Topics in Macroeconomics

Dissertation in order to acquire the doctoral degree
from the Faculty of Economic Sciences,
at the Georg-August-Universität Göttingen

Submitted by

Viacheslav Yakubenko

Born in Leningrad, USSR

Göttingen, 2018

First Supervisor: Prof. Dr. Holger Strulik

Second Supervisor: Prof. Inmaculada Martínez-Zarzoso, Ph.D

Third Supervisor: Prof. Dr. Klaus Prettner

Date of oral examination: 02.11.2018

Contents

List of Figures

List of Tables

1 Giants and Midgets: the Effect of Public Goods' Provision on Urban Population Concentration

This paper analyses explains population distribution within a system of cities with availability of public goods in the country and provides an explanation for the existence of urban giants in developing economies. The article argues that (1) differences in public goods' provision cause excessive growth of the primate cities; (2) better provision of public goods at the national level decreases the difference between growth rates of the primate agglomeration and the hinterland; (3) consequently, low provision of public goods at the national level leads to the emergence of urban giants. Regression analysis suggests that lower provision of public goods at the national level causes excessive population growth of the largest city of the country, and a subsequent further growth of urban giants can potentially attract even more migrants to the primate cities and halt urbanisation process in the rest of the country. These findings are especially relevant for developing countries, where rapid urbanisation is currently taking place.

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

Keywords: city, public goods, growth, urbanization, primacy

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 world's population was living in urban areas, this number increased to half of the mankind by 2014. By 2050 two thirds of the world's population are expected to be urbanised (United Nations, 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 (United Nations, 2015). The connection between urbanisation rates and income levels is clear and not new to the literature. As countries move from agricultural to industrialised economies, more people move to the cities. However, why do we observe more *megacities* in developing countries than in developed ones, as United Nations (2015) indicate? This paper argues that disproportionate provision of public goods might lead to excessive growth of population in the largest (primate) cities and impede growth of the remaining urban structure.

Modern theory finds many reasons for cities to emerge. Spatial concentration decreases transaction costs, provides easier and cheaper access to production infrastructure and gives enterprises a larger pool of skilful employees (see Duranton and Puga (2004) for a more detailed overview of urban agglomeration economies). However, large cities are associated with higher pollution and crime levels simply due to their increased size. Besides that, higher rents and commuting costs, crowding and decreasing returns to scale might have negative influence on the economy (Mohring, 1961; Duranton and Turner, 2011; Glaeser, 2014). 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 (Henderson, 1974; Fujita, 1989; Fujita et al., 1999; Duranton and Puga, 2001; Henderson, 2005; Au and Henderson, 2006; Glaeser, 2008; Desmet and Rossi-Hansberg, 2013). It is obvious that neither too small, nor too large cities are desired. If a city is undersized, it does not fully employ urban economies and further population increase can make its residents better off. The opposite holds for a city that is too big: disadvantages outweigh overall urban benefits. However, the size of the primate city affects utility in other areas of the country, too (Ades and Glaeser, 1995; Davis and Henderson, 2003; Desmet and Rossi-Hansberg, 2013; Glaeser, 2014). This paper analyses how the degree of population concentration in the largest city affects population distribution in the whole country. Understanding this mechanism is important for more efficient and equal economic growth for the residents of all areas of the country. Countries currently experiencing the urbanisation process can prevent the emergence of urban giants and attain higher aggregate utility level, if public resources are distributed equally across locations (Henderson, 2003). Alternatively, over-proportionate concentration of people in one location can lead to utility dead-weight losses and substantial share of resources turns out to be wasted.

This article demonstrates how urban giants can distort urbanisation in the whole nation. First, the paper presents a simple theoretical application that describes the relationship between public goods and a city size. If a primate city can offer substantially more amenities than the rest of the country, it attracts additional migrants from other areas or ensures relatively higher population growth rate of the local residents. This process diverts potential migrants to other urban areas and impedes population growth in the hinterland cities, so we can argue that the excessive growth of primate cities happens at the expense of the hinterland. Moreover, if we allow for positive correlation between population size and the amount of public goods in the location (e.g., due to more disposable tax funds), we can expect a self-inducing growth of the primate agglomeration. This process might lead to the emergence of urban giants that co-exist with immensely small towns in the same country. Such an unbalanced distribution of population across locations can entail dead-weight utility losses for the whole nation: congestion in the primate city (e.g., higher rents, crime, etc.) harms potentially higher utility from increased productivity and additional public goods, while the lack of residents in smaller towns constrains productivity spillovers and infrastructure development. This theoretical argument is supported by empirical evidence obtained using estimates from a panel of 88 countries followed for 25 years (1990-2014). To obtain robust results in the presence of endogeneity I employ Fixed Effects and 2SLS estimators. Empirical findings support the predictions of the theoretical application: lower provision of public goods in the country is positively associated with an excessive population growth of the primate cities making them grow faster than the rest of the country and turning them into urban giants.

Given that cities foster individual productivity, we can expect urbanisation to be positively associated with economic development. Income rises when countries move from agriculture to industrial production and this process requires concentration of workers in the cities (Harris and Todaro, 1970; Caselli, 2005; Caselli and Feyrer, 2007; Lagakos and Waugh, 2013; Glaeser, 2014; Gollin et al., 2016). On the other hand, spatial concentration of workers fosters innovations and, hence, further productivity growth (Duranton and Puga, 2001; Henderson, 2010; Michaels et al., 2012; Desmet and Rossi-Hansberg, 2014). As a result, we can expect a self-fueling process of urbanisation-induced growth. However, this appears not necessarily true for large, but poor cities in developing countries (Fay and Opal, 2000; Jedwab and Vollrath, 2015, 2018; Glaeser and Henderson, 2017; Chauvin et al., 2017). 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 (Jedwab and Vollrath, 2018). A range of studies provides different explanations for the emergence of urban giants with a particular focus on developing countries: e.g., Ades and Glaeser (1995) and Davis and Henderson (2003) highlight the importance of political factors; Gustavsson (1999) provides evidence for trade barriers increasing the concentration in the third-world urban centres; Jedwab and Vollrath (2015) and Jedwab et al. (2017) demonstrate that the demographic transition is largely responsible for the population growth of urban giants; finally, Desmet and Rossi-Hansberg (2013) show that higher efficiency and better amenities lead to larger cities. This paper emphasises the importance of public amenities' distribution across the whole country for the emergence of large cities in developing countries.

Leaving aside the differences between rural and urban income, population distribution across the existing urban areas might have a significant effect on the economy. Every country has its primate city, but degrees of spatial concentration can vary substantially. First, urban concentration has a positive effect on individual productivity for the whole nation, but after a certain point further concentration of workers in one city can harm the economy due to inefficient allocation of resources (Williamson, 1965; Henderson, 2003; Bruelhart and Sbergami, 2009; Castells-Quintana and Royuela, 2014). Knowing the effect of urban concentration on growth, it is important to explain which factors determine the degree of spatial concentration. Besides productivity increase, city sizes can be affected by other factors: political reasons (Ades and Glaeser, 1995; Davis and Henderson, 2003; Henderson and Wang, 2007; Chen et al., 2017), better access to trade infrastructure (Glaeser, 2014; Henderson et al., 2018), natural disasters (Ager et al., 2015; Kocornik-Mina et al., 2015) and climate (Barrios et al., 2006; Henderson et al., 2017). This paper contributes to the literature emphasising the importance of local amenities for locational choices (Knapp and Gravest, 1989; Brueckner et al., 1999; Chen and Rosenthal, 2008; Lall et al., 2009; Albouy, 2016; Castells-Quintana, 2017) by drawing particular attention to water and sanitation facilities. This paper considers the access of population to improved water and sanitation facilities due to the number of reasons. First of all, water and sanitation are crucial for human well-being. Secondly, concentrating on the *access* instead of actual consumption, we can assume these amenities to be as close to the theoretical definition of public goods, as possible. Finally, it is possible to obtain data for a wide pool of countries. The choice of water and sanitation as proxies for public goods will be explained more thoroughly in Section 2.

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) to shape urbanisation optimally (Ades and Glaeser, 1995; Au and Henderson, 2006; Henderson and Wang, 2007). However, such an interference might be rather harmful for the country as a whole. Despite higher utility from better public goods in the primate city, congestion also rises, while the rest of the country can face decline in population and subsequent decrease in public funds. It is important to understand that improving public infrastructure in the largest city of the economy can be a very cost-efficient intervention in the short term: providing amenities that improve utility of a large number of people. However, in a longer perspective subsequent excessive population growth increases congestion in the primate city and constrains the development of the hinterland lowering utility in both locations. As a result, the beneficial effect of public amenities improvement in the capital might be substantially decreased or even outweighed by congestion associated with excessive population, while the rest of the country has not benefited from it at all. An alternative way to distribute funds is to spread investments equally across locations. This might significantly improve the aggregate well-being and provide basis for more equal and sustainable growth.

This paper is organised as follows. After the introduction I describe the theoretical application that links population size and public goods in different locations of the country, present how I measure public goods and excessive population, and, finally explain the empirical methodology. The third section presents the regression results. The conclusion section summarises the findings of the paper.

1.1 Empirical strategy

1.1.1 Theoretical application

In this section I present a simple theoretical application to demonstrate the connection between the relative population size of a city[1], public goods provision and migration. The main aim is not to derive particular optimal conditions, but rather illustrate the complex relationship between the accessibility of basic amenities provided in the urban agglomeration and the number of citizens to put some structure and motivate the empirical part. For this reason I employ the theoretical framework from Galiani and Kim (2011).

The model considers two regions: the mainland (i.e., the primate city) and the hinterland. In both locations individuals provide one unit of labour for production of homogeneous output according to the following function:

$$X_i = A_i L_i^{\alpha} G_i^{1-\alpha}, \quad \alpha \in (0,1), \tag{1.1}$$

where A_i is the local productivity parameter, L_i is the population and G_i is the amount of public goods in a location i. Given only two regions in the model i takes values of C for the primate city and H for the hinterland. Having prices of output normalised to 1, the profit maximisation problem will yield the real wage w_i:

$$w_i = \alpha A_i \left(\frac{G_i}{L_i}\right)^{1-\alpha}. \tag{1.2}$$

The real wage rate in Eq.(1.2) is decreasing in population, reflecting congestion effects (Mohring, 1961; Troesken, 2004; Henderson, 2005; Duranton and Turner, 2011; Desmet and Rossi-Hansberg, 2013). However, larger areas can still be attractive for the residents due to higher productivity or the larger variety of public goods provided there. Assuming unconstrained migration between the locations, we can expect the real wages to equalise in both locations:

$$w_C(A_C, L_C, G_C) = w_H(A_H, L_H, G_H).$$

Employing Eq.(1.2) we can derive the relative size of the primate city C as a function of public goods and productivity in both locations:

$$\frac{L_C}{L_H} = \frac{G_C}{G_H}\left(\frac{A_C}{A_H}\right)^{\frac{1}{1-\alpha}}. \tag{1.3}$$

Unlike Galiani and Kim (2011) I deliberately do not model public goods endogenously to

[1] It is necessary to define the term "*city*" as it is used in this paper. Modern literature usually refers to "*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".

avoid any assumptions about the determinants of their distribution. In my interpretation of Eq.(1.3) provision of particular public amenities is a result of a complicated political process that I take as exogenously given[2]. In other words, I do not aim to explain why public goods are distributed this way, but analyse how this allocation affects population distribution. Moreover, modelling public goods as a function of the population size, as it was done by Galiani and Kim (2011), is not necessarily appropriate for the cases of access to water and sanitation that are considered in the current article. As we consider the share of population having access to a particular amenity, the effect of population is dual: on the one hand, more people can provide more funds for construction of required infrastructure, on the other, higher congestion levels constrain access to the existing infrastructure[3].

The main purpose of the theoretical part and the paper as a whole is to demonstrate that the level of public goods' provision at the national level negatively affects the relative size of the primate city. For empirical testing of this relationship I consider L_C as a number of residents in the largest city of the economy, and L_H is the population in other urban agglomerations. As Eq.(1.3) demonstrates, the relative size of the primate city is expected to be negatively affected by the amount of public goods in the hinterland.

In other words, a decrease in G_H has to be accompanied by the growth of the relative size of the primate city. There are two reasons for this. First of all, amenity-seeking agents can migrate to the largest city, where public goods are expected to be better provided (as will be demonstrated later). Second, lower level of public goods provision can entail higher mortality or lower fertility rates in the hinterland. This relationship holds even if G_C and G_H are completely independent from each other. If we allow for negative correlation between the two characters (e.g., public amenities in both regions are funded from one limited source), this effect will be only amplified. Besides public goods, population distribution is expected to be affected by local productivity parameters. Unfortunately, there exist no data of this type. However, I assume that productivity ratio does not change over 25 years considered in the study and can be controlled for by inclusion of country fixed effects.

1.1.2 Measuring public goods

This paper does not rule out the possibility that other factors, like productivity differences, affect migration, but focuses on the effect of particular amenities. Clearly, we cannot measure all public goods provided in a city. For this reason I need to find a set of proxies for local amenities or concentrate on those public goods that are utmost important for individuals. In this paper I choose shares of national population with access to improved water source and sanitation as the variables indicating public goods provision. First of all, these amenities are simply important for survival and further economic development (Cain and Rotella, 2001; Troesken, 2004; Costa and Kahn, 2006; Ashraf et al., 2016; Beach et al., 2016), so they might have more significant effect on migration decisions than, for example, reputation of the local opera theatre. If we refer to the Maslow's hierarchy of

[2]Existing studies typically explain provision of public goods by political favouritism (Ades and Glaeser, 1995; Henderson and Wang, 2007; Chen et al., 2017). However, applying similar patterns for cross-country analysis might be misleading, and does not add predicting power for the purposes of the model.

[3]Finally, possible reverse causality will be controlled for using instrumental variables strategy described later.

needs (Maslow, 1943), water and sanitation serve the “physiological” needs category, hence, typically individuals will first aim to fulfil these necessities. Public goods like healthcare or education correspond to “higher” levels of the hierarchy (the safety needs for healthcare and the esteem or self-actualisation for education). Inclusion of public goods fulfilling the needs of higher orders can help to explain migration decisions even better, especially in the countries at later stages of development. However, having developing countries in the centre of this research, “physiological” needs can still be a good predictor of amenity-seeking migration. Secondly, access to water and sanitation is, probably, the closest example to the formal definition of public goods (Mas-Colell et al., 1995). Given that the variables measure access and not actual consumption, we can, at least, assume some effects of their availability to be non-rivalrous and non-excludable. Even if an individual has no access to water or sanitation himself (or has to compete for it), the use of these facilities by the neighbours increases his utility through cleaner neighbourhood and better disease environment. Finally, data are available for a wide range of countries. In other words, I have the data for G_H from Eq.(1.3) for a large number countries and years and can investigate how these values affect population distribution.

It is important to note that access to improved water and sanitation facilities is quite likely to be a good proxy for other amenities, especially at the early stages of development. The two variables are highly correlated with each other and can be very well predicted by public per-capita expenditures, as will be demonstrated later. Given this, we can expect that if water and sanitation are well-provided, other public goods might be, too. The Demographic and Health Surveys Program (DHS) provides data on access to a set of amenities that to some degree can be considered public goods. Variables like share of population living in households with an improved water source or improved, non-shared toilet facilities can be considered rather infrastructural amenities. Net primary and secondary school attendance rate serve as proxies for education. Even though schooling can be commonly paid privately in some countries, public expenditures are still considered extremely important for efficient functioning of educational systems, especially in developing countries (UNESCO, 2015). The same holds for healthcare: without public spending developing countries cannot achieve universal access to this type of facilities (World Health Organisation, 2010). Health-related public goods are proxied by the shares of women that received antenatal care from a skilled provider and had assistance during delivery from a skilled provider. Table 1.1 demonstrates that correlation coefficients between the six variables are quite high. In other words, values of access to water and sanitation can be also considered as an indication of general level of public goods provision in the country.

It is important to clarify whether residents of big cities have better access to public goods. Unfortunately, data on public goods provision in particular cities are scarce. The World Bank data suggest that developed countries have close-to-universal access to water and sanitation at the national level, but we know little about the distribution of these amenities in the countries where provision is substantially below 100%. However, Urban Data from UN-Habitat programme and the DHS Program provide figures on shares of population in a particular location with access to some basic amenities available in different regions of the country. In Table 1.2 I compare provision of public goods in primate cities and other regions of the countries. As the results of a paired t-test of means suggest, in a given sample all of the mentioned proxies for public amenities are on average better provided in the largest cities compared to the rest of the country.

Table 1.1: Correlation coefficients between water, sanitation and other public goods

	Infrastructure		Education		Healthcare	
	water	toilet	primary	secondary	prenatal	birth attended
water	1.0000					
	(9941)					
toilet	0.4082	1.0000				
	(8661)	(8837)				
primary	0.4777	0.2197	1.0000			
	(5317)	(5301)	(5493)			
secondary	0.6145	0.5749	0.4919	1.0000		
	(5317)	(5301)	(5493)	(5493)		
prenatal	0.4032	0.1166	0.4598	0.2779	1.0000	
	(9475)	(8533)	(5473)	(5473)	(9651)	
birth attended	0.6266	0.4202	0.5221	0.6684	0.6512	1.0000
	(8936)	(7870)	(5424)	(5424)	(8986)	(9112)

Number of observations is in parentheses. Data are at the level of within-country regions. All correlations are statistically significant at 1%-level. The list of countries is included in the Appendix. All data used are from the DHS Program.

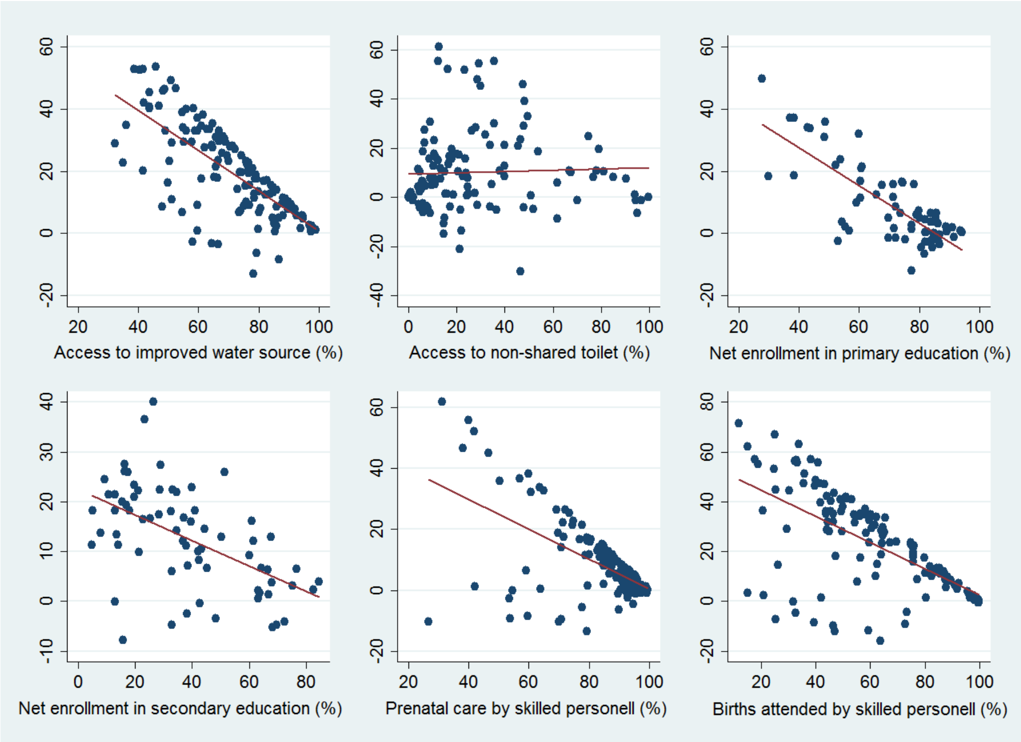
Table 1.2: Access to public goods in the primate cities and other regions of the country

	mean level of access in:			
Public good	primate city	other regions	difference	N of obs. (countries)
Panel A: Data from UN-Habitat programme				
Improved water	90.527	85.827^a	-4.7***	49(48)
Improved sanitation	78.420	66.835^a	-11.586***	49(48)
Panel B: Data from the DHS Program				
Improved water	90.453	70.941	-19.512***	136(56)
Improved non-shared toilet	37.678	27.616	-10.062***	122(56)
Primary education	79.884	71.232	-8.652***	69(43)
Secondary education	50.967	38.272	-12.694***	69(43)
Prenatal care	92.143	83.657	-8.486***	137(57)
Birth attended	83.752	60.363	-23.389***	127(55)

"Other regions" category is defined as "'other cities/towns" in the UN-Habitat data, and "Total" in the DHS data. *** $p < 0.01$

This stylised fact supports the following assumption: when public goods are not universally provided, their level of provision is typically higher in the biggest urban agglomeration of the country. Even if lower level of public amenities provision nationally can be explained by lower urbanisation rate, the disparity between the primate and other cities cannot be explained by this factor.

Figure 1.1: Provision of public goods in primate city and hinterland



Note: X-axis: total level of provision of a respective public good; Y-axis: difference between the level of provision in the primate city and the total level. Red lines show linearly fitted values. Sources: the DHS Program and author's calculations.

Finally, I can demonstrate how the provision of public goods nationwide correlates with the difference between the level of provision in the primate city and the total country level. Figure 1 demonstrates that as countries move towards universal provision of public goods, the disparity between the primate city and the hinterland shrinks. This finding supports the hypothesis that in countries with less developed infrastructure incentives to migrate to the primate city can be higher compared to relatively more developed economies, as the utility gain associated with migration is expected to be high.

1.1.3 Measuring excessive population

Similar to the data on public goods distribution, large-scale information on migration or region-specific mortality and fertility is rarely available. For this reason I construct "synthetic" populations of the primate cities. Population size data are available for the whole nation, total urban areas and the primate cities for a wide range of countries. Subtracting the size of the primate city from the overall urban population yields the size of all urban agglomerations except for the largest one. Then, we can calculate the annual growth rate of this estimated variable to see how an average (non-primate) city in the country was evolving over time. Afterwards, I take 1990 as the initial period ($t = 0$ in terms of the model) and apply the calculated growth rate year by year to the size of the primate city.

Hence, we can evaluate *how the primate city would have grown, if it was just an average city*[4]. Next, I subtract the simulated size of the primate city from its real size to obtain the "excessive population" of the largest urban area. Finally, to use this number for analysis, one needs to scale it. Larger countries simply have more citizens that can potentially migrate compared to the small ones. To control for this scale difference of various countries I divide the calculated "excessive population" ($\hat{\mu}_{B,it}$) by the actual size of the national population (N_{it}):

$$m_{it} = \frac{\hat{\mu}_{B,it}}{N_{it.}} \tag{1.4}$$

As a result, m_{it} indicates the share of national population of country i that turned out in the primate city in year t, additionally to the amount that is expected to emerge in an average urban area given a set of country characteristics (e.g., urbanisation rate, level of economic development, etc.). Of course, public goods can contribute to higher growth rates in the largest agglomerations not only with migration, but also with better survival rates (e.g., due to better disease environment). However, this does not undermine the explanatory power of m_{it}. Moreover, it amplifies the migration argument: it is natural to believe that people try to migrate to areas where they or their future offspring has better chances of survival. In other words, we can consider m_{it} as the share of total national population that migrated, was successfully born or avoided death in the primate city due to better provision of public goods there.

1.1.4 Empirical methodology

The theoretical application presented above predicts that primate cities can attract more migrants and ensure higher population growth rate than other locations in the country, especially if we consider developing states with low national level of provision of public goods. At the same time, an increase in the size of the primate urban agglomeration allows provision of more public goods and can intensify further population growth. Moreover, when citizens abandon a location of origin, it might reduce disposable funds of the local government (for example, cut in central government subsidies), so provision of public goods in this hinterland region decreases. As a result, we can see a self-fuelling migration process that can lead to the emergence of urban giants and depopulation or slower growth of hinterland areas. The empirical section of this paper suggests that: (1) provision of public goods at the national level is significantly negatively associated with the excessive population share in the largest city in the economy; (2) population concentration in the largest city is positively and significantly correlated with excessive population share; (3) excessive population increases the size of the primate city (both absolute and relative).

The baseline regression specification is as follows:

$$m_{it} = \beta_0 + \beta_1 \bar{\theta}_{it} + \beta_2 \Omega_{it} + \lambda_t + u_i + \varepsilon_{it}, \tag{1.5}$$

where $\bar{\theta}_{it}$ is the national level of public goods provision, λ_t and u_i are year and country fixed effects, respectively and ε_{it} is a residual term. It is important to note that m_{it} is

[4] This is especially important for developing countries, where urbanisation actively takes place. Alternatively, one could have applied the national population growth rate. However, this would not take into account migration from rural areas that naturally happens at this stage of economic development.

calculated using the difference between the values of population sizes in periods t and $t+1$. In other words, it indicates the excessive population share that emerged in the primate city during the period t. That is why I add controls for the period t.

Ω_{it} is a set of time-varying country characteristics that might affect migration. First of all, I need to control for the size of the primate city. As the evidence above suggests, bigger cities can provide more public goods, hence, are more attractive for migrants compared to smaller locations. I can control for the actual size of the primate city, but if other cities in the country are only marginally smaller this approach might be not very useful. Moreover, the city size per se might be misleading, if we do cross-country comparisons. New York agglomeration has a population above 20 million people making it one of the largest cities in the world. However, the US is a highly urbanised country and the existence of other large agglomerations, like Los Angeles (15.1 mil.) or Chicago (9.1 mil.) (Demographia, 2016) makes New York much less of a giant, with only of 6% of the total U.S. population residing in the city. On the contrary, Luanda had around 7 million inhabitants in 2016 according to Demographia (2016). Nevertheless, more than 20% of the total population was residing in Angola's capital. To compare how much a particular primate city dominates other areas, it makes sense to normalise its size by the total population of a country. In other words, employing relative size of the biggest cities allows more accurate comparisons between the countries at different stages of economic development.

Other controls included in Ω are national population size, absolute size of the primate agglomeration, speed of urbanisation, population density and log of per capita GDP. m_{it} already accounts for the fact that larger populations can provide more migrants. However, population size can still affect migration through other ways: the more people reside in the country, the more likely is the existence of other large agglomerations, that might be worth migrating to. Absolute size of the largest urban agglomeration is added to control if migrants are attracted by the absolute size of the primate city instead of the relative one. Furthermore, if we consider urbanisation as a process of migration from rural to urban areas, intensiveness of this process might affect excessive migration to the primate city. To control for this effect I include speed of urbanisation (growth rate of the share of population living in urban areas) into the regression. Another control variable is population density, that shows how many people on average reside per 1 squared kilometre. Higher density might ease migration to the primate city, as potential migrants might face lower relocation costs due to shorter distances they have to cover. Finally, I include natural logarithm of per capita GDP into the regression analysis. If individuals lack some public goods, in locations where they currently reside, higher income might allow them to get it privately instead of undertaking the move to the primate city or dying without it.

It is important to consider that access to water and sanitation, even measured at the country level, can be affected by population concentration in the largest city. If citizens are concentrated in one city, it might be easier to provide them with water and sanitation, as the local government has to only expand existing infrastructure rather than build if from scratch. According to this argument, share of total population living in the primate city can be positively associated with public goods provision at the national level. On the other hand, if population is concentrated in the largest city, hinterland areas might lack funds to develop necessary infrastructure. Following this logic, we can expect population concentration to be negatively associated with the share of population that has access to water and sanitation. Overall, the share of population living in the primate city can affect the water and sanitation access variables, raising problem of multicollinearity. However, as Table A1.2 demonstrates, the correlation coefficient between the share of national population living in the largest city and water and sanitation is 0.39 and 0.46, respectively, way below the rule of thumb level of 0.8 suggested by Gujarati (2003, p. 359).

Furthermore, I must ensure that provision of public goods is not affected by the share of excessive population in the primate city. An increase in the city size can provide the local governments with more funds for the provision of public goods. Again, this improves access to water and sanitation in the primate city, but decreases it in the hinterland, so the effect on the national level of provision is unambiguous. Nevertheless, this indicates an endogeneity problem that can bias the estimated coefficients. For this reason I employ an instrumental variable (IV) estimation strategy. I choose per capita governmental expenditures as an instrument for the share of population with access to water and sanitation at the national level. I assume public expenditures at the country level to be not determined by water and sanitation provision, i.e. the causality runs only one way. The potential effect of these particular public amenities is indirect and is likely to be diluted by more powerful determinants, such as GDP per capita or institutional quality. Furthermore, *average* public expenditures at the national level can be assumed to be not driven by population concentration or migration flows. Even if capitals get disproportionately more funds per one citizen at the expense of hinterland areas, this will not affect the mean level of governmental expenditures.

After demonstrating how public goods and population concentration in the largest city affect excessive population growth of the primate urban agglomeration, I can analyse the effect of m_{it} on the size of the largest city in the country. In other words I try to analyse if this *additional* source of population in the primate city is substantial for an already big urban agglomeration. For this I will estimate the following regression equation:

$$\text{primate city size}_{it} = \alpha_0 + \alpha_1 m_{it} + \alpha_2 \Omega_{it} + \kappa_t + v_i + \eta_{it}, \tag{1.6}$$

where Ω_{it} is the same set of controls, as in Eq.(1.5), but extended by the share of population living in urban areas, κ_t and v_i are time and country fixed effects, respectively, η_{it} is a residual term.

It is important to note that population concentration is not the same as the actual size of the largest city. Large primate agglomeration does not universally imply that the major share of population resides there. Moreover, Table A1.2 shows that the absolute size of the primate city and the share of national population living in it are rather independent from each other with the coefficient of correlation close to -0.01, not significant at 10-% level. As the estimation results in the next section suggest, excessive population share in the primate city is determined by its *relative* and not *absolute size*, if we control for other driving factors. As a result, I conclude that I found no significant evidence of the presence of reverse causality between excessive population share in the primate city and its absolute size.

