Explaining urban density change

Koomen, Eric; Bank, Arjen; van der Wielen, Thijmen; van Vliet, Jasper

2023

Explaining urban density change;
review and global analysis of driving forces

Eric Koomen
Arjen Bank
Thijmen van der Wielen
Jasper van Vliet

August 30, 2023

VU VRIJE UNIVERSITEIT AMSTERDAM

COLOPHON

TITLE

Explaining urban density change; review and global analysis of driving forces
Spinlab Research Memorandum SL-21

AUTHORS

Eric Koomen, Arjen Bank, Thijmen van der Wielen
School of Business and Economics (SBE), Vrije Universiteit Amsterdam.
Jasper van Vliet
Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam.

CONTACT

Vrije Universiteit Amsterdam

Faculty of Economics and Business Administration
Department of Spatial Economics
De Boelelaan 1105
1081 HV Amsterdam
Netherlands
Phone: +31 20 5986095
Email: e.koomen@vu.nl
Website: https://spinlab.vu.nl/

This study was funded by and prepared for PBL Netherlands Environmental Assessment Agency, Den Haag.

Contents

1 Introduction

More than half of the world population now lives in urban areas and this number is on the rise (UN, 2019a). During the 1950s roughly one third of the total population was considered urban, by 2018 this was 55% and it is expected that the urban population share increases to 68% in 2050. The increase from 55% to 68% is equivalent to 2.5 billion new urban residents (UN, 2019b). Although urban land covers less than about 1% of the Earth's land area (this estimate depends on the definition of urban area and the input data used, see amongst others: Liu et al., 2014), its negative impacts on hydrological systems, climate, biogeochemistry, biodiversity, and ecosystem productivity can be substantial (Grimm et al., 2008; Seto et al., 2012). Such impacts are often irreversible as land is highly unlikely to convert into another type of use once urbanised. This differs from other transitions, such as those from natural area to cropland, that are more easily reversed (Seto et al., 2011).

The growth in urban population will result in an expansion of urban areas but estimates of the growth in urban area diverge strongly. A selection of these estimates is compiled by PBL and depicted in Figure 1. The projected total urban for the middle of the road SSP2 and A2 scenarios ranges from 0,6 to 1,9 million km^2 in 2030. These estimates differ for several reasons and an extensive analysis of the underlying causes is beyond the scope of this report. Yet, a few major aspects are likely to be of influence. First, the definitions of urban area and the base data used to represent this may deviate. This is especially apparent when looking at the different estimates for the year 2000 compiled by Potere and Schneider (2007)[1] that range from 0.3 to 0.7 million km^2. These differences largely depend on the resolution of the base data and the definition of urban area. While some studies report all observed built-up area, others are restricted to urban areas only, applying different definitions in terms of, for example, population or area size. The terms urban and built-up area are thus often used interchangeably. In this report we try to be consistent in distinguishing between them and refer to urban area when discussing larger contiguous built-up areas with a minimum size. When discussing other studies, we will, however, stick to the terminology used by the authors. The differences in observed urban area will, of course, influence subsequent simulations of future change. The simulations are, furthermore, determined by the scenario or trend-based assumptions describing future trajectories in terms of, for example, assumed population totals and urbanisation rates (expected shares of urban population) that together determine the total future urban

[1] From their inventory we excluded the GRUMP estimate of 3.5 million km^2 as that is described to clearly overestimate city size (Potere et al., 2009).

population. Finally, the estimates differ in the expected future land take per person that translates urban population change into an urban area demand. This land take, often expressed as its inverse (urban population density in, for example, persons per km^2), varies widely between regions and its future state is highly uncertain (Angel et al., 2011; Li et al., 2022; Seto et al., 2011; Seto et al., 2012). This uncertainty depends on household size, the housing density in which new urban extensions are developed and the degree to which existing urban areas densify by accommodating part of the growth in urban population (Broitman & Koomen, 2015). Angel et al. (2021) show that around 25% of the population added to 200 cities worldwide between 1990 and 2014 was accommodated within their built-up area of 1990. The remaining 75% expanded the city areas with much lower density than the initial cites, thus resulting in globally declining densities. Decreasing densities alone are responsible for 125,000 km^2 of urban land conversion (Güneralp et al., 2020), or around 40% of the total increase in built-up land between 1975 and 2015 (Li et al., 2022). The increasing land take per person may even result in urban expansion in regions where population is declining, as is exemplified in many municipalities of East Germany (Li et al., 2022). So, urban development is not solely linked to (urban) population growth. The causes for urban density change have been studied for many individual cities, but not yet consistently on a global scale (Xu et al., 2020).

This study aims to enrich our understanding of urban expansion at a global scale and focusses on the importance of changes in urban density to explain urban area development. The presented analysis is intended to support the development of country-specific, future projections of total urban area that can be used in global assessment models. Such projections typically calculate urban area demand as a function of total urban population and an assumed urban land take per person. While population projections are well-established (e.g., Chen et al., 2020), urban density projections have received much less research attention.

This report consists of two main parts. We start by reviewing existing literature to identify the drivers that potentially explain urban density changes at a global scale. This review pays specific attention to availability of data sets that characterise these driving forces and that can be used in explanatory analyses of urban density change at the national level (i.e., at country scale). In the second phase, drivers found in the first phase are applied in an empirical analysis to study their impact on explaining urban density. In this analysis we keep an eye on the potential application of our results in the 2UP-model that simulates urban area development and population change (Koomen et al., 2023; van Huijstee et al., 2018).

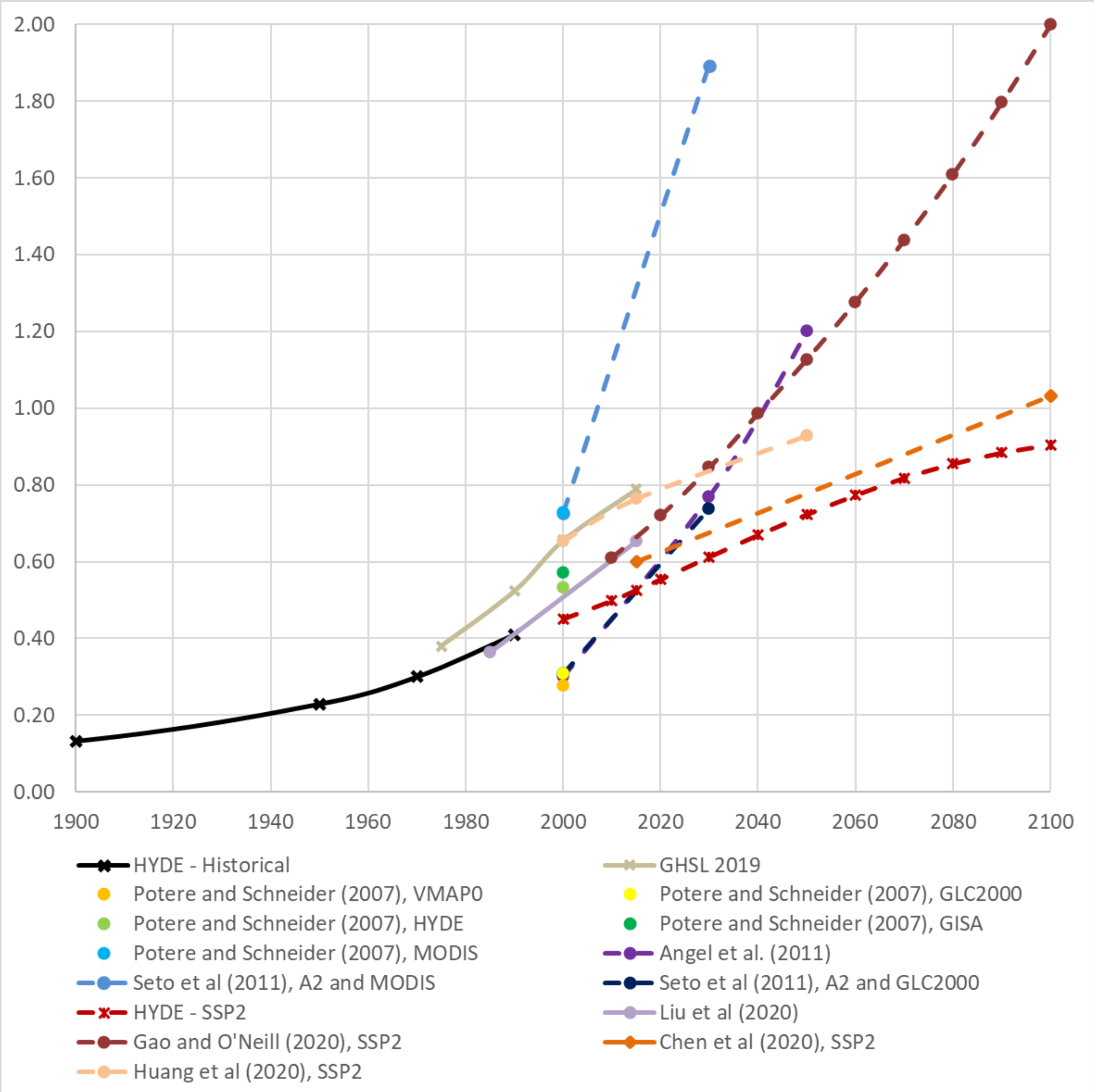

Figure 1 Future global built-up area estimates for the SSP2 scenario in million km² for a selection of global studies (adapted from: PBL, 2021).

2 Identifying drivers of urban density change

In the first phase of this project, studies that explain changes in urban density are examined with a focus on identifying drivers of these changes. Literature is derived from Web of Science[tm] publisher-independent global citation database and selected based on several criteria as discussed in the next section.

2.1 Methods

2.1.1 Selecting papers

Since papers reviewing the forces that drive urban density change at the global scale are absent, we start by selecting more general papers focussed on drivers behind urban land area. From these papers we infer the insights that link to urban density change. To select papers about drivers of urban land area changes, we apply two sets of keywords on Web of Science as indicated in Table 1. The first keywords are chosen to select papers focussing on changes in urban land, following the approach of Colsaet et al. (2018). The second set of keywords filters out articles that do not include causal relationships and articles that only map urban land changes (instead of explaining it).

Table 1 Keywords used to search for articles on Web of Science.

Keyword 1	+ Keyword 2
Urbanisation	Driver
Urban land change	Driving
Urban land take	Factor
Urban land expansion	Determinant
Urban sprawl	Cause
	Explain
	Explanatory
	Influence

From the initial selection we exclude studies on specific cases by only including review papers and meta-analyses that base their conclusions on the insights described in other papers[2]. The resulting subset of papers is further refined by selecting papers that: 1) are published in a journal with a Journal Impact Factor (JIF) greater than 1.0 and a Journal Citation Indicator (JCI) greater than 0.5 (based on journal data of 2021); and 2) were published between 2010 and 2022 to obtain the most recent scientific insights. The selection criteria yielded, several hundreds of articles that were ranked by the number of citations. The 100 most-cited papers were further assessed on their relevance for our analysis. This assessment was done by reading the title and abstract and applying an extra set

[2] This classification is based on the Web of Science 'Review Article' quick filter, to which we added the much-cited meta-analysis by Seto et al. (2011).

of criteria ensuring that: 1) the research must have a global scope (excluding studies focussing on individual continents, countries or cities); 2) the article must explain urban land changes (which means that articles using urban changes to explain other factors are discarded); and 3) the study does not solely model current urban land or make projections of future urban land. Following this approach, four relevant and frequently cited articles were selected. To this collection we add two recent review papers on drivers of urban land change (published in 2020 and 2022) to ensure the most recent insights form literature are included. Our final selection of six papers is included in Table 2.

