Beauty in the Eye of the Home-Owner:
Aesthetic Zoning and Residential Property Values

Thies Lindenthal[*]

This paper empirically confirms one core motivation for architectural zoning: Shape homogeneity among neighboring homes increases the value of residential buildings. Drawing on large-scale shape and transaction data, this study first develops a data-driven measure of architectural similarity, condensing three-dimensional shapes to univariate shape distributions. These algorithm-based similarity estimates are good predictors of human perceptions of shape similarity and are linked to property attributes and transaction prices. For the city of Rotterdam, a price premium of approximately 4 percent is estimated for row houses within very homogeneous ensembles over buildings facing heterogeneous neighbors.

One eminent objective of urban planning is the safeguarding of present and future property values (e.g. Levy, 2013, p. 43), which motivates not only the close regulation of the location or scope of new development but, additionally, the external appearance of new structures. Very explicitly, urban planners posit that too much or too little variation in architectural forms depress property values. The zoning ordinance for Eastchester, NY, is one of many manifestations of this wide-spread belief as it aims to prevent *"monotonous and unsightly uniformity of building development or unsightly structures of incongruous or inappropriate form that might tend to depress surrounding property values"* (Town of Eastchester NY, 2000).[1] Following this proposition, lawmakers and courts have been significantly curtailing land owner's rights to independently chose designs for new buildings and justified this stark intervention with assumed welfare gains at the neighborhood and city level (Anderson, 1960; Regan, 1990; Rubin, 1975). The far reaching limitation of architectural freedom warrants the question: Does the regulation of architectural diversity indeed bolster property values?

A wealth of studies has established solid evidence for tight land use regulations being associated with rising in property values and, simultaneously, declining levels of new construction (Glaeser and Ward, 2009; Hilber and Vermeulen, 2014; Ihlanfeldt, 2007; Koster et al., 2012; Mayer and Somerville, 2000; Quigley and Raphael, 2005). However, the extent to which the increase in prices is caused by land use regulation imposing supply constraints or by the planning process creating value directly is less understood.

* University of Cambridge, Department of Land Economy. 19 Silver Street, Cambridge, UK, CB3 9EP, UK. E-mail: *htl24@cam.ac.uk*

1 Similar regulations can be found throughout the US and also in the UK or continental Europe. Japan's zoning, in contrast, allows large variations in architectural design at lot level.

Overall, the identification of economic effects of policies regarding architectural designs and levels of homogeneity between neighboring building remains difficult at the *policy level*, since geographic and temporal variation in architectural regulation is challenging to quantify. Empirical research on the price effect of land use policies is plagued by regulation being "astonishingly vague" (Glaeser and Ward, 2009). Architectural regulation is no exception to this ambiguity. For instance, the Dutch city of Rotterdam requires new construction to adhere to "high design standards", to use "high quality materials" and to follow the "shape and appearance of surrounding buildings" (City of Rotterdam, 2016). In addition, the enforcement of architectural policies is commonly delegated to architecture review boards for case by case decisions, blurring any boundaries and regulations even more.

At the *building-level*, however, the assumed link between architectural design and the value of surrounding properties have been explored to some extent: Historic landmark buildings have been found to support the value of surrounding properties (Ahlfeldt and Maennig, 2010; Leichenko et al., 2001; Listokin et al., 1998). Similarly, dwellings designed by famous architects leads to other buildings in the vicinity being more highly-valued. For instance, homes within 50 m of a residential building by Frank Lloyd Wright in Oak Park, Illinois, enjoy a price premium of 8.5 percent over buildings further away (Ahlfeldt and Mastro, 2012). Inversely, buildings can also impair the value of surrounding buildings very directly, for instance by protecting or blocking a sought-after view.[2] At the neighborhood level, perceived beauty of the built environment is one of the main determinants of the resident's satisfaction, alongside economic factors, school quality, and the perceived opportunity of social interactions (Florida et al., 2009). This body of literature suggests that a building's architectural quality indeed creates externalities.

To our knowledge, only one study has investigated the direct effect of architectural *similarities within ensembles* of buildings on their respective sales values: For 19th century Boston, Moorhouse and Smith (1994) find that properties which look different from their neighbors sell for more. In a sample of (relatively homogeneous) rowhouses from Boston's South End, properties with facade styles different from other fronts close by carry a price premium. The observed premia for "sticking out" become smaller with each additional building in the proximity sharing the same architectural style. These findings are strong but not conclusive, since they rely on relatively few observations from one neighborhood only (and from more than 150 years ago).

---

2   An unobstructed sea view will increase property prices by 15% in Singapore (Yu et al., 2007) while positive values for viewsheds on nature and historical buildings have also been documented for Kyoto (Yasumoto et al., 2011).

The lack of empirical research on the economics of building shapes is not surprising since shape is difficult to capture quantitatively. So far, existing studies on the economic value of architecture relied on on-site inspections by experts and the classifications of styles and shapes into a limited number of categories (Asabere et al., 1989; Fuerst et al., 2011; Moorhouse and Smith, 1994; Vandell and Lane, 1989). Less palpable dimensions like the silhouette, massing, roof forms, proportions, angles or similarities to surrounding properties remain unrecorded and ignored.

This study is the first to estimate the effect of architectural homogeneity within ensembles of three or more adjacent buildings on observed sales prices using a large, city-wide dataset on the three-dimensional shapes. Using an automatic, algorithm-driven evaluation of similarity between buildings, it unlocks property-level data for *all* buildings in a city and also on the degree of architectural homogeneity between them. The full-city approach drastically increases the number of observations available for analysis compared to subsamples based on sales, mortgage originations or valuations for tax purposes only. Furthermore, the direct context of neighboring buildings can be analyzed which is not always possible when relying on subsamples instead of the universe of structures.

The remainder of the paper first refines a method to convert three-dimensional building shape data into a numerical representation that can be fed into empirical pricing equations. For the city of Rotterdam, form-related building attributes are found to vary extensively between buildings, as individualistic developer preferences, architectural creativity and amendments during the life-time of a building constitute a source of constant diversity. At the same time, economies of scale during construction, architectural preferences for harmony, overall fashion trends and zoning induce similarities between buildings.

Subsequently, we verify whether these quantitative representations of shape can help to explain recent transaction prices using a large database of housing transactions. Controlling for location and property quality, a positive effect of architectural similarity at the ensemble level on transaction prices is identified, providing empirical evidence for one core justification of architectural control: Home owners indeed pay an economically significant premium for buildings that are surrounded by similar properties.

Methodology

The shape and the hedonic configuration of a building are inextricably tangled as "form follows function" (Sullivan, 1896). The function or use of a building determines its shape and, simultaneously,

its value. A three-dimensional model of building shells not only reveals the location, type and spaciousness of a dwelling but also a wealth of hedonic attributes for the place (as in Jensen and Cowen, 1999). Trained observers might be able to estimate the year of construction from the architecture, the height of rooms from the location of windows and other element in the facade. Additional amenities like green spaces, garages or balconies are directly observable. Also, certain shapes might be perceived as more aesthetically pleasing than others and therefore carry a direct architectural premium.

The heterogeneity and multi-dimensionality of building shapes renders their classification a non-trivial challenge. Broad categories can describe roof forms, the 2D shape of the ground plates or overall dimensions. Still, classifications relying on a manageable number of categories cannot provide a finely grained view and the variation in shape within each of the classes remains high. In addition, even objective shape measures and large sample sizes do not fully differentiate the aesthetic side of architecture from its functional aspects. However, as both form and price depend on the stream of services provided by a building, we hypothesize that prices of buildings with *similar shapes* tend to have *similar prices*. Estimating pairwise shape similarities between buildings circumvents the problem of finding a meaningful classification system for property forms.

Thanks to advances in satellite and airplane based surveying technology, shape data availability is less of a limiting factor anymore. Advances in the interpretation of remotely sensed data has lead to a surge of large and spatially consistent data sets with detailed three-dimensional information at building level. New York, Paris, Singapore, Tokyo and many other cities can be explored digitally, while the municipalities of Berlin or Rotterdam openly share semantic city models[3]. So far, these models have been put to use in a wide range of research areas, including urban planning (Ranzinger and Gleixner, 1997; Wu et al., 2010), disaster management (Kwan and Lee, 2005), law enforcement (Wolff and Asche, 2009), navigation (Rakkolainen and Vainio, 2001), facility management and building information models (Nagel et al., 2009), or emission and other environmental modeling (Nichol and Wong, 2005). An investigation linking building shapes and property prices is still lacking.

