CESifo WORKING PAPERS

8660
2020

October 2020

The Geographical Determinants of Within-City Heterogeneity in Urban Density

Melanie Krause, André Seidel

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8660*

The Geographical Determinants of Within-City Heterogeneity in Urban Density

Abstract

This paper studies the impact of building land limitations on within-city variation in urban density and its components crowding, residential coverage, and building height. We utilise geographical obstacles like steep inclines or water bodies as exogenous source of building land limitations within parts of cities. We combine novel high resolution (10m x 10m) geo-spatial data on geography, building height and footprints with Norwegian register data. Our unit of observation are neighborhoods with an average size of 0.3 sqkm. The data indicates a high heterogeneity among the components of urban density for similar total density levels. Our main finding is that local building land limitations increase local urban density and all of its components, with the effect being particularly strong for building heights. Hence, we find support for policies that use building land restrictions to alter urban density within parts of cities. Moreover, we show that geography is another important source of inner-city heterogeneity in urban density, in addition to distance to the city center.

JEL-Codes: R230, R310, R210, C800.

Keywords: urban density, building heights, geography, neighborhoods, inner-city differences.

Melanie Krause
University of Hamburg
Department of Economics
Von-Melle-Park 3
Germany – 20146 Hamburg
melanie.krause@uni-hamburg.de

André Seidel
University of Bergen
Department of Economics
Fosswinckels gate 14
Norway – 5007 Bergen
andre.seidel@uib.no

October 23, 2020

1 Introduction

Urban population density varies across different parts of cities. In some neighborhoods, apartments are smaller, buildings are higher and/or packed more tightly next to each other. Urban density has been linked to welfare outcomes ranging from productivity to crime and pollution (see for example Ahlfeldt and Pietrostefani, 2019, Brownstone and Thomas, 2013, Ciccone and Hall, 1996, Larsson, 2014), most recently also the spread of covid-19 (Rocklöv and Sjödin, 2020). But it is still an open question what determines inner-city differences in density and its components. Distance to the central business district (CBD) has been shown to be an important factor, both theoretically in the classical Alonso-Muth-Mills model (Brueckner, 1987) and empirically (Bertaud and Malpezzi, 2014, Zielinski, 1980). Yet, a large part of inner-city heterogeneity in density remains unaccounted for. In this paper, we argue that building land limitations play a key role.

We analyze the impact of building land limitations on the within-city variation in urban density and its components crowding, building height, and residential coverage. Limitations in built-up land affect land prices, which is one of the reason why land might be determined not only be nature but also by policies. Furthermore, land without built-up still offers uses within the urban fabric as public recreational area. The public good character of such open spaces commonly also induces policy makers to regulate built-up. Both issues make it hard to study causal effects (Duranton and Puga, 2015, Fischel, 2004, Glaeser and Kahn, 2004). We overcome this issue by utilising the distribution of geography-induced building land limitations within cities such as step inclines or water bodies and by using fine-grained data that allows to account for exiting built-up regulation.

To answer our research question, we assemble a novel high-resolution geo-spatial data set at the neighborhood level for Norway. With an average residential area of 0.3 sqkm and average population of 665 inhabitants, this unique data has the necessary granularity to study inner-city differences in density. In addition to providing data on the geographical variables, we construct measures on building footprints and height. To this goal, we make use of high-resolution (10×10m) radar images on total elevation and ground elevation from the National Detailed Altitude Model project provided by the Norwegian mapping authority. From the combination of this data set with data on residential built-up from the European Settlement Map, we are able to compute a 10m $\times$ 10m raster reflecting building heights for entire Norway. Based on this, we derive our final data set on neighborhood-level values of geographical variables such as natural elevation and slope as well as density and its components.

To our knowledge, we are the first paper to empirically study neighborhood-level differences not only in urban density but also in its components crowding, building height, and residential coverage. Our data reveals a substantial heterogeneity in the components of urban density even between those part of the city that have the same overall level of urban density. For example, looking at the urban density gradients reveals that building height decreases more uniformly with distance to the CBD than crowding.

We obtain three main results in our paper, all related to the effects of building land restrictions. First, limiting available building land within parts of cities increases urban density in these parts. Second, limiting available building land affects all three density components, with the effect being particularly strong for building height. Finally, a heterogeneous distribution of geographical obstacles that limit built-up leads to inner-city heterogeneity in urban density. Neighborhoods with the same distance to the CBD can have very different urban density when geography is highly heterogeneous. Our results are also robust to numerous specifications, including the inclusion of socio-demographic and income variables or the definition of the CBD or the unit of observation.

Our identification rests on the exogeneity of geography-induced natural built-up limitations and the ability to control for built-up regulation. Housing supply is typically determined by geographical and regulatory factors in terms of zoning laws (Duranton and Puga, 2015, Glaeser and Gyourko, 2018). For example, Hilber and Vermeulen (2016) use British data on local planning authorities and find that regularity constraints affect the house price-earning elasticity more strongly than uneven topography. Shertzer et al. (2018) show that that zoning laws established in Chicago in 1923 still affect the inner-city variation of population density in that city nowadays. Green et al. (2005) find that the price elasticity of housing supply varies significantly across U.S. metropolitan areas according to their regulatory regime. In our empirical study, we can account for regulatory effects as we can utilise variation at the neighborhood level, while building regulations are set at a higher level of aggregation in Norway (Kommunal- og Moderniseringsdepartementet, 2008).

In fact, the unique set-up in Norway is particularly appropriate for studying our research questions: (i) We have high-resolution data available which allow us to calculate both density and its components as well as geography and link it to socio-economic characteristics at the neighborhood level. Such fine-grained level is not available for many other countries. (ii) We also have a high inner-city variation in geographical features which provide an ideal testing ground for our hypothesis. (iii) As mentioned above, the precise regulatory set-up in Norway makes the direct building regulation effect at the neighborhood level much less of an issue than in other countries. Despite these special conditions, Norwegian cities share many features with other agglomerations around the world, from their development around historic market places to the life-cycle based sorting behavior of their inhabitants (Andersen, 2011, Baum-Snow and Hartley, 2017, Helle et al., 2006, Jedwab et al., 2020). This makes our results readily generalizable to other countries.

Taking together, our findings yield important implications for policymakers and urban planners: While geography per se is given, understanding the mechanisms is vital for using the appropriate policy instruments to shape the city. In particular, open space can be regulated, with the corresponding effects on urban density and its components.

Our results add to the empirical literature on how building land limitations impact cities. Our work is most directly related to the strand of the literature that studies how geography impacts overall city density (Saiz, 2010) and shape (Harari, 2020). In contrast to these studies, we consider how within-city variation in geography leads to local variations in density rather than looking at overall city size and density. Our findings therefore indicate that building land limitations can be a tool for policy makers to alter urban density in specific parts of cities, rather than at the city level. In this respect, we also add to the literature investigating the determinants of urban sprawl (Burchfield et al., 2006, Glaeser and Kahn, 2004).

Another strand of literature to which we contribute is the so far small number of studies on the components of urban density. In terms of case studies, Angel et al. (2019) look at average crowding, building height and residential coverage in selected world cities. In particular, the economics of building heights is quickly garnering interest, see Ahlfeldt and Barr (2020) for a literature overview. Recent applications include firm productivity in tall commercial buildings (Liu et al., 2018), the land price elasticity of skyscrapers in Chicago (Ahlfeldt and McMillen, 2018), as well as slums and building heights in Jakarta (Harari and Wong, 2018) and Nairobi (Henderson et al., 2019). To the best of our knowledge, our paper is the first to study the variation of density and all of its components at the within-city level across several cities, as well as examining its geographical determinants.

We also add to the literature investigating the impact of geographical amenities on population densities. At the cross-city level, cities with more desirable geographical amenities, such as warm climate and ocean access, are known to have higher population densities (Albouy and Stuart, 2014, Carlito and Saiz, 2019, **?**). But within cities, other factors are coming into play. Because households are willing to pay for them, (geographical) amenities can explain the spatial income distribution within cities, see Brueckner et al. (1999) for theoretical considerations and Lee and Lin (2018) for empirical results on how geographical amenities anchor the rich persistently to certain parts of the city. Our findings suggest that the clear patters that can be found across cities tend to be less sharp within cities, for example more hours of sunshine are associated with less crowding but higher residential coverage, which makes the overall effect on urban density ambiguous.

We anchor our contribution to the classical literature of urban economic models. To frame our research question, we provide a simple way of incorporating geographical constraints on land suitability into the standard model by Alonso (1964), Muth (1969) and Mills (1967). Theoretical contributions going into a similar direction are the work by Brueckner (1983) on private yard space and Turner (2005) on open spaces. In our conceptual framework, we combine geographical constraints on the supply side with consumers' utility from open space, e.g. areas with no built-up. This stylized model yields testable predictions about the effects of building land restrictions on the components of density.

Finally, we provide a contribution on the methodological front. Our procedure to derive high-resolution building height data at the 10 by 10 meter level adds to the remote sensing literature. The method we propose uses the digital surface and terrain models provided by the Norwegian mapping authority. Similar data can also be obtained from Airbus, who commercially distribute the high resolution TanDEM-X data generated by the European Space Agency (ESA). Hence, with sufficient funding our method could be used to obtain building heights data for every city of the world.

The remainder of this paper is organized as follows. Section 2 lays out the conceptual framework for our analysis. We describe the construction of our geo-spatial data set in Section 3 and present descriptive statistics of our variables in Section 4. The estimation strategy in Section 5 is followed by the presentation and discussion of our results in Section 6. Section 7 concludes. The Online Appendix contains derivations of our theoretical model (Online Appendix A), more details on the data preparation process (Online Appendix B), supplementary descriptive statistics (Online Appendix C) as well as additional robustness checks (Online Appendix D).

2 Conceptual Framework

2.1 A Small Theoretical Model

To frame our empirical analysis, we introduce geographical constraints on land suitability into the standard Alonso-Muth-Mills style urban economic model. In the following, we will lay out our main model assumptions and results. The whole model with all the derivations can be found in Online Appendix A.

Let $G \in (0,1)$ be the spectrum of geography-based land properties, ranging from 0 (perfectly suitable for building) to 1 (completely unsuitable for building). While G will influence building supply, our model also features a corresponding component on the demand side, households' preference for public open space or recreational area.

2.1.1 Demand Side

Households receive an income y, live in different rings with distance x from the city center and have to pay a transport cost τ to get to their jobs there. As in the standard model, they derive utility from the numeraire consumption good c and housing q, which is measured in square meters and costs the rental price p. The new feature is households' disutility from the degree of built-up b within the city ring they live in:

Assumption 1. *Consumers derive a disutility from the share of built-up area b*

$$\frac{\partial v}{\partial b} < 0 \tag{1}$$

The built-up in an area depends on x as well a second exogenous component, geography g in a given ring. Remember that G denotes the land-plot specific geographical constraints that builders face, while g denotes the overall geography within a ring. Our measure for g will later be a key explanatory variable in our empirical analysis. While g can differ between different rings of the city, it is independent from x and purely determined by nature. Households maximize utility

$$v(c, q, b) = u, \tag{2}$$

while c, q and b all depend on distance x as well as geographical constraints g. Freedom to move implies equalized utility within and across cities.

2.1.2 Supply Side

As in the standard model, building firms compete for land L and use capital K to build houses with a concave production function H that is homogenous of degree one. In particular, concavity $\frac{\partial^2 H(L,K)}{\partial K^2} < 0$ implies that higher buildings are increasingly more expensive to build. As in the standard model, we normalize by dividing by L and will work with $h = \frac{H}{L}$. Note that this means that "developers are indifferent to the value of L; the size of housing complexes is indeterminate" (Brueckner, 1983, p.219). The capital-land-ratio $S = \frac{K}{L}$ is an "index for building height" (Brueckner, 1987). Let us now include land-plot-specific geographical constraints G on land suitability into the production function h: $h(S, G)$.

Assumption 2. *Geographical constraints decrease building output and make capital less productive in the building production function:*

$$\frac{\partial h(S,G)}{\partial G} < 0; \frac{\partial^2 h}{\partial S \partial G} < 0 \tag{3}$$

As usual, capital is rented at an exogenously given rate i. Building firms lease land at a rate r, which depends on location x and geographical constraints G. Firms' profit is then given by

$$\Pi = p(x, b) \cdot H - i \cdot K - r(x, G, b) \cdot L. \tag{4}$$

We assume that builders do not consider their impact on total built-up b when deciding to build a house on a piece of land by setting $S > 0$. Each individual firm believes their effect on b to be marginal and therefore not influencing p or r. We can then derive that

$$\frac{\partial S}{\partial G} = -\underbrace{\frac{\partial^2 h}{\partial S \partial G}}_{<0} \cdot \underbrace{\left(\frac{\partial^2 h(S,G)}{\partial S^2}\right)^{-1}}_{<0} < 0. \tag{5}$$

The first factor is negative because of capital's diminishing return in building process, while the second factor is negative as geographical constraints make building more expensive (Assumption 2). This shows us that building heights get shorter in more geographically constrained land-slots.

Let us now combine the individual decisions of building firms in order to analyze their effect on total built-up. For this, we first assume that the distribution of G leads to the density function $f(G, g)$. In our empirical analysis, g increases the frequency of land plots with high geographical obstacles in a given ring. Hence we assume that $\frac{\partial f(G,g)}{\partial g} < 0$. With this we now can derive b by looking at the marginal $\tilde{G}$ for which a construction firm would be indifferent to build houses ($S = 0$):

$$p \cdot \frac{\partial h(0, \tilde{G})}{\partial S} - i = 0. \tag{6}$$

On all land with geographical constraints $G < \tilde{G}$, there will be built-up. Therefore we can write

$$b = f(\tilde{G}, g). \tag{7}$$

From eq. 6 we can implicitly determine the relation between rental prises, built-up, distance and level of geographical obstacles g.

2.1.3 Comparative Statics on g

We analyze how our key variables change with increasing geographical obstacles g within a ring of the city. In particular, we find that an increase in g raises rental prices

$$\frac{\partial p}{\partial g} = \underbrace{\frac{\partial f}{\partial g}}_{<0} \cdot \Big[q \cdot \underbrace{\frac{\partial v}{\partial c}}_{>0} \underbrace{\Big(\frac{\partial v}{\partial b}\Big)}_{<0}^{-1} + \underbrace{\frac{\partial f}{\partial \tilde{G}}}_{>0} \cdot \underbrace{\frac{\partial h(0,\tilde{G})}{\partial S}}_{>0} \cdot \frac{1}{p} \cdot \underbrace{\Big(\frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G}\Big)}_{<0}^{-1} \Big]^{-1} > 0. \tag{8}$$

This result is in line with findings from the real estate literature that the size of and distance to lakes and other natural recreational areas increase the attractiveness of a location, as reflected in house prices (see for instance Krumm, 1980, Mahan et al., 2000).