All data used are from the The World Bank (2018b). I limit the sample to only complete observations over the period between 1991 and 2014, so that the panel is strongly balanced. The 88 countries that are included in the analysis are listed in the Appendix. In 2014 the total number of people living in the countries included into the sample was 6315 million people (87% of the world's population, according to (The World Bank, 2018b)). A brief overview of the analysed variables is presented in Table A1.1. Due to the fact that water and sanitation provision are highly correlated with each other (correlation coefficient 0.88, statistically significant at 1-% level), including them into one regression will induce multicollinearity. For this reason I run separated regressions for water and sanitation. As will be seen later, choice of a proxy for public goods provision does not change the general intuition.

1.2 Estimation results

1.2.1 Public goods provision and excessive population

First of all, I need to demonstrate that the level of public goods provision has significant effect on excessive population share in the largest urban agglomeration in the economy. Tables 1.3 and 1.4 show that access to improved sanitation and water facilities have significant negative effect on excessive population in the primate city. This finding supports the hypothesis that the higher is the aggregate level of public goods provision, the more likely the citizens to find them in the location they currently live in. As a result, availability of public goods reduces incentives to migrate to the largest city for the representative individual living outside the primate city and decreases mortality in the hinterland. If we compare the coefficients for sanitation and water access, we can notice that the former variable has a bigger effect, both in terms of magnitude and significance. An increase in the share of aggregate population with access to improved sanitation facilities by one standard deviation on average results into a decrease of excessive population share by one standard deviation. At the same time, improving access to water by one standard deviation decreases excessive population share by only one third of a standard deviation. This finding can be explained the following way. Firstly, on average access to water is substantially better provided in the sample than to sanitation, as Table A1.1 indicates. Moreover, a paired t-Test of the equality of means yields a p-value of 0.000, indicating that the difference between the two variables is highly significant. Given this, excessive population is expected to be less sensitive to the access to a water source, as generally chances to get it are quite high throughout the whole country (hence, lower migration incentives) and the average level is enough to survive even in the hinterland. On the contrary, improved sanitation facilities are comparatively rare, so migrating to the primate city, where public infrastructure is expected to be better developed, can bring high utility gains and provide way healthier disease environment. These results are robust to the inclusion and exclusion of additional controls, as Tables 1.3 and 1.4 demonstrate.

Another important question that needs to be tested: does an increased size of a city allows provision of more public goods? Consequently, we can expect that an increased concentration of national population in the largest urban agglomeration intensifies the migration process to this city and increases natural population growth rate. The share of national population residing in the primate city has significantly positive effect on excessive population variable. The size of the coefficient does not vary much between Tables 1.3 and 1.4. More importantly, the magnitude of this effect is very strong: one standard deviation increase in national population concentration in the primate city translates into 2.7 standard deviations stronger excessive migration inflow. The absolute size of the primate city is positively associated with excessive population, but this variable loses its significance, if we add other controls. Overall, the estimation results suggest that the bigger the primate city compared to other locations in the country, the more public goods can be expected there, so we can expect the self-fuelling growth process to take place. This can be a sign of a reverse causality problem, that will be controlled for by the IV estimation later in the paper.

Table 1.3: Sanitation provision and excessive population.

	Depedent variable: m_{it}					
	(1)	(2)	(3)	(4)	(5)	(6)
sanitation	-0.100**	-0.124***	-0.091**	-0.105**	-0.134***	-0.126***
	(0.044)	(0.046)	(0.040)	(0.043)	(0.050)	(0.044)
share of pop. in primate	1.057***	1.077***	1.066***	1.057***	0.980***	1.023***
	(0.166)	(0.161)	(0.159)	(0.164)	(0.168)	(0.176)
nat. population		0.017***				0.007
		(0.006)				(0.006)
population density		0.002				0.003
		(0.003)				(0.002)
speed of urbanisation			0.251			0.204
			(0.197)			(0.194)
ln(GDP per capita)				0.603*		0.448
				(0.326)		(0.363)
abs. size of primate					0.387**	0.225
					(0.155)	(0.215)
year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2112	2112	2024	2112	2112	2024
Number of countries	88	88	88	88	88	88
adj. R-squared	0.384	0.408	0.393	0.390	0.413	0.424
within R-squared	0.391	0.416	0.400	0.398	0.420	0.433

Fixed effects regressions. All regressions include a constant term. Data are from The World Bank (2018b). Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Summing up the results of the first set of regressions (described by Eq. (1.5)), I can conclude that increased provision of public goods at the national level has a significant negative correlation on excessive population share in the primate urban agglomeration. The results support the hypothesis that availability of access to improved sanitation and water source decreases incentives to migrate to the largest city and provide conditions for a higher population growth in the hinterland areas. Furthermore, relative size of the primate city has significant positive effect on excessive population share. This finding supports the hypothesis that increased concentration of citizens in the largest urban agglomeration intensifies further excessive population growth and can lead to a slower growth or decrease in sizes (both absolute and relative) of other locations in the country.

Table 1.4: Water access and excessive population.

	Depedent variable: m_{it}					
	(1)	(2)	(3)	(4)	(5)	(6)
water	-0.057*	-0.069**	-0.057*	-0.060*	-0.062**	-0.070**
	(0.030)	(0.031)	(0.031)	(0.031)	(0.030)	(0.032)
share of pop. in primate	1.053***	1.069***	1.071***	1.053***	1.003***	1.073***
	(0.179)	(0.179)	(0.165)	(0.178)	(0.188)	(0.191)
nat. population		0.013***				0.010
		(0.004)				(0.007)
population density		0.002				0.003
		(0.003)				(0.003)
speed of urbanisation			0.326			0.324
			(0.198)			(0.199)
ln(GDP per capita)				0.499		0.482
				(0.314)		(0.365)
abs. size of primate					0.232*	0.026
					(0.128)	(0.199)
year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2112	2112	2024	2112	2112	2024
Number of countries	88	88	88	88	88	88
adj. R-squared	0.362	0.377	0.378	0.366	0.373	0.397
within R-squared	0.369	0.385	0.385	0.374	0.381	0.406

Fixed effects regressions. All regressions include a constant term. Data are from The World Bank (2018b). Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.2.2 Excessive population and size of primate city

After demonstrating that increased population concentration is positively correlated with excessive population share in the primate city, I can show what effect migrants have on city sizes. It is straightforward that higher immigration to the city, higher birth and lower mortality rates are positively associated with its size. However, the effect of excessive population on the city size is not necessarily strong and significant by construction. Moreover, if the country in general shows a high population growth rate or urbanisation is currently actively taking place, excessive population share may play a very little or absolutely no role in explaining the primate city size. For this reason I run a set of regressions that demonstrate the significance of migration as a determinant of primate city sizes. First, I present the result of the estimation of Eq.(1.6) with the absolute size of the primate city as a dependent variable. The results are presented in Table 1.5.

Coefficients for the excessive population share in Table 1.5 indicate a positive and significant correlation between the variable and the absolute size of the primate urban agglomeration. The effect is quantitatively strong, as one standard deviation increase in excessive population share on average increases the population of the primate city by 300.000 inhabitants. Furthermore, this correlation is robust to the inclusion of additional controls, especially the

Table 1.5: Excessive population and absolute primate city size.

	Depedent variable: *absolute size of primate city*					
	(1)	(2)	(3)	(4)	(5)	(6)
m_{it}	0.122***	0.097***	0.120***	0.203***	0.112***	0.150***
	(0.037)	(0.031)	(0.037)	(0.046)	(0.037)	(0.031)
nat. population		0.039***				0.035***
		(0.011)				(0.009)
population density		0.000				0.001
		(0.002)				(0.002)
speed of urbanisation			0.047			0.065
			(0.056)			(0.042)
% of pop. in urban				0.164***		0.114***
				(0.057)		(0.033)
ln(GDP per capita)					0.755	0.122
					(0.498)	(0.187)
year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2112	2112	2112	2112	2112	2112
Number of countries	88	88	88	88	88	88
adj. R-squared	0.388	0.693	0.388	0.486	0.413	0.740
within R-squared	0.395	0.696	0.395	0.492	0.420	0.744

Fixed effects regressions. All regressions include a constant term. Data are from The World Bank (2018b). Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

size and share of national population living in urban areas. Obviously, small countries can simply lack people to inhabit large cities, so we can expect large urban agglomerations to be located in bigger countries. The same holds for urbanisation rate: if a country is mostly rural, people have not yet migrated to the cities (including the primate one). Nevertheless, inclusion of these two variables into the regression does not decrease the magnitude of the migration coefficient and still keeps it significant.

However, significance of the coefficients for national population and urbanisation rate in Table 1.5 suggest that comparing absolute sizes of the primate cities might be misleading, especially in the case of countries, where urbanisation process is actively ongoing, as was mentioned before. Furthermore, if public goods in a primate city are provided at the expense of hinterland areas, excessive population growth it the primate city might constraint the growth of other locations. To analyse the effect of excessive population share in the largest urban agglomeration on the urbanisation process in the whole country I can scale the size of the primate agglomeration by the size of other urban areas. For this purpose I use the urban primacy variable: the share of total urban population living in the largest city[5]. To see the effect of excessive population in the primate city on the relative size of the

[5] Henderson (2003) mentions two alternative concentration indicators: Herfindahl-Hirschman Index and Pareto parameter. However, urban primacy is highly correlated with these measures and does not require as much data for calculation as the two other indicators, what allows to carry out the analysis for a larger set of country-years. See Rosen and Resnick (1980), Gabaix (1999) and Henderson (2003)

largest urban agglomeration I run the same regression specification as described in Eq.(1.6) using urban primacy as a dependent variable. The estimation results are presented in Table 1.6.

Table 1.6: Excessive population and relative primate city size.

	Depedent variable: *urban primacy*					
	(1)	(2)	(3)	(4)	(5)	(6)
m_{it}	0.941***	0.944***	0.929***	0.938***	0.940***	0.928***
	(0.076)	(0.076)	(0.071)	(0.073)	(0.077)	(0.070)
nat. population		-0.000				-0.001
		(0.001)				(0.002)
population density		-0.002				-0.001
		(0.001)				(0.001)
speed of urbanisation			0.247**			0.240**
			(0.109)			(0.102)
% of pop. in urban				-0.007		-0.003
				(0.033)		(0.030)
ln(GDP per capita)					0.111	0.114
					(0.124)	(0.120)
year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2112	2112	2112	2112	2112	2112
Number of countries	88	88	88	88	88	88
adj. R-squared	0.845	0.847	0.853	0.845	0.846	0.854
within R-squared	0.847	0.848	0.855	0.847	0.847	0.856

Fixed effects regressions. All regressions include a constant term. Data are from The World Bank (2018b). Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As was expected, employing relative size of the largest urban agglomeration allows more accurate comparison of countries with diverse sizes and at different stages of urbanisation process. Urban primacy is not driven by the size of national population or urbanisation rate, as the respective regression coefficients in Table 1.6 suggest. Nevertheless, we see positive and significant correlation between excessive population share in the largest city and urban primacy. It is important to note that the significance of the excessive population coefficient does not necessarily follow from the construction of the m_{it} variable. The excessive population in the largest city, $\hat{\mu}_{it}$, in Eq.(1.4) is calculated as the difference between the actual size of the primate agglomeration and the one simulated using the growth rate of other urban agglomerations. Hence, if the primate city grows relatively faster than other urban areas in the country, it increases both m_{it} and primacy. However, this effect can be quantitatively and statistically insignificant, if an increase in absolute size of the primate city is very small compared to an absolute increase of all other urban areas. Instead, we see a statistically significant relationship that is quantitatively strong: a one standard deviation increase in m_{it} on average increases urban primacy by 3 percentage points.

for a more detailed overview of these measures.

However, urban primacy per se might be not so harmful. Do we really have to care about it? Henderson (2003) has demonstrated that after a certain point a further increase in urban primacy can be harmful for economic growth. Moreover, the theoretical application presented in this paper suggests that a bigger largest city, ceteris paribus, decreases population in the hinterland. The reasons for it are numerous. First of all, better public goods in the primate urban agglomeration attract migrants from the rest of the country, both from rural areas and hinterland cities. Moreover, if we allow for positive relationship between the size of the region (in this example the hinterland) and the amount of public goods there, outgoing migration from the hinterland might reduce local budgets and, hence, increase mortality and lower fertility. That is why I also estimate the regressions described by the Eq.(1.6) with a share of total national population living in all urban areas, except the primate one. The results of these estimations are presented in Table 1.7

Table 1.7: Excessive population and hinterland city size.

	Depedent variable: *share of national population in other urban areas*					
	(1)	(2)	(3)	(4)	(5)	(6)
m_{it}	-0.810***	-0.825***	-0.805***	-0.443***	-0.829***	-0.466***
	(0.069)	(0.067)	(0.071)	(0.052)	(0.069)	(0.046)
nat. population		0.031				0.011*
		(0.025)				(0.007)
population density		-0.003				-0.001
		(0.002)				(0.001)
speed of urbanisation			-0.095			0.072
			(0.142)			(0.111)
% of pop. in urban				0.744***		0.724***
				(0.046)		(0.036)
ln(GDP per capita)					1.390	0.296*
					(0.874)	(0.153)
year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	2112	2112	2112	2112	2112	2112
Number of countries	88	88	88	88	88	88
adj. R-squared	0.695	0.722	0.695	0.966	0.707	0.971
within R-squared	0.699	0.725	0.699	0.966	0.710	0.971

Fixed effects regressions. All regressions include a constant term. Data are from The World Bank (2018b). Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.7 demonstrates that excessive population in the primate urban agglomeration decreases the size of other urban agglomerations in the country. In other words, I can argue that excessive growth of the largest city in the economy happens at the expense of other urban areas. As can be expected, the magnitude of this effect drops after adding urbanisation rate, but still remains significant. This size of the effect is not to be underestimated: if excessive population in the primate city increases by one standard deviation, the share of population in hinterland areas decreases by almost 2 percent.

Summing this subsection up, I can conclude that excessive population has a sizeable effect on both absolute and relative sizes of the primate urban agglomeration. Furthermore, negative relationship between the share of excessive population in the largest city and the relative size of other urban areas suggests that the increased growth of the primate agglomeration can take place at the expense of slower population increase in the hinterland cities. As a result, excessive population in the primate cite can substantially affect the urbanisation processes in developing countries.

1.2.3 Robustness checks

In the previous sections I have demonstrated that better provision of public goods in the whole country is associated with a smaller excessive population share in the largest urban agglomeration. Furthermore, lower share of excessive population in the primate city results into smaller primate cities, measured both in absolute and relative terms. However, as was already mentioned above, size of the largest urban agglomeration can affect provision of public goods at the national level. To control for this type of endogeneity, I use per capita public expenditures as an instrumental variable for public goods. Public expenditures are also measured at the national level, but, unlike the case of public goods, there is no reason to consider the mean of public expenditures to be affected by the population distribution. The results of the IV regression estimations presented in Table 1.8 confirm the validity of public expenditures as an instrumental variable for public goods provision. Due to the fact that standard errors are clustered at the country level and are not independent and identically distributed, I employ Kleibergen-Paap statistics (Kleibergen and Paap, 2006). Lagrange multiplier (LM) statistic suggests that public expenditures are a relevant instrument for public goods provision, as the null hypothesis that the equation is underidentified is rejected at 1-% level. Moreover, public expenditures can also be considered not weak, as Kleibergen-Paap F statistic exceeds the rule-of-thumb value of 10 suggested by Stock and Yogo (2005) in both cases.

It is important to note that the estimations of the first stage of 2SLS regressions do not show significant relationship between the share of population living in the primate city and public goods provision at the national level. As was already mentioned above, this effect is ambiguous: excessive growth of the largest city with a better infrastructure allows more people to get access to public goods, but might constrain development of the hinterland areas and decrease public goods' provision there. Given this, I cannot make any particular conclusion on how population concentration in the largest city affects provision of public goods, rather than both of the above mentioned effects can be present and none of them prevails.

The results of the second stage of IV regressions go in line with the OLS estimations presented in Tables 1.3 and 1.4. Both access to improved sanitation and water facilities have significantly negative effect on excessive population in the largest urban agglomeration. Moreover, this effect is becoming quantitatively stronger if I control for endogeneity. In other words, an omitted variable (relative primate city size in this case) biases the estimated coefficients upwards, making them less negative, while in reality an improvement in provision of public goods decreases excessive population share even more. This bias is especially evident in case of access to water. Besides water and sanitation, Table 1.8 demonstrates

Table 1.8: Public goods and excessive population.

Panel A: *First stage*		
Depedent variable:	*sanitation*	*water*
	(1)	(2)
per capita public expenditures	-1.062***	-1.350***
	(0.194)	(0.244)
% of pop. in primate	0.163	0.401
	(0.245)	(0.361)
year FE	Yes	Yes
controls	Yes	Yes
adj. R-squared	0.567	0.576
within R-squared	0.573	0.582
Kleibergen-Paap LM statistic	19.851***	22.104***
Kleibergen-Paap F statistic	29.891	30.481
Panel B: *Second stage*		
Depedent variable:	m_{it}	
	(1)	(2)
sanitation	-0.217***	
	(0.063)	
water		-0.170***
		(0.050)
nat. population	0.021***	0.016***
	(0.008)	(0.006)
% of pop. in primate	1.112***	1.145***
	(0.147)	(0.160)
speed of urbanisation	0.188	0.380**
	(0.216)	(0.192)
population density	0.001	0.001
	(0.001)	(0.002)
ln(GDP per capita)	0.599	0.632*
	(0.377)	(0.383)
year FE	Yes	Yes
N	2112	2112
Number of countries	88	88
adj. R-squared	0.374	0.352

2SLS regressions with country fixed effects. Controls used at the first stage are: national population, speed of urbanisation, population density and log(GDP per capita). Data are from The World Bank (2018b). Standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

that the share of national population living in the largest urban agglomeration still has a significant effect. The regression coefficient does not change substantially both in terms of statistical significance and the magnitude, compared to the estimates in Tables 1.3 and 1.4. This result supports the hypothesis about the possible presence of self-reinforcing process of excessive growth of the primate urban agglomerations.

1.3 Conclusion

As it has already been shown in the literature, there are many factors that can facilitate growth of particular areas in a country. This paper contributes to the existing research emphasising the role of public goods. Larger cities can be expected to be more financially endowed to provide to their citizens a wider and, probably, better range of public goods, as the data suggest. However, an increased quality of public goods attracts more migrants, which add up to the local population, and improves survival chances in the location. This can potentially bring more resources to the city to allocate for public goods' provision. However, growth of the primate city and its budget might harm other regions of the country. This, however, might facilitate migration even more. As a result, excessive population in the primate city might deprive development of the hinterland and give rise to the congested urban giants. Such a situation would lead to dead-weight utility losses for the whole nation. The described effect is expected to be especially prominent for developing countries, where infrastructure is less developed and, hence, spatial variations in public goods provision are more pronounced.

I estimated a set of regression equations that describe (1) share of national population in the largest city of the country excessive to the amount that is expected to live there given the set of country characteristics, (2) absolute and relative (compared to the total urban population) size of the largest city, and (3) public goods' provision at the national level. To measure the share of excessive population I estimated synthetic populations of the primate cities by applying the growth rate of urban areas (excluding the primate city) in the country to the initial size of the largest city. As a result, I obtained the value that tells how the primate city would be expected to develop if it were just an "average" city in this country. The difference between simulated and actual population sizes is denoted as "excessive population". After scaling it to the national population to control for country's capacity to provide newcomers to the primate city, it was included into the regression analysis. As a result, the excessive population in the largest city in the economy is expected to be significantly influenced by the share of national population living in the primate city and provision of public goods at the national level. Higher share of total country's population living in the primate city is also positively and significantly associated with *excessive* population in the largest city, supporting the hypothesis that urban giants attract disproportionately more migrants and increase survival probability. Provision of public goods at the national level (measured by access to water and sanitation) is negatively associated with the share of excessive population, implying that when individuals can find sufficient public goods outside the primate city, they are less likely to move there and have not lower survival probability. Excessive population, in its turn, increases both absolute and relative size of the largest urban agglomeration. On the contrast, the share of excessive population in the primate city is significantly, but negatively correlated with the share of national population living in other urban areas. This finding supports the hypothesis that excessive growth of the primate areas can take place at the expense of the hinterland cities. However, no certain conclusions can be made about the effect of increased concentration of people in the primate city on shares of national population having access to water and sanitation. Larger sizes of primate cities can either provide access to these public goods to a larger number of citizens or halt infrastructure development in the hinterland areas. Distinguishing between these two effects is an important topic for the future research.

Appendix

Data sources

All data used for Table 1.1 and Figure 1 are from the DHS Program available for the following 57 countries:

Afghanistan, Angola, Armenia, Azerbaijan, Bangladesh, Benin, Bolivia, Brazil, Burkina Faso, Burundi, Cambodia, Cameroon, Chad, Colombia, Comoros, Congo, Congo Democratic Republic, Cote d'Ivoire, Dominican Republic, Egypt, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guatemala, Guinea, Guyana, Haiti, Indonesia, Jordan, Kazakhstan, Kenya, Kyrgyz Republic, Lesotho, Liberia, Madagascar, Malawi, Maldives, Mali, Moldova, Morocco, Mozambique, Namibia, Nepal, Nicaragua, Niger, Nigeria, Peru, Philippines, Rwanda, Senegal, Sierra Leone, South Africa, Tajikistan, Tanzania, Timor-Leste, Togo, Turkey, Turkmenistan, Uganda, Ukraine, Uzbekistan, Vietnam, Yemen, Zambia, Zimbabwe.

All data used for estimations are from The World Bank (2018b). Data for access to water and sanitation are: "Improved water source (% of population with access)" and "Improved sanitation facilities (% of population with access)", respectively. Regressions were run using the data years 1991-2014; available for the following 88 countries:

Algeria, Argentina, Australia, Austria, Bahrain, Bangladesh, Belarus, Belgium, Benin, Bolivia, Brazil, Bulgaria, Burkina Faso, Cameroon, Canada, Central African Republic, Chad, Chile, China, Colombia, Dem. Rep. of Congo, Costa Rica, Cote d'Ivoire, Czech Republic, Denmark, Dominican Republic, Ecuador, Arab Rep. of Egypt, El Salvador, Finland, France, Georgia, Germany, Ghana, Greece, Guatemala, Guinea, Honduras, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kenya, Kyrgyz Republic, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mexico, Morocco, Mozambique, Namibia, Nepal, Netherlands, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Peru, Philippines, Portugal, Romania, Russian Federation, Rwanda, Saudi Arabia, Senegal, Slovak Republic, South Africa, Spain, Sudan, Sweden, Switzerland, Tanzania, Thailand, Tunisia, Turkey, Uganda, United Kingdom, United States, Uruguay, Vietnam, Zimbabwe.

Table A1.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
m_{it}	-0.957	3.615	-22.912	11.334
sanitation (% of population with access)	53.501	30.166	4.2	100
water (% of population with access)	77.445	17.325	29.1	100
national population (mln.)	71.627	209.992	0.51	1364.27
population in largest city (mln.)	4.034	4.716	0.127	24.953
share of total population in primate (%)	13.847	8.352	0.72	49.701
share of urban population in primate (%)	29.963	11.522	2.637	59.543
urbanisation rate (%)	47.414	21.043	5.491	95.152
urbanisation speed (%)	1.17	1.341	-0.785	16.301
population density (people per sq. km)	111.216	205.123	1.781	1768.74
ln(GDP per capita)	7.128	1.204	4.612	10.122
per capita public expenditures (1000 USD)	0.413	0.682	0.006	6.382

All data used are from The World Bank (2018b).

Table A1.2: Correlation matrix of used control variables

	water	sanitation	share of total population in primate	national population	population density	urbanisation speed	ln(GDP p.c.)	p.c. public expenditures	population in largest city
water	1								
sanitation	0.8874	1							
share of total population in primate	0.387	0.4588	1						
national population	0.0254	-0.0735	-0.3093	1					
population density	0.1772	0.1364	-0.0467	0.1378	1				
urbanisation speed	-0.3689	-0.4044	-0.3038	0.1489	0.0907	1			
ln(GDP per capita)	0.8162	0.8473	0.4015	-0.0667	0.1092	-0.4119	1		
per capita public expenditures	0.5156	0.5631	0.1768	-0.0771	0.1029	-0.2819	0.7931	1	
population in largest city	0.2269	0.162	-0.0093	0.5052	0.1555	-0.0417	0.1772	0.0476	1

All data used are from The World Bank (2018b).

2 The Puzzle of Urban Sprawl. Now in 3d

As city populations increase, urban areas are also expected to grow to host new citizens. Nevertheless, low incomes might constrain further growth of formal housing causing squatting and congestion. Moreover, cities can expand in different dimensions: grow tall or sprawl wide to accommodate new residents. Existing theory does not explain the following paradox: cities occupy larger areas in developed countries, where land is relatively more expensive than capital, while urban agglomerations in Africa and Asia do not sprawl wide despite low land prices. This paper presents a model analysing the markets for formal housing and its production factors. The main innovation of the model is distinguishing between agricultural and housing land. The abundance of public amenities in the city determines the intensiveness of the land improvements that have to be done privately before erecting houses. Inclusion of land improvements into the analysis affects the price ratio of inputs used for housing construction and explains the capital-land paradox.

JEL Codes: O18, R31, R5, R38, H4

Keywords: urban agglomerations, public goods, land area, population, density

Introduction

People have been forming settlements since the origin of humanity. Living close to the neighbours increased security, made possible organising some basic facilities, like a common permanently maintained fire place or a primitive shelter. Moreover, it fostered long-term economic development. Spatial concentration of households facilitated labour specialisation and made exchange of goods much easier and faster. After the industrial revolution urbanisation has become crucial for further development. Growing industries required more workers that had to be spatially concentrated in the areas, where the enterprises were located. Eventually, population growth has led to a forming of towns and cities of different sizes. Nowadays it is more correct to consider the metropolitan areas, rather than single cities, as despite different names areas can be parts of a single urban agglomeration. A person can de jure live in one city, but work in another one de facto still belonging to the same metropolitan area. Modern means of transportation allow to travel fast and far away, so an everyday 30-minutes journey to work can bring people formally to another city. For example, New York and Jersey City are located in different states, but in reality belong to the same urban agglomeration.

However, cities and even urban agglomerations cannot grow infinitely. Increasing population size entails more crime, traffic or pollution (Mohring, 1961; Duranton and Turner, 2011; Glaeser, 2014). Some of these drawbacks cannot be properly assessed, while, for example, higher rents are relatively easy to quantify. Disadvantages of urbanisation, regardless if we can estimate them or not, might have an impact on economic outcomes, too. Transport costs and rents are the most obvious, but far not the only factors that might limit the further growth of the cities. Given that both advantages and disadvantages of the urban lifestyles increase, as the city's population grows, there is expected a certain number of residents that will maximise individual productivity in the given city. This topic is not new to the literature and a range of studies is already devoted to it (Henderson, 1974; Fujita, 1989; Fujita et al., 1999; Duranton and Puga, 2001; Henderson, 2005; Au and Henderson, 2006; Glaeser, 2008; Desmet and Rossi-Hansberg, 2013, 2014). Nevertheless, a set of city-related issues remain not that extensively researched. One of them is the concentration of people *within* a city.

Cities of similar population sizes have different shapes and occupy various areas. Moreover, ideally, as population increases, housing stock should grow, at least, at a similar rate, to avoid congestion. However, this does not always happen in reality. In this paper I attempt to determine the factors that influence aggregate density of citizens within the city limits and explain the overall sprawl of the urban area. This feature of urban agglomerations should be not ignored. Even if one supposes that population density has no *direct* effect on productivity growth, there is no doubt that it has a strong influence on life quality of those living in the cities. Slums are an example of extremely dense settlements, where sanitation norms and comfort are drastically low. Moreover, even in the case when stock of formal housing increases proportionately to the number or citizens, population density might still rise. Additional dwellings can be created in two ways: buildings can grow taller or agglomerations can sprawl wider. If people spread across the area, population dilutes, lowering the concentration of citizens per unit of area (say, sq. km.). The existing models (Mohring, 1961; Mills, 1980; Brueckner, 1987; Batty and Kim, 1992; Brueckner and Zenou, 1999; Brueckner et al., 1999) generally aim to explain the population distribution within the city: areas further from the centre are typically expected to be less densely populated. At the same time, to the best of my knowledge, there is no study explaining the ratio between area and population of a city. What are the factors affecting the 3-dimensional *shape* of cities? What are the determinants of urban sprawl? Why do densely-populated settlements emerge in particular countries? These questions are addressed in this paper.

It is not a new fact that the majority of the most populated cities are located in the developing countries. This holds true for the built-up areas: only 6 out of 30 largest urban agglomerations are in high income countries. Despite being widely used in the literature to describe how big the city is, number of residents is far not the only size characteristic of an urban agglomeration. Moreover, concentrating solely on population sizes, researchers ignore the important features of the settlements, like degrees of comfort or congestion. Land area occupied by a city appears equally important characteristic of an urban agglomeration in this context. Some cities occupy much more land than the others. When we rank all urban areas over 500.000 residents according to their land areas, we see that the pattern is completely different. Out of 30 cities occupying the largest land areas 14 are located in the U.S., and 5 more are in the other highly-developed countries (see Table (A2.5)). This is the first fact contradicting the standard theory: urban areas with the largest territory are mostly present in the countries with high agricultural rents. One could argue that such prevalence of the U.S. cities in the list is driven by some cultural features or the abundance of free land in the country. However, presence of Tokyo-Yokohama or Essen-Düsseldorf in the top-30 largest urban land areas might imply that there might be some other factors affecting the physical size of the agglomerations. However, comparing ranks does not really present the objective picture. For this reason it might be more reasonable to consider the population densities in the cities. This allows to compare both number of residents and land area in the cities simultaneously. When we take a look at the list of 30 or 40 (or even 150) most densely populated built-up areas in the Table (A2.6), none of the cities is located in high-income regions. More formally this fact is presented on Fig.2.1. This formulates an interesting paradox that goes against general economic intuition: cities sprawl wide when they are expected to grow tall. Assuming that land and capital are the inputs used for housing construction, we would expect the relatively cheaper production factor to be used more intensively. However, land is expensive and capital is cheap in developed countries, meaning that following standard economic intuition, it is more reasonable to build taller buildings rather than cultivate new land. Instead, cities sprawl wide in high-income countries marking a discrepancy between existing theory and real data. Explaining this paradox is the main objective of the paper.