Measuring shape similarity

Methods on measuring shape similarity both in 2D and 3D have been suggested in computer graphics, computer vision, biology and other disciplines. For a general review please refer to Cardone, Gupta, & Karnik (2003) or Tangelder & Veltkamp (2008). This paper builds on the *shape distribution*

3   3D city data for Berlin is available for download at http://www.businesslocationcenter.de/en/downloadportal and for Rotterdam at http://www.rotterdam.nl/rotterdam_3d.

approach put forward by Osada, Funkhouser, Chazelle, & Dobkin (2001): A large number of random points are drawn from the surface of each shape and pairwise distances between these points are calculated. The estimated probability density functions (EDF) of these distances represent building-specific *shape signatures* that can be stored and compared efficiently for large numbers of buildings. The mean of the distances is a proxy for the volume of buildings. The distributions can be normalized by dividing all pairwise distances by the average distance for each shape.

======= Insert Figure 1 about here =========

Figure 1 illustrates that differences in the shapes lead to distinct differences in the corresponding density functions. Three stylized building shapes are constructed by combining two base shapes, cubes and triangular prisms. The shape distribution of a single cube exhibits a single distinct peak while the distribution for a cuboid, formed by joining two cubes, has a long tail to the right. Adding a triangular roof to the cube changes the resulting shape distribution yet again: The "house with saddle roof" representation differs strongly from the other two examples.

While it is easy to reduce 3D objects to univariate shape distributions, it is not possible to do the reverse. The skewness of the distribution might give a rough indication of the overall compactness of a structure but backing out shape details from shape distributions is not feasible. However, similar shapes will lead to similar distributions. Intuitively, if the area between two plots of shape distributions is small, then the original shapes can be considered similar. A pairwise measure of similarity $S_{i,j}$ for shapes i and j is calculated from the respective EDFs (similar to Osada et al., 2001):

$$S_{i,j} = 1 - \int_{d=0}^{d=D} \left| edf_i(d) - edf_j(d) \right| \qquad (1)$$

Obviously, $S_{i,j} = S_{j,i}$.

Shape distributions possess several advantageous characteristics: they can be calculated for solid and non-solid 3D shapes like surfaces and 2D shapes alike and are tolerant to errors in the underlying geometries (Ohbuchi et al., 2005). This robustness is crucial when working with shape data for large numbers of buildings that have been automatically derived from areal scans and oftentimes comprised of non-solid shapes for individual buildings (Alam et al., 2013), caused by small gaps between walls or missing walls between adjacent buildings in the resulting models. In a sense, the building models that will be later used in this study are "drafty". If one printed these models on a 3D printer only few houses would be reasonably airtight. The share of non-solid building-level models derived from 3D city models has been documented to be as high as 95% (Boeters, 2013), which rules out any approach

requiring input shapes to be solid.

The accuracy and relevance of the suggested estimate of shape similarity S is first tested directly: Are buildings, that are known to have identical forms, recognized as being similar? In real cities, the most basic architectural form is probably a cube, which is also the easiest to identify using just a few geometric rules: Cube-buildings feature exactly four walls, a roof and a ground plate which are all squares of the same area. For a subset of buildings know to be cubic, the estimate of pairwise similarity S is expected to be close to 1, with 1 representing perfect identity. Across dissimilar shapes, S is hypothesized to be significantly smaller.

The mapping of shapes to shape distributions is not a bijective function. Shape distributions are invariant to rotation, mirroring and, if normalized, also to scaling (Osada et al., 2001). While a shape is converted into exactly one shape distribution, one distribution can be the shape signature of multiple 3D shapes. For example, a cube balancing on one of its corners will have exactly the same distribution as one resting flat on one face. Combining the three-dimensional similarity measure with an estimate of similarities of the 2D ground plates, estimated in the same way as S but in two dimensions only, reduces the odds of false positives when searching for similar shapes. In addition, other dimensions like the overall volume of the properties can also be (re-)introduced to account for large deviations in scale.

On a different note, human perceptions of similarity are likely to be a nonlinear function of S. For example, a decrease in S from a high 0.95 to 0.85 might change the perceived similarity of two buildings dramatically while moving from 0.35 to 0.25 might not. Translating S into a binary variable that classifies pairs of buildings as either similar or dissimilar accounts for non-linearities effectively.[4]

In this paper, a similarity matrix WS contains the pairwise similarity estimate for all $n \times n$ pairwise combinations of buildings in a sample of size n. Each element $ws_{i,j}$ is defined to be 1 if buildings i and j are sufficiently similar in shapes (high S in 3D), ground plates (high S in 2D) and volumes, or 0 otherwise:

$$ws_{i,j} = \begin{cases} 1, \text{ if S\_3D}_{i,j} > a \wedge \text{S\_2D}_{i,j} > b \wedge \dfrac{Volume_i}{Volume_j} \in [v_{low}, v_{high}] \\ 0 \end{cases} \tag{2}$$

The similarity estimate is symmetrical, as $ws_{i,j} = ws_{j,i}$. As mentioned before, the volume of a building is

4   An either/or classification also resonates well with the vocabulary available when describing similarity of shapes: We only have words for the extremes and cannot describe "somewhat similar" or other more nuanced degrees of similarity with single words.

approximated by the average distance between random points on the surface of each building.

To achieve a similar/dissimilar classification that resembles the perceptions of shape similarity by humans as closely as possible, values for the parameters *a*, *b*, v_{low} and v_{high} are selected based on a web-based survey on shape similarity. In that survey, students are repeatedly presented pairs of 3D model visualizations of buildings and asked to classify them as either "rather similar" or "rather dissimilar"[5]. Drawing from this unique dataset of *similarity perceptions*, threshold values are selected that lead to a good fit between human classifications and the algorithm based classifications in **WS**. With pre-compiled 2D- and 3D-shape signatures, a pairwise similarity matrix **WS** can be estimated fast and without consuming excessive computing resources even for large samples.

Does shape similarity translate into similarity in hedonic characteristics and property values? The economic relevance of shape similarity is investigated by linking the shape information to data from residential property transactions featuring information on sales prices and building attributes. In an ad-hoc test, all properties are assigned to 10 broad shape categories applying the k-means clustering algorithm to the shape similarity matrix **WS**. The distributions of prices and differences in hedonic attributes for properties across these clusters are compared. Any cross-cluster differences in observable hedonic characteristics can be read as evidence for the shape similarity data being correlated to similarities in hedonics also.

Alternatively, the intrinsically arbitrary classificantion into *N* categories is avoided by estimating a hedonic spatial error model (SEM) which investigates the relationship between transaction prices for single family homes and set of explanatory variables including property characteristics, the year of transaction, the location of each building and the transaction prices of similar properties in a generalized method of moments (GMM) regression:

$$\ln(P_i) = \alpha + \boldsymbol{B} \boldsymbol{X}_i + \boldsymbol{G} \boldsymbol{Year}_i + \mu_i \tag{3}$$

$$\mu = \lambda_1 \boldsymbol{W} \mu + \lambda_2 \boldsymbol{WS} \mu + \epsilon \tag{4}$$

The natural logarithms of transaction prices *P* for building *i* is explained by a vector of hedonic attributes $\boldsymbol{X}_i$ and a vector of dummies variables $\boldsymbol{Year}_i$ for the year of transaction. The vectors **B** and **G** contain regression coefficients. The error terms *μ* are correlated with one another for nearby observations and for similar shapes. The elements in the $n \times n$ spatial weight matrix **W** are defined to be 1 for all corresponding properties which are closer than 100 m and 0 otherwise. The coefficient λ_1 is

5   Details on the survey design and all response data are available from the author upon request.

expected to be positive, since properties that are geographically close share the same unobserved location amenities. In a similar spirit, the error terms of similar buildings (indicated by **WS**) are expected to be correlated as well, since they share unobserved attributes. If the coefficient of shape correlation λ_2 is found to be significant and positive, then prices paid for properties that share the same shape are correlated beyond the factors explained by hedonics, time, or location.

