

EPC Labels and Building Features: Spatial Implications over Housing Prices

*Original*

EPC Labels and Building Features: Spatial Implications over Housing Prices / Barreca, A., Fregonara, E., Rolando, D.. - In: SUSTAINABILITY. - ISSN 2071-1050. - 13:5(2021), p. 2838. [10.3390/su13052838]

*Availability:*

This version is available at: 11583/2873980 since: 2021-03-28T13:08:40Z

*Publisher:*

MDPI

*Published*

DOI:10.3390/su13052838

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

## Article

# EPC Labels and Building Features: Spatial Implications over Housing Prices

Alice Barreca , Elena Fregonara  and Diana Rolando 

Architecture and Design Department, Politecnico di Torino, 10125 Turin, Italy; elena.fregonara@polito.it (E.F.); diana.rolando@polito.it (D.R.)

\* Correspondence: alice.barreca@polito.it; Tel.: +39-011-090-4370

**Abstract:** The influence of building or dwelling energy performance on the real estate market dynamics and pricing processes is deeply explored, due to the fact that energy efficiency improvement is one of the fundamental reasons for retrofitting the existing housing stock. Nevertheless, the joint effect produced by the building energy performance and the architectural, typological, and physical-technical attributes seems poorly studied. Thus, the aim of this work is to investigate the influence of both energy performance and diverse features on property prices, by performing spatial analyses on a sample of housing properties listed on Turin's real estate market and on different sub-samples. In particular, Exploratory Spatial Data Analyses (ESDA) statistics, standard hedonic price models (Ordinary Least Squares—OLS) and Spatial Error Models (SEM) are firstly applied on the whole data sample, and then on three different sub-samples: two territorial clusters and a sub-sample representative of the most energy inefficient buildings constructed between 1946 and 1990. Results demonstrate that Energy Performance Certificate (EPC) labels are gaining power in influencing price variations, contrary to the empirical evidence that emerged in some previous studies. Furthermore, the presence of the spatial effects reveals that the impact of energy attributes changes in different sub-markets and thus has to be spatially analysed.

**Keywords:** energy and economic sustainability; real estate market; energy performance certificate; green attributes; hedonic price model; spatial econometric models; spatial error model; existing buildings; energy retrofitting; Turin



**Citation:** Barreca, A.; Fregonara, E.; Rolando, D. EPC Labels and Building Features: Spatial Implications over Housing Prices. *Sustainability* **2021**, *13*, 2838. <https://doi.org/10.3390/su13052838>

Academic Editor: Colin Jones

Received: 29 December 2020

Accepted: 24 February 2021

Published: 5 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Changes in consumers' demand behaviours, expressed through the willingness to pay for housing properties or for interventions to improve their energy performance, are the core of a wide number of studies in the real estate economics literature. Charalambides et al. [1] note that energy efficiency improvement is one of the principal reasons for retrofitting a house. Contextually, in the recent years the influence of the building energy performance—by means of the Energy Performance Certificate (EPC) labels—on prices has been deeply explored, showing evidence of the growing importance of EPC labels on selling/rental prices formation process [2,3]. In the meanwhile, the linkage between the building energy performance and the architectural, typological, and physical-technical attributes seems to be less studied, although these building features, besides the location variable, are often relevant in price formation processes [4].

At the European level, the relationship between EPC label and selling prices has been studied in other urban contexts through the use of hedonic regression models, and results demonstrated the presence of a positive, albeit often limited, influence of the EPC on prices [5–7]. Furthermore, as highlighted by the most recent literature, it is worth analysing the effects of dwelling attributes—and particularly the “green” ones—on property prices, by assuming their spatial variability and by considering the specific market segments to which the properties belong.

Assuming that housing unit intrinsic attributes have to be better analysed also in relation to the building types and typological components, the aim of this paper is to explore the influence of existing housing properties' relevant attributes and energy performance on prices, by also investigating their spatial component. Implicitly, the objective is also to identify which factors are able to strengthen the enhancement process through energy retrofit interventions, aiming both at improving the properties' energy efficiency and at increasing their economic values. Notice that this study analyses a data sample of real estate listings, which includes not only prices and energy attributes, but also other property characteristics, given that the energy retrofit interventions can affect also the enhancement of other building features.

Methodologically, this study is based on spatial analyses of data points. Firstly, the Exploratory Spatial Data Analyses (ESDA) are performed to verify the eventual presence of spatial autocorrelation. The global and local spatial association is detected, by means of the Local Moran's Index; then, by applying the Local Indicator of Spatial Association (LISA), cluster maps are produced in order to identify the presence of (eventual) spatial clusters. Secondly, two regression models—the Ordinary Least Squares (OLS) model and the Spatial Error Model (SEM)—are applied on the whole data sample. Thirdly, the same analyses are performed on three different sub-samples.

A sample of 2092 housing properties listed on the Turin's (Northern Italy) real estate market in the 2015–2018 time period is considered as the case study. After the analysis of the whole data sample, this study presents a focus on two different sub-samples (territorial clusters), and, finally, on a sub-sample representative of buildings with low EPC labels constructed between 1946 and 1990, which are the most critical from an energy viewpoint.

In general, the results of the analyses performed on the complete sample allow to argue that the effects of the EPC, as mandatory by law, are starting to impact on the real estate market prices, in line with some international studies [8–11]. More precisely, the results of the analyses on the whole sample show that, in comparison to some previous studies on the Turin's real estate market [12,13], the EPC label effect on prices is slightly increasing. Particularly, it is confirmed that lower EPC labels (E, F, G) negatively affect listing prices in terms of marginal prices; furthermore, for the first time higher EPC labels (A–A4, B) slightly but positively influence property prices.

The analyses performed on the two spatial sub-samples, one in the Northern zone and one in the Centre zone of the city of Turin, reveal that the high EPC labels (A–A4, B) have a positive influence on prices only in the central sub-sample, while the low EPC labels (E, F and G) negatively affect prices only in the northern sub-sample. Finally, the analyses performed on the low energy-efficient dwellings sub-sample, with EPC label (E, F and G) and built in the 1946–1990 time period, confirm that the attributes related to buildings' energy behavior, for the existing buildings stock, are in many cases not related to the building quality.

Furthermore, the research findings open the way to interpretations, illustrated in the paper, which allow also for updating and extending the framework on the influence of the EPC on prices by comparing with the results of two previous studies [12,13], in which the impact of EPC on listing prices and transaction prices was analysed in the same Italian city.

Finally, the result of this study can support construction companies, real estate developers and property owners in their investment decisions: for example, in selecting the retrofit interventions particularly effective on the building/dwelling performance and, consequently, increase the property values.

Assuming these premises, the paper is divided into the following parts: Section 2 presents the considered literature and scientific background on the topic. In Section 3 the methodological approach is presented. In Section 4 the case study and data sampling are illustrated. Section 5 presents the results of the application and, finally, Section 6 concludes.

## 2. Literature and Background

### 2.1. Analysis of Pricing in Housing Economics

A significant part of housing economics literature is devoted to exploring the determinants of real estate prices. In general, the methods adopted by scholars to analyse the residential real estate prices include the following (notice that the quoted researches are referred specifically to the recent exploration of green attributes contribution on prices, but the methodologies illustrated are commonly used in real estate analysis): Hedonic Price Model (HPM) [14], Hedonic Price Model combined with spatial specification [15,16], regression and multivariate analysis [17], binary logit regression models in conjunction with a Polytomous Universal Model [18], quantile regression [19] and evolutionary polynomial regression [20]. Furthermore, the Structural Equation Models are explored as an alternative to the regression models for exploring the presence of latent variables [21–23]. Besides, other researches focus on the use of Neural Networks and Genetic Algorithms [24–26].

Among the wide variety of methods, two families of tools emerge from the literature: the first one which is founded on the consolidated Econometric approach through the Hedonic method, and a second one, which opens up to the Spatial Econometrics.

As known since the second half of the 1960s, the first method has represented for many years the privileged methodology for the real estate market analysis, and, specifically, for detecting the impacts of building/dwelling attributes on pricing processes. Firstly, Rosen [27] explores the hedonic models, operatively solved through Multiple Regression Analysis, for detecting the price determinants in the real estate housing properties, and even before Ridker [28] he studied the relation between housing prices and environmental qualities. It must be stressed that the hedonic model was the most frequently applied among the traditional approaches on housing market analysis specifically for detecting the influence of structural, environmental and neighbourhood attributes, as demonstrated by Malpezzi [29] who provides a review of traditional hedonic regression models. In many works the impact of location on prices is empirically demonstrated [30]. Contextually, the importance of geographical segmentation in price prediction is underlined in many researches [31–35].

The pilot studies have been followed in the last decades by a number of publications worldwide, oriented towards the treatment and modeling of the spatial effects. In fact, the well-known criticalities due to the spatial heterogeneity (heteroskedasticity) and the spatial autocorrelation (interdependence) are explored [32], with the support of tests for the detection of spatial effects [36–38], opening the way to the Spatial Econometrics. Implicitly, the recent spreading in spatial analyses is related to the weight caught by the location variable in values formation.

According to the microeconomics viewpoint, for applying the Spatial Econometrics methods it is fundamental to explore the (eventual) presence of a spatial dependence in the property values. In fact, as is well known in the scientific research context, since property prices are spatial data, in addition to depend on multiple factors, they are influenced by values of the nearby locations. This phenomenon is well known in literature as the spatial autocorrelation of property prices [39]. Given the importance of taking into consideration the spatial dependence in data and the issues generated by spatial autocorrelation, many studies carry out spatial analyses and spatial regressions. For example, Wilhelmsson [40] discussed the importance of adopting spatial econometrics as a way to explore the size of bias that can occur in parameters when spatial effects are considered. The study starts from the theory of spatial econometrics in the real estate economics; through an empirical analysis, it demonstrated (among other conclusions) that the spatial hedonic model is able to explain a higher percentage of the price variation, and, above all, that the economic interpretation changes by including a spatial structure in the hedonic model.

