DESIGNING A PRICE INDEX FOR THE SPANISH COMMERCIAL REAL ESTATE MARKET (*)

Matias Lamas
BANCO DE ESPAÑA

Sara Romaniega (**)
EUROPEAN CENTRAL BANK

(*) The authors are grateful for the comments of Roberto Blanco, Carmen Broto, Ángel Estrada, Luis Gonzalo Fernández, Jorge Galán, Salomón Stalin García, José Manuel González, Dino López, David Martínez-Miera, Javier Mencía, Carlos Pérez, Irene Roibás, Lucio San Juan and Alberto Urtasun, as well as the suggestions of those attending the meetings of the Banco de España’s Real Estate Market Monitoring Group and Financial Stability and Macroprudential Policy Department, whose contributions have helped improve this paper. The authors would also like to extend particular thanks to Joaquín Hidalgo, of the Spanish Association of Registrars’ Statistical Processes Centre, for his excellent explanations about records of property sales. Lastly, the authors would like to thank the National Statistics Institute for its comments on this paper. Any errors are the authors’ sole responsibility.

(**) Sara Romaniega worked in the Banco de España’s Financial Stability and Macroprudential Policy Department between March and July 2021, as part of a work experience programme for students.
Abstract

This paper proposes a price index for the Spanish commercial real estate (CRE) market and its main segments. No official price indices currently exist for these assets and non-official statistics are thin on the ground and offer limited coverage. The estimated index fills this statistical gap, providing for enhanced monitoring of CRE market developments. The price indicator draws on the methodology used to calculate the Spanish National Statistics Institute’s housing market price indices, but factoring in certain elements specific to the CRE sector. Various alternative indices are also considered, although the indicator proposed is that which strikes the best balance between the different statistical properties analysed here.

Keywords: commercial real estate market, price indices, hedonic regression model.

JEL classification: R33.
Resumen

En este documento se propone un índice de precios para el mercado inmobiliario comercial de España y para los principales segmentos que lo componen. En la actualidad no existen índices de precios oficiales para estos activos, y la disponibilidad de estadísticas no oficiales es escasa y tiene una cobertura limitada. Por lo tanto, el índice estimado viene a cubrir esta laguna estadística y permite realizar un mejor seguimiento de la evolución de este mercado. El índice de precios se apoya en la metodología de cálculo de los índices de precios para el mercado de la vivienda del Instituto Nacional de Estadística, si bien tiene en cuenta algunos elementos específicos del sector inmobiliario comercial. Se han considerado, además, diferentes índices alternativos, aunque el indicador propuesto es el que presenta un mejor balance entre varias propiedades estadísticas analizadas en este trabajo.

Palabras clave: mercado inmobiliario comercial, índices de precios, modelo de regresión hedónica.

Código JEL: R33.
Contents

Abstract  5

Resumen  6

1 Introduction  8

2 CRE market. Definition and considerations  11

3 Price index methodology  13
   3.1 Creating cells with homogeneous properties (stratification)  13
   3.2 Estimating prices in each cell (hedonic regression model)  15
   3.3 Aggregating the data in a single price index  17

4 Price indices for the CRE market  19
   4.1 Overall index  19
   4.2 Indices for each sub-segment  20

5 Assessment of the proposed overall price index against other alternatives  21

6 Conclusions  23

References  24

Annex 1 Regression exercises for the offices and industrial buildings segments  25

Annex 2 Use of various quarters in the hedonic regression models  27

Annex 3 Assessment of the price index against other alternatives  29
1 Introduction

This paper presents a price index for the commercial real estate (CRE) market. This segment comprises properties in which some form of commercial activity is pursued, such as shops or restaurants, workplaces (offices) and industrial buildings (i.e. establishments in which, inter alia, products are manufactured and/or logistical tasks are performed). However, properties under construction are not considered part of this market, nor does it include rented property, whether for professional purposes or otherwise.

Developments in this market can have implications for financial stability. For example, the European Systemic Risk Board (ESRB), the body tasked with overseeing risks to the European Union’s financial system, notes that default rates on CRE mortgages are higher than in the case of residential mortgages (European Systemic Risk Board, 2015). This state of affairs may be explained by factors such as the limited liability of some of the debtors under these mortgages or the fact that entrepreneurial activities are more sensitive to fluctuations in the business cycle.1

Before the pandemic, the ESRB had itself already warned of the growing risks to financial stability stemming from the situation of the CRE sector, some of which appear to have materialised during the crisis (European Systemic Risk Board, 2018). Moreover, the ESRB also identified significant data gaps that prevented a rigorous analysis of developments in this sector, leading to the approval of Recommendation ESRB/2016/14, subsequently amended by Recommendation ESRB/2019/3.2 Both Recommendations point to the need for the different supervisory authorities to improve the availability of data, as well as the tools for analysing risks in the real estate market, and specifically the commercial segment.

Interest in developments in this market has grown during the COVID-19 pandemic. In a recent analysis of global financial stability, the International Monetary Fund (IMF) noted that the containment and social distancing measures set in place may negatively impact demand for these types of establishments, particularly those used for food service activities (International Monetary Fund, 2021). Meanwhile, the accelerated shift towards remote working arrangements may have dampened interest in office space. The IMF also notes that institutional investors and cross-border capital have a role to play in how this sector performs, at least in certain economies. Faced with the heightened uncertainty brought about by the pandemic, changes in investor sentiment or patterns of capital flows could shape the demand for CRE and, by extension, its price.

