Are urban areas really bad places to live in?
Evidence from developing Asia

Racquel Claveria[a,b]

Oscar Claveria[a]

Rosina Moreno[a]

ABSTRACT

We assemble a rich dataset that maps geocoded microdata on various measures of living standards, point-specific geographical attributes, and remotely-sensed pollution data onto gridded population data to determine whether urbanization yields greater benefits or costs. We focus on developing Asia where the world's largest proportion of urban population is projected to come from. We find evidence that population density, our chosen measure of urban scale, provides greater benefits relative to costs, indicating that urbanization brings net positive gains to the developing world. In particular, our causal estimates show that higher population density leads to greater household wealth, improved hygiene and sanitation, and better access to utilities, health services, clean energy, and private goods and services. On the other hand, our results suggest that density either reduces or does not significantly affect urban costs such as indoor pollution, space congestion, diseases, and crime. The only exception is outdoor pollution which deteriorates with population density.

Keywords: urban growth, developing countries, causal inference, instrumental variables, identification
JEL classification: R11, O18, C26, C3

1. INTRODUCTION

Urbanization in the developing world often evokes images of overcrowded cities, poverty-stricken areas, and dismal environmental conditions (World Bank 2008). Hence, it is unsurprising that low-income countries have been inclined to curb urban growth, primarily due to the perceived perils of urbanization and authorities' limited capacity to deal with those challenges (Glaeser and Henderson 2017; Glaeser and Xiong 2017). Unlike their richer counterparts, developing countries have urbanized rapidly which gave them limited room to raise funds and build institutions that are essential for a reasonable quality of urban life (Henderson 2002). As a consequence, poor world cities address the downsides of urbanization, such as contagious diseases, pollution, congestion, and crime, with limited financial resources and public capacity (Bryan et al. 2020; Glaeser 2020; Henderson and Turner 2020). The COVID-19 pandemic painfully depicts the constraints faced by developing world cities in grappling with the challenges of urbanization. With high rates of infection amid overcrowded living conditions, and limited access to vaccines due to inadequate funds, cities in poor countries are at the losing end of the pandemic (*The Economist* 2021a, 2021b). Against this backdrop, should national governments in the developing world restrain urbanization?

The answer to this pressing policy question rests critically on the estimates of the impact of urbanization on the benefits and costs it gives rise to. Indeed, urbanization generates benefits and costs, dubbed as the "fundamental trade-off of the spatial economy" (Fujita and Thisse 2002). As is well established in the literature, urban activity yields agglomeration economies or the productivity gains derived by firms and workers located in the same area (Combes et al. 2019). Moreover, urbanization provides greater accessibility to public and private goods and services (Duranton and Puga 2020; Overman and Venables 2005; World Bank 2008). On the flipside, urban dwellers face greater exposure to contagious diseases, crime and pollution, in addition to higher housing, transportation, and consumption costs that urbanites incur (Glaeser and Gottlieb 2009).

For policymaking purposes, estimates of the impact of urbanization on its associated benefits and costs need to go beyond the latter's mere correlation with urban scale, and should thus reflect the true causal effect of urbanization (Combes and Gobillon 2015). These causal estimates are particularly useful for the developing world which has limited financial and institutional capacity, and hence needs to prioritize which urban benefits or costs to address (Bryan et al. 2020). In addition, an inaccurate reckoning of urban benefits or costs in poor countries can lead to urban policy mistakes that could in turn result in even more economic, social and environmental harm (United Nations 2019ab).

Yet, a host of possibilities confound the causal relationship between urbanization and its associated benefits and costs. People may be drawn to or avoid urban areas because of certain urban benefits (e.g., greater accessibility to goods and services) or urban costs (e.g. higher incidence of crime). Unobserved characteristics such as ambition and motivation can drive both an urban benefit (for instance, productivity) and the degree of urbanization. More educated people tend to reside in cities which further raise productivity and urban population. These possibilities pertain to the identification issues that are repeatedly raised in the literature such as reverse causality, omitted variables, and sorting (more able people locating themselves to more productive areas). Such issues need to be addressed in order to extract the true causal impact of urbanization (Ahlfeldt and Pietrostefani 2019; Combes and Gobillon 2015; Duranton and Puga 2020).

Nevertheless, the literature on the causal estimates of the impact of urbanization on urban benefits and costs appears to be limited in scope and coverage. A number of studies mainly focus on urban benefits, particularly on productivity, and primarily cover the developed world (see reviews and meta-analyses by Ahlfeldt and Pietrostefani 2019; Combes and Gobillon 2015; Combes et al. 2012; Melo et al. 2009; Moretti 2011; Puga 2010; and Rosenthal and Strange 2004). At the same time, scholars lament the paucity of research and lack of systematic evidence on urban costs (Ahlfeldt and Gobillon 2021; Combes and Gobillon 2015; Combes et al. 2019; Duranton 2008; Duranton and Puga 2020; Overman and Venables 2005).

Hence, this paper seeks to fill the current gap in the literature by providing, to the extent possible, causal estimates of the impact of urbanization on a broad range of urban benefits and costs, particularly those that are most relevant in the developing world context. Such causal estimates can shed light on whether urbanization generates greater benefits or costs, and therefore provide a useful guide on whether policies that limit urbanization are warranted or not.

To cover a wide array of urban benefits and costs, we take advantage of the emergence of high-resolution data and advances in mapping technology. In particular, following Gollin et al. (2021), we assemble a rich dataset that spatially links geocoded microdata on various measures of living standards, point-specific geographical attributes, and remotely-sensed pollution data with gridded population data. We employ population density as measure of urban scale, and treat density space as a continuum. This avoids subjective definitions of "urban areas" which vary across jurisdictions and limit meaningful cross-country analysis (Henderson and Turner 2020; OECD/European Commission 2020). Moreover, we focus on Asia, home to more than half of the world's urban population (United Nations 2019b), and where the largest proportion of urban population is expected to come from. Based on the availability of GPS coordinates and timeliness of survey data, we cover 11 developing Asian countries such as Armenia, Bangladesh, Cambodia, India, Kyrgyz Republic, Myanmar, Nepal, Pakistan, Philippines, Tajikistan, and Timor Leste. These countries represent 30 percent of the estimated total urban population in the region in 2020 (United Nations 2018).

Our data sources enable us to employ indicators and proxy variables for a wide spectrum of urban benefits such as household wealth (a longer-term measure of living standards relative to wages and TFP), accessibility to private and public goods and services, and clean energy; and urban costs such as indoor and outdoor pollution, space congestion, diseases, and crime. At the same time, our dataset allows us to instrument population density with its corresponding 40-year lag in order to deal with the issues of reverse causality and omitted variables.

While we find that historical population density is a relevant and exogenous instrument, we also account for its possible endogeneity by controlling for local permanent characteristics that could influence both historical population density and an urban benefit or cost of interest, as proposed by Combes and Gobillon (2015). These include proximity to national borders and bodies of water, annual rainfall, and temperature volatility.

Meanwhile, we address sorting of more able and educated individuals to urban areas by including individual and household controls such as age, sex and education of women and household head, among others. Finally, we recognize that some urban benefits such as the availability of public goods and services may be supply-driven, plausibly due to the preferential treatment given to national capitals which are highly urbanized areas. We account for this urban bias by controlling for whether a household is in a national capital or not.

The causal estimates that we obtain tend to show that population density provides greater benefits relative to costs, implying that urbanization brings net positive gains to the developing world. Even after addressing the endogeneity of population density and controlling for household and individual characteristics and geographical attributes, living standards appear to improve with density in developing Asia. In particular, our causal estimates suggest that higher population density leads to greater household wealth, improved hygiene and sanitation, and greater access to utilities, health services, clean energy, and private goods and services. On the other hand, density either reduces or does not significantly affect the urban costs covered in this study such as indoor pollution, space congestion, diseases, and crime. The only exception is outdoor pollution which deteriorates with density.

Our finding that urbanization appears to generate greater benefits relative to costs squares well with earlier results that living standards improve with population density (Ahlfeldt and Pietrostefani 2019, de Weerdt et al. 2021, Gollin et al. 2021, Henderson and Turner 2020 and Henderson et al. 2020). Nevertheless, the results obtained from these studies indicate associations, and do not estimate the causal effects of population density, which this paper seeks to provide. Hence, we contribute to the literature by obtaining causal estimates of the impact of population density on various urban benefits and costs. We show evidence that the net causal gains from density in the developing world are significant.

This paper is close to Gollin et al. (2021), Henderson and Turner (2020), and Henderson et al. (2020) who also linked geocoded microdata with gridded population data in order to determine the relationship between urbanization and various urban outcomes in the developing world, with particular focus on Africa. However, we deviate from the aforementioned studies in two main aspects. Firstly, we cover Asia which houses the greatest proportion of global and developing world urban population. Secondly, we pay close attention to key identification concerns in order to estimate the causal impact of density on various urban benefits and costs. To the best of our knowledge, this is the first paper to estimate the causal effects of density on a broad set of urban benefits and costs.

This study likewise contributes to other strands in the urban economics literature. It helps bridge the gap between the advanced knowledge on developed world cities and the nascent research on developing world urbanization, as pointed out by Bryan et al. (2020) and Glaeser and Henderson (2017)). Moreover, it also adds to the emerging literature on the application of point-specific and high-resolution data in analyzing the effects of urbanization (see previews by Duranton and Rosenthal (2021) and Ahlfeldt and Gobillon (2021)).

The rest of the paper is organized as follows. Section 2 provides a simple motivating framework for my empirical strategy and interpretation of results. Section 3 presents the data. Section 4 lays down my empirical framework. Section 5 discusses the results. Finally, Section 6 concludes and offers some policy recommendations.