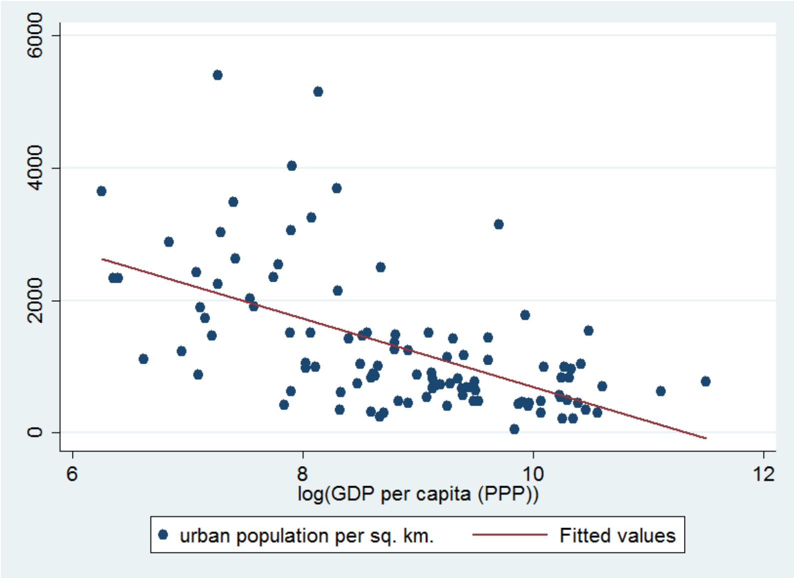


Figure 2.1: Urban population density and income

This paper presents a theoretical model that is capable of explaining the land/capital paradox mentioned above. One of the main innovations of the model is distinguishing between agricultural and housing land. I argue that before a unit of land is suitable for living, some improvements have to be undertaken. In other words, housing land has to be produced using agricultural land and physical capital. However, the intensiveness of necessary capital investments depends on the abundance of public goods provided by the city: the more and better amenities are present in the city, the less has to be financed privately. Considering physical capital aspect determining the price of land might change the capital/land ratio, as will be demonstrated further in the paper. Finally, the model demonstrates that financial constraints of cities in developing countries can limit the urban sprawl and increase the population density; while larger city budgets in developed countries allow amenities provision at a larger territory, hence, more spatially spread populations.

This paper is organised as follows. Section 2 present an overview of existing studies on the topic. It highlights the gap in the literature that this paper aims to fill and justifies the importance of the matter. Section 3 describes the model itself. Starting from general assumptions it gradually explains the particular details of the model setting. Section 4 conducts a simple empirical testing of the model. A brief summary section in concludes the paper.

2.1 Existing literature

To start with, it is necessary to define explicitly, what is the object of this study. To state the motivation of this study I use the data from World Urban Areas atlas by Demographia (2016). The atlas aggregates the data from different sources providing information about land areas, populations and, hence, densities of the largest metropolitan areas of the world. It is important to note that the data is provided for built-up urban areas, a *"continuously built up land mass of urban development that is within a labour market (metropolitan area or metropolitan region). An urban area contains no rural land (all land in the world is either urban or rural)"* (Demographia, 2016, p.4). Defining built-up areas this way allows to concentrate on the functions that a city performs and not on the bureaucratic definition of a city. Modern infrastructure and transport means allow people to live further from their work and commute longer distances in shorter time. Moreover, some cities have grown in size dramatically over years and now de jure different cities physically have merged into one urban area. This leads to the situation that formally an individual might reside in one city and work in the other, even though they de facto belong to the same economic entity. Employing the built-up areas as the main object of the study puts emphasis on the functions of the city instead of formal division into different municipalities.

The topic of this paper has not yet been studied extensively in the literature. Michaels et al. (2012) study the relationship between initial population density and growth both in urban and rural areas. Their study uses the past density to predict the growth and current population in Minor Civil Divisions of the U.S., but does not consider current city population density explicitly. However, there are two big strands of research that partially consider the problem of population density in urban areas. First of all, there is a set of papers that model the population distribution within a city. Starting from Clark (1951), studies of this strand have used the distance from central business district (CBD) to predict population density in a particular part of the city (Mills, 1980; Henderson, 2005). Brueckner and Helsley (2011) or Brueckner et al. (1999) go further and demonstrate how mobile residents cause rise or decay of particular city areas. Results of these studies are important for city developers, but they do not really explain the differences in aggregate population densities across urban agglomerations. Another strand of research considers population distribution across existing cities. Studies by Ades and Glaeser (1995), Gabaix (1999), Henderson (2003), Henderson and Wang (2007), Castells-Quintana (2017), Skorobogatov (2018) and many others explain patterns of urban population distribution across the system of cities and demonstrate possible effects of population size of one city on the whole national economy. However, these studies also do not explain, why, given the same number of citizens, some urban areas occupy much bigger areas than others.

Alternatively, one could refer to the studies linking city sizes to productivity to see if variations in population density can be explained with theoretical models of these articles. However, a range of studies by Mohring (1961), (Duranton and Puga, 2001), Henderson (2005) or Desmet and Rossi-Hansberg (2013) link area to population size. Cities are typically assumed to be flat circles, commuting and housing costs depend on the distance from the CBD and exogenous rents. However, these two assumptions make the models silent about population density by construction. Assuming urban areas to be flat by definition applies similar density in all cities. In other words, all urban areas have similar population density: one person per one standard unit of land. This approach might be suitable to explain why cities with larger populations have higher living costs or productivity spillovers, but completely ignores the population density in the city. Further in this paper I present the model of a city that does not depend on the two above mentioned assumptions.

However, some studies model housing rents in a more detailed way. Brueckner and Zenou (1999) has added the land market to Harris-Todaro model. In this formulation of the model, city sizes are affected by agricultural rents. As city population increases, new land has to be converted from agricultural into urban. Due to the fact that land can be universally used, housing and agricultural rents should be equal at the city edge. This model does not address the issue directly, but one can derive that higher agricultural rents should limit the further growth of the city (by simply reducing the maximum distance from CBD). Thus, between country variation in agricultural productivity (hence, land rents) could potentially explain different city sizes. More developed countries have higher agricultural rents and are expected to have spatially smaller cities, while those countries that are developing at the moment can afford reallocating more (cheap) land from agriculture to housing. In reality, this pattern does not really hold, as will be demonstrated later.

A range of classical urban models (Wheaton, 1974; Brueckner, 1987) deal with population densities more explicitly. In the unified version of Muth-Mills model presented by Brueckner (1987) buildings vary in size: higher land rents in the centre make the houses taller and flats inside them smaller, while at the outskirts, where land is cheaper, houses are flatter and individuals can afford bigger dwellings. Brueckner (1987) also considers population density within the city. As housing costs are lower further from the centre and people can afford larger dwellings there, population density also decreases with the distance to CBD. However, this model also ties the housing rents to the agricultural rent. As could be expected, higher agricultural rents make cities spatially smaller, buildings taller and dwellings tinier and more expensive. As a result, population density also increases in agricultural rents. However, this would imply that more densely populated areas are expected to appear in the countries with high agricultural rents. In other words, this formulation of Muth-Mills model predicts that slums are more likely to emerge in developed countries. However, productivity differences between developing and developed countries are even larger in agriculture, compared to non-agriculture (Lagakos and Waugh, 2013).

Given that, agricultural rents are also expected to be bigger in the high-income countries, where in reality cities are less densely populated, as will be demonstrated later.

As was already pointed above, the mentioned studies were not designed to explain the phenomena of population density and urban sprawl. However, as a by-product they would predict that cities occupy smaller areas in developed countries, making us doubt the validity of the mentioned models for explaining the emergence of heavily populated settlements. Obviously, there are two main forces that make city populations more dense. First, already mentioned, high agricultural rents decrease the amount of land available, hence, the spatial size of the urban area. Second, urbanisation is positively associated with economic development, so, ceteris paribus, one simply expects more urban citizens in industrialised countries (Harris and Todaro, 1970; Henderson, 2010; Glaeser, 2014; Gollin et al., 2016). As a result, many existing theoretical studies predict that overpopulated areas should prevail in developed countries due to larger urban populations residing in smaller areas. However, this contradicts the fact that urban population density is higher in developing countries (see Fig. 2.1).

As was already suggested, there might be some proneness for the overpopulated areas to appear in the developing countries. Jedwab and Vollrath (2018) explain this pattern by the process of post-WWII demographic transition: countries that were already industrialised by the end of the war had lower birth rates and did not experience rapid population growth in the cities; while urban agglomerations in the less developed countries went on growing due to still high birth and already low mortality rates. However, this still does not answer the question, why did cities not sprawl in terms of land area. In this paper I come up with one possible explanation of high population density. One of the main limitations of the models mentioned above is equating housing land rents to agricultural rents. As a result, these studies cannot model the population density correctly. If housing land rents are derived solely from the agricultural ones, they are expected to be higher in developed countries, hence, cities should occupy smaller area there. As this is not the case in the real world (as Fig.2.2 demonstrates), there is a need for an alternative explanation. As will be demonstrated further in this paper, distinguishing between agricultural and housing land might solve the issue and explain the variations in population densities across the cities of the world better.

To finish the review of existing literature on the topic, I have to explain, why is it important to do research on heavily populated areas. Costa and Kahn (2006) demonstrate that larger city sizes alone are associated with higher child mortality. It must be noted that their results are derived from U.S. data from the first half of the 20th century (before demographic transition (Haines, 2001)), but their results are robust to controlling for health expenditures, meaning that an increase in population alone might be associated with higher risk of disease. However, this fact does not explain the exact mechanisms that affect child

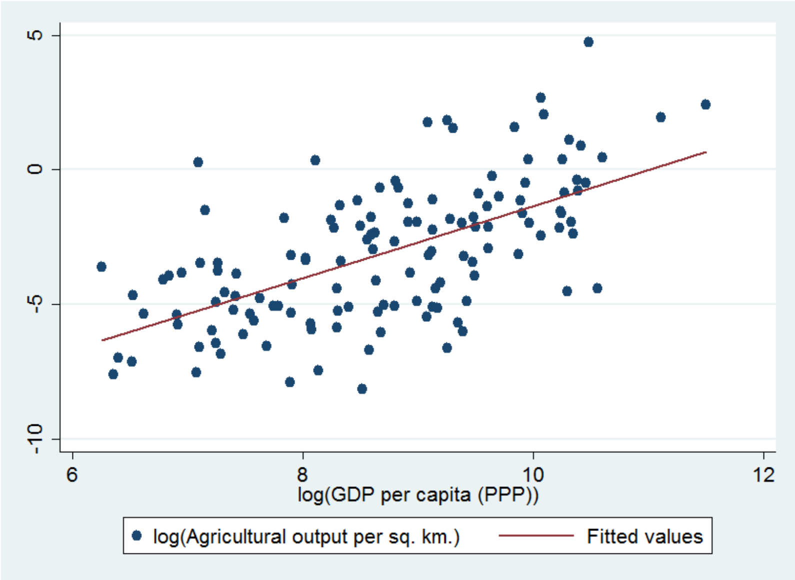


Figure 2.2: Agricultural productivity and income

mortality in bigger cities. One of the obvious reasons for such a difference would be better disease environment in smaller cities. Troesken (2004) convincingly demonstrates that worse access to water and sewage systems results into higher incidence of disease, hence, higher mortality. Moreover, this hold true not only for water-borne diseases, but also for the other types. Connecting the infrastructure and slums, the most extreme example of highly populated areas, we know that provision of public goods is generally substantially worse in the poor areas (UN-Habitat, 2016). Furthermore, this holds not only for infrastructure goods: it is quite challenging to maintain security or well-functioning road network in the areas with high population concentration (Duranton and Puga, 2004; Duranton and Turner, 2011). As a result, low provision of various public goods in the slums leads not only to lower life quality, but also worse economic outcomes and some kind of a poverty trap in these areas. Formally these effects will be modelled in the next section of the paper.

2.2 The model setting

In this section I present a simple model that explains urban population density. As was mentioned above, cities can grow tall or sprawls wide to host its citizens. This section presents a static model that does not consider population growth, so the number of citizens is fixed. The model demonstrates how wages, interest rates highlights public expenditures as an important determinant of urban density.

2.2.1 Production and income in the City

Let us assume the City to be populated by workers endowed with human capital. Similar to Henderson (2005), aggregate production of the City is described by the following monotonic and increasing function of human capital:

$$X = f(H), \tag{2.1}$$

where X is total output, H is aggregate stock of human capital[1].

Due to the fact that human capital is the only factor used in production of goods, workers retain the whole output. The wage rate (per unit of input), w, then can be written as:

$$w = \frac{f(H)}{H} = \frac{X}{H}. \tag{2.2}$$

As workers do not necessarily have identical skills, they are paid accordingly: each worker j provides h_j units of human capital, so that his labour income, y_j, is expressed as:

$$y(h_j, w) = h_j w \equiv y_j. \tag{2.3}$$

Such formulation of y_j allows the existence of different levels of income within the city. This feature will be utilised later to explain different patterns of housing consumption. The aggregate labour income, Y is defined as:

$$Y \equiv w \sum_{j=1}^{N} h_j = wH, \tag{2.4}$$

where N is the number of workers, i.e., total population of the City. As one can see from Eq. (2.4), the aggregate labour income in the City is increasing in two components: larger population (quantity) or higher average level of human capital (quality).

2.2.2 Consumption in the city

Individuals have to pay an income tax to the local government at a rate τ. The rest of the individual income is spent on consumption of private goods, c_j, and housing, m_j. I intendedly neglect qualitative features of housing. For example, price-differentiating between locations within the city does not substantially add explanatory power to the

[1] Henderson (2005) enriches production function with productivity spillovers from an increased population of the city. However, this would unnecessarily complicate the model or require us either to assume all workers to be identical. The predictions of the model presented in this paper do not depend on the presence or absence of productivity spillovers.

model, but leads to a loss of generality. Instead, I concentrate on the size of the dwellings (quantity) individuals rent. As citizens yield utility from both types of consumption, the choice between housing and private goods can be seen as a standard utility maximisation problem[2]:

$$\max_{c_j, m_j} U(c_i, m_j) = a_c \log(c_j - \bar{c}) + a_m \log(m_j)$$
$$\text{s.t.} \quad (1-\tau)y_j = c_j + r m_j, \tag{2.5}$$

where is the subsistence consumption level, r is the rent per unit of housing (e.g., squared meter) and a_c, a_m are relative preferences for private goods' and housing consumption, respectively. Price of c_j is normalised to unity. To survive individuals need to consume a certain amount of private goods (e.g., food), so $\bar{c}$ is the subsistence consumption level. As a result, the optimal housing consumption of a utility-maximising citizen is:

$$m_j = m(y_j) \equiv \begin{cases} \frac{(1-\tau)y_j - \bar{c}}{\left(1+\frac{a_c}{a_m}\right)r}, & if \quad (1-\tau)y_j > \bar{c} \\ 0 & if \quad (1-\tau)y_j = \bar{c}. \end{cases} \tag{2.6}$$

Those individuals that cannot afford consumption above the subsistence level end up in a sub-standard housing. This statement differs to some of the studies on the topic (Jimenez, 1985; Hoy and Jimenez, 1991; Turnbull, 2008; Brueckner and Selod, 2009; Shah, 2014; Brueckner and Lall, 2015). Jimenez (1985) models ending up in a squatter housing as a choice rather than a necessity. Squatters can be evicted from their dwellings with some probability and forced to enter the formal housing market, where they will face higher rents. However, the quality of housing is ignored in this formulation of a problem. The model presented in this paper assumes that individuals always prefer to reside in formal dwellings, unless their income is too low to afford it. Additionally, we can determine the two major factors determining the emergence of sub-standard housing. Recalling Eq. (2.3) and plugging it into inequality conditions of net income from Eq. (2.6), we can derive the necessary condition for an individual to reside in the formal sector:

$$h_j > \frac{\bar{c}}{(1-\tau)w} \equiv \tilde{h}. \tag{2.7}$$

Citizens that can afford formal residency are endowed with human capital, h_j, above a certain level that yields them income enough to survive ($c_j > \bar{c}$) and rent, at least, some amount of housing. Moreover, from Eq. (2.7) we can see that a higher per hour wage rate, w, decreases the necessary level of h_j. In other words, individuals in developed countries (where salaries are generally higher) do not need to posses exceptional skills to afford proper

[2]Utility function does not consider public goods explicitly, as individuals have no choice on their optimal provision level, when no migration is allowed in the model.

housing. On the contrast, if w is low, citizens need substantially more individual human capital not to end up in sub-standard housing. I denote μ as the share of citizens that can afford official housing:

$$\mu \equiv \int_{\tilde{h}}^{h_{max}} p(h)dh, \tag{2.8}$$

where $p(h)$ is a probability density function of human capital. Eq. (2.8) suggests that we can expect μ to be substantially higher in developed countries, due to higher average level of h (i.e., $p(h)$ is skewed to the right).

The local government collects taxes to finance its activities (basically, to provide public goods)[3]:

$$B = \tau \sum_{j=1}^{N} y_j = \tau w H. \tag{2.9}$$

The larger governmental budget, B, the wider range and the better quality of amenities, G, the City can offer:

$$G \equiv G(B, \xi), \quad \frac{\partial G(B,\xi)}{\partial B} > 0 \quad \text{and} \quad \frac{\partial G(B,\xi)}{\partial \xi} > 0, \tag{2.10}$$

where $\xi > 0$ is a parameter that determines how efficiently the governments converts funds into public goods.

2.2.3 Housing land production

Existing studies on agglomerations by (Wheaton, 1974), Brueckner (1987) and Brueckner and Zenou (1999) assume perfect transformability of agricultural land into housing land. However, this leads to an interesting paradox contradicting basic economic theory. Suppose housing is produced using physical capital and land. Then, profit-maximising construction firms should use more intensively the production factor that is relatively cheaper. In other words, in the countries with cheap land cities are expected to sprawl wide, and grow tall when land is relatively more expensive. Lagakos and Waugh (2013) demonstrate that productivity differences between developing and developed countries are higher in agricultural sector than in industry. Hence, price of land is expected to be lower in low-income countries, but capital should be cheaper in high-income countries. As a result, we would expect cities in developing countries sprawl wide due to the presence of cheap land and scarce capital. The opposite holds for developed countries with expensive land

[3] The ultimate goal of the paper does not require me to model the tax rate explicitly. In this setting taxes are simply added to explain how public goods are funded. Alternatively, I could model τ as a result of a maximisation problem of the government that aims to stay incumbent: $\tau = arg\max[u(G(\tau)) - d(\tau)]$, where $u(G(\tau))$ is the utility a median voter obtains from public goods and $d(\tau)$ is the utility lost due to the paid taxes.

and abundant capital. However, in reality the pattern is just the opposite. The largest urban land areas are located mostly in developed countries (see Table A2.5). This pattern becomes even more obvious if we look at population densities (Table A2.6): non of the top-50 most populated agglomerations lies in developed countries.

This paradox can be explained by introducing the distinction between housing and agricultural land. I argue that modern houses cannot be build in the plain field, but rather need a special type of land. Agricultural land needs to be processed in a particular way to become urban. The range of necessary land improvements is not limited solely by infrastructure. We can also think of access to public goods in general as a required land upgrade. Besides connection to pipelines and electricity network citizens want to have a healthcare facility or a school within a reasonable distance. Of course, the set of land improvements that citizens consider satisfactory might differ across countries. However, I assume that the mechanism holds in general: houses are build on a special type of land that has to be produced using plain (agricultural) land and capital. Physical capital is needed to connect a new land lot to the existing infrastructure. This can be not only already mentioned pipelines and roads, but also less material amendments, like, for example, security services. For instance, if tenants have concerns about the efficiency of police performance, they might hire a private company to secure their dwellings. This increases a price of land, but also makes it more suitable for citizens. Moreover, density of available public goods around a land lot determines the required investments. If a unit of land is surrounded by different infrastructural objects, connecting it is generally cheaper compared to the situation, when long pipelines have to be laid to the remote lots. In other words, the more public goods are available, the less capital you need to invest to convert it into urban land. More formally this can be presented the following way:

$$l_m = \min\{l_a, ek_l\}, \quad e \equiv e(G), \quad \frac{\partial e(G)}{\partial G} > 0, \tag{2.11}$$

where l_a is agricultural land that you need to convert into housing land using k_l, physical capital used for land improvements. The use of Leontief production function is dictated by the fact that agricultural land and capital are complimentary inputs. We cannot employ more physical capital to create more land units: one lot of plain (agricultural) land will always yield one lot of housing land. The variable e represents the "productivity" of private capital investments into housing land. In other words, the amount of capital resources that land developing firms have to devote to the transformation of one land unit depends on the amount of public goods provided by the city government. It is important to note that both qualitative and quantitative features of public goods matter for e. The larger number of amenities is provided by the local government, the less should be generated privately "on the spot". The same holds for quality or density of public goods provision: the denser is the coverage in the city, the less has to be invested into connecting to the network.

Producing according to the Eq. (2.11) land developers hire agricultural land and physical capital at the national market to create housing land l_m that they sell at a price p_l[4]. Assuming that housing land market is perfectly competitive, profits of a land developer, π_l, are described the following way:

$$\pi_l = p_l l_m - l_a p_a - k_l i = 0, \tag{2.12}$$

where p_a is a price of one lot of agricultural land, k_l is a physical capital used for land improvements and i is the interest rate. The optimal ratio of inputs in the Eq. (2.11) is $l_m = l_a = ek_l$. Given this optimality condition and plugging it into Eq. (2.12) land developers determine the price of one lot of housing land:

$$p_l = p_a + \frac{i}{e}. \tag{2.13}$$

Eq. (2.13) brings a new component to the price of land suitable for housing. Existing studies have typically neglected the second part of the expression, assuming $p_l = p_a$. Such formulation of housing land prices is not suitable for unambiguous cross-country comparisons: higher agricultural rents (hence, p_a) are expected to be observed in countries with higher e. However, to analyse the directions of urban sprawl we have to compare not p_l's, but $\frac{p_l}{i}$ ratios, as will be demonstrated later, and this is already a feasible task.

2.2.4 Housing construction

Once the agricultural land was processed into a housing one, dwellings can be build on it using physical capital, k_m. Unlike production of housing land itself, construction has substitutable inputs: you can produce more dwellings employing more capital. Buildings occupying a single lot of land can just grow taller. This technology is described by the following production function:

$$m = k_m l_m. \tag{2.14}$$

Housing market is also assumed to be perfectly competitive. The profit of construction firms, π_m, is described by the following function:

$$\pi_m = k_m l_m r - k_m i - l_m p_l = 0. \tag{2.15}$$

[4]In the modern western World the distinction into capitalists and workers is not that pronounced. However, inequality still persists in all countries, indicating that some individuals posses much more assets than the others (especially in developing countries). Given that, we can assume workers to be numerous, while capitalists are few (negligible in the total population of the City compared to the number of workers).

A profit-maximising agent chooses the optimal $\frac{k_m}{l_m}$ ratio that can also be interpreted as a "relative height" of buildings in the City:

$$\frac{k_m}{l_m} = \frac{p_l}{i} = \frac{p_a}{i} + \frac{1}{e}. \tag{2.16}$$

From Eq. (2.16) we see the main purpose of introducing the concept of housing land. Previous studies would explain the tallness of buildings in a city solely by $\frac{p_a}{i}$ causing a paradox described above. We know that productivity differences between high- and low-income countries are even higher in agriculture than in non-agriculture (Lagakos and Waugh, 2013). Moreover, interest rates are higher in developing regions (The World Bank, 2018b). Given this, we can assume that $p_a^h > p_a^l$ and $i^h < i^l$, where the indices h and l stand for high-income and low-income countries, respectively. This drives $\frac{k_m}{l_m}$ ratio in developed countries up. This fact is not necessarily wrong: buildings in high-income regions might be really taller compared to those in developing regions. However, if land areas of high-income agglomerations are substantially larger, standard theory would predict the buildings to be overproportionally higher than in developing countries. Inclusion of the e component decreases the predicted difference in capital to land ratio between high- and low-income countries. As B's are expected to be higher in developed countries (implying higher G, thus, e) and $\lim_{e\to\infty} \frac{k_m}{l_m} = \frac{p_a}{i}$, capital to land ratio can be a good approximation for $\frac{p_a}{i}$ in high-income countries. However, in developing regions, where B's are relatively low and housing land needs more improvements, the e component becomes more important for determining the capital to land ratio, making it higher compared to the case, when e is absent. As a result, adding e as a determinant of $\frac{k_m}{l_m}$ would generally predict relatively more required capital investments into housing construction (from scratch) compared to a standard $\frac{p_a}{i}$ formulation, but potential bias is higher particularly for developing countries.

As was mentioned above, Eq. (2.16) does not necessarily imply that buildings are taller in developing countries. Besides the height of the buildings larger population densities can be explained by lower aggregate demand for housing (measured in sq. meters, but not number of dwellings). This issue will be considered in the next section of the paper.

2.2.5 Aggregate housing demand

The lack of public funds available to expand the urban areas might be not the only factor limiting the growth of cities in developing countries. It might be the case that citizens in poorer countries have lower demand for housing. Combining Eq. (2.4), Eq. (2.6) and Eq. (2.8) we can derive the aggregate demand function for housing, M:

$$M = \frac{(1-\tau)\mu(Y - N\bar{c})}{\left(1 + \frac{a_c}{a_m}\right) r}. \tag{2.17}$$

Higher wage rates and human capital levels in developed countries, which imply higher μ and Y, drive M upwards. Obviously, if an individual earns more, he can afford to spend more on housing. On the other hand, higher earnings might be accompanied with a higher cost of housing. However, data suggests that as countries develop, incomes rise, at least, as fast as rents. Fig. 2.3 presents the data from International Comparison Program (ICP 2011) for 180 countries to show how variations in per capita income affect the share of overall expenditures devoted to housing payments. We see that there is no significant positive correlation between these two variables, but even a slight negative one (correlation coefficient of -0.1659, statistically significant at 5% level), meaning that the share of income individuals spend on housing does not increase as countries get better-off. Moreover, in general the dispersion of expenditure shares is surprisingly low across the world: households in half of the countries spend between 10% and 20% of their disposable income on housing, as Fig. (2.4) demonstrates. Unfortunately, no data on actual consumption of housing for the whole world is available. Instead, the OECD provides data on 35 member-countries and Brazil, Russia and South Africa. Fig. 2.5 demonstrates that for this set of countries we see a very strong positive correlation between disposable income and housing consumption (measured in number of rooms per person): the correlation coefficient of 0.8235, statistically significant at the 0.1% level. Overall, the facts mentioned in this section imply that in terms of our model the average $\frac{y_j}{r}$ ratio from Eq. (2.6) is expected to increase in individual income. In other words, as countries develop, average value of y_j grows at a rate that is higher than r. As a result, we can expect individual housing demand, m_j, to be higher in developed countries, as Fig. (2.5) actually demonstrates.