Against this backdrop, tracking the prices of commercial establishments is key to properly analysing the risks to the CRE sector. In this respect, the information on CRE prices is much harder to come by and, in general, of lower statistical value than the information

---

1 Unlike mortgages extended to individuals, who are liable with all of their present and future assets in the event of default (with certain limits), a business’ shareholders are generally only liable up to the capital invested.
that can be found for the residential market, hence the above ESRB Recommendations. For example, the commercial property price indices drawn up by consultants and other private sector operators tend to use information on sales in a small slice of the market (typically the most expensive or prime areas). Meanwhile, some of these indices are constructed from property appraisal information rather than actual transaction prices. Carried out by the real estate businesses themselves, as opposed to independent appraisal companies, such valuations may not accurately reflect the market value of the properties, particularly where calculated using smoothing techniques (making it hard to identify price inflection points) or not regularly updated. It is also worth bearing in mind that such operators have their own interests in the market; accordingly, private sector price indices may be biased (Bank for International Settlements, 2019; Eurostat, 2017).

Calculating price indicators for the real estate sector poses certain challenges, particularly in the commercial segment. First, a definition of what is meant by a property of this type is needed, a question that has sparked some debate (European Systemic Risk Board, 2015). Section 2 of this paper offers a working definition of this type of property. Second, a price index needs to account for any changes in the composition of the establishments bought or sold in each period. Thus, the heterogeneous characteristics of these establishments (taking in everything from offices and other work spaces to food service outlets or industrial buildings) add an extra layer of difficulty when designing these indices.

This paper proposes a price index for the Spanish CRE sector that seeks to address some of the challenges detailed above, as well as the possible drawbacks of the price indicators offered by the private sector. First, it draws on the databases of the Spanish Association of Registrars (CR, by its Spanish abbreviation), which contain information on all registered property sales, and not just one portion of the market. Then, based on the CR data itself and other databases, highly granular information is gathered on the properties sold (such as their physical attributes or location). As explained below, this is key to constructing a solid index.

The index has been prepared in line with the best practices for constructing property price indicators (Eurostat, 2013 and 2017). Specifically, the methodology used by the National Statistics Institute (INE, by its Spanish abbreviation) to calculate the House Price Index (IPV, by its Spanish abbreviation) has been replicated, albeit making some adjustments to account for the characteristics of the CRE segment. The INE index has been updated quarterly since 2007 Q1, and draws on highly detailed methodological handbooks (National Statistics Institute, 2015).

The main advantage of indices calculated using a methodology in line with the IPV approach is that any changes in the quality of a property are controlled for, and price

---

3 Eurostat has launched an initiative to develop price indices with homogeneous characteristics for the CRE sector in European countries.
4 The index does not include unregistered transactions. Nonetheless, given the legal certainty provided by the property registry, it is safe to assume that the vast majority of sales have been registered.
movements are therefore calculated using properties with homogeneous characteristics. Such indices are calculated in three steps. First, strata or cells are created for the different types of property. Each cell contains properties with homogeneous characteristics. The price of each of these cells is then estimated, using hedonic regression models that assume the existence of an implicit price for each of a property’s characteristics or attributes. The explanatory variables in these specifications are the same as those used to define the cells, meaning that a price can be estimated for each one. Lastly, the resulting information is summarised in several price indices for the key segments of this sector (retail premises, offices and industrial buildings), from which an overall price index is then derived. To this end, the cells are weighted in line with the IPV methodology.

This paper is structured as follows. Section 2 provides a working definition of the CRE sector and sets out a series of preliminary considerations for the construction of the index, concerning the scale and composition of this market in Spain. Section 3 then describes the methodology used to construct the overall index and the sub-indices for the main segments, which are detailed in Section 4. Section 5 assesses the overall index against other alternatives. Finally, Section 6 sets out the main conclusions of the study.
2 CRE market. Definition and considerations

The price index presented in this paper summarises the prices of completed commercial property. In other words, premises on which construction is complete and that are equipped for the pursuit of a commercial activity, whether by the owner or by third parties (where the property is rented out for use). This definition of CRE sector thus excludes establishments under construction, which other organisations do include in this market.5

Compared to the housing market, the Spanish CRE market is relatively small. According to micro data provided by the CR, between 2014 and 2019 CRE transactions amounted to around 14% of housing market transactions (see Chart 1.1). This proportion fell to 12% in 2020, as a result of the greater impact of the pandemic on this segment of the real estate market.6 In terms of the type of establishment, the CR classifies CRE into three categories, based on their principal use on notarisation of the sale. Most transactions involved what the CR refers to as “retail premises”, a category that encompasses a broad-ranging group of properties, including, inter alia, restaurants, supermarkets and department

---

5 The ESRB defines CRE as “income-producing real estate, either existing or under development, including rental housing; or real estate used by the owners of the property for conducting their business, purpose or activity, either existing or under construction; that is not classified as RRE (Residential Real Estate); and includes social housing” (European Systemic Risk Board, 2019). CRE is therefore any type of establishment able to generate income, whether used for the pursuit of a commercial activity by the owner or rented out to third parties, or where there are plans to sell it on completion of the development or construction process.

6 The smaller share of this segment can also be seen in the mortgage exposures of financial institutions. Thus, mortgages secured by properties of this type accounted for 15% of the total mortgage portfolio in the first quarter of 2021 (this total includes residential and commercial mortgages). For a more detailed analysis of the CRE sector in Spain, see Fernández et al. (2021).
stores. “Industrial buildings” (properties used predominantly for industrial purposes or the storage of products) and “offices” (or work spaces) constitute the second and third types of commercial establishments referred to in this analysis (see Chart 1.2).
3 Price index methodology

Quality changes in the sample of properties represents one of the biggest challenges when designing indices for the real estate sector. The methodology used in this paper (very similar to that used to construct the INE’s IPV) enables a comparison of homogeneous properties, since it accounts for their main characteristics. Thus, any changes in the composition of the establishments bought or sold in each period are factored in.