Operatively, spatial analyses were founded on the relevance of space in influencing the real estate prices. Dubin [41] introduced the spatial effects in the hedonic model, assuming, in the hedonic regression, the autocorrelation of the error term. LeSage and Pace [42] discussed the adjacency effect considered as an influence (or spill-over effect)

among neighbouring housing prices [43]. By implementing the hedonic model with the management of the spatial effects the Spatial Auto Regressive Model (SAR), the Spatial Lag Model (SLM) and Spatial Error Model (SEM) were obtained [44,45] with lattice data. Notice that the wide variety of spatial regression models were developed due to the growing availability of data technologies [46]. Besides, widely used methods are the Global Indicators of Spatial Autocorrelation (GISA) and the Local Indicators of Spatial Autocorrelation (LISA), among ESDA techniques [33,40,47,48].

In the following sub-section, the background of the study is illustrated, mainly considering some of the studies strictly related to the topics of this research.

## 2.2. A brief Overview on the European and Italian Regulatory Framework

The buildings' energy performance assessment was introduced in the European regulatory framework in 2002 and is expressed by different indicators that are part of the Energy Performance Certificate (EPC) of buildings and residential units. The European regulatory framework is presented in Table 1.

**Table 1.** European Directives on Energy performance of Buildings (Source: Author's elaboration).

European Directive	Description
European Directive 2002/91/EC [49]	Also known as Energy Performance of Buildings Directive (EPBD), it indicates the framework for European countries to develop regulations for the energy efficiency of buildings.
European Directive 2010/31/EU [50]	Also known as EPBD recast, it updates the previous European Directive and makes the EPC mandatory for all European Union buildings. The certification represents a crucial aspect within the EPBD for securing transparency in the calculation of the energy performance of buildings and, furthermore, for informing consumers about the energy performance of dwellings.
Guidelines accompanying Commission Delegated Regulation (EU) n. 244/2012 of 16 January 2012 [51]	They supplement the European Directive 2010/31/EU on the energy performance of buildings by establishing a comparative methodology framework for calculating cost-optimal levels of minimum energy performance requirements for buildings and building elements. The cost-optimal methodology has been recently explored [52].
European Directive 2018/844/EU [53]	It is aimed at accelerating the cost-effective renovation of existing buildings, towards a decarbonized building stock by 2050 and the mobilization of investments.

Concerning the Italian context, laws and standards have been produced aiming at transposing the European Directives. In Italy, the EPBD was adopted firstly by Legislative Decree n. 192 on 19 August 2005, according to which buildings are rated on a scale from A+ to G. The EPC labels are calculated in each Italian region, according to the standards UNITS 11300 parts 1 and 2 procedure for evaluating the building energy performance index. As an assumption, the highest label A+ corresponds to the lowest Energy Performance index for Heating (EPH). Coherently, the lowest label G indicates the highest EPH. With the Legislative Decree n. 28 dated 3rd March 2011, since 1st January 2012, the EPC is a mandatory requirement for dwellings listed in the selling/rental market. With the Ministerial Decree dated 22nd November 2012, the owners cannot self-certify the lowest EPC label G.

The EPBD recast was then adopted by Legislative Decree n. 63 dated 4th June 2013, and with the Law n. 90 dated 3rd August 2013 introduced a new methodology for the calculation of the energy performance of buildings, later adopted by the Ministerial Decree dated 26th June 2015. This last norm includes the National Guidelines for buildings' EPC and introduced a new EPC format, based on ten homogeneous labels at the national level. The highest level is A4 and corresponds to the lowest EP, whilst the lowest level is G with the highest EP.

### 2.3. The Energy Performance Certificate Effects on Housing Prices

Contextually, with the evolution of the regulatory framework and the consequent evolution of the methods for applying the norms and rules, a growing literature at the international/national level has been produced. According to the real estate economics perspective, many studies aim to detect the impact of EPC labels on pricing processes and on market dynamics. Operatively, on the same line as the widespread literature produced, a great number of these studies utilize the Hedonic Price Method.

More recently, a wide amount of literature has explored the potentialities of spatial econometrics in real estate market analyses, even by comparing spatial approaches with traditional Hedonic methods. Special attention is posed on the detection of green attributes' influence on pricing processes. Among the most recent literature, some studies can be cited as examples in the context of the present research due to the object of the analyses.

In a study by Bottero et al. [8], the conjoint use of a spatial econometric model and of the Hedonic Price Method is analysed for estimating the implicit marginal price as a measure of the willingness to pay for building energy consumption in the city of Turin. More precisely, the study tries to estimate the differential of building energy performance in monetary terms. It is assumed necessary to evaluate the social cost of energy wastes. As the authors report, the results show an evident necessity to pay particular attention towards the coherence between the spatial and econometric approaches.

In a recent study, Chen and Marmolejo-Duarte [9,54] reflect on the differences between dwellings whose energy efficiency impacts price formation and dwellings whose energy efficiency does not influence pricing processes, in Barcelona's (Spain) residential market. Starting from studies about the energy efficiency marginal prices methodologically founded on the use of hedonic models, the authors assumed the premises that, in some cases, the increase of the relevance of EPC labels, in terms of marginal prices, is not constant for all the classes (in Spain, A–G). They propose to analyse the differences in terms of architectural and location characteristics between dwellings that show an increase in EPC rating marginal prices, and, at the same time, dwellings for which EPC seems not to be relevant for pricing processes. More precisely, the study proposes a pooled Spatial Error Model for exploring the prices of multifamily houses in the period 2014–2016. Among others, the following results are particularly interesting for the present work: firstly, in general there is a correlation between spatial and houses characteristics, and the attributes can reveal a weight varying in function of the building location (spatial effects); secondly, architectural attributes and location contributions to the green premium must be further explored.

In a study by Dell'Anna et al. [11], a comparison between two European cities belonging to two different climate zones, Barcelona (Spain) and Turin (Italy), is presented. Methodologically, datasets of listing prices in the residential market are analyzed with the Hedonic Price Method and Spatial Econometric Models (specifically, Spatial Autoregressive Model and Spatial Error Model), for detecting the contribution (in terms of marginal price) of green attributes, and, contextually, for controlling the spatial correlation among the prices. As highlighted by the authors, the results related to the two cities are different (in Italy, EPC is more relevant than in Spain, where single characteristics are more appreciated). Furthermore, they deduce that the EPC implementation is still irregular in European States, and also that it can be reinforced by introducing a standardized rating model for EPC. Obviously, the importance to introduce the effects of location when analyzing green labels is confirmed, and consequently, so is the necessity to strengthen the Hedonic Price Model with Spatial Econometric Models.

Besides the abovementioned researches, other recent experiences on the topic are assumed in the background as being specifically aimed to calculate the spatial autocorrelation in property prices in the city of Turin; in these studies, spatial analyses are proposed to detect the presence of spatial dependence between different kinds of indicators and to manage the spatial latent variables in the property price determination process [55–58].

For concluding the scientific background, a comparative reading with the results of two previous studies in the Italian context focused on the relevance of the building

EPC label on listing/transaction prices formation, and on supply and demand behaviors, is fundamental.

The first study by Fregonara et al. [12] aimed at investigating the economic effects of the Italian statutory provisions, related to the energy performance of buildings, on the listing behaviors. More precisely, the study was directed towards the measuring of the impact of the EPC labels on listing prices, for analyzing whether the EPC label, besides the mandatory requirement, is considered as a significant aspect able to influence dwelling prices. The study assumes that listing prices and dwelling characteristics are the initial and fundamental information considered by sellers/buyers during a first preliminary analysis [59,60]. A data sampling of 577 listing prices collected in Turin in the year 2012 (source: real estate advertisement websites) is analyzed by means of a log-linear Hedonic regression model, applied through hedonic regression analyses. The results demonstrate that a low level of EPC (EPC label F) is found to be a significant influence; the same results were confirmed in the model tested with EPC label clusters. In conclusion, EPC labels only partially explain listing prices.

In the second study by Fregonara et al. [13], the impact of the EPC label on dwelling prices and on market liquidity was analyzed, considering both listing and transaction prices. Particularly, the time on the market and the difference between listing prices and transaction prices were considered. A sample of 879 transactions of old apartments in Turin in the period 2011–2014 was considered, to analyze the effect of a set of variables on the listing and transaction prices, and, furthermore, on time on the market and bargaining outcome. Besides the EPC labels, building construction period and the main dwelling characteristics were modelled. The study revealed that low EPC labels (E, F and G) are priced in the market but explain only 6–8 per cent of price variation; then, including the dwelling principal characteristics, EPC labels have no impact on prices (more precisely, the G rating is weakly significant. Notice that the G level is to be considered with special attention, due to the possibility, at least for a certain period, to self-certify the dwelling). Finally, a focus on old buildings built in the period of 1940–1989 confirms the previous results. From the study emerges the clear influence of the building construction period and of assets' location on prices/real estate market dynamics.

As a general conclusion, the results of these last two studies suggest the presence of latent variables that are able to catch the explicative power of EPC label, which must be carefully explored in line with the literature. This consideration encourages us to address the research towards the deepening of the power characteristics, not only in the pricing process, but also in explaining the consumers' behavior and the relative spatial real estate dynamics. Above all, the examination of the recent literature on the topic reveals the necessity to introduce spatial econometrics beside the traditional Hedonic Price Method to manage the influence of spatial effects on pricing processes.

### 3. Methodological Approach

The analyses performed in this paper were mainly based on two widely known methods: Exploratory Spatial Data Analyses (ESDA) and the Spatial Error Model (SEM). The spatial regression model was applied to investigate the influence of the Energy Performance Certificate label and other property intrinsic features on listing prices. The analyses were focused on a data sample of real estate listings and three different sub-samples, which were selected on the basis of two different approaches.

Methods and approaches are briefly illustrated in the following subsections.

#### 3.1. ESDA Statistics

Before performing regression models, the correlation among the considered variables has to be analysed by means of Spearman correlation test. Subsequently, Exploratory Spatial Data Analyses are necessary to investigate the presence of spatial autocorrelation that usually affects the real estate market [30,48,61,62]. When property prices and a series of characteristics are analysed across territorial units, the global and local spatial association

that exists between each unit and the neighbouring ones has to be detected. Therefore, the Local Moran's Index is calculated to assess the spatial autocorrelation level [63], as follows (1):

$$I_i = z_i \sum_j w_{ij} z_j, \quad (1)$$

where  $z_i$  is the standardized spatial weight and the summation over  $j$  is such that only neighbouring values  $j \in J_i$  are included. For ease of interpretation, the weights  $w_{ij}$  may be in row standardized form, and by convention  $w_{ij} = 0$ .