Table 2 The six identified articles included in this review with their number of citations and references, journal impact factors and citation indicators.

Reference	Title	Nr. of citations / references	JIF/ JCI
Allan et al. (2022)	Driving forces behind land use and land cover change: A systematic and bibliometric review	0 / 127	3.905/ 0.83
Angel et al. (2011)	The dimensions of global urban expansion: Estimates and projections for all countries, 2000-2050	562 / 40	6.063/ 1.40
Colsaet et al. (2018)	What drives land take and urban land expansion? A systematic review	70 / 22	6.189/ 1.37
Kim et al. (2020)	A review of driving factors, scenarios, and topics in urban land change models	11 / 175	3.905/ 0.83
Seto et al. (2010)	The new geography of contemporary urbanization and the environment	325 / 146	17.909/ 1.21
Seto et al. (2011)	A meta-analysis of global urban land expansion	1167 / 22	3.752/ 0.88

Information about the content, spatial and temporal scale, approach and number of articles or observations of the six articles is shown in Table 3. The spatial scale ranges from local (intra-city) to global; the temporal scale ranges from the 1970s to recent times (2022); five different methods are used in the articles; and the number of articles or observations per study vary between 40 and 326. The paper by Angel et al. – albeit being classified as review article by Web of Science – differs from the other papers in its more empirical nature. It is deemed highly relevant, however, as it is the only paper that explicitly address the drivers of urban density for a large number of case study areas.

Table 3 Content of the identified articles: number of observations, spatial scale, temporal scale and approach.

Reference	Nr. of articles / observations	Scale range	Period covered	Methodological approach
Allan et al. (2022)	110	Local/ global	2012-2022	Systematic and bibliometric review
Angel et al. (2011)	142*	National	2000	Multiple regression model
Colsaet et al. (2018)	193	Local/ supra-national	1993-2017	Systematic review
Kim et al. (2020)	164	Local/ global	1974-2020	Systematic review
Seto et al. (2010)	40**	Local/ national	1973-2010****	Non-systematic review
Seto et al. (2011)	326***	Local	1970-2000	Meta-analysis and multiple regression model

*Value refers to the number of observations (countries) used in the multiple regression model.
**Not mentioned but based on references (79-109) in text about non-demographic drivers of contemporary urbanization.
***This value refers to the number of case studies derived from 181 papers used in the meta-analysis.
****Years not mentioned but based on publication years of references (79-109) in text about non-demographic drivers of contemporary urbanization.

2.1.2 Defining main driving forces

In a first step to synthesise the results from the selected papers we made an overview of the main and sub-categories of the forces driving urban land area changes that are listed in each article. This overview is included in Annex 1. Main categories are used here as the broadest term that describes a coherent set of urban land change drivers, referring to drivers such as demography or economic factors. Sub-categories are more specific individual drivers as, for example, population size or GDP per capita. Subsequently, the sub-categories of the individual papers are classified according to the main categories of drivers identified by Colsaet et al. (2018)[3]. The categorisation of this paper is helpful as it is sufficiently broad to encompass most of the individual drivers mentioned in the other papers. The paper distinguishes the following main categories of driving forces: (1) demography, (2) social processes, (3) economy, (4) infrastructure and transport, (5) geographical constraints, (6) path dependency and neighbouring effects and (7) policies and institutions. We adapted the names of some categories slightly, to better characterise the individual drivers listed in the other papers. The term 'social processes' is adjusted to 'socio-cultural factors' because we believe that a process focuses on a change over time, whereas the new name also includes static social states. In addition, 'geographical constraints' is altered to 'geographical configuration' because constraints only set limits to possible new urban areas, but it is also possible that geography enhances the possibility of a location to become urban. The new term includes both the constraints and stimulation of urbanisation.

2.1.3 Grouping individual drivers

In a second step to provide a more concise overview of this categorisation process, we group the individual driving forces of the selected papers under the seven renamed, main categories of driving forces. Some individual drivers are grouped together, because they have similar names and refer to the same process. When grouping individual drivers, we applied specific labels that allow a more direct interpretation in relation to the impact on urban area change. For instance, Allan et al. (2022) have a sub-category 'demographic', but Angel et al. (2011) refer to 'total urban population'. In this case, we include the latter term as it is more specific and allows a more direct interpretation than the former. This has the added value of allowing us to determine whether a driver has a positive or negative impact on urban area density. So, we list the drivers that are found to steer urban land change and characterise their impact on urban density changes.

[3] This categorisation of main and subcategories is visualised in Annex 1 with a colour coding. The colours correspond to the main categories in the paper by Colsaet et al. (2018).

The categorisation of individual driving forces aims to support understanding how they relate to each other and define the overarching processes they represent. While it is possible to debate the exact categorisation we propose here, this exercise is helpful to provide a structured overview of the driving forces that steer urban development and thus should be considered for selection as explanatory factors in the statistical analysis of phase 2. The results of this extensive review are listed in Annex 2 and in this section, we list the most important findings in detail per main category. The subsections motivate the selection of individual driving forces (as included in the annex) and summarise their impact on urban density change. A positive impact (+) means that when the value of a driver increases (a higher number, a larger share etc.), urban density increases too. When the paper does not specify a link between a driver and density, this is indicated as not applicable (n.a.). While interpreting these relationships it is important to note that we do this from a national perspective, i.e., does the state of an individual driver impact density at the country level. Many studies report on the local impact of drivers and, for example, find that higher elevation or larger distance from specific amenities limit the chances of urban development. Such relations are not necessarily relevant or true at a national scale.

2.2 Results

2.2.1 Demography

Demographic changes are an obvious driver for urban development; a growing population requires more land for housing. Yet, their relationship with urban density change is less straightforward. The larger the total population size of an area, the higher the density (Table 4). The same holds true for the size of urban population. However, solely focussing on the growth of population, this shows a mixed or negative impact on density. Colsaet et al (2018) point out that this mixed impact can be the result of a specific location (large city or small village) within a region where the population growth takes place. While larger households may require larger houses, the resulting density in terms of people per hectare is usually higher than for smaller households. Lastly, when a larger share of the total population consists of children, this results in less dense urban regions, whereas a larger share of elderly does not necessarily seem to increase or decrease the density (i.e., this shows mixed results). Regional and internal migration are often mentioned to influence urban development, but this is especially true for specific regions (e.g., the cities that people from the rural hinterland migrate to). For our national-level analysis these are less relevant as we focus on the factors that steer the average urban density in a country.

Table 4 Driver groups for demography, their impact on urban density and references.

Main category	Driver	Impact*	Ref**
Demography	Urban population	+ / + / ±	1/2/3
	Population growth	± / -	3/6
	Total population	+	3
	Household size	+	3
	Proportion of children	-	3
	Proportion of elderly	±	3

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).

2.2.2 *Socio-cultural factors*

Lifestyle and preferences for detached housing and other larger housing types are typical socio-cultural factors limiting average housing densities. We combine these under the label luxurious lifestyle demand, that we define as a preference for larger houses goods and services. Larger shares of people living in informal settlements typically yield high densities and thus less demand for urban area. The time in history of the start of urban planning is known to affect the total amount of urban land, but the impact on density is not described. This links to policies and institutions of course but this driver is included here as it also reflects socio-cultural preferences at large. Lastly, spatial segregation may lead to inhabitants fleeing out of urban cores, thereby decreasing urban density. On the other hand, at least in the United States, larger shares of minorities increase density because land take per person is lower.

Table 5 Driver groups for socio-cultural factors, their impact on urban density and references.

Main category	Driver	Impact*	Ref**
Socio-cultural factors	Luxurious lifestyle demand (spacious houses / second homes / aspiration for detached housing / demand for high-end consumer goods/services)	- / - / n.a.	1/3/5
	Share of population in informal settlements	+	2
	Spatial planning culture	±	3
	Spatial segregation (number of distinctive groups / minorities)	±	3

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).

2.2.3 *Economy*

The drivers belonging to the main category economy are subdivided into the primary, secondary and tertiary sector, attractiveness for multinational firms, market organisation and national income. The reviewed papers highlight the importance of the following aspects of the primary sector: higher agricultural profitability results in less urban land take. A thriving agricultural sector yields higher land rents and will thus be better able to compete with urban functions likely limiting land take and thus increasing density. Yet, its effect on density is not specified in the paper.

The secondary (industry) sector is mentioned in the reviews in relation to, for instance, technological progress, industrial transformation (e.g., as rate of industrialisation), number of factories. Obviously, increasing importance of the secondary sector results in urban expansion. This type of expansion is typically captured in the built-up area maps that describe urban development, but it is not directly linked to population growth. Strong growth in industrial land use may thus result in declining residential densities, so it is important to include reference to changes in this sector when studying density change, although a clear effect is not found in literature.

The tertiary or service sector encompasses a wide range of activities including transport, logistics, retail, financial services, entertainment, tourism. Many of these activities induce urban expansion, but research on the whole of the tertiary sector has mixed or unclear effects on urban density. This will partly relate to the fact that an increase in the service sector may correlate with a decline in the industrial sector and thus a densification in land use. In addition, it is known that when the tourism sector develops, density declines.

The attractiveness for multinationals of an area/country is often studied and therefore, listed as separate driver. Research indicates that more globalisation and foreign direct investment stimulates urban development, because of for example service infrastructural needs of firms. Furthermore, urban expansion correlates with a strong presence of international real estate developers, the effect that these multinational firms have on the real estate market and demands of multinationals. One can think of demands such as a stable electricity and internet network, water, telecommunication facilities, security, etc. It may seem that these drivers decrease density, but it may also attract more people, thereby increasing density. All in all, a density effect of multinationals is not specified.

The fifth set of drivers is market organisation and includes drivers such as market power and incentives, land and housing prices (rents), market failures, speculation and distance to jobs. Market power and incentives drive changes in urban structure by affecting housing development and density. However, a clear directional (positive or negative) impact on density is not indicated. When land and housing prices are high, developers prefer to build there because of site advantages which induces densification. Low land value areas, on the other hand, are known to yield more scattered, low-density urban areas. Market failures, such as imperfect information, speculation, and externalities, enhance urban land take too, but a density effect is unclear. The state of the economy also affects urban land expansion, mostly by its effect on the number of jobs in different sectors. When the economy is performing poorly and unemployment increases, urban growth decreases. Although, such events result in a declining city core population, this does not necessarily affect national urban density, at least, this is not described. All in all, these drivers shape the market and in turn affect urbanisation, but for most drivers, their effect on urbanisation is unclear.

Finally, national income is considered as a separate driver, as it is included in many studies and important to differentiate between development levels. Income is often expressed as Gross Domestic Product (GDP) per capita, which controls for differences in population size between countries. Most studies find a negative relationship between GDP per capita and density meaning that urban land take per capita increases when income increases.

Table 6 Driver groups for economy, their impact on urban denstiy and references.