2.2.6 Urban land and density

However, higher individual housing consumption does not necessarily imply higher population density. For example, people can comfortably live in spacious flats in skyscrapers that occupy very little land. However, the framework developed above allows us to express the amount of land that urban areas occupy and population density within these agglomerations. Having both housing supply and aggregate demand functions determined, we can combine Eq. (2.15), Eq. (2.16) and Eq. (2.17) to determine the total land area occupied by the City, L:

$$L = \frac{(1-\tau)\mu(Y - N\bar{c})}{2(1+\frac{a_c}{a_m})(p_a + \frac{i}{e})}. \tag{2.18}$$

We see that L is increasing in μ and Y, the two parameters determining demand for formal housing. The more people are able to afford a dwelling in the City, and the more they can spend, the larger area this city occupies. On the other hand, both agricultural land price, p_a, and interest rate, i, decrease L. Having the urban land area defined, we can finally

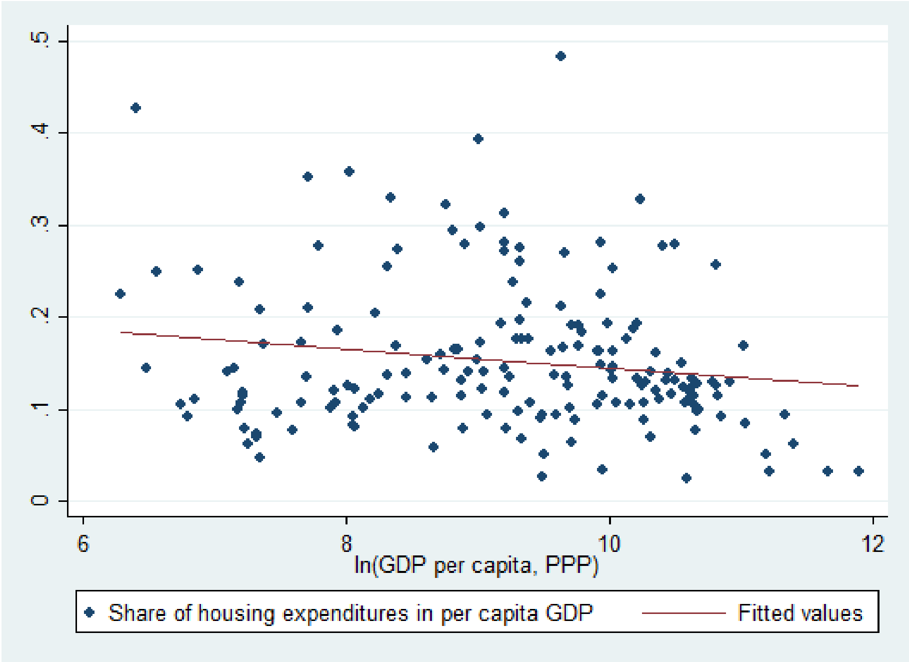


Figure 2.3: Housing expenditure shares and per capita GDP

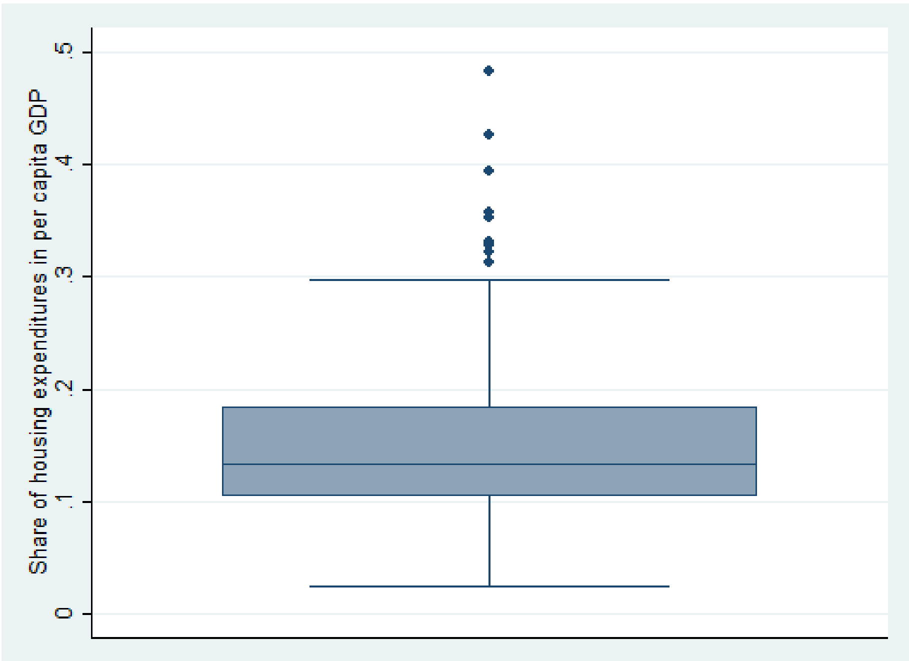


Figure 2.4: Box plot of share of housing expenditures in per capita GDP

express population density in the City:

$$\frac{N}{L} = 2\left(1 + \frac{a_c}{a_m}\right)\left(\frac{p_a}{\bar{y}} + \frac{i}{e\bar{y}}\right), \quad \textit{where} \quad \bar{y} \equiv \frac{\mu(1-\tau)(Y - N\bar{c})}{N}. \tag{2.19}$$

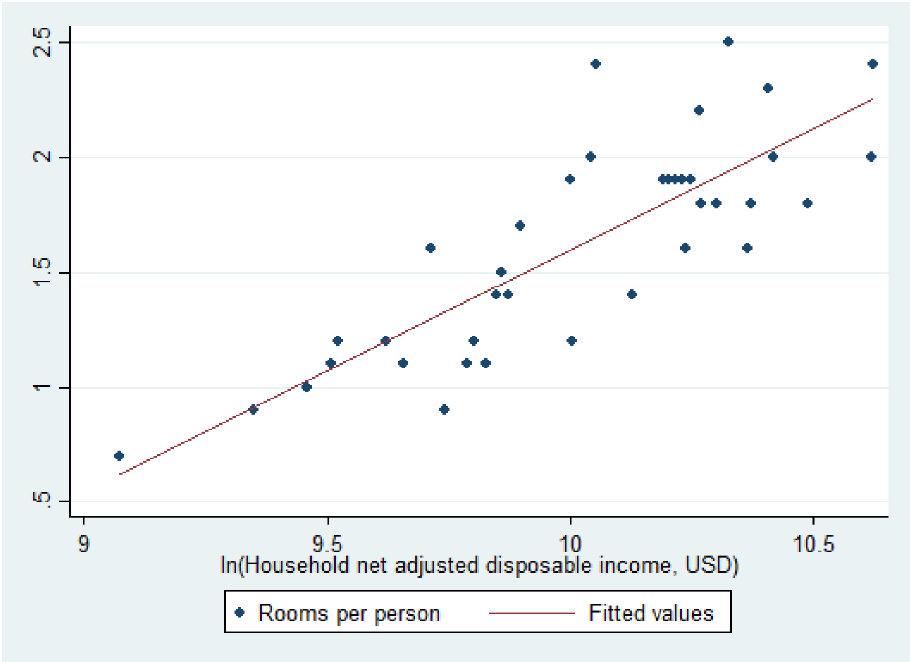


Figure 2.5: Housing consumption and disposable income

Eq. (2.19) presents number of citizens per one unit of land as a function of input prices relative to the average disposable income, $\bar{y}$. If agricultural productivity is high compared to the average disposable income of an urban citizen, we would expect, ceteris paribus, more spatially concentrated cities. In case of physical capital, one has to compare interest rates not only with $\bar{y}$, but also consider various efficiency of capital investments determined by public goods' availability.

Formulation of urban densities in Eq. (2.19) allows some comparison between developing and developed countries. By definition, average income, is higher in developing regions. Moreover, as was explained above, μ is also expected to be higher, as population gets richer. However, economic development increases productivity not evenly across sectors. Caselli (2005), Restuccia et al. (2008) and Lagakos and Waugh (2013) demonstrate that productivity differences are larger in agriculture than in non-agriculture. Models that do not consider housing land improvements, hence, the e character in the denominator of Eq. (2.19), would simply predict higher $\frac{p_a}{\bar{y}}$ ratio and, hence, lower density for developing countries. However, in reality the pattern is just the opposite, as was argued above. The variable e in combination with interest rates in Eq. (2.19) provides one possible explanation for higher urban densities in developed countries, what is consistent with the data.

2.3 Empirical evidence

Once potential determinants of urban density are revealed, it is possible to test some of the theoretical predictions empirically. Unfortunately, most of the variables of interest are hard to measure. Moreover, I have not made many assumptions about the functional forms of several relationships: $G(B, \xi)$ or $e(G)$. We know that government expenditures increase the amount and quality of provided public goods, but no assumptions were made about exact elasticities. Similar holds for the efficiency of private capital investments into production of housing land. Finally, agricultural productivity is an important, but not the only determinant of agricultural land price. As a result, it is feasible to estimate only conditional correlations of some variables with urban population density. This approach can, nevertheless, demonstrate some support for a theoretical model presented in the previous section.

2.3.1 Empirical strategy

To start with, I have to explain how the dependent is calculated. To estimate urban population density one has to scale the number of citizens residing in an urban agglomeration with its area. Data for population sizes is widely available, for example, from administrative statistics or censuses. However, measuring urban land area is a more complicated task, especially in developing countries, where urbanisation process is actively taking place at the moment. The data used in this study is obtained from the Global Human Settlement Layers (GHSL) by Pesaresi and Freire (2016). The GHSL is a geocoded dataset that presents the data on land use according to the degree of urbanisation (DEGURBA) concept by Eurostat (2018) for years 1975, 1990, 2000 and 2015. Due to data availability, for this paper I used the all the years, except 1975. The GHSL classifies all inhabited land as cities (densely populated areas), towns and suburbs (intermediate density) and rural areas (scarcely populated areas) at 1-km scale. I consider all pixels that are classified as cities and towns urban. Counting all urban pixels (size of 1 km^2) at the country level yields the size of the area that is occupied by urban settlements, regardless of their population density. Then, I use the Gridded Population of the World dataset (CIESIN 2016) to obtain the number of people residing in each grid cell that is considered urban. Summing up the number of residents of urban grid cells I obtain the total urban population of the country. Once I have both total urban land area and total urban population, I can calculate average urban population density. This approach allows to analyse the degree of urban sprawl: do people live close to each other or spread more remotely?

Now I can list the set of variables that can potentially explain variations in urban population densities. As was already mentioned above, unfortunately, no data is available for any of the determinants as they are for a wide range of countries. Instead, I come up with a set of proxies for each right-hand-side variable. Speaking about agricultural land, its price is expected to increase if land becomes productive. I measure land productivity with the value added in agriculture per one hectare of land. For the price of capital i I come up with several proxies. First of all, the World Bank provides lending interest rate, that should be the actual price of capital that economic agents of the country face. However, the data for this variable is available for a limited amount of countries. Moreover, inflation rates can play a significant role, when it comes to actual borrowing. Real interest rates provided by the World Bank account for inflation rate, but are also not available for all countries in the sample. Another way of measuring capital price could be marginal product of capital, if we assume perfectly competitive market for physical capital. The Penn World Table (Feenstra et al., 2015) provides data for the total output (GDP), physical capital stock and share of labour compensation in GDP. Assumin Cobb-Douglas production technology, calculation of marginal product of capital appears to be a feasible task. However, the data for the share of labour compensation in GDP is also not universally available. Finally, capital productivity that is available for the largest number of countries can be used as an (imperfect) proxy for the price of physical capital. Further, regressions employing all four measures of capital price will be demonstrated. The results are robust to the choice of a particular indicator.

As Eq.(2.19) demonstrates, population density is expected to decrease in the efficiency of private capital investments, e. According to Eq. (2.11) e is higher if public goods are well provided. In other words, we can expect private land investments to be more efficient in the areas with high level of public amenities. As a result, population density is expected to decrease in the amount of public goods. I proxy public goods with an index of infrastructural goods provided in urban areas of the country. To estimate this index, I calculate a simple average of the following variables: share of urban population with access to improved water source, sanitation facilities, electricity and non-solid fuel. The intuition behind this approach is quite simple: if the existing infrastructural networks in urban areas are adequate (provide access to these public goods to everyone who requires it), the city can sprawl further. If the level of infrastructure is too low to satisfy the demand for all citizens, new residents will aim to concentrate around the existing infrastructural focal points. Of course, the four types of public goods do not represent the full set of amenities that a city can offer, but they are easy to measure and are assumed to be correlated with other types of public goods (see (Yakubenko, 2018) for a detailed argumentation of this statement). Due to the fact that levels of public goods provision are typically correlated between each other, I cannot include them simultaneously and aggregate using a simple average. This approach is based on the assumption of similar preferences towards each amenity, but allows for some degree of substitution between particular goods (e.g., “I have no electricity, but, at least water is there”).

The last determinant of urban population density from Eq. (2.19) is the average disposable income of an urban citizen, $\bar{y}$. A simple GDP per capita can also capture the productivity in agriculture, hence, correlate with land productivity. I assume that urban residents are primarily employed in manufacturing and services, while agriculture is the primary activity in rural areas. Then, to separate urban and rural incomes, I subtract value added in agriculture from total GDP and divide this value by the total urban population. The obtained value is not a precise estimate of the average income of an urban citizen, if we think about non-agricultural activities of rural residents. However, the calculated variable can be assumed to be independent from agricultural productivity.

Regression estimates presented in this article are obtained through ordinary least squares estimator (OLS). Ideally, one should use a country fixed effect (FE) estimator that will allow to control for unobserved time-invariant country characteristics that can affect city sprawl. However, the data structure does not fit the FE technique. First of all, the time dimension of the data is available only for three periods and the panel is unbalanced, so that there is not enough variation for many countries of the sample. Secondly, the main purpose of this exercise is to explore the variation between developed and developing countries. However, level of public goods provision in cities of many developed countries has reached 100%-level by the first period of the panel (1990). Absence of time-variation of the main variable of interest for a substantial part of the sample is another reason not to use a FE estimator. Instead, I employ plenty of control variables that capture various country characteristics.

The model presented above assumes no migration to focus on the importance of public goods for determination of urban sprawl, but in reality one has to account for some population dynamics. Urbanisation rate can be a good indicator of how many new citizens are expected to arrive to urban areas. If a substantial share of population resides in rural areas, massive migration to urban areas appears more likely. Once cities face a large amount of migrants, they might lack time to sprawl wide to accommodate new citizens. In this context, we also need to consider total population of a country: the more people there are, the larger is the pool of potential migrants. Other control variables employed in the regression analysis can be split into three categories listed below:

1. Land features

Agricultural productivity can be not the only factor that affects the price of agricultural land and, hence, the extent of urban sprawl. First of all, due to a limited amount of land available, we can expect the market for agricultural land to be inelastic. As a result, land scarcity can also affect the price of agricultural land. To control for land supply (currently and potentially), I add the amount of agricultural and total land of the country. Moreover, not all land can be used for agriculture, so to consider this I include land suitability for agriculture and its Gini coefficient variables. Higher land suitability is expected to increase the land price, as agricultural use option appears more profitable. Besides availability of public goods there are other factors that determine the capital intensity of processing land for construction. One of the factors that can complicate the process of land conversion is terrain roughness. Moreover, elevation might be a factor that limits the urban sprawl. However, a country might be just located on an elevated plateau. So, besides average elevation level, one has to consider how much it varies, so the standard deviation of elevation is also included into the regression.

2. Geographic features

Other geographic variables, not directly related to the land characteristics, might affect urban sprawl. To control for climate, I include variables of latitude and longitude (geographic centroids), average temperature and precipitation. Besides that, Bleakley and Lin (2012) have demonstrated that proximity to a river can be crucial for economic development of a city and population density of the area. To control if average proximity to water has an effect of urban sprawl I add the percentage of land near a waterway variable.

3. Cultural features

What if living in densely-populated settlements is just a cultural feature? Part of this can be captured by the share of housing expenditures in total expenditures of a household. In some countries individuals might just have higher relative preference for housing (a_m in terms of Eq. (2.5). Besides this, "favouring congestion" habit might be correlated with some other features that constitute the national culture. For example, in case of ex-colonies, the share of population of European descent can be a way to compare their cultural similarity to the European states. Furthermore, ethnic fractionalisation can motivate people to live further from a rival ethnic group, sprawling the city wider. Finally, religion is a very important determinant of culture, so I control for its effect including shares of population attributed to one of the main religious confessions.

As a result, the estimated regression equation looks the following way:

$$y_{it} = \beta_0 + \beta' X_{it} + \gamma' Z_{it} + \lambda_t + \varepsilon_{it}, \tag{2.20}$$

where y_{it} is urban population density, X_{it} is a vector of main explanatory variables: land productivity, capital price, public goods and non-agricultural GDP. Z_{it} is a vector of auxiliary controls, explained above, λ_t is a year fixed effect and ε_{it} is an error term.

2.3.2 Empirical results

Tables 2.1 and 2.2 demonstrate the results of the estimation of the regression model described in Eq.(2.20) without auxiliary controls. The estimated coefficients follow the predictions of the theoretical model presented above. Agricultural productivity, that is assumed to raise the price of land, increases urban population density. Better provision of public goods decreases density, supporting the hypothesis that better development of infrastructure allows cities to sprawl wider. These results are robust to inclusion of non-agricultural per capita GDP, that also has the predicted sign - higher income allows citizens to consume more housing, thus, increasing area that cities occupy. Interestingly, price of capital has a non-significant coefficient, regardless of the proxy employed[5]. Tables 2.3 and 2.4 demonstrate that the results are generally robust to inclusion of additional controls, especially if we concentrate on the primary variable of interest - index of public goods. However, land productivity and non-agricultural income lose significance, once I add urbanisation rate. I suppose that the three variables are correlated with each other and can be considered an indicator of overall development of a country. When land becomes more productive (partially due to better quality of labour), less employers are needed in agriculture, so more people can migrate to urban areas to concentrate on non-agricultural activities. Simultaneously, when income in a city is higher, it becomes more attractive to migrate there from rural areas.

To sum up, this section has demonstrated some support for the predictions of the theory section: urban population density is increasing in land productivity, that serves as a proxy for price of agricultural land, while public goods and urban income decrease it. However, no effect of capital price variables employed in this study was detected. There might be several reasons for this result. First of all, used proxies can be bad indicators of the price for physical capital. Secondly, one can expect both land and labour to become more productive if capital endowment increases. In other words, cheap capital enables higher productivity, hence, increased wages of other production factors. Even though Table A2.2 demonstrates that land productivity, non-agricultural GDP and each of capital price indicators demonstrate correlation coefficient quantitatively way below the threshold level of 0.8 implying multicollinearity (Gujarati, 2003), one can still expect the effect of capital price to be captured by the other two variables. Disentangling these effects and more thorough empirical testing of the theoretical model are indeed important topics for future research. This paper, nevertheless, argues that lack of public goods can limit sprawl of urban areas, thus increasing urban population density.

[5] Tables A2.3 and A2.4 present the results of marginal product of capital and real interesting rate, respectively, as the indicators of capital price. Results remain robust to the choice of the proxy.

Table 2.1: Determinants of urban density. OLS estimates.

Dependent variable:	urban population density			
	(1)	(2)	(3)	(4)
land productivity	0.105**	0.099**	0.132***	0.126***
	(0.051)	(0.050)	(0.047)	(0.045)
lending interest rate	3.121	4.217	0.622	1.900
	(3.822)	(4.361)	(4.000)	(4.447)
public goods index	-19.546***	-19.070***	-15.438**	-14.731**
	(6.438)	(6.474)	(6.646)	(6.705)
non-agricultural GDP p.c.			-0.016**	-0.016**
			(0.007)	(0.007)
year FE	No	Yes	No	Yes
N	210	210	209	209
Number of countries	107	107	107	107
Adj. R-squared	0.112	0.119	0.143	0.154

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.2: Determinants of urban density. OLS estimates.

Dependent variable:	urban population density			
	(1)	(2)	(3)	(4)
land productivity	0.084	0.079	0.130***	0.125***
	(0.051)	(0.051)	(0.045)	(0.043)
capital productivity	20.262	5.215	-133.468	-140.992
	(142.612)	(162.739)	(132.757)	(151.919)
public goods index	-22.008***	-21.969***	-17.382***	-17.138***
	(4.552)	(4.555)	(4.694)	(4.686)
non-agricultural GDP p.c.			-0.016***	-0.017***
			(0.003)	(0.003)
year FE	No	Yes	No	Yes
N	329	329	327	327
Number of countries	128	128	128	128
Adj. R-squared	0.150	0.151	0.212	0.219

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.3: Determinants of urban density. OLS estimates.

Dependent variable:	urban population density							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
land productivity	0.081*	0.125***	0.113**	0.143**	0.101**	0.053	0.125***	0.035
	(0.044)	(0.042)	(0.046)	(0.055)	(0.045)	(0.047)	(0.043)	(0.058)
capital productivity	-185.420	-117.152	-185.168	-111.744	-188.175	-243.920*	-137.116	-84.149
	(148.905)	(149.379)	(140.302)	(150.307)	(164.208)	(133.118)	(156.413)	(156.703)
public goods index	-15.069**	-17.307***	-10.543*	-15.349***	-17.160***	-24.080***	-16.922***	-26.698***
	(7.404)	(4.734)	(5.571)	(4.565)	(6.269)	(3.852)	(5.229)	(6.819)
non-agricultural GDP p.c.	-0.007***	-0.019***	-0.013***	-0.014***	-0.006*	-0.007*	-0.017***	-0.001
	(0.003)	(0.004)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.004)
urbanisation rate						4226.994***		2921.605***
						(830.279)		(756.477)
population						0.000		0.000
						(0.000)		(0.000)
housing share							-235.943	-1126.329
							(1633.173)	(1681.115)
year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
continent FE	Yes	No	No	No	No	No	No	Yes
land features	No	Yes	No	No	No	No	No	Yes
geographic controls	No	No	Yes	No	No	No	No	Yes
legal origin	No	No	No	Yes	No	No	No	Yes
culture	No	No	No	No	Yes	No	No	Yes
N	327	323	325	327	322	327	327	318
Number of countries	128	126	127	128	125	128	128	123
Adj. R-squared	0.292	0.269	0.280	0.270	0.367	0.417	0.217	0.556

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.4: Determinants of urban density. OLS estimates.

Dependent variable:	urban population density							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
land productivity	0.087*	0.117***	0.100**	0.161***	0.085*	0.127***	0.058	0.004
	(0.050)	(0.033)	(0.042)	(0.060)	(0.051)	(0.044)	(0.047)	(0.057)
lending interest rate	0.304	3.128	4.506	1.210	5.411	2.481	4.909	3.003
	(4.582)	(4.855)	(4.773)	(4.194)	(4.292)	(4.288)	(3.813)	(4.832)
public goods index	-19.453*	-15.413**	-9.971	-15.257**	-17.968*	-13.503*	-21.772***	-33.507***
	(9.962)	(7.033)	(7.970)	(6.459)	(9.352)	(7.163)	(4.894)	(8.628)
non-agricultural GDP p.c.	-0.008	-0.016**	-0.013**	-0.013*	-0.003	-0.018***	-0.008	-0.002
	(0.006)	(0.008)	(0.006)	(0.008)	(0.005)	(0.007)	(0.006)	(0.008)
housing share						-1621.953		-2533.674
						(1636.231)		(1686.333)
urbanisation rate							4376.568***	2914.625***
							(1027.564)	(818.367)
population							0.000	0.000*
							(0.000)	(0.000)
year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
continent FE	Yes	No	No	No	No	No	No	Yes
land features	No	Yes	No	No	No	No	No	Yes
geographic controls	No	No	Yes	No	No	No	No	Yes
legal origin	No	No	No	Yes	No	No	No	Yes
culture	No	No	No	No	Yes	No	No	Yes
N	209	205	207	209	208	209	209	204
Number of countries	107	105	106	107	106	107	107	104
Adj. R-squared	0.230	0.173	0.165	0.208	0.305	0.158	0.399	0.543

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.4 Conclusions

This paper presents a model of an urban agglomeration that explains the differences in urban population densities between developed and developing countries. First of all, citizens in poorer countries are more likely to be financially constrained to afford formal housing. On the contrast, higher average wages and human capital levels in high-income regions prevent individuals from settling informally. Secondly, there is some evidence that higher income is associated with more housing consumption. This relation holds despite the fact that per unit housing rents are expected to increase as countries develop. The straightforward explanation for this fact would be that incomes grow at a higher rate than rents. Overall, lower population densities in the cities of developed world can be partially explained by larger housing stock. However, this factor alone cannot explain the physical shapes of the cities. As populations of cities increase, urban areas are expected to expand. A city can grow tall or sprawl wide to host the same number of residents. Assuming that land and physical capital are needed to produce housing, and employing general economic intuition we would expect tall buildings in the areas with relatively expensive land, and flat, but sprawled agglomerations, when land is cheap. However, in reality cities in developed countries, where capital-land price ratio is supposed to be low, occupy larger land areas than agglomerations in developing countries, where land is expected to be much cheaper than capital. This paper points out and explains this obvious discrepancytr between the existing literature on the matter and the real data. The theoretical model described in this paper considers necessary improvements to agricultural land that are required before erecting houses. Abundance of public infrastructure in the city determines the intensity of land improvements that are needed to be done privately. If a local government provides public goods at a high scale, connecting additional lot of land to the existing network is relatively cheap. However, when amenities are scarce, substantial investments are needed to be done privately in order to make a land unit suitable for housing. As a result, the presented model is capable of explaining what makes a city sprawl wide or grow tall.

This paper fills the gap in the existing theoretical literature. These findings can be useful for future research and also can help shaping efficient policy for city government. Extremely densely populated urban areas waste utility and might be physically harmful for its residents. This model formally demonstrates potential ways to deal with the problem of overconcentration. Local government can allocate funds in different ways to decrease population concentration. First of all, higher incomes allow more housing consumption. This fact is quite obvious, but increase in the dwellings size not necessarily leads to population dilution, as buildings can just grow taller. Moreover, the model demonstrates that population concentration might even increase. If a local budget grows slower than individual incomes, due to whatever reason, a city will grow tall, increasing the number of residents per a unit of land. Furthermore, the local taxes have an ambiguous effect: lower taxes allow citizens spend more on housing, but decreased public funds constrain the urban sprawl. This aspect is an important topic for the future research. Finally, a city government can facilitate land area growth in two ways. First option is to subsidise cultivation of new urban land (namely, decrease the price of "raw", agricultural land or lower the interest rates). However, further development of public infrastructure can serve the same purpose, as it increases the effectiveness of necessary private investments into new land. Determining which of these two strategies is more efficient and more rigorous empirical testing of the model are promising topic for the further research.

Appendix

Data sources of used variables

Variable urban population density is calculated by the author using the Global Human Settlement Layers by Pesaresi and Freire (2016) the Gridded Population of the World dataset (CIESIN 2016) using Google Earth Engine. Land productivity is calculated using the data for value added in agriculture and agricultural land provided by FAOSTAT. Lending interest rate and real interest rate, public goods variables, total population are taken from The World Bank (2018b). Marginal product of capital and capital productivity are calculated using the data from the Penn World Table (Feenstra et al., 2015). Urban population is calculated from the Gridded Population of the World dataset (CIESIN 2016) using Google Earth Engine - this variable was not employed explicitly in the analysis, but was used for calculation of non-agricultural per capita GDP and urbanisation rate. Housing share is calculated using the ICP data (The World Bank, 2018a). Major religion shares, legal origin, average land elevation, standard deviation of land elevation, terrain roughness, land suitability for agriculture, Gini of land suitability, geographic centroid latitude and longitude, average temperature and precipitation levels, share of land area near to a waterway, share of population of European descent and ethnic fractionalisation are taken from the Ashraf and Galor (2013) dataset. Total land area is taken from FAOSTAT.

Table A2.1: Summary statistics of selected variables

Variable	Obs	Mean	Std. Dev.	Min	Max
urban density	504	2642.436	1440.685	0	11901.78
land productivity	437	2095.995	8913.347	2.390	143271.5
lending interest rate	308	32.071	271.535	1.143	4774.525
real interest rate	302	7.191	13.554	-97.616	67.159
MPK	381	0.333	1.867	0.042	35.026
capital productivity	504	0.711	3.981	0.116	85.164
public goods	383	82.169	21.360	25.906	100
urban GDP	362	11458.57	21233.62	65.079	185342.1
urbanisation rate	498	0.713	0.164	0	1
population (mln.)	498	36.7	133	59326	1370
housing share	493	0.151	0.077	0.024	0.482

Table A2.2: Correlation matrix of selected variables

	density	land productivity	lending interest rate	real interest rate	MPK	capital productivity	public goods	urban GDP	urbanisation rate	population	housing share
density	1.000										
land productivity	0.344	1.000									
lending interest rate	0.024	-0.124	1.000								
real interest rate	0.000	0.050	-0.122	1.000							
MPK	-0.002	0.283	0.025	0.095	1.000						
capital productivity	-0.008	0.275	0.008	0.078	0.999	1.000					
public goods	-0.378	0.298	-0.033	-0.012	-0.332	-0.290	1.000				
urban GDP	-0.332	0.380	-0.281	0.015	-0.243	-0.262	0.411	1.000			
urbanisation rate	0.396	-0.005	-0.005	0.024	-0.169	-0.1370	0.192	-0.163	1.000		
population	0.134	-0.026	-0.014	-0.013	-0.021	-0.018	0.067	-0.043	0.181	1.000	
housing share	-0.127	-0.088	-0.038	0.041	-0.016	-0.003	0.157	-0.201	-0.034	-0.100	1.000

Table A2.3: Determinants of urban density. OLS estimates.

Dependent variable:	urban population density			
	(1)	(2)	(3)	(4)
land productivity	0.070	0.064	0.110**	0.105**
	(0.054)	(0.054)	(0.045)	(0.044)
public goods index	-24.825***	-24.760***	-19.967***	-19.586***
	(5.043)	(5.074)	(5.293)	(5.318)
MPK	91.384	79.590	-140.601	-118.856
	(293.322)	(331.692)	(286.248)	(322.155)
non-agricultural GDP p.c.			-0.014***	-0.016***
			(0.003)	(0.003)
year FE	No	Yes	No	Yes
N	266	266	266	266
Number of countries	104	104	104	104
Adj. R-squared	0.185	0.187	0.254	0.267

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.4: Determinants of urban density. OLS estimates.