Furthermore, the Local Indicator of Spatial Association (LISA) [64] cluster maps are generated to identify eventual spatial clusters and the related spatial association classification between the following four categories:

- spatial cluster defined by high values of the investigated phenomenon with a high level of similarity with their surroundings named "high-high" or "hot spots";
- spatial cluster defined by observations with low values and a high level of similarity with its surroundings named "low-low" or "cold spots";
- spatial outliers defined by observations with high values surrounded by low ones named "high-low";
- spatial outliers defined by observations with low values surrounded by high ones named "low-high".

Spatial clusters are identified on the basis of a Row Standardized Queen Contiguity-First Order Weight matrix ( $W$ ), representative of the degree of the spatial autocorrelation of each territorial unit and its surroundings, by means of the GeoDa software (software tool devised by Centre for Spatially Integrated Social Sciences—CSISS) [65].

### 3.2. The Spatial Error Model (SEM)

Spatial hedonic models aim to manage the spatial components of dependent and explanatory variables to reach more precise and unbiased results. All spatial hedonic models are based on the classical ordinary least squared (OLS) regression model. Commonly, the OLS model is preliminary applied to detect the influence of a set of independent variables (observable attributes) on the dependent variable, which in the real estate market are usually represented by the property price (listing price or transaction price, according to the data availability). The hedonic model which, in turn, is founded on the OLS algorithm estimation, can be formally represented as follows (2):

$$Y = \alpha_k + \sum_{i=1}^n \alpha_i X_{ik} + \dots + \sum_{i=1}^n \beta_i Z_{im} + \varepsilon \quad (2)$$

where  $Y$  stands for the dependent variable,  $\alpha_k$  stands for the model intercept,  $X_{ik}$ , with  $k = 1, \dots, K$ , and  $Z_{im}$ , with  $m = 1, \dots, M$  stand for the variables introduced for each of the  $n$  observable characteristics,  $\alpha_i$  and  $\beta_i$  represent the hedonic weights assigned to each variable, i.e., the contribution given by each single characteristic level to the price value, and  $\varepsilon$  represents the error term.

In performing OLS models, the presence of spatial dependence has to be tested; if tests results are significant, the model is probably biased. The statistics are the simple Moran's  $I$  on the spatial autocorrelation of errors, the simple Lagrange Multiplier (LM-lag) test for a missing spatially lagged dependent variable, the simple LM test for error dependence (LM-error), and their robust variants: Robust LM (lag), Robust LM (error) and a portmanteau test (SARMA), composed by Lagrange Multiplier (error) and Robust LM (lag). Therefore, the spatial autocorrelation has to be managed by means of the spatial regression models described in the following subsections [63,66].

Assuming that the spatial component of prices has to be managed in order to make regression models unbiased [67] in the presence of spatial autocorrelation of the residuals, two possible spatial models can be used: Spatial Lag Model (SLM) and Spatial Error Model (SEM). In order to identify the best model, the results of Moran's  $I$  test and the

Lagrange Multiplier tests (LM-lag and LM-error) must be compared in the early phases of the analysis [62].

In this study, price variability shows a non-linear relation with explicative variables, thus the spatial autocorrelation is managed by applying SEM, which is estimated by the Maximum Likelihood Estimator (MLE) algorithm, given the assumption of normality in the error term. This approach is suggested when errors in measuring the locational characteristics are present, or when the errors of the hedonic Model are correlated due to spatial effects. This effect can be reduced by introducing a correction of the error in the model (in fact, SEM is also known as Spatial Error Correction Model).

The SEM can be specified as (3):

$$y_i = X_i\beta + \lambda w_i\varepsilon_i + u_i \quad (3)$$

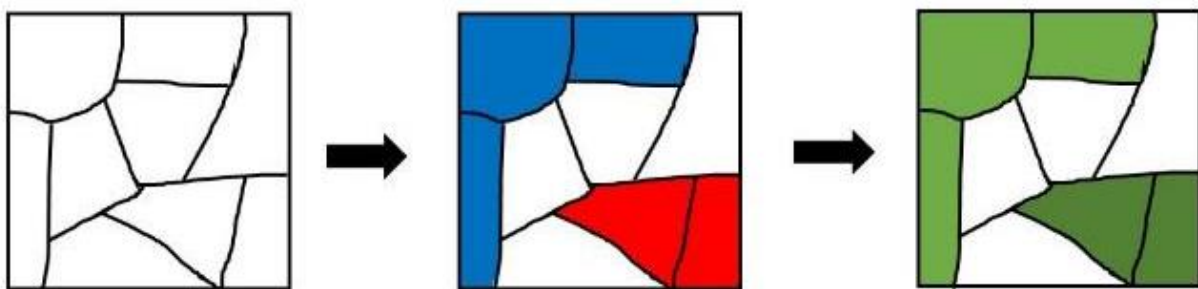
where  $u_i$  is the random error (independent identically distributed—i.i.d.), and the spatially structured error is composed of the added spatial error coefficient ( $\lambda$ ) and the original error term ( $\varepsilon$ ) weighted by a weight matrix  $w_i$  ( $W$ ). If there is no spatial correlation between errors, then  $\lambda = 0$ . If  $\lambda \neq 0$ , OLS is unbiased and consistent, but the standard errors will be wrong and the  $\beta$  will be inefficient. Notice that a positive and significant value of  $\lambda$  means that the model fit is good.

### 3.3. Sample Stratification and Spatial Clustering

The analyses on the whole data sample can be focused by applying two different stratification approaches.

Firstly, two different sub-samples can be spatially identified, starting from the results of the Local Indicator of Spatial Association (LISA) cluster maps, generated by analysing the listing price variable.

The first sub-sample can be identified by including the listing prices classified as “high-high” clustered and their surroundings, within a maximum 500 m distance. The second sub-sample instead can be identified by assuming the listing prices “low-low” clustered and their surroundings, within a maximum 500 m distance (Figure 1).



**Figure 1.** Data subsampling process (Source: Authors' Elaboration).

Secondly, another attribute (not spatial) sub-sampling approach can be assumed to create a different sub-sample, based on intrinsic features, representative of the most energy inefficient housing units. In particular, a set of housing units can be selected on the basis of the following assumptions related to building features:

- Construction period: in the 1946–1990 time period, which represents the post-war period, many medium Italian cities were characterized by a consistent population growth, a great urban development and a series of economic policies finalized to support housing construction activities at the municipal level, even if not yet sensitive to the building energy issues. In fact, buildings built during that period can be considered particularly “energy voracious” due to their typological and technological characteristics (one-layer walls, single glass windows, low quality of materials, etc.);

- EPC label: when precise information about the building physical features is not available, this variable can represent a suitable proxy of the energy performance.

It is worth mentioning that this second approach is based on two building intrinsic features (construction period and EPC label) that do not generate spatial clusters. These are significant characteristics for the study area of city of Turin, but they could differ in other urban contexts and thus they could be modified according to the built heritage specificities.

#### 4. Case Study, Data and Statistics

The real estate market of the city of Turin (Northern Italy) was assumed as a case study for this study. A data sample of existing housing units listed on the market in the 2015–2018 time period was used by assuming a database from the Turin Real Estate Market Observatory (TREMO). The TREMO was founded in 2000 by a partnership between the Politecnico di Torino and the Municipality of Turin; it constantly monitors and analyses housing prices and the related intrinsic or extrinsic features, assuming the different territorial segments of the city of Turin, which correspond to different real estate submarkets.

The data sample consisted of 2092 real estate listings published on one of the most important Italian real estate market advertisement websites, which were punctually analysed in order to study the listing prices and a series of features related both to the housing unit and to the building.

Obviously, the data set results from a previous phase of selection and cleaning of eventual outliers or incomplete data. The unitary mean listing price of the sample is 2210 Euro per square meter, while the standard deviation is 983Euro per square meter.

As is known, in Italy selling prices of real estate properties are not public information. Thus, selling prices are difficult to observe. Nevertheless, in many studies listing prices are used by researchers, appraisers and real estate companies to perform market analyses and to estimate the value of properties.

In fact, as shown from studies and literature, listing prices are capable of representing a fundamental aspect of the asset value formation process, specifically for their influence on selling processes and price prediction [68,69]. Furthermore, studies demonstrate the impact of listing prices on assets liquidity [70] and price spreads, represented by the difference between the listing price and selling price [71]. In the Italian context, a study analyses whether and to what extent listing prices can be considered a proxy for selling prices [72].

For these reasons, in this study listing prices are assumed as the main information available to analyse house pricing processes. A fundamental aspect is that listing prices are a function of house attributes, able to influence the bargaining processes [73] and opening to the analysis of the impact of green attributes and EPC labels on prices/values, as explored in previous studies [12,13].

##### 4.1. Variables and Descriptive Statistics

In addition to the listing price variable, a set of relevant characteristics was selected and analysed.

As shown in Table 2, this set includes 22 variables related to both building and housing unit features. A first group refers to the building physical-distributive features: building type (BLT), building category (BLC), construction time period (CTP), presence of concierge (CSS), classification of the housing unit as a penthouse (PNT), presence of car box (CBX) and presence of lift (LFT). A second group refers to the housing unit features: number of rooms (NRM), number of bathrooms (NBT), presence of terrace (LTR), number of views (NVW), allocation level (ALL) and maintenance level (MTL).