The price index is prepared in line with the steps shown in Figure 1. First, each of the three segments in this sector (retail premises, offices and industrial buildings) are divided into strata or cells containing homogeneous properties. In other words, such cells have similar characteristics (e.g. same built floor area or same type of location, among other criteria). The price of each stratum is then estimated in accordance with a hedonic regression model, which takes as explanatory variables the same attributes used to split the market into cells. The coefficient estimates in the model (calculated quarterly) enable a price to be attributed to the cells defined in the preceding step. Lastly, such information is summarised in a single price index for each segment, based on a statistical procedure. The overall index for the commercial market is obtained by weighting each of the above sub-indices based on the proportion of total sales accounted for by each property type. Each of these steps is described in further detail below.

3.1 Creating cells with homogeneous properties (stratification)

As noted above, the first step in preparing the indices consists of splitting the market into cells in which the characteristics of the properties are very similar. To this end, inter alia, physical and location-related variables are used. This stratification process enables the price of homogeneous units to be compared over time. For example, one of the cells constructed for the retail premises sub-group contains establishments of between 150 and 240 square metres and located in high-income neighbourhoods, among other attributes. All properties with these characteristics are placed in this cell, yielding a homogeneous stratum.
The variables chosen for such stratification are similar to those used by the INE in the IPV, albeit replacing certain attributes for which no information is available and adding others to account for specific features of the CRE sector. To facilitate the creation of strata, all of the variables defined are categorical in nature (i.e. they are indicators that take a limited number of values). Table 1 lists these variables.

The first six indicators in the table are taken from the CR micro data. For the most part, these are binary variables (they include two categories of property), save for the built floor area variable and the variable classifying properties by the province in which they are located (more than two categories). The location variable for prime properties identifies transactions involving premises located in the central business district of various large cities (following the criteria used by various real estate consultancies; see Banco de España, 2020). Other location variables that can shape the price of such properties are the province, the share of commercial and hotel establishments out of total premises in the same municipality, and whether they are located in a tourist area, in line with the INE criteria.

Meanwhile, and as in the INE’s IPV, a categorical variable is used to proxy for the type of neighbourhood in which a property is located. To this end, information on average gross household income (available for the post code in which an establishment is located) is obtained from the tax authorities. It is assumed that the higher (lower) the household income in a post code, the more (less) advantageous the property’s surroundings, and, by extension, the higher (lower) its price. It is worth noting that while the CR does state a property’s post code, this field was very rarely reported until 2014. Thus, the price indices estimated in this paper cannot be reliably calculated before that year.

7 The cities in which prime establishments are located are Barcelona, Bilbao, Madrid, Malaga, Palma and Valencia.
As noted by the INE (2017), the more granular the stratification, the more homogeneous the properties in each cell. Nonetheless, at least one transaction must be included in each cell to avoid creating strata containing non-existent properties. For example, it would make no sense to create a prime establishments cell for low-income postcodes, since such properties are to be found in areas in which average household income is high. With this constraint in mind, all of the possible combinations of the categorical variables in Table 1 yield a total of 12,346 cells in the retail premises segment, 1,034 cells in the offices segment and 1,925 in the industrial buildings segment. Each of these cells refers to properties with homogeneous attributes.

3.2 Estimating prices in each cell (hedonic regression model)

Once the cells have been generated (for the three segments considered), a price is estimated for each one. The literature in this respect offers several alternatives (Eurostat, 2013). The most straightforward of these consists of calculating the average or median price of the transactions in each cell. Though simple, this method may prove unsound where there are “empty” cells or cells with few transactions in certain periods. This is either because the price of an empty cell cannot be calculated for a given period, or because representativeness is limited where a cell contains few sales for a given moment. Elsewhere, Case and Shiller (1987) propose the repeat sales method, which compares sale prices for the same property over time, thus yielding a high level of homogeneity. The drawback to this approach is that only a portion of the sample is used to estimate prices (properties sold more than once over the period under analysis), and it may therefore be skewed.

A hedonic regression model is the method chosen in this paper to allocate prices to the cells. This type of model assumes that the price of a property is a function of its characteristics. In other words, there is an implicit cost associated with each characteristic of a commercial establishment, meaning that the price of a commercial property is the sum total of the “costs” of all of its characteristics. Regressions are estimated for each period, incorporating as explanatory variables the same set of categorical variables used to define the cells or strata into which the market has been split. Thus, the hedonic model yields potentially more reliable estimates of the price of cells, even where the number of observations available for each one is small, and even where there are empty cells or cells without registered sales at a given moment (albeit not for the sample as a whole).

The hedonic regression model is captured by equation [1]:

\[
\ln P_{i,e}^t = X_{i,e}^t \beta^t + \epsilon_{i,e}^t
\]

---

8 In general, there are more transactions per cell in the retail premises segment, since this accounts for the majority of sales. Moreover, more granular built floor area and household income variables (more groups) have been considered for this type of property. This yields a greater number of cells in this segment, as well as a greater number of parameters to be estimated in the regressions (Section 3.2 discusses the regression exercise used to estimate the price indices in this paper).
where the dependent variable, $\ln P_{t,e}^i$, is the natural logarithm of the price per square metre built of property $i$, pertaining to cell $e$, in period $t$. The explanatory variables, grouped in vector $X_{t,e}$, capture the characteristics of the properties. The parameters linking price sensitivity to the different attributes of a property are defined in vector $\beta$. The relationship between prices and characteristics in the cross-section of cells is thus variable over time. To improve the predictive power of this simple model, certain interactions between the explanatory variables of vector $X_{t,e}$ are included. Moreover, to estimate the model, the observations with the largest prediction error or residual are eliminated. Lastly, the White method for obtaining robust standard errors is used to correct potential heteroscedasticity problems in the model.

