3. Individuals will move from A to B only if the utility gain will be high enough to compensate the resistance term ρ . In the equilibrium situation both cities yield the same utility level creating no incentives to migrate (figure 3a). We do not need to assume that initially both agglomerations are at the optimal size. A is slightly smaller than the optimum, B is slightly bigger and utility is just equalised at some national level \bar{U} . In this case it is still possible to increase the aggregate utility of the nation: a certain number of people have to be reallocated from B to A, so both cities are closer to the maximum utility. However, that type of migration requires enormous coordination efforts and is quite likely to be not voluntary: there are no direct incentives for an individual to move to the city currently yielding the same (or similar) utility level.

Now consider the situation, when the utility levels vary between the cities. For example, once a shock occurs, and one of the cities (say, B) suddenly becomes more attractive. It plays absolutely no role, where does the shock take place: utility decrease in A, or a sudden increase in B, or both (presented in the figure 3b). If the utility gap between the two cities is bigger than ρ , it triggers intercity migration. One possible example of such a shock was partly described in Ades & Glaeser (1995). City authorities of a single city (often, but not necessary, the capital) might have more power in redistribution of national tax funds. As a result, θ of the city gets higher than the $tDh^{\phi+\psi}n^{1+\delta}$ value. This alone raises utility in the “preferred city”. Moreover, the difference between the city’s revenue and expenditures should be financed from some sources, so the θ s of other cities, on the contrary, fall below their tax incomes. As a result, the utility disparities between the cities occur. Due to the substantial differences in well-being individuals move from A to B. This migration raises the utility of migrants, but decreases the utility levels in both cities. Population reduction in A decreases productivity and level of public goods provision. Even despite lower commuting costs, overall utility goes down. The opposite holds for the city B: increased congestion cannot be compensated by higher productivity or better amenities. As a result, both cities move further from the optimal size (figure 3c) until utility levels converge or the gap becomes smaller than ρ . Due to the fact that A is already below the \bar{U} , the utilities can converge only below the initial utility level.

Why do not all people move out from the less attractive city? There exist several reasons for it. First of all, a model with two cities is just a simplification,

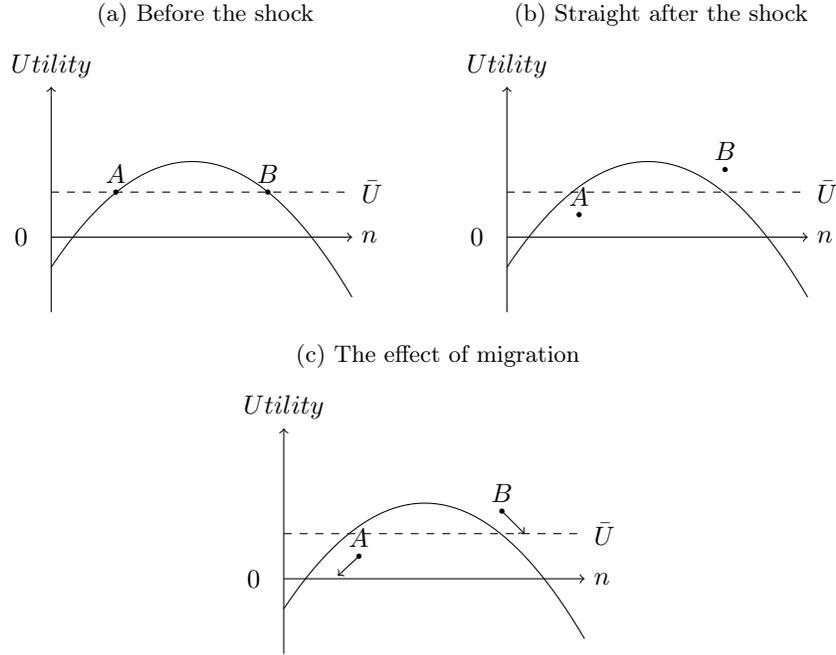


Figure 3: City size and utility before and after the shock

while in reality there might be more cities in the economy, so the same mechanics are realised in a more complex system. Complete abandonment of a certain cities is not always necessary for the utility levels to converge. Instead, each of the smaller cities will supply migrants to the primate city. Moreover, residents of rural areas can also consider the move to the more attractive city. This case will be considered later in this section in a bigger detail. Furthermore, even cities with different utility levels can coexist in the same economy. The migration mechanism described above rests on the assumption of free unconstrained migration. However, it can be easily violated in a number of ways. For example, individuals might have not all information or, what is even more likely, be financially constrained to migrate (high ρ). Hence, utility-equalising migration does not occur at all or occurs at a lower rate letting cities with different utility rates exist in the same economy.