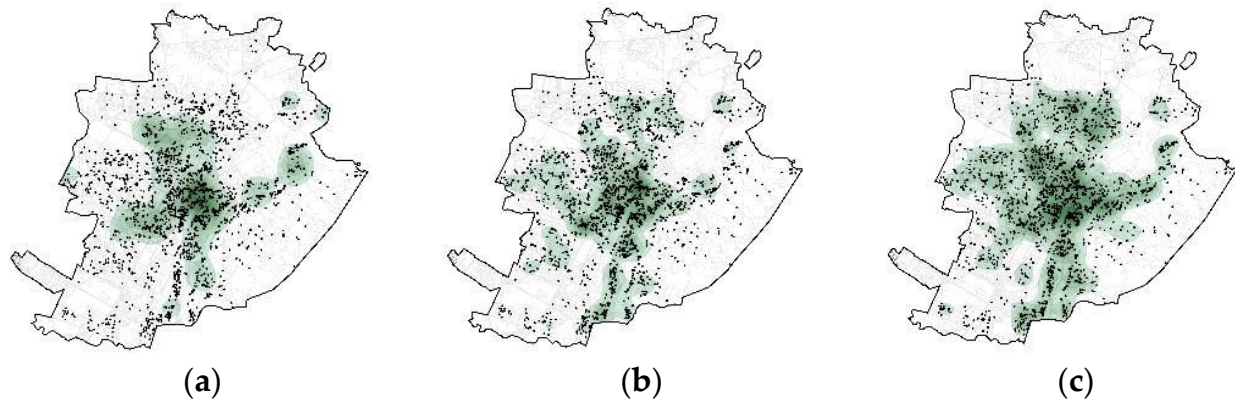
**Table 2.** List of variables and Descriptive statistics (Source: Author’s Elaboration).

		Data Typ.	min	Media	max	dev.st
<b>Dependent variables</b>						
TLP	Total Listing Price	Euro	21,000	278,181.835	5200.000	290,331.123
GS	Gross surface	Mq	15.000	113.346	1250.000	76.771
LP	Unitary Listing Price	Euro/mq	538.462	2209.727	6933.333	983.464
LogLP	Log listing price	Log e	6.289	7.603	8.844	0.447
<b>Independent variables</b>						
<b>Building features:</b>						
BLT	Building Type (Levels 0=condominium, 1 = detached house/villa)	dummy	0.000	0.928	1.000	0.258
BLC	Building category (Levels, 0 = NA, 1 = tenement, 2 = Economic, 3 = Medium, 4 = Noble, 5 = Prestigious)	dummy	0.000	3.087	5.000	1.028
CTP	Construction time period (Levels, 0 = NA, 1 = ≤ 1945, 2 = 1946–1975, 3 = 1976–now)	dummy	0.000	0.764	3.000	0.954
CSS	Custodian Service (Levels, 0 = no; 1 = yes)	dummy	0.000	0.166	1.000	0.372
PNT	Penthouse (Levels, 0 = no; 1 = yes)	dummy	0.000	0.037	1.000	0.190
CBX	Car Box (Levels, 0 = no; 1 = yes)	dummy	0.000	0.194	1.000	0.396
LFT	Lift (Levels, 0 = no; 1 = yes)	dummy	0.000	0.749	1.000	0.434
<b>Housing unit features:</b>						
NRM	Number of rooms	numeric	0.000	2.064	9.000	1.178
NBT	Number of bathrooms	numeric	0.000	1.506	5.000	0.731
LTR	Large Terrace (Levels, 0 = no; 1 = yes)	dummy	0.000	0.144	1.000	0.351
NVW	Number of views	numeric	0.000	2.061	4.000	0.639
ALL	Allocation Level (Levels, 0 = NA, 1 = lower, 2 = intermediate, 3 = higher)	dummy	1.000	2.086	3.000	0.825
MTL	Maintenance Level (Levels, 0 = NA; 1 = to renovate; 2 = partially to renovate; 3 = good; 4 = renovated)	dummy	0.000	2.868	4.000	1.034
ARC	Air-conditioning system (Levels, 0 = no; 1 = yes)	dummy	0.000	0.107	1.000	0.309
EPC	EPC label group (Levels, E,F,G = 1, C and D = 2, A-A4 and B = 3)	dummy	0.000	1.064	3.000	0.799

Lastly, two “green variables” were considered in order to analyse the energy performance level and the technological equipment of the building: the presence of air-conditioning system (ARC), and the EPC label. The analysis of latter variable—the EPC label—is particularly relevant considering the aim of the present study, since it can be considered a proxy for a series of other variables related to technological and physical-technical characteristics. Furthermore, being strictly related to the energy consumption, if considered jointly with other characteristics (e.g., level of maintenance and age of the building), it can also represent an “indicator of energy voracity” of a building. This indicator might give a more complete picture of the building features that could be appreciated by buyers and might be one part of the comprehensive “quality performance” of the buildings themselves.

In this study, the EPC labels were grouped in three levels: “high EPC labels” (A–A4, B), which correspond to the 3% of the sample; “medium EPC labels” (C, D), which correspond to the 39% of the sample; and “low EPC labels” (E, F, G), which correspond to the 58% of the sample. In Figure 2, the three groups are spatially represented: high labels are a

small part of the sample and are mainly located in the city centre. On the contrary, medium and low labels constitute the main part of the sample and are more scattered all over the city. In particular, the representation of the “low EPC labels” group confirms that it cannot correspond to a proper spatial cluster.



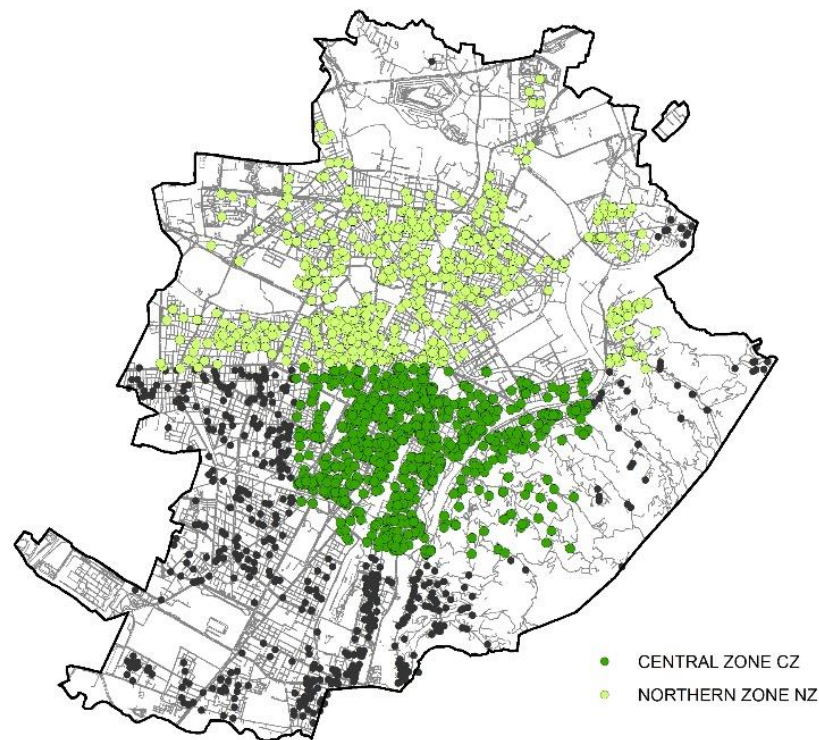
**Figure 2.** Density Point (Heat) maps of Energy Performance Certificate (EPC) label groups—heat maps: (a) high EPC labels (A-A4, B); (b) medium EPC labels (C, D); (c) low EPC labels (E, F, G) (Source: Author’s elaboration).

#### 4.2. Data Sub-Sampling

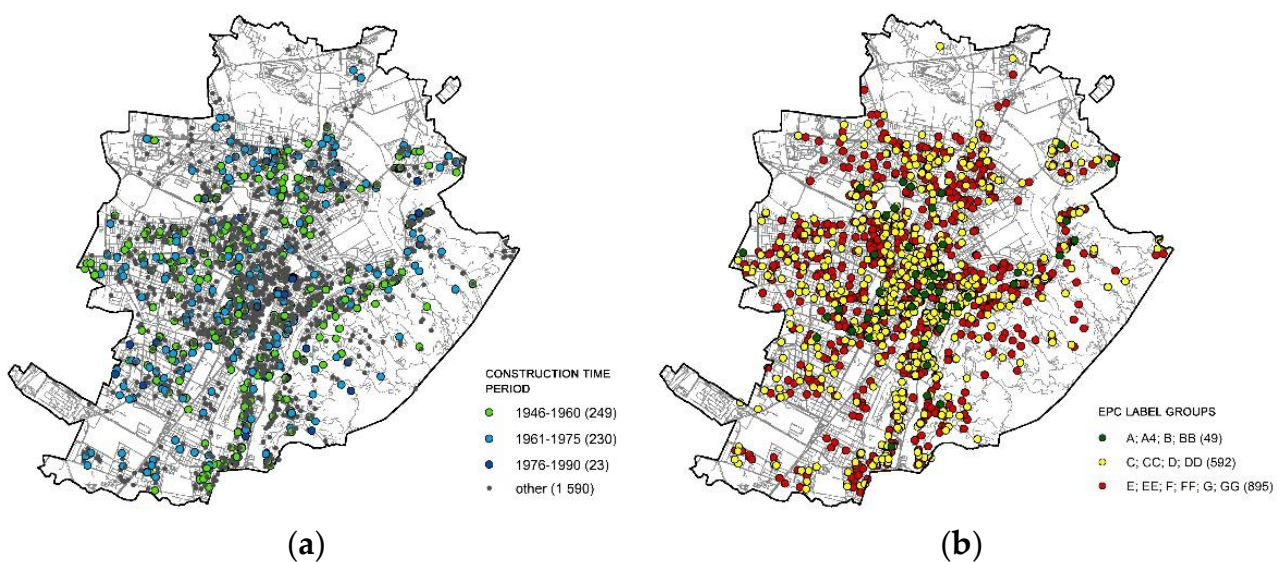
The whole data sample was split in different sub-samples by applying the two above-mentioned different stratification approaches.

Firstly, two spatial sub-samples were selected, on the basis of the LISA “high-high” and “low-low” cluster maps (in Section 5.1, Figure 5), to in-depth analyse the different influence of the EPC label in different areas of the city. The so called “Central Zone” (CZ) included both spatially correlated values (LISA = 1) and the probable outliers (LISA = 3 and 4) and those “not significant” (LISA = 0) that fell in the same zone, while the “Northern Zone” (NZ) cluster included “low-low” values (LISA = 2), the probable outliers (LISA = 3 and 4) and the “not significant” ones (LISA = 0) that fall in this zone (as previously explained in Section 3.3, Figure 1). In Figure 3, the result of the first spatial sub-sampling is presented: the Central Zone (CZ) sub-sample (dark green dots) and the Northern Zone (NZ) sub-sample (light green dots).