Our model also yields that geographical constraints increase building heights:

$$\frac{\partial S}{\partial g} = -\frac{\partial h(S)}{\partial S} \cdot \frac{1}{p} \cdot \underbrace{\Big(\frac{\partial^2 h(S,G)}{\partial S^2}\Big)}_{<0}^{-1} \cdot \underbrace{\frac{\partial p}{\partial g}}_{>0} > 0. \tag{9}$$

Moving on to the floor space consumption q, we assume that demand can be described by a non-further specified function depending negatively on price, as in the standard model. This yields a negative effect, such that apartments are smaller in areas with geographical obstacles:

$$\frac{\partial q}{\partial g} = \nu \cdot \frac{\partial p}{\partial g} < 0. \tag{10}$$

Finally, We look at total built-up and obtain a negative effect:

$$\frac{\partial b}{\partial g} = \underbrace{\frac{\partial f}{\partial g}}_{<0} \cdot \Big[1 + \underbrace{\frac{\partial f}{\partial \tilde{G}}}_{>0} \cdot \underbrace{\frac{\partial h(0,\tilde{G})}{\partial S}}_{>0} \cdot \underbrace{\frac{\partial v}{\partial c}}_{>0} \underbrace{\Big(\frac{\partial v}{\partial b}\Big)}_{<0}^{-1} \Big(p \cdot q \cdot \underbrace{\frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G}}_{<0} \Big)^{-1} \Big]^{-1} < 0. \tag{11}$$

2.2 Model Predictions: The Effect of g on Density

The focus of our empirical analysis will be the effect of ring-specific geographical obstacles, as captured by g, on urban density and its components. Let us relate the model variables to urban density and formulate hypothesis. Following Angel et al. (2019), we define urban density as the ratio of population to the urban extent and split it up as follows:

$$\text{Urban Density} = \text{Crowding} \cdot \text{Building Height} \cdot \text{Residential Coverage} \tag{12}$$

$$\frac{\text{Pop}}{\text{Urban Extent}} = \frac{\text{Pop}}{\text{Floor Area}} \cdot \frac{\text{Floor Area}}{\text{Footprint}} \cdot \frac{\text{Footprint}}{\text{Urban Extent}} \tag{13}$$

With q denoting apartment space per person, crowding can be expressed as $\frac{1}{q}$. Building height is directly given by S. Using eq. 10 and eq. 9, we can based on our stylized model predict the sign of the effect of g on two of the components of urban density directly:

Proposition 1. *The effects of an increase in geographical constraints in a certain ring of a city on urban density in this ring are as follows:*

(a) Crowding increases,

(b) Building height increases.

To make a prediction for the total effect of g on urban density we have to identify the effect on the third component, residential coverage. Our model does not directly allow to make such a prediction. Empirically, the urban extent consists of, among others, building footprint, public recreational areas like parks, private recreational areas like yards, as well as roads and walkways. While our model predicts a decrease in the built-up share, it is silent on the components. The proportion between roads and building footprints might conceivably be fixed, but for building footprints and private yard space, this is far from clear.

Brueckner (1983) examines yard space and its relation to other goods, such as apartment size in more detail in an extension of the Alonso-Muth-Mills model. While yard space does not explicitly feature in our model, we may argue along similar lines and explore the case that households regard yard space and open space as substitutes rather than complements. When public open space outside the neighborhood increases in response to g, building footprints may increase as people reduce their demand for private yard spaces given the availability of public open space. In this case, there would be an increase in residential coverage resulting from an increase in geographical constraints. Making use of this argument as well as eq. 12 we can state:

Proposition 2. *(a) If households consider private yard space and public open space as substitutes, it is possible that residential coverage increases when geographical constraints g increase. (b) Under these conditions, the total effect of geographical constraints on urban density will unambiguously be positive.*

2.3 Discussion of the Model and Its Limitations

The main goal of our model is to serve as a framework for our empirical analysis. It is very simple and stylized in order to focus on what we perceive to be the main mechanism through which geography determines heterogeneity in urban density and its components. In fact, models of the traditional Alonso-Muth-Mills style such as ours are known to capture various features of real-life cities (Brueckner, 1987). But compared to more sophisticated models such as Turner (2005), Ahlfeldt et al. (2015) and Murphy (2018), we necessarily neglect a number of components which are worth discussing.

For example, there is no income heterogeneity in our model. Income is known to be correlated with desirable geographical amenities (Brueckner et al., 1999, Lee and Lin, 2018) and therefore might affect our results. In the empirical analysis, we will take this into account by using mean neighborhood income as well as other socio-demographic variables as controls.

Moreover, our model neglects the overall city-level effects. A reduction in overall available building land commonly increases urban density in the entire city as shown by Saiz (2010). Our model deliberately abstracts from the effect that the distribution of geography in one ring has on the overall land available in the city and the subsequent influence on the rental price level across the city. This is done with a view to the empirical analysis, where we will include both city and ring-level fixed effects to account for political economy issues. As these fixed effects will in addition absorb the overall city-level effects, we do not model them from a theoretical perspective.

Finally, our model does not formally include private yard space in households' utility function. Brueckner (1983) analyzes in depth how households might trade-off apartment size vis-à-vis yard space depending on whether they are considered as substitutes or complements. Including this in our model would sacrifice analytical tractability, as one would not only have to impose assumptions on the trade-off between apartment size and private yard space, but also on the trade-off between private yard space and public open space. We briefly explore what imposing one such assumption might mean under certain circumstances (Proposition 2), but in the end, this is an empirical questions.

We have seen that our model yields a number of testable predictions for our empirical analysis. Our theoretical considerations show that local geography can be another source of within-city density heterogeneity in addition to distance from the CBD. Geographical heterogeneity can lead to various non-standard city forms and density patterns that are at first glimpse not in line with the what the most simple theory would predict. For example, it is possible that building height first decreases, then increases and afterwards decreases again when moving further away from the CBD. This would be the case if there is a ring in between with vary harsh geographical conditions relative to all other rings. In such a ring the geography effect might overcompensate the distance effect so that density is higher in that ring than in the adjacent one. To what extent these effects are empirically relevant will be explored next.

3 Data

Norwegian cities have a unique inner-city variation in geography, therefore providing an excellent testing ground for our hypothesis of building land limitations on density. At the same time, their natural shape with coasts, mountain slopes and islands makes Norwegian cities particularly complex from the point of view of simple circular and monocentric urban economics models. If we can empirically confirm key model predictions in such a setting, it bodes well for other cities.

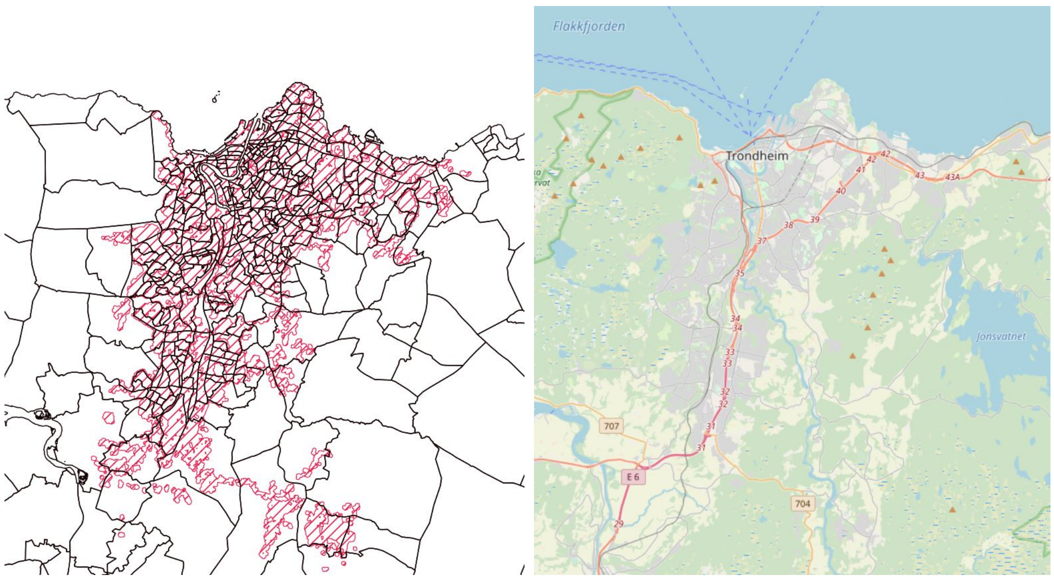
We will in the following present our novel data set, which has been constructed from various high-resolution geo-spatial variables from a number of sources. In particular, we combine data on geographical features and ground elevation with building footprints and height, as well as administrative data on income and socioeconomic characteristics. More details on the data and the process of data preparation is contained in Appendix Online Appendix B.

3.1 Unit of Observation: Neighborhood

Our unit of observation is the neighborhood. We define this based on the smallest administrative unit in Norway, the *grunnkrets*, of which there are approximately 14000. Working with Norwegian data at this level has two advantages: (i) We can obtain average pretax yearly income, as well as other socio-economic characteristics, at the neighborhood level from the population and income register, using data from 2013. (ii) The regulation influencing urban built-up is decided on the next higher administrative level, the *kommune* (Kommunal- og Moderniseringsdepartementet, 2008). In our study, this allows us to account for large proportion of omitted variables related to the political economy of built-up regulation, whose importance has been shown in other studies (Duranton and Puga, 2015, Hilber and Vermeulen, 2016).

With a view to our research question, we define a neighborhood as the *residential built-up area of an urban grunnkrets*. We combine the information on continuous built-up area from the European Settlement Map (ESM) with urban classification from the Global Human Settlement Layers (GHSL) of 2015. Figure 1 shows the *grunnkrets* borders in black and the urban residential built-up areas in red (on the left), compared to the area of Trondheim in the OpenStreetMap project (on the right).

Figure 1: Neighborhoods in Trondheim



Note: The figure shows shows the *grunnkrets* borders in black and in red the urban residential built-up areas (on the left), compared to the area of Trondheim in the OpenStreetMap project (on the right).

3.2 Urban Density and Its Three Components

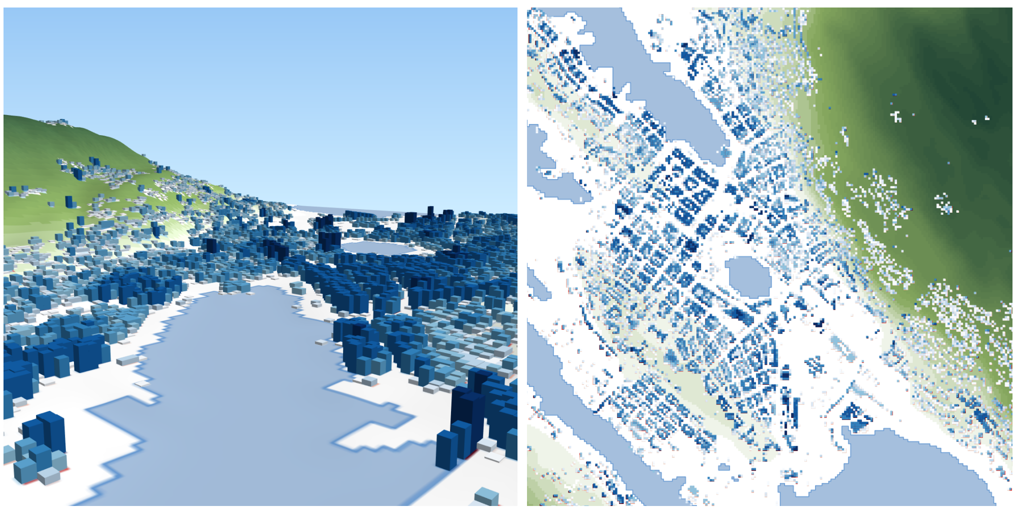
Urban density, our main outcome variable, is calculated as the number of people per sqkm of the urban, residential extent.

3.2.1 Building Height and Footprint

For urban building footprint and height, we rely on the ESM data and high resolution laser telemetry data from the National Detailed Altitude Model project provided by the Norwegian mapping authority. The data was collected from 2014 to 2016 by aircraft or helicopter-mounted laser scanners. The vertical resolution is of 10 m $\times$ 10 m; the horizontal resolution lies in the realm of centimeters. The output is the Norwegian Digital Surface Model (DSM), which includes all elevation, both natural and man-made. In addition, the Norwegian mapping authority also provides the so-called Digital Terrain Model (DTM) which reflects only ground elevation. Our approach is to take the difference between DTM and DSM, which yields non-ground elevation, including the height of man-made objects like houses as well as natural objects such as trees. We require the features to have a minimum height of 1 meter and to be marked as residential built-up in the ESM data, ending up with a 10 m × 10 m raster reflecting building heights for entire Norway. Defining a building footprint as an area where we measure a positive building height, we also obtain the final building footprint map.

To illustrate our approach, Figure 2 shows the 3D view of the old port of Bergen (Brugen) from the sea (on the left) and an eagle's view on the city center (on the right). In both figures, blue indicates built-up, with a darker blue indicating higher buildings.

Figure 2: Building height and footprint in Bergen



Note: The figure shows the 3D view of the old port of Bergen (Brugen) from the sea (on the left) and an eagle's view on the city center (on the right). In both figures, blue indicate built-up, with a darker blue indicating higher buildings.

3.2.2 Crowding and Residential Coverage

According to eq. 12, crowding is given by the number of people by the floor area in square meters. We infer the floor area by the building volume divided by 3m (assumed to be the average floor height). The building volume is, in turn, calculated as the product of building height and the building footprint:

$$\text{Crowding} = \frac{\text{Pop}}{\text{Floor Area in } m^2} = \frac{\text{Pop}}{\frac{1}{3} \cdot \text{Building Height} \cdot \text{Footprint}}. \tag{14}$$

Residential coverage, the final component in eq. 12, is given by the building footprint divided by the urban, residential extent. This means that a neighborhood with more parks, streets and/or private yards will have a lower residential coverage than a neighborhood where buildings are tightly packed next to each other.

3.3 Geography

Our main explanatory variable g is an index of the average geographical constraints within a ring of a city, which we construct based on four individual components. They are all calculated from the 10m $\times$ 10m laser telemetry data from the National Detailed Altitude Model. We expect these variables both to affect built-up within a given neighborhood and overall density across adjacent neighborhoods:

(i) **Slope mean**, the mean slope within a neighborhood measured in degrees. Higher slopes are known to increase building costs; see Saiz (2010), who declares inclines of more the 15% as unsuitable for built-up. With slightly less than 10% of all neighborhoods in our data set showing built-up despite being located at a steeper slope than 15%, we use this as a cutoff value.

(ii) **Slope COV**, the variation of the slope between 100m $\times$ 100m grid cells. It captures the irregularity of the terrain, which makes consistent built-up particularly difficult. Less the 10% of built-up neighborhoods in Norway have a higher slope coefficient of variants then 0.6938003, so this will be our cutoff in this category.

(iii) **Elevation mean**, the mean elevation of a neighborhood. Higher altitudes increase built-up costs because raw materials have to transported further up. Less the 10% of neighborhoods in Norway have built-up and are higher then 173.455m, which will serve as our cutoff value for land suitability.

(iv) **Ocean**, classified as every bit of land below the mean sea level. Building on or close to water is particularly challenging in the Norway fjords, where the sea beds become deep very quickly. We assume that ares on water are unsuitable for built-up.