The described model considers urban-urban migration and population distribution. However, it can also account for rural-urban migration. Returning to the figure 3a, in this situation a potential migrant from the rural area is indiffer-

ent between the two cities. IF we assume that initially migration flow is equally split between A and B, the subsequent population growth will raise utility in A and decrease it in B. After the point, where utility gap is large enough, residents of B start leaving for A, making population of both cities better off. As a result, both cities converge in population and utility. Note, that in extreme cases both cities might end up on the right side from the optimum (become overpopulated), if the number of migrants is too large. This fact partially motivates the existence of *hukou* system in China (Au & Henderson 2006). Urban citizens are in favour of imposing migration restrictions on the rural population anticipating that higher number of citizens will decrease the life quality of the city. In other words, unconstrained rural-urban migration is seen as a factor that might shift the cities further to the right from the optimum.

However, the rural-urban migration can be not only utility-equalising. The existence of distortions described above might influence this mechanism. In the situation shown on 3b, migrants from the rural areas are not indifferent between the two cities any more. They go to B, as the potential utility is higher there. Population inflow drives the city further from the optimal size and utility gap becomes smaller or disappears completely: the same process, as shown on the figure 3c. However, the size of B grows at a higher rate, due to the larger pool of potential migrants (residents of A and rural population). As a result, both cities end up with similar utility levels, below the initial \bar{U} value.

The cases mentioned above consider cities that initially lie on the different sides from the optimal size. The model also allows to study the situations, when both cities are smaller or bigger than the optimum point and the urban system is not at the equilibrium. If both cities are underpopulated, the migration to the city with a higher utility will first move it closer to the optimal point pushing the smaller one further down in utility terms. However, once the “more attractive” city reaches the maximum point, it does not hold there, as new migrants are still attracted by a higher utility, so the city goes on growing until the utility levels of the two equalise (not at the optimal point) or the smaller city gets completely abandoned. A slightly different picture can be observed for the pair of cities that are overpopulated compared to the optimum. Migration to the city with a higher utility will make it even more overpopulated and push it further from the optimum. As a result, the utility levels in both cities converge to some non-optimal level and cities remain overpopulated.

To sum it up, this section of the papers has presented the theoretical model that explains the population distribution across a system of cities. Citizens are

assumed to move to the areas, where they will derive the highest possible utility level. The utility function of an individual consists of two components: private consumption (disposable income) and public goods available in the city. Bigger cities are associated with higher gross income and more public amenities coming at a cost of higher congestion costs. Hence, neither under-, nor over-populated cities are desired. Utility levels throughout the country are equalised through migration, but possible distortions in public goods provision lead to sub-optimal equilibria, when some share of potential utility is lost in congestion. As a result, alternatively both utility and private consumption could be higher, if public goods are distributed more evenly.

4 Empirical testing

After describing the theoretical mechanism, we can test it empirically. There is a set of potential hypotheses to check within the model framework using cross-country data. Unfortunately, we cannot measure utility directly. Moreover, the data for public goods provision at a regional level might be unavailable or not fully reliable. For these reasons, I need to develop such a strategy that will clearly demonstrate the link between amenities and urban population distribution using the available data. First of all, we need to determine the hypothesis that will support or doubt the validity of the model.

To begin with, I state the particular implications of the model that we want to test. According to the model, presented in the previous section, utility is increasing in public goods and disposable income, but at decreasing rates, as presented on the figure 4. This implies that when both productivity and public goods are at low levels, even minor variations have large marginal effect. In other words, in the countries with low income and poor public goods provision (stage I on fig. 4) even slight improvements in one city might cause a significant utility gain and, hence, attract potential migrants much more. Contrary, due to a high level achieved, fluctuations in net wages and amenities in developed countries entail rather small utility variations, especially compared to the migration costs (stage II on fig. 4). As a result, the model predicts that individuals in the countries with lower level of urban amenities are more likely to be attracted to the primate cities. When the amenities on average are provided at a high level (that still might differ within the country), individuals are not sensitive to the variations in public goods, hence, reluctant to migrate. These facts imply that