Dependent variable:	urban population density			
	(1)	(2)	(3)	(4)
land productivity	0.100*	0.094*	0.132***	0.125***
	(0.050)	(0.050)	(0.047)	(0.045)
real interest rate	3.646	3.338	4.087	3.532
	(5.512)	(5.610)	(5.748)	(5.833)
public goods index	-19.806***	-19.417***	-15.418**	-14.767**
	(6.459)	(6.457)	(6.671)	(6.705)
non-agricultural GDP p.c.			-0.016**	-0.017***
			(0.006)	(0.006)
year FE	No	Yes	No	Yes
N	210	210	209	209
Number of countries	107	107	107	107
Adj. R-squared	0.112	0.117	0.145	0.155

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2.5: World's 40 largest urban land areas

	Country	Urban Area	Population estimate	Land area (km^2)	Population density
1	United States	New York	20,685,000	11,642	1,800
2	Japan	Tokyo-Yokohama	37,750,000	8,547	4,400
3	United States	Chicago	9,185,000	6,856	1,300
4	United States	Atlanta	5,120,000	6,851	700
5	United States	Los Angeles	15,135,000	6,299	2,400
6	United States	Boston	4,490,000	5,325	800
7	Russia	Moscow	16,570,000	5,310	3,100
8	United States	Dallas-Fort Worth	6,280,000	5,175	1,200
9	United States	Philadelphia	5,595,000	5,131	1,100
10	United States	Houston	6,005,000	4,828	1,200
11	China	Beijing	20,390,000	3,937	5,200
12	China	Shanghai	22,685,000	3,885	5,800
12	Japan	Nagoya	10,035,000	3,885	2,600
14	China	Guangzhou-Foshan	18,760,000	3,820	4,900
15	United States	Detroit	3,660,000	3,463	1,100
16	United States	Washington	4,950,000	3,424	1,300
17	Indonesia	Jakarta	31,320,000	3,225	9,700
17	United States	Phoenix	4,295,000	3,225	1,300
19	Japan	Osaka-Kobe-Kyoto	16,985,000	3,212	5,300
20	United States	Miami	5,820,000	3,209	1,700
21	United States	Seattle	3,475,000	2,989	1,100
22	France	Paris	10,870,000	2,845	3,700
23	United States	San Francisco-San Jose	5,955,000	2,797	2,100
24	Brazil	Sao Paulo	20,605,000	2,707	7,600
25	Argentina	Buenos Aires	14,280,000	2,681	5,300
26	Germany	Essen-Dusseldorf	6,675,000	2,655	2,500
27	United States	Minneapolis-St. Paul	2,795,000	2,647	1,000
28	South Africa	Johannesburg-East Rand	8,655,000	2,590	3,300
28	South Korea	Seoul-Incheon	23,575,000	2,590	9,100
28	Thailand	Bangkok	15,315,000	2,590	5,900
31	Australia	Melbourne	3,955,000	2,543	1,500
32	United States	Tampa-St. Petersburg	2,660,000	2,479	1,000
33	United States	St. Louis	2,195,000	2,393	900
34	United States	Pittsburgh	1,730,000	2,344	700
35	Canada	Toronto	6,550,000	2,287	2,900
36	US: Puerto Rico	San Juan	2,135,000	2,246	1,000
37	India	Delhi	25,735,000	2,163	11,900
38	Malaysia	Kuala Lumpur	7,365,000	2,124	3,500
39	Mexico	Mexico City	20,230,000	2,072	9,800
40	United States	Orlando	2,125,000	2,046	1,000

All data is from Demographia (2016). Population densities indicate number of citizens per km^2.

Table A2.6: World's 40 most densely populated areas

	Country	Urban Area	Population estimate	Land area (km^2)	Population density
1	Bangladesh	Dhaka	16,235,000	368	44,100
2	Pakistan	Hyderabad	2,990,000	73	41,200
3	India	Vijayawada	1,775,000	57	31,200
4	Bangladesh	Chittagong	3,250,000	111	29,200
5	India	Mumbai	22,885,000	881	26,000
6	China: Hong Kong	Hong Kong	7,280,000	285	25,600
7	India	Aligarh	1,050,000	41	25,300
8	China: Macau	Macau	655,000	26	25,300
9	Syria	Hamah	1,300,000	52	25,100
10	Somalia	Mogadishu	2,265,000	91	25,000
11	India	Surat	5,685,000	233	24,400
12	Pakistan	Karachi	22,825,000	945	24,100
13	Syria	Al-Raqqa	745,000	31	24,000
14	India	Kannur	2,160,000	91	23,800
15	India	Ranchi	1,280,000	57	22,500
16	India	Malegaon	675,000	31	21,700
17	Congo (Dem. Rep.)	Tshikapa	730,000	34	21,700
18	India	Gwalior	1,235,000	57	21,700
19	India	Madurai	1,615,000	75	21,500
19	India	Rajkot	1,615,000	75	21,500
19	Philippines	General Santos	1,615,000	75	21,500
22	Indonesia	Tasikmalaya	830,000	39	21,400
23	India	Ahmadabad	7,410,000	350	21,200
24	India	Guntur	765,000	36	21,100
25	India	Kozhikode	2,495,000	119	20,900
26	India	Muzaffarnagar	595,000	28	20,900
27	India	Asansol	1,335,000	65	20,600
27	India	Salem	1,015,000	49	20,600
29	Syria	Hims	1,705,000	83	20,600
30	India	Firozabad	585,000	28	20,500
31	India	Saharanpur	790,000	39	20,300
32	Pakistan	Faisalabad	3,675,000	181	20,300
33	Pakistan	Larkana	520,000	26	20,100
34	Congo (Dem. Rep.)	Bukavu	880,000	44	20,000
34	India	Moradabad	1,035,000	52	20,000
36	Egypt	Al Mahallah al Kubra	510,000	26	19,700
37	India	Bhavnagar	660,000	34	19,600
38	Congo (Dem. Rep.)	Kinshasa	11,380,000	583	19,500
39	Syria	Lattakia	795,000	41	19,200
40	India	Siliguri	840,000	44	19,100

All data is from Demographia (2016). Population densities indicate number of citizens per km^2.

3 Burnt by the sun. Explaining suicide in developing countries

More than 800.000 people commit suicide every year - around one person every 40 seconds. Despite the popular stereotypes, around 75% of suicides occur in low- and middle-income countries. Nevertheless, the vast majority of studies of suicide is conducted in developed countries undermining their external validity. This paper takes advantage of the macro data and analyses fundamental factors determining cross-country variations in suicide rates. The main novelty of the study is introduction of ultra-violet radiation as a key factor affecting incidence of suicide. As a result, this paper documents a striking contrast between the major determinants of suicide in developed and developing countries. While biology explains a substantial part of cross-country variation in high-income countries, suicide rates in the developing world are primarily driven by a range of social factors (e.g., adult mortality). These findings are crucial for designing an efficient suicide-preventing policy.

JEL codes: I12, I18, Q51, Q54
Keywords: suicide, health, solar radiation, adult mortality

Introduction

Suicide is a personal tragedy. It affects not only individuals that have decided to take their own lives, but also their families, friends, colleagues. The taboo and social stigma around suicide keeps the topic away from an active public discussion and fosters the emergence of unjustified and misleading stereotypes. However, suicide is clearly a public health problem, as 800,000 people die by suicide every year (WHO, 2014). Carefully tailored mental health plans can reduce the number of self-murders. To develop an efficient strategy to tackle suicide we need to better understand the factors causing it. This paper focuses on the fundamental determinants of suicide that can explain the variation in suicide rates across countries.

Despite the fact that suicide is a personal decision taken at the micro level, there is still a number reasons to study this phenomenon at the cross-country level. First of all, one has to consider the difference in factors that can be captured at different levels of analysis. There exist several studies conducted at the individual level (Hakko et al., 1998; Cutler et al., 2001; Phillips et al., 2002; Andres and Halicioglu, 2010; Daly et al., 2013; Case and Deaton, 2015a), but, due to numerous reasons (e.g., privacy concerns or local ethics norms), individual data on suicide is rarely available, so more studies employ some degree of data aggregation. For example, Koo and Cox (2008), Chan et al. (2014), Madianos et al. (2014) conduct their analyses using the evolution of national suicide rates over time, while Ahlburg and Schapiro (1984), Mathur and Freeman (2002) and Case and Deaton (2015b) analyse the variation at the U.S. state level. Studies employing aggregated data (including this paper) do not aim to answer the question *why those particular people* committed suicide, but rather *why so many* did it in a certain region. In other words, cross-country analysis does not identify the individual triggers, but allows to capture some more fundamental factors, like ultra-violet radiation (UVR) or mortality considered in this paper, that determine the probability of an average citizen of the country to commit suicide.

Moreover, studying suicide from a cross-country perspective can help to overcome concerns about the out-of-sample validity of the results. The majority of related existing studies are conducted using figures from developed countries (e.g., Andres 2005; Andres et al. 2011; Okada and Samreth 2013; Jalles and Andresen 2015), since statistical capacities there provide a wider range of data. However, WHO (2014) argues that 75% of the World's suicides occur in low- and middle-income countries. Partially this can be noticed if we plot suicide rates on the map of the World (Fig.3.1). Moreover, conclusions drawn from studies conducted in high-income countries are not necessarily applicable to developing ones. Therefore, the driving forces of suicide can differ across countries even more than within a country, especially if we speak about such fundamental factors as culture or religion (Durkheim, 1951; Helliwell, 2007). A range of prominent studies by Gunnell and

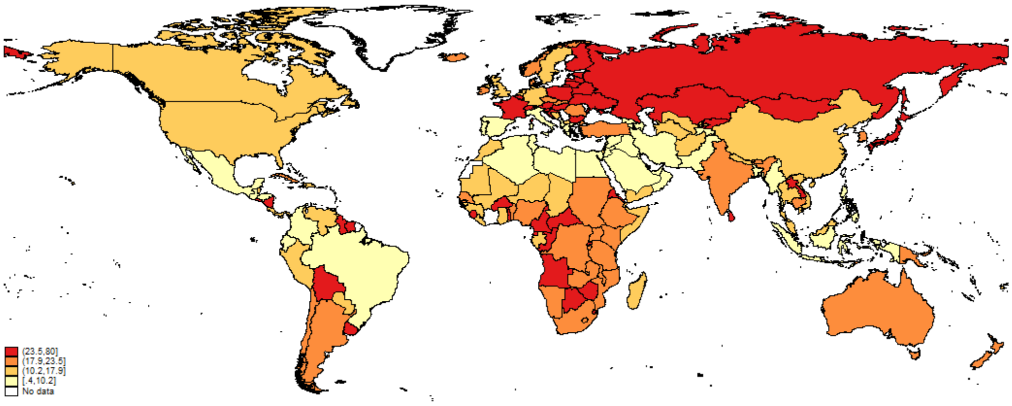


Figure 3.1: Age-standardised male suicide rate in 2000

Eddleston (2003), Das (2011) Gruere and Sengupta (2011), Mahfoud et al. (2011), Toprak et al. (2011), Patel et al. (2012), Ramim et al. (2013), Qaisar et al. (2014) and Machado et al. (2015) do a great effort to overcome this problem focusing on developing countries, but they typically concentrate either on particular drivers or countries. However, this approach limits the external validity of the findings for the states not included into the analysis. Similar to Neumayer (2003); Milner et al. (2012) this article considers differences between countries at different stages of economic development for explaining variations in suicide rates.

This article contributes to the literature that links natural factors with health outcomes (e.g., Dalgaard and Strulik (2015); Andersen et al. (2016); Dalgaard and Strulik (2016)) and human behaviour (e.g., Acemoglu et al. (2001); Spolaore and Wacziarg (2009); Ashraf and Galor (2013); Strulik (2017))[1]. I argue that between-country variation in UVR exposure is a potential fundamental determinant of suicide that acts through biological channels that are, unlike cultural aspects, valid in all countries of the world. This effect is robust to inclusion of a set of socio-economic variables, like income, employment, religion or mortality. Nevertheless, some socio-economic variables, especially adult mortality, also have a significant effect on suicide rates. Moreover, this study argues that the lack of UVR distorts normal functioning of the body and promotes suicide in developed countries that mostly lie in darker areas, while adult mortality contributes to higher suicides in developing countries, where people expect to have shorter lives.

This paper is organised the following way. Section 2 summarises existing medical studies to distinguish factors that can potentially affect suicide rates and integrates them into a simple economic model that simplifies the understanding of complex biological processes for the purpose of this study. Section 3 describes empirical methodology and data. Section 4 presents the results of regression analysis. Finally, Section 5 concludes the paper.

[1]For a more detailed review see Spolaore and Wacziarg (2013)

3.1 Theory of suicide

Suicide can be explained by a wide variety of factors. Moreover, it is attempted to be explained by different fields of science. On one hand, medicine (namely psychiatry) and psychology try to explain the phenomenon from physiological point of view. On the other hand, social sciences as sociology and economics try to analyse human behaviour and factors affecting it. Sometimes these two strands of science go against each other. In this paper I attempt to link major biological factors established in the medical literature with existing economic theories of suicide.

3.1.1 Biology of suicide

Despite often being a striking event for families and friends, suicides do not happen completely out of the blue. Typically, victims of suicide indicate their need for help in some ways even though that might be not noticed by surrounding people (WHO, 2014). Given this fact I can suppose that individuals committing suicide are likely to be affected by a particular condition that does not kill them immediately. Medical literature suggests that one of the major factors developing suicidal behaviour is depression. There exist numerous studies demonstrating positive link between depression and suicide (see Beskow (1990) for a detailed summary of the literature on the topic). People affected by depression are at a high risk of committing suicide (Lynch and Duval, 2011). As a result, I can consider suicide as a complication of depression. Even though there is no evidence for existence of one major cause of depression, there are several factors associated with it. Once we identify these potential causes of depression, we can analyse if the same drivers can be used to predict suicide rates.

First of all, depression is normally provoked by some stressful event (Burrows, 1977; Risch et al., 2009). However, individuals can react to these occasions differently: not everyone becomes depressed after loosing a job or having a divorce. Besides individual preferences, the likelihood of a certain stressful experience initiating depression might be affected by genetics. All events in our life provoke an influx of serotonin into the space between neurons to activate serotonin (5-hydroxytryptamine, 5-HT) receptors (Beliveau et al., 2017). Some of the 5-HT receptors are responsible for such functions as regulation of anxiety or mood(Tatarczynska et al., 2004; Young, 2007). Different combinations of these two functions can yield emotions that people describe as "happiness" or "unhappiness". However, a functional polymorphism (basically, difference in forms) of the serotonin transporter (5-HTT) can regulate the sensitivity of individuals to different stressful events (Caspi et al., 2003; Kendler et al., 2005; Goldman et al., 2010; Karg et al., 2011). In other words, the same event can be very stressful and cause depression of one individual, but not affect the other and the reason for it is individual's genetics. As a result, we can expect a particular (genetically determined) form of 5-HTT to be suicide-promoting (Lee et al., 2015).

Besides genetics, there exist other factors that make people to be more prone to depression. Some studies argue that lower availability of serotonin may be a crucial factor in depression(Meltzer, 1989; Kambeitz and Howes, 2015). However, currently there is a big debate in the literature if this is truly the case (Cowen and Browning, 2015; Healy, 2015). I do not aim to contribute to this discussion, but rather use *potential* theoretical mechanism that has been successfully applied for therapy. A range of studies suggests that serotonin is positively associated with light. By "light" we can mean sun light in general (Petridou et al., 2002; Lambert et al., 2002; Sansone and Sansone, 2013) or more precisely defined ultra-violet radiation (UVR) (Iyengar, 1994; Zawilska et al., 2007). If we believe serotonin to be efficient for depression prevention, we can expect exposure to solar radiation to be negatively correlated with depression. Moreover, the existing literature suggests that even if serotonin levels play no role in depression occurrence, there must be other ways for light to affect people's mood. Currently, bright light therapy (BLT) is an accepted treatment option for depression (Mårtensson et al., 2015). Both exposure to BLT (Eastman et al., 1998; Young, 2007; Al-Karawia and Jubair, 2016) or solar radiation (Eastman, 1990; Gambichler et al., 2002) can be efficient for treating various types of depression. Similar beneficial effects of the sun light are observed if suicide rates are considered (Lambert et al., 2003). As a result, regardless of the particular channel, we can expect higher levels of UVR exposure to be negatively associated with suicide rates[2].

To sum this section up, the existing studies of depression and suicide claim that various stressful life events can provoke depression. Acute forms of depression can lead to suicide. However, individual vulnerability depends on a set of various biological factors. In this paper I concentrate on the two potential drivers, that I can, at least, try to measure at the population level: genetics and environment (namely, UVR). Firstly, some people (or populations) might be genetically predetermined to be more stressed, hence, more vulnerable to depression. Secondly, lack of UVR can contribute to higher depression rates that, in turn, result into more suicides.

3.1.2 Economics of suicide

Even though biology plays crucial role in the mechanisms of suicide, we have to consider other factors that can be potential drivers. Different economic and social conditions can be considered as stress factors that trigger depression or suicide. Since Durkheim economists

[2]It is important to remember that excessive exposure to solar radiation can be harmful for general health: skin cancer or cataract are the most common threats (Lucas et al., 2006). Even though these diseases do not necessarily cause immediate depression, they can raise stress levels that, in turn, cause depression. This issue will be addressed in the Section 3.3.1.

and sociologists tend to explain suicides with various models of human behaviour. These theories do not necessarily oppose the biological mechanisms, moreover, they can be integrated into physiological models of suicide. Existing economic literature often lists low incomes (Helliwell, 2007; Andres and Halicioglu, 2010; Daly et al., 2013) and unemployment (Andres, 2005; Chang et al., 2013; Jalles and Andresen, 2015) among the drivers of suicide. From biological point of view, these can be seen as stressful life events. If we talk about stress, we normally imagine some sudden and unexpected changes in life, however, persistent states (e.g., long-term poverty) can also raise stress levels (Baum, 1990). Furthermore, Minkoff et al. (1973); Kovags et al. (1975); Beck et al. (1985) identified hopelessness (more precisely defined as the cognitive element of negative expectations), which is a component of the depression syndrome, as a stronger indicator of suicidal intent than depression itself. In other words, when individuals do not see perspective in their life, they are more likely to commit suicide.

Of course, hopelessness is the state that should be described individually. However, there are some factors that might be quite universal for the majority of the people. For example, the World Values Survey indicates that more than 62 percent of the total surveyed across the world are worried about losing their job and more than 87 percent find that work plays an important role in their lives (Inglehart et al., 2014). Assuming that to an extent generalisation of preferences is reasonable, we can recall the seminal paper by Hamermesh and Soss (1974) that presents a model of suicide behaviour. In this paper I slightly modify the original model to adjust it for the purpose of this study.

First, let us assume that the utility an individual derives at time m is a function of income Y:

$$U_m = \alpha Y_m, \tag{3.1}$$

where $\alpha \in (0, 1)$ is a biological parameter. I plug income into the utility function, as it allows consumption of more and wider variety of goods that should lead to a higher utility level. However, by Y we can also consider other things that derive us positive utility, like family, friends or achievements of the favourite football team, so it should be rather considered as a vector of variables. However, individuals also incur costs of maintaining themselves alive at particular age a: $K(a)$. As people become older, the age-related disutility becomes larger: medical expenditures, physical or emotional pain rise with age ($\frac{\partial K}{\partial a} > 0$). Summing it up, the individual's lifetime utility at a certain age a looks the following way[3]:

$$Z(a,Y) = \int_a^{\omega} e^{-r(m-a)} U_m P(m) dm - K(a), \tag{3.2}$$

where r is the private discount rate, ω is the highest attainable age and $P(m)$ is the probability to survive until m. However, individuals do not necessarily kill themselves, when their lifetime utility is negative. Hamermesh and Soss (1974) assume that individuals might have some taste for living, b, making them refrain from suicide even when utility is negative. If we translate b for the whole country, this variable can include such major factors as religion that typically forbids suicide or social norms ("What will the neighbours say?"). As a result, we can expect a representative individual to commit suicide when:

$$Z(a,Y) + b < 0. \tag{3.3}$$

Combining Eq.(3.1) and (3.2) we can analyse how different variables affect lifetime utility and, hence, potential decision of a respective individual to commit suicide. For example, as income is supposed to increase lifetime utility, we can expect that higher incomes are associated with less suicide. On the contrast, older populations are expected to be more prone to suicide, as they have less time to enjoy consumption and higher maintaining costs. The relationship between the suicide rate and its potential drivers will be tested empirically in the next section. It is important to note that even though Eq.(3.3) models suicide as a rational decision. If people are so unhappy why do not we just let them kill themselves? However, as it was demonstrated above, perception of various life events (in this case, utility-enhancing) can be distorted by the environment. As a result, from an individual's point of view committing suicide can be perfectly rational, but this rationality is based on a set of determinants, some of which can be changed. Identifying these factors can be useful for suicide prevention in the future.

3.2 Empirical strategy

The main goal of this paper is to identify some fundamental drivers of suicide that can affect suicide rates at the country level. Even though many factors are purely individual, some still might be valid for a large amount of people residing in the country. First of all, I process the equations of the previous section and extract those potential drivers that can

[3] The only difference to Hamermesh and Soss (1974) is that I include into lifetime utility only the cost of maintaining oneself at age a: $K(a)$. In the original paper authors have included costs as a negative factor directly into utility at every time m (here, in Eq.(3.1)), so that lifetime utility Z included lifetime maintaining costs, too. I consider my formulation slightly more realistic, as it prevents young individuals from committing suicide even if they are aware that at some point in life maintaining themselves alive will be extremely tough.

be measured at the country level.

To start with, I analyse Eq.(3.1) to determine the factors that might affect utility. Income is measured with GDP per capita, but I can also analyse how it translates into utility. For this purpose I take a closer look at α. This biological parameter can be affected in several ways. First of all, genetics might play a role. Some populations are more vulnerable to particular genetic diseases (Nakatsuka et al., 2017), like sickle-cell disease (Weatherall and Clegg, 2001; Asnani et al., 2011) or Tay-Sachs disease (Myerowitz and Costigan, 1988; McDowell et al., 1992). Given this, we can expect a particular type of 5-HTT polymorphism that distorts normal serotonin circulation might be more prevalent in some populations. Unfortunately, I cannot obtain data on prevalence of this particular type of serotonin transporter for representative samples of countries' populations. However, I can use some proxy that might capture the probability of presence of a certain type of 5-HTT polymorphism that increases the likelihood of committing suicide. For this I employ the ancestry-adjusted predicted genetic diversity variable from Ashraf and Galor (2013). Ramachandran et al. (2005) have demonstrated that migratory distance from Eastern Africa can be a good indicator of the genetic diversity of indigenous populations. To account for more recent migration the genetic diversity predicted by the distance from Africa is adjusted by applying the World Migration Matrix by Putterman and Weil (2010). As a result, adding the genetic diversity variable I attempt to answer the following question: does genetic diversity decrease the presence of the suicide-promoting form of 5-HTT polymorphism in the national genetic pool? In other words, if people in the country have a broader genetic pool, the incidence of a particular gene that determines the certain form of 5-HTT (so that it increases the probability of committing suicide) can be relatively lower. This hypothesis will be tested in the regression analysis.

However, as was described in the previous section, besides genetics the α parameter can be influenced by the insensitivity of UVR in the country. If we take the hypothesis that serotonin levels are negatively associated with depression and, hence, suicide as valid, we can expect UVR (that promotes serotonin production) to be suicide-decreasing. If the serotonin hypothesis is not valid, we cannot neglect beneficial effect of light for depression alleviation. Regardless of the mechanism, the previous section has shown motivation for the following hypothesis: higher levels of UVR increase utility through bigger α and decrease suicide rates. To test this statement I employ the average daily ambient ultraviolet radiation level data from the World Health Organisation. This variable is adjusted for population distribution across countries and is measured in Joules per square metre. WHO provides data only for 2004, however Andersen et al. (2016) have argued in their paper that the UVR levels at the surface are relatively constant and "current comparative UVR levels are likely to be an excellent indicator of UVR conditions a few centuries ago"(Andersen et al., 2016, p.1349). The concerns about the depletion of ozone layer should not cause serious bias for my estimations. The point is that there are three types of UVR that reach the Earth: UVC with wavelengths from 200 to 280 nm, UVB - from 280 to 315 nm, and UVA - from 315 to 400 nm (Calbo and Gonzalez, 2005). The smaller the wavelength, the deeper the radiation can penetrate human bodies and more harm it can cause. Fortunately for us, UVC radiation is completely filtered out by the atmosphere (Calbo and Gonzalez, 2005). UVB is partially filtered by the ozone layer, while UVA is mostly unaffected by ozone depletion (Rousseaux et al., 1999). Out of the three types of UVR UVA is particularly important for my research, as it seems to enhance serotonin production (Gambichler et al., 2002; Zawilska et al., 2007). The WHO data indicates the amount of energy reaching the ground and from NASA (2002)[4] we know that most of the energy reaching the surface is coming exactly from UVA radiation. Finally, depression normally takes some time to develop, especially to the stage that leads to suicide, but the timing might differ across individuals. Given that, I do not expect fluctuations in values (also lagged ones) of UVR in a particular year to explain between-year variations in suicide rates. Instead, I aim to capture the deterministic effect of the general UVR level on suicides. For this reason, time-invariant measure of solar exposure appears appropriate.

The two variables already described in this section serve as potential determinants of α - the parameter that regulates the extent to which income and other factors translate to utility in Eq(3.1). However, from Eq.(3.2) we can expect other factors to be affecting lifetime utility. First of all, the probability of survival to the age m. If an individual expects to live a short life, he might value his future utility way lower. Alternatively, this can be thought about from the psychiatric point of view: a higher probability of dying might raise stress levels and provoke depression. I proxy $P(m)$ with adult mortality rate (per 100 people). The higher mortality rate would indicate a lower probability of a respective individual surviving to any age m. Besides that, mortality also affects the maximum attainable age ω, so it captures this channel, too. To avoid any potential reverse causality of suicide on adult mortality and bad control problem, I subtract suicides from the total mortality rate. As a result, adult mortality employed further in this paper accounts for all causes except suicide. The final component of the lifetime utility function from Eq.(3.2) is the cost of maintaining oneself alive. To measure the cost in terms of utility, rather than money, I employ the variable "years lived with disability" (YLD) provided by WHO. It captures the severity of the illnesses individuals live with before death, so can serve as an indicator of maintenance costs. WHO reports this data in absolute terms, so to make the values comparable I scale them down by the total population in the country (separately for men and women). As a result I get a per capita YLD variable that indicates how many years of a healthy life a representative individual loses due to disability. To test how income inequality affects suicide rates I include it into the regression equation. To measure inequality I use data for

[4] Last accessed on 08.06.2108

Gini coefficients from the World Bank[5]. Speaking about the reservation component b from Eq.(3.3), I come up with a set of proxies to measure it at the country level. As was already mentioned in the previous section unemployment rates might be important for prediction of suicide rates. However, besides lowering effect on expected income or raising stress levels, unemployment might have the opposite effect. For example, the depressed, but employed person might refrain from committing suicide if he knows that he has to support his family, especially at the times when jobs are scarce. Which effect prevails will be tested in the regression analysis. Finally, religion is a very important factor that might affect the suicide decisions. I control for it constructing a set of religious dummies. Pew Research Center's Forum on Religion & Public Life (2012) provides data for shares of population affiliated with one of the world's major religious groups: Christians, Muslims, Hindus, Jews, Folk, Other or not affiliated to any. I assign countries to a particular religion if more than 50 percent of their population are affiliated to this confession. Hence, I split countries into Christian, Muslim, Buddhist, Jewish, Hindu and Unaffiliated. Countries without prevailing religion (none of the religious groups constitutes 50 percent) are defined as Mixed. The latter group of countries is excluded from the regressions as a base group.

Finally, I need to explain the dependent variable used for the regression analysis. The WHO provides several variables measuring suicide rate in the country. First of all, I take values for males and females separately. This allows to increase precision and analyse if determinants of suicide are in any way gender-specific. Secondly, it is natural to expect that countries with larger populations simply experience more suicides, so the rates are calculated as number of suicides per 100.000 people. Finally, if we expect older people to kill themselves more often, due to higher maintaining costs and lower remaining lifetime utility, we have to address this fact. WHO (2014) provides data on suicide rates disaggregated between age groups. Unfortunately, this data is available only for 2012, but for a large number of countries. Table 3.1 presents crude suicide rates disaggregated between age cohorts separately for OECD[6] and non-OECD countries. Advantage of using crude suicide rates is that they can also be considered as simple probabilities of a respective age group to commit suicide. In Table 3.1 suicide rates generally increase with age for males and females both in OECD and non-OECD countries, as existing studies predict. Interestingly, children (aged 4-15) and elderly (70+) tend to commit suicide more often in developing countries, while adults between 30 and 69 kill themselves more often in OECD member-states. Overall, crude suicide rates for the whole population (i.e., not controlling for age structure) are significantly higher in developed countries. However, these numbers are misleading. For example, teenagers and young adults (aged 15-29) on average have similar suicide rates both

[5]Due to the fact that this data has lots of gaps in the study period, I extend the dataset employing five-year average values. I assume that inequality is persistent enough and does not vary dramatically within 5 years. This assumption is supported by the studies of Mookherjee and Ray (2003); Williamson (2010); Bertola et al. (2010); Covarrubias et al. (2015).

[6]See the list of OECD countries in Appendix.

in OECD and non-OECD. In contrast, this age cohort generally has a higher probability of committing suicide than a respective citizen of their country (regardless of the age - "all" line in Table 3.1) in non-OECD countries, and smaller one in OECD. This observation points at the fact that the crude suicide rate of the whole population is likely to be downward biased in developing (predominantly young) countries. Kids and teenagers (4-15 years) commit suicide extremely rarely regardless of the OECD status, but constitute a substantial share of population in developing countries. At the same time, elderly people are generally more prone to suicide and they are relatively over-represented in developed countries.

Table 3.1: Crude age-specific suicide rates (per 100,000), 2012

	OECD countries			non-OECD countries			
Age cohort	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Δ
Panel A: Males							
4-15	35	0.720	0.753	135	1.341	1.628	-0.621^{**}
15-29	35	14.963	6.447	135	13.578	11.436	1.385
30-49	35	24.934	10.810	135	16.004	15.671	8.931^{***}
50-69	35	30.851	16.257	135	20.887	18.934	9.964^{***}
70+	35	39.689	32.771	135	42.061	35.324	-2.372
all	35	21.311	9.766	135	12.098	10.833	9.214^{***}
Panel B: Females							
4-15	35	0.377	0.409	135	0.947	1.118	-0.570^{***}
15-29	35	4.534	3.046	135	5.253	5.183	-0.719
30-49	35	6.766	4.005	135	4.322	4.039	2.443^{***}
50-69	35	9.151	5.076	135	7.487	6.545	1.664
70+	35	9.869	11.401	135	19.914	24.305	-10.046^{**}
all	35	6.337	3.810	135	4.372	3.565	1.965^{***}

Δ indicates the difference of means between OECD and non-OECD countries. P-values are obtained after running a t-test of means. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To enable comparison of countries with different demographic profiles WHO provides age-standardised rates: a weighting function is applied such that all countries are standardised to the WHO Standard Population (WHO, 2014). This data is provided for years 2000, 2005, 2010 and 2015 for virtually all countries of the World, so it allows to cover a very substantial part of the globe. It is important to mention that suicide might be not properly reported, especially if we consider developing countries, where vital registration systems are not well-developed and social and religious norms might lead to under-reporting (or maybe over-reporting) of the actual suicide cases. The WHO evaluates the quality of the mortality data and assigns each country a value from 1 to 4, where smaller values indicated higher quality of the data. I include this parameter into regression to check if the countries with low-quality of statistics systematically under- or over-report suicides.