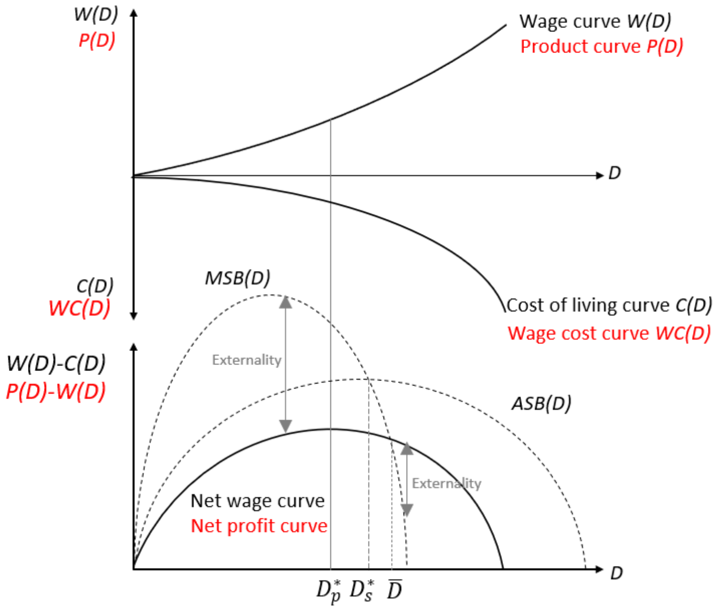
2. CONCEPTUAL FRAMEWORK

To motivate our empirical strategy and to guide the interpretation of results, we synthesize the diagrammatic frameworks of Combes et al. (2005) and Overman and Venables (2005) on the formation of cities, which both draw inspiration from Henderson (1974). We illustrate graphically the relationship between urban benefits and costs, and population density.[1] The graphical illustration likewise depicts the trade-off between the urban costs and benefits, both from the perspectives of the worker and the firm.

A well-established urban benefit in the literature is the higher productivity of workers and firms (see meta-analysis by Ahlfeldt and Pietrostefani (2019)). Urban increasing returns arise due to more efficient sharing of local infrastructure, risks and gains; better matching between employers and employees, or buyers and suppliers; and greater learning through the transmission and accumulation of skills or the development and adoption of new technologies and business practices (Duranton and Puga 2020). These mechanisms of sharing, matching and learning which were formally articulated in Duranton and Puga (2004) occur in the markets for labor, intermediate inputs, and knowledge, as in Marshall (1890). From the point of view of the worker, higher productivity is reflected in the wage curve $W(D)$ which is upward sloping with respect to population density D (see Figure 1).

[1] In contrast with Combes et al. (2005), and Overman and Venables (2005), we utilize population density in lieu of total population in the conceptual framework above.

Figure 1. Diagrammatic framework: urban benefits and costs, and population density



Note: The figure depicts the costs and benefits brought forth by population density, and the trade-off between the two, both from the perspective of the worker (labeled in black) and the firm (marked in red).
Source: Authors' construction based on Combes et al. (2005), and Overman and Venables (2005).

Meanwhile, the worker likewise faces a cost of living curve *C(D)*, plotted at the inverse Y-axis. The *C(D)* curve represents transport, housing and consumption expenses, and likewise increases with *D*. The positive slope of the cost of living curve rests on the assumption that higher population per area leads to higher commuting costs as more travelers congest roads and make commuting time longer, higher housing prices as greater number of residents drive up land prices, and higher consumption expenses as more expensive land prices feed into retail prices (Duranton 2008; Combes et al. 2019). The difference between the wage curve and the cost of living curve is represented by the bell-shaped net wage curve in the lower panel of Figure 1. Net wages increase initially as wages dominate living costs at low population densities then decline as the cost of living exceeds wages at high population densities (Combes et al. 2005).

On the other hand, from the perspective of the firm, the productivity benefits of urbanization are represented by the product curve *P(D)*, measured as the value of output per worker. The product curve is likewise upward sloping with respect to *D* due to higher output per worker in urban areas. At the same time, the firm needs to compensate the more productive workers with higher wages, and thus the firm faces the wage cost curve *WC(D)*. It is also increasing in *D* as more workers per area entail greater wage payments. Subtracting the wage cost curve from the product curve yields the bell-shaped net profit curve of firms. It initially rises due to increasing value of workers' output, and eventually declines as more workers per area need to be compensated by higher wages.

Nevertheless, urbanization is riddled with market failures. Urban proximity gives rise to externalities, both positive and negative (Bryan et al. 2020; Duranton and Puga 2020;

Glaeser 2020; Glaeser and Xiong 2017b; World Bank 2008). For instance, the concentration of economic activity in high-density areas provides greater accessibility to various private goods and services. Public goods are also more likely to be available in cities due to the lower cost of providing them in dense areas, and the government's propensity to yield to the political power of the urban people who demand better public services, or what is referred as "urban bias" (Gollin et al. 2021; Overman and Venables 2005). On the other hand, dense areas yield negative externalities that naturally arise from people locating close to one another. Standard examples include diseases (closeness allows bacteria and viruses to spread more easily) and congestion (higher population increases the demand for and constrains the supply of land and housing) (Glaeser and Xiong 2017b). In addition, crime thrives in cities because criminals find victims more easily on crowded streets and use urban anonymity to escape detection (Glaeser 2020). Moreover, the congregation of households and industries in urban areas exacerbates air pollution.

Meanwhile, the bell-shaped net wage/net profit curves represent private returns. Yet, private and social returns, in general, do not coincide (Duranton 2008). A plausible reason is that individual objectives diverge from social objectives. Even when both goals converge, workers and firms may not correctly reckon the social gains from urbanization (Combes and Gobillon 2015). As such, the lower panel of Figure 1 depicts the average social benefit curve *ASB(D)* distinctly from the net wage/net profit curves, following Overman and Venables (2005). The *ASB(D)* represents the total returns in the urban area, expressed per individual. The marginal social benefit curve *MSB(D)* is derived from the *ASB(D)*, with the former intersecting the latter at the maximum of *ASB(D)*, or at point D_S^* where social returns are maximized. On the other hand, optimal private returns are reached at D_P^*.

The difference between marginal social benefits and private returns, respectively, correspond to externalities brought forth by population density. To the left of point $\overline{D}$ (where marginal social benefits and private returns coincide), density brings greater positive externalities relative to negative externalities, while the converse occurs to the right of point $\overline{D}$. In sum, to the left of point $\overline{D}$, density yields greater urban benefits which include the productivity benefits, as embodied in the wage and product curves, as well as positive urban externalities (e.g., better accessibility to public and private goods and services). Meanwhile, to the right of point $\overline{D}$, density creates greater urban costs which consist of transport, housing, consumption and wage costs, as represented by the cost of living and wage cost curves, as well as negative urban externalities (e.g. pollution, congestion, and crime).

This simple conceptual framework suggests that when the net benefits generated by density are large (i.e. to the left of D_S^*, the point where the optimal level of social benefits is reached), there is scope to increase population density in order to maximize social benefits. Conversely, when net urban costs engendered by density are sizeable (i.e. to the right of D_S^*), curbing urban growth may be warranted. Hence, from a policy perspective, it is worth determining whether density brings larger urban benefits or costs. To do so, two tasks are in order. First, a wide range of urban costs and benefits need to be considered. Second, it should be ascertained that such benefits and costs are indeed brought forth by and not merely associated with density. We seek to achieve these tasks in the succeeding sections.

3. DATA

To analyze a broad range of urban benefits and costs, we take advantage of the availability of high-resolution data and advances in mapping technology. Following Gollin et al. (2021), we assemble a rich dataset that maps geocoded microdata on various measures of living standards, point-specific geographical attributes, and remotely-sensed pollution data onto gridded population data.

3.1 Main databases

We extract granular data from the Global Human Settlement Population Grid (GHS-POP), Demographic and Health Survey (DHS), and National Aeronautics and Space Administration - Socioeconomic Data and Applications Center (NASA - SEDAC).

3.1.1 Global Human Settlement Population Grid (GHS-POP)

We employ population density as our measure of urban scale and as the key independent variable. Following Henderson and Turner (2020), we treat density space as a continuum and avoid subjective definitions of "urban areas" which vary from country to country. We obtain population density data from the GHS-POP which estimates population in each grid cell of 250 by 250 meters for the years 1975, 1990, 2000 and 2015 (Schiavina et al. 2019; Florczyk et al. 2019). The GHS-POP is one of the global grids in the Global Human Settlements Layer (GHSL) produced by the European Commission's Joint Research Center. The GHSL remains the only detailed grid spanning 40 years and the only disaggregated global dataset relying on a single, time-specific, and consistent built-up areas. For our estimations, we employ the 1975 and 2015 population density grids.

3.1.2 Demographic and Health Surveys (DHS)

The DHS are one of the largest survey programs in the developing world. They are nationally-representative household surveys that provide microdata for a wide range of indicators such as population, health, nutrition, household assets, and women's and children's conditions, among others (see Appendix A).

Since 2003, the DHS provided GPS coordinates for selected surveys allowing researchers to spatially link DHS data with external georeferenced data (Henderson et al. 2020). To protect the confidentiality of survey respondents, the GPS coordinates are randomly displaced by up to two kilometers for urban clusters, and up to five kilometers for rural clusters (with one percent of the rural clusters displaced up to 10 kilometers) (Burgert et al. 2013). As proposed by Perez-Haydrich et al. (2013), we construct a five-kilometer buffer around each urban and rural cluster to account for the random displacement made by the DHS (see Appendix B). We then extract the average population density and average pollution concentration (as described in 3.1.3 below) within the buffer surrounding the DHS cluster location.

In this study, we include Asian countries that have: (1) DHS GPS coordinates, and (2) latest survey not earlier than 2012. As a result, we cover 11 developing Asian countries, namely, Armenia, Bangladesh, Cambodia, India, Kyrgyz Republic, Myanmar, Nepal, Pakistan, Philippines, Tajikistan, and Timor Leste. In sum, our sample consists of more than 700,000 households and 900,000 individuals (see Table A1).