We avoid the circulatory argument when measuring the availability of land suitable for built-up, by working with neighborhoods and artificial neighborhoods outside of the original neighborhoods. For this we randomly locate points within the circumference of the urban residential built-up areas and generate Voronoi polygons with similar geometric properties as the actual neighborhoods (see Appendix for an example). Note that we only use artificial neighborhoods when measuring geography across neighborhoods, but we do not use them in our main empirical analysis as dependent variable.

We classify all neighborhoods as unsuitable for built-up that have at least one of four characteristics above the cutoff value. Hence, for neighborhood i in ring r in *kommune* k located in city c, we define

$$\lambda_{irck} = \begin{cases} 1 & slope\,mean_{irck} > 8.5308 \cup slope\,COV_{irck} > 0.6938003 \\ & \cup\, elevation\,mean_{irck} > 173.455 \cup ocean_{irc} > 0 \\ 0 & \text{else.} \end{cases} \tag{15}$$

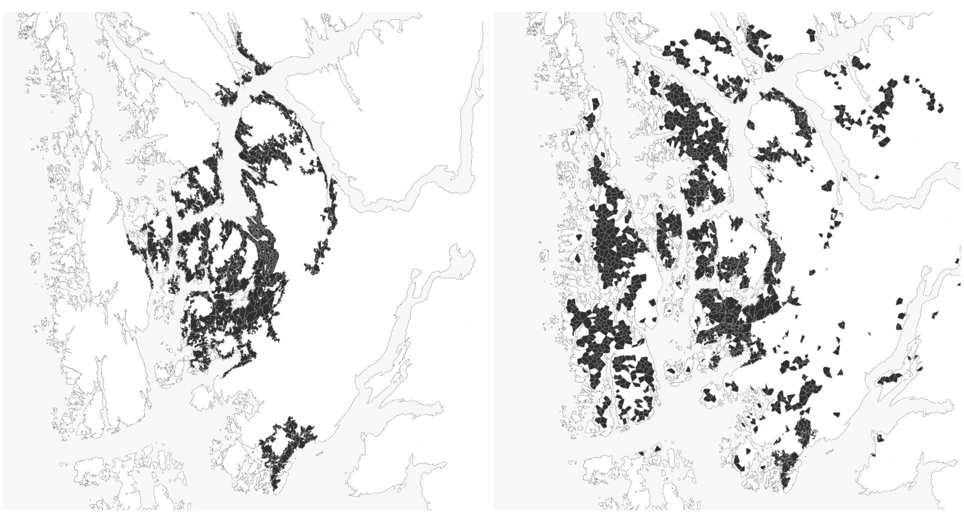
This allows us to calculate the share of land that is unsuitable for built-up within a ring of the city, witch refer to as **geography** in our empirical analysis, as

$$g_{rck} = \frac{\sum_{i \in r}(1 - \lambda_{irck}) \cdot area_{irck}}{\pi \cdot (r^2 - (r-1)^2)}, \tag{16}$$

where $area_{irck}$ is the area of the neighborhood.

In Figure 3 we compare the neighborhoods in Bergen that have urban built-up (on the left) with those that are suitable by our definition for built-up (on the right). At the city center, there are areas with built-up that our algorithm would declare as unsuitable, while at the outskirts the opposite is the case. This is well in line with our theoretical framework: Higher rental prices in the city core make it attractive to invest in built-up even if building cost are higher than at alternative plots outside of the city.

Figure 3: Bergen observed urban area vs potential built-up land



Note: The figure shows neighborhoods within the circumference of the metropolitan area of Bergen. On the left areas in black indicate neighborhood with urban build up. On the right areas in black indicate neighborhood with a geography that is in average suitable for buildup.

The geography surrounding a neighborhood is likely to affect urban density indirectly through other channels as well: Mountains affect the sunniness of a neighborhood, the view, and transport costs to the CDB. To isolate these mechanisms from the effect of building land limitations, we measure these characteristics and include them as control variables in our regressions.

In particular, we generate the variable **sun hours** as the sunshine hours at equinox based on the surrounding terrain and longitude and latitude. Sunshine is an important amenity in the cross-city literature (Albouy and Lue, 2015), while for example data from New Zealand have shown that an extra daily hour of sunlight raises house prices by 2.3% (Fleming et al., 2018). In Norway, light in the winter is particularly precious and our data show that in Norwegian cities some of the neighborhoods close to the cite center belong literally to the dark side of town. This is different at the outskirts of the cities with more sunny places (see Appendix).

We compute both **distance to the ocean** in *km* as the crow flies, as well as well **ocean view**, which is fulfilled, if more than 8 points on the ocean surface - approximately half a *sqkm* of ocean - are on average visible from the neighborhood. With these variables, we follow the real estate literature that has studied the effect of natural amenities on individual house prices for a long time (Davies, 1974, Nelson, 1972), see for instance Benson et al. (1998) and Bourassa et al. (2004) on ocean view and Lee and Lin (2018) as well as Carlito and Saiz (2019) on proximity to the ocean.[1] Our data show that close proximity to the ocean is not always securing a view on the ocean (see Appendix).

3.4 Distance to the CBD

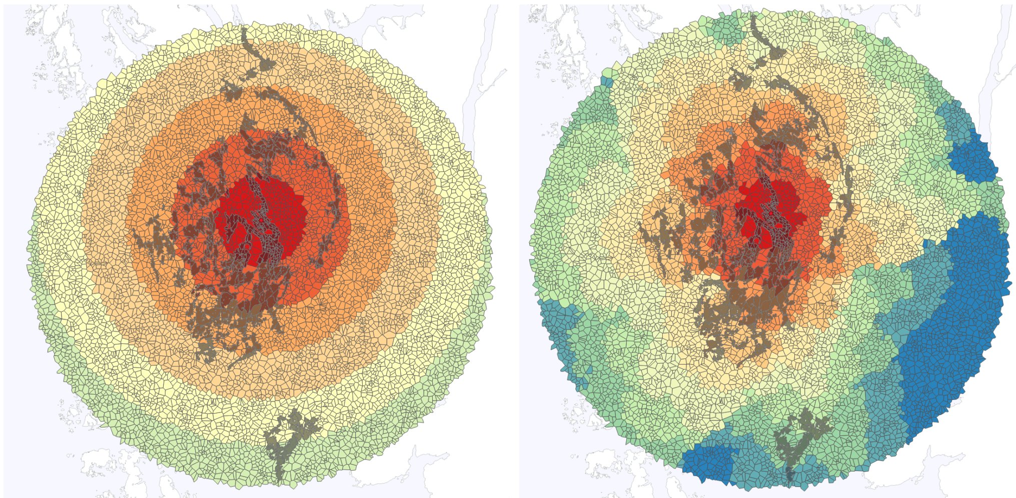
Distance to the CBD is a key determinant of urban density in all classical Alonso-Muth-Mills stlye models and their extensions because it constitutes commuting costs for households (Brueckner, 1987, Davies, 1974). The goal of our analysis is to study geography as a source of within-city density heterogeneity in addition to distance from the CBD, so it is important how we measure the latter.

In particular, we measure distances through the geography-based shortest travel path. Abstaining from actual travel distances on existing roads due to a policy bias, we calculate travel path based on the incline of the terrain (see Appendix for more detail). With this we further ensure that our measures of average ring geography do not capture differences in commuting distances. The comparison of the shortest travel path distance (right) with the

[1]Hypothetically, ocean view is one reason why hilly neighborhoods are empirically correlated with high incomes in many cities, the so-called 'Beverly Hills effect', see for instance Ye and Becker (2019).

Euclidean distance (left) in Figure 4 for the different rings around the CBD of Bergen shows that controlling for the impact of geography on commuting distances seem relevant.

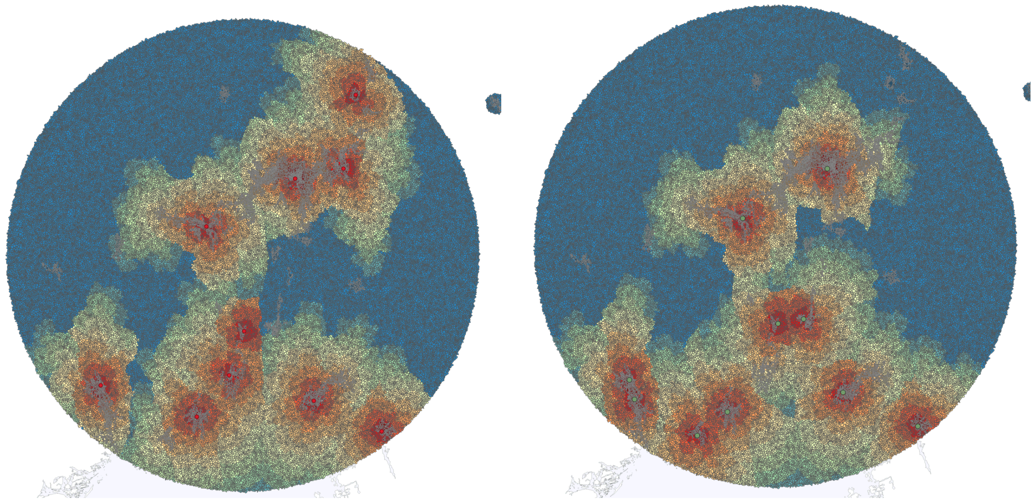
Figure 4: Bergen: Distance from the rings to the CBD



Note: The figure shows neighborhoods within the circumference of the metropolitan area of Bergen. Color from red to blue indicates in increasing order the distance to the CBDs in 5km intervals. On the left: Euclidean distances. On the right: Distances based on the shortest path given the terrain. Gray borders indicate neighborhoods with urban built-up.

This leads to the question how to adequately define the CBD in the first place. We will use two different measures of the CBD: (i) based on the density of cafés recorded in the OpenStreetMaps data, (ii) based on the World Port Index. Assuming that where people work they have to consume food and beverages, implies that a high density of cafes signal high levels of business activity. This is also in line with recent work linking cafés and restaurants as endogenous amenities to the city center (Aguiar and Bils, 2015, Baum-Snow and Hartley, 2017). Yet, one potential drawback of café density as an indicator of the CBD is precisely the endogeneity. This is why our alternative indicator of the CBD relies on ports. In Norway, ports are natural harbors, and in an economy strongly driven by fishing, sea trade and more recently oil, they correlate strongly with historical city centers (Helle et al., 2006). Note that according to our definitions, large metropolitan areas are allowed to have several CBDs: In Figure 5, we see the 10 distinct CBDs in Oslo based on the café density (left) and 9 CDBS based on ports. Details on the calculation are described in the Appendix.

Figure 5: Oslo central business districts



Note: The figure shows neighborhoods within the circumference of the metropolitan area of Oslo. Color from red to blue indicates in increasing order the distance to the CBDs measured by the shortest path. On the left, CBDs are defined by café density; on the right based on port locations. Gray borders indicate neighborhoods with urban built-up.

4 Descriptive Statistics

Having gathered all the data, let us now take a look at some statistics of our final data set. The summary statistics in Table 1 are calculated across the 3506 neighborhoods in our sample. These are located in 13 urban clusters and 66 different *kommuner*. While Lillehammer is the smallest urban cluster with around 14,000 inhabitants, one CBD and 30 neighborhoods in one *kommune*, the largest cluster is Oslo with around 1,400,000 inhabitants, 10 CBDs and 2020 neighborhoods in 34 *kommuner* (for more details see Table C-1 in the Appendix).

The average neighborhood has a mean of 665 inhabitants, reflecting the fine-grained nature of our analysis. Even the largest neighborhood, Skadberg in Stavanger, has only 5725 inhabitants. The residential area of the average neighborhood is 0.27 sqkm, ranging from Østerås-Eiksmarka in Oslo with 1.5 hectares to Torgård in Trondheim with 2.6 sqkm.

Turning to urban density, we find strong variation: The average neighborhood has an urban density of 0.0041 people per sqm - or 41 people per hectare, while the most densely populated neighborhood has ten times as many people people per area (Kampen rode 5 close to the main harbour of Oslo). The average crowding is 0.011 people per sqm of floorspace, the largest crowding implies an apartment size of 20 sqm per person in Lysskar in Haugesund. Residential coverage is given by the share of the urban extent covered by the building footprint: While it is 20.6% in the average neighborhood, it goes from a mere 3.0% in Torgård in the outskirts of Trondheim to 65.6% in the center of Oslo (Uranienborg rode 6). The average building height in the average neighborhood is 1.80 floors, but the neighborhood with the highest average building height is Solfjellet in Oslo with 6.16 floors.

Comparing the names of neighborhoods of the extremes reveals the huge heterogeneity between the components of urban density. None of the neighborhoods is named twice. For example, Solfjellet is a neighborhood with many highrises but also a lot of parks and hence a low residential coverage. Kampen rode 5 is a neighborhood with a high level of residential coverage and high buildings, but crowding is close to the average. In contrast, the neighborhood with the highest level of crowding, Lysskar, is a suburb with a high share of children, small detached houses with plenty of yard space which is why urban density is even below average. This is also reflected in the correlations between urban density and its components that range from .76 to .02 (for details see Table C-2 in the Appendix). The descriptive statistics therefore indicate the need to study not just urban density but all so its components as they can deviate from one another significantly.

Our index of geography measures the share of land with geographical constraints, which is, on average 67.0%. We can see that this key explanatory variable has a lot of variation, as some neighborhoods only have a share of 18.0% of geographical constraints, while others have 100%. On average, our neighborhoods are located 10.11 km away from their CBD, measured in terrain-based travel distance.

A look at the other geographical variables yields additional insights: For example, the mean elevation of the average neighborhood is 75.6m, with the mean slope varying considerably across neighborhoods (5.68 degrees to 28.58 degrees). Equinox sunshine hours range from 6.05 to 12. We see the importance of the ocean for Norwegian settlement structures: The average distance to the ocean is 6.69 km, and 70.6% of neighborhoods have ocean view.

Finally, we turn to income, converted into 10,000s of US dollars, as well as other socio-economic and demographic variables. The average neighborhood has a yearly income of 74,000 USD, while the wealthiest neighborhood is at more than 200,000 US dollars (Sentrum 3 /rode 4, the neighborhood closest to the yacht harbour in downtown Oslo). The poorest neighborhood is Hatleberget at the outskirts of Bergen with an average income of 23,000 USD. We also include the coefficient of variation of income as an inequality indicator. There is obviously a correlation between income and average age of the neighborhood, which has a mean value of 39.7 years but varies considerably across neighborhoods. Looking at further demographics, we see that the share of the retired population (aged 62 years and above) varies from 0 to 92.3%. The share of children and teenagers (under 18) is on average 20.0%, while the share of migrants (defined as those without Norwegian nationality as well as Norwegian nationals born abroad) is on average 17.4%, but goes up to 83.9%. Finally, we include a health indicator, the number of yearly sick notes per working population.