LeSage (2014) advises to "avoid the pitfall of multiple weight matrices" in spatial models, since, among other concerns, covariances between multiple weight matrices are restricted to be zero (LeSage and Pace, 2011). When estimating Eq. (3), alternative specifications of **WS** are therefore tested that explicitly have a correlation of zero with the spatial weight matrix **W**, circumventing any covariance restrictions.

Estimating the value of architectural homogeneity

At the neighborhood-level, architectural homogeneity in residential real estate has been traditionally associated with large-scale developments of affordable and mass produced homes. Examples are "monotonous" post-WWII home building schemes (Gartman, 2009) for returning veterans in the US or aesthetically bland suburbs where few large developers continue to produce "more of the same" (Peiser, 2014). Affluent neighborhoods, on the contrary tend to exhibit more variety in architecture.

We control for neighborhood and unobserved building quality effects by looking at homogeneity within small ensembles of rowhouses within close geographic bounds. Ensembles comprise three or more adjacent rowhouses that are identified to have (almost) identical shapes and that are therefore very likely to stem from the same development and to share very similar hedonic characteristics. Due to their close proximity, location specific amenities are also comparable within each ensemble, which ensures that all buildings from that ensemble are almost perfect substitutes. Remaining differences in upkeep and interior amenities of buildings within the ensemble are assumed to be distributed randomly.

A systematic difference between otherwise homogeneous ensemble buildings is introduced whenever the ensemble directly borders a house of distinctly different architectural shape. To illustrate, picture a row of four houses (A, B, C, D) containing an ensemble of three substitutable structures (A, B, C) next to an architecturally diverse house D. In this example, C differs from B only in terms of its location within the ensemble as differences in location and hedonic attributes are negligible. C is subject to the architecture externalities of D, while B is surrounded by homogeneous properties. Comparing transaction prices of buildings within the ensemble (B) to prices of buildings from the

periphery of the same ensemble (C), singles out the value of homogeneity: If prices *within* are higher than prices *at the periphery*, then homogeneity in architecture is preferred over shape variety.

Any price premium (or discount) for homogeneity is hypothesized to depend on the degree of architectural impairment by building D. If the shape difference between C and D is large, then any price effect is expected to be highest, while small difference matter less. Additionally, for small and affordable rowhouses ensemble effects are likely to account for a larger share of total value than for larger dwellings.

This identification approach translates into a regression estimation in which the ratio of the sales price from a periphery-of-ensemble property C over a within-ensemble building B transaction price (*PriceRatio$_{CB}$ = Price$_C$/Price$_B$*) is regressed against an intercept α and a linear combination of the shape similarity between C and D *(Similarity$_{CD}$)* and the interior floor space of C *(Size$_C$)*:

$$PriceRatio_{CB,i} = \alpha + \beta_1 Similarity_{CD,i} + \beta_2 Size_{C,i} + \beta_3 Similarity_{CD,i} * Size_{C,i}$$
$$+ \sum_{Yb} \sum_{Yc} \delta_{b,c} Y_{b,c,i} + \beta_4 IntSpaceRatio_{CB,i} + \epsilon_i \qquad (5)$$

A set of dummy variables $Y_{b,c}$ accounts for different years of sale for B and C. $Y_{b,c}$ are defined to be 1 for all pairs *i* where B was sold in year *b* and C in year *c*, -1 if B was sold in year *c* and C in year b – and 0 otherwise. The ratio of interior floorspace of C over B's interior floor space *(IntSpaceRatio$_{CB}$)* accounts for any remaining differences in the interior floorspace that might exist due to different floor plans within similar external shapes or differently used basements or attics. The β's and δ_s are regression coefficients to be estimated and the error term ε is assumed to be independently and identically distributed.

Data

This paper relies on four sources of data. First, the City of Rotterdam provides a three-dimensional model of all buildings in the city[6], which has been calculated from surface scanning data captured from helicopters in April 2010. The accuracy of the spatial data is high: At least 30 points per square meter have been scanned in the city center and 65 percent of these points are within 10 cm of the true location (95 percent within 15 cm), and the confidence intervals around height estimates are even narrower (City of Rotterdam, 2015). The virtual representation of Rotterdam is distributed in the CityGML (Level of Detail 2) format, which is an open data model for the storage and exchange of three-

6   Available for download at http://www.rotterdam.nl/links_rotterdam_3d

dimensional city information. A building's shape is defined by a set of polygons, each representing a wall, part of a roof or the ground plate. One can compare this to building a model of a house by cutting two-dimensional shapes out of cardboard and gluing them together: any structure can be approximated but very fine architectural nuances might be lost. Demarcations of buildings that share walls have been added based on land registry records (City of Rotterdam, 2015). After dropping small structures with a ground plate of less than 3 m^2, 185,914 properties remain in the database.

Second, data on residential transactions in Rotterdam is acquired through the Association of Dutch Realtors (NVM). About 70% of all transactions in the Netherlands are facilitated by members of the NVM[7]. The NVM database contains 29,948 observations for Rotterdam in the years 2006-2013. For each sale, the sales price, the exact address and a basic set of quality attributes for the property like interior floor space, dwelling type, year of construction, number of bedrooms, number of bathrooms/WC and the building's volume are recorded. The street address can be translated into geographic coordinates using the geocoding service of the Dutch land register[8]. Based on these coordinates, sales can be matched with buildings in the 3D model.

Third, the Dutch land registry maintains a national register of all buildings (*Basisregistraties Adressen en Gebouwen,* BAG) which offers information on the number of units within each building. Fourth, a web-based survey of perceived similarity is conducted.

Combining the 3D data, the sales database and the building registry gives a sample of 6,717 transactions of individual structures that contain only one residential unit and for which at least one sale has been recorded. Multi-unit buildings are excluded because their 3D shape cannot be assigned to individual sales reliably. Further, observations with extreme or potentially wrongly coded values are dropped whenever the transaction price is below 30,000 EUR or above 1 million EUR, the value for interior floor space is below 30 m^2 or above 500 m^2, a lot size above 5,000 m^2 or an estimate of the building's volume below 30 m^3 or above 5,000 m^3. The adjusted final sample is comprised of 6,126 transactions.

==== *Insert Figure 2 about here* ====

Figure 2 gives an overview of the spatial distribution of the sample within the borders of the Rotterdam municipality. The gray areas indicate all buildings from Rotterdam's 3D city map, including residential, industrial and commercial properties. The black areas represent the final sample of single

7   https://www.nvm.nl/over_nvm/english.aspx
8   More information on the geocoding webservice is available at https://www.pdok.nl/nl/service/openls-bag-geocodeerservice .

family homes for which transaction data is available in 2006-2013. Solid lines mark the official neighborhood boundaries. The majority of residential transactions can be found in the residential neighborhoods in Rotterdam proper in the east, while the west is dominated by harbor, infrastructure, warehouses and industrial properties.

Results

Shape Similarity Measures

For all buildings in Rotterdam, the 3D shape distributions and 2D shape distributions of ground plates are calculated.[9] The computation of the shape distributions for a single building takes only a fraction of a second on a contemporary PC.

To verify that the suggested shape similarity measure S holds up in a real world application, the distribution of S for buildings that are known to be similar is compared to the overall distribution of S. Cube-shaped buildings can be easily identified as they have exactly four walls, a roof and a ground plate which are all squares of the same area. For 1,229 (out of 185,914) buildings, these conditions are met reasonably well. For all pairwise combinations of cube-buildings, the average value of S is 0.95, which is close to the ideal value of 1. In contrast the distribution of S for all buildings has a mean of 0.76. It is re-assuring that the difference in means between cube and non-cube buildings is large and statistically significant (t-value: 3,828). Overall, S passes the initial test of being able to tell similar from distinct shapes.

==== *Insert Table 1 about here* ====

==== *Insert Figure 3 about here* ====

The 6,129 shape distributions displayed in *Figure 3* exhibit substantial heterogeneity, indicating a large diversity in the shapes of the single-family homes in Rotterdam. At the same time, the darker areas in the figure show a clustering around typical distributions – despite all uniqueness, building exteriors appear to be variations of a limited number of typical architectural forms.