The rich DHS database provides indicators and proxy variables for a wide range of urban benefits such as a wealth index, access to utilities, hygiene and sanitation facilities, and clean energy, as well as child inoculations. Travel time to reach a high-density urban center (as proxy for availability of private goods and services) along with controls for first-nature characteristics are derived from DHS geospatial covariates generated by Mayala et al. (2018). Likewise, we obtain from the DHS the measures of urban costs such as indoor pollution, space congestion, and prevalence of diseases.

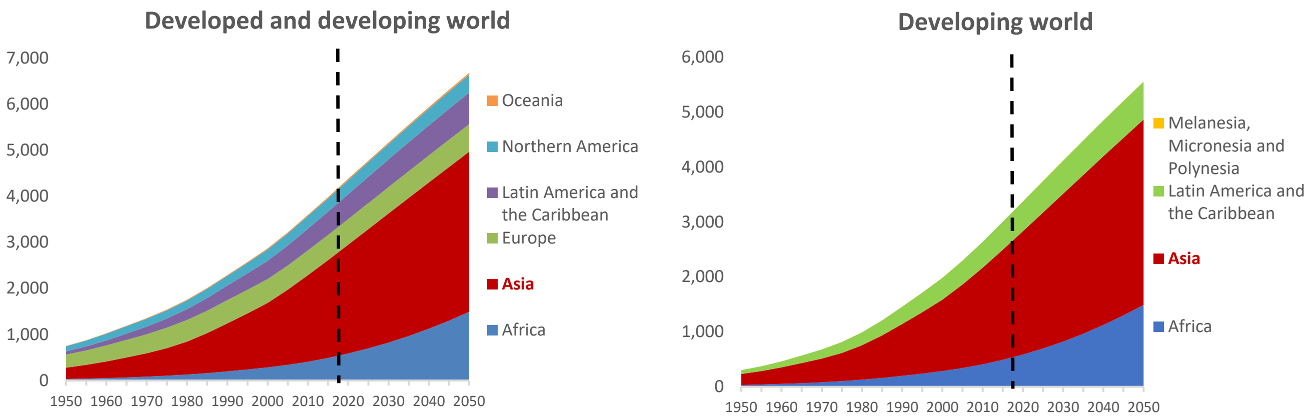
3.1.3 National Aeronautics and Space Administration - Socioeconomic Data and Applications Center (NASA-SEDAC)

The NASA-SEDAC synthesizes earth science and socioeconomic data in support of policymaking and applied research. Among its knowledge products are satellite-derived environmental indicators. We extract from the NASA-SEDAC, through van Donkelaar et al. (2018) and Geddes et al. (2017), remotely-sensed data on air pollutants that are usually generated by urban activities, such as ground-level fine particulate matter (PM) and nitrogen dioxide (NO_2) concentrations. For instance, PM, a common proxy indicator for air pollution, is produced by various industrial processes as well as burning of fossil fuels for heating or energy production, and through dispersion of dust on streets and tire wear of cars (Borck and Schrauth 2021). It affects more people than any other pollutant, particularly those with a diameter of 2.5 microns or less (PM2.5) which can penetrate the lung barrier and enter the blood system (WHO 2018). Concentrations of NO_2 are likewise produced from the combustion of fuel and from emissions from cars, trucks and buses, power plants, and off-road equipment.

3.2 Country Coverage

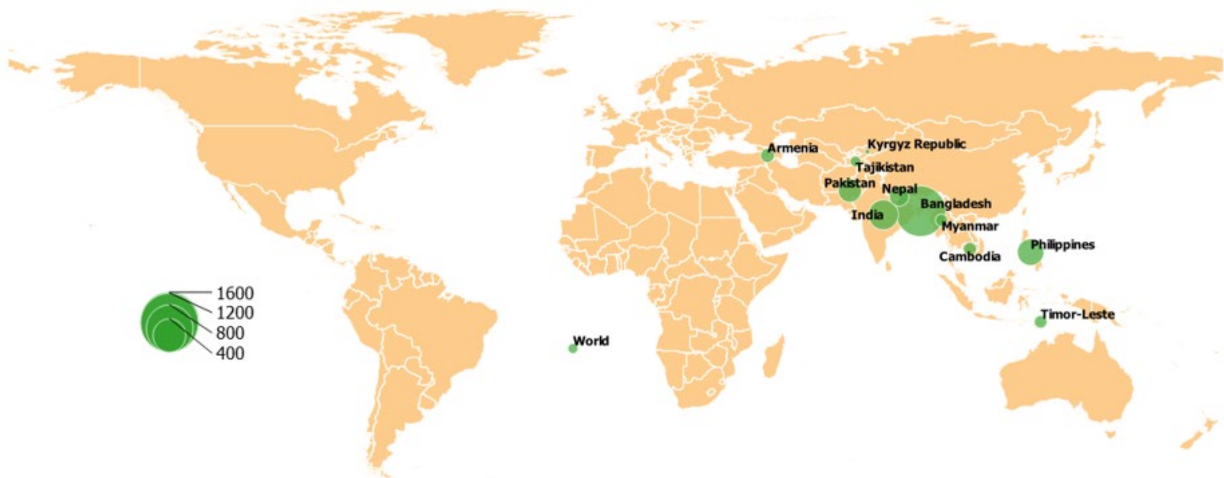
Previous studies that examined urban living standards in the developing world mainly focused on Africa (see de Weerdt et al. 2021; Gollin et al. 2021; Henderson and Turner 2020; Henderson et al. 2020). Yet more than half of the global and developing world urban population is expected to come from Asia (see projections in Figure 2). Moreover, the percentage of population living in urban areas in Asia is nearing 50%, higher than the 40-percentage urban in Africa (United Nations 2019b). For these reasons, we focus our analysis on the Asian region. As mentioned in Section 3.1.1, we select 11 countries in developing Asia based on the availability of GPS coordinates and timeliness of survey data in the DHS. Data from United Nations (2018) indicate that these countries comprise 30% of the estimated urban population in Asia, and about 20% in the entire developing world in 2020. According to the World Bank's classification, all the countries in our sample are categorized as lower-middle-income countries, except for Nepal which is considered low-income (United Nations 2019b). Moreover, average population density in these countries is 278 persons per square kilometer, almost five times the world average of 60 persons per square kilometer (Figure 3).

Figure 2. Urban population at mid-year, 1950-2050 (millions)



Note: Actual population estimates lie to the left of the dashed line; projected population to the right.
Source: Authors' construction using data from United Nations, Department of Economic and Social Affairs, Population Division (2018). World Urbanization Prospects: The 2018 Revision, Online Edition. Accessed 27 April 2021.

Figure 3. Population density (persons per square km), 2020



Source: Authors' construction using data from United Nations, Department of Economic and Social Affairs, Population Division (2018). World Urbanization Prospects: The 2018 Revision, Online Edition. Accessed 15 April 2021.

3.3 Variables and descriptives

Table C1 provides the description and sources of the variables employed in this study while Tables C2 and C3 show the corresponding descriptive statistics. The majority of the variables are drawn from survey data, and hence most of them take on a value of one if the variable applies to the household or individual surveyed, and zero otherwise. It should be noted, however, that sampling weights are not accounted for in the computation of the descriptive statistics. Meanwhile, sampling weights are applied in the variation of the variables across population density quartiles Tables D1 and D2.

3.3.1 Population density

The average population density that we extract from the GHS-POP for more than 700,000 households covered in this study appears to have almost doubled from 1975 to 2015. This is consistent with the data from the United Nations (2019) which showed that the population density of most countries in our sample increased by more than twofold during the 40-year period.

3.3.2 Indicators for urban benefits and costs

Given the wealth of data at our disposal, we employ indicators and proxies for a wide array of urban benefits such as household wealth, hygiene and sanitation, and access to utilities, health services, clean energy and private goods and services; and urban costs such as indoor and outdoor pollution, space congestion, diseases, and crime. We summarize in Table 1 and discuss in Appendix E how our selection of variables represents or proxies for the urban benefits and costs that we cover in this study.

3.3.3 Control variables

We also employ a broad set of control variables consisting mainly of household and individual characteristics, and first-nature characteristics, as indicated in Table C3. It is worth noting that a woman's age ranges from 15 to 49 which reflects the age range of eligible women (i.e. women of reproductive age) for the DHS. The woman's questionnaire forms the central part of the DHS questionnaires and covers all topics of the survey. On the other hand, children who are eligible for the DHS are those aged below five. Hence, a child's minimum age in years is 0 while the maximum is 4. Meanwhile, household locations display huge disparity in terms of geographical attributes, as seen from the wide range of values for distance to national borders and to bodies of water, amount of annual rainfall, and volatility of temperature.

Table 1. Indicators and proxy variables for urban benefits and costs

	Indicator/proxy	Description/unit of measure	Basis for choice of indicator/proxy	Data source
Urban benefits				
Household wealth	Wealth index	Composite index of a household's ownership of selected assets (e.g. televisions and bicycles), materials used for housing construction, and types of water access and sanitation facilities	Ratledge et al. (2021) Rustein and Johnson (2004)	DHS
Access to public utilities	Whether a household has electricity Whether a household has access to safe drinking water	Binary	Croft et al. (2008)	DHS
Hygiene and sanitation	Whether a household has access to improved toilet facilities Whether a household has handwashing place with soap and water	Binary	Henderson et al. (2019)	DHS
Accessibility to private goods and services	Inverse of travel time to reach high-density urban center	Minutes	Ahlfeldt and Pietrostefani (2019) Duranton and Puga (2020) Overman and Venables (2005)	Mayala et al. (2018)
Access to clean energy	Whether a household uses clean fuel for cooking	Binary	Borck and Schrauth (2021) Duranton and Puga (2020)	DHS
Access to health services	Child vaccination: Whether a child has received diphtheria vaccine (third dose) Whether a child has received measles vaccine (first dose)	Binary	Henderson and Turner (2020)	DHS
Urban costs				
Indoor pollution	Exposure to indoor smoke: Whether a household uses solid fuels for cooking Whether a household member/s is/are smoking tobacco on a daily, weekly or monthly basis	Binary	Gollin et al. (2021)	DHS
Outdoor pollution	PM2.5 concentration	Micrograms per cubic meter	Borck and Schrauth (2021) Gollin et al. (2021)	van Donkelaar et al. (2018)
	NO_2 concentration	Parts per billion		Geddes et al. (2017)
Space congestion	Ratio of no. of household members to no. of sleeping rooms	No. of persons	Ahlfeldt and Pietrostefani (2019) Overman and Venables (2005)	DHS
Prevalence of diseases	Child diseases (whether a child had any of the following in the last 24 hours or in the last two weeks): Fever Cough Diarrhea	Binary	Henderson et al. (2019)	DHS
Crime	Violence against women (whether a woman has experienced any of the following): Physical violence Sexual violence Emotional violence Justified beating by spouse /partner	Binary	Glaeser (2020)	DHS

Source: Authors' compilation.