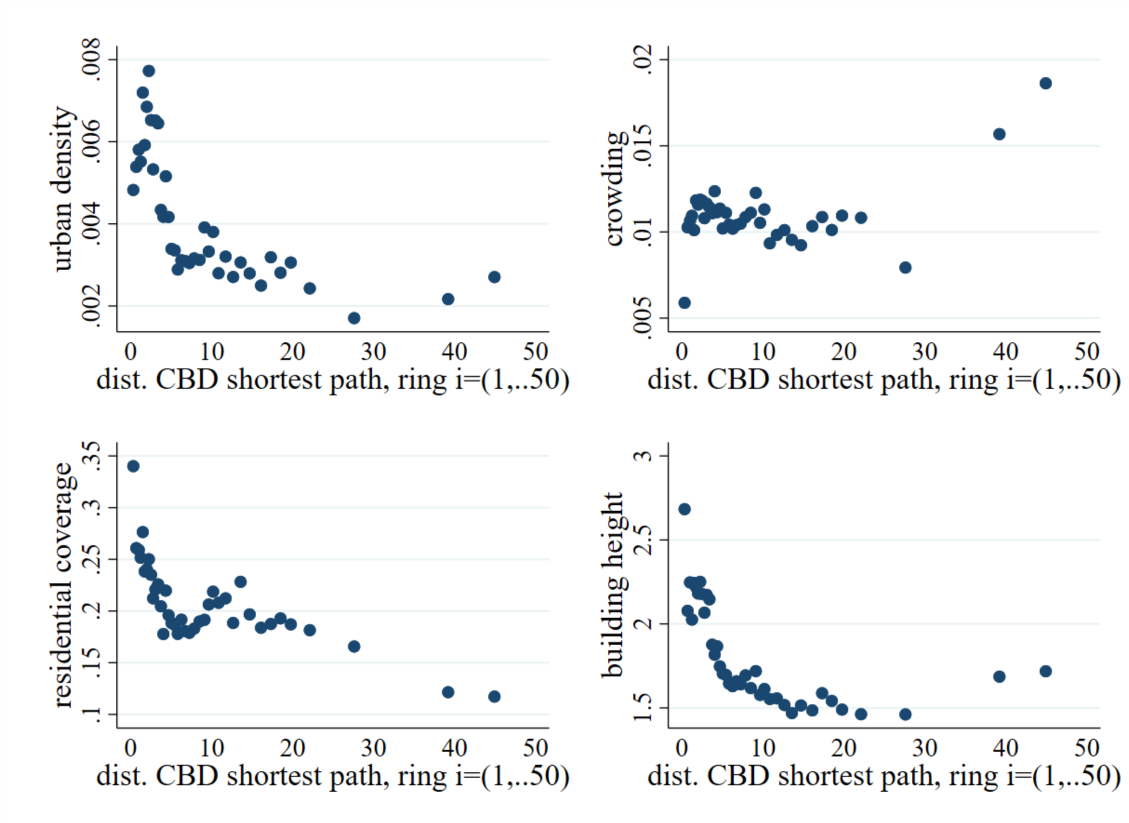
Table 1: Neighborhood Descriptive Statistics

Variable	(1) Obs	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max	(6) units
population	3,506	664.7379	491.2321	101	5725	pop
area	3,506	272077.9	270796.5	15031.62	2607329	m^2
footprint	3,506	45573.62	40034.44	1200	411400	m^2
urban density	3,506	.0040947	.0046336	.0001419	.0405212	$\frac{pop}{m^2}$
crowding	3,506	.0110969	.0070484	.001698	.0562575	$\frac{pop}{m^2}$
residential coverage	3,506	.2058337	.108413	.0300691	.6561644	*share*
building height	3,506	1.798155	.7611536	.93041	6.160101	*floors*
geography	3,506	.6703498	.1697338	.1804923	1	*share*
dist. shortest path CBD	3,506	10.11036	11.72687	.0996686	87.69476	*km*
elev mean	3,506	75.56864	61.41623	1.771931	451.5524	*m*
slope mean	3,506	4.497416	3.167414	.0567857	28.58431	*grad*
slope COV	3,506	.4245391	.1965823	.0012988	2.112248	*grad*
sun hours	3,506	10.49633	.8872226	6.052083	12	*hours*
dist. ocean	3,506	6.692824	20.91549	.0341827	135.6913	*km*
ocean view	3,506	.7065031	.455429	0	1	
income p.c.	3,506	7.417467	1.712979	2.337588	20.25347	$\frac{10.000}{pop}\$$
income p.c. cov	3,506	.8945699	.4558725	.358167	12.44372	$\frac{10.000}{pop}\$$
age	3,506	39.72397	5.537227	23.74138	82.91525	*years*
age cov	3,506	.5624136	.0566396	.1334791	.7497294	*years*
retired	3,506	.1871159	.098329	0	.9322034	*share*
kid	3,506	.2005931	.065887	0	.4419831	*share*
migrant	3,506	.1741686	.1214481	0	.8385461	*share*
sick leave. p.w.c.	3,506	1.076515	.2754443	.1228861	2.361702	*share*

Note: Descriptive statistics are for the final sample that is limited to actual neighborhoods within a 50km radius to the nearest CBD. There are 13 urban clusters, 25 CBDs and 66 *kommuner*.

The main goal of our paper is to examine whether building land limitations can be one explanatory factor behind the inner-city heterogeneity in urban density. One other well-known factor is distance to CBD. Apart from the theoretical treatment in the Alonso-Muth-Mills models and its extensions, there are numerous empirical papers finding population density to be a downward-sloping function of distance to the CBD (Batty and Longley, 1994, Bertaud and Malpezzi, 2014, Zielinski, 1980). Figure 6 shows bin scatter plots of gradients of neighborhood urban density - and its components, with distance to the CBD calculated based on the shortest path given the terrain. In fact, while gradients have been discussed extensively for density, this is - to our knowledge - the first time such gradients are being studied for the density components of crowding, residential coverage and building height.

Figure 6: Distance Gradients



Note: The figure displays without controls the binscatter plot of urban density and the distance to the CBD measured by the shortest path given the terrain.

For overall density, we receive an exponential decay pattern in line with Bertaud and Malpezzi (2014) and other works. Interestingly, this pattern appears to be mostly driven by the building height component (lower right panel), which decays notably in a similar way with distance to the CBD. Decreasing building height is a prediction from the standard Alonso-Muth-Mills model (Brueckner, 1987), but it has rarely been empirically verified in this way.[2]

[2] Ahlfeldt and Barr (2020) provide downward-sloping building height gradients for New York City and Chicago based on high-rise data from the Emporis database.

As regards crowding (upper right panel), the pattern is less clear. There might be a slight downward slope up to the 30th ring around the CBD, while the few observations further away are not supportive of this hypothesis and point towards an increase in crowding. Finally, for residential coverage (lower left panel), we find again a downward slope, which, however, does not appear monotonous. We conclude that, with distance to the CBD, urban density and most of its components decrease, but this behavior is not uniform and other factors than distance might play a role. This leads us to our empirical analysis about building land limitations induced by geography.

5 Estimation Strategy

In order to study the effect of within-city building land limitations arising from geography on within-city heterogeneity in urban density, we estimate the following equation:

$$ln(\mathbf{\Gamma}_{irck}) = \beta_1 \cdot ln(g_{rck}) + \beta_2 \cdot ln(x_{irck}) + \mathbf{\Sigma}_{irck} + \rho_r + \kappa_k + \zeta_c + \epsilon_{irck}, \tag{17}$$

where $ln(\mathbf{\Gamma}_{irck})$ is the log of the vector of urban the density measures discussed in Section 3.2 in neighborhood i, ring r (in Euclidean 1 km spacing), *kommune* k and city c; $ln(g_{rck})$ is the measure for the level of restrictions imposed by geography on built-up in ring r as discussed in Section 3.3, $ln(x_{irck})$ is the distance to the CBD as discussed in Section 3.4, $\mathbf{\Sigma}_{irck}$ is a vector of additional geography- and socio-demographic controls, ρ_r is a ring fixed effect, κ_k is a *kommune* fixed effect, ζ_c is a city fixed effect and ϵ_{ij} is the error term.

In line with Propositions 1, we expect building-land limitations induced by geography to have a positive effect on density, so that we should have $\beta_1 > 0$. This should hold irrespective of whether we use total density or its components (see Proposition 2.)

The additional geography-based controls are included to make sure that our measures of building- land limitations across a specific ring do not proxy simply neighborhood level geography. Controlling for the neighborhood level of average elevation ($ln(elev\ mean)$), the slope of the terrain ($ln(slope\ mean)$) and its coefficient of variation ($ln(slope\ COV)$) helps us to avoid an omitted variable bias: These properties are driven by the average ring geography but directly affect neighborhood built-up as they increase building costs. In the same way, we also use distance from the CBD as measured by the shortest path through the terrain in order to rule out that the distance effect would incorrectly be attributed to geography.

Moreover, as described in Section 3.3, we control for the average hours of sunshine ($ln(sun\ hours)$), distance to the coast ($ln(dist\ ocean)$) ocean view ($view$). This captures a possible geographical effect on density which does not work through building land limitations but through the amenity aspects brought about by geographical features.

As motivated in Section 2.3, we also include socio-demographic controls, such as income, age and population composition to take care of the empirical correlation between income, life-cycle and desirable geographical amenities.

Finally, we control for a large set of fixed effects mostly with the aim of accounting for any influence by the political economy of building regulations. Most of the regulation in Norway happens either on the city or *kommune* level (Kommunal- og Moderniseringsdepartementet, 2008). Neighbourhood residents only have a very limited scope to influence new built-up. Most indirect effects of regulation should therefore be captured by the *kommune* and city fixed effects. We also include ring-specific fixed effects, mainly to account for the fact that rings mechanically increase in size with distance to the CBD. We do not want this potential size effect to interfere when measuring relative built-up limitation induced by ring geography. To account for different levels of urban density across urban clusters of different size, we use city level fixed, defined for neighborhoods with the same CBD. This allows us also to separate the local effect of ring geography from the overall effect that comes with a different geography across the entire city.

6 Results

6.1 Main results

Table 2 contains our main results, namely the effects of geographical constraints on density (column 1) and its components (columns 2-4). In this very parsimonious specification without additional controls, we find that geographical constraints increase density, an effect which in this specification is statistically significant at the 10% level. A 10% increase in the share of geographical constraints raises urban density by 2.76%. We also note a positive sign for all of its three components, although only the effect on building height is significant at the 99% level. A doubling in the share of geographical constraints leads to a 12% increase in average building height (which at the mean would be 0.2 floors, around half a meter). Overall, these effects are in line with the predictions of our model, see Proposition 1. The positive response of residential coverage to geographical constraints can be interpreted in the light of households' trade-off between private yard space and public open spaces, see Proposition 2.

Table 2: Neighborhood Urban Density vs. Ring Geography

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.276*	0.063	0.090	0.123***
	(0.145)	(0.120)	(0.084)	(0.036)
ln(dist. shortest path CBD)	0.022	0.266***	-0.165***	-0.078**
	(0.034)	(0.041)	(0.031)	(0.031)
Constant	-4.907***	-4.750***	-1.150***	0.993***
	(0.176)	(0.120)	(0.062)	(0.040)
Observations	3,506	3,506	3,506	3,506
R-squared	0.441	0.281	0.504	0.507
Kommune, CBD & Ring FE	YES	YES	YES	YES

Note: The table reports regression results of eq. 17. CBD is determined based on café density. CBD, *kommune* and ring fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

Interestingly, we do not find a negative effect of distance on density. In part this is due to the ring level fixed effects. When omitting them, effects are significant and confirm our observations on the gradients summarised in Figure 6. In terms of the components, we find that building height still decreases with distance to the CBD, in line with theory. We also note a highly statistically significant decrease in residential coverage. Yet, there is a strong increase in crowding with distance to the CBD, which drives down the overall effect of density. We observe this with and without applying ring level fixed effects. One might explain this behavior of density with reference to Brueckner (1983), who argues that crowding and residential coverage might move into opposite directions if households consider apartment size and yard space as substitutes.[3]

From this specification without geographic and demographic controls, we conclude that there is evidence of geographical constraints driving up urban density and its components. The goodness of fit of our regression is high in comparison to the related literature, suggesting

[3]Brueckner (1983) writes: "Under the Cobb-Douglas assumptions, yard space per dwelling is always increasing in x, while floor space per dwelling may be increasing, constant, or decreasing in x depending on the relationship between production and utility function parameters. Note that since intuition suggests that floor and yard space will in fact be substitutes rather than complements, the type of peculiar attribute behavior found in this example is a conceivable outcome in real-world cities."

that heterogeneity in geography, distance to the CBD, and the fixed effects alone can explain a substantial proportion of inner-city differences in density and its components. In particular, the building height component can be explained very well (R^2 of 50.7%), while for crowding other factors seem to play a part (R^2 of 28.1%).

Table 3: Neighborhood Urban Density vs. Ring Geography with Geographic Controls

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.298**	0.055	0.143*	0.099***
	(0.114)	(0.089)	(0.077)	(0.034)
ln(dist. shortest path CBD)	0.049	0.265***	-0.152***	-0.065**
	(0.034)	(0.037)	(0.026)	(0.025)
ln(elev mean)	-0.026	0.061*	0.029	-0.116***
	(0.055)	(0.034)	(0.026)	(0.028)
ln(slope mean)	-0.075**	-0.064**	0.024	-0.035
	(0.033)	(0.028)	(0.024)	(0.022)
ln(slope COV)	-0.402***	-0.177***	-0.154***	-0.072***
	(0.038)	(0.031)	(0.021)	(0.022)
ln(sun hours)	0.466	-0.694***	1.195***	-0.034
	(0.512)	(0.199)	(0.343)	(0.188)
ln(dist ocean)	0.067	0.031	-0.010	0.045**
	(0.048)	(0.025)	(0.019)	(0.022)
ocean view	0.020	-0.028	0.023	0.025
	(0.047)	(0.048)	(0.032)	(0.027)
Constant	-6.325***	-3.429***	-4.284***	1.388***
	(1.324)	(0.523)	(0.864)	(0.438)
Observations	3,506	3,506	3,506	3,506
R-squared	0.515	0.324	0.544	0.553
Kommune, CBD & Ring FE	YES	YES	YES	YES

Note: The table reports regression results of eq. 17. CBD is determined based on café density. CBD, *kommune* and ring fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

In Table 3 we include geographical controls. While the overall regression fit improves, the magnitude of our main coefficient estimates stays mostly the same compared to the first, more parsimonious specification. In fact, the statistical significance increases: The positive effect of geography-induced building land limitations on urban density is now statistically significant at the 95%. The positive effect of residential coverage is weakly significant at the 90% level and that of building height at the 99% level. The coefficient of the distance variable also remains overall unaffected by the inclusion of the geographic controls. Looking at the control variables themselves, we note, inter alia, that elevation has a negative effect on building height. Yet, it is slope rather than elevation - and in particular the coefficient of variation of slope - which has a strong and negative effect on urban density and all of its complements. Neighborhoods where the terrain is very uneven have less crowding, less residential coverage and shorter buildings. As regards the amenities, we note that sunshine hours decrease crowding - possibly an income effect - and increase residential coverage. Distance to the ocean increases building height, while ocean view itself, after controlling for all the other variables has no significant effect on density.