Although there is some disparity across the segments, CRE prices are largely determined by the average household income in the postcode area and the built floor area of the property (see Chart 2, showing estimates for the retail premises price model; the charts

---

**Chart 2**

**ESTIMATED COEFFICIENTS FOR THE RETAIL PREMISES PRICE MODEL (a)**

![Chart showing estimated coefficients for retail premises price model](chart2)

**SOURCE:** Banco de España calculations.

(a) The chart shows the estimated coefficients (triangles) and the associated confidence intervals (vertical lines, with a confidence level of 95%) of the regression described by equation [1]. The sample used captures all sales of retail premises since 2014. The model also incorporates some interactions between the explanatory variables whose coefficients are not shown in the chart in the interests of brevity. “Buyer” indicates that the buyer is a legal person; “Seller”, that the seller is a legal person; “Resale” indicates that the property has changed hands in the past (approximates to the concept of second-hand property); “m2 (group x)” refers to the floor area bracket (the higher groups/categories have more m2 of floor area); “Prime” indicates that the property is located in a prime area; “Prov. group x” refers to the province group where the property is located (four groups in total); “High CRE” and “High hotel” indicate whether the number of commercial or hotel properties in the same municipality as the property is above median; “Income (group x)” refers to the bracket of average gross household income in the same postcode as the property (the higher the income, the higher the value of x in the chart); lastly, “Tourist mun.” shows the coefficient associated with the property being located in a “tourist municipality” according to the INE. A positive value for any of the estimated coefficients indicates that the characteristic in question is conducive to a higher price per square metre, while a negative value indicates the opposite. For example, in the chart the estimated coefficient for the “Tourist mun.” characteristic is close to 0.25, indicating that property prices in those municipalities are, all other things being equal, 25% higher (approximately) than those elsewhere.

---

9 The number of interactions is different for each segment of the CRE sector. The model with the most interactions is the one on the retail premises segment, for which the floor area and household income variables are more granular. In general, the interactions considered are statistically significant.

10 In other words, to begin with, the model is estimated, eliminating the observations with the largest prediction error. The model is then estimated again using the smaller set of transactions. The parameters estimated at this stage are those obtained without eliminating anomalous observations or observations subject to greater prediction error, the resulting price index is nonetheless more stable.

11 The robust error calculation is also used in the preceding step, identifying the anomalous observations.
in Annex 1 show the estimates for offices and industrial buildings, respectively). Specifically, higher income levels are associated with higher prices, while floor area and price per square metre are negatively correlated. Prices are likely to be higher for prime establishments or those located in busy shopping or tourist areas. The transaction type (first sale or resale) and the transaction parties (natural or legal persons) appear to play a lesser role in determining prices.  

Equation [1] is estimated for each quarter using data for the current quarter and a few preceding quarters. Specifically, for the retail premises and industrial buildings indices, reference period \( t \) includes sales in the current quarter and the immediately preceding quarter, whereas the offices index includes transactions in the current quarter and the four preceding quarters. A rolling window spanning several quarters is used in order to provide estimates for the parameters of vector \( \beta \) that are more stable over time; this has the advantage of mitigating the volatility of the price index (see Annex 2). Further, this approach yields more observations, which helps to reduce the uncertainty surrounding the estimates.

### 3.3 Aggregating the data in a single price index

Once the cell prices are estimated based on equation [1], the data are aggregated in a chained Laspeyres index. To this end, price indices are calculated for each cell (elementary price indices):

\[
I_{t,y}^{e} = \frac{P_{t,y}^{e}}{P_{4,y-1}^{e}} \times 100
\]

where \( I_{t,y}^{e} \) is the index for cell \( e \) in quarter \( t \) of year \( y \), \( P_{t,y}^{e} \) is the estimated price for quarter \( t \) of year \( y \), and \( P_{4,y-1}^{e} \) is the price estimated for cell \( e \) in the fourth quarter of the preceding year. The reference period for the elementary indices is 2014.

The information from the elementary indices is summarised in a single index by allocating weights \( (W_{t}^{e}) \) to each cell:

\[
W_{t}^{e} = \frac{\text{Supcons}_{e}^{1} \times P_{4,y-1}^{e}}{\sum_{e} \text{Supcons}_{e}^{1} \times P_{4,y-1}^{e}}
\]

where \( \text{Supcons}_{e}^{1} \) the average built floor area in square metres of cell \( e \) (these weights are determined using 24-quarter rolling windows) and the remaining values are known.

---

12 The regression model underlying the price indices in this section is estimated period to period, whereas in Chart 2 and the charts in Annex 1 the regressions draw on information from the entire sample. Therefore, the charts show average effects associated with each of the property attributes across the entire analysis period, rather than the effects at each point in time.

13 In the cell price calculation, the estimator developed by El-Shaarawi and Viveros (1997) is evaluated. This estimator corrects the bias stemming from the estimations of the proposed log-linear model.

14 The weighting criterion is similar to that proposed by the INE (2015) to estimate the IPV. Other options that might introduce biases are ruled out, such as calculating the weights using only the floor area of the properties, since larger establishments have a far lower price per square metre (see Chart 2 for the negative correlation between price per square metre and floor area). \( W_{t}^{e} \) is determined using the price estimated for each cell \( e \) in the fourth quarter of the preceding year rather than the cell prices in each quarter, which would result in somewhat more volatile weights over time.
The aggregate index for quarter $t$ is obtained as follows:

$$I^t_A = \sum W^t_e \times I^t_e$$  \[4\]

that is, in each segment the elementary price index of the cell is multiplied by its relative importance (or weight).