For regression analysis I primarily employ OLS estimator. For robustness checks I also use weighted least squares and fixed effects estimators that will be described in the next section. The main regression specification can be described by the following equation:

$$y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 \Gamma_{it} + \lambda_t + \mu_r + \varepsilon_{it}, \tag{3.4}$$

where y_{it} is the suicide rate in country i at year t, a vector X includes variables of primary interest for this paper: UVR (the average daily ambient ultraviolet radiation level), predicted genetic diversity, GDP per capita, Gini coefficient, unemployment and adult mortality rates, YLD and religious shares. The vector Γ includes a set of auxiliary controls: data quality parameter, latitude, temperature and precipitation. λ_t is a year fixed effect, μ_r is a set of regional dummies (the World Bank regions) and ε_{it} is an error term. The results of the estimations are presented in the next section.

3.3 Estimation results

Table 3.2 demonstrates the results of the estimation of the regression model described in Eq.(3.4). We see negative correlation between UVR and suicide rates, both for males and females. However, the relationship gets weaker, if we add additional controls and even disappears for females in column (6). One possible issue is that the relationship between suicide rates and UVR might be non-linear in the used dataset. If I run a non-parametric estimation (locally weighted scatterplot smoothing) of the relationship between suicide rates and UVR, I get clearly non-linear pattern of distribution of fitted values, as Fig.(3.2) demonstrates. Graphs for both male and female suicide rates remind the U-shaped relationship between UVR and suicides. We can easily test if this is actually the case, by adding the squared UVR term into regression described in Eq.(3.4).

Table 3.3 demonstrates that once we allow for a non-linear pattern, the relationship between UVR and suicides becomes strongly significant and robust to inclusion and exclusion of additional controls and fixed effects. Tables A3.1 and A3.2 in Appendix demonstrate that single coefficients are also generally robust to inclusion of other explanatory variables not included in the list of control variables. The difference in coefficients between genders should not be misleading: females generally have way lower suicide rates, so the coefficients partially account for this difference. Moreover, the results of the test of a U-shaped relationship described by Lind and Mehlum (2010) indicate statistically significant U shape in all regression specifications. Moreover, the UVR$_{min}$ values indicating the minimum of the function appear quite similar for males and females. Besides UVR values, suicide rates are significantly positively correlated with adult mortality. Higher adult mortality decreases

Table 3.2: Determinants of suicide. OLS estimates.

	male suicide rate			female suicide rate		
	(1)	(2)	(3)	(4)	(5)	(6)
UV-radiation	-3.692***	-5.367***	-4.690***	-0.679**	-1.194***	-0.981**
	(1.259)	(1.636)	(1.654)	(0.292)	(0.380)	(0.466)
genetic diversity	-40.533	-84.607*	-71.532	-29.175	-52.705**	-49.783*
	(36.980)	(45.870)	(46.983)	(18.151)	(25.077)	(25.619)
log(GDP per capita)	0.436	0.365	0.387	0.135	0.063	0.073
	(0.959)	(1.023)	(1.018)	(0.251)	(0.243)	(0.242)
inequality	-0.076	-0.208*	-0.172	-0.016	-0.038	-0.038
	(0.100)	(0.115)	(0.117)	(0.024)	(0.030)	(0.031)
unemployment	-0.339***	-0.289**	-0.303**	-0.069***	-0.056***	-0.062***
	(0.117)	(0.129)	(0.123)	(0.019)	(0.021)	(0.022)
adult mortality	0.804***	0.769***	0.764***	0.173***	0.128***	0.130***
	(0.129)	(0.133)	(0.130)	(0.042)	(0.041)	(0.042)
YLD	7.974	2.587	35.324	30.215	20.650	30.913*
	(92.964)	(87.129)	(84.614)	(19.141)	(19.021)	(18.638)
data quality	-1.751	-2.540*	-3.026**	0.107	-0.279	-0.375
	(1.271)	(1.521)	(1.504)	(0.327)	(0.372)	(0.363)
Christian	-2.414	-3.177	-4.554	-2.202	-2.097	-2.557*
	(3.909)	(3.648)	(3.835)	(1.831)	(1.453)	(1.405)
Muslim	-3.389	-2.816	-3.464	-1.093	-1.089	-1.312
	(4.069)	(4.026)	(4.096)	(1.860)	(1.472)	(1.449)
Buddhist	10.448	12.775*	11.906*	2.822	3.265*	3.052
	(8.251)	(6.801)	(6.818)	(2.379)	(1.960)	(2.016)
Jewish	-0.282	0.795	-0.645	-1.873	-2.319	-2.707
	(4.723)	(5.699)	(5.729)	(2.010)	(1.849)	(1.815)
Hindu	-4.919	-7.491	-8.797	0.863	-0.887	-1.288
	(4.013)	(7.257)	(7.472)	(2.202)	(2.686)	(2.689)
unaffiliated	-1.758	-2.685	-4.575	-1.107	-1.199	-1.668
	(4.383)	(4.067)	(4.044)	(2.221)	(2.050)	(2.027)
year FE	No	Yes	Yes	No	Yes	Yes
regional FE	No	Yes	Yes	No	Yes	Yes
geographic controls	No	No	Yes	No	No	Yes
N	420	420	415	420	420	415
Number of countries	142	142	140	142	142	140
Adj. R-squared	0.478	0.499	0.518	0.361	0.405	0.418

Unemployment, adult mortality and YLD stand for male unemployment, adult mortality and YLD in columns (1)-(3) and female unemployment, adult mortality and YLD in columns (4)-(6). Geographic controls include: temperature, precipitation and latitude. Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

evaluated lifetime utility: people expect to yield utility for a shorter timespan, and also evaluate future utility less, as the probability of living until then is rather low. Religious dummies have different effects: while predominantly Christian countries typically have less suicides than with mixed religious profile, male residents of Buddhist countries tend to commit suicide more often.

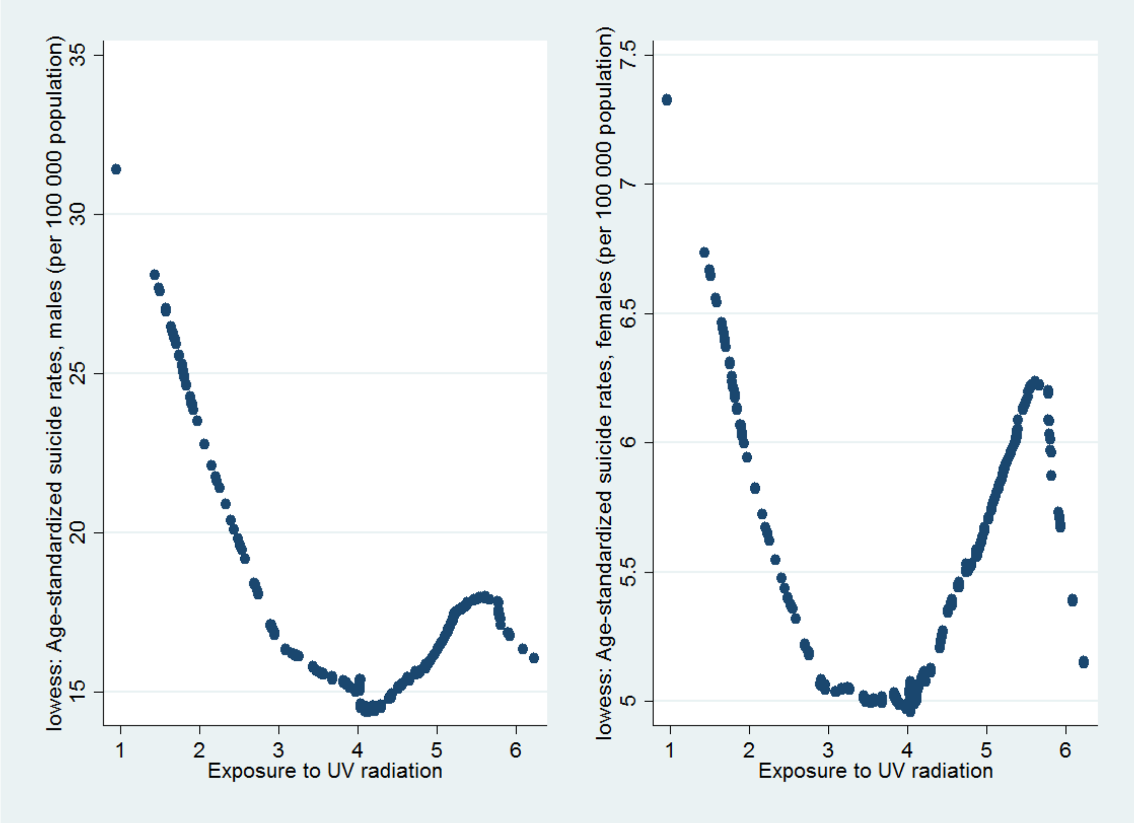


Figure 3.2: Suicide and UVR. Non-parametric estimation

Another interesting result of the regression estimations presented in Tables 3.2 and 3.3 is that coefficients for the genetic diversity variable indicate no significant effect on suicide rate. However, this result does not indicate whether genetics has an effect on suicide, but rather demonstrates that genetic diversity is not a suitable indicator of suicide-promoting genetic profile of the respective individual in a country. The incidence of suicide-increasing 5-HTT form might be not dependent on overall genetic diversity of the population and be simply more prevalent in some countries, as suggested by the medical literature on the topic. Goldman et al. (2010) present a detailed survey of the studies that analyse the distribution of 5-HTT polymorphism genotypes across different countries and ethnic groups. Their results suggest that the incidence of particular forms of the 5-HTT-related genes rather depends on the race. Given that genetic diversity is not solely determined by the race (for example, Kazakhstan and the Gambia have similar genetic diversity values and different racial profiles), the latter can be a more deterministic factor for suicide-related 5-HTT polymorphism. The countries' racial composition is partially captured by the regional effects and will be later fully controlled in the country fixed effects specification.

It is worth noting that Table 3.3 does not demonstrate strongly significant effect of income. First of all, GDP per capita is correlated with adult mortality. Even though the correla-

Table 3.3: Determinants of suicide. OLS estimates.

	male suicide rate			female suicide rate		
	(1)	(2)	(3)	(4)	(5)	(6)
UVR	-20.459***	-26.178***	-28.893***	-4.584***	-5.876***	-6.758***
	(3.607)	(4.181)	(5.608)	(1.043)	(1.061)	(1.454)
UVR^2	2.194***	2.603***	2.908***	0.514***	0.588***	0.696***
	(0.459)	(0.512)	(0.648)	(0.141)	(0.131)	(0.178)
genetic diversity	-34.598	-38.513	-26.253	-29.460*	-44.007*	-40.312*
	(33.147)	(45.181)	(41.608)	(17.518)	(24.836)	(23.594)
log(GDP per capita)	-0.139	-0.792	-0.953	0.015	-0.190	-0.236
	(0.943)	(1.061)	(1.131)	(0.249)	(0.253)	(0.280)
inequality	0.094	-0.010	-0.021	0.019	0.004	-0.003
	(0.096)	(0.112)	(0.117)	(0.024)	(0.030)	(0.031)
unemployment	-0.254**	-0.167	-0.227*	-0.045**	-0.031	-0.045**
	(0.113)	(0.125)	(0.122)	(0.021)	(0.021)	(0.021)
adult mortality	0.720***	0.719***	0.735***	0.153***	0.120***	0.126***
	(0.127)	(0.130)	(0.130)	(0.040)	(0.039)	(0.040)
YLD	-31.134	-75.226	-70.494	28.029	13.520	14.697
	(83.252)	(82.702)	(90.981)	(18.808)	(18.953)	(19.186)
data quality	-1.832	-3.120**	-3.267**	0.131	-0.393	-0.425
	(1.150)	(1.380)	(1.261)	(0.305)	(0.350)	(0.323)
Christian	-4.674	-4.747	-7.159**	-2.760	-2.519*	-3.231**
	(3.861)	(3.178)	(3.533)	(1.779)	(1.354)	(1.392)
Muslim	-3.750	-2.865	-3.989	-1.262	-1.144	-1.491
	(3.891)	(3.500)	(3.591)	(1.799)	(1.361)	(1.397)
Buddhist	11.477	11.333**	10.987*	2.940	2.814	2.665
	(7.627)	(5.525)	(5.617)	(2.255)	(1.822)	(1.884)
Jewish	1.457	0.663	-0.144	-1.253	-1.933	-2.249
	(4.512)	(4.742)	(4.737)	(2.008)	(1.655)	(1.684)
Hindu	-2.653	-6.493	-7.856	1.313	-0.726	-1.135
	(3.885)	(6.127)	(6.131)	(2.248)	(2.461)	(2.527)
unaffiliated	-4.771	-6.108	-7.511*	-1.831	-2.041	-2.447
	(4.493)	(3.695)	(4.216)	(2.347)	(2.033)	(2.221)
UVR_{min}	4.662	5.028	4.968	4.457	4.999	4.854
U-test p-value	0.004	0.012	0.007	0.012	0.022	0.020
year FE	No	Yes	Yes	No	Yes	Yes
regional FE	No	Yes	Yes	No	Yes	Yes
geographic controls	No	No	Yes	No	No	Yes
N	420	420	415	420	420	415
Number of countries	142	142	140	142	142	140
Adj. R-squared	0.534	0.560	0.578	0.402	0.448	0.466

Unemployment, adult mortality and YLD stand for male unemployment, adult mortality and YLD in columns (1)-(3) and female unemployment, adult mortality and YLD in columns (4)-(6). Geographic controls include: temperature, precipitation and latitude. Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

tion coefficient of -0.68 for males and -0.71 for females is above the threshold value of -0.8 indicating multicollinearity, as Gujarati (2003) suggests. Nevertheless, Tables A3.1 and A3.2 in Appendix demonstrate that coefficients for per capita GDP lose significance after including adult mortality into the regression. However, this highlights potential transmission channel: people might feel stressed not because they are just poor, but rather because they are more likely to die due to being poor. Furthermore, aggregate values of income might be not a very good predictor of suicide rate in the country, as we do not know who exactly commits suicide. As was mentioned above, financial hardship might be a stress factor that increases the number of suicides. At the same time, poverty might be a limiting factor, as the breadwinner might feel obliged to stay alive and support his family. Another issue in this context is matching incomes with suicide cases: mean income does not allow us to distinguish if suicides are committed over-proportionately by rich or poor. Finally, income inequality coefficients demonstrate no significant effect on suicide rates. Moreover, Table A3.3 demonstrates similar non-significant outcomes, once I control for non-linear relationship of inequality in income (e.g., do status concerns cause more stress in poor or rich countries?).

Surprisingly, unemployment coefficients presented so far demonstrate negative or not significant effects on suicide rates. There might be several explanations for this result. First of all, besides satisfaction work can provoke long-term stress and thus contribute to higher suicide rates. Secondly, when jobs are scarce, people might try to keep their jobs (also by staying alive) to support their families. To explore the relationship between suicide and unemployment, Table 3.4 demonstrates the regression results with different combinations of unemployment measures. Columns (1) and (4) include 5 year lags of unemployment rates to control for long-term effects of unemployment-induced stress for males and females, respectively. The general level of unemployment is proxied by 5-year average levels of unemployment rates in columns (2) and (5). The effect of the other gender's unemployment rates is evaluated in columns (3) and (6). The results presented in Table 3.4 rather indicate again that if there is an effect, unemployment might be suicide-decreasing. Results for the other gender unemployment rates support the hypothesis that unemployment can act as a reservation factor: when more males (females) have hard times finding the job, females (males) from the same households might feel obliged to stay alive to support the family. However, these results should be treated with caution, as the effect of unemployment on suicide rate can be better analysed at the micro level.

Finally, a very important finding is the significance of the data quality coefficient in some specifications considered so far. Negative effect indicates that countries with bad data quality typically report lower number of suicides. One possible explanation is that countries with low quality of vital registration statistics share some characteristics, besides the ones controlled for in the regression analysis so far, which prevent individuals from committing suicide. However, it can also be the sign of under-reporting of suicides, for example, due to religious and cultural norms.

Table 3.4: Suicide and unemployment. OLS estimates.

	male suicide rate			female suicide rate		
	(1)	(2)	(3)	(4)	(5)	(6)
UVR	-27.968***	-29.016***	-28.425***	-6.747***	-6.783***	-6.916***
	(5.632)	(5.615)	(5.699)	(1.461)	(1.450)	(1.442)
UVR^2	2.824***	2.927***	2.865***	0.691***	0.698***	0.713***
	(0.648)	(0.650)	(0.657)	(0.178)	(0.177)	(0.176)
genetic diversity	-24.532	-27.190	-20.381	-41.571*	-40.399*	-41.885*
	(41.406)	(41.936)	(42.154)	(23.775)	(23.402)	(23.629)
log(GDP per capita)	-0.763	-1.026	-0.779	-0.246	-0.249	-0.282
	(1.146)	(1.135)	(1.150)	(0.284)	(0.278)	(0.279)
inequality	-0.049	-0.031	-0.033	-0.005	-0.004	-0.002
	(0.119)	(0.120)	(0.116)	(0.032)	(0.032)	(0.032)
male unemployment	-0.238					-0.048
	(0.371)					(0.030)
female unemployment			-0.181**	0.050		
			(0.087)	(0.076)		
$\text{unemployment}_{(t-1)}$	-0.106			-0.173		
	(0.625)			(0.111)		
$\text{unemployment}_{(t-2)}$	0.295			0.136		
	(0.542)			(0.107)		
$\text{unemployment}_{(t-3)}$	-1.089			-0.068		
	(0.713)			(0.105)		
$\text{unemployment}_{(t-4)}$	1.295*			-0.024		
	(0.777)			(0.100)		
$\text{unemployment}_{(t-5)}$	-0.319			0.035		
	(0.319)			(0.055)		
5-year unemployment		-0.184			-0.044*	
		(0.133)			(0.023)	
adult mortality	0.739***	0.732***	0.734***	0.127***	0.126***	0.126***
	(0.130)	(0.132)	(0.132)	(0.040)	(0.040)	(0.040)
YLD	-76.540	-72.606	-77.736	13.144	14.506	14.860
	(91.999)	(91.870)	(91.060)	(19.295)	(19.194)	(19.125)
data quality	-3.169**	-3.317***	-3.286**	-0.427	-0.430	-0.425
	(1.264)	(1.263)	(1.266)	(0.327)	(0.323)	(0.325)
UVR_{min}	4.952	4.956	4.961	4.884	4.856	4.853
U-test p-value	0.008	0.007	0.008	0.024	0.020	0.017
year FE	Yes	Yes	Yes	Yes	Yes	Yes
regional FE	Yes	Yes	Yes	Yes	Yes	Yes
geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
religious dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	415	415	415	415	415	415
Number of countries	140	140	140	140	140	140
Adj. R-squared	0.581	0.575	0.578	0.462	0.466	0.464

5-year unemployment stands for 5-year average male unemployment in column (2) and 5-year average female unemployment in column (5). Lagged unemployment values stand for lagged male unemployment in column (1) and lagged female unemployment in column (4). Adult mortality and YLD stand for male adult mortality and YLD in columns (1)-(3) and female adult mortality and YLD in columns (4)-(6). Geographic controls include: temperature, precipitation and latitude. Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.3.1 Explaining the U-shape

The relationship between the UVR and suicides presented on Fig.(3.2) appears quite puzzling in spite of the presented biological mechanism. The downward slope of the curve to the left from 4.5 kJ/m^2 point perfectly fits into the described logic: higher levels of exposure to solar radiation increase serotonin circulation lowering the number of suicides. However, the upward slope of the curve after the minimum point cannot be explained by this mechanism. I found no evidence suggesting that excessive exposure to solar radiation might harm serotonin turnover or provoke suicide directly through other biological channels. However, we can also analyse the indirect channels: for example, the negative effect of excessive solar exposure on general health might serve as a stress factor or induce hopelessness (Andersen et al., 2016). To test this line of argument I evaluate the effect of sun-related diseases on suicide. WHO claims that excessive UVR can cause a range of diseases, while skin cancer and eye diseases are the most prevalent between them (Lucas et al., 2006). I suppose that the prevalence of sun-caused diseases might raise stress levels or cause the state of hopelessness and, hence, raise the suicide rates. WHO provides data for the number of disability-adjusted life years (DALY) for sun-related diseases. It allows me to capture both incidence and severity of diseases and test if they really contribute to higher suicide rates.

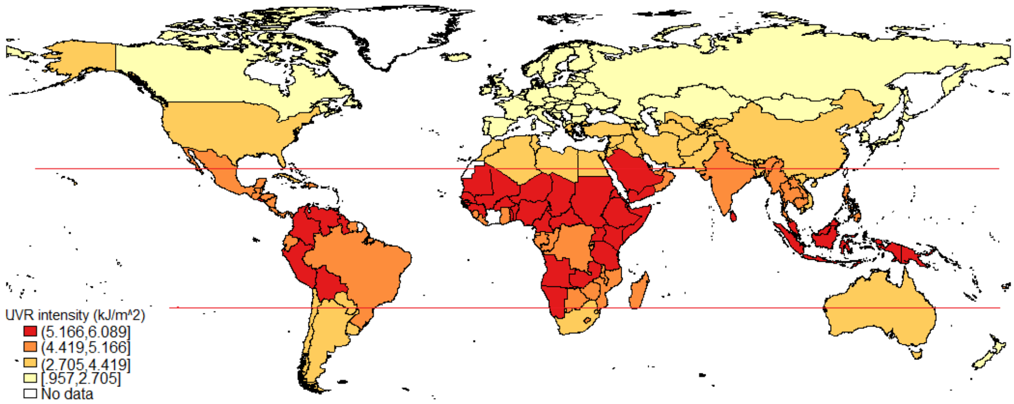
Another factor to consider in this context is the adaptability of individuals to the environment. To help people cope with solar radiation, over generations skin pigmentation has been changing to adjust for the amount of ambient UVR (Lucas et al., 2006). Lighter skin allows to absorb as much solar energy as possible in UVR-scarce regions, while in UVR-intensive areas an increased amount of melanin in the skin protect cells from excessive amounts of solar radiation. In other words, evolution has adjusted skin types to optimise absorption of solar radiation. However, once the discovery of the New World has triggered mass migration, many individuals were relocated to the places, where ambient solar radiation does not fit their skin colour. To check if this has an effect on suicide rates nowadays, I construct the difference between the actual UVR exposure and the ancestry adjusted one. To do so, I employ the World Migration Matrix by Putterman and Weil (2010) and calculate ancestry-weighted UVR exposure variable. This number tells me to how much solar radiation an average individual in the country would have been exposed to, absent the post-1500 migration. The difference between this calculated number and the actual UVR exposure might tell how much more or less UVR a respective individual gets in his country given his ancestry.

However, the results presented in Table 3.5 suggest that both sun-related diseases and the difference between actual and historical UVR levels do not have a significant effect on male suicide rates. Similar insignificant results are observed for females. This finding suggests

Table 3.5: Suicide and sun-related diseases. OLS estimates.

	male suicide rate			
	(1)	(2)	(3)	(4)
UVR	-28.881***	-29.028***	-28.854***	-28.811***
	(5.642)	(5.730)	(5.610)	(5.642)
UVR^2	2.906***	2.924***	2.904***	2.878***
	(0.651)	(0.663)	(0.648)	(0.660)
genetic diversity	-26.150	-26.816	-25.705	-34.733
	(41.610)	(41.681)	(41.674)	(42.404)
log(GDP per capita)	-0.952	-0.970	-0.953	-1.071
	(1.132)	(1.131)	(1.133)	(1.117)
inequality	-0.020	-0.027	-0.019	-0.038
	(0.117)	(0.118)	(0.117)	(0.121)
unemployment	-0.228*	-0.224*	-0.228*	-0.227*
	(0.122)	(0.123)	(0.121)	(0.121)
adult mortality	0.735***	0.736***	0.735***	0.738***
	(0.130)	(0.130)	(0.130)	(0.129)
data quality	-3.266**	-3.265**	-3.262**	-3.082**
	(1.262)	(1.259)	(1.259)	(1.329)
YLD	-70.206	-72.942	-69.259	-68.397
	(91.386)	(92.586)	(90.814)	(90.780)
UVR DALYs	-0.000			
	(0.006)			
skin cancer		0.006		
		(0.017)		
cataract			-0.002	
			(0.008)	
UV difference				-2.602
				(2.885)
UVR_{min}	4.969	4.963	4.967	5.005
U-test p-value	0.007	0.007	0.007	0.012
year FE	Yes	Yes	Yes	Yes
regional FE	Yes	Yes	Yes	Yes
geographic controls	Yes	Yes	Yes	Yes
religious dummies	Yes	Yes	Yes	Yes
N	415	415	415	415
Number of countries	140	140	140	140
Adj. R-squared	0.577	0.577	0.577	0.578

Unemployment, adult mortality, YLD and UVR DALYs indicate data for males. Geographic ontrols include: temperature, precipitation and latitude. Standard errors clustered at the country level are in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



Note: Horizontal red lines mark 25th parallels north and south.

Figure 3.3: Average levels of UVR exposure

that environmental factors that might increase suicide rates even through indirect biological channels do not seem to play significant role. Hence, there must be other factors driving the suicide rates upwards as we move above the 4.5 kJ/m^2 point. To better understand what those factors might be, we first need to identify what countries are exposed to UVR levels around 4.5 kJ/m^2 and above.

First of all, I check if results are robust to the splitting of the sample between OECD and non-OECD countries. At the same time it allows to check if the effects of UVR and adult mortality established so far hold in both groups of countries. For OECD countries I cannot include regional dummies, as majority of the countries lie in Europe and Central Asia (23 out of 28 included into the regression) and their inclusion distorts the regression results. To ensure comparability, Column (2) of Table 3.6 provides results of a regression without regional dummies for non-OECD countries. Coefficients in Columns (1) and (2) of Table 3.6 do not differ substantially from the results obtained so far for the whole sample of countries. Moreover, the minimum point also lies within the same range, so the results are robust to the exclusion of OECD countries. Speaking about OECD countries, only 1 country (Mexico) is exposed to UVR levels above the level of 4 kJ/m^2. This explains insignificant UVR coefficients in Column (3): there are simply no countries that can reverse the negative effect of UVR. As a result, linear specification in Column (4) appears more justified.

Table 3.6 demonstrates that U-shape relationship between UVR and suicide is not driven by the presence of OECD countries in the full sample. However, there might be other country characteristics that provide this pattern besides the level of economic development. Figure 3.3 demonstrates that the threshold line after which an increase in UVR is associated with higher suicide rates lies approximately 25 degrees of latitude north and south from the equator. Countries within this belt can share other common features that can bias the

Table 3.6: Determinants of suicide in non-OECD and OECD countries. OLS estimates.

	male suicide rate			
	(1)	(2)	(3)	(4)
sample of countries:	non-OECD		OECD	
UVR	-41.121***	-45.974***	-2.001	-9.861**
	(5.660)	(6.608)	(6.996)	(4.165)
UVR^2	4.479***	4.835***	-1.617	
	(0.654)	(0.753)	(1.133)	
genetic diversity	-5.437	-49.185	-368.445***	-304.235**
	(37.328)	(46.197)	(121.664)	(114.460)
log(GDP per capita)	0.532	0.675	0.496	1.126
	(1.221)	(1.264)	(3.078)	(3.583)
inequality	0.116	-0.022	-0.207	-0.158
	(0.102)	(0.108)	(0.358)	(0.372)
unemployment	-0.414***	-0.319***	-0.094	0.020
	(0.118)	(0.118)	(0.153)	(0.166)
adult mortality	0.581***	0.597***	1.341***	1.287***
	(0.123)	(0.122)	(0.282)	(0.280)
YLD	-99.028	-199.531**	60.499	216.610*
	(84.064)	(90.183)	(190.438)	(122.013)
data quality	-1.354	-2.651**	5.033	5.521
	(1.186)	(1.123)	(3.195)	(3.438)
Christian	-3.129	-4.611	-11.876***	-17.297***
	(2.274)	(3.338)	(3.959)	(2.992)
Muslim	-0.903	-2.644	-9.021*	-13.738***
	(2.277)	(3.509)	(5.008)	(4.871)
Buddhist	14.564**	13.287**		
	(6.793)	(5.240)		
Jewish	3.679	-1.680		
	(2.824)	(4.293)		
Hindu	0.838	-7.109		
	(3.313)	(6.155)		
unaffiliated	-6.583**	-8.867**	-11.273***	-16.157***
	(2.832)	(3.825)	(3.554)	(2.158)
UVR_{min}	4.590	4.754	-0.619	
U-test p-value	0.000	0.000	TR	
year FE	Yes	Yes	Yes	Yes
regional FE	No	Yes	No	No
geographic controls	Yes	Yes	Yes	Yes
N	328	328	87	87
Number of countries	112	112	28	28
Adj. R-squared	0.619	0.662	0.800	0.793

"TR" stands for trivial rejection of a U-shape hypothesis, as the extreme point lies outside the interval of UVR values in the sample. Unemployment, adult mortality and YLD indicate data for males. Geographic controls include: temperature, precipitation and latitude. Standard errors clustered at the country level are in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

regression results. Given that the 25th parallel south mostly passes through the oceans, I expect the reversal of the effect of UVR on suicides to be mostly driven by the countries on the 25th parallel north. Moreover, if we take a closer look at the map of Figure 3.1, we see that countries on the 25th parallel north generally have relatively low suicide rates, on the contrast to those lying further to the south. Once I have identified the geographical region that can bias the regression estimates, I can attempt to determine why are countries there special in terms of suicide rates. Solar radiation in the countries within this belt is definitely quite high: a person with skin type I (fair skinned Caucasian) needs only 200 J/m^2 to gain a perceptible sunburn (Lucas et al., 2006). What if people residing south of this region, hence, closer to the equator avoid sun and just stay indoors? UVR variable indicates only the characteristic of the environment that individuals face, but does not say how they actually react to it. As a proxy for the amount of time individuals actually stay under the sun, I use the share of population employed in agriculture. Contrast to factory workers or office clerks, people working in the fields have to stay outdoors, at least, part of their time, when the sun is already out. This measure is not perfect, as farmers in sunny regions may refrain from working in the hours of peak solar activity. However, exposure in relatively less intensive periods still allows the body to gain, at least, some amount of UVR. Once I interact UVR with the share of population employed in agriculture, I capture both quality (UVR intensity) and quantity (time people stay outdoors) aspects of exposure to solar radiation. However, as columns (1) and (4) of Table 3.7 demonstrate, controlling for actual exposure is not enough to explain the U-shaped relationship between UVR an suicide, so there must be other factors determining this pattern.