3.4 Variables across density quartiles

Tables D1 and D2 take into account survey sample weights and show how the variables employed in this study vary across population density quartiles. Four interesting patterns stand out. First, all indicators and proxies for urban benefits except for the presence of place for handwashing, rise progressively with population density quartile. Regarding urban costs, only outdoor pollution and incidence of child fever increase across density quartiles. Second, the proportion of household head, women, and mothers with secondary education and higher, rises with population density. Third, the fraction of households located in the national capital likewise increases across density quartiles. Fourth, households whose locations are closest to national borders and bodies of water, and have the most annual rainfall and least temperature volatility, belong to the highest population density quartile. We look into these patterns more closely in the next section.

4. EMPIRICAL FRAMEWORK

4.1 Identification strategy

The conceptual framework in Section 2 underscores the need to estimate the causal impact of density on urban benefits and costs in order to determine if density brings greater benefits or costs and consequently, to evaluate whether restricting urbanization is warranted or not. Nevertheless, the causal relationship between density and urban benefits and costs is beset by a host of identification issues (Bryan et al. 2020; Combes and Gobillon 2015; Duranton 2015; Duranton and Puga 2020). The patterns that we highlight in the Section 3.4 help illustrate the identification issues that confound the causal effect of density on urban costs and benefits. Foremost of these identification issues is reverse causality. The strong positive association between population density and indicators and proxies for urban benefits shown in Table D1 may imply that higher density indeed brings greater urban benefits. However, it is also possible that the numerous urban benefits that rise with density in the previous section, such as better access to basic utilities and greater availability of public services, have attracted more people to reside in an area.

Another identification concern relates to omitted variables that influence both density and an urban benefit or cost. For instance, places with natural endowments such as favorable weather and proximity to bodies of water are more likely to generate greater household wealth and, at the same time, draw higher population. Still, there are unobserved factors that affect density and urban outcomes simultaneously. Examples include local factors such as cultural heritage and social life, and individual traits such as ambition and motivation (Combes and Gobillon 2015; Henderson and Turner 2020).

Finally, the sorting or self-selection of more able individuals into denser and wealthier areas likewise complicate the causal relationship between density and urban costs and benefits. Greater household wealth and better access to various public and private goods and services in higher density quartiles, as seen in the previous section, could reflect the presence of more educated people which also rises with density in our sample. These individuals have higher potential to accumulate wealth and greater influence to demand better goods and services (Henderson et al. 2020).

As prescribed in previous studies, we address the issues of reverse causality and omitted variables via instrumentation (Combes et al. 2010; Combes and Gobillon 2015; Duranton 2015; Duranton and Puga 2020). Meanwhile, the literature recommends controlling for individual characteristics, or whenever possible imposing fixed worker effects to deal with the issue of sorting (Duranton 2015). As noted by Combes and Gobillon (2015), panel data is not usually available for developing countries and it may not be feasible to account for unobserved individual heterogeneity. Hence, we utilize a wide array of household and individual characteristics as control variables to address the sorting of more able individuals into denser areas. We also recognize that some urban benefits, specifically public goods and services, are made more available in cities due to lower cost of providing them in denser areas along with political considerations (Henderson and Turner 2020). We account for this urban bias by controlling for the location of the household, i.e. whether it is in a national capital or not. These strategies are discussed in greater detail in Appendix F.

4.2 Estimation strategy

Given the identification strategy above, we estimate two sets of regressions. The first set consists of baseline estimations, while the second set pertains to instrumental variable (IV) estimations.

The baseline model takes the following form:

$$y_i = \alpha log(D_i) + X_i\beta + U_i + e_i \tag{1}$$

where y is the urban benefit or cost of interest, D is 2015 population density, X is the vector of household and individual controls, e is the error term, and i pertains to household or individuals. We consider U as a vector that incorporates all unobserved components that also influence an urban benefit or cost.

Meanwhile, the IV model is represented by the following system of equations:

$$log(D_i) = Z_i\zeta + X_i\theta + \omega_i \tag{2}$$

$$y_i = \gamma log\left(\widehat{D_i}\right) + X_i\delta + \varepsilon_i \tag{3}$$

where Equation 2 is the first-stage regression, and Equation 3 is the second-stage regression, Z is a vector of instruments (in this paper, the main instrument is the log of 1975 population density), $\widehat{D}$ are the fitted values of D derived from Equation 2, while ω and ε are the corresponding error terms for the first- and second-stage equations.

As seen in Table C1, the indicators and proxies for urban benefits and costs that we cover in this paper consists of both binary and continuous variables. To recall, binary variables relate to access to public utilities, hygiene and sanitation, access to clean energy, child inoculations (as proxy for health services), indoor pollution, child diseases (as proxy for overall prevalence of diseases), violence against women (as proxy for crime). Meanwhile, continuous variables consist of wealth index, inverse of travel time to reach high-density urban center (as proxy for access to private goods and services), outdoor pollution, and ratio of number of household members to number of sleeping rooms (as proxy for space congestion). We employ linear regression for continuous variables, and probit regression for dichotomous variables which I solve via least squares and maximum likelihood estimations, respectively.

Moreover, we estimate the baseline model under three specifications: without controls, with controls for sorting (i.e., includes variables representing household and individual characteristics), and with controls for sorting and urban bias (i.e., includes a dummy variable equal to 1 if a household is located in a national capital, zero otherwise). Meanwhile, as discussed in the preceding section, we seek to address the issues of reverse causality and omitted variables through the IV model which we estimate under four specifications: without controls, with controls for sorting, with controls for sorting and urban bias, and with controls sorting, urban bias and first-nature characteristics. The latter set of controls consists of local permanent attributes such as proximity to bodies of water and national borders, average rainfall and temperature volatility, that may influence both my instrument i.e., historical density, and an urban benefit or cost. Meanwhile, we include in all specifications the number of household members to control for household size, and country dummies to account for country heterogeneity. In addition, standard errors are clustered at the level of the primary sampling unit or the survey cluster of the DHS which is my main database. This is in line with the advice of Colin Cameron and Miller (2015) to be conservative and avoid bias by using more aggregate clusters as possible.

The last specification mentioned above i.e., IV estimation with controls for sorting, urban bias and first-nature characteristics, addresses the identification issues that confound the causal relationship between the key independent variable, $log(D_i)$, and the urban benefit or cost of interest, y_i. This permits the interpretation of γ in Equation 3 as the causal impact of population density on an urban benefit or cost. The conceptual framework in Section 2 brings to the fore the question of whether population density confers greater benefits or greater costs. Hence, the sign and significance of γ over the range of urban benefits and costs covered in this paper can provide an indication on whether density brings net benefits or costs. For instance, a significantly positive γ indicates a positive causal effect of population density on an urban benefit or cost. In addition, a relatively greater number of urban benefits relative to costs that have significantly positive γ may suggest that population density generates net urban benefits. Reverting to the conceptual framework, this could refer to areas to the left of D_S^* which implies that there is still scope to increase urban population in order to maximize the social benefits of urbanization. The converse holds in the case of a relatively larger number of urban costs with significantly positive γ.

As the continuous dependent variables in this study are expressed in logarithms (e.g., wealth index and concentrations of PM2.5 and NO_2), the magnitude of γ may be interpreted as the elasticity of an urban benefit or cost with respect to density. For binary dependent variables, we compute the marginal effect from the probit regressions to estimate the change in the probability of an urban benefit or cost being applicable to a household or individual for every percent change in density, while keeping the other explanatory variables at their mean values.

5. RESULTS

Tables 2 and 3 summarize the coefficients of the log of population density in the regressions on indicators and proxies for the urban benefits and costs covered in this study. Columns 1 to 3 of both tables report the results from the baseline estimations while columns 4 through 7 present the outcomes from the IV estimations. Particularly, the specification that corresponds to column 7 addresses the identification issues (such as sorting, reverse causality and omitted variables), and also controls for urban bias and local permanent characteristics that can influence an instrument of density and an urban outcome simultaneously. Hence, the said column shows the causal estimates of the impact of the log of population density on an urban benefit or cost of interest.

Meanwhile, detailed regression results are shown in Tables G1 through G22 in the appendix. The first-stage estimates contained in these tables further confirm the relevance of historical density as instrument for contemporary density. In particular, population density in 1975 is highly and positively correlated with population density in 2015, with an average coefficient of about 0.75 in all regressions. Moreover, the R-squared values in the first-stage linear regressions are also large at around 0.80.

5.1 Urban benefits

A cursory look at Table 2 indicates that the coefficient of the log of population density is highly significant and positive across all specifications, and for all indicators and proxies of urban benefits except for child measles vaccination, one of the proxy variables for access to health services.