Once the indices are calculated for each sub-segment, the overall price index for the CRE sector is obtained by weighting each of the previous indicators according to the weight of each property type (retail premises, industrial buildings and offices) in the market as a whole. The weighting criteria is the number of transactions.
4 Price indices for the CRE market

4.1 Overall index

Chart 3.1 shows developments in the overall price index for the CRE sector in Spain. The index is estimated from 2014 Q1, since earlier information is not available for some of the variables used to formulate the index.

Prices for this type of establishment declined between 2015 Q2 and 2016 Q4, before holding relatively stable until the onset of the pandemic. The pandemic (beginning in 2020 Q1) has seen additional price adjustments, although recently there have been signs of stabilisation (in 2021 Q2).

Chart 3
COMMERCIAL REAL ESTATE MARKET PRICE INDICES

SOURCES: Association of Registrars. Latest observation: 2021 Q2. The data for this quarter are provisional.
4.2 Indices for each sub-segment

Charts 3.2 to 3.4 show price indices for the market’s main sub-segments.

Prices for retail premises have closely tracked the overall index (see Chart 3.2), since this segment accounts for the bulk of CRE transactions (the retail premises index makes up just over 75% of the overall index). Specifically, there was a sharp price correction during the pandemic, possibly driven by factors such as the steady rise of e-commerce, which may have undermined demand for such establishments. The containment measures during the crisis may also have contributed to these developments (see International Monetary Fund, 2021).

The crisis has had a more moderate impact on the price of industrial buildings and offices (see Chart 3.2). Indeed, prices in these segments have stabilised or risen somewhat in recent quarters, meaning they may be less sensitive to the containment measures and other adverse developments stemming from the crisis. Some industrial buildings, such as logistics hubs and distribution centres, may even benefit from the recent increasing prominence of online shopping channels.

The pandemic has also been a turning point for properties in the prime segment (see Chart 3.3, comparing the overall price index with an index of properties located in prime areas). Although these establishments had risen notably in value prior to the crisis, prices fell sharply during the early stages of the pandemic. As with offices and industrial buildings, these properties have shown recent signs of price recovery.

Chart 3.4 summarises the above information and calculates the cumulative price change for each real estate category in two periods: (i) 2015 Q1 to 2019 Q4, prior to the outbreak of the coronavirus crisis, and (ii) from 2020 Q1 onwards, capturing the impact of the crisis. As has been noted, retail premises registered the sharpest price correction during the crisis, although this segment had already been underperforming the others before the pandemic. Prime retail premises also saw prices move negatively during the crisis, in contrast with the pre-pandemic trend.
5 Assessment of the proposed overall price index against other alternatives

One of the aims in formulating these indices is to obtain price measures that show a clear trend. To this end, the price indicator cannot behave erratically (stability). At the same time, it is imperative that the indicator provides a true picture of prices (representativeness). That is, the estimated property price must approximate to the real price. Likewise, the index must reflect the price situation at the current juncture.

To assess the above characteristics (stability and representativeness), Chart 4 compares the overall index for the commercial sector against several alternatives. First, the index is compared with another index capturing a simple average property price in each period (therefore not including the harmonising effect provided by the hedonic indicator). The proposed indicator is also contrasted against alternatives that estimate prices using different time windows. As indicated previously, the selected indicator – in Chart 4, the series represented by the legend “Proposed index (two-quarter window)” – is calculated using windows of data spanning two quarters: the current quarter and the immediately preceding quarter. The alternative indices, meanwhile, cover periods ranging from one quarter (the standard for indices such as the IPV) to five periods for the index with the longest time lag.

The proposed indicator strikes an appropriate balance between stability and representativeness. Compared with the alternative indices with longer lag times (i.e. those using a higher number of quarters to estimate property prices), the reference index is

---

**Chart 4**

PRICE INDICES. PROPOSED INDEX AND ALTERNATIVES

PROPOSED PRICE INDEX (TWO-QUARTER WINDOW) AND ALTERNATIVES

![Graph showing price indices comparison](image)

**SOURCE:** Association of Registrars. Latest observation: 2021 Q2.

---

15 Although more quarters of information are used to estimate prices in the offices segment, the impact on the overall index is very minor owing to this market’s low weight.
more contemporaneous (by construction) and does not prove to be much more unstable. The unlagged index (the “one-quarter window” in the chart, which estimates the price of commercial properties using a single quarter) captures more contemporaneous information. However, it is far more volatile than the proposed indicator while the turning points do not appear to differ significantly. Lastly, the average indicator is obtained by calculating the simple average price for these properties and eliminating the outliers. Since this indicator is not based on a hedonic regression model, it would be less representative. Further, it is more unstable than the above options.

Annex 3 sets out a more formal analysis of the properties of the price index and of the alternatives considered, in addition to assessing the goodness of fit of the models underlying each of the indices in the chart. These confirm the descriptive evidence covered in this section.

---

16 In Annex 3, dispersion measures are calculated for these indicators using the index return series. Specifically, the proposed index has a far smaller standard deviation (0.019) than the unlagged index (0.035).
6 Conclusions

The crisis prompted by the COVID-19 pandemic has highlighted the need for building the right tools to analyse the CRE market. This paper furthers that end by proposing a new price index for the sector in Spain.