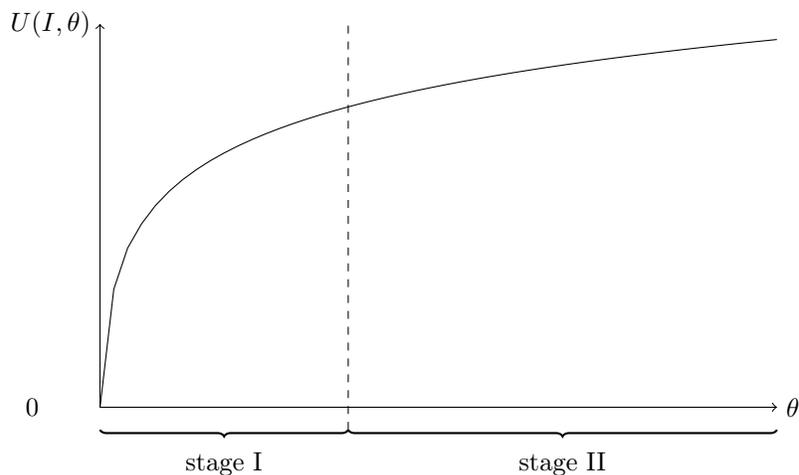


Figure 4: Utility and public goods

urban population is expected to be more concentrated (relatively larger primate cities), when public goods are scarce, and more dispersed (relatively smaller primate cities) when amenities are abundant. To sum up, the main hypothesis to be tested is: “urban population concentration in developing countries is higher due to the lower level of public goods provision”. This hypothesis does not show the direct effect of urban population distribution on utility. Instead, it analyses the causality from another side. The relationship between income and urban concentration is not new to the literature. However, public goods, which appear to be extremely important in this model, might be another explanation to the existence of urban giants in the developing countries.

4.1 Data

Having the hypothesis set, it is important to determine the data that will be used for the empirical testing. First of all, I need to determine the dependent variable. Urban concentration is the extent to which resources are spread between the largest city and the hinterland. To capture this I employ the *urban primacy* variable. It shows the share of total urban population living in the largest city. However, alternatively population concentration can be measured by a number of measures. Henderson (2003) mentions three concentration indicators. First one is Herfindahl-Hirschman Index (HHI). It is calculated as a

sum of squared shares of total urban population living in each city. Because the share of the largest city is usually quite high compared to the smaller ones, the largest city will outweigh shares of the others. Hence, HHI is expected to be heavily dominated by the share of the primate city, HHI values will be strongly correlated with urban primacy value, which is much easier to compute. Another option is to use Pareto parameter “looking at the distribution of city sizes within a country, which measures how quickly size declines as we move from up to bottom in the size distribution, or the overall degree of disparity in size distribution” (Henderson 2003, p.50). Rosen & Resnick (1980) shows that estimated Pareto parameters are also highly correlated with urban primacy. Given the existence of Zipf’s law, all cities are expected to have similar value for the product of a size of the city times its rank in the distribution (Gabaix 1999). Then, the size of the largest city is again sufficient to determine population concentration. Moreover, both HHI and Pareto distribution require a lot of data for calculation, which is not always available or reliable. However, urban primacy is easy to calculate and is highly correlated with these measures, thus, it appears the appropriate measure of urban concentration.

After determining the way to measure urban concentration, I need to choose the variables that might affect it. Basically, I have to come up with a set of indicators that will reflect how well the public goods are provided. For this purpose I use two different strategies. First, we can simply measure provision of some public goods in the cities. I arbitrary choose a set of variables that indicate how likely individuals living in urban areas are to have access to these amenities. If a public good is well provided in all the cities of a country, it should have no effect on population concentration. However, when it is scarce, individuals have incentives to move to the areas, where they will have an opportunity to use this good. The variables used under this strategy are: access to electricity, non-solid fuel and improved sanitation facilities. All the three variables are calculated as a share of urban population having access to these public goods. It is important to note that these indicators do not show real consumption of the amenities, but rather measure their accessibility in a city. For example, electricity consumption is not a public good. It is clearly excludable, individuals normally have to pay to use it. However, the *access to electricity* variable measures the capability to have an option of using (and paying) for the electricity. Thus, the selected variables serve as indicators of the development of some basic public infrastructure in a city, hence, can serve as proxies for the overall *provision* of public goods and not necessarily of their actual consumption.

The second way to measure the public goods provision is to use the level of public expenditures data, assuming that government spendings are converted into amenities. Instead of referring to particular public goods, as it is done in the previous strategy, we can consider the aggregate level of their provision. Despite the fact that public goods are considered non-rival, using overall expenditures appears to be misleading. Countries with larger populations are expected to have higher GDP and, hence, public expenditures. However, in per capita terms, the values can be much lower compared to the smaller, but better developed states. Apparently, hardly any good can be considered perfectly non-rival. At some point congestion decreases the quality of amenities, so it appears more reasonable to consider per capita expenditures. Nevertheless, this fact does not contradict the theoretical model. As it is seen from the equation 10, every additional citizen increases both overall and per capita tax revenue preventing possible dilution of public goods.