In Table 4, we expand the set of controls even further by including socio-demographic controls. All the effects of geography and distance on density remain qualitatively unchanged. For example, the coefficient estimate of the effect of geographical constraints on urban density is now 0.240 compared to 0.298 in Table 3, both statistically significant at the 95% level. We also note some interesting effects of the socio-economic and demographic variables on density: Urban density strongly decreases with income per capita, with a 10 % increase in income per capita decreasing urban density by 5.96 %. This works mainly through crowding rather than building height. We also note a negative and significant effect of mean age on urban density. In areas with more older rather than younger people, apartments are larger and residential coverage is lower. It is important to refrain from interpreting the coefficients of these socio-demographic variables in a causal way, as the relation between density and socio-economic variable is known to be highly endogenous. We merely include them into this specification to ensure that our main result, the positive effect of geography-induced land limitations on urban density is robust to a large number of controls.

Next, we consider a different definition of the CDB: While our main regression relied on the café density, we now use the CBD definition based on ports (see Section 3.4). As Table 5 shows, this leaves the main results of geography and distance on density and its components unchanged.

In Online Appendix D, we conduct a set of further robustness tests. Our main results about the positive effects of geography on density and its components are robust to (i) dropping all neighborhoods within a 5 km radius of the CBD, (ii) dropping all neighborhoods

Table 4: Neighborhood Urban Density vs. Ring Geography with Socio-Demographic Controls

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.240**	0.030	0.159**	0.051**
	(0.092)	(0.080)	(0.070)	(0.024)
ln(dist. shortest path CBD)	0.049*	0.230***	-0.148***	-0.032**
	(0.028)	(0.032)	(0.023)	(0.012)
ln(elev mean)	0.076*	0.089***	0.026	-0.039**
	(0.040)	(0.025)	(0.029)	(0.018)
ln(slope mean)	-0.045	-0.037	0.016	-0.024*
	(0.029)	(0.032)	(0.025)	(0.013)
ln(slope COV)	-0.355***	-0.160***	-0.157***	-0.038**
	(0.029)	(0.027)	(0.017)	(0.016)
ln(sun hours)	0.740*	-0.525**	1.150***	0.115
	(0.037)	(0.024)	(0.021)	(0.014)
ln(dist ocean)	-0.015	-0.024	-0.003	0.012
	(0.041)	(0.025)	(0.022)	(0.021)
ocean view	0.054	0.021	0.005	0.028
	(0.048)	(0.048)	(0.027)	(0.018)
ln(income p.c.)	-0.596***	-0.653***	0.077	-0.021
	(0.124)	(0.099)	(0.085)	(0.050)
ln(income p.c. cov)	-0.207***	-0.124***	0.012	-0.095***
	(0.048)	(0.043)	(0.040)	(0.016)
ln(age mean)	-0.048***	-0.005	-0.015	-0.029***
	(0.016)	(0.011)	(0.011)	(0.007)
ln(age cov)	-0.246	-0.128	0.291	-0.408**
	(0.495)	(0.354)	(0.331)	(0.190)
ln(retired)	1.472**	0.076	0.286	1.110***
	(0.712)	(0.503)	(0.506)	(0.237)
ln(kid)	-2.002***	0.994**	-0.770	-2.226***
	(0.695)	(0.438)	(0.469)	(0.407)
ln(migrant mean)	-0.152	-0.310	-0.266	0.423***
	(0.173)	(0.257)	(0.229)	(0.092)
ln(sick notes p.c.)	0.175***	0.247***	-0.133***	0.061**
	(0.061)	(0.059)	(0.043)	(0.023)
Observations	3,506	3,506	3,506	3,506
R-squared	0.572	0.392	0.556	0.701
Kommune, CBD & Ring FE	YES	YES	YES	YES

Note: The table reports regression results of eq. 17. CBD is determined based on café density. CBD, Constant, *kommune* and ring fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

Table 5: Neighborhood Urban Density vs. Ring Geography with CBDs based on ports

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.307***	0.048	0.158***	0.101***
	(0.111)	(0.101)	(0.053)	(0.037)
ln(dist. shortest path CBD)	-0.086	0.149**	-0.175***	-0.060***
	(0.059)	(0.059)	(0.054)	(0.019)
ln(elev mean)	-0.006	0.056**	0.040	-0.102***
	(0.060)	(0.027)	(0.039)	(0.026)
ln(slope mean)	-0.110***	-0.059**	-0.001	-0.050***
	(0.041)	(0.025)	(0.026)	(0.017)
ln(slope COV)	-0.411***	-0.174***	-0.162***	-0.075***
	(0.041)	(0.031)	(0.021)	(0.021)
ln(sun hours)	0.385	-0.728***	1.151***	-0.037
	(0.570)	(0.241)	(0.359)	(0.162)
ln(dist ocean)	0.061	0.044*	-0.022	0.039*
	(0.049)	(0.023)	(0.026)	(0.023)
ocean view	0.033	-0.024	0.035	0.022
	(0.040)	(0.040)	(0.036)	(0.028)
Constant	-6.294***	-3.229***	-4.394***	1.329***
	(1.484)	(0.595)	(0.900)	(0.391)
Observations	3,306	3,306	3,306	3,306
R-squared	0.510	0.309	0.530	0.554
Kommune, CBD & Ring FE	YES	YES	YES	YES

Note: The table reports regression results of eq. 17. CBD is determined based on ports. CBD, *kommune* and ring fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

farther away than 10 km from the CBD, (iii) leaving out ring fixed effects,[4] (iv) merging neighborhoods with the same *kommune* in the same ring. All this ensures us that our results are not driven by specifics of the city center, the outskirts or the administrative processes behind the definition of a *grunnkrets*.

6.2 Implications

Having established the effect of geography-induced building land restrictions on density and its components, we are now going to set our results into a broader context. Urban density is thought to affect a number of socio-economic outcomes (see for example Brownstone and Thomas, 2013, Ciccone and Hall, 1996, Larsson, 2014). In the meta-study by Ahlfeldt and Pietrostefani (2019), elasticities of cross-city density and various outcome variables are provided: For instance, density is associated with both higher wages (elasticity of 4%) and higher wage inequality (elasticity of 3.5%), a higher mortality risk (elasticity of 9%) and higher subjective well-being (elasticity of 0.4%). Yet, all these elasticities we are aware of have been computed at the cross-city level. With our neighborhood-level data, we are now in a position to study the association of inner-city density with various outcome variables which we can also observe at the neighborhood level. This allows us to investigate to what extent the cross-city patterns of density and its covariates hold within cities. When looking at the following results, one should be careful not to interpret the associations as causal effects and rather see them as associations in the vein of (Ahlfeldt and Pietrostefani, 2019). To keep estimates simple and comparable to those of the literature, we estimate elasticities without any other controls than the *kommune* fixed effects.

In Table 6 we see that urban density is associated with lower income p.c. (elasticity of 6.9%), which stands against the positive elasticity between density and wages found by the literature in the cross-city setting. Panel B reveals that this result is driven by crowding and building height, which have a highly statistically significant and negative association with income per capita. Similarly, we find a negative elasticity between urban density and income inequality of 10.8%, while the consensus elasticity from the cross-city literature is positive. This suggests that different economic mechanisms are at play at the inner-city than at the cross-city level. While an in-depth analysis is beyond the scope our own study, the productivity-enhancing effects of density (Ciccone and Hall, 1996, Rosenthal and Strange,

[4] While the signs of the effects of geography on density are unaltered, the omission of ring fixed effects leads to a strongly negative effect of distance on density, in line with the standard model and the gradients in Figure 6.

2004) can be thought to play a larger role at the city level, while neighborhood-level density is obviously also influenced by sorting and residential choice (Albouy and Lue, 2015, Kuminoff et al., 2013).

Table 6: Elasticities of Inner-City Density and Its Components with Outcomes

Depend.Var:	(1) ln(income p.c.)	(2) ln(income p.c. cov)	(3) ln(age)	(4) ln(age cov)	(5) ln(migrant share)	(6) ln(sick notes)
Panel A: Urban density						
ln(urban density)	-0.069***	-0.108***	-0.006	-0.032***	0.181***	-0.000
	(0.008)	(0.019)	(0.010)	(0.009)	(0.028)	(0.026)
constant	1.576***	-0.807***	3.636***	-0.771***	-0.899***	0.039
	(0.048)	(0.113)	(0.056)	(0.049)	(0.167)	(0.150)
R^2	0.386	0.142	0.113	0.091	0.261	0.179
Panel B: The component urban density						
ln(crowding)	-0.091***	-0.172***	-0.025***	0.011*	0.051	0.095***
	(0.020)	(0.032)	(0.009)	(0.006)	(0.042)	(0.028)
ln(residential cover.)	0.047**	0.060***	-0.023**	-0.000	-0.004	-0.105***
	(0.021)	(0.022)	(0.009)	(0.006)	(0.090)	(0.027)
ln(building height)	-0.251***	-0.298***	0.065	-0.189***	0.819***	0.004
	(0.019)	(0.033)	(0.041)	(0.009)	(0.102)	(0.068)
Constant	1.770***	-0.721***	3.485***	-0.432***	-2.154***	0.297***
	(0.064)	(0.120)	(0.033)	(0.028)	(0.172)	(0.107)
R^2	0.477	0.217	0.139	0.277	0.343	0.249
Observations	3,506	3,506	3,506	3,506	3,506	3,506
Kommune FE	YES	YES	YES	YES	YES	YES

Note: The table reports regression results outcome variables on, respectively, density (Panel A) or its individual components (Panel B). Standard errors are clustered at the *kommune* level. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

As regards further variables, we see that age has a negative association with the average crowding of a neighborhood, while density overall has a negative elasticity with the age covariance. This suggests that in dense neighborhoods, particularly those with high building height, inhabitants are, ceteris paribus, of similar age. Life-cycle based housing decisions, with families with children moving to less dense suburbs, might play a role (Andersen, 2011, Kim et al., 2005). We also note a strongly significant and positive elasticity of 18.1% between urban density and the migrant share, which again is driven by building height. Finally, we look at health outcomes. While the literature points to a positive elasticity between cross-city density and mortality, we can analyze sick notes per working population at the neighborhood level. The elasticity between density and sick notes is nearly zero (see column 6), but its component reveal two highly significant and opposite effects: Residential coverage is negatively associated with the number of sick notes, but crowding exhibits a strongly positive elasticity (9.5%). Infectious diseases might play a role here, along the lines of Rocklöv and Sjödin (2020), who link the spread of covid-19 to urban density.

Taken together, our elasticity analysis shows that understanding urban density and its effects is important for policymakers. Density might have different associations with socio-economic outcome variables at the neighborhood than at the cross-city level and these might be driven by particular density components. With our paper we have shed light into how geography-induced building-land limitations and distance to the city center determine density. Being aware of these mechanisms, policymakers can shape urban density with a view to the socio-economic outcomes.

7 Conclusion

Urban density varies strongly within cities. While the theoretical and empirical literature has mostly focused on distance to the CBD as the main explanatory factor of density variation, we discuss the role of local geography in its implication for available building land. Exploiting fine-grained geo-spatial data at the neighborhood level from Norway, we are able to show a positive effect of geographical build-up constraints on urban density. This result is robust to various different specifications and supported by a theoretical framework as a motivation.

By combining geographical data with building footprints and high-resolution elevation data, this is - to our knowledge - also the first paper to split urban density into its components of crowding, building height and residential coverage at the neighborhood level. We provide evidence that all three components increase as response of geographical constraints, with the effect on building height strongest. In addition, we analyze how the density components behave as a function of distance to the CBD. Both the gradients and the regression results with controls suggest that building height and residential coverage react in a more uniform way than crowding. The behavior of crowding might be explained by the trade-off between apartment size and yard space discussed by Brueckner (1983) in his theoretical model. This calls for further empirical research in this direction.

From a policy perspective our findings allow for the first time to make a prediction on how building land restrictions affect urban density and its components. Our study overcomes the bias resulting from political economy already influencing observed urban density. We do so by using the high exogenous variation in geography limiting built-up in Norway, as well as the fine-grained data that allows us to apply fixed effects for neighborhoods influenced by the same building regulations. Our findings indicate that if policy makers aim to increase urban density in parts of a city they can do so by regulating the existence of open public spaces, for example by dedicating space to parks [5].

One of the limitation of our study is that we do not disentangle the supply and demand side effect of building land limitations. From our current empirical set-up we only identify the overall effect of geography induced by built-up limitations. It is an important avenue for future work to disentangle the effect of building limitations that arises form the demand for open public spaces and the supply of available built-up land.

Proost and Thisse (2019, p.615) call it ”surprising” that so few papers have worked with building heights, ”given the importance of the subject matter.” We hope that our approach of deriving high-resolution building height data will open the door to many more applications on urban densities, its components, effects and determinants in cities around the world.

[5] Note that for US cities there is evidence that parks can become a public bad in the presence of high levels of crime (Albouy et al., 2020). If such effects persist over a long period, they might alter the effect of open space on urban density.

References

Aguiar, M. and M. Bils (2015). Has Consumption Inequality Mirrored Income Inequality? *American Economic Review 105*, 2725–2756.

Ahlfeldt, G. and J. Barr (2020). The Economics of Skyscrapers: A Synthesis. Cesifo working paper 8427-2020.

Ahlfeldt, G. and D. McMillen (2018). Tall Buildings and Land Values: Height and Construction Cost Elasticities in Chicago, 1870–2010. *Review of Economics and Statistics 100*(5), 861–875.

Ahlfeldt, G. and E. Pietrostefani (2019). The Economic Effects of Density: A Synthesis. *Journal of Urban Economics 111*, 93–107.

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015). The Economics of Density - Evidence from the Berlin Wall. *Econometrica 83*(6), 2127–2189.

Albouy, D., P. Christensen, and I. Sarmiento-Barbieri (2020). Unlocking Amenities: Estimating Public Good Complementarity. *Journal of Public Economics 182*, 104–110.

Albouy, D. and B. Lue (2015). Driving to Opportunity: Local Rents, Wages, Commuting, and Sub-Metropolitan Quality of Life. *Journal of Urban Economics 89*, 74–92.

Albouy, D. and B. Stuart (2014). Urban Population and Amenities: The Neoclassical Model of Location. NBER Working Paper No. 19919.

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent.* Harvard University Press.

Andersen, H. (2011). Motives for Tenure Choice During the Life Cycle: The Importance of Non-Economic Factors and Other Housing Preferences. *Housing, Theory and Society 28*, 183–207.

Angel, S., P. Lamson-Hall, and Z. G. Blanco (2019). Anatomy of Density I: Six Measurable Factors that Together Constitute Urban Density. NYU Marron Institute of Urban Management Working Paper No. 43.

Batty, M. and P. Longley (1994). *Fractal Cities: A Geometry of Form and Function.* San Diego, CA and London: Academic Press.

Baum-Snow, N. and D. Hartley (2017). Accounting for Central Neighborhood Change, 1980-2010. Federal Reserve Bank of Chicago Working Paper WP 2016-09.

Benson, E., J. Hansen, A. Schwartz, and G. Smersh (1998). Pricing Residential Amenities: The Value of a View. *Journal of Real Estate Finance and Economics 16*(1), 55–73.

Bertaud, A. and S. Malpezzi (2014). The Spatial Distribution of Population in 57 World Cities: The Role of Markets, Planning, and Topography. Manuscript, University of Wisconsin-Madison.