As reported in column 1, the raw estimates on the coefficient of log of density (i.e., those generated without controls and instruments employed in the regressions) are positive and significant at 1% level for all indicators and proxies for urban benefits. When household and individual characteristics are controlled for (column 2), the coefficient of log of density remains significant but declines slightly for all variables, except for the first dose of measles vaccination. Meanwhile, the coefficient of log of density on the proxy for accessibility to private goods and services i.e., the log of inverse of travel time to reach urban center remains unchanged at 0.089, while the corresponding coefficient in the probit regression on the first dose of measles vaccination turns insignificant when controls for household and individual characteristics are introduced.

The detailed results for household variables such as the wealth index (Tables G1 and G2), having access to electricity and safe drinking water (Tables G3 and G4), having improved toilet facilities and place for handwashing (Tables G5 and G6), and use of clean fuels for cooking (Table G8) show that the coefficient of household head's education is positive and significant at 1%, with values ranging from 0.107 to 0.827. Similarly, for child indicators that proxy for the availability of health services i.e., having received the third dose of diphtheria vaccine and first dose of measles vaccine, the coefficient of the education of the child's mother, as well as that of the household head, is positive and highly significant (Tables G9 and G10). The strong positive association between an individual's education and favorable urban outcomes lends credence to the sorting of more educated and skilled people into wealthier and more livable places. It also supports the notion that education can be considered as a proxy for an individual's earning potential, as put forth in Appendix F2. As such, more educated individuals are more likely to afford access to public utilities, sanitation facilities, clean energy, and health services. They are also possibly equipped with more knowledge on practicing proper hygiene, environmental awareness, and healthy habits (such as getting required vaccinations).

Meanwhile, as reported in column 3 of Table 2, including additional controls on whether a household is in a national capital retains the significance, but further reduces the magnitude of the coefficient of log of density for indicators of household wealth, access to public utilities and clean energy, hygiene and sanitation, and for one of the proxy variables for access to health services (i.e., third dose of diphtheria vaccine). The detailed regression results for these variables except for having a place for handwashing (see column 3 of Tables G1 to G10), indicate that the coefficient of the national capital dummy is positive and significant, except for having a place for handwashing. This implies that urban bias plausibly holds. The greater provision of public goods and infrastructure in administrative centers, which are high-density areas, can create a conducive environment for the accumulation of wealth, as well as for the availability of a wide range of utilities and private goods and services.

On the other hand, addressing reverse causality and omitted variables via IV estimation likewise maintains the significance, and raises the magnitude of the coefficient of log of density relative to its raw estimates in most indicators of urban benefits. This points to the negative bias in the raw estimates of the coefficient of log density that comes from reverse causality between density and the variables mentioned above, and from unobservable variables that are negatively correlated with density and the said urban benefit indicators. The opposite occurs for the indicator of household wealth (log of wealth index) and the proxy for accessibility to private goods and services (log of inverse of travel time to urban center) where the coefficient of log density, while remaining significant, decline in the IV estimation. This suggests a positive bias in the raw estimates emanating from reverse causality between density, and household wealth and accessibility to private goods and services, as well missing variables that are positively correlated with density and the said urban benefits.

Similar to the baseline estimation results, adding controls for household and individual characteristics to the IV estimations in order to address sorting and urban bias reduces the coefficients of log of density for all variables (columns 5 and 6), except for the first dose of measles vaccination which loses its significance. In columns 5 and 6 of Tables G1 to G10, the education of the household head and the child's mother (as applicable) remain significant for all variables of urban benefits, except for log of inverse of travel time to urban center. This implies that sorting of more able individuals into areas with greater wealth and better amenities likely persists even when the issues of reverse causality and omitted variables are addressed. In addition, the dummy variable on whether a household is in a national capital or not maintains its significance in the IV estimations on all indicators of urban benefits, except for having improved toilet facilities. A probable reason for less improved sanitation facilities in administrative centers in developing countries despite the higher provision of water- and sewer-related infrastructure is the unwillingness of poor households to defray the costs of connecting to sewerage systems (Glaeser 2020).

Finally, accounting further for first-nature characteristics or geographical attributes that can affect both the instrument, population density in 1975, and an urban benefit (see column 7 of Table 10) generally increases the coefficient of log of density across urban benefit variables. On receiving the first dose of measles vaccine, the coefficient of log of density even turned significant when first-nature characteristics are controlled for. A closer scrutiny of the results (column 7 of Tables G1 to G10) points to temperature volatility as the potential source of the negative bias in the coefficient of log of density in column 6 of Table 10. In particular, the log of temperature volatility is significantly negative for most indicators and proxies of urban benefits. This suggests that the negative correlation between temperature volatility and urban benefits such as household wealth, having access to electricity, improved hygiene and sanitation facilities, clean energy, and health services, as well as accessibility to private goods and services.

The results in column 7 of Table 2 account for identification issues such as reverse causality and omitted variables, and address sorting, urban bias and first-nature characteristics that can render our instrument endogenous. Hence, they correspond to the causal estimates of the impact of density on urban benefits covered in this paper, as mentioned previously. Specifically, for continuous indicators of urban benefits, the results suggest that doubling density leads to a 14 percent increase in household wealth and about 2 percent rise in accessibility to private goods and services. Meanwhile, for dichotomous variables that relate to urban benefits, I derive the marginal effects from the probit estimations and summarize them in Table 4. As reported in column 7, the causal estimates imply that doubling density raises the likelihood of having access to electricity and safe drinking water by 0.17, having improved toilet facilities by 0.25, having place for handwashing by 0.13, having access to clean energy by 0.35, and having access to health services by 0.02, while keeping other explanatory variables at their mean values.

5.2 Urban costs

In contrast to urban benefits, the coefficient of log of density in the regressions on indicators and proxies for urban costs, as summarized in Table 3, is either significantly negative or unstable in the specifications. This implies that urban costs covered in this paper either decline or do not change with density. The only exception is outdoor pollution (as measured by PM2.5 and NO_2 concentrations) where the log of density is highly significant and positive under all specifications, indicating that outdoor pollution increases with density.

Going into greater detail, the coefficient of log of density in the probit regressions on the proxy variables for exposure to indoor pollution namely, whether a household uses solid fuels for cooking and whether a household member smokes tobacco inside the home, is negative and significant at 1% level across specifications. This points to a lower probability of being exposed to indoor pollution in denser areas. Controlling for household and individual characteristics shows a negative bias in the raw estimates (column 2). The negative and highly significant coefficient of household head's education in the detailed results (column 2 of Tables G11 and G12) reflects the possible sorting of more able individuals into areas with less exposure to indoor smoke. Households with higher-skilled members are more likely to have the financial capacity to purchase clean fuels for cooking, and to possess greater knowledge on the benefits of using clean fuel and on the detriments of tobacco smoking. For exposure to indoor smoke from cooking, the significantly negative coefficient of the national capital dummy indicates that urban bias plausibly holds. Areas close to or within administrative centers are plausibly equipped with more public goods and infrastructure that promote the use of more advanced and environment friendly cooking methods. Meanwhile, IV estimation results (column 4 of Tables G11 and G12) suggest a positive bias in the raw estimate arising from reverse causality and unobserved variables that are both positively correlated with density and the probability of being exposed to indoor smoke. To determine the magnitude of the causal impact of density on exposure to indoor smoke from cooking, we derive the marginal effect from the probit estimations. As reported in column 7 of Table 4 which summarizes the marginal effects of density on binary indicators and proxies for urban costs, the probability of being exposed to indoor smoke from cooking and indoor tobacco smoke decreases by 38.6 percentage and 2.8 percentage points, respectively when density doubles, while keeping other independent variables at their mean values.

With regard to outdoor pollution as gauged by PM2.5 and NO_2 concentrations, the coefficient of log of density is positive and significant at 1% level under all specifications (Table 3). This suggests that outdoor pollution is greater in higher density areas. For PM2.5 concentrations, a common proxy indicator for air pollution, including household and individual controls (column 2) suggests evidence of sorting of more able individuals to places with better air quality. This is shown by the significantly negative coefficient of household head's education in the detailed regression results (column 2 of Table G13). Nevertheless, the opposite occurs for NO_2 concentration. This implies that exposure to NO_2 concentration is greater for more able individuals. A probable reason is that one of the sources of NO_2 are emissions from cars which are likely to be afforded by higher-skilled individuals. Meanwhile, addressing endogeneity via IV estimation yields higher density elasticity of PM2.5 and NO_2 concentrations (column 4 of Table 3). However, controlling further for sorting, urban bias, and first-nature characteristics lowers the estimated density elasticities. Finally, as reported in Table 5, the marginal effects that we derive indicate that doubling density increases PM2.5 and NO_2 by 8% and 12%, respectively. The density elasticity that we obtain for PM2.5 coincides with magnitude derived by Brock and Schrauth (2021) for Germany.

Meanwhile, the coefficient of log of density in the regressions for congestion of space as proxied by the log of ratio of the number of household members to the number of sleeping rooms in a home is significantly negative in all specifications (Table 3). This runs counter to the expectation that space congestion worsens with population density. Nevertheless, the detailed results in Table G15 point to some indication of sorting as depicted by the negative and highly significant coefficient of household head's education. This implies that more educated individuals are more likely to afford larger living spaces.

For the rest of the urban costs that we cover in this paper such as prevalence of diseases (as proxied by child diseases) and crime (as proxied by violence against women), the coefficient of log of density is either unstable or insignificant across specifications, as seen from Table 3.

For our chosen proxies for crime, the raw estimate of the coefficient of log of density for physical violence against women is significantly negative (column 1). This indicates that the probability having experienced physical violence by a woman is lower at denser areas. However, the coefficient begins to lose its significance when controls for household and individual characteristics are introduced in the baseline and IV estimations. Meanwhile, the coefficient of log of density is insignificant in nearly all specifications for sexual violence against a woman, and unstable for emotional violence against a woman (see Table 3). On the other hand, for justified beating of woman by her spouse or partner, the coefficient of log of density is significantly negative in all specifications except for last one which estimates the causal impact of log of density (column 7). Probing into the detailed results for my proxies for crime shows that the coefficients of the education of the household head and the woman are both significantly negative (Tables G19 to G22), indicating that the probability of a woman having experienced physical violence is lower when the household head or the woman reached secondary education or higher. This points to the possible sorting of more educated people into safer places. At the same time, it is also plausible that more able individuals are equipped with greater means and knowledge on how to protect themselves from crime. Moreover, the detailed results also indicate that the national capital dummy is significantly negative for my proxy variables for crime, pointing to lower incidence of crime in national capitals. Areas that are close to the seat of power are more likely to be provided with public infrastructure and services that safeguard the security of their residents.