There are several difficulties in constructing such an index, first and foremost the marked heterogeneity of commercial properties. To control for the changing composition of the establishments bought or sold in each period, this paper turns to methodologies that are established best practices in the estimation of price indices in the housing market. Specifically, the overall price index and the sub-indices described here take as baseline the methodology used by the INE to formulate the IPV, although the idiosyncrasies of the CRE sector are taken into account in the analysis. First, the market is divided into cells or strata of commercial properties with identical attributes. Regressions are then estimated for each quarter to find the relationship between these attributes and property prices, thus allowing prices to be imputed to the cells established in the previous step. The resulting hedonic index or indices serve to track the prices of homogeneous properties, which represents an improvement on other types of statistics such as average or median prices.

The price index shows that CRE market prices have declined in recent years. This downward trend was accentuated following the onset of the pandemic. By segment, price performances have been relatively heterogeneous in the most recent period. Specifically, during the health crisis the price correction has been more pronounced for retail premises and prime establishments than for other properties. It is not clear whether this owes to a temporary restructuring of supply and demand in the real estate market or a more persistent adjustment.

Generally speaking, the price indicators presented in this paper complement the tools available for analysing this market in Spain (Fernández et al., 2021). However, the indices are merely descriptive and do not indicate how high or low commercial property prices are in relation to their theoretical or equilibrium price. Therefore, in the future it might be useful to supplement this work with analysis of these properties’ fundamental equilibrium value. This could be done by capitalising on the methodologies that have been developed to assess the degree of over- or undervaluation in the housing market (see, for example, Martínez and Maza, 2003, or Hiebert and Sydow, 2011).
References


European Systemic Risk Board (2015), Report on commercial real estate and financial stability in the EU.

European Systemic Risk Board (2016), Recommendation of the European Systemic Risk Board of 31 October 2016 on closing real estate data gaps.

European Systemic Risk Board (2018), Report on vulnerabilities in the EU commercial real estate sector.


Annex 1  Regression exercises for the offices and industrial buildings segments

This annex sets out the coefficients estimated for the price models for offices (see Chart A1.1) and for industrial buildings (see Chart A1.2). Compared with retail premises - the most common type of property in the commercial real estate (CRE) sector -, fewer strata or cells are included in the analysis so as to reduce the number of empty cells (no price data) in each period. Accordingly, a somewhat lower number of coefficients is estimated than for retail premises (for further details, compare the results of this section with Chart 2).

Generally speaking, the effects identified do not differ greatly across the segments (for instance, properties are more expensive in higher income districts regardless of whether they are retail premises or offices). However, the magnitude of the coefficients varies somewhat in each sub-sector, which bears out the decision to estimate prices for each market segment separately.

Estimated coefficients for the office price model

Chart A1.1

<table>
<thead>
<tr>
<th>Buyer</th>
<th>Seller</th>
<th>Resale</th>
<th>m² (group 2)</th>
<th>Prime</th>
<th>Prov. group 2</th>
<th>Prov. group 3</th>
<th>Prov. group 4</th>
<th>High CRE</th>
<th>High hotel</th>
<th>Income (group 2)</th>
<th>Income (group 3)</th>
<th>Tourist mun.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SOURCE: Banco de España calculations.

a  The chart shows the estimated coefficients (triangles) and the associated confidence intervals (vertical lines, with a confidence level of 95%) of the regression described by equation [1]. The sample used captures all sales of offices since 2014. The model also incorporates some interactions between the explanatory variables whose coefficients are not shown in the chart in the interests of brevity. "Buyer" indicates that the buyer is a legal person; "Seller", that the seller is a legal person; "Resale" indicates that the property has changed hands in the past (approximates to the concept of second-hand property); "m² (group x)" refers to the floor area bracket (the higher groups/categories have more m² of floor area); "Prime" indicates that the property is located in a prime area; "Prov. group x" refers to the province group where the property is located (four groups in total); "High CRE" and "High hotel" indicate whether the number of commercial or hotel properties in the same municipality as the property is above median; "Income (group x)" refers to the bracket of average gross household income in the same postcode as the property (the higher the income, the higher the value of x in the chart); lastly, "Tourist mun." shows the coefficient associated with the property being located in a "tourist municipality" according to the INE. A positive value indicates that the characteristic in question is conducive to a higher price per square metre, while a negative value indicates the opposite. For example, in the chart the estimated coefficient for the "Tourist mun." characteristic is close to 0.20, indicating that property prices in those municipalities are, all other things being equal, 20% higher (approximately) than those elsewhere.
Chart A1.2

ESTIMATED COEFFICIENTS FOR THE INDUSTRIAL BUILDINGS PRICE MODEL (a)

SOURCE: Banco de España calculations.

The chart shows the estimated coefficients (triangles) and the associated confidence intervals (vertical lines, with a confidence level of 95%) of the regression described by equation [1]. The sample used captures all sales of industrial buildings since 2014. The model also incorporates some interactions between the explanatory variables whose coefficients are not shown in the chart in the interests of brevity. “Buyer” indicates that the buyer is a legal person; “Seller”, that the seller is a legal person; “Resale” indicates that the property has changed hands in the past (approximates to the concept of second-hand property); “m2 (group x)” refers to the floor area bracket (the higher groups/categories have more m2 of floor area); “Prime” indicates that the property is located in a prime area; “Prov. group x” refers to the province group where the property is located (four groups in total); “High CRE” and “High hotel” indicate whether the number of commercial or hotel properties in the same municipality as the property is above median; “Income (group x)” refers to the bracket of average gross household income in the same postcode as the property (the higher the income, the higher the value of x in the chart); lastly, “Tourist mun.” shows the coefficient associated with the property being located in a “tourist municipality” according to the INE. A positive value indicates that the characteristic in question is conducive to a higher price per square metre, while a negative value indicates the opposite. For example, in the chart the estimated coefficient for the “Tourist mun.” characteristic is close to 0.20, indicating that property prices in those municipalities are, all other things being equal, 20% higher (approximately) than those elsewhere.
Annex 2   Use of various quarters in the hedonic regression models

As indicated in Section 3.2, the hedonic regression models used to estimate cell prices include transaction data from more than one quarter (two quarters for retail premises and industrial buildings, five for offices). Thus, the estimates for the parameters of vector \( \beta \) are more stable across periods, yielding less volatile price indices.