Bourassa, S., M. Hoesli, and J. Sun (2004). What's in a View? *Environment and Planning A 38*(8), 1427–1450.

Brownstone, D. and F. Thomas (2013). The Impact of Residential Density on Vehicle Usage and Fuel Consumption: Evidence from National Samples. *Energy Economics 40*, 196–206.

Brueckner, J. K. (1983). The Economics of Urban Yard Space: An "Implicit Market" Model for Housing Attrubutes. *Journal of Urban Economics 13*, 216–234.

Brueckner, J. K. (1987). The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model. *Handbook of Regional and Urban Economics 2*, 821–845.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (1999). Why is Central Paris Rich and Downtown Detroit Poor?: An Amenity-Based Theory. *European Economic Review 43*, 91–107.

Burchfield, M., H. Overman, D. Puga, and M. Turner (2006). Causes of Sprawl: A Portait from Space. *Quarterly Journal of Economics 121*(2), 587–633.

Carlito, G. and A. Saiz (2019). Beautiful City: Leisure Amenitities and Urban Growth. Federal Reserve Bank of Philadelphia Working Paper 19-16.

Ciccone, A. and R. Hall (1996). Productivity and the Density of Economic Activity. *American Economic Review 86*(1), 54–70.

Davies, G. (1974). An Econometric Analysis of Residential Amenity. *Urban Studies 11*, 217–225.

Duranton, G. and D. Puga (2015). Urban Land Use. In G. Duranton, V. Henderson, and W. Strange (Eds.), *Handbook of Regional and Urban Economics*, pp. 467–560. Amsterdas: Elsevier.

Fischel, W. (2004). An Economic History of Zoning and a Cure for its Exclusionary Effects. *Urban Studies 41*, 317–340.

Fleming, D., A. Grimes, L. Lebreton, D. Maré, and P. Nunns (2018). Valuing Sunshine. *Regional Science and Urban Economics 68*, 268–276.

Glaeser, E. and J. Gyourko (2018). The Economic Implications of Housing Supply. *Journal of Economic Perspectives 32*(1), 3–30.

Glaeser, E. and M. Kahn (2004). Sprawl and Urban Growth. In V. Henderson and J. Thisse (Eds.), *Handbook of Regional and Urban Economics*, pp. 2481–2527. Amsterdam: Elsevier.

Green, R., S. Malpezzi, and S. Mayo (2005). Metropolitan-Specific Estimates of the Price Elasticity of Supply of Housing, and Their Sources. *American Economic Review 95*(2), 334–339.

Harari, M. (2020). Cities in Bad Shape: Urban Geometry in India. *American Economic Review* (forthcoming).

Harari, M. and M. Wong (2018). Slum Upgrading and Long-run Urban Development: Evidence from Indonesia. Manuscript University of Pennsylvania.

Helle, K., F.-E. Eliassen, J. E. Myhre, and O. S. Stugu (2006). *Norsk Byhistorie: Urbanisering Gjennom 1300 År*. Pax.

Henderson, V., T. Regan, and A. Venables (2019). Building the City: Urban Transition and Institutional Frictions. Oxford Economics Series Working Papers 891.

Hilber, C. and W. Vermeulen (2016). The Impact of Supply Constraints on House Prices in England. *The Economic Journal 126*(591), 358–405.

Jedwab, R., N. Johnson, and M. Koyama (2020). Medieval Cities Through the Lens of Urban Economic Theories. Iiep working paper 2020-9.

Kim, T., M. Horner, and R. Marans (2005). Life Cycle and Environmental Factors in Selecting Residential and Job Locations. *Housing Studies 20*, 457–473.

Kommunal- og Moderniseringsdepartementet (2008). Lov om Planlegging og Byggesaksbehandling (LOV-2008-06-27-71).

Krumm, R. (1980). Neighbourhood Amenities: An Economic Analysis. *Journal of Urban Economics 7*, 208–224.

Kuminoff, N. V., V. K. Smith, and C. Timmins (2013). The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets. *Journal of Economic Literature 51*, 1007–1062.

Larsson, J. (2014). The Neighborhood or the Region? Reassessing the Density-Wage Relationship Using Geocoded Data. *Annals of Regional Science 52*, 367–384.

Lee, S. and J. Lin (2018). Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income. *Review of Economic Studies 85*, 663–694.

Liu, C., S. Rosenthal, and W. Strange (2018). The Vertical City: Rent Gradients and Spatial Structure. *Journal of Urban Economics 106*, 101–122.

Mahan, B., S. Polansky, and R. Adams (2000). Valuing Urban Wetlands: A Property Price Approach. *Land Economics 76*, 100–113.

Mills, E. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *American Economic Review 57*(2), 197–210.

Murphy, A. (2018). A Dynamic Model of Housing Supply. *American Economic Journal: Economic Policy 10*(4), 243–267.

Muth, R. (1969). *Cities and Housing: The Spatial Patterns of Urban Residential Land Use*. University of Chicago Press.

Nelson, R. (1972). Housing Facilities, Site Advantages and Rent. *Journal of Regional Science 12*(2), 249–259.

Proost, S. and J.-F. Thisse (2019). What Can Be Learned from Spatial Economics? *Journal of Economic Literature 57*(3), 575–643.

Rocklöv, J. and H. Sjödin (2020). High Population Densities Catalyse the Spread of COVID-19. *Journal of Travel Medicine 27*(3), 1–2.

Rosenthal, S. and W. Strange (2004). Evidence on the Nature and Sources of Agglomeration Economics. In J. Henderson and J. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2119–2171. Elsevier North Holland.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics 125*, 1253–1296.

Shertzer, A., T. Twinam, and R. Walsh (2018). Zoning and the Economic Geography of Cities. *Quarterly Journal of Economics 105*, 20–39.

Turner, M. (2005). Landscape Preferences and Patterns of Residential Development. *Journal of Urban Economics 57*, 19–54.

Ye, Y. and C. Becker (2019). Moving Mountains: Geography, Neighborhood Sorting, and Spatial Income Segregation. Manuscript.

Zielinski, K. (1980). The Modelling of Urban Population Density: A Survey. *Environment and Planning A 12*(2), 135–154.

A Theory Appendix

This appendix derives the theoretical model which we use in the paper as the framework for our empirical analysis. We introduce geographical constraints on land suitability into the standard Alonso-Muth-Mills style urban economic model. In particular, let $G \in (0, 1)$ be the spectrum of geography-based land properties, ranging from 0 (perfectly suitable for building) to 1 (completely unsuitable for building). While G will influence building supply, our model also features a corresponding component on the demand side, households' preference for open space or recreational area. Put differently, they derive a disutility from a high share of built-up area b.

A.0.1 Demand Side

Households receive an income y, live in different rings with distance x from the city center and have to pay a transport cost τ to get to their jobs there. As in the standard model, they derive utility from the numeraire consumption good c and housing q, which is measured in square meters and costs the rental price p. The new feature is households' disutility from the degree of built-up b within the city ring they live in:

Assumption 1. *Consumers derive a disutility from the share of built-up area b*

$$\frac{\partial v}{\partial b} < 0 \tag{(A-1)}$$

The built-up in an area depends on x as well a second exogenous component, geography g ,in a given ring. Remember that G denotes the land-plot specific geographical constraints that builders face, while g denotes the overall geography within a ring. While g can differ between different rings of the city, it is independent of x and purely determined by nature. Households maximize utility

$$v(c(x, g), q(x, g), b(x, g))) = u \tag{(A-2)}$$

The budget constraint is

$$y = t \cdot x + p(x, g) \cdot q(x, g) + c \cdot 1 \tag{(A-3)}$$

To ease notation we drop dependencies from now on. Utility maximization leads to the following first-order condition:

$$\frac{\partial v}{\partial q} = p \cdot \frac{\partial v}{\partial c} \tag{(A-4)}$$

To equalize utility across the city regardless of distance to the CBD requires the total differential of eq. (A-2) with respect to x to equal zero:

$$\frac{\partial v}{\partial c} \cdot \frac{\partial c}{\partial x} + \frac{\partial v}{\partial q} \cdot \frac{\partial q}{\partial x} + \frac{\partial v}{\partial b} \cdot \frac{\partial b}{\partial x} = 0 \tag{(A-5)}$$

Making use of eq. (A-3) allows to rewrite eq. (A-5) as

$$\frac{\partial v}{\partial c} \cdot \left(-t - q\frac{\partial p}{\partial x} - p\frac{\partial q}{\partial x}\right) + \frac{\partial v}{\partial q} \cdot \frac{\partial q}{\partial x} + \frac{\partial v}{\partial b} \cdot \frac{\partial b}{\partial x} = 0 \tag{(A-6)}$$

Dividing eq. (A-5) by $\frac{\partial v}{\partial c}$ and plugging in eq. (A-4) yields

$$-t - q\frac{\partial p(x)}{\partial x} - p\frac{\partial q}{\partial x} + p \cdot \frac{\partial q}{\partial x} + \frac{\frac{\partial v}{\partial b}}{\frac{\partial v}{\partial c}} \cdot \frac{\partial b}{\partial x} = 0. \tag{(A-7)}$$

The second and third term cancel out, so that we can solve for the dependence of p on x:

$$\frac{\partial p}{\partial x} = -\frac{1}{q(x)} \cdot \left(t - \frac{\frac{\partial v}{\partial b}}{\frac{\partial v}{\partial c}} \cdot \frac{\partial b}{\partial x}\right) \tag{(A-8)}$$

Compared to the standard model, where $\frac{\partial p}{\partial x}$ is unambiguously negative at first sight, we have an additional term involving the dependence of the built-up share b on x. In fact, eq. (A-8) is similar in spirit to the modeling of amenities in Brueckner et al. (1999).

Not only do households living far from the CBD have to be compensated for the transport costs, but they also derive disutility from built-up. We still have to determine how built-up varies with x: To derive $\frac{\partial b}{\partial x}$, we have to study the supply side. Then we will be able to return to eq. (A-8) and determine the sign of $\frac{\partial p}{\partial x}$

Before doing so, we can already examine how p varies with g, keeping x fixed. The first-order condition of households' utility maximization problem with respect to g yields

$$\frac{\partial v}{\partial c} \cdot \frac{\partial c}{\partial g} + \frac{\partial v}{\partial q} \cdot \frac{\partial q}{\partial g} + \frac{\partial v}{\partial b} \cdot \frac{\partial b}{\partial g} = 0 \tag{(A-9)}$$

From the budget constraint eq. (A-3) we get

$$\frac{\partial c}{\partial g} = -\frac{\partial p}{\partial g} \cdot q - p \cdot \frac{\partial q}{\partial g} \tag{(A-10)}$$

Plugging eq. (A-10) and eq. (A-4) into eq. (A-9), and dividing by $\frac{\partial v}{\partial c}$, we obtain:

$$-\frac{\partial p}{\partial g} \cdot q - p \cdot \frac{\partial q}{\partial g} + p \cdot \frac{\partial q}{\partial g} + \frac{\frac{\partial v}{\partial b}}{\frac{\partial v}{\partial c}} \cdot \frac{\partial b}{\partial g} = 0 \tag{(A-11)}$$

Again, the second and third term cancel out. Solving for $\frac{\partial p}{\partial b}$ then yields

$$\frac{\partial p}{\partial g} = \frac{\frac{\partial v}{\partial b}}{\frac{\partial v}{\partial c}} \cdot \frac{\partial b}{\partial g} \cdot \frac{1}{q} \tag{(A-12)}$$

How rental prices vary with geography again depends on how geography affects the built-up share, which is determined on the supply side. As in eq. (A-8), we need to determine $\frac{\partial b}{\partial g}$ before we know the sign of $\frac{\partial p}{\partial g}$ with certainty.

A.0.2 Supply Side

As in the standard model, building firms compete for land L and use capital K to build houses with a concave production function H that is homogenous of degree one. In particular, concavity $\frac{\partial^2 H(L,K)}{\partial K^2} < 0$ implies that higher buildings are increasingly more expensive to build. As in the standard model, we normalize by dividing by L and will work with $h = \frac{H}{L}$. Note that this means that "developers are indifferent to the value of L; the size of housing complexes is indeterminate" (Brueckner, 1983, p.219). The capital-land-ratio $S = \frac{K}{L}$ is an "index for building height" (Brueckner, 1987). Let us now include land-plot-specific geographical constraints G on land suitability into the production function h: $h(S, G)$.

Assumption 2. *Geographical constraints decrease building output and make capital less productive in the building production function:*

$$\frac{\partial h(S,G)}{\partial G} < 0; \frac{\partial^2 h}{\partial S \partial G} < 0 \tag{(A-13)}$$

As usual, capital is rented at an exogenously given rate i. Building firms lease land at a rate r, which depends on location x and geographical constraints G. Firms' profit is then given by

$$\begin{aligned} \Pi &= p(x,b) \cdot H - i \cdot K - r(x,G,b) \cdot L & \text{((A-14))} \\ &= L \cdot \Big(p(x,b) \cdot h(S,G) - i \cdot S - r(x,G,b)\Big) & \text{((A-15))} \end{aligned}$$

We assume that builders do not consider their impact on total built-up b when deciding to build a house on a piece of land by setting $S > 0$. Each individual firm believes their effect on b to be marginal and therefore not influencing p or r. Given this, the first order condition of eq. (A-14) with respect to building height S is

$$\frac{\partial \Pi}{\partial S} = L \cdot \Big(p(x,b) \cdot \frac{\partial h(S,G)}{\partial S} - i\Big) = 0 \tag{(A-16)}$$

To identify how built-up varies with geographical constraints, we use the total differential of eq. (A-16) with respect to G:

$$p \cdot \Big(\frac{\partial^2 h(S,G)}{\partial S^2} \cdot \frac{\partial S}{\partial G} + \frac{\partial^2 h(S,G)}{\partial S \partial G}\Big) = 0 \tag{(A-17)}$$

$$\tag{(A-18)}$$

We can solve for $\frac{\partial S}{\partial G}$ and see how a change in geography affect the optimal investment in land and thereby building height:

$$\frac{\partial S}{\partial G} = -\underbrace{\frac{\partial^2 h}{\partial S \partial G}}_{<0} \cdot \underbrace{\Big(\frac{\partial^2 h(S,G)}{\partial S^2}\Big)}_{<0}^{-1} < 0 \tag{(A-19)}$$

The first factor is negative because of capital's diminishing return in building process, while the second factor is negative as geographical constraints make building more expensive (Assumption 2). We conclude that building heights get shorter with more geographical constraints on the given land plot.

Let us now combine the individual decisions of building firms in order to analyze their effect on total built-up. For this, we first assume that the distribution of G leads to the density function $f(G,g)$. The parameter g increases the frequency of land plots with high geographical obstacles in the given ring. Hence we assume that $\frac{\partial f(G,g)}{\partial g} < 0$. With this we can now derive b by looking at the marginal $\tilde{G}$ for which a construction firm would be indifferent to build houses $S > 0$:

$$p \cdot \frac{\partial h(0,\tilde{G})}{\partial S} - i = 0. \tag{(A-20)}$$

On all land with geographical constraints $G < \tilde{G}$, there will be built-up. Therefore we can write

$$b = f(\tilde{G}, g) \tag{(A-21)}$$

From eq. (A-20) we can implicitly determine the relation between rental prises, built-up, distance and level of geographical obstacles.