Similarity in shapes comes with similarity in hedonic attributes (Table 2). When dividing the buildings into 10 dominant shape clusters (the dark thick traces in Figure 3) using the k-means algorithm, stark contrasts in building attributes can be observed. For instance, Cluster 8 features the most affordable transaction price (EUR 194,000), the smallest average interior floor space (101 m^2), a

---

9   Both the code to draw large numbers of random points from the exterior of a building model and the stored shape distributions are available from the author upon request.

low volume (281 m^3) and the most recent average year of construction (1975). It is comprised almost exclusively of terraced houses (99.1%). Cluster 10, in contrast, features the highest share of detached homes (11.7%), the highest average values for sales price (EUR 365,000) and volume (450 m^3), and high values for interior floor size (153 m^2). The test for equal means in a one-way layout shows that the differences in cluster means are statistically significant, with F-values of 22 and higher (num. df = 9, denom. df = 2,163). Also, housing types are not equally distributed across clusters (X^2 =3333.4, df = 36).

==== *Insert Table 2 about here* ====

Finding a link between shapes and building characteristics corroborates the underlying assumption of this paper: Shape information can be used as a proxy for observable – and more interestingly – otherwise unobservable building attributes. However, shape-related estimates remain difficult to interpret as they represent both an effect for a specific shape and jointly the contribution of unobserved hedonic variables correlated with specific shapes.

Strict zoning in combination with economies of scale in large developments of multiple units with similar designs enforce a high degree of homogeneity in buildings' forms and appearances at block or street level. This combination of strong regulations and market forces induces high levels of spatial correlation in any measure of building shape for Rotterdam. The data support this expectation: The odds of observing buildings from identical shape clusters within 100 m are 2.8 times higher than expected under the assumption of random spatial distributions. The same-shape joint count test statistics with nonfree sampling (Cliff and Ord, 1981; Upton and Fingleton, 1985) are highly significant. With such strong spatial correlations present in shape, strong spatial controls are indispensable in the subsequent analysis.

==== *Insert Table 3 about here* ====

Reassuringly, the automatic classification of buildings into similar and dissimilar pairs corresponds well with the perception of building similarity by human. Overall, 374 combinations of Rotterdam building models have been presented to human survey participants, who were then asked whether they would consider these buildings as being "pretty much the same" or "different"[10]. The automatic classification suggested in this paper can predict the human classifications well: 116 out of 125 combinations that have been classified as being "rather similar" by human respondents are also

10   An example of the comparisons presented in survey is shown in Appendix 1.

classified as similar in WS.[11] Only 9 (or 7%) are not. For pairs that are perceived as being different by humans, the match is a little lower: 193 out of 249 combinations flagged as "rather dissimilar" by humans are also considered dissimilar by the automatic classification (76%). A highly significant chi-squared statistic of 163 (with 1 degree of freedom) confirms that the automatic identification of similar buildings is highly correlated with human classifications.

==== *Insert Table 4 about here* ====

Shape distributions and property values

Table 5 presents coefficients for four independent GMM regressions. First a reduced version of Equation 3 is estimated in Model I, which explains the natural logarithm of transaction prices by dummies for the year of transaction and a traditional spatial weight matrix W (in which element $w_{i,j}$ is set to one if buildings i and j are less than 100 m apart, and 0 otherwise) only. The fit of this rudimentary model is surprisingly good (adj. R^2: 0.716) due to the fine-grained spatial weights capturing the variation in location amenities and building characteristics. The coefficient of spatial correlation, λ_w, is large (0.8) and statistically significant.

==== *Insert Table 5 about here* ====

Adding a second weight matrix WS based on shape similarity (Model II) boosts the explanatory power further. The adj. R^2 reaches 0.74, reducing the unexplained variation by 8.5% ((1-0.74)/(1-0.716)=0.915). The coefficient of similar-shape correlation, λ_{ws}, is relevant in size (0.286) and also found to be significantly different from zero. In the sample, three quarters of all variation in transaction prices can be attributed to the overall market, location and similarities in shapes – all variables that can be remotely observed without on-site inspections and which represent low-hanging fruits for mass appraisal systems.

The coefficient estimates for the hedonic variables in III and IV do not surprise: Detached homes are valued most as all other types carry significant negative discounts. Terraced houses, for instance, are about 25% more affordable. Interior floor space and volume have positive elasticities, which add up to a little below 1. The elasticity of lot size is a low 0.039. The 1960s through 1980s vintages carry a significant discount, while newer homes command a premium over historic homes built before 1906.

Interestingly, the spatial correlation coefficient λ_w is the highest (0.805) in Model I, and drops

---

11 The 98th percentiles for 3D-*S* and 2D-*S* are used as thresholds a and b, and v_{low} and v_{high} are set to 0.83 and 1.2, respectively, when calculating *WS*.

sizably (to 0.699) after controlling for shape similarities in Model II. This suggests that spatial correlations in a traditional SER model does not exclusively capture micro-location related amenities but also includes a sizable share of property attribute information. Similarly, adding hedonic control variables directly (Model III) reduces the magnitude of the spatial correlation estimate (0.637), while the "purest" spatial correlation estimate (0.609) can most likely be observed in Model IV, which includes both direct hedonic variables and the indirect controls for unobserved variables through the shape similarity weight matrix *WS*.

Including hedonic control variables reduces the coefficient for shape similarity λ_{ws}, by more than half, again indicating that form and function are correlated. Still, even with strong hedonic controls, finely-grained spatial weight matrices and strong model fits (R^2 exceeding 0.8!), shape similarity correlation estimates remain statistically significant (p-value of 0.03).

Robustness tests find that buildings of similar shapes exhibit a common structure in regression error terms even if observations are located far apart: setting all elements of *WS* for buildings that are less than 5 km from each other to 0 does not change the λ_{ws} estimates substantially. This is interesting for buyers and sellers of properties: Using shape information, one can identify relevant comparables, even if they are at the other end of town.

Finally, to rule out that the shape similarity matrix *WS* is not solely a disguised fixed effect for buildings by the same developer (which happen to have similar designs), the weight matrix is again manipulated. Assuming that buildings from different vintages have been realized by different developers, all elements of *WS* for buildings built less than 15 years from each other are set to 0. Again, the λ_{ws} estimate remains robust in magnitude. As long as shapes are similar, differences in building age do not matter when looking for comparables.

The Value of Homogeneity

For 320 ensembles of at least three similar rowhouses adjacent to a differently shaped builiding, sales prices for buildings within the ensemble and also at its periphery have been recorded in the years 2000 through 2013. For pairs of sales from the same year, the ratio of periphery-over-within ensemble sales *(PriceRatio$_{CB}$)* is on average 0.97. A relatively close standard deviation of 0.09 confirms that the quality and location characteristics of the homes are controlled for effectively. The 25th and 75th percentiles of the *IntSpaceRatio$_{CB}$* are 0.96 and 1.04, respectively.

========= *Insert Table 6 about here* =========

Table 6 presents the regression estimates of Equation 5 based on an ordinary least squares regression with robust standard errors (White's estimator). The dependent variable *PriceRatio (Price$_C$/Price$_B$)* is normalized around mean 0 and all independent variables are standardized (mean 0, SD 1). Panel (I) displays the estimated coefficients from a reduced version of Eq. 5. The negative constant indicates a 4% price discount of ensemble-buildings at the periphery versus closely comparable buildings within an ensemble.

The full model (II) confirms this negative constant and an overall discount (-3.8%) for direct proximity to a differently shaped building. In addition, *larger homes* appear to be less sensitive to the influence of difference in architecture: An one standard deviation increase in the interior floor space of the ensemble buildings offsets the negative effect already (+3.3%). Also, small rowhouses are more sensitive to the *degree of shape similarity* to neighboring structures. An one standard deviation increase in similarity to an adjacent non-ensemble building in combination with an one standard deviation decrease in size reduces the discount by 2.7%. Conversely, a small building next to a distinctly different building experiences an even higher price discount.

For about half of the ensembles, the sales sample contains information on property values for the adjacent, non-ensemble building. In a robustness test, we check if the heterogeneity discount is lower if the neighboring building is of higher value than the ensemble buildings. No significant effect of any spillover from high-value to low-value properties can be estimated, however. Differences in the volume of the ensemble buildings versus the different shape building (again for all 320 ensembles) do not lead to difference in the heterogeneity discount either.

Conclusion

The far reaching question of how the architecture of a building and its neighbors codetermines their respective values has so far not been addressed in a large-scale and data-driven study. This paper shows that it is not only feasible but also worthwhile to empirically analyze the shape of buildings. Existing research on property values has eschewed three-dimensional building models as an information source since these data do not come in convenient bite-size formats but have unwieldy "Big Data" properties. City-wide shape data sets tend to be massive in size, exceeding the computational limits of traditional regression-style empirics. Furthermore, the data is unstructured and needs interpretation before derived information on shapes can be linked to other property characteristics.