This outcome may owe to two factors. First, the significant heterogeneity of CRE, which cannot be fully captured by the model variables. Second, possible inter-period compositional changes in the sample of properties. Indeed, if the properties sold in one quarter are of higher (lower) quality and this is not fully captured by the indicators used, the regression coefficients could be higher (lower) without actually being associated with an increase (decrease) in the implicit price of the properties’ attributes, which is what the hedonic regression models seek to measure.

The problem of quality change in the sample seems particularly acute in the commercial sector. Chart A2.1 shows the distribution of a variable that captures the change in the weight of the cells between the 2020 Q3 and Q4 in the retail premises

---

**Chart A2.1**

QUARTER-ON-QUARTER CHANGE IN THE WEIGHT OF THE CELLS (2020 Q4) IN DIFFERENT SEGMENTS (a)

 SOURCE: Banco de España calculations, based on the Association of Registrars.

a The chart shows the distribution of two variables that capture the change in the weight of each stratum or cell of the commercial real estate market (histogram) and the housing market (density function) between 2020 Q3 and Q4. Both measures are divided by the average weight of the cells in each market to make them comparable. The density function is approximated using a kernel estimator, which enables a non-parametric estimation of the density function, providing a continuous and smooth graphical representation thereof.
segment (similar results are found for other segments and periods). A similar variable has been calculated for the housing market, whose distribution is likewise set out in the chart.¹

The distribution is far more concentrated around zero in the variable calculated for the housing market, while that for the commercial market shows greater dispersion. This bears out the hypothesis that the sample of commercial properties changes significantly between quarters, thereby justifying the use of several quarters in the estimations.

---

¹ Since there is no model for estimating cell prices in the housing market, the weight of these is obtained by dividing the number of transactions in each cell by the total, rather than calculating the product of the built floor area in the cell and the cell price, as proposed in Section 3.3. The same methodology is used for the retail premises market. The variable of change in the weight of each cell (distribution shown in the chart) is likewise divided by the average cell weight (1/N, where N is the total number of cells obtained for each market). This makes the index more comparable across the two markets, since there are more cells in the housing market than in the commercial market.
Annex 3 Assessment of the price index against other alternatives

As indicated above, a good price index must display a relatively stable performance and be representative at the same time. This annex sets out a more formal analysis of these characteristics, both for the proposed index and for the alternatives detailed in Section 5.

Regarding the first of the criteria (stability), various statistics have been calculated based on the price index return (IR), as per the Eurostat recommendations (2013), where the return is the rate of price index change between quarter \( t \) and quarter \( t-1 \) (\( \text{IR}_t = \frac{\text{Index}_t}{\text{Index}_{t-1}} - 1 \)). These statistics are the standard deviation of \( \text{IR}_t \) and the correlation coefficient between \( \text{IR}_t \) and \( \text{IR}_{t-1} \), calculated using information for the entire sample period. A smaller standard deviation and a positive autocorrelation are desirable characteristics in price indicators (greater stability and lower quarter-on-quarter volatility).

Further, the representativeness of the indices is assessed by calculating for each one a goodness-of-fit measure for the estimations: the root-mean-square error (RMSE). A lower RMSE indicates a smaller deviation between the observed price and that predicted by the model; accordingly, indices with lower RSMEs will be preferred. In the proposed index, the estimations factor in not only the current period but also previous periods. Therefore, consideration should be given to the extent to which the relationship between property prices and their attributes is contemporaneous. Hence, a price index that uses only information from the current period is preferable to those with a longer time lag.

Table A3.1 sets out the above statistics for the proposed index and the estimation period. The same information is shown for the alternative indices described in Section 5.\(^1\) As noted in that section, the proposed index is rather more unstable than the alternatives with

<table>
<thead>
<tr>
<th>Index</th>
<th>Standard deviation</th>
<th>Autocorrelation</th>
<th>RMSE</th>
<th>Time window</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed index</td>
<td>0.019</td>
<td>0.018</td>
<td>22.349</td>
<td>t-1</td>
</tr>
<tr>
<td>Index with additional lag</td>
<td>0.013</td>
<td>0.364</td>
<td>22.411</td>
<td>t-2</td>
</tr>
<tr>
<td>Index with two additional lags</td>
<td>0.012</td>
<td>0.443</td>
<td>22.443</td>
<td>t-3</td>
</tr>
<tr>
<td>Index with three additional lags</td>
<td>0.011</td>
<td>0.525</td>
<td>22.462</td>
<td>t-4</td>
</tr>
<tr>
<td>Unlagged index</td>
<td>0.035</td>
<td>-0.424</td>
<td>22.166</td>
<td>t</td>
</tr>
<tr>
<td>Average index</td>
<td>0.043</td>
<td>-0.273</td>
<td>—</td>
<td>t</td>
</tr>
</tbody>
</table>

SOURCE: Banco de España calculations.