In summary, our causal estimates that density seem to bring greater benefits relative to costs in the developing world. Reverting to the conceptual framework in Section 3, the sizeable net benefits brought forth by density correspond to the area to the left of N_S^* or the point where the social benefits of urbanization are maximized. This implies urban population growth may be allowed until the optimal level of the social benefits of urbanization is reached.

Our results also provide evidence of sorting of more able individuals into more livable places as depicted by the strong positive association between an individual's education and urban benefits such as household wealth, and access to utilities, improved hygiene and sanitation, clean energy, and health services; and by the former's significant negative relationship with urban costs such as exposure to indoor and outdoor pollution (as measured by PM2.5 concentration), space congestion, and crime. In addition, we also find evidence of urban bias in the provision of public goods and services as shown by the significant positive coefficient of the national capital dummy in the regressions on urban amenities such as access to utilities, improved hygiene and sanitation, clean energy, and health services; and by its significant negative coefficient in the regressions on urban disamenities such as exposure to indoor pollution coming from home cooking, space congestion, and crime.

5.3 Country heterogeneity analysis

The results that we obtain above apply to our sample of 11 developing Asian countries as a whole. While the sampling weights are adjusted for individual country population to make each country representative at the aggregated level, and country dummies are included in the regressions to account for country heterogeneity, it is worth looking at the causal impact of density on urban benefits and costs for each country. Despite all belonging to low to lower-middle income category of the World Bank, the countries in our sample are diverse in various aspects, such as institutional development, political system, and cultural background, that can influence how density affects each country's urban outcomes. Figure H1 plots the coefficient of log of population density in the regressions on indicators and proxies for urban benefits under the full specification i.e., IV estimation with controls for sorting, urban bias, and first-nature characteristics, for the whole sample and for each country. Hence, the coefficients correspond to the causal impact of density on the urban benefits and costs covered in this study.

In general, the figures show that the coefficient of log of density for Nepal is insignificant in almost all urban benefit and cost variables. It is only significant in the proxies for access to utilities (i.e., having electricity and having access to safe drinking water). Moreover, the error bands of the coefficient of log of density are widest for Nepal compared to other countries in the sample. It is interesting to note that Nepal is the only country in the sample that is classified as low income while the rest are categorized as lower-middle income by the World Bank. The insignificant causal impact of density on urban outcomes (which can be considered as indicators of income levels) in the case of Nepal, may possibly reflect the positive relationship between urbanization and growth as shown in various studies (e.g. Bertinelli and Black 2004; Henderson 2005; Spence et al. 2009). The gains from urbanization may have not yet translated into growth in the case of Nepal. Nevertheless, this hypothesis requires further research.

Meanwhile, individual country coefficients converge to the overall sample results for most indicators and proxies for urban benefits and costs, except for access to clean energy, exposure to indoor pollution due to smoke from cooking, and exposure to outdoor pollution from NO_2 concentration, where country outliers can be seen (Figure H2). Meanwhile, in contrast with result that I obtain for the overall sample, the causal impact of log of density on space congestion is positive and in accord with expectations (i.e., density constrains space) for Bangladesh and the Philippines which are two of the densest countries in our sample.

Table 2. Impact of log(population density) on indicators and proxies of urban benefits

	Model	Estimation method	Baseline specifications			IV specifications			
			No controls	With controls		No controls	With controls		
				Household and individual characteristics	Household and individual characteristics + Located in national capital		Household and individual characteristics	Household and individual characteristics + Located in national capital	Household and individual characteristics + Located in national capital + First-nature characteristics
			(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(wealth index)	Linear regression	LS	0.155***	0.146***	0.140***	0.154***	0.142***	0.136***	0.144***
			(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.004)
Access to public utilities									
Electricity	Probit	ML	0.166***	0.145***	0.139***	0.196***	0.172***	0.166***	0.170***
			(0.007)	(0.007)	(0.007)	(0.011)	(0.011)	(0.012)	(0.007)
Safe drinking water	Probit	ML	0.142***	0.136***	0.132***	0.147***	0.142***	0.136***	0.171***
			(0.007)	(0.007)	(0.007)	(0.010)	(0.010)	(0.010)	(0.007)
Hygiene and sanitation									
Improved toilet facilities	Probit	ML	0.240***	0.218***	0.216***	0.270***	0.245***	0.244***	0.246***
			(0.006)	(0.005)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)
Place for handwashing	Probit	ML	0.151***	0.127***	0.125***	0.163***	0.135***	0.133***	0.134***
			(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)
Log(accessibility to urban center)	Linear regression	LS	0.0889***	0.0891***	0.0770***	0.0845***	0.0843***	0.0695***	0.0243*
			(0.012)	(0.012)	(0.012)	(0.015)	(0.014)	(0.015)	(0.012)
Access to clean energy	Probit	ML	0.366***	0.347***	0.331***	0.398***	0.375***	0.359***	0.353***
			(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.006)
Child vaccination									
Diphtheria (third dose)	Probit	ML	0.0429***	0.0168***	0.0150**	0.0452***	0.0182**	0.0162*	0.0207***
			(0.006)	(0.005)	(0.005)	(0.006)	(0.007)	(0.007)	(0.006)
Measles (first dose)	Probit	ML	0.0242***	0.00759	0.00547	0.0336***	0.0101	0.00762	0.0151*
			(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.006)

LS = Least squares, ML = Maximum likelihood. *** p<0.01, ** p<0.05, * p<0.1.
Notes: Robust standard errors in parentheses, clustered by primary sampling unit (survey cluster). Country dummies are included in all specifications.

Table 3. Impact of log(population density) on indicators and proxies of urban costs

	Model	Estimation method	Baseline specifications No controls	With controls Household and individual characteristics	Household and individual characteristics + Located in national capital	IV specifications No controls	With controls Household and individual characteristics	Household and individual characteristics + Located in national capital	Household and individual characteristics + Located in national capital + First-nature characteristics
			(1)	(2)	(3)	(4)	(5)	(6)	(7)
Indoor pollution									
Smoke from cooking	Probit	ML	-0.387***	-0.370***	-0.355***	-0.425***	-0.404***	-0.390***	-0.386***
			(0.011)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.006)
Tobacco smoke	Probit	ML	-0.0505***	-0.0273***	-0.0290***	-0.0561***	-0.0291***	-0.0314***	-0.0283***
			(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)
Outdoor pollution									
Log(PM2.5 concentration)	Linear regression	LS	0.0835***	0.0848***	0.0647***	0.0855***	0.0870***	0.0644***	0.0787***
			(0.005)	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)	(0.004)
Log(NO_2 concentration)	Linear regression	LS	0.143***	0.141***	0.121***	0.151***	0.149***	0.127***	0.118***
			(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)
Space congestion									
Log(no. of household members/no. of sleeping rooms)	Linear regression	LS	-0.0175***	-0.00650***	-0.00824***	-0.0223***	-0.00932***	-0.0115***	-0.00714***
			(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Child diseases									
Fever	Probit	ML	0.0187**	-0.00376	-0.000880	-0.00732	-0.0131	-0.0101	-0.00817
			(0.006)	(0.005)	(0.005)	(0.007)	(0.007)	(0.008)	(0.005)
Cough	Probit	ML	0.0162*	-0.00491	-0.00197	-0.00789	-0.0130	-0.00990	-0.00574
			(0.007)	(0.005)	(0.005)	(0.007)	(0.007)	(0.008)	(0.005)
Diarrhea	Probit	ML	-0.00224	0.00280	0.00216	-0.00143	-0.00367	-0.00474	-0.0137*
			(0.006)	(0.005)	(0.005)	(0.007)	(0.007)	(0.008)	(0.006)
Violence against women									
Physical violence	Probit	ML	-0.0245***	-0.00322	0.00434	-0.0287***	-0.00488	0.00468	0.000240
			(0.006)	(0.006)	(0.007)	(0.008)	(0.008)	(0.009)	(0.008)
Sexual violence	Probit	ML	-0.0132	0.00213	0.0152	-0.0204*	-0.00280	0.0135	0.000653
			(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)
Emotional violence by husband/partner	Probit	ML	-0.0376***	-0.0216**	-0.00960	-0.0515***	-0.0342***	-0.0204*	-0.0156
			(0.007)	(0.007)	(0.008)	(0.009)	(0.009)	(0.010)	(0.009)
Justified beating by husband/partner	Probit	ML	-0.0522***	-0.0380***	-0.0317***	-0.0605***	-0.0432***	-0.0364***	-0.0101
			(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.005)

LS = Least squares, ML = Maximum likelihood. *** p<0.01, ** p<0.05, * p<0.1.

Notes: Robust standard errors in parentheses, clustered by primary sampling unit (survey cluster). Country dummies are included in all specifications.