The theoretical model presented in this paper leaves aside some variables that might affect urban population concentration. Obvious factors would be land and population size. The larger is the country area, the more space is potentially available for foundation of new cities. Moreover, spatially bigger countries are more likely to have attractive locations. Pirenne (2014) finds the origins of the medieval cities in trade. Bleakley & Lin (2012) demonstrate that river trade routes were an important determinant for the emergence of the U.S. cities. Hence, bigger countries that were more likely to have a higher number of trade routes coming through their territory and, as a result, had more cities. Furthermore, larger areas have higher probability to be endowed with natural resources, creating another necessity for the foundation of cities. Population is also expected to be an important factor considering the population distribution across the urban agglomerations. The more people live in the country, the more cities can be potentially populated. Another control variable in the analysis is GDP per capita. Income is expected to be highly correlated with public-goods-related variables, but it is added to demonstrate that results are not driven by the variations in per capita GDP. Moreover, exclusion of income does not alter the results.

4.2 Methodology

The panel dataset employed in this paper is unbalanced. The variation in the data is across countries and time. Given the set of chosen explanatory variables,

Table 1: Summary statistics

Variable	mean	Std. Dev.	min	max
primacy	33.10	15.40	2.61	89.39
urb. electr	85.57	23.65	0.1	100
urb. fuel	74.86	35.15	1.99	100
urb. sanitation	74.65	26.46	6.3	100
per capita exp	1093.20	2254.00	0	21897.05

I cannot use the fixed effects framework, as it would cause a conflict with land, which is constant for most of the countries in the sample. However, even after excluding land from the regression specification, the Hausman test does not reject the null-hypothesis implying that random effects estimator is consistent. Another notable fact is that coefficients do not substantially differ under two estimators. Standard errors are clustered at the national level. The estimated equation can be presented the following way:

$$primacy_{it} = \alpha + \beta_1 public\ good_{it} + \beta_2 population_{it} + \beta_3 land_i + \beta_4 GDPp.c._{it} + \varepsilon_i + \mu_{it}, \quad (12)$$

where i and t are country and time indicators, respectively; ε_i is a country-specific part of the error term and μ_{it} - an idiosyncratic one.

The model assumes that all types of public goods are financed from one source - tax revenues. This implies that the more funds available, the higher the provision of each public good can be. Hence, all types of amenities are expected to be correlated between each other. As a result, it appears to be not correct to put them into one regression. Due to this reason, each of the estimated equations has only one public variable.

A possible issue arising during the estimation could be endogeneity in the case when we measure the actual provision of some amenities. One could argue that it is less complicated to provide public goods in the cities if urban population is more concentrated. For example, it might be easier to extend the existing sewage system in the primate urban area than build it from scratch in a smaller city. I try to control for it using lagged variables on the right hand side of the equations, as current primacy rates should have no effect on the previous values of amenities. The similar intuition can be followed in the case of public expenditures variable. Henderson (2003) has demonstrated that urban population primacy might affect productivity growth. Countries at higher stages of

economic development have the opportunity to spend more on public goods, hence primacy cannot be considered perfectly exogenous. However, as in the previous case, using the time lags solves the issue. Below the results without the time lags are presented, but the estimates do not significantly change, when lagged variables are used, as presented in the Appendix.

4.3 Estimation results

The table 2 presents the regression estimates for three variables indicating the effect of provision of public goods on primacy. The results suggest that higher access of urban population to electricity and non-solid fuel is associated with lower primacy levels. In other words, in the countries, where these amenities are well-provided, urban population is expected to be less concentrated in the largest city. This fact supports the validity of the model presented above: when overall provision of public goods is high, there are less incentives to migrate to the primate city. Controls for land and population are also significant and have the expected negative signs, implying that larger countries (both in size and population) have lower urban concentration. Another notable fact is that the coefficient for GDP per capita is not significant. (Henderson 2003) has demonstrated that the best degree of urban concentration for productivity growth maximisation varies with the level of development, measured by GDP per worker. My findings suggest that the level of economic development of a country affects primacy (not in the context of productivity growth) through the infrastructure, but not directly through income.