Secondly, a third sub-sample, including only the energetically inefficient buildings, was created by applying the abovementioned attribute (not spatial) sub-sampling approach. Therefore, this last sub-sample, named “Low Energy Efficient (LEE),” was not based on a spatial clustering, and it included housing units in buildings built between 1946 and 1990 with a low EPC label (EPC = E, F, G), which are randomly located in the city of Turin, as Figure 4 shows.



**Figure 3.** Spatial sub-samples—Listing price variable (Source: Authors' Elaboration on Turin Real Estate Market Observatory (TREMO) data).



**Figure 4.** Low Energy Efficient (LEE) sub-sample: (a) buildings built between 1946 and 1990; (b) housing units with a low EPC label (EPC = E, F, G) (Source: Authors' Elaboration on Turin Real Estate Market Observatory (TREMO) data).

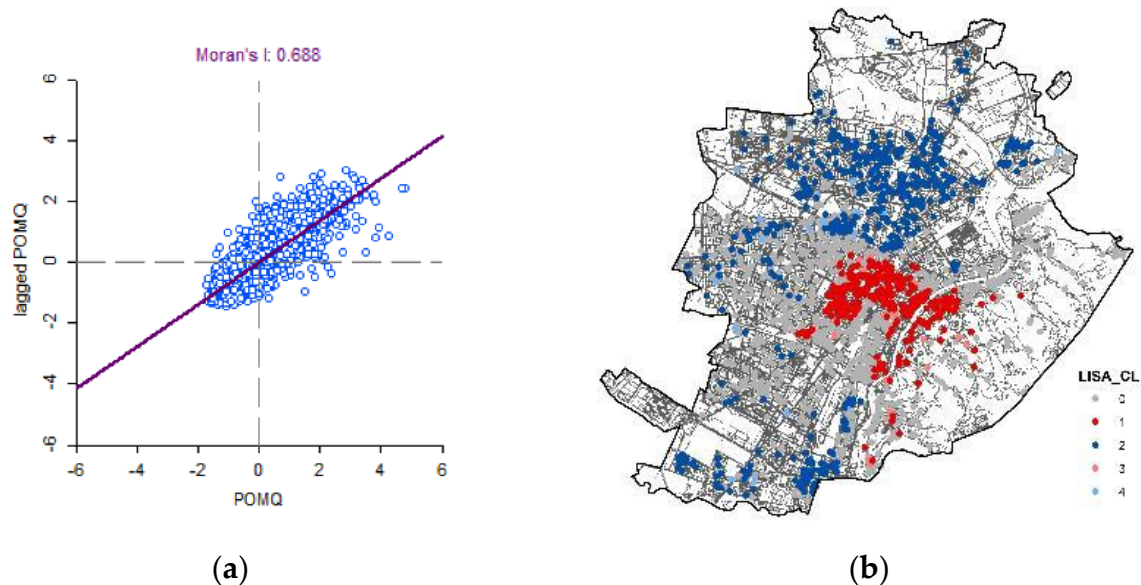
## 5. Results

By applying the methodological approach on the case study presented in the Section 4, the dataset was processed and tested by means of the GeoDa software (1.14.0—24 August 2019), and by the open-source software “R” (by R Foundation for statistical computing, software version 3.6.1 (5 July 2019)); spatial analyses and maps were produced by the

ArcGIS Desktop software package (software version 10.8.0.12790, Esri, Redlands, CA, USA). The results are illustrated and commented in the following sub-sections.

### 5.1. ESDA Statistics and Data Sub-Samples

Firstly, the Spearman correlation test was performed: results showed the absence of significant correlation among the considered variables. Therefore, to verify the presence of spatial correlation in the dataset, the Moran's Index was calculated and represented in the scatterplot and in the corresponding LISA cluster map in Figure 5.



**Figure 5.** Moran's Index scatter plot (a) and Local Indicator of Spatial Association (LISA) Cluster Map of Listing Prices per square meter—LP (b) (Source: Authors' elaboration).

Results showed that most of the observations fall in the II and IV quadrants, suggesting the presence of a positive spatial autocorrelation between LP values and their lagged (Moran's  $I = 0.688$ ). The Local Indicator of Spatial Association (LISA) was calculated to explore the significance of the spatial clusters and to produce the related map. The LISA cluster map confirmed the presence of clusters with the highest and lowest concentration of highest and lowest LP values, whereas the significance calculation (99 permutation) on the basis of Monte Carlo statistics confirmed the significance of the clusters, with a  $p$ -value between 0.001 and 0.05.

A first cluster called "high-high," collecting 170 data points with a positive autocorrelation of high values, was identified in the historical city centre and in the nearby part of the Turin's hillside, which are very classy and rich areas. A second "low-low" cluster, collecting 216 data points with a positive autocorrelation of lower values, was identified in the northern and southern outskirts of the city, where there is a high population with low income and the building quality is low (Figure 5b).

### 5.2. Hedonic Regression Models

A traditional logarithmic hedonic model was applied on the complete sample (2092 existing housing units listed on the market in 2015–2018 time period). The unitary listing price (LP), measured in Euro/m<sup>2</sup>, was assumed as the dependent variable. On the basis of the normality test of the LP variable, the logarithmic transformation was preferred: thus, the dependent variable of the OLS model was LogLP. The whole set of explanatory variables was considered to assess the influence of the variables on price variation. All Lagrange Multiplier tests (both simple and robust) on the OLS application were significant and

showed the presence of spatial dependence between variables (Moran's I = 0.346), thus the OLS model could be biased.

Furthermore, results showed that, above the spatial regression models, the more suitable spatial model tested was the Spatial Error Model (SEM), since AIC and Log likelihood values are respectively lower and higher than in the Spatial Lag Model (SLM). The Breusch-Pagan test on the spatial effects, calculated for testing the homoscedasticity hypothesis, showed that the null hypothesis was confirmed (Table 3).

**Table 3.** Ordinary Least Squared (OLS) and Spatial Error Models (SEM), complete sample (Source: Authors' elaboration).

Variables	New Housing Listing Prices per Square Meter (LogLP)			New Housing Listing Prices per Square Meter (LogLP)		
	Ordinary Least Squares (OLS)			Spatial Error Model (SEM)		
	Coefficients	Probability		Coefficients	Probability	
NRM	−0.009	0.331		−0.021	0.001	***
NBT	0.070	0.000	***	0.034	0.001	***
LTR = 0			Omitted			
LTR = 1	0.081	0.000	***	0.085	0.000	***
NVW	0.017	0.170		0.031	0.001	***
CSS = 0			Omitted			
CSS = 1	0.128	0.000	***	0.047	0.001	***
PNT = 0			Omitted			
PNT = 1	0.092	0.014	*	0.090	0.000	***
CBX = 0			Omitted			
CBX = 1	0.019	0.293		0.057	0.000	***
LFT = 0			Omitted			
LFT = 1	0.069	0.000	***	0.111	0.000	***
CTP = 1	0.016	0.392		−0.034	0.013	*
CTP = 2	−0.100	0.000	***	−0.070	0.000	***
CTP = 3			Omitted			
MTL = 1	−0.053	0.031	*	−0.057	0.001	***
MTL = 2			Omitted			
MTL = 3	0.099	0.000	***	0.096	0.000	***
MTL = 4	0.199	0.000	***	0.181	0.000	***
BLC = 1			Omitted			
BLC = 2	0.194	0.000	***	0.100	0.001	***
BLC = 3	0.388	0.000	***	0.159	0.000	***
BLC = 4	0.606	0.000	***	0.205	0.000	***
BLC = 5	0.831	0.000	***	0.251	0.000	***
BLT = 0			Omitted			
BLT = 1	0.088	0.005	**	0.096	0.000	***
ARC = 0			Omitted			
ARC = 1	0.054	0.017	*	0.045	0.004	**
EPC = 3	0.025	0.585		0.062	0.050	*
EPC = 2			Omitted			
EPC = 1	−0.033	0.018	*	−0.027	0.005	**
LAMBDA	-	-		0.797	0.000	***
Constant	6.900	0.000		7.184	0.000	
Number of observations	2092.000			2092.000		
Log likelihood	−478.616			22.744		
AIC	1001.230			−1.487		
<b>R square</b>	0.537			<b>0.754</b>		
Adj R square	0.532			-		
Breusch-Pagan test	98.471	0.000		84.693	0.000	
Likelihood ratio test				1002.719	0.000	
DIAGNOSTICS FOR SPATIAL DEPENDENCE FOR WEIGHT MATRIX: LP_2015–2018 2092_ (row-standardized weights)						
TEST	MI/DF	VALUE	PROB			
Moran's I (error)	0.346	27.118	0.000			
Lagrange Multiplier (lag)	1.000	1279.516	0.000			
Robust LM (lag)	1.000	560.604	0.000			
Lagrange Multiplier (error)	1.000	722.496	0.000			
Robust LM (error)	1.000	3.584	0.058			
Lagrange Multiplier (SARMA)	2.000	1283.100	0.000			

Signif codes:  $p \leq 0.001$  '\*\*\*',  $p \leq 0.01$  '\*\*',  $p \leq 0.05$  '\*',  $p \leq 0.1$  '.';  $p \leq 1$  ''.

Results showed that the SEM model can explain 75% of the price variation (R squared = 0.754) and regression residuals were not clustered. This result is rather good considering that the aim of this study was not to predict property values, but just to study the influence

of a set of characteristics on property prices. The significant variables with the higher marginal prices were the presence of elevator (LFT (1) = 0.111), the renovated condition of the unit (MTL (4) = 0.181), the category of the building, that means prestigious buildings (BLC (5) = 0.250) and the coefficient on the spatially correlated errors (LAMBDA = 0.797), which were able to manage the spatial dependence in the model.

On the other hand, in relation with the characteristics of the housing units, a clear, even if small, influence of the EPC label and the presence of an air conditioning system on housing prices emerged. In particular, low EPC labels from G to E (EPC (1) = −0.027) significantly and negatively affected housing prices, while high EPC labels from B to A4 (EPC (3) = 0.062) had a positive influence on them, and the presence of air conditioning had a positive small marginal coefficient (ARC = 0.045). These results are in line with the recent literature related to the perception of the EPC label in the real estate market, which is an ongoing research field [74]. For example, Chen and Marmolejo-Duarte highlighted that in Barcelona, an energy performance improvement from label G to label A brought in a growth of 8.6% of housing prices in 2014, and an increase of 10.6% from label G to label B in 2016 [75]. Moreover, Dell'Anna et al. estimated an increase given by the EPC of 6.33% for each rating level from G to A in Turin for the 2014–2018 time period [11].