Table 4. Marginal effects derived from probit estimations of log(density) on binary indicators and proxies of urban benefits

	Baseline specifications			IV specifications			
	No controls	With controls		No controls	With controls		
		Household and individual characteristics	Household and individual characteristics + Located in national capital		Household and individual characteristics	Household and individual characteristics + Located in national capital	Household and individual characteristics + Located in national capital + First-nature characteristics
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Access to public utilities							
Electricity	0.0327***	0.0275***	0.0263***	0.196***	0.172***	0.166***	0.170***
	(0.001)	(0.001)	(0.001)	(0.011)	(0.011)	(0.012)	(0.007)
Safe drinking water	0.0154***	0.0148***	0.0143***	0.147***	0.142***	0.136***	0.171***
	(0.001)	(0.001)	(0.001)	(0.010)	(0.010)	(0.010)	(0.007)
Hygiene and sanitation							
Improved toilet facilities	0.0767***	0.0652***	0.0646***	0.270***	0.245***	0.244***	0.246***
	(0.002)	(0.001)	(0.001)	(0.007)	(0.007)	(0.007)	(0.006)
Place for handwashing	0.0539***	0.0433***	0.0426***	0.163***	0.135***	0.133***	0.134***
	(0.002)	(0.001)	(0.001)	(0.006)	(0.006)	(0.006)	(0.005)
Access to clean energy	0.114***	0.0992***	0.0938***	0.398***	0.375***	0.359***	0.353***
	(0.002)	(0.002)	(0.002)	(0.010)	(0.010)	(0.010)	(0.006)
Child vaccination							
Diphtheria (third dose)	0.0165***	0.00484***	0.00433**	0.0452***	0.0182**	0.0162*	0.0207***
	(0.002)	(0.001)	(0.001)	(0.006)	(0.007)	(0.007)	(0.006)
Measles (first dose)	0.00947***	0.00206	0.00149	0.0336***	0.0101	0.00762	0.0151*
	(0.002)	(0.001)	(0.001)	(0.005)	(0.006)	(0.007)	(0.006)

*** p<0.01, ** p<0.05, * p<0.1.

Notes: Robust standard errors in parentheses, clustered by primary sampling unit (survey cluster). Country dummies are included in all specifications.

Table 5. Marginal effects derived from probit estimations of log(density) on binary indicators and proxies of urban costs

	Baseline specifications			IV specifications			
	No controls	With controls		No controls	With controls		
		Household and individual characteristics	Household and individual characteristics + Located in national capital		Household and individual characteristics	Household and individual characteristics + Located in national capital	Household and individual characteristics + Located in national capital + First-nature characteristics
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Indoor pollution							
Smoke from cooking	-0.119***	-0.104***	-0.0989***	-0.425***	-0.404***	-0.390***	-0.386***
	(0.002)	(0.002)	(0.002)	(0.011)	(0.010)	(0.010)	(0.006)
Tobacco smoke	-0.0193***	-0.0100***	-0.0106***	-0.0561***	-0.0291***	-0.0314***	-0.0283***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.003)
Child diseases							
Fever	0.00522**	-0.000978	-0.000228	-0.00732	-0.0131	-0.0101	-0.00817
	(0.002)	(0.001)	(0.001)	(0.007)	(0.007)	(0.008)	(0.005)
Cough	0.00440*	-0.00122	-0.000492	-0.00789	-0.0130	-0.00990	-0.00574
	(0.002)	(0.001)	(0.001)	(0.007)	(0.007)	(0.008)	(0.005)
Diarrhea	-0.000403	0.000456	0.000352	-0.00143	-0.00367	-0.00474	-0.0137*
	(0.001)	(0.001)	(0.001)	(0.007)	(0.007)	(0.008)	(0.006)
Violence against women							
Physical violence	-0.00753***	-0.000961	0.00129	-0.0287***	-0.00488	0.00468	0.000240
	(0.002)	(0.002)	(0.002)	(0.008)	(0.008)	(0.009)	(0.008)
Sexual violence	-0.00145	0.000230	0.00163	-0.0204*	-0.00280	0.0135	0.000653
	(0.001)	(0.001)	(0.001)	(0.010)	(0.010)	(0.011)	(0.011)
Emotional violence by husband/partner	-0.00808***	-0.00442**	-0.00196	-0.0515***	-0.0342***	-0.0204*	-0.0156
	(0.002)	(0.002)	(0.002)	(0.009)	(0.009)	(0.010)	(0.009)
Justified beating by husband/partner	-0.0106***	-0.00750***	-0.00625***	-0.0605***	-0.0432***	-0.0364***	-0.0101
	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)	(0.007)	(0.005)

*** p<0.01, ** p<0.05, * p<0.1.

Notes: Robust standard errors in parentheses, clustered by primary sampling unit (survey cluster). Country dummies are included in all specifications.

6. CONCLUSION

Authorities in the developing world have long expressed their aversion to urbanization due to limited resources and capacity to deal with the "grime, crime, and time" costs of urban growth. As a result, governments in low-income countries aim to minimize the downsides of urbanization by slowing migration to urban areas, decongesting dense places by establishing new cities, and making the biggest cities centers for cleaner high-technology activities.

In this paper, we seek to determine whether such fear of urbanization and inclination to curb urban growth in developing countries are warranted or not. We focus on developing Asia where the world's largest proportion of urban population is projected to come from. To do so, we perform a granular and systematic analysis in order to estimate the causal impact of density on a broad set of urban benefits and costs which are most relevant in the developing world context. To carry out our granular analysis, we assemble a rich database that maps geocoded microdata on various measures of living standards, point-specific geographical attributes, and remotely-sensed pollution data onto gridded population data. We then estimate the causal impact of density on my chosen set of urban benefits and costs by addressing the identification issues that complicate the causal relationship between density and urban benefits and costs, such as reverse causality, omitted variables and sorting, through instrumental variable estimation. we also account for possible urban bias, and first-nature characteristics that may render our instrument, historical population density, endogenous.

Our results suggest that density provides greater benefits relative to costs, indicating that urbanization brings net positive gains to the developing world. In particular, the causal estimates that we obtain show that higher population density potentially leads to greater household wealth, improved hygiene and sanitation, and better access to utilities, health services, clean energy, and private goods and services. On the other hand, density appears to either reduce or do not significantly affect urban costs such as exposure indoor pollution, space congestion, diseases, and crime. The only exception is outdoor pollution which seems to deteriorate with population density.

The findings in this paper carry important implications for urban policy in the developing world. The significant net positive gains from density call for allowing rather than restricting urbanization so that the latter's social benefits are maximized. In addition, the urban bias that we detect in the provision of public goods and services implies that the availability of some urban benefits such as utilities, improved hygiene and sanitation, clean energy, and health services may be supply-driven. Moreover, the sorting of more able individuals into more livable areas illustrates the positive relationship between education and favorable urban outcomes. Hence, in lieu of artificially restricting urban growth, authorities in low-income countries may consider providing public services such as adequate water, sanitation, health care, and schools in both rural and urban areas so that people are pulled to urban areas by urban increasing returns, and not pushed out by the lack of essential infrastructure, basic utilities, and educational facilities in rural areas.

Meanwhile, we recognize that the roster of urban benefits and costs covered in this paper far is from exhaustive. They are based on the indicators of urban outcomes that we obtain from the Demographic and Health Surveys (DHS) which are one of the largest nationally representative surveys for the developing world. In addition, the composition of countries in my sample is determined by the availability of GPS coordinates and timeliness of DHS data, albeit countries in this study already comprise about one-third of the estimated urban population in Asia, and around one-fifth of the entire developing world. Hence, future studies may consider other urban outcomes and other developing countries to ascertain whether urbanization indeed deliver net benefits to the developing world.

Moreover, we share the view of Combes et al. (2005) that diagrammatic frameworks, similar to the one that we employ in this paper, do not supersede formal models. This calls for the development of formal models on the trade-off between urban benefits and costs that are most applicable to developing countries. Furthermore, the net positive causal gains from density that we find in this research point to higher utility levels in urban areas. This appears to run counter to the Rosen-Roback notion of spatial equilibrium where people will move between rural and urban locations until they have equalized utility levels in the two areas (Roback 1982; Rosen 1989). To this end, a formal modelling of whether of spatial equilibrium holds in the developing world constitutes an exciting area of future research.

CRediT author statement

Racquel Claveria: Conceptualization, Methodology, Software, Formal Analysis, Data Curation, Writing – Original Draft. **Oscar Claveria:** Writing – Review & Editing, Supervision. **Rosina Moreno:** Writing – Review & Editing, Supervision.

Declarations of interest

None.

Acknowledgements

The authors are grateful to the participants at the University of Barcelona PhD in Economics Workshop 2021, and the 15th Ruhr Graduate School Doctoral Conference in Economics for helpful comments. R. Claveria thanks the PREDOCS-UB scholarship program (which did not review this paper) for financial support.

REFERENCES

Ahlfeldt, Gabriel M., and Laurent Gobillon. 2021. "Introduction to the Special Issue: 'Emerging Trends in Urban Economics.'" *Regional Science and Urban Economics* 90. doi: 10.1016/j.regsciurbeco.2021.103715.

Ahlfeldt, Gabriel M., and Elisabetta Pietrostefani. 2019. "The Economic Effects of Density: A Synthesis." *Journal of Urban Economics* 111. doi: 10.1016/j.jue.2019.04.006.

Bertinelli, Luisito, and Duncan Black. 2004. "Urbanization and Growth." *Journal of Urban Economics* 56(1). doi: 10.1016/j.jue.2004.03.003.

Borck, Rainald, and Philipp Schrauth. 2021. "Population Density and Urban Air Quality." *Regional Science and Urban Economics* 86. doi: 10.1016/j.regsciurbeco.2020.103596.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis. 2020. "Cities in the Developing World." *Annual Review of Economics* 12(1):273–97. doi: 10.1146/annurev-economics-080218-030303.

Burgert, Clara R., Josh Colston, Thea Roy, and Blake Zachary. 2013. *Geographic Displacement Procedure and Georeferenced Data Release Policy for the Demographic and Health Surveys*. Calverton, Maryland, USA: ICF International.

Colin Cameron, A., and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster- Robust Inference." *Journal of Human Resources* 50(2). doi: 10.3368/jhr.50.2.317.

Combes, Pierre Philippe, Gilles Duranton, and Henry G. Overman. 2005. "Agglomeration and the Adjustment of the Spatial Economy." *Papers in Regional Science* 84(3). doi: 10.1111/j.1435-5957.2005.00038.x.