Extracting shape information is not "Big Data" wizardry, however. Condensing building models to shape distributions reduces the complexity while preserving sufficient information to estimate the degree of similarity between properties. These algorithm-based similarity estimates are good predictors of human perceptions of similarity. This opens up new avenues of research not only in real estate finance and economics, but also in the domain of architecture, urban planning or sustainability.

Ultimately, this paper presents empirical support for the notion that architectural homogeneity is positively valued in residential property markets. Rowhouses surrounded by other buildings of the same shape carry an economically and statistically significant premium of several percentage points vis-à-vis comparable buildings in heterogeneous rows of houses, which can be interpreted as evidence for benefits of enforcing coordination between developers of new buildings and owners of existing stock. After all, the value of buildings we live, work and are invested in largely depends on the architectural choices of neighbors.

Whether this preference for shape homogeneity is specific to Dutch home buyers or whether the positive attitude towards ensembles of similar shapes is universal remains a question for follow-up studies in other markets and cultures. Additionally, considering the shape of properties in empirical price estimations could put a price tag on certain architectural forms and could lead to more accurate marginal price estimates for attributes like housing type, year of construction, or location which are closely correlated with architecture.

References

Ahlfeldt, G., Mastro, A., 2012. Valuing Iconic Design: Frank Lloyd Wright Architecture in Oak Park, Illinois. Hous. Stud. 27, 1079–1099. doi:10.1080/02673037.2012.728575

Ahlfeldt, G.M., Maennig, W., 2010. Substitutability and Complementarity of Urban Amenities : External Effects of Built Heritage in Berlin. Real Estate Econ. 38, 285–323.

Alam, N., Wagner, D., Wewetzer, M., Coors, V., Pries, M., 2013. Towards automatic validation and healing of CityGML models for geometric and semantic consistency, in: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Istanbul, pp. 27–29.

Anderson, R.M., 1960. Architectural Control. Syracuse Law Rev. 26. doi:10.1525/sp.2007.54.1.23.

Asabere, P.K., Hachey, G., Grubaugh, S., 1989. Architecture , Historic Zoning , and the Value of Homes. J. Real Estate Financ. Econ. 2, 181–195. doi:10.1007/BF00152347

Boeters, R., 2013. Automatic enhancement of CityGML LoD2 models with interiors and its usability for net internal area determination. TU Delft.

Cardone, A., Gupta, S.K., Karnik, M., 2003. A Survey of Shape Similarity Assessment Algorithms for Product Design and Manufacturing Applications. J. Comput. Inf. Sci. Eng. 3, 109. doi:10.1115/1.1577356

City of Rotterdam, 2015. Rotterdam 3D [WWW Document]. URL http://www.rotterdam.nl/rotterdam_3d

City of Rotterdam, 2016. Welstandscirteria Centrumgebied.

Cliff, A.D., Ord, J.K., 1981. Spatial Processes: Models & Applications. Pion, London.

Florida, R., Mellander, C., Stolarick, K., 2009. Beautiful Places : The Role of Perceived Aesthetic Beauty in Community Satisfaction, Martin Prosperity Research.

Fuerst, F., McAllister, P., Murray, C.B., 2011. Designer buildings: estimating the economic value of "signature" architecture. Environ. Plan. A 43, 166–184. doi:10.1068/a43270

Gartman, D., 2009. From Autos to Architecture: Fordism and Architectural Aesthetics in the Twentieth Century. Princeton Architectural Press, New York, NY.

Glaeser, E.L., Ward, B. a., 2009. The causes and consequences of land use regulation: Evidence from Greater Boston. J. Urban Econ. 65, 265–278. doi:10.1016/j.jue.2008.06.003

Hilber, C. a. L., Vermeulen, W., 2014. The Impact Of Supply Constraints On House Prices In England. Econ. J. 126, n/a–n/a. doi:10.1111/ecoj.12213

Ihlanfeldt, K.R., 2007. The effect of land use regulation on housing and land prices. J. Urban Econ. 61, 420–435. doi:10.1016/j.jue.2006.09.003

Jensen, J., Cowen, D., 1999. Remote sensing of urban suburban infrastructure and socio-economic

attributes. Photogramm. Eng. Remote Sensing 65, 611–622.

Koster, H.R. a., van Ommeren, J., Rietveld, P., 2012. Bombs, boundaries and buildings. Reg. Sci. Urban Econ. 42, 631–641. doi:10.1016/j.regsciurbeco.2012.02.007

Kwan, M.-P., Lee, J., 2005. Emergency response after 9/11: the potential of real-time 3D GIS for quick emergency response in micro-spatial environments. Comput. Environ. Urban Syst. 29, 93–113. doi:10.1016/j.compenvurbsys.2003.08.002

Leichenko, R., Coulson, E., Listokin, D., 2001. Historic Preservation and Residential Property Values: An Analysis of Texas Cities. Urban Stud. 38, 1973–1987. doi:10.1080/00420980120080880

LeSage, J.P., 2014. What regional scientists need to know about spatial econometrics. Available SSRN (Aug. 8, 2015). doi:10.2139/ssrn.2420725

LeSage, J.P., Pace, R.K., 2011. Pitfalls in higher order model extensions of basic spatial regression methodology. Rev. Reg. Stud. 41, 13–26.

Levy, J.M., 2013. Contemporary Urban Planning. Routledge, New York City, NY.

Listokin, D., Listokin, B., Lahr, M., 1998. The contributions of historic preservation to housing and economic development. Hous. Policy Debate 9, 431–478. doi:10.1080/10511482.1998.9521303

Mayer, C.J., Somerville, C.T., 2000. Land use regulation and new construction. Reg. Sci. Urban Econ. 30, 639–662. doi:10.1016/S0166-0462(00)00055-7

Moorhouse, J.C., Smith, M.S., 1994. The Market for Residential Architecture: 19th century Row Houses in Boston's South End. J. Urban Econ. 35, 267–277. doi:10.1006/juec.1994.1016

Nagel, C., Stadler, A., Kolbe, T.H., 2009. Conceptual Requirements for the Automatic Reconstruction of Building Information Models from Uniterpreted 3D Models, in: The International Archives of Photogrammetry, Remote Sensing and Spatial Information Science.

Nichol, J., Wong, M.S., 2005. Modeling urban environmental quality in a tropical city. Landsc. Urban Plan. 73, 49–58. doi:10.1016/j.landurbplan.2004.08.004

Ohbuchi, R., Minamitani, T., Takei, T., 2005. Shape-similarity search of 3D models by using enhanced shape functions. Int. J. Comput. Appl. Technol. 23, 70. doi:10.1504/IJCAT.2005.006466

Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D., 2001. Matching 3D models with shape distributions, in: Proceedings International Conference on Shape Modeling and Applications. IEEE Computer Society, pp. 154–166. doi:10.1109/SMA.2001.923386

Peiser, R., 2014. Decomposing sprawl. Town Plan. Rev. 72, 275–298.

Quigley, J.M., Raphael, S., 2005. Regulation and the high cost of housing in California. Am. Econ. Rev. 95, 323–328. doi:10.1257/000282805774670293

Rakkolainen, I., Vainio, T., 2001. A 3D City Info for mobile users. Comput. Graph. 25, 619–625. doi:10.1016/S0097-8493(01)00090-5

Ranzinger, M., Gleixner, G., 1997. GIS Datasets for 3D Urban Planning. Comput. Environ. Urban Syst. 21, 159–173.

Regan, K., 1990. You can't build that here: The Constitutionality of Aesthetic Zoning and Architectural Review. Fordham Law Rev. 58, 1013–. doi:10.1525/sp.2007.54.1.23.

Rubin, B.A., 1975. Architecture, Aesthetic Zoning and the First Amendment. Stanford Law Rev. 179, 179–201.

Sullivan, 1896. The Tall Office Building Artistically Considered. Lippincott's Mag.

Tangelder, J.W., Veltkamp, R.C., 2008. A survey of content based 3D shape retrieval methods. Multimed. Tools Appl. 39, 441–471. doi:10.1007/s11042-007-0181-0

Town of Eastchester NY, 2000. Local Law No 5 - 2000.