However, the negative primacy – amenity does not hold, when sanitation variable is considered. Apparently, the mechanics is more complicated in this case. One possible explanation would be the non-linear relationship between the two variables. The column (1) of the table 3 presents the results of a regression with a squared sanitation term included. Coefficients of both linear and squared sanitation terms are significant and have different signs, implying some type of the inverted-U relationship. Primacy reaches its global maximum when the level of urban sanitation is around 70%. In other words, these results demonstrate that in the countries, where improved sanitation facilities in urban areas are rather scarce, concentration of population in the primate cities is higher. The columns (2), (3), (4) and (5) of the table 3 demonstrate the results of the linear regression specification run for different samples. As countries develop better systems of sanitation provision in the cities, urban population concentration

Table 2: Urban primacy and public goods. Random Effects estimations.

	Primacy		
	(1)	(2)	(3)
electricity	-0.118 (0.000)***		
non-solid fuel		-0.035 (0.038)**	
sanitation			-0.054 (0.323)
land (1000 sq. km.)	-0.002 (0.000)***	-0.002 (0.000)***	-0.002 (0.000)***
population (mil.)	-0.018 (0.008)***	-0.020 (0.000)***	-0.000 (0.417)
GDP per capita	-0.000 (0.137)	-0.000 (0.501)	-0.000 (0.137)
R^2	0.23	0.17	0.18
number of obs.	552	280	3384
number of countries	143	141	142

All data is from the World DataBank. P-values in parenthesis.

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

goes down. Furthermore, the effect of this type of amenity is decreasing (both in size and significance), as countries move towards the universal provision of improved sanitation facilities in the cities. This partly suggests the theoretical model presented above. As sanitation provision rises, the incentives to migrate to the primate cities decrease. Finally, as the column (6) of the table 3 shows, above the 70% threshold improved sanitation facilities no longer have an effect on primacy. In terms of the theoretical model that would mean that the utility is already in the IIInd, flat part of the curve shown on the figure 4.

Why does access to electricity and non-solid fuel affect primacy in a slightly different way compared to improved sanitation facilities? One possible explanation to this question might be the specificity of use of these public goods. As was said above, access to electricity and non-solid fuel rather represent the capability of using these amenities, what can be referred to as a public good. However, the actual consumption is a perfectly excludable private good and has very little spill-over effects (if any), so individuals having no own access

to electricity or non-solid fuel cannot really mitigate their absence and have to move to the areas, where they will be able to use these public goods. On the contrast, improved sanitation facilities have positive externalities. Better sanitation not only brings utility to the users, but also makes the whole city cleaner also improving the life of non-users. In the context of the empirical findings presented above, when improved sanitation facilities are scarce (below 70%), individuals are willing to migrate to the cities with better infrastructure not only to have access to improved sanitation, but also to live in better sanitary conditions and safer disease environment. However, once improved sanitation facilities are better provided (above 70%), they have much smaller marginal effect on utility (hence, primacy), as citizens have less incentives to move to the better developed areas.

Table 3: Urban primacy and sanitation. Random Effects estimations.

	Primacy					
	(1) whole sample	(2) sanitation < 40	(3) < 50	(4) < 60	(5) ≤ 70	(6) > 70
sanitation (%)	-0.335 (0.040)**	-0.297 (0.020)**	-0.239 (0.017)**	-0.168 (0.033)**	-0.144 (0.095)*	0.012 (0.850)
(sanitation) ²	0.002 (0.027)**					
land (1000 sq. km.)	-0.002 (0.000)***	-0.004 (0.262)	-0.002 (0.968)	-0.008 (0.060)**	-0.004 (0.033)**	-0.002 (0.000)**
population (mil.)	-0.005 (0.354)	-0.109 (0.039)**	-0.059 (0.019)**	-0.006 (0.590)	-0.005 (0.545)	-0.005 (0.605)
GDP per capita	-0.000 (0.114)	-0.001 (0.664)	-0.000 (0.167)	-0.001 (0.031)**	0.000 (0.779)	-0.000 (0.135)
R^2	0.15	0.20	0.07	0.15	0.14	0.14
number of obs.	3384	576	740	940	1164	2217
number of countries	142	28	36	44	56	98

All data is from the World DataBank. P-values in parenthesis.