Main category	Driver	Impact*	Ref**
Economy	Primary sector	n.a.	3
	Secondary sector	n.a. / n.a.	1/3
	Tertiary sector	n.a. / ±	1/3
	Attractiveness for multinationals	n.a. / n.a. / n.a.	1/3/5
	Market organisation	± / n.a. / +	1/3/4
	Income per capita	- / - / ± / n.a.	1/2/3/6

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).

2.2.4 *Infrastructure and transport*

The main category infrastructure and transport differentiates between car usage inducers and costs, public transportation and distance to international transport. Car usage inducers include bridges, traffic service and road density, because these drivers facilitate car usage. These drivers (might) decrease urban density. Higher costs for car usage, e.g., for gasoline or parking prices, on the other hand is reported to increase density. The provision of public transport (e.g., the density of train, metro, light-rail networks, or the proximity to such transport facilities) also has a mainly positive impact on density. Distance to international transport hubs is identified as a separate set of drivers, because cars and public transport are used for shorter distances than airplanes (airport) and cargo ships (harbour), for example. The impact on density is, however, not described in literature.

Table 7 Driver groups for infrastructure and transport, their impact on urban density and references.

Main category	Driver	Impact*	Ref**
Infrastructure and transport	Car usage inducers	n.a. / - / n.a.	1/3/4
	Car usage costs	+ / +	2/3
	Public transportation	n.a. / + / n.a.	1/3/4
	Distance to international transport hubs (airport/harbour)	n.a.	1

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).

2.2.5 Geographical configuration

Geographical configuration refers to five groups of individual drivers: elevation factors, distance to water bodies, flood prone areas availability of resources and arable land per capita. Elevation factors include slope, elevation, and ruggedness. They are shown to negatively influence urban expansion, but their impact on density is not determined. The distance to water bodies includes the distance to many types of water: coast, rivers, lakes, wetlands, and ponds. Most articles mention these as opportunities for trade. On the other hand, water can also pose a danger to human lives and damage urban infrastructure. Maybe because of this ambiguous character, no consistent impact is documented in the reviewed papers. Availability of resources refers to a broad set of natural conditions (e.g., presence of oil and mineral reserves or other natural amenities) that might attract people and thus stimulate urban development. While this may enhance densities, resulting economic activities such as mining also take up space. Therefore, its effects are mixed. The fifth driver is arable land per capita which is described to decrease urban density. When there is more arable land available, this makes it easier to convert a portion of it to urban, because the land is less expensive. Conversely, when little arable land is available, agricultural land rent increases and it is found that densities increase. Finally, it should be pointed out that study areas size has a significant negative impact on urban expansion (i.e., the larger the area, the less urban expansion). Yet, this is not included as a driver, because the article points out that its effect is negligibly small. Furthermore, climate is excluded. Although there are large countries where climate might affect densities within that country, we assume that this will merely change the location of dense and dispersed urban land and not the density for a country as a whole.

Table 8 Driver groups for geographical configuration, their impact on urban density and references.

Main category	Driver	Impact*	Ref**
Geographical configuration	Elevation factors	n.a. / n.a. / n.a.	1/3/4
	Distance to water bodies	n.a. / n.a. / n.a. / n.a.	1/3/4/6
	Flood prone areas	n.a.	1
	Availability of resources (forestry / ecological / oil / minerals / ecosystem services)	n.a. / ± / n.a.	1/3/4
	Arable land per capita	-	2

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).

2.2.6 Path dependency and neighbouring effects

Drivers of this category are split into path dependency and neighbouring effects. Path dependency refers to the phenomenon that future development trajectories to a large extent depend on historic developments. So, the current density in a country is the result of many different factors and these conditions strongly determines which future changes are likely or even possible. This can most easily be incorporated in analysis by referring to prior densities when analysing subsequent changes.

Neighbouring effects refers to explaining urban development from local-level interactions with, for example, neighbouring land use, distance to the city centre or built-up areas in general. When a country has many dense urban areas, it is likely that it will also develop denser urban areas when urban expansion takes place. One could argue that these drivers are only applicable to the local scale and should therefore be excluded in this review, but by averaging distances over countries it is also possible to include this driver in national-level analyses. Every country has, for instance, an average distance from hospitals/medical care. On the other hand, urban growth in neighbouring cities and construction of residential areas are local phenomena and hence, these are not included. For the path dependency it is found that historically dense areas tend to have high densities in the future as well (positive effect), which then results in less urban land take. In addition, Seto et al. (2011) used a 1980s indicator to mark a period in which, globally, more urban land expansion took place than during other decades. So, this also applies to the path dependency. However, no impact on density is specified.

Table 9 Driver groups for path dependency and neighbouring effects, their impact on urban density and references.

Main category	Driver	Impact*	Ref**
Path dependency and neighbouring effects	Path dependency	+ / n.a.	3/6
	Neighbouring effects	n.a. / + / n.a.	1/3/4

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).

2.2.7 Policies and institutions

The final main category of drivers is subdivided into: planning and regulation; fees and subsidies; non-governmental actors; administrative structure. Planning and regulations includes drivers such as planning by governments (e.g., controlling land ownership, urban growth area regulation (e.g., building permit regulation), urban containment, weak, incomplete or unstable planning, low-density zoning and protection of certain areas. Urban containment or urban growth boundaries are known to increase density, whereas low-density zoning or clustering regulations decrease density. Obviously, the density effect is dependent on the types of policies and regulations in place. This is difficult to characterise at the national level and close to impossible to describe consistently for sets of countries.

Fees and subsidies are a second set of drivers that, for example, contain property taxes impact fees and subsidies for, amongst others, agriculture, and urban renewal. Only for property tax, literature discusses its effect on density: it may result in a dwelling size reduction (increase of density) or it may increase urban sprawl when the tax base is the improvement value of a house, for example. So, results are again mixed. For other fees and subsidies, their impact on density is not described.

Third, the administrative structure of the government is important. Drivers in this set include administrative fragmentation, administrative hierarchical strength, or the amount of state-level control over local planning, state-led development, differences in regulation, reliance on local taxes or insufficient support from national governments to municipalities, financial problems of municipalities, the capacity (in terms of human, technical and financial resources) of planning agencies and decentralization. For none of these drivers, an impact on density has been identified.

Finally, non-governmental actors comprise participation of different users such as private enterprises, developers and real estate agencies, pressure of interest groups, corruption, private governments and the voluntary sector. All of them have different interests and it is found that more pressure of different interest groups, exacerbates urban growth. For density, results are mixed for the participation of users: its effect on density totally depends on if and how much land they supply, if and how they finance and design projects and if, how and where they construct large-scale projects. To conclude, its effect is mixed. For the other drivers, a directional impact is lacking, yet.

Table 10 Driver groups for policies and institutions, their impact on urban density and references.

Main category	**Driver**	**Impact***	**Ref****
Policies and institutions	Planning and regulation	± / ± / n.a. / n.a.	1/3/4/5
	Fees and subsidies	n.a. / ± / n.a. / n.a.	1/3/4/6
	Administrative structure	n.a. / n.a. / n.a.	1/3/5
	Non-governmental actors	± / n.a. / n.a.	1/3/5

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).

3 Assessment of relevant data sources

To include the driving forces reviewed in the previous chapter in an empirical analysis of urban density change, we need to collect data sources that document the state of these drivers per country and per year. Therefore, we made an inventory of potentially relevant data sources that can represent them. This review is presented in Table 11 and contains data sets that cover the full globe. In practice this means they should refer to the national level and be available for at least 100 major countries following the World Bank countries list[4]. In addition, the data needs to document changes over time and preferably be available for several decades. Ideally from 1975 onwards as that is also the year for which the first global land cover data are available.

The table lists the drivers identified in the previous section and describes indicators that are available to operationalise them and the units in which they are expressed. We only included the drivers for which data is available. So, for example, housing preferences, spatial planning culture, and market organisation were excluded as no global data sets on these drivers could be found. Path dependency and neighbouring effects are excluded because the review indicates there are only of importance on a local/regional scale and not on the national level. In addition, we excluded climate because its impact is inconclusive. For some drivers we indicate several, alternative indicators.

The table, furthermore, lists the number of included countries. Note that many indicators have country value 217 (the total number of countries identified by the World Bank). The column 'initial year' indicates the first year for which observations are available. It should be noted that this is not necessarily the year that all countries have their first observation. The column 'final year' shows the last year an observation is made. Here too, it may be the case that this is not the year that all countries have their last observation, i.e., the last observation for certain countries can be further back in time. The column 'time interval' specifies how frequently data is updated. Ideally, this would be yearly, thus having an interval value of 1, to allow the best match with the years for which land cover and population data are available. The final column shows the source of the dataset with number that is further detailed in Annex 3.

[4] https://data.worldbank.org/country

Table 11 Operationalization of the identified drivers indicating maximum number of included countries, period covered and time interval (in years) for which data is available. Full sources are listed in Appendix 3.

Driver*	Indicator (unit)	Nr of countries	Initial year	Final year	Time interval	Source
Demography						
Urban population (increase)	Urban population at mid-year (thousands)	217	1950	2050	5 years	1
Population growth	Total population at mid-year (thousands)	217	1950	2050	5 years	1
	Population change (thousands)	236	1950	2021	1 year	2
Total population	Total population at mid-year (thousands)	217	1950	2050	5 years	1
	Total population as of 1 January (thousands)	236	1950	2021	1 year	2
Household size	Household size (number of members)	171	1960	2016	Inconsistent*	3
Proportion of children	Percentage of total population by select age group, both sexes combined 0-17	236	1950	2021	1 year	4
Proportion of children	Percentage of total population by select age group, both sexes combined 0-17	236	1950	2021	1 year	4
Socio-cultural factors						
Luxurious lifestyle demand	Households and NPISHs consumption expenditure per capita (constant 2015 US$)	185	1960		2021	5
Share of population in informal settlements	Average of share unimproved drinking water and sanitation service (percent of population)	217	2000	2020	1 year	6
Spatial planning culture	No data set found	-	-	-	-	-
Spatial segregation	Ethnic groups (number)	217	1972	2022	No time series	7
Economy						
Primary sector	Share of agriculture in total GDP (%)	214	2010	2020	10 years	8
	Share of agriculture in total GDP (%)	217	2000	2017	No time series	9
Secondary sector	Share of industry/manufacturing in total GDP (%)	214	2010	2020	10 years	8
	Share of industry in total GDP (%)	217	2000	2017	No time series	9
Tertiary sector	Share of services in total GDP (%)	214	2010	2020	10 years	8
	Share of services in total GDP (%)	217	2000	2017	No time series	9
Attractivity for multinationals	Foreign direct investment, net inflows (million current US$)	217	1970	2020	1 year	10
	Inward direct investment positions (million US$)	129	2009	2021	Inconsistent*	11
	Global attractiveness index (GAI score)	148	2016	2022	1 year	12
Market organisation	Primary production intensity (current US$)	207	1970	2022	1 year	13
Income per capita	GDP per capita (current US$)	217	1960	2021	1 year	14
Infrastructure and transport						
Car usage inducers	Road density (km/km^2)	217	1999	2022	No time series	15
	Road density; GRIP (km/km^2)	217	2010	2016	No time series	16
Car usage costs	Fuel price (US cents/litre gasoline)	176	2018	2018	No time series	17
Public transportation	Railway density (km/km^2)	137	2008	2022	No time series	18
Distance to international hubs	Airports density (number/km^2)	217	2021	2021	No time series	19
Geographical configuration						
Elevation factors	Mean elevation (m)	217	2021	2021	Static feature	20
	Mean slope (%)	217	2012	2012	Static feature	21
	Terrain Ruggedness Index (elevation differences)	217	2012	2012	Static feature	21
Distance to water bodies	Median population distance to water (km)	159	2011	2011	Static feature	22
Flood prone areas	Inundation (probability)	217	2017	2017	Static feature	23
	Floodability (index)	217	2023	2023	Static feature	24
	Annual average population affected by river floods (number of people)	163	2015	2015	Static feature	25
	Share of population exposed to high flood risk (%)	188	2022	2022	Static feature	26
Availability of resources	Resources of commercial importance (number)	217	2021	2021	No time series	27
	Arable land (thousand ha)	217	1961	2020	1 year	28
	Fertile soil (percent of total land)	217	2012	2012	No time series	29