By comparing the results of the current analysis with two previous researches based on the same study area (Turin) with the same source of data (TREMOS) [12,13] and with other recent literature [54,75,76], it is possible to confirm that the attention on the EPC label rose slightly in recent years and is starting to be monetized by the real estate market, not only for the lower labels (E, F, G) but also for the higher ones (A–A4, B).

This means that the international energy policies and regulations transposed into the national laws and legal constraints (EPC is mandatory for all transaction contracts since 2012, and minimum energy requirements are mandatory for all types of building interventions in Turin since 2007) applied on the residential building stock in the city of Turin are finally starting to be recognized by the real estate market.

#### 5.2.1. Spatial Clustering: “Northern Zone” and “Central Zone” Sub-Samples

A further step was carried out by applying the SEM on both the CZ and the NZ sub-samples, which were previously defined. The CZ sub-sample included 886 observations, while the NZ sub-sample included 660 observations. Both spatial regressions assumed the logarithm of unitary listing price (LogLP) as a dependent variable, while, on the basis of the significance of the independent variables (Stepwise selection), different intrinsic and extrinsic characteristics were included in the two final models (Table 4).

By comparing the results of the two models, the following variables were significant both in the NZ and in the CZ sub-sample: the presence of Terrace and Lift, the Maintenance Level, the Building Type and the 1946–1975 Construction Time Period (CTP (2)). Furthermore, the presence of Custodian Service and Air Conditioning variables were significant only in the NZ sub-sample, while the Allocation Level, the Number of Bedrooms and Bathrooms, the Number of Views, and the classification of the unit as a Penthouse were significant in the CZ sub-sample. Although the Building Category represents one of the most significant variables, there are also some differences in its significance in the two sub-samples: in the NZ sub-sample almost all levels were significant, while in the CZ only higher levels (Noble (4) and Prestigious (5)) were significant. The analysis of the EPC label highlighted that the high EPC labels (3) were significant only in the CZ sub-sample, while low EPC labels (1) had a negative influence on the price formation process only in the NZ sub-sample.

**Table 4.** Spatial Error Model, Central Zone (CZ) and Northern Zone (NZ) sub-samples (Source: Authors' Elaboration).

Listing Prices per Square Meter (LogLP)CENTRAL ZONE				Listing Prices per Square Meter (LogLP)NORTHERN ZONE			
Variables	Spatial Error Model (SEM)			Variables	Spatial Error Model (SEM)		
	Coefficients	Probability			Coefficients	Probability	
NRM	−0.033	0.000	***	CSS = 1	0.082	0.025	*
NBT	0.043	0.001	***	CBX = 1	0.063	0.025	*
LTR = 1	0.093	0.000	***	LTR = 1	0.105	0.001	***
NVW	0.029	0.019	**	LFT = 1	0.144	0.000	***
ALL = 2	0.06	0.001	***	MTL = 1	−0.093	0.002	***
ALL = 3	0.054	0.001	***	MTL = 3	0.089	0.001	***
PNT = 1	0.096	0.003	**	MTL = 4	0.21	0.000	***
LFT = 1	0.091	0.000	***	BLT = 1	0.217	0.000	***
MTL = 3	0.101	0.000	***	BLC = 4	0.266	0.000	***
MTL = 4	0.154	0.000	***	BLC = 5	0.281	0.001	***
BLT = 1	0.072	0.036	*	BLC = 2	0.125	0.009	**
BLC = 4	0.06	0.000	***	BLC = 3	0.18	0.000	***
BLC = 5	0.138	0.000	***	CTP = 2	−0.069	0.001	**
CTP = 2	−0.045	0.015	*	EPC = 3	0.039	0.544	
EPC = 3	0.079	0.054	.	EPC = 1	−0.037	0.049	*
EPC = 2	0.031	0.034	*	ARC = 1	0.083	0.014	*
ARC = 1	0.024	0.245		LAMBDA	0.627	0.000	***
LAMBDA	0.65	0.000	***	Constant	6.919	0.000	***
Constant	7.581	0.000	***				
N of observations	886	-		N of observations	660	-	
Log likelihood	119.010	-		Log likelihood	−20.489	-	
AIC	−202.021	-		AIC	74.977	-	
R square	0.583	-		R square	0.603	-	
Breush-Pagan test	36.143	0.004		Breush-Pagan test	36.118	0.003	
Likelihood ratio test	209.338	0.000		Likelihood ratio test	167.224	0.000	

Signif. codes:  $p \leq 0.001$  '\*\*\*';  $p \leq 0.01$  '\*\*';  $p \leq 0.05$  '\*';  $p \leq 0.1$  '.';  $p \leq 1$  ''.

In conclusion, the results confirmed that different spatial clusters behave as different sub-markets: they both include different listing prices and building categories and are able to differently monetize the intrinsic characteristics of the buildings.

Firstly, there is a clear connection between the impact of EPC label on housing prices and the quality of the buildings and apartments. Namely, dwellings with a high energy-label (EPC = A), which slightly impacts housing prices, are more expensive and boast the best architectural attributes. Moreover, those dwellings with high EPC labels are located in buildings that are older than 60 years and are probably totally refurbished, both being mainly historical buildings listed by Superintendence of Archaeology, Fine Arts and Landscape, and being regularly under ordinary maintenance by the owners, often representative of the richest population.

Secondly, some attributes totally differed in the two sub-samples. The characteristic related to the air conditioning system is emblematic: it is not monetized by the city centre sub-market as it is a common dwelling's feature, in order to guarantee a high internal comfort to almost all residential units (mainly classy or prestigious). On the opposite hand, air conditioning is a stronger discriminating factor for dwelling buying decisions in the northern suburbs, where the housing quality is lower and for which the "housing added services" seems to compensate by acquiring the greatest importance.

The other dimensional and physical features seem related to the micro-context of the dwelling. In fact, in the city centre the building oldness and density allow the presence of non-ordinary apartment plans, with no balconies or with a solely principal view. Furthermore, the building density means that classy buildings are widely closed, and medium or low floor levels have low lighting and face the opposite buildings. Thus, the presence of a

terrace (in the penthouse) or more than one view are appreciated features, which emerged in previous studies [77].

Even these last considerations which emerged from the study seem to confirm the results reported in the most recent literature (previously mentioned in the Section 2), which reveal an uneven importance of building features—green features included—in view of the building energy performance and, jointly, to the reference sub-market. Then, the results can allow us to argue that, on the one side, some attributes are able to capture a part of the EPC and/or the energy performance features, in terms of marginal prices, for reasons that should be furtherly explored (for example, by analysing the awareness degree of the EPC potentialities among buyers); on the other side, to argue the presence of “latent” variables able to act as a proxy of the building/dwelling energy rating, due to the knowledge degree of the EPC potentialities which could be surely reinforced.

### 5.2.2. EPC Clustering: Low Energy-Efficient Dwellings Sub-Sample

A further step of the analysis was conducted by considering the LEE sub-sample, consisting of 234 observations related to housing units with low EPC labels (E, F and G), built in the 1946–1990 time period. The identification of this not-spatial sub-sample, based on an attribute selection, moved from the results of the previous spatial regressions, which highlighted that the low EPC labels (EPC = 1) and the buildings built between 1946 and 1990 (CTP = 2), to two variables that significantly and negatively influence the listing price formation.

A spatial regression (SEM) was performed also on this sub-sample, assuming LogLP as the dependent variable and different intrinsic features as independent variables. The results (final model with only significant variables) are illustrated in Table 5.

**Table 5.** Spatial Error Model (SEM) Low Energy Efficient (LEE) Sub-Sample. (Source: Authors' elaboration).

Variables	New Housing Listing Prices per Square Meter (LogLP)		
	Spatial Error Model (SEM)		
	Coefficients	Probability	
LTR = 1	0.149	0.001	***
CSS = 1	0.146	0.003	**
CBX = 1	0.114	0.004	**
LFT = 1	0.090	0.016	*
MTL = 3	0.116	0.000	***
MTL = 4	0.189	0.000	***
BLC = 3	0.128	0.000	***
BLC = 4	0.214	0.000	***
BLC = 5	0.387	0.001	***
ARC = 1	0.143	0.005	**
LAMBDA	0.739	0.000	***
Constant	7.082	0.000	***
Number of observations	234		
Log likelihood	0.356		
AIC	21.287		
R square	0.718		
Breush-Pagan test	9.599	0.476	
Likelihood ratio test	81.314	0.000	

Signif. codes:  $p \leq 0.001$  '\*\*\*';  $p \leq 0.01$  '\*\*';  $p \leq 0.05$  '\*';  $p \leq 0.1$  '.';  $p \leq 1$  ''

Above the numerous variables and their related levels, the final SEM can explain the 72% (R squared = 0.718) of the price formation process; results showed the spatial LAMBDA indicator (Lambda = 0.739), the prestigious (BLC (5) = 0.387) to medium level (BLC (3) = 0.128) of building category the high level of maintenance (MTL (4) = 0.189), and the presence of terrace, lift, box and custodian service as the most influential variables.

In conclusion, in this case the spatially correlated errors indicator (expressed by the LAMDA variable) and the “building quality” (expressed by the prestigious building category) are able to explain the greatest part of the price variability. It is interesting to notice that in this sub-sample of energy inefficient housing, refurbished dwellings are also included with a good level of maintenance and prestigious building category, confirming that the energy efficiency and the techno-physical features related to the energy behavior of a building, in the case of existing and historical buildings, are often unrelated to the building architectural quality. Moreover, also for the sub-sample of existing assets, the “housing added services” such as the presence of the appurtenance car box, the presence of lift and custodian service are attributes appreciated by the market.

Finally, it is worth highlighting the presence of two variables probably more related with the present climate changes: the presence of terrace/balcony and the presence of air conditioning. Since the average temperatures increased significantly in Turin from +8.07 °C (in 1963) to +10.69 °C (in 2018) [78], devices (passive and active) related to home cooling are increasingly appreciated by the market.