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

Table 4: Urban primacy and public expenditures

	Primacy
per capita expenditures (US \$1000)	-0.207 (0.066)*
land (1000 sq. km.)	-0.002 (0.000)***
population	-0.006 (0.078)*
R^2	0.16
number of obs.	5876
number of groups	142

All data is from the World DataBank. P-values in parenthesis.
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

As it was said above, besides measuring some particular types of public goods I could use the data for public expenditures as the indicator of provision of all amenities. Unfortunately, data at the city level is rarely available, so it is hard to make any judgements about the distribution of the funds across existing cities. Alternatively, it is possible to test if higher overall budget expenditures affect primacy. We can test the effect of government spendings on primacy analogously to the levels of certain public goods provision: higher overall public expenditures increase amenities and individuals find themselves in the IInd part of the utility curve shown on the figure 4 having less incentives to migrate to the primate city. The results of the estimations under this strategy are presented in the table 4. As expected, higher governments spendings are associated with lower primacy levels. Due to the multicollinearity issues, it is not correct to include GDP per capita as a control into this regression. However, the signs for the land and population coefficients are negative and significant, as it was expected. These findings also provide support for the theoretical model presented above. Larger funds spent on public goods increase the amenities-related part of utility and decrease the marginal utility. As a results, migration to the areas with better provision of public goods does not pay off, as the associated utility gain is not enough to compensate the resistance term ρ .

5 Conclusion

As it was already demonstrated in the literature, cities increase the productivity of their residents. Moreover, larger populations pay more taxes allowing city authorities to provide more public goods. Both income and city amenities increase utility of an individual, but at a decreasing rate. On the other hand, larger urban agglomerations are associated with higher congestion costs. As a result, there exist certain city sizes that maximise consumption of private goods and utility of the residents. It is important to notice that disposable income and overall utility reach its maximum at different values of the population. This paper presents a theoretical model of the whole urban systems. It allows to better explain the population distribution patterns within the country and links the utility levels of different urban agglomerations. Migrants moving within the country affect productivity and utility both in the cities they leave and the destination points. Moreover, such kind of migration might make both cities worse off.

The important feature of the model is that it considers not solely individual incomes and costs, but also includes the concepts of utility and public goods to the analysis. It allows us to better explain concentration of urban population in the largest cities. This can serve as another explanation for the existence of urban giants in the developing countries. Due to a lower level of public goods even slight variations in their provision cause significant utility differences across urban agglomerations and, hence, migration to the areas with better amenities. Another important feature of the model is the fact that utility-maximising individuals do not only chase higher incomes while making the migration decisions. Moreover, citizens can give up a share of their income in order to achieve higher overall utility level.

The model described in this paper is especially relevant for the case of developing countries. As less developed countries typically have lower levels of productivity and public goods provision, they are expected to be more sensitive to even minor deviations across the cities due to higher marginal utility of migration in the less-developed countries. Combined with lower migration costs, intercity disparities in wages and amenities cause stronger incentives to change the place of residence. As a result, regional inequalities can lead to the emergence of urban giants in poor countries.

The empirical testing of the theoretical model has demonstrated that provision of amenities in the urban areas and public expenditures have negative and

significant effect on population concentration. This fact does not directly prove the validity of the model, but supports it in an indirect way. The theoretical part implies that better provision of public goods lowers the marginal utility of each additional unit of amenities. Hence, individuals are less attracted to the primate cities, as the utility gains of this move are not high enough to trigger migration. Moreover, citizens are simply more likely to find access to the necessary amenities at their place of residence. However, when public goods are scarce, even slight improvements cause incentives to migrate to the primate city. As a result, the model predicts that countries with better amenities, measured both in units of particular infrastructural goods and aggregate public expenditures, are expected to have lower concentration of the urban population in the primate areas. Results presented in the section 4 tend to support this hypothesis: better access to electricity and non-solid fuel is associated with lower urban primacy rates. Access to improved sanitation facilities follow a slightly different pattern having a significant negative effect on urban population concentration until the coverage reaches the range around 70%. After this point access to improved sanitation does not have a significant effect on primacy. One possible explanation presented this pattern might be better possibilities to share the access to these facilities compared to electricity or non-solid fuel, so that universal coverage of sanitation is not required to satisfy the demand. Public expenditures rather indicate the aggregate amount of amenities offered in the country and are also negatively associated with urban primacy rates. Overall, the empirical results tend to support the theoretical model presented in this paper. Better amenities provision, measured both in coverage of particular public goods and aggregate public expenditures, is associated with lower urban primacy rates. Moreover, economic development affects concentration of urban population through the achieved level infrastructure, but not income, as implies from the fact that the coefficient for GDP per capita is not significant. As it was expected, land size and total population of the country affect urban primacy negatively.