Path dependency and neighbouring effects						
	City density (number of cities/km^2)	217	2011	2011	Static feature	30
	University density (number of universities/km^2)	185	2023	2023	Static feature	31
	Hospital density (number of hospitals/km^2)	137	2013	2013	Static feature	32
	Population density (people/km^2)	215	1961	2020	1 year	33
Policies and institutions						
Planning and regulation	Government effectiveness index (-)	212	1996	2021	Inconsistent*	34
	Indicators of regulatory governance (-)	186	2017	2018	No time series	35
	Political stability index (-)	214	1996	2021	Inconsistent*	36
Fees and subsidies	Government expenditure on agriculture, forestry and fishing (% of total expenditure)	148	2001	2019	Inconsistent*	37
	Subsidies and other transfers (% of total expense)	169	1972	2020	Inconsistent*	38
Non-governmental actors	NGO's (number)	217	2022	2022	No time series	39
	Firms (number)	123	2006	2020	Inconsistent*	40
	Global 500 firms (% of 500)	217	1995	2022	Inconsistent*	41
Administrative structure	Administrative divisions (number)	217	2021	2021	No time series	42

* Reported inconsistently over time and between countries. So, time intervals are irregular and differ per country. For many countries observations before 2000 are lacking.

The table shows that only a few indicators are available from 1975 onwards. Many datasets are only available from a much later starting year and/or have only one observation. So, for many drivers we have a good understanding of the current situation but a limited view on past developments. This implies that they are not very useful for our statistical analysis focussed on explaining how densities change over time.

4 Empirical analysis

4.1 Methods

As we are interested in understanding changes in urban density over time, we have set up a time series representing the vast majority of countries[5] around the globe for which we describe urban density and several explanatory variables. In this case the observations of the dependent variable (urban density) are related across time because each individual country is included multiple times. This makes it inappropriate to apply regular (OLS) regression analysis. Instead, we use a panel regression approach with country-fixed effects:

$$D_{ct} = \beta X_{ct} + F_c + \varepsilon_{ct}$$

In which D_{ct} is the urban density for country c and year t, β represents the set of coefficients that describe the impact of the explanatory variables X_{ct} that have observations for each country and year, F_c is a fixed effect per country and ε_{ct} the error term that captures the residual variation per country and year. This setup fits our purpose as we are not interested in explaining the variation in density levels between countries but want to know what is driving changes in density. So, our focus is on understanding what is affecting changes in density levels over time and not so much on replicating how differences between countries arose.

The presented regressions are intended to be applicable in PBL's 2UP model and we use the total urban area, total urban population, total land area and total population included in that model (for more details, see: Koomen et al., 2023). This implies that we inherit some operational decisions. First, we rely on the global built-up area and population data distribution as provided by the European Commission's-Joint Research Centre data. This so-called Global Human Settlement Layer (GHSL) offers a fairly long and consistent time series capturing four moments in time: 1975, 1990, 2000 and 2015 (Florczyk et al., 2019). While this data set is known to contain inaccuracies (see, e.g., Kuffer et

[5] We focus on 134 countries with a land area larger than approximately 1000 km², having more than around 100 km² urban area in 2015. Together they account for 98% of the world's urban area. Excluding the remaining small island and city states, prevents skewing country-based results. To verify this, we also report regression results using alternative restrictions on urban area in Annex 4. In addition, we lack one or more observations for the economic variables for 32 data-poor or newly formed countries: Afghanistan, Angola, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Cambodia, Croatia, Czech Republic, Ethiopia, Guinea, Haiti, Iraq, Kazakhstan, Kyrgyzstan, Lebanon, Libya, Lithuania, Macedonia, Poland, Qatar, Russia, Serbia, Slovakia, Slovenia, Tajikistan, Tanzania, Uganda, Ukraine, Uzbekistan, Vietnam, Yemen. These countries are included in the basic regression models but may be excluded (for one or more years) when including the economic variables. We complemented missing values by linear interpolation or - when impossible - by taking the nearest available value (if less than 5 years apart) for 5 countries: Afghanistan, Bosnia and Herzegovina, Cambodia, Haiti, and Kuwait.

al., 2022), we consider this the best possible data set for our analysis. Second, we follow the definition of urban area included in 2UP: any 30 x 30 arc seconds grid cell where the percentage built up of the total land area is 50% or higher. These grid cells are rectangles of approximately 920x920 metres at the equator, but they become smaller and elongated towards the poles. So, our urban areas are clusters of built-up area pixels in the original GHSL data set (with a 38-metres resolution) that cover at least around 0.5 km^2 within a larger grid cell of circa 1 km^2. Smaller clusters are considered to represent non-urban areas. All population within the urban cells is counted as urban population, while the population in the remaining grid cells qualifies as non-urban population. This approach has the advantage that we do not follow administrative definitions of urban areas that may differ per country. As any distinction, however, this imposes an artificial threshold on an essentially continuous phenomenon. Using built-up area clusters with a minimum size of around 0.5 km^2 in our analysis, implies that our results refer to all but the smaller built-up areas and human settlements. This representation underestimates the total global built-up area by about 40% (Koomen et al., 2023)[6]. Locally these may represent even larger shares of the built-up area (for a more extensive discussion, see: van Vliet et al., 2019).

In addition to the (urban) area and population data originating from the GHSL-source, our regressions include a few economic variables that we can match to the four years observed in the GHSL dataset (1975, 1990, 2000 and 2015). As noted in the previous chapter, only a limited number of datasets fully capture this period. In addition to GDP corrected for purchasing power (PPP) and its derivative GDP|PPP per capita[7], we also computed a variable that reflects the primary production intensity in a country. This variable takes the total contribution to the GDP|PPP of agriculture, fishery and forestry (in 2005 US$) and divides this by the agricultural land area as reported by the Food and Agriculture Organization of the United Nations (FAO). This variable is intended to proxy non-urban land rents as also applied by Jedwab et al. (2022). The assumption is that higher primary production intensity adds more value to land, making it more costly to be converted into urban area. This, of course only captures part of the processes that influence land values and, for example, includes the impact of zoning regulations.

[6] The excluded built-up area is likely to represent a smaller share of the population associated with built-up areas as these include small and dispersed villages with relatively low densities.

[7] Unfortunately, the GDP data are available at 10-year intervals only. So, for 1975 we had to use 1980 values and for 2015 the 2010 values. For the few countries that lacked GDP observations for 2010 we used 2010 starting year from the SSP predictions database.

Applying fixed effects per country yields coefficients (F_c) that control for any unobserved heterogeneity per country, in addition to absorbing any time-invariant variables. This makes it impossible (and irrelevant) to include specific reference to local geographic conditions such as average elevation, climatic conditions etc. The summary statistics of the variables included in our regressions are shown in

Table 12. The table highlights that the value ranges for most variables are large. This is also true for our dependent variable urban density, reinforcing the need for a fixed effects approach that captures this variation. To limit the extreme variation in our independent variables and better approach a normal distribution in the observed values we use natural logarithms of most variables in the regression analysis.

Table 12 Summary statistics of the basic GHSL data and variables used in the panel regression.

Variable	**Description (unit)**	**N**	**Mean**	**St.dev.**	**Min**	**Max**
Area	Total land area (km^2)	568	872,333	2,080,125	1,496	16,247,000
Population	Total population (persons)	568	39,765,717	132,750,701	163,149	1,375,700,000
Density	Population density (persons/km^2)	568	130.78	360.92	1.10	4,897.00
Urban area	Total urban area (km^2)	568	2,352.13	8,080.81	0.86	104,812.00
Urban population	Urban population (persons)	568	11,130,345	31,920,940	172	399,286,000
Urban density	Urban population density (persons/km^2)	568	6,117.38	4,291.55	117.90	33,859.28
GDP|PPP	Gross domestic product by purchasing power parity (2005 bn. US$)	500	341.07	1,077.87	0.76	13,087.12
GDP|PPP per capita	Gross domestic product by purchasing power parity, per capita (2005 bn. US$/pers.)	500	0.00001	0.00002	0.00000	0.0002
Urban population fraction	Urban population as fraction of total population	568	0.30	0.16	0.00	0.80
Primary product. intensity	Primary production (agriculture, fishery, forestry) per agricultural land area (2005 bn. US$/$km^2$)	470	0.0002	0.001	0.000	0.009

Note: the number of observations (N) represents the number of included unique country-year combinations. At most this equivalates 134 countries and 4 years.

4.2 Results

The results of our panel regression are listed in Table 13. Columns (1) and (2) report our baseline specifications using pooled OLS and country fixed effects estimations respectively. Since pooled OLS does not allow for individual heterogeneity and violates the assumption that error terms are not correlated across observations, these results are likely biased compared to the fixed-effects specifications. We include the results from the OLS-approach as a reference, however, to allow more straightforward interpretation and get an impression of the bias it offers when comparing it to the fixed-effects approach. In addition, we can explore the effect of country-specific conditions. The basic model only includes urban population and national income (GDP|PPP) and indicates that countries with a larger urban population, on average have higher densities. An effect that is countered by increased income. We include further explorations of the OLS-model in Annex 4 (Table A4-1) applying several controls to indicate how time-invariant factors such as topography influence density level across our sample. These results suggest that higher land values as proxied by our primary production intensity indeed yield higher urban densities, whereas more rugged terrain seems to result in lower densities.

For econometric reasons, however, we prefer the panel regression with country-fixed effects. The basic panel regression (2) captures country-specific average densities in its country fixed-effects (not reported) and focusses on explaining change over time. This approach shows that with increasing urban population, urban density increases further. The density decreases, however, with increasing national income, suggesting that higher incomes enable suburbanisation with more dispersed urban areas.

Table 13 Selected panel regression results for explanation of natural log of urban density.