A.0.3 Comparative Statics

Comparative statics on x

We now can continue with our analysis of comparative statics related to changes in x. From eq. (A-21) we obtain $\frac{\partial b}{\partial x}$ as

$$\frac{\partial b}{\partial x} = \frac{\partial f}{\partial \tilde{G}} \cdot \frac{\partial \tilde{G}}{\partial x} \tag{(A-22)}$$

Hence, the change of total built-up with distance x from the CBD depends on the effect of x on the marginal $\tilde{G}$ for which a construction firm would be indifferent to build. The total differential of eq. (A-20) with respect to x is

$$\frac{\partial p}{\partial x} \cdot \frac{\partial h(0, \tilde{G})}{\partial S} + p \frac{\partial^2 h(0, \tilde{G})}{\partial S \partial G} \cdot \frac{\partial \tilde{G}}{\partial x} = 0 \tag{(A-23)}$$

From this we can solve for

$$\frac{\partial \tilde{G}}{\partial x} = -\frac{\partial h(0, \tilde{G})}{\partial S} \cdot \frac{1}{p} \cdot \Big(\frac{\partial^2 h(0, \tilde{G})}{\partial S \partial G}\Big)^{-1} \frac{\partial p}{\partial x} \tag{(A-24)}$$

Plugging eq. (A-24) into eq. (A-22), we obtain

$$\frac{\partial b}{\partial x} = -\frac{\partial f}{\partial \tilde{G}} \cdot \frac{\partial h(0, \tilde{G})}{\partial S} \cdot \frac{1}{p} \cdot \Big(\frac{\partial^2 h(0, \tilde{G})}{\partial S \partial G}\Big)^{-1} \cdot \frac{\partial p}{\partial x} \tag{(A-25)}$$

This allows us to return to eq. (A-8) and determine the sign of the effect of distance on rental prices. Plugging eq. (A-25) into eq. (A-8) yields

$$\begin{aligned} \frac{\partial p}{\partial x} &= -\frac{1}{q} \cdot \Big[t - \frac{\frac{\partial v}{\partial b}}{\frac{\partial v}{\partial c}} \cdot \frac{\partial f}{\partial \tilde{G}} \cdot \Big(-\frac{\partial f}{\partial \tilde{G}} \cdot \frac{\partial h(0, \tilde{G})}{\partial S} \cdot \frac{1}{p} \Big(\frac{\partial^2 h(0, \tilde{G})}{\partial S \partial G}\Big)^{-1} \cdot \frac{\partial p}{\partial x}\Big] \\ &= -t \cdot \Big[q + \underbrace{\frac{\partial v}{\partial b}}_{<0} \cdot \underbrace{\Big(\frac{\partial v}{\partial c}\Big)}_{>0}^{-1} \cdot \underbrace{\frac{\partial f}{\partial \tilde{G}}}_{>0} \cdot \underbrace{\frac{\partial h(0, \tilde{G})}{\partial S}}_{>0} \cdot \frac{1}{p} \cdot \underbrace{\Big(\frac{\partial^2 h(0, \tilde{G})}{\partial S \partial G}\Big)}_{<0}^{-1}\Big]^{-1} < 0 \end{aligned} \tag{(A-26)}$$

Although more variables are involved than in the standard model, we still obtain that rental prices unambiguously decrease with distance to the CBD. Not only does the transport cost play a role, but also the disutility households derive from high built-up areas, which again depends on the geographical constraints.

With this result, we can now examine the effect of distance x on building height S and floor space q. Starting with building height, we take the total differential of eq. (A-16) with respect to x:

$$\frac{\partial p}{\partial x} \cdot \frac{\partial h(S,G)}{\partial S} + p \cdot \frac{\partial^2 h(S,G)}{\partial S^2} \cdot \frac{\partial S}{\partial x} = 0 \tag{(A-27)}$$

As in the standard model, we obtain

$$\frac{\partial S}{\partial x} = -\frac{\partial h(S)}{\partial S} \cdot \frac{1}{p} \cdot \underbrace{\left(\frac{\partial^2 h(S,G)}{\partial S^2}\right)}_{<0}^{-1} \cdot \underbrace{\frac{\partial p}{\partial x}}_{<0} < 0 \tag{(A-28)}$$

Geographical constraints on land do not alter the results that building heights decrease towards the outskirts of the city.

Moving on to floor space consumption, we assume that demand can be described by a non-further specified function depending negatively on price, as in the standard model. This yields the result that apartment size increases with distance from the CBD:

$$\frac{\partial q}{\partial x} = \underbrace{\eta}_{<0} \cdot \underbrace{\frac{\partial p}{\partial x}}_{<0} > 0 \tag{(A-29)}$$

This is equivalent to a decrease in crowding $\frac{1}{q}$.
Also note that we can determine the effect of x on total built-up share from eq. (A-25):

$$\frac{\partial b}{\partial x} = -\underbrace{\frac{\partial f}{\partial \tilde{G}}}_{>0} \cdot \underbrace{\frac{\partial h(0,\tilde{G})}{\partial S}}_{>0} \cdot \frac{1}{p} \cdot \underbrace{\left(\frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G}\right)}_{<0}^{-1} \cdot \underbrace{\frac{\partial p}{\partial x}}_{<0} < 0 \tag{(A-30)}$$

Comparative Statics on g

We now continue our analysis of comparative statics with changes in the ring-specific geographical parameter g. As the components of density all depend on the rental price g, we need to know how p varies with g. eq. (A-12) involves the term $\frac{\partial p}{\partial g}$ which we can now determine further. The total differential of eq. (A-20) with respect to g is

$$\frac{\partial p}{\partial g} \cdot \frac{\partial h(0,\tilde{G})}{\partial S} + p \cdot \frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G} \cdot \frac{\partial \tilde{G}}{\partial g} = 0 \tag{(A-31)}$$

From this we obtain how the marginal $\tilde{G}$, where construction firms are just willing to build, depends on the geography parameter g

$$\frac{\partial \tilde{G}}{\partial g} = -\frac{\partial p}{\partial g} \cdot \frac{\partial h(0,\tilde{G})}{\partial S} \cdot \frac{1}{p} \cdot \Big(\frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G}\Big)^{-1} \tag{(A-32)}$$

Taking the total differential of eq. (A-21) with respect to g and plugging in eq. (A-32) yields

$$\frac{\partial b}{\partial g} = \frac{\partial f}{\partial \tilde{G}} \cdot \frac{\partial \tilde{G}}{\partial g} + \frac{\partial f}{\partial g} = \frac{\partial f}{\partial \tilde{G}} \cdot \Big[-\frac{\partial p}{\partial g} \cdot \frac{\partial h(0,\tilde{G})}{\partial S} \cdot \frac{1}{p} \cdot \Big(\frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G}\Big)^{-1}\Big] + \frac{\partial f}{\partial g} \tag{(A-33)}$$

Returning to eq. (A-12), we can plug in eq. (A-33):

$$\frac{\partial p}{\partial g} = \frac{\frac{\partial v}{\partial b}}{\frac{\partial v}{\partial c}} \cdot \Big[\frac{\partial f}{\partial \tilde{G}} \cdot \Big(-\frac{\partial p}{\partial g} \cdot \frac{\partial h(0,\tilde{G})}{\partial S} \cdot \frac{1}{p} \cdot \Big(\frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G}\Big)^{-1}\Big) + \frac{\partial f}{\partial g}\Big] \cdot \frac{1}{q} \tag{(A-34)}$$

$$= \underbrace{\frac{\partial f}{\partial g}}_{<0} \cdot \Big[q \cdot \underbrace{\frac{\partial v}{\partial c}}_{>0} \underbrace{\Big(\frac{\partial v}{\partial b}\Big)^{-1}}_{<0} + \underbrace{\frac{\partial f}{\partial \tilde{G}}}_{>0} \cdot \underbrace{\frac{\partial h(0,\tilde{G})}{\partial S}}_{>0} \cdot \frac{1}{p} \cdot \underbrace{\Big(\frac{\partial^2 h(0,\tilde{G})}{\partial S \partial G}\Big)^{-1}}_{<0}\Big]^{-1} > 0 \tag{(A-35)}$$

Now that we have established that rental prices increase with the geography constraint parameter g, we can analyze the effects of g on building height, floor space consumption and total built-up. Starting with building height, we take the total differential of eq. (A-16) with respect to g:

$$\frac{\partial p}{\partial g} \cdot \frac{\partial h(S,G)}{\partial S} + p \cdot \frac{\partial^2 h(S,G)}{\partial S^2} \cdot \frac{\partial S}{\partial g} = 0 \tag{(A-36)}$$

This yields

$$\frac{\partial S}{\partial g} = -\frac{\partial h(S)}{\partial S} \cdot \frac{1}{p} \cdot \underbrace{\Big(\frac{\partial^2 h(S,G)}{\partial S^2}\Big)^{-1}}_{<0} \cdot \underbrace{\frac{\partial p}{\partial g}}_{>0} > 0 \tag{(A-37)}$$

On average, geographical constraints in the given ring increase building heights.

Moving on to the floor space consumption q, we assume that demand can be described by a non-further specified function depending negatively on price, as in the standard model. This yields a negative effect

$$\frac{\partial q}{\partial g} = \nu \cdot \frac{\partial p}{\partial g} < 0 \tag{(A-38)}$$

Crowding $\frac{1}{q}$ increases.

Finally, looking at total built-up we can plug eq. (A-12) into eq. (A-33) and obtain

$$\frac{\partial b}{\partial g} = \underbrace{\frac{\partial f}{\partial g}}_{<0} \cdot \big(1 + \underbrace{\frac{\partial f}{\partial \tilde{G}}}_{>0} \cdot \underbrace{\frac{\partial h(0, \tilde{G})}{\partial S}}_{>0} \cdot \underbrace{\frac{\partial v}{\partial c}}_{>0} \Big(\underbrace{\frac{\partial v}{\partial b}}_{<0}\Big)^{-1} \big(p \cdot q \cdot \underbrace{\frac{\partial^2 h(0, \tilde{G})}{\partial S \partial G}}_{<0}\big)^{-1}\big)^{-1} < 0 \qquad \text{((A-39))}$$

B Data Appendix

This data appendix complements Section 3 in the paper by providing more detail on the data and the process of data preparation, including additional illustrations.

B.1 Unit of Observation: Neighborhood

We define neighborhood as the *residential built-up area of an urban grunnkrets*. The administrative boundaries of the *grunnkretser* reflect the status in 2013. Our measure of residential built-up area is based on high resolution remote sensing data (10m $\times$ 10m) indicating residential built-up which we extract from the European Settlement Map (ESM) of 2015.[1] We calculate a buffer with a radius of 50m around all areas with residential built-up (DN=255). To distinguish between consumption and production, we deliberately do not account for built-up that is clearly industrial and hence labeled with DN=250 in the ESM data. We drop all non-contiguous areas where the ratio of built-up area to urban area is less than one to ten. The latter step removes small standalone housing settlements far away from the agglomeration. We do so because *grunnkretser* at the fringe of urban agglomerations are typically more extended than in the core and might include very small remote house groups that we do not think belong to the urban agglomeration.

To identify urban residential built-up areas we match the residential built-up area with Global Human Settlement Settlement Model (GHS-SMOD) grid data from 2015. The GHS-SMOD data indicates on a 1 km $\times$ 1 km grid level the 'degree of urbanization' as defined by EUROSTAT. We keep all residential built-up areas that are within or adjacent to areas that are classified as urban in the GHS-SMOD data (DN>20). This includes the urban core but also urban peripheral areas like suburbs.

This way, we arrive at the 3507 neighborhoods to be included in our final sample. At the *grunnkrets* level, we also have access to the number of residents and their average socioeconomic characteristics. We extract this data from population and income register of Norway. It contains information on the pretax yearly income of all residents of all *grunnkretser* that have more then 100 inhabitants in the year 2013. The minimum restriction is imposed by the authorities to secure privacy regulations. It does not constitute a problem for our analysis because it only leads to the loss of a handful of *grunnkretser* in the Northern Finnmark region which are far from any urban area and therefore not in our sample.

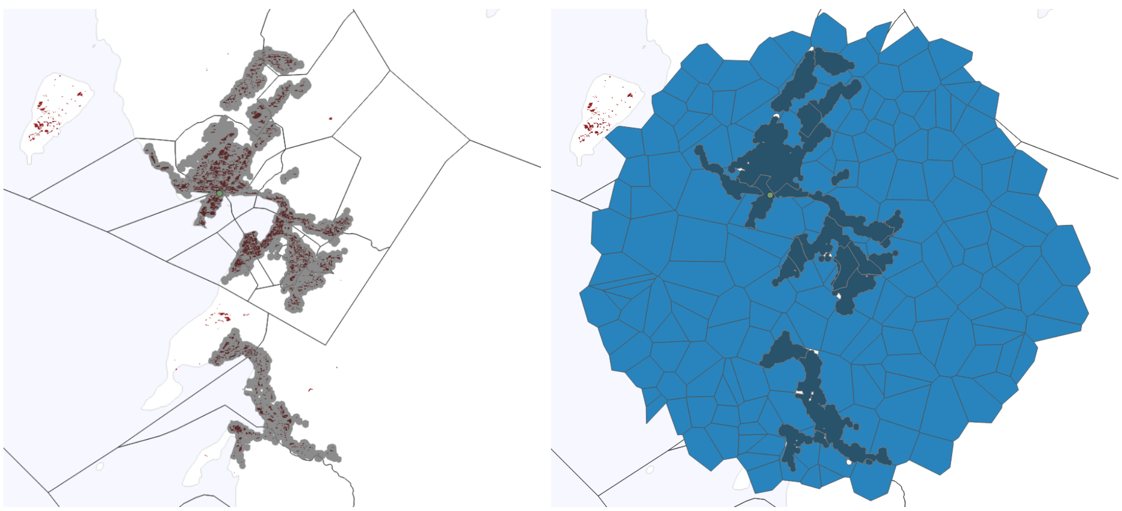
[1] ESM data is derived via machine learning applied to the Copernicus VHR_IMAGE_2015 data set based on the satellite images from Pleiades, Deimos-02, WorldView-2, WorldView-3, GeoEye-01 and Spot 6/7 ranging from 2014 to 2016.

B.2 Geography

When analyzing the suitability of a certain neighborhood for built-up, we have to avoid the circulatory argument of looking only at built-up areas. In the construction of g, we therefore work with both the original neighborhoods and artificial neighborhoods outside of the original neighborhoods. For this we randomly locate points within the circumference of the urban residential built-up areas and generate Voronoi polygons with similar geometric properties as the actual neighborhoods.

To illustrate our approach, Figure B-1 presents the case of Hammerfest (though not part of the final data set because it is not classified as urban agglomeration). On the left, the urban residential built-up areas (gray) and built-up (red) are displayed; on the right, one can see the artificial Voronoi neighborhoods.