This is a working paper and some changes are expected in the future. This especially applies to the empirical testing section. At the current stage the analysis is made at the aggregate level ignoring the distribution of public goods and funds across the country. As was said above, this type of data is available for a very limited number of countries. However, I can try to overcome this obstacle in two ways. First, I can employ data on fiscal decentralisation. In spite of the model presented in this paper, one would expect countries with more empowered local budgets to have lower primacy rates. When more public funds are

distributed at the regional level, amenities might be more evenly distributed across the country, causing no incentives to migrate to the primate cities. Alternatively, when the largest share of the national public expenditures is spent by the central government, it is hard to provide even and efficient allocation of public goods. The second way to capture the distribution of amenities across cities would be through luminosity data. Lights intensity can be also interpreted as the amount of public goods offered on the territory. Assuming this, we can compare the light intensity in the primate city to the rest of the country and try to see if the difference has effect on urban population distribution. These two additional strategies will be implemented in the later versions of the paper.

References

- Ades, A. F. & Glaeser, E. L. (1995), ‘Trade and circuses: Explaining urban giants’, *The Quarterly Journal of Economics* (110), 195–227.
- Au, C.-C. & Henderson, J. V. (2006), ‘Are chinese cities too small?’, *Review of Economic Studies* (73), 549–576.
- Bleakley, H. & Lin, J. (2012), ‘Portage and path dependence’, *The Quarterly Journal of Economics* **127**(2), 587–644.
- Davis, J. C. & Henderson, J. V. (2003), ‘Evidence on the political economy of the urbanization process’, *Journal of Urban Economics* (53), 98–125.
- Desmet, K. & Rossi-Hansberg, E. (2013), ‘Urban accounting and welfare’, *American Economic Review* **103**(6), 2296–2327.
- Duranton, G. & Puga, D. (2001), ‘Nursery cities: Urban diversity, process innovation, and the life cycle of products’, *American Economic Review* **91**, 1454–1477.
- Duranton, G. & Puga, D. (2004), *Handbook of Urban and Agglomeration Economics*, North-Holland, Amsterdam, chapter Micro-foundations of urban agglomeration economies, pp. 2003–2118.
- Fujita, M. (1989), *Urban economic theory: land use and city size*.
- Fujita, M., Krugman, P. & Venables, A. (1999), *The spatial economy: cities, regions, and international trade*.
- Gabaix, X. (1999), ‘Zipf’s law for cities: An explanation’, *The Quarterly Journal of Economics* **114**(3), 739–767.
- Glaeser, E. L. (2008), *Cities, Agglomeration and Spatial Equilibrium*, Oxford University Press.
- Gustavsson, R. (1999), ‘Explaining the phenomenon of third world urban giants’, *Journal of Economic Integration* **14**(4), 625–650.
- Henderson, J. V. (2003), ‘The urbanization process and economic growth: The so-what question’, *Journal of Economic Growth* (8), 47–71.
- Henderson, J. V. (2005), *Handbook of Urban and Agglomeration Economics*, North-Holland, Amsterdam, chapter Urbanization and Growth.

- Henderson, J. V. (2010), ‘Cities and development’, *Journal of regional science* **50**(1), 515–540.
- Henderson, J. V. & Wang, H. G. (2007), ‘Urbanization and city growth: the role of institutions’, *Regional Science and Urban Economics* **37**(3), 283–313.
URL: <http://www.econ.brown.edu/faculty/henderson/papers/Urbanization%20and%20City%20Growth0406%20revised%20-%20Hyoungh0906.pdf>
- Jedwab, R. & Vollrath, D. (2015), The mortality transition, malthusian dynamics, and the rise of poor mega-cities.
- Mohring, H. (1961), ‘Land values and measurement of highway benefits’, *Journal of Political Economy* (49), 236–249.
- Pirenne, H. (2014), *Medieval Cities: Their Origins and the Revival of Trade*, Princeton University Press.
- Rosen, K. T. & Resnick, M. (1980), ‘The size distribution of cities: An examination of the pareto law and primacy’, *Journal of Urban Economics* **8**(2), 165–186.
- Simon, H. A. (1955), ‘A behavioral model of rational choice’, *The Quarterly Journal of Economics* **69**(1), 99–118.
- Statistics Canada (2016), ‘Population and dwelling count highlight tables, 2011 census’.
URL: <http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hltfst/pd-pl/Table-Tableau.cfm?LANG=EngT=201S=3O=DRPP=150>
- Sveikauskas, L., Gowdy, J. & Funk, M. (1988), ‘Urban productivity: city size or industry size’, *Journal of Regional Science* **28 pages**(2), 185–202.
- The State Committee of Republic Uzbekistan on statistics (2015), ‘Uzbekistan in figures’.
URL: <http://www.stat.uz/en/publikatsii/65-informatsiya/publikatsii/150-uzbekistan-v-tsifrah>
- United Nations, Department of Economic and Social Affairs, Population Division (2015), *World Urbanization Prospects: The 2014 Revision*.