	OLS	*Fixed effects*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(urban population)	0.400***	0.416***	0.372***	0.508***	0.508***	0.366***	0.498***
	(0.050)	(0.047)	(0.040)	(0.050)	(0.049)	(0.043)	(0.049)
ln(GDP|PPP)	-0.304***	-0.151***					
	(0.044)	(0.043)					
Ln(GDP|PPP per capita)			-0.207***	-0.127**	-0.152***	-0.248***	-0.175***
			(0.054)	(0.049)	(0.048)	(0.062)	(0.055)
Urban population fraction				-2.064***	-7.802***		-7.072***
				(0.404)	(1.654)		(1.650)
ln(urban population)					0.350***		0.309***
* urban population fraction					(0.098)		(0.097)
ln(primary product.						0.056	0.048
intensity)						(0.054)	(0.044)
Constant (OLS only)	3.821***						
	(0.626)						
Observations	500	500	500	500	500	470	470
R^2	0.363	0.671	0.679	0.762	0.780	0.694	0.784
Adjusted R^2	0.361	0.550	0.560	0.673	0.697	0.570	0.694
F Statistic	141.759***	371.947***	384.917***	387.490***	321.290***	252.276***	240.636***

*Notes: Significance coding: * p<0.1; ** p<0.05; *** p<0.01. Standard errors (between brackets) are clustered at country level.*

If we consider a specification with per capita GDP|PPP, reported in column (3), we find similar coefficients but a slightly better fit then when using GDP|PPP alone. So, correcting GDP|PPP for differences in population size is helpful and clearly indicates that increased prosperity coincides with lower urban densities. Finally, columns (4) through (7) report several specifications using the share of urban population relative to the total population. This results in a better fit and significant coefficients indicating that increasing levels of urbanity correspond to lower densities. A result that may hint at increased low-density suburbanisation when countries become more urbanised. The basic effect of the urban population fraction gets more pronounced when we add an interaction with total urban population. The interaction term itself indicates that the urban population fraction impact is less prominent for countries with larger urban populations.

Models (6) and (7) include our primary production intensity variable that, however, does not yield a significant coefficient in the panel regression setting. The models suggest a positive impact on density as expected, but this relation is not consistent across all countries. We also estimated several alternative specifications using other variables, such as levels of foreign direct investment (FDI) and time-lagged variables (referring to values in the previous period). These are not reported here as they did not yield significant results but included in Annex 4 (Table A4-2). Based on these results, we settle for model 5 as our preferred specification as this combines a fairly high explanatory power with limited number of significant drivers.

In addition, we tested the sensitivity of our results for adapting the set of observations. First, we added the countries with an urban area between 10 and 100 km^2, to see whether the results depend on this selection. This is not the case as we obtain very similar coefficients and explanatory power (Table A4-3). We only lose significance for the interaction between urban population and urban population fraction, which is not surprising as there is substantial variation in these variables for the relatively small countries with an urban area between 10 and 100 km^2.

When excluding the countries with very large urban areas (over 1500km^2 in 2015, such as China, India, USA, Japan and many other developed countries) we again retain very similar results (see Table A4-4). This is interesting as it indicates that the observed relations are not specific for these more urbanised countries. We further explore the impact of selections based on total urban area in Table A4-5 where we estimate our preferred model separately for categories of countries with a similar urban area size. Again, we find that our main conclusion hold: urban densities increase with the size of the urban population, decrease with increasing GDP|PPP per capita and urban population fraction, with the latter impact being smaller for countries with large urban population. When looking at the groups of countries separately, we find that the countries with smallest urban area (between 10 and 100 km^2) show the least consistent and significant results. For the remaining categories GDP impacts are strongest and more significant for the countries with larger urban areas, while the impact of urban population fraction is strongest for the countries with smaller urban areas.

In a final robustness check we exclude the 1975 observations, since we fear these may have a lower accuracy as the satellite imagery underlying the GHSL data for this year is coarser and the statistical records underlying the population and GDP values for many countries are likely to be less reliable. Yet also these results are very similar (see Table A4-6), so there is not much reason to believe this has a strong impact on our findings.

To illustrate the outcomes of our analysis we compare the observed urban density values to those predicted by our preferred model (Column 5) for the worlds' three most populous countries and two characteristic countries on other continents (Figure 2). Together these five countries represent a substantial part of the variation in our sample in terms of absolute urban density values and their change over time. The figure illustrates that the statistical model can capture different trajectories. The continuous increase in Nigeria, the continuous decrease in the Netherlands, the absence of variation in the USA and the initial increase, followed by a decrease in China are all replicated by the panel regression results. Only for India the model is unable capture the general trend, suggesting a much stronger increase than observed. This may very well relate to the rather strict height restrictions in Indian cities that foster outward expansion and limit a further increase in density (Brueckner & Sridhar, 2012).

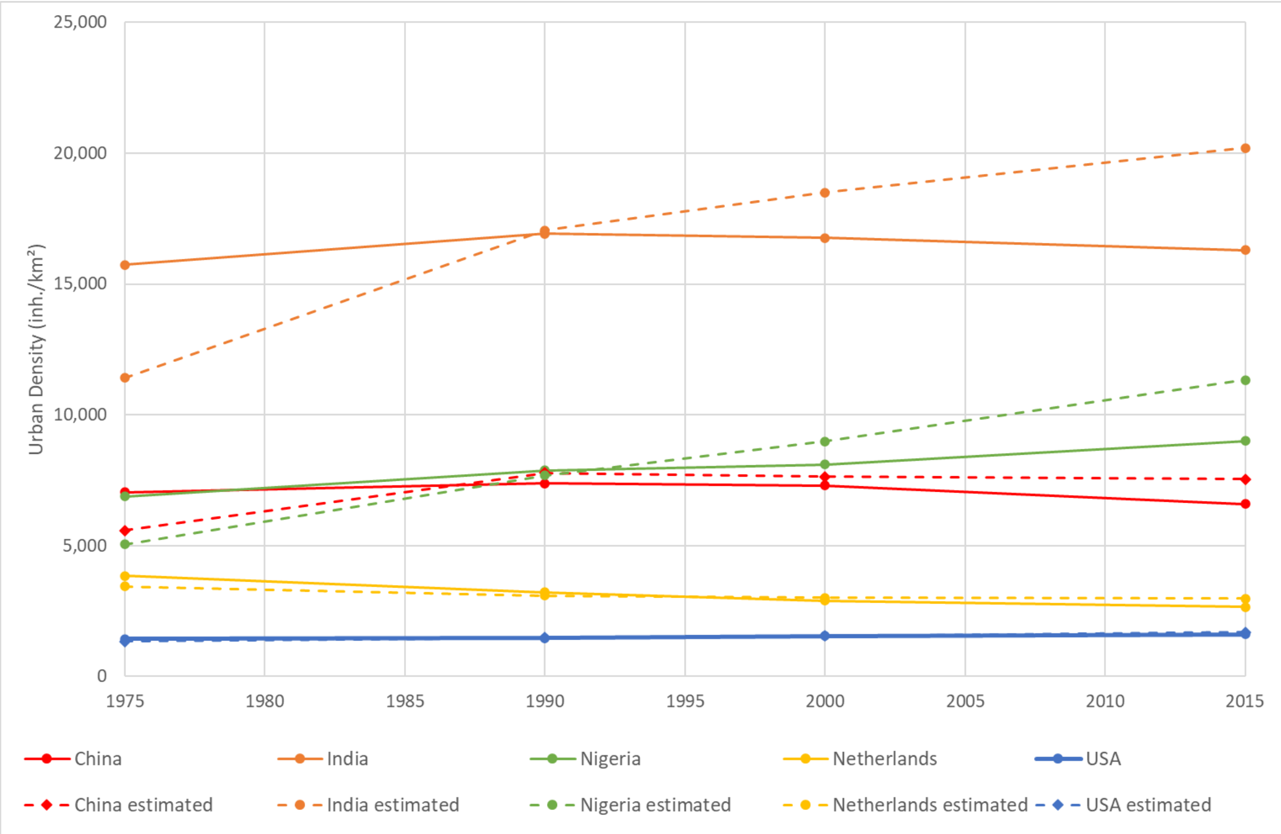

Figure 2 Observed and estimated urban density trajectory for a selection of countries.

5 Conclusion

Using a fixed effect panel regression, we can explain a large part of the substantial variation in country-specific urban densities around the globe. We set up an explanatory model that shows that urban densities increase with the size of the urban population and decrease with increasing GDP|PPP per capita and urban population fraction. The urban population fraction impact is smaller for countries that have larger total urban population.

The explanatory model can be used to project future urban densities using the estimated coefficients and projected future conditions of our explanatory variables. The SSP-database describes, amongst others, future population totals, the urban population fraction and GDP|PPP per capita for all countries in the world per decade for a range of SSP's and can thus be used to estimate future urban densities. Figure 3 shows the resulting total global urban area for our selection of 134 countries when applying the SSP2 scenario values in comparison with several recent future projections for the same scenario.

The estimated total urban area based on our explanatory model shows an increase in line with past developments that is a bit more modest than most other recent future projections for the same SSP. There are several explanations for this outcome. First, we rely on the urban area as contained in the 2UP model. This is a relatively low estimate as it applies a 50% threshold at a 1-km resolution, producing a substantially lower estimate than the GHSL2019 built-up area data it is based on. So, our new estimate has a lower starting point than some of the other studies. Second, the 2UP-based past developments show a diminishing growth trend, as opposed to the sharper increasing trend in, for example, the GHSL2019 data. This provides a more modest growth path to our analysis that is reproduced in the regression analysis. It is interesting to note, however, that the recently released GHSL2023 description of built-up area which is based on more detailed satellite imagery also shows a more modest growth trajectory.

As with any analysis, the validity of the outcomes depends on the quality of the data. Although we use the best available data, it is obvious that some aspects influence our outcomes. First, there is temporal mismatch between some of the datasets. This affects the two extremes of our time-series: the 1975 urban area and population data from GHSL is linked to 1980 GDP|PPP data from the SSP-database, while the 2015 GHSL data is linked to 2010 GDP|PPP data. Second, the GDP|PPP time series is incomplete for the first two observation years, implying that the regression analysis for these earlier years is based on a smaller number of countries. Moreover, the data for the earliest observation year is likely to be less accurate than the later years as the remote sensing images used to produce the GHSL data were coarser and statistical records on population and GDP may have been less complete for many countries.

A more general comment relates to the validity of extrapolating trends as we do here. This approach assumes that the drivers underlying observed, past developments will exert the same influence in the future. While this is likely for the short-term future, it is well possible that their impact will change over longer time periods. This is certainly a risk when extrapolating beyond the observation period as we do here. So, it may be better to limit projections of future change to a similar period similar as for the observations, which in our case would be until around 2050.

In conclusion we think that the presented analysis offers a helpful approach to characterise urban future densities. To further enhance the analysis, we suggest repeating it with the newly released GHLS data to assess whether the observed relationships are robust. This new analysis could make use of more recent, revised GDP data to provide a more extensive set of observation. In addition, the analysis could be repeated for total built-up area (and not just the urban area as defined in 2UP) to exclude the impact of reformatting the data to represent coarser urban areas.