Combes, Pierre-Philippe, and Laurent Gobillon. 2015. "Chapter 5 - The Empirics of Agglomeration Economies." Pp. 247–348 in *Handbook of Regional and Urban Economics*. Vol. 5, edited by G. Duranton, J. V. Henderson, and W. C. Strange. Elsevier. https://doi.org/10.1016/B978-0-444-59517-1.00005-2.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2019. "The Costs of Agglomeration: House and Land Prices in French Cities." *The Review of Economic Studies* 86(4):1556–89. doi: 10.1093/restud/rdy063.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. 2012. "The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection." *Econometrica* 80(6):2543–94. doi: https://doi.org/10.3982/ECTA8442.

Croft, Trevor, Aileen Marshall, Courtney Allen, et al. 2018. *Guide to DHS Statistics*. Rockville, Maryland, USA: ICF.

de Weerdt, Joachim, Luc Christiaensen, and Ravi Kanbur. 2021. *When Distance Drives Destination, Towns Can Stimulate Development*. The World Bank. https://doi.org/10.1596/1813-9450-9622.

Duranton, Gilles. 2008. “Viewpoint: From Cities to Productivity and Growth in Developing Countries.” *Canadian Journal of Economics* 41(3): 689-736. https://doi.org/10.1111/j.1540-5982.2008.00482.x.

Duranton, Gilles. 2015. “Growing through Cities in Developing Countries.” *World Bank Research Observer* 30(1). doi: 10.1093/wbro/lku006.

Duranton, Gilles, and Diego Puga. 2004a. “Chapter 48 - Micro-Foundations of Urban Agglomeration Economies.” Pp. 2063–2117 in *Handbook of Regional and Urban Economics*. Vol. 4, edited by J. V. Henderson and J.-F. Thisse. Elsevier. https://doi.org/10.1016/S1574-0080(04)80005-1.

Duranton, Gilles, and Diego Puga. 2020. "The Economics of Urban Density." *Journal of Economic Perspectives* 34(3):3-26. doi: 10.1257/jep.34.3.3.

Duranton, Gilles, and Stuart S. Rosenthal. 2021. “Special Issue on Delineation of Urban Areas.” *Journal of Urban Economics* 125. https://doi.org/10.1016/j.jue.2021.103352

[dataset] Florczyk, Aneta, Christina Corbane, Daniele Ehrlich, Sergio Freire, Thomas Kemper, Luca Maffenini, Michele Melchiorri, Martino Pesaresi, Panagiotis Politis, Marcello Schiavina, Filip Sabo, and Luigi Zanchetta. 2019. GHSL data package 2019 EUR 29788 EN. Publications Office of the European Union, Luxembourg. ISBN 978-92-76-13187-8 doi:10.2760/0726, JRC 117104.

Fujita, Masahisa, and Jacques-Francois Thisse. 2002. *Economics of Agglomeration*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511805660.

[dataset] Geddes, Jeffrey, Randall Martin, Brian Boys, and Aaron van Donkelaar. 2017. Global 3-year running mean ground-level nitrogen dioxide (NO2) grids from GOME, SCIAMACHY and GOME-2. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4JW8BTT. Accessed 21 February 2021.

Glaeser, Edward L., and J. Vernon Henderson. 2017. “Urban Economics for the Developing World: An Introduction.” *Journal of Urban Economics* 98. doi: 10.1016/j.jue.2017.01.003.

Glaeser, Edward L. 2020. “Urbanization and Its Discontents.” *Eastern Economic Journal* 46(2). doi: 10.1057/s41302-020-00167-3.

Glaeser, Edward L., and Joshua D. Gottlieb. 2009. “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States.” *Journal of Economic Literature* 47(4). doi: 10.1257/jel.47.4.983.

Glaeser, Edward L., and Wentao Xiong. 2017. “Urban Productivity in the Developing World.” *Oxford Review of Economic Policy* 33(3):373–404. doi: 10.1093/oxrep/grx034.

Gollin, Douglas, Martina Kirchberger, and David Lagakos. 2021. “Do Urban Wage Premia Reflect Lower Amenities? Evidence from Africa.” *Journal of Urban Economics* 121. doi: 10.1016/j.jue.2020.103301.

Henderson, J. Vernon. 1974. “The Sizes and Types of Cities.” *American Economic Review* 64(4):640-656.

Henderson, J. Vernon. 2002. “Urbanization in Developing Countries.” *The World Bank Research Observer* 17(1):89–112. doi: 10.1093/wbro/17.1.89.

Henderson, J. Vernon. 2005. “Chapter 24 - Urbanization and Growth.” Pp. 1543–91 in *Handbook of Economic Growth*. Vol. 1, edited by P. Aghion and S. N. Durlauf. Elsevier. https://doi.org/10.1016/S1574-0684(05)01024-5.

Henderson, J. Vernon, Vivian Liu, Cong Peng, and Adam Storeygard. 2020. *Demographic and health outcomes by degree of urbanization: perspectives from a new classification of urban areas.* Luxembourg: European Commission.

Henderson, J. Vernon, and Matthew A. Turner. 2020. “Urbanization in the Developing World: Too Early or Too Slow?” *Journal of Economic Perspectives* 34(3):150–73. doi: 10.1257/jep.34.3.150.

Marshall, Alfred. 1890. *Principles of Economics (8th ed.)*. London: Macmillan.

[Dataset] Mayala, Benjamin, and, David Eitelberg, and Trinadh Dontamsetti. 2018. *The DHS Program Geospatial Covariate Datasets Manual (Second Edition)*. Rockville, Maryland.

Melo, Patricia C., Daniel J. Graham, and Robert B. Noland. 2009. “A Meta-Analysis of Estimates of Urban Agglomeration Economies.” *Regional Science and Urban Economics* 39(3). doi: 10.1016/j.regsciurbeco.2008.12.002.

Moretti, Enrico, 2011. "Local Labor Markets." in: O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics,* edition 1, volume 4, chapter 14, pages 1237-1313, Elsevier. https://doi.org/10.1016/S0169-7218(11)02412-9

OECD/European Commission. 2020. *Cities in the world: a new perspective on urbanization.* OECD Urban Studies, Paris: OECD Publishing. https://doi.org/10.1787/d0efcbda-en.

Overman, Henry G., and Anthony J. Venables. 2005. "Cities in the Developing World." CEPDP (695). Centre for Economic Performance, London School of Economics and Political Science, London, UK. ISBN 0753018756.

Perez-Haydrich, Carolina, Joshua L. Warren, Clara R. Burgert, and Michael E. Emch. 2013. *Guidelines on the Use of DHS GPS Data*. Calverton, Maryland, USA: ICF International.

Puga, Diego. 2010. "The Magnitude and Causes of Agglomeration Economies." *Journal of Regional Science* 50(1). doi: 10.1111/j.1467-9787.2009.00657.x.

Ratledge, Nathan, Gabriel Cadamuro, Brandon De la Cuesta, Matthieu Stigler, and Marshall Burke. 2021. "Using satellite imagery and machine learning to estimate the livelihood impact of electricity access." National Bureau of Economic Research Working Paper 29237 September. doi: 10.3386/w29237.

Rosenthal, Stuart S, and William C. Strange. 2004. "Chapter 49 - Evidence on the Nature and Sources of Agglomeration Economies." Pp. 2119–71 in *Handbook of Regional and Urban Economics*. Vol. 4, edited by J. V. Henderson and J.-F. Thisse. Elsevier. https://doi.org/10.1016/S1574-0080(04)80006-3.

Rutstein, Shea, and Kiersten Johnson. 2004. *The DHS wealth index*. DHS Comparative Reports No. 6. Calverton, Maryland: ORC Macro.

[dataset] Schiavina, Marcello, Sergio Freire, and Kytt MacManus. 2019. GHS population grid multitemporal (1975-1990-2000-2015), R2019A. European Commission, Joint Research Centre (JRC) [Dataset] doi:10.2905/0C6B9751-A71F-4062-830B-43C9F432370F PID: http://data.europa.eu/89h/0c6b9751-a71f-4062-830b-43c9f432370f

Spence, Michael, Patricia Clarke, Annez, and Robert M. Buckley. 2009. *Urbanization and Growth*. Commission on Growth and Development. World Bank. © World Bank. https://openknowledge.worldbank.org/handle/10986/2582 License: CC BY 3.0 IGO.

The Economist. 2021a. "Economically, Covid-19 has hit hard-up urbanites hardest." 19 June. https://www.economist.com/international/2021/06/19/economically-covid-19-has-hit-hard-up-urbanites-hardest.

The Economist. 2021b. "Millions of lives depend on how the pandemic ends." 16 October. https://www.economist.com/leaders/2021/10/16/millions-of-lives-depend-on-how-the-pandemic-ends

[dataset] United Nations, Department of Economic and Social Affairs, Population Division. 2018. *World Urbanization Prospects: The 2018 Revision*, Online Edition. Accessed 27 April 2021.

United Nations, Department of Economic and Social Affairs, Population Division. 2019a. *World Urbanization Prospects 2018: Highlights* (ST/ESA/SER.A/421). New York: United Nations.

United Nations, Department of Economic and Social Affairs, Population Division. 2019b. *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. New York: United Nations.

[dataset] van Donkelaar, A., R. V. Martin, M. Brauer, N. C. Hsu, R. A. Kahn, R. C. Levy, A. Lyapustin, A. M. Sayer, and D. M. Winker. 2018. Global annual PM2.5 grids from MODIS, MISR and SeaWiFS aerosol optical depth (AOD) with GWR, 1998-2016. Palisades NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4ZK5DQS. Accessed 21 February 2021.

World Bank. 2009. *World Development Report 2009: Reshaping Economic Geography*. World Bank. © World Bank. https://openknowledge.worldbank.org/handle/10986/5991 License: CC BY 3.0 IGO.