6 Appendix

6.1 Data sources

Regressions for urban access to electricity were run using the data from four years: 1990, 2000, 2010 and 2012; available for the following countries 143: Afghanistan, Algeria, Angola, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahrain, Bangladesh, Belarus, Belgium, Benin, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Burkina, Faso, Burundi, Cambodia, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Congo, Dem. Rep., Congo, Rep., Costa Rica, Cote d’Ivoire, Croatia, Cuba, Czech Republic, Denmark, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Eritrea, Estonia, Ethiopia, Finland, France, Gabon, Gambia, The Georgia, Germany, Ghana, Greece, Guatemala, Guinea, Guinea-Bissau, Haiti, Honduras, Hungary, India, Indonesia, Iran, Islamic Rep., Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea, Rep., Kuwait, Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Liberia, Libya, Lithuania, Macedonia, FYR, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mexico, Moldova, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Puerto Rico, Qatar, Romania, Russian Federation, Rwanda, Saudi Arabia, Senegal, Serbia, Sierra Leone, Slovak Republic, South Africa, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syrian Arab Republic, Tajikistan, Tanzania, Thailand, Togo, Tunisia, Turkey, Turkmenistan, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Venezuela, RB, Vietnam, Yemen, Rep., Zambia, Zimbabwe.

Regressions for urban access to non-solid fuel were run using the data from 2010 and 2012; available for the same set of countries as electricity data, except for Puerto Rico and Syrian Arab Republic.

Regressions for urban access to improved sanitation facilities were run using the data from 1990-2014 (25 years); available for the same countries as electricity data, except for New Zealand.

Regressions for per capita public expenditures were run using the data from 1961-2014 (54 years); available for the same countries as electricity data, except for Myanmar.

6.2 Robustness checks

Table A1: Urban primacy and lagged public goods. Random Effects estimations.

	Primacy		
	(1)	(2)	(3)
l.electricity	-0.104 (0.001)***		
l.non-solid fuel		-0.039 (0.031)**	
l.sanitation			-0.044 (0.434)
land (1000 sq. km.)	-0.002 (0.000)***	-0.002 (0.000)***	-0.002 (0.000)***
population (mil.)	-0.011 (0.012)**	-0.020 (0.000)***	-0.005 (0.345)
GDP per capita	-0.000 (0.166)	-0.000 (0.697)	-0.000 (0.148)
R^2	0.22	0.17	0.18
number of obs.	548	279	3256
number of countries	143	141	142

All data is from the World DataBank. P-values in parenthesis.

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

Table A2: Urban primacy and lagged sanitation. Random Effects estimations.

	Primacy					
	(1) whole sample	(2) sanitation < 40	(3) < 50	(4) < 60	(5) ≤ 70	(6) > 70
l.sanitation (%)	-0.288 (0.093)*	-0.286 (0.028)**	-0.222 (0.017)**	-0.148 (0.062)*	-0.125 (0.146)	0.023 (0.702)
(l.sanitation) ²	0.002 (0.060)*					
land (1000 sq. km.)	-0.002 (0.000)***	-0.004 (0.227)	-0.001 (0.974)	-0.008 (0.025)**	-0.004 (0.033)**	-0.002 (0.000)**
population (mil.)	-0.007 (0.296)	-0.109 (0.043)**	-0.059 (0.016)**	-0.007 (0.522)	-0.006 (0.486)	-0.005 (0.539)
GDP per capita	-0.000 (0.121)	-0.000 (0.753)	-0.000 (0.144)	-0.001 (0.028)**	0.000 (0.795)	-0.000 (0.139)
R^2	0.15	0.21	0.07	0.15	0.15	0.14
number of obs.	3256	551	710	904	1116	2140
number of countries	142	28	36	43	54	98

All data is from the World DataBank. P-values in parenthesis.

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

Table A3: Urban primacy and lagged public expenditures

	Primacy
per capita expenditures (US \$1000)	-0.205 (0.072)*
land (1000 sq. km.)	-0.002 (0.000)***
population	-0.006 (0.076)*
R^2	0.16
number of obs.	5823
number of groups	142

All data is from the World DataBank. P-values in parenthesis.

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