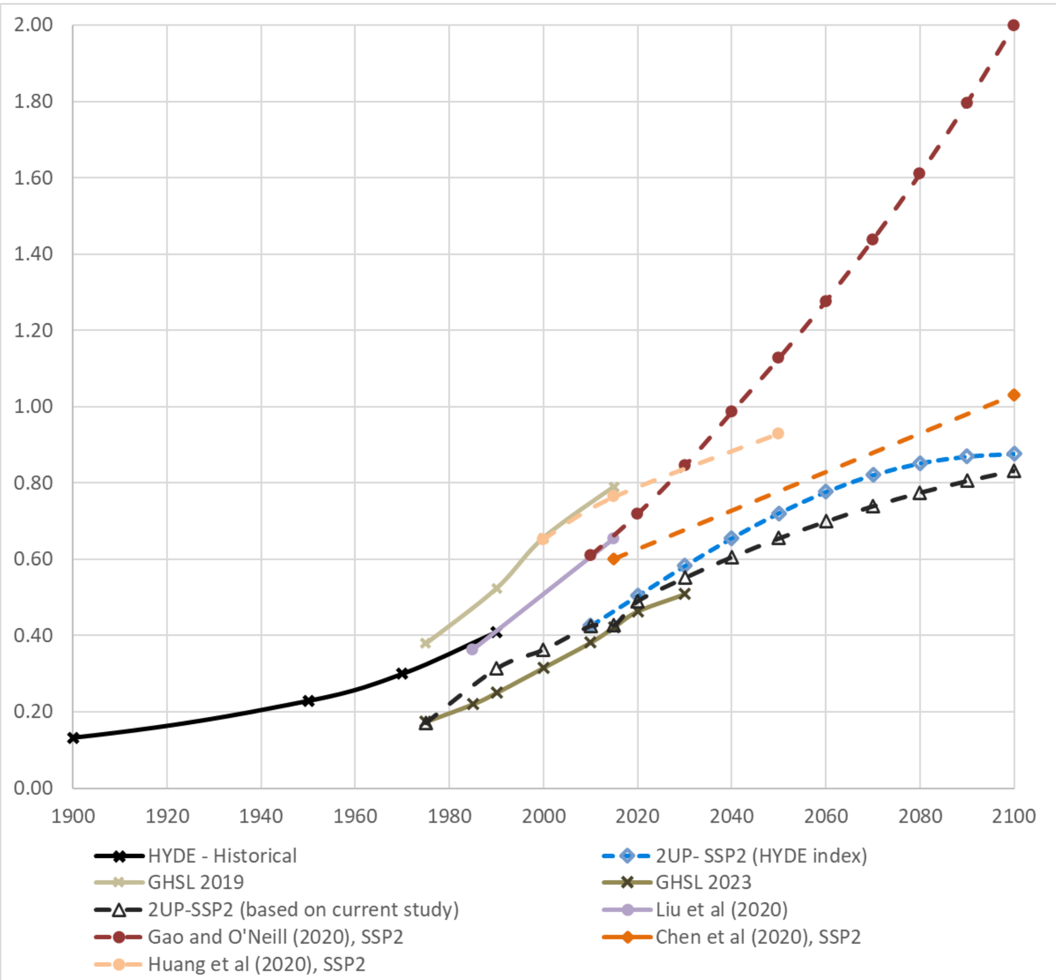

Figure 3 The estimated total global urban area based on the current analysis, compared to a selection of other projections. All future estimates follow the SSP2 scenario.

References

Allan, A., Soltani, A., Abdi, M. H., & Zarei, M. (2022). Driving Forces behind Land Use and Land Cover Change: A Systematic and Bibliometric Review. *Land*, *11*(8). https://doi.org/10.3390/land11081222

Angel, S., Lamson-Hall, P., Blei, A., Shingade, S., & Kumar, S. (2021). Densify and Expand: A Global Analysis of Recent Urban Growth. *Sustainability*, *13*(7), 1-28. https://doi.org/10.3390/su13073835

Angel, S., Parent, J., Civco, D. L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000-2050. *Progress in Planning*, *75*(2), 53-107. https://doi.org/10.1016/j.progress.2011.04.001

Broitman, D., & Koomen, E. (2015). Residential density change: Densification and urban expansion. *Computers Environment and Urban Systems*, *54*, 32-46. https://doi.org/10.1016/j.compenvurbsys.2015.05.006

Brueckner, J. K., & Sridhar, K. S. (2012). Measuring welfare gains from relaxation of land-use restrictions: The case of India's building-height limits. *Regional Science and Urban Economics*, *42*(6), 1061-1067. https://doi.org/10.1016/j.regsciurbeco.2012.08.003

Chen, G., Li, X., Liu, X., Chen, Y., Liang, X., Leng, J., Xu, X., Liao, W., Qiu, Y. a., Wu, Q., & Huang, K. (2020). Global projections of future urban land expansion under shared socioeconomic pathways. *Nature Communications*, *11*(1), 537. https://doi.org/10.1038/s41467-020-14386-x

Colsaet, A., Laurans, Y., & Levrel, H. (2018). What drives land take and urban land expansion? A systematic review. *Land Use Policy*, *79*, 339-349. https://doi.org/10.1016/j.landusepol.2018.08.017

Florczyk, A. J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., Sabo, F., & Zanchetta, L. (2019). *GHSL Data Package 2019, EUR 29788 EN* (L. Publications Office of the European Union, Ed.). Publications Office of the European Union, Luxembourg. https://doi.org/doi:10.2760/062975

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global change and the ecology of cities. *Science*, *319*(5864), 756-760. https://doi.org/10.1126/science.1150195

Güneralp, B., Reba, M., Hales, B. U., Wentz, E. A., & Seto, K. C. (2020). Trends in urban land expansion, density, and land transitions from 1970 to 2010: a global synthesis. *Environmental Research Letters*, *15*(4). https://doi.org/10.1088/1748-9326/ab6669

Jedwab, R., Barr, J., & Brueckner, J. K. (2022). Cities Without Skylines: Worldwide Building-Height Gaps and their Possible Determinants and Implications. *Journal of Urban Economics*, *132*, 103507. https://doi.org/10.1016/j.jue.2022.103507

Kim, Y., Newman, G., & Guneralp, B. (2020). A Review of Driving Factors, Scenarios, and Topics in Urban Land Change Models. *Land (Basel)*, *9*(8), 1-22. https://doi.org/10.3390/land9080246

Koomen, E., van Bemmel, M. S., van Huijstee, J., Andrée, B. P. J., Ferdinand, P. A., & van Rijn, F. J. A. (2023). An integrated global model of local urban development and population change. *Computers, Environment and Urban Systems*, *100*, 101935. https://doi.org/10.1016/j.compenvurbsys.2022.101935

Kuffer, M., Owusu, M., Oliveira, L., Sliuzas, R., & van Rijn, F. (2022). The Missing Millions in Maps: Exploring Causes of Uncertainties in Global Gridded Population Datasets. *ISPRS International Journal of Geo-Information*, *11*(7). https://doi.org/10.3390/ijgi11070403

Li, M. M., Verburg, P. H., & van Vliet, J. (2022). Global trends and local variations in land take per person. *Landscape and Urban Planning*, *218*, 104308. https://doi.org/10.1016/j.landurbplan.2021.104308

Liu, Z., He, C., Zhou, Y., & Wu, J. (2014). How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion. *Landscape Ecology*, *29*. https://doi.org/10.1007/s10980-014-0034-y

PBL. (2021). *Overview of global built-up area for the IMAGE-2UP project; internal memo*.

Seto, K. C., Fragkias, M., Guneralp, B., & Reilly, M. K. (2011). A meta-analysis of global urban land expansion. *PLoS ONE*, *6*(8), e23777. https://doi.org/10.1371/journal.pone.0023777

Seto, K. C., Guneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc Natl Acad Sci U S A*, *109*(40), 16083-16088. https://doi.org/10.1073/pnas.1211658109

Seto, K. C., Sanchez-Rodriguez, R., & Fragkias, M. (2010). The New Geography of Contemporary Urbanization and the Environment. *Annual Review of Environment and Resources, Vol 35*, *35*, 167-194. https://doi.org/10.1146/annurev-environ-100809-125336

UN. (2019a). *World Urbanization Prospects 2018: Highlights*. United Nations (UN) Department of Economic Social, Affairs. https://doi.org/10.18356/6255ead2-en

UN. (2019b). *World Urbanization Prospects: The 2018 Revision*. United Nations (UN) Department of Economic Social, Affairs. https://doi.org/10.18356/b9e995fe-en

van Huijstee, J., van Bemmel, B., Bouwman, A., & van Rijn, F. (2018). *Towards an urban preview; Modelling future urban growth with 2UP*.

van Vliet, J., Verburg, P. H., Grădinaru, S. R., & Hersperger, A. M. (2019). Beyond the urban-rural dichotomy: Towards a more nuanced analysis of changes in built-up land. *Computers, Environment and Urban Systems*, *74*, 41-49. https://doi.org/10.1016/j.compenvurbsys.2018.12.002

Xu, G., Zhou, Z. Z., Jiao, L. M., & Zhao, R. (2020). Compact Urban Form and Expansion Pattern Slow Down the Decline in Urban Densities: A Global Perspective. *Land Use Policy*, *94*(January 2019), 104563-104563. https://doi.org/10.1016/j.landusepol.2020.104563

Annex 1 Main and subcategories of driving forces per paper

Overview of main and subcategories of the driving forces mentioned in the selected papers. Individual sub-categories are coloured according to the main categories of paper 4. Main categories are only colour coded when all related subcategories belong to the same main category.

Main and subcategories of driving forces include in paper 1: Allan et al. (2022)

Main category	Subcategory
Urban growth factors	Transport infrastructure
	Industry
	Accessiblity
	Residence
Policy and regulation factors	Urban / land use policies
	Regulations
Economic and financial factors	Investment
	Urban economy
Contextual factors	Demographic
	Socio-economic factors
	Environment and natural resources

Main and subcategories of driving forces include in paper 2: Angel et al. (2011)

Main category	Subcategory
	Total urban population
	Income (GDP/capita)
	Arable land per capita
	Price of gasoline
	Share of urban population in informal settlements

This paper only mentions specific drivers in the regression models that are listed here in the sub-category column.

Main and subcategories of driving forces include in paper 3: Colsaet et al. (2018)

Main category	Subcategory
Demography	Increase in population
	Increase in urban population
	Internal migrations
	Population size
	Smaller household size
	Proportion of children
	Proportion of elderly
Social processes	Aspiration for detached housing
	Spatial planning culture (e.g., focus on land rights of individuals or collective interests of saving open space)
	Spatial segregation
	Inner city/downtown problems
Economy	Rising GDP/income
	Manufacturing sector
	Tertiary sector
	Tourism sector
	Globalization
	Higher agricultural profitability
	Market failures or imperfections
	Speculation
Infrastructure and transport	Proximity to transport facilities
	Development of transport infrastructure
	Road/highway density
	Automobile reliance

	Gasoline price/parking price
	Density of public transportation
Geographical constraints	Slope
	Elevation
	Rugged terrain
	Attractive climate
	Proximity to water
	Share of water
	Share of green and blue areas
Path dependency and neighbouring effects	Path dependency
	Urban growth in neighbouring cities
Policies and institutions	Regulation
	Incentives
	Governance

Main and subcategories of driving forces include in paper 4: Kim et al. (2020)

Main category	**Subcategory**
Natural environment	Topography
	Amenity
Built environment	Transportation
	Land use
	Job
	Service
Socio-economy	Population density
	Property value
	Landuse plan
	Plan policies

Main and subcategories of driving forces include in paper 5: Seto et al. (2010)

Main category	**Subcategory**
Global capital	International real estate developers
	Demands of multinational firms
	Housing type
Governance	Decentralization
	Financial and institutional national governmental support
	Private governments
	Human, technical and financial resources
	In/exclusion local municipalities in decision-making
Institutions	Multinational corporations affecting real estate market
	Spatial planning by local authorities
	Voluntary sector

Main and subcategories of driving forces include in paper 6: Seto et al. (2011)

Main category	**Subcategory**
	Population growth rate
	GDP growth rate
	Farm subsidy
	Coastal zone location
	1980s indicator
	Study area size

This paper only mentions specific drivers in the regression models that are listed here in the sub-category column.