Upton, G., Fingleton, B., 1985. Spatial data analysis by example: point pattern and quatitative data. John Wiley, Chhichester, UK.

Vandell, K.D., Lane, J.S., 1989. The Economics of Architecture and Urban Design: Some Preliminary Findings. AREUEA J. 17, 235–260.

Wolff, M., Asche, H., 2009. Towards Geovisual Analysis of Crime Scenes – A 3D Crime Mapping Approach, in: Sester, M., Bernard, L., Paelke, V. (Eds.), Advances in GIScience. Springer, Berlin and Heidelberg, pp. 429–448. doi:10.1007/978-3-642-00318-9

Wu, H., He, Z., Gong, J., 2010. A virtual globe-based 3D visualization and interactive framework for public participation in urban planning processes. Comput. Environ. Urban Syst. 34, 291–298. doi:10.1016/j.compenvurbsys.2009.12.001

Yasumoto, S., Jones, A.P., Nakaya, T., Yano, K., 2011. Computers , Environment and Urban Systems The use of a virtual city model for assessing equity in access to views. Comput. Environ. Urban Syst. 35, 464–473. doi:10.1016/j.compenvurbsys.2011.07.002

Yu, S.-M., Han, S.-S., Chai, C.-H., 2007. Modeling the value of view in high-rise apartments: a 3D GIS approach. Environ. Plan. B Plan. Des. 34, 139–153. doi:10.1068/b32116

Figures and Tables

Figure 1: Basic solid geometries and their representation as a shape distribution

House with saddle roof   Cube   Rectangular form
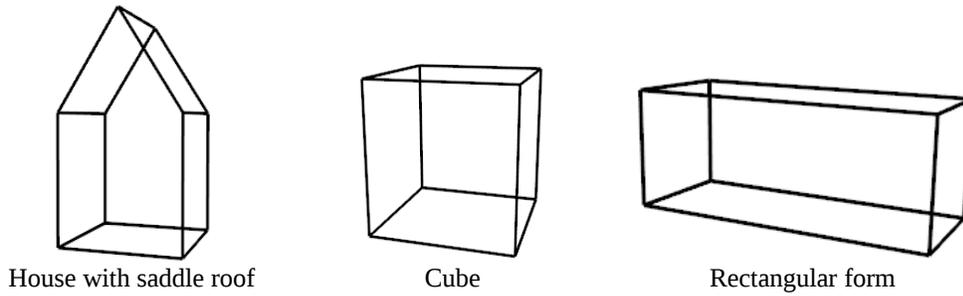
*EDF for distances between randomly selected points on hull of shape*
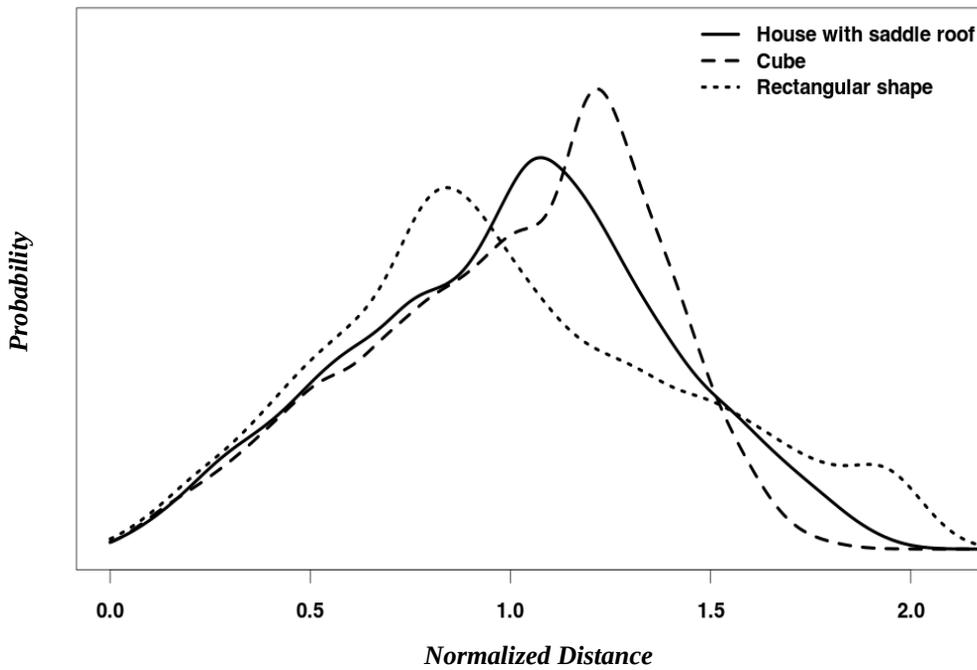
Notes: The distributions are normalized by dividing all distances by average distance per shape before estimation of the density functions. The estimated density functions of three basic shapes show very distinct profiles with cubes having the most pronounced peak. Rectangular shapes exhibit flatter space distributions with a hump in the right tail.
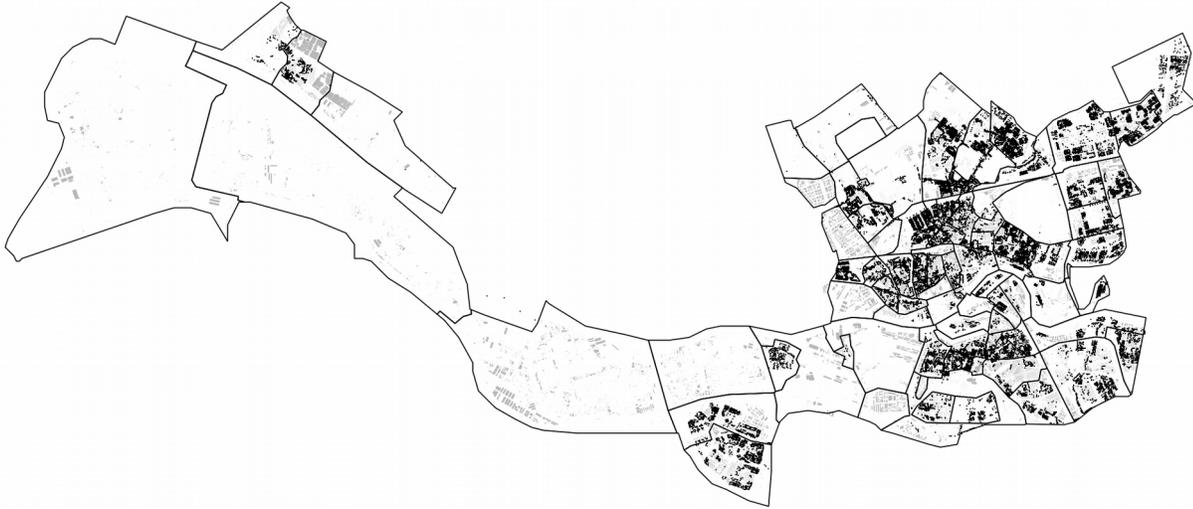
Notes: The gray areas indicate all buildings from Rotterdam's 3D city map. The black areas represent the final sample of single family homes for which transaction data is available in 2006-2013. Solid lines mark the official neighborhood boundaries. The majority of residential transactions can be found in the residential neighborhoods in Rotterdam proper in the east, while the west is dominated by the harbor, infrastructure, warehouses and industrial properties.