## 6. Conclusions

This paper aimed to explore the pricing processes of the real estate residential properties, considering the influence of the EPC labels and relevant building/dwelling attributes on prices in their spatial context. The study was developed by analysing a sample of residential listing prices in Turin. Exploratory Spatial Data Analyses (ESDA) and two regression models (the Ordinary Least Squares model and the Spatial Error Model) were firstly applied on the whole data sample.

The results highlighted that low EPC labels (E, F and G) significantly and negatively affect housing prices, while high EPC labels (B, A1, A2, A3 and A4) have a lower but positive influence on them. Moreover, some intrinsic building/dwelling features emerged as characteristics particularly able to influence the property price formation: the building category and the housing unit maintenance level.

Secondly, the regression models were applied on three different sub-samples: two of them were generated by a spatial approach—based on LISA cluster maps of LP—while the third was identified by applying an attribute (not spatial) approach based on intrinsic features, representative of the most energy inefficient housing units. The results showed that high EPC labels are significant only in the “high-high” spatial cluster characterized by a positive autocorrelation of high LP values, while low EPC labels have a negative influence on the price formation process only in the “low-low” spatial cluster, characterized by a positive autocorrelation of low LP values. By comparing the influence of the building/dwelling attributes on property prices in these two spatial sub-samples, some variables resulted as always significant, but several differences emerged, which means that different spatial clusters behave as different sub-markets.

The results achieved in the third sub-sample, which included low energy efficient dwellings, on the one hand confirmed the great influence of prestigious “building category” and high levels of maintenance, and on the other highlighted the relevance of other variables, such as the presence of a terrace, lift, box, custodian service and air-conditioning system. Therefore, the results allow for arguing that the EPC labels are acquiring an explicative power on property prices, in line with some international studies. Specifically, by comparing the results with two previous studies conducted in the city of Turin in two different periods [12,13], a growing effect of EPCs on prices can be seen.

Furthermore, some building features related to the energy performance emerged which deserve to be deeply explored in future, being connected to two relevant aspects: on one side the capacity of certain features, including the “green” ones, to capture a part of the EPC’s explicative power; on the other side, the eventual presence of “latent” variables acting as a proxy of the building/dwelling energy rating.

Operatively, the methodological approach illustrated in this work can be considered a support for building/dwelling retrofitting policies, considering the potentialities of

technological interventions in terms of dynamics in sub-markets and prices. Being in a period of deep changes in the real estate market, this study demonstrates that spatial analyses jointly with the Hedonic pricing models represent important approaches to study the real estate sub-markets and explore their different pricing processes. In fact, spatial regression models permit to manage the spatial autocorrelation and to study the influence of latent variables on pricing processes, which represent important factors able to change all spatial hierarchies.

**Limitations:** In the Italian context property transaction prices are not public information. Thus, researchers, real estate companies and public administrations are used to study and analyze listing prices to perform market analyses and to estimate the property values. Even if it represents a key limitation of this study, it is worth mentioning that previous studies demonstrated that listing prices can be considered a proxy for transaction prices [79] and that they can influence the selling processes and prices prediction [68,69]. Moreover, the data sample consists in a set of property listings containing detailed information that in some cases are not complete: this is the case of the building construction time period and also of the EPC label.

**Author Contributions:** This paper is to be attributed in equal parts to the authors. Conceptualization, A.B., E.F. and D.R.; data curation, A.B. and D.R.; formal analysis, A.B. and D.R.; investigation, A.B., E.F. and D.R.; methodology, A.B., E.F. and D.R.; supervision, E.F.; writing—original draft, A.B., E.F. and D.R.; writing—review & editing, A.B., E.F. and D.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** TREMO data (housing listing prices and related characteristics), used to support the findings of this study, have not been made available because of third-party rights.

**Acknowledgments:** Acknowledgement goes to the anonymous reviewers who have contributed with their suggestions to improve the quality of this paper, to the academic editor and to all MDPI staff for their valuable work.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Charalambides, A.G.; Maxoulis, C.N.; Kyriacou, O.; Blakeley, E.; Frances, L.S. The impact of Energy Performance Certificates on building deep energy renovation targets. *Int. J. Sustain. Energy* **2019**, *38*, 1–12. [[CrossRef](#)]
2. Brounen, D.; Kok, N. On the economics of energy labels in the housing market. *J. Environ. Econ. Manag.* **2011**, *62*, 166–179. [[CrossRef](#)]
3. Fuerst, F.; McAllister, P.; Nanda, A.; Wyatt, P. Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England. *Energy Econ.* **2015**, *48*, 145–156. [[CrossRef](#)]
4. Fuerst, F.; McAllister, P.; Nanda, A.; Wyatt, P. Energy performance ratings and house prices in Wales: An empirical study. *Energy Policy* **2016**, *92*, 20–33. [[CrossRef](#)]
5. Davis, P.T.; McCord, J.; McCord, M.J.; Haran, M. Modelling the effect of energy performance certificate rating on property value in the Belfast housing market. *Int. J. Hous. Mark. Anal.* **2015**, *8*, 292–317. [[CrossRef](#)]
6. Murphy, L. The influence of the energy performance certificate: The Dutch case. *Energy Policy* **2014**, *67*, 664–672. [[CrossRef](#)]
7. Högberg, L. The impact of energy performance on single-family home selling prices in Sweden. *J. Eur. Real Estate Res.* **2013**, *6*, 242–261. [[CrossRef](#)]
8. Bottero, M.; Bravi, M.; Dell'Anna, F.; Mondini, G. Valuing buildings energy efficiency through Hedonic Prices Method: Are spatial effects relevant? *Valori Valutazioni* **2018**, *21*, 27–39.
9. Chen, A.; Marmolejo-Duarte, C. How different are dwellings whose energy efficiency impacts price formation? *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *603*, 042015. [[CrossRef](#)]
10. Marmolejo-Duarte, C.; del Moral, C.; Delgado, L.; Spairani, S.; de Botton, J.; Pérez, C.; Chen, A.; Gyurkovich, M. Energy efficiency in the residential market and implications for architecture education in Spain. *World Trans. Eng. Technol. Educ.* **2019**, *17*, 225–230.
11. Dell'Anna, F.; Bravi, M.; Marmolejo-Duarte, C.; Bottero, M.C.; Chen, A. EPC green premium in two different European climate zones: A comparative study between Barcelona and Turin. *Sustainability* **2019**, *11*, 5605. [[CrossRef](#)]
12. Fregonara, E.; Rolando, D.; Semeraro, P.; Vella, M. The impact of Energy Performance Certificate level on house listing prices. First evidence from Italian real estate. *Aestimium* **2014**, *65*, 143.

13. Fregonara, E.; Rolando, D.; Semeraro, P. Energy performance certificates in the Turin real estate market. *J. Eur. Real Estate Res.* **2017**, *10*, 149–169. [[CrossRef](#)]
14. Wilkinson, S.J.; Sayce, S. Decarbonising real estate: The evolving relationship between energy efficiency and housing in Europe. *J. Eur. Real Estate Res.* **2020**, *13*, 387–408. [[CrossRef](#)]
15. Dell’Anna, F.; Bottero, M. Green premium in buildings: Evidence from the real estate market of Singapore. *J. Clean. Prod.* **2021**, *286*, 125327.
16. Bisello, A.; Antoniucci, V.; Marella, G. Measuring the price premium of energy efficiency: A two-step analysis in the Italian housing market. *Energy Build.* **2020**, *208*, 109670. [[CrossRef](#)]
17. Brounen, D.; Alexander, M.G.; Martin, H. The value effects of green retrofits. *J. Eur. Real Estate Res.* **2020**, *13*, 10–20. [[CrossRef](#)]
18. McCord, M.; Davis, P.; McCord, J.; Haran, M.; Davison, K. An exploratory investigation into the relationship between energy performance certificates and sales price: A polytomous universal model approach. *J. Financ. Manag. Prop. Constr.* **2020**, *25*, 247–271. [[CrossRef](#)]
19. McCord, M.; Haran, M.; Davis, P.; McCord, J. Energy performance certificates and house prices: A quantile regression approach. *J. Eur. Real Estate Res.* **2020**. [[CrossRef](#)]
20. Morano, P.; Rosato, P.; Tajani, F.; Di Liddo, F. An analysis of the energy efficiency impacts on the residential property prices in the city of Bari (Italy). In *Values and Functions for Future Cities*; Springer: Cham, Switzerland, 2020; pp. 73–88.
21. Manganelli, B.; Morano, N. Comparative performance of Structural Equations System and Neural Networks for real estate appraisal. In *Proceeding of the 7th SIGEF Congress, New Logis For the New Economy*; Edizioni Scientifiche Italiane: Roma, Italy, 2001.
22. Liu, Y.; Wu, Y.X. Analysis of residential product’s value based on structural equation model and hedonic price theory. In *Proceedings of the 2009 International Conference on Management Science and Engineering*, Moscow, Russia, 14–16 September 2009; pp. 1950–1956.
23. Freeman, J.; Zhao, X. An SEM Approach to Modeling Housing Values. In *Data Analysis and Applications 1: Clustering and Regression, Modeling-estimating, Forecasting and Data Mining*; Wiley: Hoboken, NJ, USA, 2019; Volume 2, pp. 125–135.
24. Morano, P.; Tajani, F. Least median of squares regression and minimum volume ellipsoid estimator for outliers detection in housing appraisal. *Int. J. Bus. Intell. Data Min.* **2014**, *9*, 91–111. [[CrossRef](#)]
25. Morano, P.; Tajani, F.; Locurcio, M. Multicriteria analysis and genetic algorithms for mass appraisals in the Italian property market. *Int. J. Hous. Mark. Anal.* **2018**, *11*, 229–262. [[CrossRef](#)]
26. Locurcio, M.; Morano, P.; Tajani, F.; Liddo, F. Di An innovative GIS-based territorial information tool for the evaluation of corporate properties: An application to the Italian context. *Sustainability* **2020**, *12*, 5836. [[CrossRef](#)]
27. Rosen, S. Hedonic prices and implicit markets: Product differentiation in pure competition. *J. Polit. Econ.* **1974**, *82*, 34–55. [[CrossRef](#)]
28. Ridker, R.G.; Henning, J.A. The determinants of residential property values with special reference to air pollution. *Rev. Econ. Stat.* **1967**, *49*, 246–257. [[CrossRef](#)]
29. Malpezzi, S. Hedonic Pricing Models: A Selective and Applied Review. In *Housing Economics and Public Policy*; Wiley: Hoboken, NJ, USA, 2002; pp. 67–89.
30. Basu, S.; Thibodeau, T.G. Analysis of Spatial Autocorrelation in House Prices. *J. Real Estate Financ. Econ.* **1998**, *17*, 61–85. [[CrossRef](#)]
31. Goodman, A.C.; Thibodeau, T.G. Housing Market Segmentation. *J. Hous. Econ.* **1998**, *7*, 121–143. [[CrossRef](#)]
32. Bourassa, S.C.; Cantoni, E.; Hoesli, M. Predicting House Prices with Spatial Dependence: A Comparison of Alternative Methods. *J. Real Estate Res.* **2010**, *32*, 139–159. [[CrossRef](#)]
33. Bourassa, S.C.; Cantoni, E.; Hoesli, M. Spatial dependence, housing submarkets, and house price prediction. *J. Real Estate Financ. Econ.* **2007**, *35*, 143–160. [[CrossRef](#)]
34. Bourassa, S.C.; Hoesli, M.; Peng, V.S. Do housing submarkets really matter? *J. Hous. Econ.* **2003**, *12*, 12–28. [[CrossRef](#)]
35. Bourassa, S.C.; Hoesli, M.; Scognamiglio, D.; Zhang, S. Land leverage and house prices. *Reg. Sci. Urban Econ.* **2011**, *41*, 134–144. [[CrossRef](#)]
36. Cajias, M.; Ertl, S. Spatial effects and non-linearity in hedonic modeling. *J. Prop. Invest. Financ.* **2018**, *36*, 32–49. [[CrossRef](#)]
37. Getis, A. A history of the concept of spatial autocorrelation: A geographer’s perspective. *Geogr. Anal.* **2008**, *40*, 297–309. [[CrossRef](#)]
38. Seymour, L. Spatial Data Analysis: Theory and Practice. *J. Am. Stat. Assoc.* **2005**, *100*, 353. [[CrossRef](#)]
39. LeSage, J.P.; Pace, R.K. An Introduction to Spatial Econometrics. *Rev. Econ. Ind.* **2008**, *123*, 19–44. [[CrossRef](#)]
40. Wilhelmsson, M. Spatial models in real estate economics. *Hous. Theory Soc.* **2002**, *19*, 92–101. [[CrossRef](#)]
41. Dubin, R.A. Spatial autocorrelation and neighborhood quality. *Reg. Sci. Urban Econ.* **1992**, *22*, 433–452. [[CrossRef](#)]
42. LeSage, J.; Pace, R.K. *Introduction to Spatial Econometrics*; Routledge: London, UK, 2009.
43. Can, A. Specification and estimation of hedonic housing price models. *Reg. Sci. Urban Econ.* **1992**, *22*, 453–474. [[CrossRef](#)]
44. Anselin, L.; Bera, A.K. Spatial dependence in linear regression models with an introduction to spatial econometrics. In *Handbook of Applied Economic Statistics*; Hullah, A., Gelis, D.E.A., Eds.; Marcel Dekker: New York, NY, USA, 1998; pp. 237–290.
45. Seya, H.; Yamagata, Y.; Tsutsumi, M. Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach. *Reg. Sci. Urban Econ.* **2013**, *43*, 429–444. [[CrossRef](#)]
46. Fernández-Delgado, M.; Sirsat, M.S.; Cernadas, E.; Alawadi, S.; Barro, S.; Febrero-Bande, M. An extensive experimental survey of regression methods. *Neural Netw.* **2019**, *111*, 11–34. [[CrossRef](#)]