Annex 2 Overview of drivers their impact on urban density

Main categories, individual drivers and their impact on urban density and references.

Main category	Driver of urban area change	Impact on density*	Ref**
Demography	Migration of rural population to urban areas	n.a.	1
	Labour migration	n.a.	1
	Migration within urban areas (internal migration)	n.a. / ±	1/3
	Urban population	+ / + / ±	1/2/3
	Population density	+ / +	1/4
	Population growth	± / -	3/6
	Total population	+	3
	Household size	+	3
	Proportion of children	-	3
	Proportion of elderly	±	3
Socio-cultural factors	Luxurious lifestyle demand (spacious houses / second homes / aspiration for detached housing / demand for high-end consumer goods/services)	- / - / n.a.	1/3/5
	Share of population in informal settlements	+	2
	Spatial planning culture	±	3
	Spatial segregation (number of distinctive groups / minorities)	±	3
	City centre problems	n.a.	3
Economy	Technological progress / industrial transformation	n.a.	1
	Industry / industrial parks/sites / manufacturing sector	n.a. / n.a.	1/3
	Foreign direct investment / globalization	n.a. / n.a.	1/3
	Investment attraction	n.a.	1
	Market power/incentives	±	1
	Land price / property value	n.a. / +	1/4
	Land price distribution skewness	n.a.	1
	Housing price	n.a.	1
	Tourism development / sector	n.a. / -	1/3
	Economic downturn / unemployment (jobs / trade / industrial / agricultural / commercial)	n.a. / n.a.	1/4
	Gross Domestic Product / income per capita	- / - / ± / n.a.	1/2/3/6
	Tertiary sector	±	3
	Agricultural profitability	n.a.	3
	(Real estate) market failures / imperfections / speculation / inefficiencies / conflict	n.a. / n.a.	3/5
	International real estate developers / property management firms	n.a.	5
	Demands (space / infrastructure / human resources) of multinational firms	n.a.	5

Infrastructure and transport	(Presence of) airports	n.a.	1
	(Presence of) bridges	n.a.	1
	(Presence of / distance to) public transport (railway / subway / light rail)	n.a. / + / n.a.	1/3/4
	Road / highway network density / quality	n.a. / - / n.a.	1/3/4
	(Presence of) wharfs / ferries / harbours / ports	n.a.	1
	Transport cost / gasoline/parking price	+ / +	2/3
	Distance to transport facilities	n.a.	3
	Development of transport infrastructure	n.a.	3
	Automobile reliance	-	3
Geographical configuration	Slope/rugged terrain	n.a. / n.a. / n.a.	1/3/4
	Elevation	n.a. / n.a.	1/3
	Climate quality	n.a. / -	1/3
	Flood prone areas	n.a.	1
	Geographical location	n.a.	1
	Distance to coast	n.a. /n.a. /n.a. /n.a.	1/3/4/6
	Distance to other water bodies (rivers / lakes / wetlands / ponds)	n.a. / n.a.	1/3
	Availability of resources (forestry / ecological / oil / minerals)	n.a. / n.a.	1/4
	Distance to ecosystem services / natural amenities/scenery	n.a. / ± / n.a.	1/3/4
	Arable land per capita	-	2
	Study area size	n.a.	6
Path dependency and neighbouring effects	((Functional) distance to) city/district/town centre / CBD	n.a. / n.a.	1/4
	((Functional) distance to) public facilities (medical care / education / research / commercial / leisure centres / parks)	n.a. / n.a.	1/4
	((Functional) distance to) hotels	n.a.	1
	Neighbouring effect (nearby dense)	n.a. / +	1/3
	((Functional) distance to) built-up areas / settlement	n.a. / n.a.	1/4
	Construction of residential areas on urban periphery	-	1
	Path dependency (historically dense)	+	3
	Land use	n.a.	4
	Distance to residential	n.a.	4
	1980s indicator	n.a.	6

Policies and institutions	Administrative fragmentation / division adjustment	n.a. / n.a.	1/3
	Administrative hierarchical strength / state-level control of local planning / exclusion of local municipalities in decision-making	n.a. / n.a. / n.a.	1/3/5
	Local government (land use) planning / policy (zoning / controlling land ownership)	± / n.a. / n.a.	1/4/5
	Land supply / financing / design / construction of large-scale projects / infrastructure (facilitated by developable land / private enterprises / participation of owners/developers/real estate agencies)	±	1
	Urban growth area regulation (municipalities / urban planning / residential land use regulation building permits)	n.a. / n.a.	1/4
	Urban containment / growth boundaries	n.a. / +	1/3
	Property tax	n.a. / ±	1/3
	Weak / incomplete / unstable planning	n.a.	3
	State-led development	n.a.	3
	Low-density zoning / clustering regulations	-	3
	Protected areas / resources / open space	n.a. / n.a.	3/4
	Subsidies for new urban development	n.a.	3
	Subsidies for highways / cars	n.a.	3
	Impact fees	n.a. / n.a.	3/4
	Subsidies for agriculture	n.a. / n.a.	3/6
	Subsidies for urban renewal	n.a.	3
	Tradeable development rights	n.a. / n.a.	3/4
	Differences in regulation	n.a.	3
	Reliance on local taxes / insufficient financial / institutional governmental support to municipalities	n.a. / n.a.	3/5
	Financial problems of municipalities	n.a.	3
	Capacity / resources (human / technical / financial) of planning agencies	n.a. / n.a.	3/5
	Pressure of interest groups	n.a.	3
	Corruption	n.a.	3
	Decentralization / regionalization	n.a.	5
	Private governments	n.a.	5
	Voluntary sector / NGOs	n.a.	5

* = +: driver has a positive effect on urban density; -: driver has a negative effect on urban density; ±: mixed results are reported; n.a.: paper does not specify a positive or negative effect.
** = Reference to the articles. 1: Allan et al. (2022); 2: Angel et al. (2011); 3: Colsaet et al. (2018); 4: Kim et al. (2020); 5: Seto et al. (2010); 6: Seto et al. (2011).
In one case, multiple scales of interest were used to determine the impact of transport cost on density (Angel et al., 2011). Here, we only show the impact that was found on a country scale (and not on local/city scale), because this report focusses on density on a national level.

Annex 3 References corresponding to the indicators of Table 11

1. United Nations Department of Economic and Social Affairs (2018). *World Urbanizations Prospects: The 2018 Revision*. Retrieved 10 February 2023 from https://population.un.org/wup/Download/

2. UN (2022). *Compact (most used: estimates and medium projections)*. United Nations (UN) Department of Economic Social, Affairs. Retrieved 3 May 2023 from https://population.un.org/wpp/Download/Standard/MostUsed/

3. United Nations Department of Economic and Social Affairs (2022). *United Nations Database on Household Size and Composition 2022I.* Retrieved 10 February 2023 from https://population.un.org/Household/index.html#/countries/

4. UN (2022). *Percentage of annual total population (both sexes combined) by select age group*. United Nations (UN) Department of Economic Social, Affairs. Retrieved 1 May 2023 from https://population.un.org/wpp/Download/Standard/Population/

5. World Bank (2023). *Households and NPISHs Final consumption expenditure per capita (constant 2015 US$).* Retrieved 3 May 2023 from https://data.worldbank.org/indicator/NE.CON.PRVT.PC.KD

6. World Health Organization & UNICEF (2020). *Households*. Retrieved 10 February 2023 from https://washdata.org/data/household#!/

7. Central Intelligence Agency (n.d.). *Ethnic groups*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/ethnic-groups/

8. World Bank (n.d.). *4.2 World Development Indicators: Structure of value added*. Retrieved 10 February 2023 from http://wdi.worldbank.org/table/4.2

9. Statistics Times (2018). *List of Countries by GDP Sector Composition*. Retrieved 10 February 2023 from https://statisticstimes.com/economy/countries-by-gdp-sector-composition.php

10. World Bank (2021). *Foreign direct investment, net inflows (BoP, current US$)*. Retrieved 10 February 2023 from https://data.worldbank.org/indicator/BX.KLT.DINV.CD.WD

11. International Monetary Fund (2021). *Table 1-i: Inward Direct Investment Positions, as of end-2021*. Retrieved 10 February 2023 from https://data.imf.org/regular.aspx?key=61227424

12. The European House Ambrosetti (2022). *The thermometer of a Country's attractiveness.* Retrieved 10 February 2023 from https://www.ambrosetti.eu/en/global-attractiveness-index/#:~:text=%EF%BB%BF%EF%BB%BF%EF%BB%BF%EF%BB%BFThe,and%20Conflict%20Exposure%2C%20added%20in

13. Food and Agriculture Organization of the United Nations (2023). *Macro Indicators*. Retrieved 12 May 2023 from https://www.fao.org/faostat/en/#data/MK

14. World Bank (n.d.). *GDP per capita (current US$)*. Retrieved 10 February 2023 from https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

15. Central Intelligence Agency (n.d.). *Roadways*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/roadways/country-comparison

16. Meijer, J. R., Huijbregts, M. A., Schotten, K. C., & Schipper, A. M. (2018). Global patterns of current and future road infrastructure. *Environmental Research Letters*, *13*(6), 064006.

17. Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (2019). *International Fuel Prices 2018/19.* Retrieved on 10 February 2023 from https://sutp.org/publications/international-fuel-prices-report/

18. Central Intelligence Agency (n.d.). *Railways*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/railways/

19. Central Intelligence Agency (n.d.). *Airports*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/airports/

20. Central Intelligence Agency (n.d.). *Elevation*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/elevation/

21. Nunn & Puga (2012). *Data and replication files for 'Ruggedness: The blessing of bad geography in Africa*'. Retrieved 10 February 2023 from https://diegopuga.org/data/rugged/

22. Kummu, M., De Moel, H., Ward, P. J., & Varis, O. (2011). How close do we live to water? A global analysis of population distance to freshwater bodies. *PloS one*, *6*(6), e20578.

23. Aires, F., Miolane, L., Prigent, C., Pham, B., Fluet-Chouinard, E., Lehner, B., & Papa, F. (2017). A global dynamic long-term inundation extent dataset at high spatial resolution derived through downscaling of satellite observations. *Journal of Hydrometeorology*, *18*(5), 1305-1325.

24. Nguyen, T. H., & Aires, F. (2023). A global topography-and hydrography-based floodability index for the downscaling, analysis, and data-fusion of surface water. *Journal of Hydrology*, *620*, 129406.

25. Winsemius, H., & Ward, P. (2015). *Aqueduct Global Flood Risk Country Rankings*. Retrieved on 4 May 2023 from https://www.wri.org/data/aqueduct-global-flood-risk-country-rankings

26. Rentschler, J., Salhab, M., & Jafino, B. A. (2022). Flood exposure and poverty in 188 countries. *Nature communications*, *13*(1), 3527.

27. Central Intelligence Agency (n.d.). *Natural resources*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/natural-resources/

28. Food and Agriculture Organization of the United Nations (2022). *Land Use*. Retrieved 10 February 2023 from https://www.fao.org/faostat/en/#data/RL

29. Central Intelligence Agency (n.d.). *Land use*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/land-use/

30. Angel, S., Parent, J., Civco, D. L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Progress in Planning*, *75*(2), 53-107.