Besides similar geographic conditions, we can also notice that the majority of the countries are predominantly Muslim and belong to the Arab World, so I can assume some degree of cultural proximity. Out of Middle-Eastern and North-African (MENA) countries the World Values Survey (2014) provides data for Algeria, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Palestine, Tunisia, Qatar and Yemen. On average 89% of surveyed individuals in these countries have attributed religion as being very important in their lives, while in the rest of the surveyed world this share is only 43.3% (Inglehart et al., 2014). Similar pattern is observed in the results of Gallup Poll in 2009: 89% of surveyed citizens of the Middle-Eastern countries has replied yes to the question “Is religion important in your daily life?”, while in the rest of the world only 71% has replied the same way[7] (Crabtree, 2010). In other words, the data suggest that not only religion, but especially the degree of religiosity, can be a serious reservation factor that prevents people from committing suicide in the MENA region. To control for the degree of religiosity in the covered sample of countries, I use the share of population that considers religion important in their daily life from the Gallup Poll 2009, as it has a much wider coverage, compared to the World Values Survey. I assume that religiosity across countries has not changed

[7] Difference between the MENA countries and the rest of the world is statistically significant at 1%-level.

dramatically between 2000 and 2015 and values of 2009 can be used for its measurement. However, columns (2) and (5) of Table 3.7 demonstrate no significant degree of religiosity on suicide rates. The reason might be different attitudes towards suicide between religions. Even though, all religions included in the analysis in one way or another discourage suicide, some might be less strict in this regard. To control for different effect of religiosity between confessions, I interact religious dummies used so far with the share of people that consider religion important in their daily life. Hence, I obtain not only nominal affiliation to a particular confession, but also some evaluation of how important religious norms are for people. This approach yields the insignificant U-shape relationship between UVR and suicides, as columns (3) and (6) of Table 3.7 show. As a result, I can argue that there is no increase in the number of suicides due to the excessive solar radiation south from the 25 degrees of latitude line, but this effect is rather driven by the presence of the MENA countries in the sample: moderately sunny, predominantly Muslim and relatively more religious states that report low suicide rates. Finally, the coefficient of a squared UVR term is still significant even though the hypothesis of the U-shape relationship can be rejected. This can be perfectly explained by non-linear (but still monotonously decreasing) effect of UVR on suicide rates. As exposure to solar radiation increases, the marginal effect of UVR decreases. In other words, we can expect some degree of saturation with UVR: in areas closer to the equator the sun is so intensive, that individuals residing in these areas can hardly avoid solar radiation in quantities sufficient for proper functioning of the body.

Table 3.7: Religiosity and suicide. OLS estimates.

	male suicide rate			female suicide rate		
	(1)	(2)	(3)	(4)	(5)	(6)
UVR	-28.156***	-24.248***	-22.052***	-6.124***	-6.046***	-5.826***
	(5.532)	(5.443)	(5.617)	(1.445)	(1.484)	(1.614)
agriculture	-0.084	-0.064	-0.135	-0.045	-0.050	-0.063*
	(0.174)	(0.189)	(0.180)	(0.032)	(0.033)	(0.036)
UVR × agriculture	-0.006	-0.010	0.001	0.008	0.009	0.012
	(0.033)	(0.037)	(0.037)	(0.007)	(0.008)	(0.008)
UVR^2	2.878***	2.517***	2.050***	0.592***	0.624***	0.570***
	(0.677)	(0.639)	(0.645)	(0.187)	(0.187)	(0.202)
genetic diversity	-28.127	-45.416	-48.407	-40.328*	-43.533*	-44.155*
	(40.705)	(43.166)	(42.860)	(23.881)	(25.422)	(25.400)
log(GDP p.c.)	-2.394	-2.657	-2.504	-0.421	-0.672*	-0.657*
	(1.609)	(1.927)	(1.625)	(0.334)	(0.374)	(0.384)
inequality	-0.031	0.012	-0.024	-0.001	0.009	0.010
	(0.120)	(0.138)	(0.132)	(0.032)	(0.030)	(0.030)
unemployment	-0.272**	-0.249*	-0.301**	-0.042*	-0.043*	-0.044**
	(0.126)	(0.128)	(0.120)	(0.021)	(0.022)	(0.020)
adult mortality	0.731***	0.708***	0.719***	0.127***	0.089**	0.088**
	(0.134)	(0.159)	(0.144)	(0.040)	(0.038)	(0.037)
YLD	-48.277	-43.583	33.424	11.102	13.522	21.484
	(91.640)	(92.564)	(81.432)	(19.945)	(21.823)	(21.769)
data quality	-2.848**	-3.204***	-1.485	-0.389	-0.562*	-0.339
	(1.237)	(1.118)	(0.907)	(0.327)	(0.309)	(0.305)

Continued on the next page

Table 3.7: – *continued from the previous page*

	male suicide rate			female suicide rate		
	(1)	(2)	(3)	(4)	(5)	(6)
religiosity		-0.080	-0.121		-0.021	-0.000
		(0.059)	(0.109)		(0.016)	(0.050)
Christian	-6.829*	-7.699	-11.475	-3.260**	-4.041**	-2.895
	(3.476)	(4.795)	(10.637)	(1.400)	(1.739)	(5.134)
C. × religiosity			0.114			-0.013
			(0.120)			(0.055)
Muslim	-4.080	-4.977	-2.583	-1.412	-2.282	0.620
	(3.583)	(4.841)	(14.428)	(1.394)	(1.753)	(5.263)
M. × religiosity			0.050			-0.034
			(0.162)			(0.059)
Buddhist	12.172**	15.276**	-2123.825***	2.804	3.817	-206.867
	(5.656)	(6.842)	(286.434)	(1.889)	(2.319)	(139.928)
B. × religiosity			21.986***			2.154
			(2.914)			(1.435)
Jewish	0.587	-3.986	3.414	-2.285	-4.158**	-3.196
	(4.698)	(5.460)	(4.771)	(1.673)	(1.848)	(2.323)
Hindu	-7.510	-8.180	254.198***	-0.946	0.935	194.432***
	(6.154)	(8.388)	(75.337)	(2.456)	(3.068)	(19.367)
H. × religiosity			-2.670***			-2.092***
			(0.833)			(0.212)
unaffiliated	-7.850*	-10.007*	-13.020	-2.590	-5.127**	-10.993
	(4.239)	(5.660)	(18.897)	(2.271)	(2.185)	(7.253)
U. × religiosity			0.111			0.351
			(0.864)			(0.331)
UVR_{min}	4.891	4.817	5.379	5.172	4.848	5.107
U-test p-value	0.016	0.020	0.148	0.126	0.064	0.140
year FE	Yes	Yes	Yes	Yes	Yes	Yes
regional FE	Yes	Yes	Yes	Yes	Yes	Yes
geo. controls	Yes	Yes	Yes	Yes	Yes	Yes
N	415	383	383	415	383	383
N of countries	140	126	126	140	126	126
Adj. R-squared	0.583	0.579	0.632	0.467	0.485	0.497

Employment in agriculture, unemployment, adult mortality and YLD stand for male employment in agriculture, unemployment, adult mortality and YLD in columns (1)-(3) and female employment in agriculture, unemployment, adult mortality and YLD in columns (4)-(6). Geographic controls include: temperature, precipitation and latitude. Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.3.2 Robustness check - Fixed Effects estimation

The results demonstrated above support the hypothesis that higher levels of UV radiation decrease the number of suicides, while potential stress factors, as higher adult mortality increase it. Shares of population attributed to particular confessions can act as reservation factors preventing suicides, especially in Protestant and Muslim countries. However, once I point out that religion might play an important role, I need to consider other aspects of culture. Moreover, variations between countries in cultural and religious norms even within one confession can be crucial in determining the "culture of suicide". For this reason I need to control for unobserved heterogeneity between countries. To do so, I employ the country fixed effects (FE) estimator. Raw version of UVR variable provided by WHO does not vary over time, hence, is not suitable for this purpose. However, the approach used in the previous section, allows to make exposure to solar radiation a dynamic variable, due to the fact that actual UVR exposure varies over time, as people shift their activities in- or outdoors over time. Having a time-varying measure of exposure to solar radiation, I can perform a fixed effects estimation, results of which are presented in Table 3.8.

Table 3.8: Determinants of suicide. Country FE estimates.

	male suicide rate		female suicide rate	
	(1)	(2)	(3)	(4)
UVR × agriculture	-0.128***	-0.146***	-0.012	-0.014**
	(0.047)	(0.048)	(0.007)	(0.007)
agriculture	0.593**	0.709***	0.056	0.070**
	(0.230)	(0.238)	(0.037)	(0.035)
log(GDP per capita)	-2.469*		-0.669*	
	(1.353)		(0.385)	
unemployment	0.027	0.063	0.001	0.013
	(0.058)	(0.057)	(0.011)	(0.011)
adult mortality	0.300***	0.309***	0.032***	0.036***
	(0.074)	(0.078)	(0.012)	(0.013)
YLD	-157.406***	-153.681***	-35.848**	-34.433**
	(52.081)	(53.906)	(15.548)	(15.471)
country FE	Yes	Yes	Yes	Yes
year FE	Yes	Yes	Yes	Yes
N	662	680	662	680
Number of countries	167	170	167	170
Adj. R-squared	0.309	0.299	0.189	0.180

Unemployment, adult mortality and YLD stand for male unemployment, adult mortality and YLD in columns (1) and (2) and female unemployment, adult mortality and YLD in columns (3) and (4). Agriculture denotes share of male working population employed in agriculture in columns (1) and (2) and female working population in columns (3) and (4). Standard errors clustered at the country level are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8 demonstrates that controlling for unobserved factors that constitute potential "culture of suicide", we still observe significant negative effect of solar exposure on suicide rates. This effect is smaller for females, but this can be explained by the fact that women generally commit suicide way less and their suicide rates are harder to predict compared to males. Moreover, the interaction of UVR and agricultural employment variables might be not equally informative for males and females, if it is meant to control for the "efficient" exposure to solar radiation. As was mentioned above, agricultural workers are assumed to spend more time outdoors, hence, to be more exposed to UVR. However, the actual "absorption" of solar radiation can also be affected by some cultural factors. Table 3.9 demonstrates that among the countries, included in the regression sample, share of women employed in agriculture is significantly higher in Muslim countries compared to the non-Muslim. This reveals an interesting fact: female employment in agriculture is higher in the countries, where women traditionally wear less open clothes, thus limiting the potential degree of exposure to UVR. In other words, employment in agriculture might be a worse predictor of actual exposure to solar radiation for females than for males. Interestingly, for males there is no significant difference between predominantly Muslim states and the rest of the World. Nevertheless, this should have no effect on actual UVR exposure for men, as they are typically not constrained in their way of clothing.

Table 3.9: Share of population employed in agriculture (%)

	Muslim countries			Rest of the World			
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Δ
Females	176	39.530	29.174	504	28.332	30.059	11.197^{***}
Males	176	33.588	24.688	504	30.972	24.932	2.616

Δ indicates the difference of means between Muslim and non-Muslim countries. P-values are obtained after running a t-test of means. *** $p < 0.01$

It is important to note that inclusion of per-capita GDP together with shares of employment in agriculture into the regression presented in columns (1) and (3) of Table 3.8 raises multicollinearity problem, but coefficients of interest remain robust nevertheless, as columns (2) and (4) demonstrate. Even though, the negative coefficient of income on suicide rates fits to the theoretical predictions of the paper, this finding should be considered cautiously, as inequality issue is completely ignored in this case. Gini coefficients are excluded from the regressions to keep the sample close-to-balanced. Moreover, as was noted above, income inequality is quite persistent, especially within 15 years, and, hence, should be controlled for by inclusion of country fixed effects. Finally, we still observe significant effect of adult mortality on suicide rates, but controlling for time-invariant country characteristics makes the years lived with disability coefficient significant. Unlike the theoretical prediction, where YLD is included to control for costs of maintaining oneself alive, the empirical result has a negative coefficient on suicide rates. There might be several explanations to this finding. On one hand, severely ill individuals can refrain from suicide as they expect to die due to natural causes in the nearest future. On the other hand, higher number of ill people in the country might indicate a larger demand for care from their relatives, what can serve as a reservation factor. Regardless of the exact channel, the magnitude of the effect is quantitatively small: one standard deviation in YLD is associated with less than 0.1 standard deviation in suicide both for males and females.

3.3.3 What does it all mean?

Previous sections of the paper have demonstrated that differences in suicide rates across countries can be driven by two fundamentally different groups of factors. The main determinants of suicide identified in this paper are UVR and adult mortality - two factors that differ in nature. While UVR is purely environmental characteristic, adult mortality is to a large extent economically or socially determined. For the purpose of this paper I categorise suicides into environmentally-driven (lack of UVR) and socially-driven (high stress levels due to high adult mortality) suicides. Given this difference in the nature of factors, we can analyse what countries are more vulnerable to a particular group of drivers. Table 3.10 presents summary statistics of UVR variables, adult mortality and suicide rates in OECD member-states (developed countries) and non-OECD states (developing countries).

Table 3.10: Summary statistics of selected variables in OECD and non-OECD states

Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: OECD countries					
male suicide rate	87	17.526	7.477	5.1	39.3
female suicide rate	87	5.224	2.614	1.1	18.4
UVR	87	2.166	0.789	1.439	4.974
UVR^2	87	5.308	4.821	2.071	24.741
adult mortality (male)	87	11.641	4.278	6.2	25.7
adult mortality (female)	87	5.915	1.678	3.3	10.8
Panel B: non-OECD countries					
male suicide rate	328	18.799	12.690	0.4	70.3
female suicide rate	328	5.645	3.318	0.1	19.2
UVR	328	4.240	1.251	1.671	6.089
UVR^2	328	19.534	9.610	2.792	37.076
adult mortality (male)	328	26.165	11.131	7	65.2
adult mortality (female)	328	18.110	11.748	3.8	63.1

Strikingly, unlike crude suicide rates presented in Table 3.1 age-standardised suicide rates both for males and females are quite similar across OECD and non-OECD countries. However, we see that UVR exposure and adult mortality are significantly higher in developing countries[8]. Majority of developed countries lie in less sunny areas, where lack of solar radiation can provoke depression and, consequently, pull suicide rates upwards. On the contrast, better living conditions, reflected by lower mortality rates, can reduce stress levels and, hence, suicide rates. As a result, I can claim that in developed countries suicides are rather driven by environmental characteristics, namely UVR levels. However, in developing countries the two groups of factors act completely different: more solar radiation should prevent suicides (and this claim is supported by a set of biological studies presented above), but higher adult mortality contributes to an increase in number of self-

[8]T-test of means indicates that the following differences between non-OECD and OECD countries: 2.074 for UVR, 14.524 for male adult mortality, 12.195 for female adult mortality, are statistically significant at 1%-level.

killings. This claim can be supported by comparing the contribution of the two variables to the explanation of the suicide rates. I do this by comparing the partial R-squared of UVR-related variables and adult mortality rates in Table 3.11. To calculate the share of variation of suicide rates, which is explained only by a single variable of interest (UVR or adult mortality), I first run the regression (similar to the ones presented in column 1 and 4 of Table 3.3) including the whole set of controls, except for the variable of interest, separately for OECD and non-OECD countries. Afterwards, I extract the residuals, obtaining only that part of variation in suicide rates that is not correlated with other controls. Then I repeat the exercise having the variable of interest as the dependent variable - to obtain the part of variation in the variable of interest not correlated with other regressors. Finally, I regress the predicted residuals of the suicide rates on the predicted residuals of a UVR or adult mortality. The R-squared of this regression tells us how much of the variation in suicide rates is explained by each of the variables of interest, once the effect of all other variables has already been controlled for.

Table 3.11: Partial R^2 of selected variables

	males		females	
	OECD	non-OECD	OECD	non-OECD
adult mortality rate	0.195	0.228	0.286	0.087
UVR	0.413	0.031	0.340	0.011

Note: Regressions of UVR in non-OECD sample do not include the squared term. Inclusion of the squared term of predicted residuals yields the partial R-squared of 0.031 for males and 0.012 for females.

As can be seen from Table 3.11, adult mortality rates explain comparable parts of variation in suicide rates both in OECD and non-OECD countries. On the contrast, while UVR can explain more than one third of variation in developed countries, it works as a very bad predictor of suicide rates in developing countries. This implies a striking difference in the nature of the major driving forces of suicides across OECD and non-OECD member-states. The empirical results suggest that we should expect more environmentally-driven suicides in developed countries and more of socially-driven in developing ones. Moreover, estimations suggest that improving mortality rate in developing countries up to the average level of OECD member-states can decrease the number of suicides roughly by half for males and by one third for females.

3.4 Conclusion

This paper has analysed cross-country variations in suicide rates to identify fundamental factors that can influence the number of people killing themselves given a set of country characteristics. As was pointed above, suicide still remains personal and often sporadic decision. Hence, not all factors can be identified at the macro level. This paper has demonstrated that the effects of genetics, income or unemployment are hard to detect at the country-level, as we do not know how the distribution of these variables relates to the distribution of suicides. Nevertheless, this paper has identified a set of fundamental factors that can form a physical environment or some cultural features that, in turn, affect suicide rates. Ultra-violet radiation and adult mortality are the examples of the factors that can be studied at the cross-country level. Solar radiation affects many biological processes in our body, including functioning of the brain. Lack of UVR can distort circulation of serotonin and cause depression, which in extreme cases leads to suicide. If a country lies in areas with less solar radiation, we can expect its residents, ceteris paribus, to be more vulnerable to suicide. Unlike income or employment, exposure to solar radiation has unambiguous effect on suicides and can be assumed to be distributed way more evenly across citizens of the country allowing some degree of generalisation. Another factor that appears appropriate for macro-level analysis is adult mortality. As was demonstrated in the theoretical section of the paper this variable affects utility of individuals through several channels, but all of them are better captured at the macro level. If one individual dies at young age, it does not necessarily set a prior for his compatriots that they die young, too. However, when many people manage to live only until a certain age, it might determine the perception of maximum attainable age and probability of attaining it for the citizens of a country. Those individuals who do not expect to live much longer can develop hopelessness (another component of depression syndrome) that increases their likelihood of committing suicide.

Another important finding of this paper is that even though similar factors affect suicide rates both in developed and developing countries, the quantitative effects of different factors vary substantially. Developed countries mostly lie in relatively darker areas (suicide increasing), but have low adult mortality rates (suicide decreasing). On the contrary, most of the developing countries (e.g., in Sub-Saharan Africa) are exposed to large amounts of solar radiation, but also have high adult mortality. As a result, we see no significant difference in suicide rates (after controlling for age structure of the population) between developed and developing countries, but similar suicide rates result from entirely different combinations of causes. Results of the regression estimations presented in this paper should not be treated as ultimate drivers of suicide or used for predictions, but rather used to motivate future studies on the topic. The findings of the paper point out at the differences in nature of major factors affecting suicide decisions across countries at different stages of economic development. Hence, results of existing studies of suicide conducted at the micro level in developed countries are not necessarily applicable in the developing World, implying that more attention should be devoted to the problems of suicide in developing countries.

Appendix

List of OECD countries

OECD countries are defined according to the membership status in the beginning ot the sampled period (2000) and therefore include the following 30 countries:

Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom and United States.

List of non-OECD countries

Non-OECD countries include countries that were not members of the organisation until 2000:

Afghanistan, Albania, Algeria, Angola, Argentina, Armenia, Azerbaijan, the Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Brunei Darussalam, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo (Dem. Rep.), Congo (Rep.), Costa Rica, Croatia, Cuba, Cyprus, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Estonia, Eritrea, Ethiopia, Fiji, Gabon, the Gambia, Georgia, Ghana, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, India, Indonesia, Iran, Iraq, Israel, Jamaica, Jordan, Kazakhstan, Kenya, Kuwait, Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Lesotho, Liberia, Libya, Lithuania, Macedonia, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Moldova, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Nepal, Nicaragua, Niger, Nigeria, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Qatar, Romania, Russian Federation, Rwanda, Samoa, Saudi Arabia, Senegal, Sierra Leone, Singapore, Slovenia, Solomon Islands, Somalia, South Africa, Sri Lanka, Sudan, Suriname, Swaziland, Syria, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkmenistan, Uganda, Ukraine, United Arab Emirates, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, Yemen, Zambia and Zimbabwe.

Summary statistics and sources of used variables

Variable	Obs	Mean	Std. Dev.	Min	Max	Source
age-standardised male suicide rate	415	18.532	11.793	0.4	70.3	World Bank
age-standardised female suicide rate	415	5.556	3.185	0.1	19.2	World Bank
UVR	415	3.805	1.442	1.439	6.089	WHO
UVR^2	415	16.552	10.554	2.071	37.076	WHO, AC
male adult mortality net of suicides	415	23.120	11.692	6.2	65.2	World Bank, AC
female adult mortality net of suicides	415	15.553	11.589	3.3	63.1	World Bank, AC
genetic diversity	415	0.725	0.029	0.628	0.774	Ashraf and Galor (2013)
log(GDP per capita, PPP)	415	8.893	1.220	6.098	11.554	World Bank, AC
inequality (Gini coefficient, 5-year average)	415	39.602	9.014	21.1	65.8	World Bank, AC
male unemployment rate	415	7.785	5.594	0.405	34.903	World Bank
female unemployment rate	415	9.640	7.329	0.215	59.165	World Bank
years lost due to disability (male)	415	0.090	0.009	0.064	0.119	WHO
years lost due to disability (female)	415	0.101	0.012	0.077	0.131	WHO
data quality	415	2.289	1.298	1	4	WHO
Christian dummy	415	0.660	0.474	0	1	Pew Research Center(2012)
Muslim dummy	415	0.212	0.409	0	1	Pew Research Center(2012)
Buddhist dummy	415	0.048	0.214	0	1	Pew Research Center(2012)
Jewish dummy	415	0.010	0.098	0	1	Pew Research Center(2012)
Hindu dummy	415	0.014	0.120	0	1	Pew Research Center(2012)
Unaffiliated dummy	415	0.024	0.154	0	1	Pew Research Center(2012)
latitude	415	21.603	25.588	-34	64	Ashraf and Galor (2013)
temperature	415	16.975	8.671	-7.929	28.639	Ashraf and Galor (2013)
precipitation	415	90.069	57.806	2.911	259.952	Ashraf and Galor (2013)
share of labour in agriculture (male)	415	31.908	24.007	0.8	88.3	World Bank
share of labour in agriculture (female)	415	31.374	29.584	0.1	96.8	World Bank
religious importance	383	72.133	24.688	16	99	(Crabtree, 2010)

AC stands for author's calculations. All incurred mistakes are therefore mine.

Table A3.1: Determinants of male suicide. OLS estimates.

	suicide rate								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
UVR	-18.036***	-20.243***	-20.846***	-25.587***	-25.358***	-23.247***	-16.639***	-16.197***	-20.459***
	(3.829)	(4.175)	(4.107)	(4.305)	(4.240)	(4.193)	(3.273)	(3.507)	(3.607)
UVR^2	2.244***	2.554***	2.454***	2.805***	2.708***	2.478***	1.655***	1.633***	2.194***
	(0.483)	(0.531)	(0.525)	(0.521)	(0.507)	(0.485)	(0.403)	(0.403)	(0.459)
genetic diversity		11.291	-3.740	-12.518	4.285	-16.798	-69.400**	-64.826**	-34.598
		(26.228)	(27.185)	(29.432)	(28.607)	(27.753)	(29.352)	(30.561)	(33.147)
log(GDP per capita)			-2.493***	-3.623***	-3.600***	-3.730***	1.350	0.971	-0.139
			(0.657)	(1.062)	(1.048)	(1.021)	(0.913)	(0.967)	(0.943)
inequality				0.329***	0.402***	0.385***	-0.028	-0.036	0.094
				(0.109)	(0.115)	(0.112)	(0.120)	(0.125)	(0.096)
unemployment					-0.262*	-0.291**	-0.354***	-0.357***	-0.254**
					(0.143)	(0.137)	(0.097)	(0.095)	(0.113)
YLD						189.454*	-59.118	-44.966	-31.134
						(98.608)	(89.960)	(106.631)	(83.252)
adult mortality							0.801***	0.801***	0.720***
							(0.125)	(0.126)	(0.127)
data quality								-0.699	-1.832
								(1.248)	(1.150)
Christian									-4.674
									(3.861)
Muslim									-3.750
									(3.891)
Buddhist									11.477
									(7.627)
Jewish									1.457
									(4.512)
Hindu									-2.653
									(3.885)
unaffiliated									-4.771
									(4.493)
N	696	636	619	420	420	420	420	420	420
Number of countries	174	159	156	142	142	142	142	142	142
Adj. R-squared	0.119	0.135	0.181	0.225	0.235	0.248	0.467	0.468	0.534

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.2: Determinants of female suicide. OLS estimates.

	female suicide rate								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
UVR	-2.896***	-3.340***	-3.461***	-4.037***	-3.526***	-2.876**	-2.131*	-2.573**	-4.584***
	(0.909)	(1.073)	(1.068)	(1.207)	(1.230)	(1.215)	(1.102)	(1.080)	(1.043)
UVR^2	0.401***	0.473***	0.425***	0.500***	0.427***	0.367**	0.223	0.256*	0.514***
	(0.125)	(0.146)	(0.147)	(0.158)	(0.162)	(0.157)	(0.148)	(0.142)	(0.141)
genetic diversity		-5.527	-11.190	-13.649	-7.139	-14.610	-34.840**	-36.876**	-29.460*
		(12.902)	(13.178)	(14.961)	(15.499)	(15.530)	(15.880)	(15.846)	(17.518)
log(GDP per capita)			-0.878***	-0.861***	-0.789***	-0.904***	0.049	0.258	0.015
			(0.216)	(0.285)	(0.284)	(0.271)	(0.251)	(0.248)	(0.249)
inequality				0.005	0.023	0.008	-0.063**	-0.046	0.019
				(0.035)	(0.035)	(0.034)	(0.030)	(0.033)	(0.024)
unemployment					-0.066**	-0.064**	-0.068***	-0.065***	-0.045**
					(0.028)	(0.027)	(0.021)	(0.020)	(0.021)
YLD						57.142**	12.226	11.060	28.029
						(24.594)	(23.500)	(21.884)	(18.808)
adult mortality							0.182***	0.159***	0.153***
							(0.042)	(0.042)	(0.040)
data quality								0.607**	0.131
								(0.304)	(0.305)
Christian									-2.760
									(1.779)
Muslim									-1.262
									(1.799)
Buddhist									2.940
									(2.255)
Jewish									-1.253
									(2.008)
Hindu									1.313
									(2.248)
unaffiliated									-1.831
									(2.347)
N	696	636	619	420	420	420	420	420	420
Number of countries	174	159	156	142	142	142	142	142	142
Adj. R-squared	0.029	0.039	0.103	0.094	0.111	0.135	0.252	0.270	0.402

Standard errors clustered at the country level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3.3: Suicide and inequality. OLS estimates.

	male suicide rate		female suicide rate	
	(1)	(2)	(3)	(4)
UVR	-27.780***	-28.024***	-5.336***	-4.937***
	(7.011)	(7.450)	(1.925)	(1.841)
UVR^2	2.750***	2.778***	0.531**	0.485**
	(0.857)	(0.890)	(0.257)	(0.243)
genetic diversity	-55.037	-54.780	-51.258**	-51.897**
	(47.812)	(47.894)	(24.828)	(24.783)
log(GDP per capita)	-1.104	-1.775	-0.230	0.797
	(1.201)	(4.328)	(0.287)	(0.895)
inequality	-0.214	-0.359	-0.044	0.191
	(0.164)	(0.832)	(0.044)	(0.194)
log(GDP) × inequality		0.018		-0.028
		(0.103)		(0.024)
unemployment	-0.034	-0.042	-0.040	-0.032
	(0.169)	(0.171)	(0.032)	(0.033)
adult mortality	0.739***	0.737***	0.114***	0.110***
	(0.129)	(0.130)	(0.039)	(0.040)
YLD	-133.569	-130.314	41.664**	40.057*
	(123.090)	(120.387)	(20.088)	(20.416)
data quality	-3.714**	-3.679**	-0.205	-0.261
	(1.468)	(1.479)	(0.330)	(0.325)
Christian	-5.242*	-5.205*	-1.566*	-1.613*
	(3.050)	(3.075)	(0.809)	(0.819)
Muslim	-2.834	-2.866	-0.319	-0.268
	(3.119)	(3.141)	(0.875)	(0.867)
Buddhist	11.660**	11.650**	4.690***	4.735***
	(4.863)	(4.883)	(1.172)	(1.174)
Jewish	-2.341	-2.379	-1.428	-1.268
	(4.737)	(4.760)	(1.192)	(1.175)
Hindu	-6.436	-6.380	0.540	0.425
	(5.942)	(6.013)	(2.458)	(2.510)
unaffiliated	-4.880	-4.823	-0.076	-0.170
	(4.147)	(4.107)	(1.926)	(1.940)
UVR_{min}	5.051	5.044	5.024	5.087
U-test p-value	0.067	0.064	0.189	0.212
year FE	Yes	Yes	Yes	Yes
regional FE	Yes	Yes	Yes	Yes
geographic controls	Yes	Yes	Yes	Yes
N	325	325	325	325
Number of countries	129	129	129	129
Adj. R-squared	0.537	0.536	0.457	0.459

Unemployment, adult mortality and YLD stand for male unemployment, adult mortality and YLD in Columns (1) and (2) and female unemployment, adult mortality and YLD in Columns (3) and (4). Geographic controls include: temperature, precipitation and latitude. Standard errors clustered at the country level are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Bibliography

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*, 1369–1401.