Figure B-1: Hammerfest neighbourhoods and artificial neighborhoods



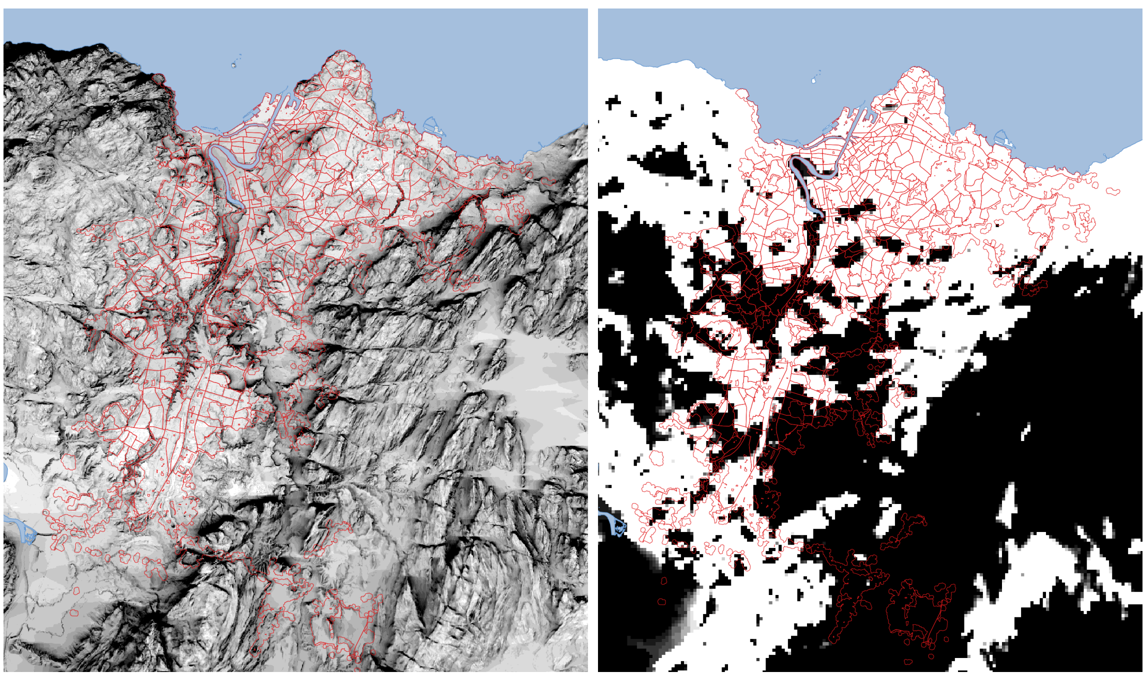
Note: The figure shows neighborhoods within the circumference of the small town of Hammerfest. On the left, the black lines indicate original *grunnkrets* borders, gray areas urban built-up areas and red areas actual built-up. The picture on the right displays the artificial Voronoi neighborhoods in blue and the actual neighborhoods defined by the urban built-up areas of the *grunnkrets* in gray.

In the following, we present illustrations of the *sun hours*, *distance to the ocean* and *ocean view* variables, which we include as controls in our regression.

Sun hours are calculated as the sunshine hours at equinox based on the surrounding terrain and longitude and latitude. The left panel of Figure B-2 shows the sunshine hours for Trondheim on a black-white scale ranging from areas with less than 5 hours (black) to those with full 12 hours (white). We can see the strong inner-city variation in sunshine determined by the terrain.

We measure *distance to the ocean* in *km* as the crow flies. Furthermore, we calculate the mean number of points located on the ocean surface with a spacing of 500m that are directly visible from the neighborhood, given the topography on the way to the ocean. We say that a neighborhood has ocean view if more than 8 points on the ocean surface (approximately half a sqkm of ocean) are on average visible from the neighborhood. We illustrate this approach for Trondheim in the right panel of Figure B-2, with white denoting ocean view and black the lack of ocean view. Comparing this figure with the left panel shows that ocean view and hours of sunshine vary considerably, given the direction of mountain lines. Moreover, close proximity to the ocean is sufficient for securing an ocean view.

Figure B-2: Sunshine hours and ocean view in Trondheim



Note: The figures show, respectively, sunshine hours and ocean view in Trondheim. Neighborhoods with urban built-up figure in red, blue areas are ocean. Left: Areas in pure black have less then 5 hours of sunshine, those in pure white 12 hours. Right: Black areas have no view of the ocean, while white areas do.

B.3 Distance to the CBD

Here we provide more information on the calculation of distance based on the shortest path through the terrain, as well as the definition of the CBD.

To calculate the shortest path, we assume that transport costs are equal to the incline of the terrain and that traveling over water has a cost equal to a 10 degree incline in a 100m $\times$ 100m raster. Comparing actual road data and shortest paths reveals that overground roads are often very close to shortest paths. Larger deviations are often associated with the extents of tunnels.

As regards the CBD, our first definition is based on the density of cafés. Using the Open Street Map data on the location of cafés, we define the gravitational center of consecutive areas that are larger than half a sqkm and have a café density of more than 5 cafés per sqkm to be a CBD. This definition allows us to define at least one CBD in all except three clusters of urban areas classified by the ESM data and our built-up data. In the cases of Halden, Haugesund and Kristiansund we had to reduce the cafe density cutoff further down to obtain at least one a CBD. In downtown Oslo we merged the CBDs that had less than 5km distance to one another. In this way, we obtain a total of 25 CBDs in all urban areas in Norway in our final sample. Most urban areas only have one CBD, but some have more and, formidably, the metropolitan area of Oslo has 10 CBDs.

In terms of the port definition of the CBD, we rely on the size of ports from the World Port Index. As the coordinates of the ports reported in the World Port Index are in some cases on land and in others on water, we unify locations using daylight satellite images by hand. Moreover, we compare pre-industrial-revolution maps of Norway with the location of ports in urban areas to prove that they are highly correlated. Hence, the location of ports capture historical - and still modern-day - CDBs. Based on the port location, we obtain 19 CBDs for all urban areas in Norway in our final sample. Most urban areas only have one CBD, but some have more and the metropolitan area of Oslo has 9 CBDs.

C Supplementary Descriptive Statistics

Here we supplement the descriptive analysis from the main text with an overview over the urban clusters as well as correlations between the variables.

Table C-1: Descriptive Statistics on Urban Clusters

Cluster name	(1) Total pop	(2) # neighborhoods	(3) # CBD (cafe)	(4) #*kommuner*
Lillehammer	14018	30	1	1
Kristiansund	17588	34	1	1
Molde	18531	26	1	1
Bodø	21233	49	1	1
Tromsø	23971	34	1	1
Haugesund	40389	92	1	2
Ålesund	43802	54	1	2
Hamar	44107	104	1	4
Kristiansand	98208	150	4	4
Trondheim	168957	312	1	3
Stavanger	211837	255	1	6
Bergen	283934	347	1	6
Oslo	1344126	2020	10	34

Table C-2: Correlations Urban Density and its Components

	(1) urban density	(2) residential coverage	(3) building height	(4) crowding
urban density	1.0000			
residential coverage	0.6269	1.0000		
building height	0.7553	0.4451	1.0000	
crowding	0.2296	-0.2999	0.0205	1.0000

D Additional Results

In this section, we repeat the main regression from eq. 17 under slightly altered specifications. In Table D-1, we drop all neighborhoods which are closer than 5 km to the CBD. This should mitigate concerns about results being driven by specifics of the city core, such as the height of historical buildings or the mixture between office and residential dwellings. It reduces the number of observations by nearly one half, but preserves the signs and magnitudes of our main estimation results. The positive effect of geography on density becomes even larger.

In Table D-2 we drop all neighborhoods which are farther away than 10 km from the CBD in order to make sure that our results are not primarily driven by the outskirts. Our results remain in place, mostly staying even at the same significance levels despite the reduced number of observation.

In Table D-3 we return to the full sample but run the regression without ring fixed effects. This leads to some interesting changes: While the effects of geography on urban density and building height remain positive and (weakly) statistically significant, the sign on crowding turns negative, but statistically insignificant. More remarkable is the effect of distance on density, which is now negative and highly significant - in line with the standard model. This helps to reconcile the negative gradient figures (without controls) with the insignificant main regression results in this respect. The clear negative association between distance and density only seems to hold in the absence of ring-specific characteristics, highlighting the various sources of inner-city heterogeneity. However, the trade-off between crowding and residential coverage seems to remain in place.

Finally, Table D-4 repeats the regression when neighborhoods from the same *kommune* and the same ring are merged. This is a robustness check against the administrative processes behind the definition of a *grunnkrets* which underlie our neighborhood unit. Although this leaves us with only 389 observations, we can replicate our main results of geography on density with similar magnitude and similar levels of significance. For distance, we again observe a strongly negative effect on density, with one source of inner-city heterogeneity reduced.

Table D-1: Neighborhood Urban Density vs. Ring Geography Without the Innermost Neighborhoods

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.589**	0.384***	0.007	0.198***
	(0.228)	(0.089)	(0.175)	(0.054)
ln(dist. shortest path CBD)	-0.236	0.082	-0.299***	-0.019
	(0.159)	(0.167)	(0.103)	(0.039)
ln(elev mean)	0.101*	0.074**	0.090**	-0.064***
	(0.060)	(0.037)	(0.043)	(0.020)
ln(slope mean)	-0.088*	-0.040	-0.006	-0.042**
	(0.049)	(0.029)	(0.031)	(0.019)
ln(slope COV)	-0.410***	-0.230***	-0.126***	-0.053***
	(0.040)	(0.035)	(0.020)	(0.018)
ln(sun hours)	0.070	-0.495*	0.784**	-0.220
	(0.564)	(0.271)	(0.359)	(0.143)
ln(dist ocean)	-0.003	0.010	-0.043*	0.030
	(0.043)	(0.026)	(0.022)	(0.019)
ocean view	-0.022	-0.060	0.015	0.024
	(0.049)	(0.051)	(0.032)	(0.025)
Constant	-6.034***	-3.939***	-3.506***	1.411***
	(1.531)	(0.605)	(1.002)	(0.418)
Observations	1,969	1,969	1,969	1,969
R-squared	0.373	0.377	0.516	0.276
Kommune, CBD & Ring	YES	YES	YES	YES

Note: The table reports regression results of eq. 17, but neighborhoods closer than 5km to the CBD are dropped. CBD is determined based on café density. CBD, *kommune* and ring fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

Table D-2: Neighborhood Urban Density vs. Ring Geography Without the Outermost Neighborhoods

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.232**	0.020	0.127*	0.085**
	(0.108)	(0.085)	(0.074)	(0.035)
ln(dist. shortest path CBD)	0.056	0.275***	-0.150***	-0.070**
	(0.034)	(0.041)	(0.026)	(0.027)
ln(elev mean)	-0.063	0.056	0.018	-0.136***
	(0.057)	(0.042)	(0.027)	(0.029)
ln(slope mean)	-0.076**	-0.083**	0.032	-0.025
	(0.036)	(0.032)	(0.027)	(0.024)
ln(slope COV)	-0.385***	-0.155***	-0.155***	-0.074***
	(0.043)	(0.031)	(0.025)	(0.026)
ln(sun hours)	0.635	-0.753***	1.323***	0.065
	(0.467)	(0.154)	(0.363)	(0.215)
ln(dist ocean)	0.113**	0.040	0.009	0.064**
	(0.046)	(0.028)	(0.019)	(0.024)
ocean view	0.062	0.013	0.033	0.016
	(0.047)	(0.046)	(0.043)	(0.026)
Constant	-6.696***	-3.324***	-4.529***	1.156**
	(1.164)	(0.409)	(0.893)	(0.498)
Observations	2,706	2,706	2,706	2,706
R-squared	0.524	0.278	0.535	0.574
Kommune, CBD & Ring FE	YES	YES	YES	YES

Note: The table reports regression results of eq. 17, but neighborhoods further away than 10 km to the CBD are dropped. CBD is determined based on café density. CBD, *kommune* and ring fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

Table D-3: Neighborhood Urban Density vs. Ring Geography Without Ring FE

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.152*	-0.019	0.103	0.067*
	(0.088)	(0.104)	(0.088)	(0.035)
ln(dist. shortest path CBD)	-0.250***	0.134***	-0.255***	-0.130***
	(0.062)	(0.030)	(0.026)	(0.020)
ln(elev mean)	-0.062	0.037	0.023	-0.122***
	(0.068)	(0.034)	(0.028)	(0.035)
ln(slope mean)	-0.084**	-0.063**	0.017	-0.038*
	(0.036)	(0.028)	(0.023)	(0.020)
ln(slope COV)	-0.428***	-0.191***	-0.160***	-0.077***
	(0.047)	(0.030)	(0.022)	(0.027)
ln(sun hours)	0.407	-0.707***	1.174***	-0.060
	(0.592)	(0.230)	(0.346)	(0.185)
ln(dist ocean)	0.069	0.038*	-0.013	0.044*
	(0.046)	(0.021)	(0.019)	(0.024)
ocean view	0.068	-0.001	0.037	0.033
	(0.051)	(0.046)	(0.036)	(0.029)
Constant	-6.371***	-3.447***	-4.329***	1.405***
	(1.479)	(0.570)	(0.865)	(0.414)
Observations	3,506	3,506	3,506	3,506
R-squared	0.487	0.296	0.536	0.542
Kommune & CBD FE	YES	YES	YES	YES

Note: The table reports regression results of eq. 17, but there are no ring fixed effects CBD is determined based on café density. CBD and *kommune* fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

Table D-4: Neighborhood Urban Density vs. Ring Geography Merged Within *Kommuner* and Rings

Depend.Var:	(1) ln(urban density)	(2) ln(crowding)	(3) ln(residential coverage)	(4) ln(building height)
ln(geography)	0.359***	0.108	0.138*	0.113**
	(0.133)	(0.131)	(0.078)	(0.046)
ln(dist. shortest path CBD)	-0.266***	0.072**	-0.202***	-0.135***
	(0.048)	(0.028)	(0.033)	(0.017)
ln(elev mean)	0.158*	0.186**	0.038	-0.066
	(0.092)	(0.084)	(0.063)	(0.041)
ln(slope mean)	-0.216*	-0.091	-0.122	-0.003
	(0.111)	(0.095)	(0.073)	(0.032)
ln(slope COV)	-0.653***	-0.313***	-0.335***	-0.005
	(0.130)	(0.089)	(0.079)	(0.050)
ln(sun hours)	-0.391	-0.612	0.049	0.172
	(1.233)	(0.751)	(0.707)	(0.300)
ln(dist ocean)	-0.085	-0.062	-0.028	0.005
	(0.072)	(0.065)	(0.044)	(0.023)
ocean view	0.070	0.026	0.020	0.024
	(0.124)	(0.109)	(0.072)	(0.038)
Constant	-5.583*	-4.268**	-1.884	0.569
	(2.979)	(1.795)	(1.738)	(0.724)
Observations	389	389	389	389
R-squared	0.633	0.659	0.768	0.697
Kommune & CBD FE	YES	YES	YES	YES

Note: The table reports regression results of eq. 17, but merging neighborhoods from the same *kommune* in the same ring. CBD is determined based on café density. CBD and *kommune* fixed effects not reported. Robust standard errors clustered on the *kommune* level. The number of urban clusters = 13, the number of CBD=25 and the number of *kommune*=66. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.