31. All University Info (2023). *Search University by Country*. Retrieved 4 May 2023 from https://www.alluniversity.info/search-by-country/

32. World Health Organization (2016). *Total density per 100 000 population: Hospitals*. Retrieved 4 May 2023 from https://www.who.int/data/gho/data/indicators/indicator-details/GHO/total-density-per-100-000-population-hospitals

33. World Bank (2023). *Population density (people per sq. km of land area).* Retrieved 4 May 2023 from https://data.worldbank.org/indicator/EN.POP.DNST

34. Global Economy (2021). *Government effectiveness – Country rankings*. Retrieved 10 February 2023 from https://info.worldbank.org/governance/wgi/

35. World Bank (2022). *Global Indicators of Regulatory Governance*. Retrieved 10 February 2023 from https://rulemaking.worldbank.org/en/rulemaking

36. Global Economy (2021). *Government effectiveness – Country rankings*. Retrieved 10 February 2023 from https://info.worldbank.org/governance/wgi/

37. Food and Agriculture Organization of the United Nations (2021). *World Food and Agriculture – Statistical Yearbook 2021*. Rome. https://doi.org/10.4060/cb4477en

38. World Bank (n.d.). *Subsidies and other transfers (% of expense)*. Retrieved 10 February 2023 from https://data.worldbank.org/indicator/GC.XPN.TRFT.ZS?end=2020&start=1972&view=chart

39. World Association of Non-Governmental Organizations (n.d.). *Worldwide NGO Directory*. Retrieved 10 February 2023 from https://www.wango.org/resources.aspx?section=ngodir

40. World Bank (2021). *Entrepreneurship database*. Retrieved 10 February 2023 from https://www-worldbank-org.vu-nl.idm.oclc.org/en/programs/entrepreneurship#total

41. Fortune (2022). *Visualize the Global 500*. Retrieved 10 February 2023 from https://fortune.com/franchise-list-page/visualize-the-global-500-2022/

42. Central Intelligence Agency (n.d.). *Administrative divisions*. Retrieved 10 February 2023 from https://www.cia.gov/the-world-factbook/field/administrative-divisions/

Annex 4 Additional regression results

Table A4-1 Alternative pooled-OLS regression results for explanation of the natural log of urban density. Results are included to indicate the impact of specific drivers on urban density. Note, however, that the panel regression with country-fixed effects is the preferred specification.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(urban population)	0.400***	0.403***	0.185***	0.172***	0.336***	0.172***	0.304***
	(0.050)	(0.050)	(0.039)	(0.041)	(0.057)	(0.035)	(0.053)
ln(GDP|PPP)	-0.304***	-0.308***					
	(0.044)	(0.044)					
ln(GDP|PPP per capita)			-0.237***	-0.260***	-0.244***	-0.289***	-0.292***
			(0.036)	(0.042)	(0.040)	(0.035)	(0.042)
Urban population fraction				0.345	10.296***		9.098***
				(0.393)	(2.415)		(2.094)
ln(primary production intensity)						0.175***	0.157***
						(0.050)	(0.045)
ln(mean terrain roughness indicator)		0.055	-0.054	-0.037	0.014	-0.173*	-0.094
		(0.081)	(0.083)	(0.085)	(0.078)	(0.091)	(0.081)
ln(urban pop.) * urban pop. fraction					-0.669***		-0.589***
					(0.155)		(0.136)
Constant	3.821***	3.758***	2.928***	2.714***	0.515	4.264***	2.011**
	(0.626)	(0.638)	(0.781)	(0.805)	(1.025)	(0.775)	(0.953)
Observations	500	500	500	500	500	470	470
R^2	0.363	0.365	0.310	0.313	0.402	0.420	0.490
Adjusted R^2	0.361	0.361	0.306	0.308	0.396	0.415	0.483
F Statistic	141.759***	94.847***	74.412***	56.490***	66.364***	84.203***	74.112***

*Notes: Significance coding: * p<0.1; ** p<0.05; *** p<0.01. Reported standard errors are clustered at country (c) level.*

Table A4-2 Alternative panel regression results for explanation of the natural log of urban density including reference to time lagged density and foreign direct investment (FDI) levels. All models with country-specific fixed effects (not listed). Our preferred model (5, from Table 13) is added as reference.

	(5)	*(7)*	*(8)*	*(9)*
ln(urban population)	0.508***	0.400***	0.364***	0.362***
	(0.049)	(0.046)	(0.040)	(0.040)
ln(GDP|PPP per capita)	-0.152***	-0.148***	-0.211***	-0.210***
	(0.048)	(0.032)	(0.079)	(0.076)
Urban population fraction	-7.802***			
	(1.654)			
ln(previous urban density)		0.077		
		(0.051)		
ln(urban pop.) * urban pop. fraction	0.350***			
	(0.098)			
FDI- second lowest group of 27 countries			0.035	
			(0.049)	
FDI- middle group of 27 countries			-0.026	
			(0.051)	
FDI- second highest group of 27 countries			-0.003	
			(0.047)	
FDI- highest group of 27 countries			-0.047	
			(0.071)	
FDI- second lowest group of 27 countries				0.013
				(0.038)
FDI- middle group of 27 countries				-0.021
				(0.055)
FDI- second highest group of 27 countries				-0.028
				(0.055)
FDI- highest group of 27 countries				-0.048
				(0.070)
Observations	470	398	466	466
R^2	0.694	0.708	0.702	0.700
Adjusted R^2	0.570	0.556	0.575	0.572
F Statistic	252.276***	210.807***	127.839***	126.561***

*Notes: Significance coding: * p<0.1; ** p<0.05; *** p<0.01. Reported standard errors are clustered at country (c) level. The bottom group of 27 countries acts as reference category for the estimations of foreign direct investment impact.*

Table A4-3 Alternative panel regression results including observations for countries with >10 km² of urban area in 2015[8]. All models with country-specific fixed effects (not listed).

	(1)	(2)	(3)	(4)	(5)	(6)
ln(urban population)	0.441***	0.392***	0.523***	0.524***	0.365***	0.487***
	(0.048)	(0.040)	(0.051)	(0.051)	(0.040)	(0.046)
ln(GDP|PPP)	-0.167***					
	(0.047)					
ln(GDP|PPP per capita)		-0.190***	-0.114**	-0.126***	-0.228***	-0.170***
		(0.050)	(0.046)	(0.048)	(0.058)	(0.052)
Urban population fraction			-2.035***	-4.743**		-6.117***
			(0.391)	(2.398)		(1.500)
ln(urban pop.) * urban pop. fraction				0.167		0.259***
				(0.148)		(0.088)
ln(primary production intensity)					0.051	0.053
					(0.050)	(0.038)
Observations	558	558	558	558	521	521
R^2	0.684	0.683	0.753	0.757	0.701	0.781
Adjusted R^2	0.566	0.565	0.660	0.665	0.578	0.688
F Statistic	439.136***	437.203***	411.751***	314.673***	287.783***	260.477***

*Notes: Significance coding: * p<0.1; ** p<0.05; *** p<0.01. Reported standard errors are clustered at country (c) level.*

Table A4-4 Alternative panel regression results excluding observations for countries with >1500 km² of urban area in 2015. All models with country-specific fixed effects (not listed).

	(1)	(2)	(3)	(4)	(5)	(6)
ln(urban population)	0.404***	0.371***	0.496***	0.493***	0.364***	0.483***
	(0.051)	(0.044)	(0.051)	(0.050)	(0.047)	(0.051)
ln(GDP|PPP)	-0.118**					
	(0.046)					
ln(GDP|PPP per capita)		-0.184***	-0.112*	-0.135**	-0.232***	-0.159**
		(0.064)	(0.059)	(0.056)	(0.072)	(0.064)
Urban population fraction			-2.199***	-7.975***		-7.229***
			(0.467)	(1.747)		(1.745)
ln(urban pop.) * urban pop. fraction				0.367***		0.324***
				(0.106)		(0.106)
ln(primary production intensity)					0.073	0.046
					(0.049)	(0.041)
Observations	385	385	385	385	360	360
R^2	0.672	0.681	0.766	0.783	0.697	0.787
Adjusted R^2	0.549	0.561	0.677	0.700	0.570	0.695
F Statistic	286.365***	298.163***	303.550***	250.238***	194.226***	185.524***

*Notes: Significance coding: * p<0.1; ** p<0.05; *** p<0.01. Reported standard errors are clustered at country (c) level.*

[8] To avoid logs of zero values we exclude three countries without urban area in 1975: Equatorial Guinea, Swaziland and Lesotho.

Table A4-5 Alternative panel regression results for subsets of countries based on their urban area in 2015 following our preferred model (5) in Table 13. All models with country-specific fixed effects (not listed).

urban area in 2015	*10km² -100km²*	*100km² -250km²*	*250km² -500km²*	*500km² -1,500km²*	*1,500km² -10,000km²*	*>=10,000km²*
	(1)	(2)	(3)	(4)	(5)	(6)
ln(urban population)	0.592***	0.515***	0.414***	0.344***	0.695***	0.394*
	(0.152)	(0.076)	(0.065)	(0.090)	(0.131)	(0.207)
ln(GDP|PPP per capita)	0.083	-0.047	0.002	-0.152**	-0.188*	-0.221***
	(0.153)	(0.146)	(0.075)	(0.069)	(0.102)	(0.050)
Urban population fraction	11.210	-12.391***	-13.520***	-16.019***	-5.498	-3.883
	(15.165)	(2.850)	(2.796)	(4.183)	(4.168)	(13.661)
ln(urban pop.) * urban pop.	-0.873	0.709***	0.747***	0.916***	0.152	0.156
fraction	(1.147)	(0.192)	(0.162)	(0.274)	(0.261)	(0.714)
Observations	58	107	108	131	115	39
R^2	0.809	0.854	0.767	0.669	0.814	0.461
Adjusted R^2	0.713	0.791	0.668	0.527	0.738	0.181
F Statistic	40.217***	108.420***	61.870***	46.008***	88.692***	5.347***

*Notes: Significance coding: * p<0.1; ** p<0.05; *** p<0.01. Reported standard errors are clustered at country (c) level.*

Table A4-6 Alternative panel regression results excluding observations for 1975. All models with country-specific fixed effects (not listed).

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
ln(urban population)	0.511***	0.451***	0.519***	0.494***	0.456***	0.494***
	(0.065)	(0.045)	(0.067)	(0.074)	(0.052)	(0.080)
ln(GDP|PPP)	-0.122***					
	(0.039)					
ln(GDP|PPP per capita)		-0.156***	-0.135***	-0.143***	-0.193***	-0.170***
		(0.033)	(0.030)	(0.030)	(0.040)	(0.037)
Urban population fraction			-1.074**	-5.048*		-4.558*
			(0.535)	(2.581)		(2.593)
ln(urban pop.) * urban pop.				0.243		0.224
fraction				(0.162)		(0.161)
ln(primary production intensity)					0.020	0.001
					(0.033)	(0.028)
Observations	398	398	398	398	373	373
R^2	0.679	0.692	0.706	0.715	0.712	0.729
Adjusted R^2	0.514	0.534	0.552	0.565	0.546	0.569
F Statistic	277.452***	294.940***	208.497***	163.388***	194.317***	125.632***

*Notes: Significance coding: * p<0.1; ** p<0.05; *** p<0.01. Reported standard errors are clustered at country (c) level.*