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

Ager, P., C. W. Hansen, and L. Lonstrup (2015). Shaking up the equilibrium: Natural disasters, immigration and economic geography. *Discussion Papers from University of Copenhagen. Department of Economics* (15-17).

Ahlburg, D. and M. Schapiro (1984). Socioeconomic ramifications of changing cohort size: an analysis of US postwar suicide rates by age and sex. *Demography 21*, 97–108.

Al-Karawia, D. and L. Jubair (2016). Bright light therapy for nonseasonal depression: Meta-analysis of clinical trials. *Journal of Affective Disorders 198*, 64–71.

Albouy, D. (2016). What are cities worth? land rents, local productivity, and the total value of amenities. *Review of Economics and Statistics 98*(3), 477–487.

Andersen, T. B., C.-J. Dalgaard, and P. Selaya (2016). Climate and the emergence of global income difference. *Review of Economic Studies 83*, 1334–63.

Andres, A. R. (2005). Income inequality, unemployment, and suicide: a panel data analysis of 15 european countries. *Applied Economics 37*(4), 439–451.

Andres, A. R. and F. Halicioglu (2010). Determinants of suicides in Denmark: Evidence from time series data. *Health Policy 98*(2), 263–269.

Andres, A. R., F. Halicioglu, and E. Yamamura (2011). Socio-economic determinants of suicide in Japan. *The Journal of Socio-Economics 40*(6), 723–731.

Ashraf, N., E. L. Glaeser, and G. A. M. Ponzetto (2016). Infrastructure, incentives, and institutions. *American Economic Review 106*(5), 77–82.

Ashraf, Q. and O. Galor (2013). The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development. *American Economic Review 103*(1), 1–46.

Asnani, M. R., A. M. McCaw-Binns, and M. E. Reid (2011). Excess risk of maternal death from sickle cell disease in Jamaica: 1998-2007. *PLoS ONE 6*.

Au, C.-C. and J. V. Henderson (2006). Are Chinese cities too small? *Review of Economic Studies* (73), 549–576.

Barrios, S., L. Bertinelli, and E. Strobl (2006). Climatic change and rural - urban migration: the case of sub-Saharan Africa. *Journal of Urban Economics 60*(3), 357–371.

Batty, M. and K. S. Kim (1992). Form follows function: Reformulating urban population density functions. *Urban Studies 29*(7), 1043–1069.

Baum, A. (1990). Stress, intrusive imagery, and chronic distress. *Health Psychology 9*(6), 653–675.

Beach, B., J. Ferrie, M. Saavedra, and W. Troesken (2016). Typhoid fever, water quality, and human capital formation. *The Journal of Economic History 76*(1), pages 41–75.

Beck, A. T., R. A. Steer, M. Kovacs, and B. Garrison (1985). Hopelessness and eventual suicide: a 10-year prospective study of patients hospitalized with suicidal ideation. *The American Journal of Psychiatry 142*(5), 559–563.

Beliveau, V., M. Ganz, L. Feng, B. Ozenne, L. Højgaard, P. M. Fisher, C. Svarer, D. N. Greve, and G. M. Knudsen (2017). A high-resolution in vivo atlas of the human brain's serotonin system. *Journal of Neuroscience 37*(1), 120–128.

Bertola, L., L. P. de la Escosura, and J. G. Williamson (2010). Latin American inequality in the long run. *Revista de Historia Economica / Journal of Iberian and Latin American Economic History 28*.

Beskow, J. (1990). Depression and suicide. *Pharmacopsychiatry 23*(3-8).

Bleakley, H. and J. Lin (2012). Portage and path dependence. *The Quarterly Journal of Economics 127*(2), 587–644.

Brueckner, J. K. (1987). *Handbook of Regional and Urban Economics*, Volume 2, Chapter 20: The structure of urban equilibria: a unified treatment of the Muth-Mills model, pp. 821–845. North-Holland.

Brueckner, J. K. and D. A. Fansler (1983). The economics of urban sprawl: Theory and evidence on the spatial sizes of cities. *The Review of Economics and Statistics 65*(3), 479–482.

Brueckner, J. K. and R. W. Helsley (2011). Sprawl and blight. *Journal of Urban Economics* (65), 205–213.

Brueckner, J. K. and S. V. Lall (2015). *Handbook of Regional and Urban Economics*, Chapter Cities in Developing Countries: Fueled by Rural-Urban MIgration, Lacking in Tenure Security, and Short of Affordable Housing, pp. 1399–1454. Elsevier.

Brueckner, J. K. and H. Selod (2009). A theory of urban squatting and land-tenure formalization in developing countries. *American Economic Journal: Economic Policy 1*(1), 28–51.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review 43*(1), 91–107.

Brueckner, J. K. and Y. Zenou (1999). Harris-Todaro models with a land market. *Regional Science and Urban Economics* (29), 317–339.

Bruelhart, M. and F. Sbergami (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics 65*, 48–63.

Burrows, G. D. (Ed.) (1977). *Handbook of Studies on Depression.* Elsevier/North-Holland Biomedical Prress.

Cain, L. P. and E. J. Rotella (2001). Death and spending: Urban mortality and municipal expenditure on sanitation. *Annales De Demographie Historique 1*, 139–54.

Calbo, J. and J.-A. Gonzalez (2005). Empirical studies of cloud effects on UV radiation: A review. *Reviews of Geophysics 43*(2).

Case, A. and A. Deaton (2015a). Rising morbidity and mortality in midlife among white non-Hispanics Americans in the 21st century. *PNAS 112*(49), 15078–83.

Case, A. and A. Deaton (2015b). Suicide, age, and wellbeing: an empirical investigation. *NBER Working Paper* (21279).

Caselli, F. (2005). *In Handbook of Economic Growth*, Chapter Accounting for Cross-Country Income Differences, pp. 679–741. Amsterdam: Elsevier Science.

Caselli, F. and J. Feyrer (2007). The marginal product of capital. *The Quarterly Journal of Economics 122*(2), 535–568.

Caspi, A., K. Sugden, T. Moffitt, A. Taylor, I. Craig, H. Harrington, J. McClay, J. Mill, J. Martin, A. Braithwaite, and R. Poulton (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science 301*(5631), 386–389.

Castells-Quintana, D. (2017). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics 98*, 158–173.

Castells-Quintana, D. and V. Royuela (2014). Agglomeration, inequality and economic growth. *The Annals of Regional Science 52*(2), 343–366.

Chan, C. H., E. D. Caine, S. You, K. W. Fu, S. S. Chang, and P. S. F. Yip (2014). Suicide rates among working-age adults in South Korea before and after the 2008 economic crisis. *Journal of epidemiology and community health 68*(3), 246–252.

Chang, S.-S., D. Stuckler, P. Yip, and D. Gunnell (2013). Impact of 2008 global economic crisis on suicide: time trend study in 54 countries. *BMJ 347:f5239*.

Chauvin, J. P., E. Glaeser, Y. Ma, and K. Tobio (2017). What is different about urbanization in rich and poor countries? cities in Brazil, China, India and the United States. *Journal of Urban Economics 98*, 17–49.

Chen, Y., J. V. Henderson, and W. Cai (2017). Political favoritism in China's capital markets and its effect on city sizes. *Journal of Urban Economics 98*, 69–87.

Chen, Y. and S. S. Rosenthal (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics 64*(3), 519–537.

Clark, C. (1951). Urban population densities. *Journal of the Royal Statistical Society. Series A (General) 114*(4), 490–496.

Costa, D. L. and M. E. Kahn (2006). *Poverty, the Distribution of Income, and Public Policy*, Chapter Public Health and Mortality: What Can We Learn from the Past?

Covarrubias, M., J. Lafortune, and J. Tessada (2015). Who comes and why? Determinants of immigrants skill level in the early XXth century US. *Journal of Demographic Economics*, 115–155.

Cowen, P. J. and M. Browning (2015). What has serotonin to do with depression? *World Psychiatry 14*(2), 158–160.

Crabtree, S. (2010). Religiosity highest in world's poorest nations. https://news.gallup.com/poll/142727/religiosity-highest-world-poorest-nations.aspx. Last accessed on 18.09.2018.

Cutler, D. M., E. L. Glaeser, and K. E. Norberg (2001). *Risky Behavior among Youths: An Economic Analysis*, Chapter Explaining the Rise in Youth Suicide, pp. 219–270. University of Chicago Press.

Dalgaard, C.-J. and H. Strulik (2015). The physiological foundation of the wealth of nations. *Journal of Economic Growth 20*, 37–73.

Dalgaard, C.-J. and H. Strulik (2016). Physiology and development: Why the West is taller than the rest. *Economic Journal 126*, 2292–2323.

Daly, M. C., D. Wilson, and N. J. Johnson (2013). Relative status and well-being: Evidence from U.S. suicide deaths. *The Review of Economics and Statistics 95*(5), 1480–1500.

Das, A. (2011). Farmers' suicide in India: implications for public mental health. *International Journal of Social Psychiatry 57*(1), 21–29.

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

Demographia (2016). Demographia: world urban areas. http://www.demographia.com/db-worldua.pdf. Last accessed on 05.02.2018.

Desmet, K. and E. Rossi-Hansberg (2013). Urban accounting and welfare. *American Economic Review 103*(6), 2296–2327.

Desmet, K. and E. Rossi-Hansberg (2014). Spatial development. *American Economic Review 104*(4), 1211–43.

Duranton, G. and D. Puga (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review 91*, 1454–1477.

Duranton, G. and D. Puga (2004). *Handbook of Urban and Agglomeration Economics*, Chapter Micro-foundations of urban agglomeration economies, pp. 2003–2118. North-Holland, Amsterdam.

Duranton, G. and M. A. Turner (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review 101*(6), 2616–52.

Durkheim, E. (1951). *Suicide: A Study in Sociology.* trans. George Simpson and John A. Spaulding (New York: Free Press).

Eastman, C. I. (1990). Natural summer and winter sunlight exposure patterns in seasonal affective disorder. *Physiology & Behavior 48*(5), 611–616.

Eastman, C. I., M. A. Young, L. F. Fogg, L. Liu, and P. M. Meaden (1998). Bright light treatment of winter depression: A placebo-controlled trial. *Archives of General Psychiatry 55*(10), 883–889.

Eurostat (2018). Degree of urbanisation. `https://ec.europa.eu/eurostat/web/degree-of-urbanisation/background`. Last accessed on 18.09.2018.

Fay, M. and C. Opal (2000). Urbanization without growth : A not-so-uncommon phenomenon. Policy Research Working Paper, No. 2412., available at `https://openknowledge.worldbank.org/handle/10986/21373`. Last accessed on 05.02.2018.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015). The Next Generation of the Penn World Table. *American Economic Review 105*(10), 3150–3182. available for download at www.ggdc.net/pwt.

Food and Agriculture Organization of the United Nations (2018). FAOSTAT. `http://www.fao.org/faostat/en/data`. Last accessed on 21.09.2018.

Fujita, M. (1989). *Urban economic theory: land use and city size.*

Fujita, M., P. Krugman, and A. Venables (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.

Galiani, S. and S. Kim (2011). *Understanding Long-Run Economic Growth: Geography, Institutions, and the Knowledge Economy*, Chapter Political Centralization and Urban Primacy: Evidence from National and Provincial Capital in the Americas, pp. 121–153. University of Chicago Press.

Gambichler, T., A. Bader, M. Vojvodic, F. G. Bechara, K. Sauermann, P. Altmeyer, and K. Hoffmann (2002). Impact of UVA exposure on psychological parameters and circulating serotonin and melatonin. *BMC Dermatology 2*(6).

Glaeser, E. L. (2008). *Cities, Agglomeration and Spatial Equilibrium.* Oxford University Press.

Glaeser, E. L. (2014). A world of cities: The causes and consequences of urbanization in poorer countries. *The Journal of the European Economic Association 12*(5), 1154–1199.

Glaeser, E. L. and J. V. Henderson (2017). Urban economics for the developing world: An introduction. *Journal of Urban Economics 98*, 1–5.

Goldman, N., D. A. Glei, Y.-H. Lin, and M. Weinstein (2010). The serotonin transporter polymorphism (5-HTTLPR): Allelic variation and links with depressive symptoms. *Depression and Anxiety 27*.

Gollin, D., R. Jedwab, and D. Vollrath (2016). Urbanization with and without industrialization. *Journal of Economic Growth 21*(1), 35–70.

Gruere, G. and D. Sengupta (2011). Bt cotton and farmer suicides in india: an evidence-based assessment. *The Journal of Development Studies 47*(2), 316–337.

Gujarati, D. N. (2003). *Basic Econometrics* (Fourth ed.). McGraw-Hill.

Gunnell, D. and M. Eddleston (2003). Suicide by intentional ingestion of pesticides: a continuing tragedy in developing countries. *International Journal of Epidemiology 32*(6), 902–909.

Gustavsson, R. (1999). Explaining the phenomenon of Third World urban giants. *Journal of Economic Integration 14*(4), 625–650.

Haines, M. R. (2001). The urban mortality transition in the United States, 1800-1940. *NBER Historical Working Paper* (104).

Hakko, H., P. R´as´anen, and J. Tiihonen (1998). Seasonal variation in suicide occurrence in Finland. *Acta Psychiatrica Scandinavica 98*(2), 92–97.

Hamermesh, D. and N. Soss (1974). An economic theory of suicide. *Journal of Political Economy 82*(1), 83–98.

Harris, J. R. and M. P. Todaro (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review 60*(1), 126–142.

Healy, D. (2015). Serotonin and depression. *British Medical Journal 350*.

Helliwell, J. F. (2007). Well-being and social capital: Does suicide pose a puzzle? *Social Indicators Research 81*, 455–496.

Henderson, J. V. (1974). The sizes and types of cities. *American Economic Review* (64), 640–656.

Henderson, J. V. (2003, March). 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*, Chapter Urbanization and Growth. North-Holland, Amsterdam.

Henderson, J. V. (2010). Cities and development. *Journal of regional science 50*(1), 515–540.

Henderson, J. V., A. Storeygard, and U. Deichmann (2017). Has climate change driven urbanization in Africa? *Journal of Development Economics 124*, 60–82.

Henderson, J. V. and H. G. Wang (2007). Urbanization and city growth: the role of institutions. *Regional Science and Urban Economics 37*(3), 283–313.

Henderson, V., T. Squires, A. Storeygard, and D. Weil (2018). The global distribution of economic activity: nature, history, and the role of trade. *The Quarterly Journal of Economics* (forthcoming).

Hoy, M. and E. Jimenez (1991). Squatters' rights and urban development: an economic perspective. *Economica 58*(229), 79–92.

Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin, and B. Puranen (Eds.) (2014). *World Values Survey: Round Six - Country-Pooled Datafile Version: http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp*. Madrid: JD Systems Institute.

Iyengar, B. (1994). Indoleamines and the UV-light-sensitive photoperiodic responses of the melanocyte network: a biological calendar? *Experientia 50*(8), 733–736.

Jalles, J. and M. Andresen (2015). The social and economic determinants of suicide in Canadian provinces. *Health Economics Review 5*(1).

Jedwab, R., L. Christiansen, and M. Gindelsky (2017). Demography, urbanization and development: Rural push, urban pull and ... urban push? *Journal of Urban Economics 98*, 6–16.

Jedwab, R. and D. Vollrath (2015). The mortality transition, malthusian dynamics, and the rise of poor mega-cities.

Jedwab, R. and D. Vollrath (2018). The urban mortality transition and poor country urbanization. *American Economic Journal - Macroeconomics.*.

Jimenez, E. (1985). Urban squatting and community organization in developing countries. *Journal of Public Economics* (27), 69–92.

Kambeitz, J. P. and O. D. Howes (2015). The serotonin transporter in depression: Meta-analysis of in vivo and post mortem findings and implications for understanding and treating depression. *Journal of Affective Disorders 186*, 358–366.

Karg, K., M. Burmeister, K. Shedden, and S. Sen (2011). The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: Evidence of genetic moderation. *Archives of General Psychiatry 68*(5), 444–454.

Kendler, K. S., J. W. Kuhn, J. Vittum, C. A. Prescott, and B. Riley (2005). The interaction of stressful life events and a serotonin transporter polymorphism in the prediction of episodes of major depression: a replication. *Archives of General Psychiatry 62*(5), 529–535.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*, 97–126.

Knapp, T. A. and P. E. Gravest (1989). On the role of amenities in models of migration and regional development. *Journal of Regional Science 29*(1), 71–87.

Kocornik-Mina, A., T. K. McDermott, G. Michaels, and F. Rauch (2015). Flooded cities. *CEP Discussion Paper* (1398).

Koo, J. and W. M. Cox (2008). An economic interpretation of suicide cycles in Japan. *Contemporary Economic Policy 26*(1), 162–174.

Kovags, M., A. T. Beck, and A. Weissman (1975). The use of suicidal motives in the psychotherapy of attempted suicides. *The American Journal of Psychotherapy 29*(3), 363–368.

Lagakos, D. and M. E. Waugh (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review 103*(2).

Lall, S. V., C. Timmins, and S. Yu (2009). Connecting lagging and leading regions: The role of labor mobility. *World Bank Policy Research Working Paper* (4843).

Lambert, G., C. Reid, D. Kaye, G. Jennings, and M. Esler (2002). Effect of sunlight and season on serotonin turnover in the brain. *The Lancet 360*(9348), 1840–1842.

Lambert, G., C. Reid, D. Kaye, G. Jennings, and M. Esler (2003). Increased suicide rate in the middle-aged and its association with hours of sunlight. *The American Journal of Psychiatry 160*(4), 793–795.

Lee, H.-Y., J.-P. Hong, J.-A. Hwang, H.-J. Lee, H.-K. Yoon, B.-H. Lee, and Y.-K. Kim (2015). Possible association between serotonin transporter gene polymorphism and suicide behavior in major depressive disorder. *Psychiatry Investigation 12*.

Lind, J. T. and H. Mehlum (2010). With or without U? The appropriate test for a U shaped relationship. *Oxford Bulletin of Economics and Statistics* (72), 109–18.

Lucas, R., T. McMichael, W. Smith, and B. Armstrong (2006). Solar ultraviolet radiation: Global burden of disease from solar ultraviolet radiation. *Environmental Burden of Disease Series* (13).

Lynch, V. A. and J. B. Duval (2011). *Forensic Nursing Science* (2nd ed.). Mosby Inc.

Mårtensson, B., A. Pettersson, L. Berglund, and L. Ekselius (2015). Bright white light therapy in depression: A critical review of the evidence. *Journal of Affective Disorders 182*, 1–7.

Machado, D., D. Rasella, and D. dos Santos (2015). Impact of income inequality and other social determinants on suicide rate in Brazil. *PLOS ONE 10*(4).

Madianos, M. G., T. Alexiou, A. Patelakis, and M. Economou (2014). Suicide, unemployment and other socioeconomic factors: evidence from the economic crisis in Greece. *The European Journal of Psychiatry 28*(1), 39–49.

Mahfoud, Z. R., R. A. Afifi, P. H. Haddad, and J. DeJong (2011). Prevalence and determinants of suicide ideation among Lebanese adolescents: Results of the GSHS Lebanon 2005. *Journal of Adolescence 34*(2), 379–384.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Oxford University Press.

Maslow, A. (1943). A theory of human motivation. *Psychological Review 50*(4), 370–96.

Mathur, V. K. and D. G. Freeman (2002). A theoretical model of adolescent suicide and some evidence from us data. *Health economics 11*(8), 695–708.

McDowell, G., E. Mules, P. Fabacher, E. Shapira, and M. Blitzer (1992). The presence of two different infantile TayŨSachs disease mutations in a Cajun population. *American Journal of Human Genetics 51*(5), 1071–1077.

Meltzer, H. Y. (1989). Serotonergic dysfunction in depression. *The British Journal of Psychiatry 155*(8), 25–31.

Michaels, G., F. Rauch, and S. J. Redding (2012). Urbanization and structural transformation. *The Quarterly Journal of Economics 127*(2), 535–586.

Mills, E. (1980). A comparison of urban population density functions in developed and developing countries. *Urban studies 17*(3).

Milner, A., R. McClure, and D. D. Leo (2012). Socio-economic determinants of suicide: an ecological analysis of 35 countries. *Social Psychiatry and Psychiatric Epidemiology 47*(1), 19–27.

Minkoff, K., E. Bergman, and A. Beck (1973). Hopelessness, depression, and attempted suicide. *The American Journal of Psychiatry 130*(4), 455–459.

Mohring, H. (1961). Land values and measurement of highway benefits. *Journal of Political Economy 49*, 236–249.

Mookherjee, D. and D. Ray (2003). Persistent inequality. *Review of Economic Studies*, 369–393.

Myerowitz, R. and F. Costigan (1988). The major defect in Ashkenazi Jews with TayŨSachs disease is an insertion in the gene for the alpha-chain of beta-hexosaminidase. *Journal of Biological Chemistry 263*(35), 18587–18589.

Nakatsuka, N., P. Moorjani, N. Rai, B. Sarkar, A. Tandon, N. Patterson, G. S. Bhavani, K. M. Girisha, M. S. Mustak, S. Srinivasan, A. Kaushik, S. A. Vahab, S. M. Jagadeesh, K. Satyamoorthy, L. Singh, D. Reich, and K. Thangaraj (2017). The promise of discovering population-specific disease-associated genes in South Asia. *Nature Genetics 49*(9), 1403–1407.

NASA (2002). SOLVE II Science Implementation. `https://cloud1.arc.nasa.gov/solveII/implement.html`. Last accessed on 20.07.2018.

Neumayer, E. (2003). Are socioeconomic factors valid determinants of suicide? controlling for national cultures of suicide with fixed-effects estimation. *Cross-Cultural Research 37*(3), 307–329.

Okada, K. and S. Samreth (2013). A study on the socio-economic determinants of suicide: Evidence from 13 european OECD countries. *The Journal of Socio-Economics 45*, 78–85.

Patel, V., C. Ramasundarahettige, L. Vijayakumar, J. Thakur, V. Gajalakshmi, G. Gururaj, W. Suraweera, and P. Jha (2012). Suicide mortality in india: a nationally representative survey. *The Lancet 379*(9834), 2343–2351.

Pesaresi, M. and S. Freire (2016). GHS Settlement grid following the REGIO model 2014 in application to GHSL Landsat and CIESIN GPW v4-multitemporal (1975-1990-2000-2015). European Commission, Joint Research Centre (JRC) [Dataset]. `http://data.europa.eu/89h/jrc-ghsl-ghs_smod_pop_globe_r2016a`. Last accessed on 18.09.2018.

Petridou, E., F. Papadopoulos, C. Frangakis, A. Skalkidou, and D. Trichopoulos (2002). A role of sunshine in the triggering of suicide. *Epidemiology 13*(1), 106–109.

Pew Research Center's Forum on Religion & Public Life (2012). The Global Religious Landscape: A report on the size and distribution of the world's major religious groups as of 2010. `http://www.pewforum.org/files/2014/01/global-religion-full.pdf`. Last accessed on 18.09.2018.

Phillips, M. R., G. Yang, Y. Zhang, L. Wang, H. Ji, and M. Zhou (2002). Risk factors for suicide in China: a national case-control psychological autopsy study. *The Lancet 360*(9347), 1728–1736.

Putterman, L. and D. N. Weil (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. *Quarterly Journal of Economics 125*(4), 1627–82.

Qaisar, F., S. Mohsin, M. Ayesha, K. Dileep, M. Awais, and M. Umer (2014). The epidemiology of deliberate self-poisoning presenting at a tertiary care hospital in Hyderabad Sindh, Pakistan. *British Journal of Medicine and Medical Research* (4), 1041–1048.

Ramachandran, S., O. Deshpande, C. C. Roseman, N. A. Rosenberg, M. W.Feldman, and L. L. Cavalli-Sforza (2005). Support from the Relationship of Genetic and Geographic Distance in Human Populations for a Serial Founder Effect Originating in Africa. *Proceedings of the National Academy of Sciences 102*(44), 15942–47.

Ramim, T., M. Mobayen, N. Shoar, M. Naderan, and S. Shoar (2013). Burnt wives in Tehran: a warm tragedy of self-injury. *International Journal of Burns and Trauma 3*(1), 66–71.

Restuccia, D., D. T. Yang, and X. Zhu (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics 55*(2), 234–250.

Risch, N., R. Herrell, T. Lehner, K.-Y. Liang, L. Eaves, J. Hoh, A. Griem, M. Kovacs, J. Ott, and K. R. Merikangas (2009). Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA 301*(23), 2462–2471.

Rosen, K. T. and M. Resnick (1980). The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics 8*(2), 165–186.

Rousseaux, M. C., C. L. Ballare, C. V. Giordano, A. L. Scopel, A. M. Zima, M. Szwarcberg-Bracchitta, P. S. Searles, M. M. Caldwell, and S. B. Diaz (1999). Ozone depletion and UVB radiation: Impact on plant DNA damage in southern South America. *PNAS 96*(26), 15310–15315.

Sansone, R. A. and L. A. Sansone (2013). Sunshine, serotonin, and skin: A partial explanation for seasonal patterns in psychopathology? *Innovations in Clinical Neuroscience 10*(7-8), 20–24.

Shah, N. (2014). Squatting on geovernment land. *Journal of Regional Science 54*(1), 114–136.

Skorobogatov, A. S. (2018). Why do newer cities promise higher wages in Russia? *Journal of Urban Economics 104*, 16–34.

Spolaore, E. and R. Wacziarg (2009). The diffusion of development. *Quarterly Journal of Economics 124*(2), 469–529.

Spolaore, E. and R. Wacziarg (2013). How deep are the roots of economic development? *Journal of Economic Literature 51*(2), 325–69.

Stock, J. and M. Yogo (2005). *Identification and Inference for Econometric Models*, Chapter Testing for Weak Instruments in Linear IV Regression., pp. 80–108. New York: Cambridge University Press.

Strulik, H. (2017). Myopic misery: Maternal depression, child investments, and the neurobiological poverty trap. *Macroeconomic Dynamics*, 1–13.

Tatarczynska, E., A. Klodzinska, K. Stachowicz, and E. Chojnacka-Wojcik (2004). Effects of a selective 5-HT1B receptor agonist and antagonists in animal models of anxiety and depression. *Behavioural Pharmacology 15*(8), 523–534.

The World Bank (2018a). International Comparison Program (ICP). `http://www.worldbank.org/en/programs/icp`. Last accessed on 21.09.2018.

The World Bank (2018b). World Bank Open Data. `http://data.worldbank.org`. Last accessed on 15.08.2018.

Toprak, S., I. Cetin, T. Guven, G. Can, and C. Demircan (2011). Self-harm, suicidal ideation and suicide attempts among college students. *Psychiatry Research 187*(1-2), 140–144.

Troesken, W. (2004). *Water, Race, and Disease*. MIT Press.

Turnbull, G. (2008). Squatting, eviction and development. *Regional Science and Urban Economics 38*(1), 1–15.

UN-Habitat (2016). *World cities report*. Nairobi, Kenya: UN-Habitat.

UNESCO (2015). *Education for All Global Monitoring Report*. United Nations Educational, Scientific and Cultural Organization.

United Nations (2015). *World Urbanization Prospects: The 2014 Revision*.

Weatherall, D. J. and J. B. Clegg (2001). Inherited haemoglobin disorders: an increasing global health problem. *Bulletin of the World Health Organisation 79*(8), 704–712.

Wheaton, W. C. (1974). A comparative static analysis of urban spatial structure. *Journal of Economic Theory 9*(2), 223–237.

WHO (2014). *Preventing suicide: A global imperative*. World Health Organization.

Williamson, J. G. (1965). Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change*, 3–45.

Williamson, J. G. (2010). Five centuries of Latin American income inequality. *Revista de Historia Economica / Journal of Iberian and Latin American Economic History 28*, 227–252.

World Health Organisation (2010). *The world health report: health systems financing: the path to universal coverage*. WHO Press.

Yakubenko, S. (2018). Giants and midgets: the effect of public goods' provision on urban population concentration.

Young, S. (2007). How to increase serotonin in the human brain without drugs. *Journal of Psychiatry & Neuroscience 32*(6), 394–399.

Zawilska, J. B., J. Rosiak, K. Trzepizur, and J. Z. Nowak (2007). The effects of near-ultraviolet light on serotonin N-acetyltransferase activity in the chick pineal gland. *Journal of Pineal Research 26*(2), 122–127.