Figure 3: Shape distributions of single family homes in sample
(Rotterdam, 2006-2013)
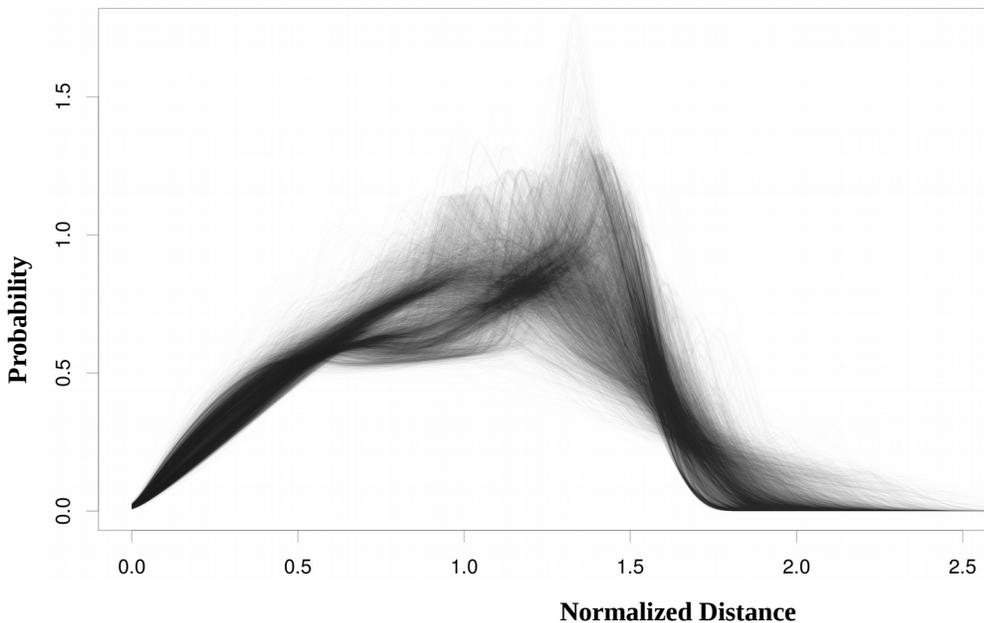
Notes: The shape distributions of all 6,129 buildings in the sample exhibit substantial heterogeneity, indicating a large diversity in the shapes of the single-family homes in Rotterdam. At the same time, the darker areas in the figure show a clustering around typical distributions – despite all uniqueness, building exteriors appear to be variations of a limited number of typical architectural forms.

Table 1: Distribution of Shape similarity S across all Rotterdam buildings

	Min	1st Quantile	Median	Mean	3rd Quantile	Max	SD
All buildings	0.21	0.69	0.77	0.76	0.84	1.00	0.11
Cube-shape buildings	0.60	0.94	0.96	0.95	0.97	1.00	0.03

Notes: The pairwise shape similarity measure *S* is calculated for all combinations of 185,914 buildings in Rotterdam. The distribution of similarity values clearly differs from the distribution for cube-shaped buildings, which display higher levels of similarity. Overall, 1,229 buildings are classified as having a cube shape: they consist of exactly 4 walls, a roof and ground plate which are all squares and of similar size. The difference in means between non-cube and cube-shape buildings is statistically significant (t-value = 3,828).

Table 2: Mean values for hedonic attributes and distribution of house types across shape clusters

Cluster	Count	Mean								% House type				
		Price (in '000)	Int. space (m²)	Lot (m²)	Volume (m³)	Dist. Betw. points	Num. rooms	Num. floors	Year of constr.	Detached	Semi-detaced	Corner	Linked-detached	Terraced
1	319	285	148	128	428	7.7	5.2	3.0	1963	0.3%	0.3%	0.6%	2.2%	96.6%
2	581	303	151	127	439	7.7	5.5	3.2	1956	0.2%	0.2%	0.9%	0.7%	98.1%
3	406	286	131	208	392	7.3	4.6	2.5	1960	7.1%	15.8%	21.2%	5.9%	50.0%
4	860	322	140	275	403	6.6	5.1	2.8	1959	7.8%	19.3%	50.5%	2.3%	20.1%
5	491	248	123	165	351	6.4	5.1	3.0	1961	0.0%	0.2%	1.0%	0.8%	98.0%
6	800	301	141	181	406	6.8	5.3	3.1	1961	0.1%	7.0%	42.5%	1.9%	48.5%
7	716	259	133	151	378	7.1	4.8	3.0	1967	0.3%	1.0%	11.7%	2.5%	84.5%
8	319	194	101	179	281	6.4	4.1	2.1	1975	0.0%	1.3%	6.6%	0.3%	91.8%
9	785	284	137	146	396	6.9	5.4	3.2	1958	0.0%	0.1%	0.3%	0.5%	99.1%
10	849	365	153	273	450	7.0	5.1	2.8	1968	11.7%	24.4%	38.9%	6.8%	18.3%
Total	**6,126**	**294**	**138**	**192**	**400**	**7.0**	**5.1**	**2.9**	**1962**	**3.3%**	**8.3%**	**21.4%**	**2.5%**	**64.54%**
One-way analysis of equal means, F-value										Chi-squared test of independence				
		89	96	62	96	159	64	145	22	X^2 =3333.4, df = 36				

Notes: Properties are grouped based on their shape distributions using the k-means clustering algorithm (k=10). The values for core hedonics differ across shape clusters. The equality of means can be rejected, with all F values exceeding 22 in one-way analysis (num. df = 9, denom. df = 2,163). Also, housing types are not equally distributed across clusters (X^2 =3333.4, df = 36).

Table 3: Spatial correlation in building shapes

Shape cluster	Same-shape statistic	Expectation	Variance	P-Value
1	42.48	8.48	1.18	0.00
2	99.19	28.19	3.60	0.00
3	60.71	13.75	1.86	0.00
4	131.95	61.79	7.21	0.00
5	85.88	20.12	2.64	0.00
6	101.44	53.46	6.36	0.00
7	119.51	42.82	5.24	0.00
8	57.81	8.48	1.18	0.00
9	153.00	51.48	6.15	0.00
10	132.11	60.22	7.05	0.00

Notes: Joint count tests under nonfree sampling (Cliff & Ord, 1981) suggest that buildings of similar shapes tend to be close to each other. The odds of observing buildings from identical shape clusters within 100 m off each other are 2.8 times higher than expected under the assumption of random spatial distributions. The same-shape statistics are statistically highly significant.

Table 4: Automatic vs. human classification

		Automatic classification (WS)	
		Different	*Similar*
Classification by survey respondents	*Different*	193	56
	Similar	9	116

Notes: Overall, 374 pairs of buildings have been classified by human subjects as either being similar or different. The corresponding values in similarity matrix WS show that the automatic shape comparison leads to classifications that are, on average, similar to classifications by humans. $X^2 = 162.81$, df = 1, p-value < 0.001.

Table 5: Regression coefficient estimates

	I				II				III				IV			
Variable	*Coeff.*	*SE*	*P-Val.*		*Coeff.*	*SE*	*P-Val.*		*Coeff.*	*SE*	*P-Val.*		*Coeff.*	*SE*	*P-Val.*	
Const.	12.510	0.015	0.000	***	12.631	0.016	0.000	***	8.399	0.069	0.000	***	8.475	0.072	0.000	***
Year of sale (vs. 2006)																
2007	0.013	0.010	0.211		0.019	0.010	0.047	*	0.018	0.008	0.019	**	0.019	0.008	0.016	**
2008	0.047	0.011	0.000	***	0.049	0.010	0.000	***	0.044	0.008	0.000	***	0.046	0.008	0.000	***
2009	-0.003	0.011	0.766		0.003	0.011	0.748		0.006	0.009	0.461		0.008	0.009	0.357	
2010	-0.005	0.012	0.668		0.000	0.011	0.993		0.008	0.009	0.382		0.009	0.009	0.311	
2011	-0.014	0.012	0.257		-0.012	0.011	0.282		0.006	0.009	0.525		0.006	0.009	0.505	
2012	-0.082	0.012	0.000	***	-0.079	0.011	0.000	***	-0.050	0.009	0.000	***	-0.050	0.009	0.000	***
2013	-0.122	0.016	0.000	***	-0.131	0.016	0.000	***	-0.102	0.013	0.000	***	-0.106	0.013	0.000	***
Type (vs. detached)																
Corner									-0.207	0.016	0.000	***	-0.212	0.016	0.000	***
Terraced									-0.269	0.016	0.000	***	-0.274	0.016	0.000	***
Semi-det.									-0.088	0.017	0.000	***	-0.095	0.017	0.000	***
Linked-det.									-0.216	0.021	0.000	***	-0.226	0.021	0.000	***
ln(int. space m^2)									0.741	0.015	0.000	***	0.725	0.016	0.000	***
ln(lot size m^2)									0.039	0.003	0.000	***	0.039	0.003	0.000	***
ln(Volume)									0.247	0.036	0.000	***	0.252	0.038	0.000	***
Year of construction (vs. before 1906)																
1906-1930									-0.013	0.016	0.413		-0.014	0.015	0.364	
1931-1944									0.018	0.017	0.303		0.006	0.017	0.738	
1945-1959									0.025	0.019	0.192		0.018	0.020	0.365	
1960-1970									-0.054	0.021	0.010	**	-0.055	0.021	0.010	**
1971-1980									-0.090	0.021	0.000	***	-0.080	0.021	0.000	***
1981-1990									-0.071	0.019	0.000	***	-0.063	0.019	0.001	***
1991-2000									0.094	0.019	0.000	***	0.086	0.019	0.000	***
Yoc ≥ 2001									0.102	0.020	0.000	***	0.095	0.021	0.000	***
Yoc unknown									0.007	0.187	0.972		-0.056	0.186	0.763	
λ_W spat.	0.805		0.000	***	0.699		0.000	***	0.637		0.000	***	0.609		0.000	***
λ_{WS} shape					0.286		0.000	***					0.120		0.030	**
R^2	0.716				0.740				0.838				0.840			
Adj. R^2	0.716				0.740				0.837				0.839			

Notes: N=6,126.