47. Goodchild, M.F. Concepts and Techniques in Modern Geography. In *Spatial Autocorrelation*; Geo Books: Norwich, UK, 1986; pp. 1–56.
48. Ismail, S. Spatial Autocorrelation and Real Estate Studies: A Literature Review. *Malays. J. Real Estate* **2006**, *1*, 1–13.
49. European Parliament. Directive 2002/91/EC of the European Parliament and of the Council 6 December 2002 on the Energy Performance of Buildings. *Off. J. Eur. Union* **2003**, *1*, 65–70.
50. European Parliament. Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings. *Off. J. Eur. Union* **2010**, *3*, 126–142.
51. European Commission. Commission Delegated Regulation (EU) No 244/2012 of 16 January 2012. *Off. J. Eur. Union* **2012**, *5*, 18–36.
52. Becchio, C.; Ferrando, D.G.; Fregonara, E.; Milani, N.; Quercia, C.; Serra, V. The cost-optimal methodology for the energy retrofit of an ex-industrial building located in Northern Italy. *Energy Build.* **2016**, *127*, 590–602. [[CrossRef](#)]
53. European Parliament. Directive (EU) 2018/844 of the European parliament and of the council. *Off. J. Eur. Union* **2018**, *6*, 75–91.
54. Marmolejo-Duarte, C.; Chen, A.; Bravi, M. Spatial implications of EPC rankings over residential prices. In *Values and Functions for Future Cities*; Springer: Cham, Switzerland, 2020; pp. 51–71.
55. Barreca, A.; Curto, R.; Rolando, D. Assessing social and territorial vulnerability on real estate submarkets. *Buildings* **2017**, *7*, 94. [[CrossRef](#)]
56. Barreca, A.; Curto, R.; Rolando, D. Housing vulnerability and property prices: Spatial analyses in the Turin real estate market. *Sustainability* **2018**, *10*, 3068. [[CrossRef](#)]
57. Barreca, A.; Curto, R.; Rolando, D. Urban vibrancy: An emerging factor that spatially influences the real estate market. *Sustainability* **2020**, *12*, 346. [[CrossRef](#)]
58. Barreca, A.; Curto, R.; Rolando, D. Is the Real Estate Market of New Housing Stock Influenced by Urban Vibrancy? *Complexity* **2020**, *2020*, 1908698. [[CrossRef](#)]
59. Robertson, K.; Doig, A. An empirical investigation of variations in real-estate marketing language over a market cycle. *Hous. Theory Soc.* **2010**, *27*, 178–189. [[CrossRef](#)]
60. Semeraro, P.; Fregonara, E. The impact of house characteristics on the bargaining outcome. *J. Eur. Real Estate Res.* **2013**, *6*, 262–278. [[CrossRef](#)]
61. Goodchild, M. *Spatial Autocorrelation*; GeoBooks: Norwich, UK, 1986.
62. Haining, R.; Wise, S.; Ma, J. Exploratory spatial data analysis in a geographic information system environment. *J. R. Stat. Soc. Ser. D Stat.* **1998**, *47*, 457–469. [[CrossRef](#)]
63. Anselin, L.; Rey, S. Properties of Tests for Spatial Dependence in Linear Regression Models. *Geogr. Anal.* **1991**, *23*, 112–131. [[CrossRef](#)]
64. Anselin, L. Local Indicators of Spatial Association—LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [[CrossRef](#)]
65. Anselin, L.; Syabri, I.; Kho, Y. GeoDa: An introduction to spatial data analysis. *Geogr. Anal.* **2006**, *38*, 5–22. [[CrossRef](#)]
66. Baltagi, B.H.; Heun Song, S.; Cheol Jung, B.; Koh, W. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *J. Econom.* **2007**, *140*, 5–51. [[CrossRef](#)]
67. Anselin, L.; Griffith, D.A. Do spatial effect really matter in regression analysis? *Pap. Reg. Sci.* **1988**, *65*, 11–34. [[CrossRef](#)]
68. Knight, J.R. Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Econ.* **2002**, *30*, 213–237. [[CrossRef](#)]
69. Horowitz, J.L. The role of the list price in housing markets: Theory and an econometric model. *J. Appl. Econom.* **1992**, *7*, 115–129. [[CrossRef](#)]
70. Knight, J.R.; Sirmans, C.F.; Turnbull, J.K. List price information in residential appraisal and underwriting. *J. Real Estate Res.* **1998**, *15*, 59–76. [[CrossRef](#)]
71. Song, S. Determinants of Bargaining Outcome in Single-family Housing Transactions: An Empirical Examination. *Urban Stud.* **1995**, *32*, 605–614. [[CrossRef](#)]
72. Curto, R.; Fregonara, E.; Semeraro, P. Prezzi di offerta vs prezzi di mercato: Un’analisi empirica. Asking Prices vs. Market Prices: An Empirical Analysis. *Territ. Ital.* **2012**, *2*, 20–40.
73. Anglin, P.M.; Rutherford, R.; Springer, T.M. The Trade-off between the Selling Price of Residential Properties and Time-on-the-Market: The Impact of Price Setting. *J. Real Estate Financ. Econ.* **2003**, *26*, 95–111. [[CrossRef](#)]
74. Marmolejo-Duarte, C.; Spairani Berrio, S.; Del Moral-Ávila, C.; Delgado Méndez, L. The Relevance of EPC Labels in the Spanish Residential Market: The Perspective of Real Estate Agents. *Buildings* **2020**, *10*, 27. [[CrossRef](#)]
75. Marmolejo-Duarte, C.; Chen, A. The Uneven Price Impact of Energy Efficiency Ratings on Housing Segments and Implications for Public Policy and Private Markets. *Sustainability* **2019**, *11*, 372. [[CrossRef](#)]
76. Bottero, M.; Bravi, M.; Dell’Anna, F.; Marmolejo-Duarte, C. Energy efficiency choices and residential sector: Observable behaviors and valuation models. In *Green Energy and Technology*; Springer: Cham, Switzerland, 2020; pp. 167–179.
77. Lo Verso, V.R.M.; Fregonara, E.; Caffaro, F.; Morisano, C.; Maria Peiretti, G. Daylighting as the driving force of the design process: From the results of a survey to the implementation into an advanced daylighting project. *J. Daylighting* **2014**, *1*, 36–55. [[CrossRef](#)]
78. Arpa Piemonte. *Città di Torino. Assessment Climatico e Scenari Futuri*; Arpa Piemonte: Torino, Italy, 2018; p. 104.
79. Curto, R.; Fregonara, E.; Semeraro, P. Asking prices vs. Market Prices: An Empirical Analysis. *Territ. Ital.* **2012**, *12*, 53–72.