Table 6: Regression Estimates for Architectural Homogeneity Premium

	I				II			
Variable	*Coeff.*	*Robust SE*	*P-Val.*		*Coeff.*	*Robust SE*	*P-Value*	
Const.	-0.040	0.015	0.008	***	-0.038	0.015	0.011	**
IntSpaceRatio$_{CB}$	0.015	0.010	0.132		0.008	0.009	0.370	
Size$_C$					0.033	0.012	0.008	***
Shape Similarity$_{CD}$					0.007	0.007	0.327	
Size$_C$*Shape Similarity$_{CD}$					-0.027	0.008	0.001	***
Controls for different years of sale *(Y$_{BC}$)*	YES				YES			

Notes: N=320, df=193 and 196. The dependent variable *PriceRatio (Price$_C$/Price$_B$)* is normalized (mean: 0). Independent variables are all standardized (mean: 0, SD: 1). Panel (I) presents the estimated coefficients from a reduced version of Eq. 5. The negative constant indicates a price discount of ensemble-buildings at the periphery. The P-values are calculated based on robust standard errors (White's estimator). The full model (II) confirms a negative constant and an overall discount (-3.8%) for a location at the periphery. Larger homes are less sensitive to the influence of difference in architecture: An one standard deviation increase in the interior floor space of the ensemble buildings offsets the effect already (+3.3%). Also, small rowhouses are more sensitive to the degree of shape similarity to neighboring structures. An one standard deviation increase in similarity to an adjacent non-ensemble building in combination with an one standard deviation decrease in size reduces the discount by 2.7%.
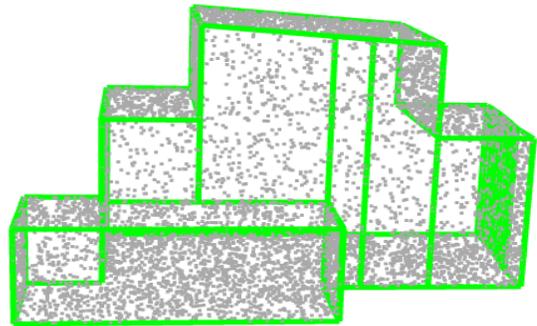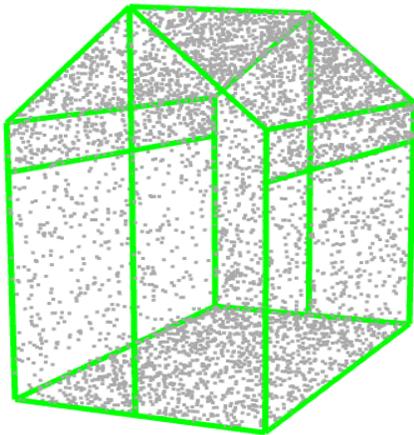
Appendix 1 (for referees)

Overall, 374 combinations (similar to the one shown above) of Rotterdam building models have been presented to human survey participants, who were then asked whether they would consider these buildings as being "pretty much the same" or "different" The automatic classification suggested in this paper can predict the human classifications well: 116 out of 125 combinations that have been classified as being "rather similar" by human respondents are also classified as similar by the automatic classification system suggested in this study. Only 9 (or 7%) are not. For pairs that are perceived as being different by humans, the match is a little lower: 193 out of 249 combinations flagged as "rather dissimilar" by humans are also considered dissimilar by the automatic classification (76%). A highly significant chi-squared statistic of 163 (with 1 degree of freedom) confirms that the automatic identification of similar buildings is highly correlated with human classifications

Appendix 2

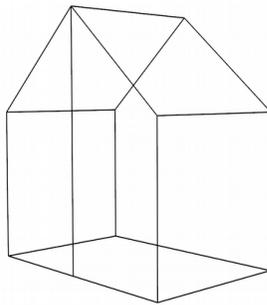*Example of a residential street in Rotterdam: Cypruslaan*

Three shapes of buildings can be observed in this section of a residential street in Rotterdam. On the left, complex single family homes with flat roofs while the right hand side features more basic semi-detached homes with saddle roofs. Some of the semi-detached houses have garages. Based on the 3D models of these houses (first panel in Table 7) , three distinct shape distributions are calculated. The estimated density functions (EDF) for distances between randomly selected points on hull of buildings show clear differences between all three shapes. In this example, the EDF for the semi-detached house with a garage (c) is situated between the EDFs for the more complex shape of the detached house (a) and the basic form of the semi-detached house (b).
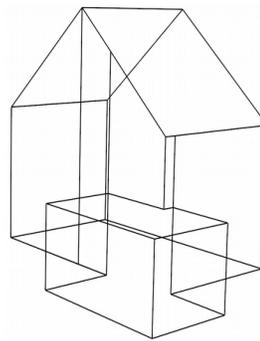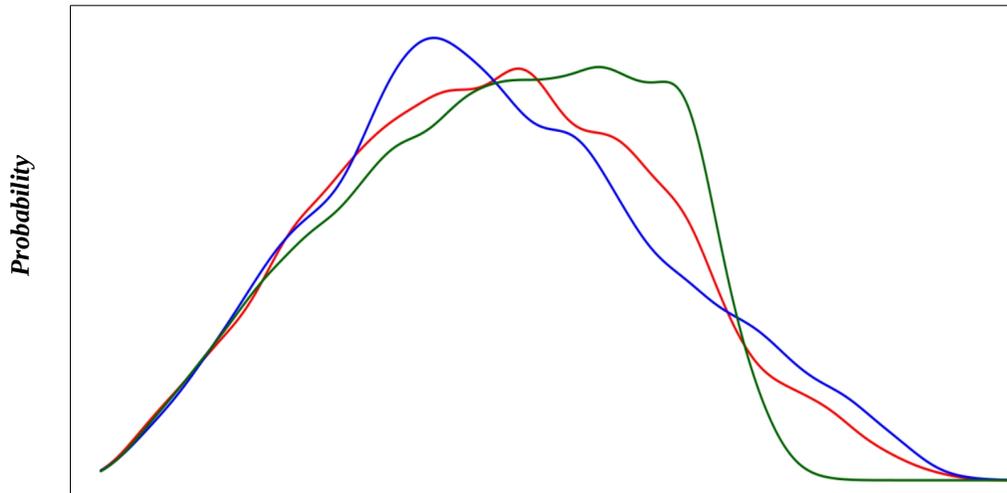
Table 7
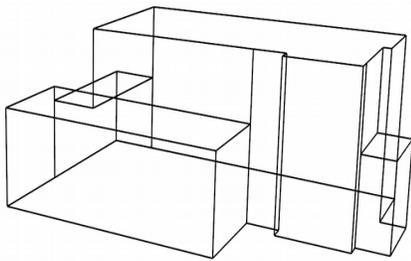
(a) Detached house	(b) Semi-detached	(c) Semi-detached, with garage

EDF for distances between randomly selected points on hull of building

Normalized distance
Blue: detached (a), green: semi-detached (b), red: semi-detach, with garage (c)

Appendix 3: Architectural homogeneity can reduce perceived shape heterogeneity

Ensemble of homogenous houses bordering a diverging shape and architecture

Photo by author

In this picture, the building on the right is perceived as the "odd one out" by respondents in a survey and also automatically classified as an outlier based on shape distributions. In the picture below, the automatic classification clearly sets the right row house apart from the other four. Survey respondents, however, do not perceive the right building to be very different since it adheres to the homogeneous architecture of the ensemble. Architectural homogeneity can reduce perceived shape heterogeneity.

Ensemble of 5 row houses with homogenous architecture but one diverging shape

Photo: Google Streetview