\(^a\) The table shows various statistics for the reference index (“proposed index”) and several alternatives. The first two statistics are the standard deviation and the autocorrelation coefficient of the index return. This return is calculated as follows: \( \text{IR}_t = \frac{\text{Index}_t}{\text{Index}_{t-1}} - 1 \), where IR is the index return, “Index” refers to the CRE index and t is the quarter. The RMSE is the root-mean-square error obtained for each estimation (not calculated for the average indicator, which is not based on a regression exercise). Since the index is constructed for each segment, the RMSE of each is weighted by the weight of the segment in the overall CRE market. Lastly, “Time window” indicates the number of quarters considered to estimate the index. The more quarters there are, the less contemporaneous the index.

\(^1\) The outliers have been eliminated from the calculation of the average index.
a longer time lag. However, it improves on the alternatives because it has a somewhat lower RMSE and is more contemporaneous (by construction). The unlagged indicator (“unlagged index”), as the most contemporaneous (by construction) and that with the best goodness of fit (lowest RMSE; more representative), is quite unstable. Thus, its IR captures a high standard deviation and a negative correlation coefficient. The latter implies that periods of price decline tend to be accompanied by periods of price growth (and vice versa), meaning the resulting signals would be misleading. The average index would be more unstable than the previous options according to the statistics constructed. Its RMSE is not calculated since it is not based on a model like the previous indices.
BANCO DE ESPAÑA PUBLICATIONS

OCCASIONAL PAPERS

2030 ÁNGEL GÓMEZ-CARRIEÑO GARCÍA-MORENO: Juan Sebastián Elcano – 500 años de la Primera vuelta al mundo en los billetes del Banco de España. Historia y tecnología del billete.

2031 OLYMPIA BOVER, NATALIA FABRA, SANDRA GARCÍA-URIBE, AITOR LACUESTA and ROBERTO RAMOS: Firms and households during the pandemic: what do we learn from their electricity consumption?

2032 JULIA BRUNET, LUCÍA CUADRO-SÁEZ and JAVIER J. PÉREZ: Contingency public funds for emergencies: the lessons from the international experience. (There is a Spanish version of this edition with the same number).


2034 PEDRO DEL RÍO LÓPEZ: A BVAR toolkit to assess macrofinancial risks in Brazil and Mexico.

2035 FRUCTUOSO BORRALLO: Una taxonomía de actividades sostenibles para Europa.

2036 ROBERTO RAMOS: The uneven impact of the health crisis on the euro area economies in 2020. (There is a Spanish version of this edition with the same number).

2037 ERNESTO JAVIER J. PÉREZ: The relationship between pandemic containment measures, mobility and economic activity. (There is a Spanish version of this edition with the same number).

2038 CORINNA GHIRELLI, MARIÁ GIL, SAMUEL HURTADO and ALBERTO URTASUN: The relationship between pandemic containment measures, mobility and economic activity. (There is a Spanish version of this edition with the same number).

2039 DMITRY KHAMETSHIN: High-yield bond markets during the COVID-19 crisis: the role of monetary policy. (There is a Spanish version of this edition with the same number).

2040 IRMA ALONSO and LUIS MOLINA: A GPS navigator to monitor risks in emerging economies: the vulnerability dashboard.

2041 JOSÉ MANUEL CARBÓ and ESTHER DIEZ GARCÍA: El interés por la innovación financiera en España. Un análisis de sus causas y su relevancia en el contexto de la crisis COVID-19.

2042 CRISTINA BARCELÓ, MARIO IZQUIERDO, AITOR LACUESTA, SERGIO PUENTE, ANA REGIL and ENRIQUE MORAL-BENITO: Real-time analysis of the revisions to the structural position of public finances.

2043 LANZUÑEJA: Los efectos del salario mínimo interprofesional en el empleo: nueva evidencia para España.

2044 ENRIQUE MORAL-BENITO: Fiscal policy measures adopted since the second wave of the health crisis: the euro area, the United States and the United Kingdom. (There is a Spanish version of this edition with the same number).


2046 ROBERTO BLANCO, SERGIO MAYORDOMO, ÁLVARO MENÉNDEZ and MARISTELA MULINO: Impacts of the COVID-19 crisis on Spanish firms’ financial vulnerability. (There is a Spanish version of this edition with the same number).

2047 JULIO DE QUINTO: The Spanish pension system: an update in the wake of the pandemic. Banco de España. Testimony before the Parliamentary Budget Committee, 2 September 2020. (There is a Spanish version of this edition with the same number).

2048 ANA GÓMEZ-LOSCOS: Dating and synchronisation of regional business cycles in Spain. (There is a Spanish version of this edition with the same number).

2049 PABLO BURRIEL, VÍCTOR GONZÁLEZ-DÍEZ, JORGE MARTÍNEZ-PAGÉS and ENRIQUE MORAL-BENITO: Incentive compatible relationship between the ERM II and close cooperation in the Banking Union: the cases of Bulgaria and Croatia.

2050 PEDRO DEL RÍO LÓPEZ: A GPS navigator to monitor risks in emerging economies: the vulnerability dashboard.

2051 MATÍAS PACCIE, ISABEL SÁNCHEZ and MARTA SUÁREZ-VAIRELA: Recent developments in Spanish retail electricity prices: the role played by the cost of CO2 emission allowances and higher gas prices. (There is a Spanish version of this edition with the same number).
MARIO ALLOZA, JAVIER ANDRÉS, PABLO BURRIEL, IVÁN KATARYNIUK, JAVIER J. PÉREZ and JUAN LUIS VEGA: The reform of the European Union's fiscal governance framework in a new macroeconomic environment. (There is a Spanish version of this edition with the same number).

MARIO ALLOZA, VÍCTOR GONZÁLEZ-DÍEZ, ENRIQUE MORAL-BENITO and PATROCINIO TELLO-CASAS: Access to services in rural Spain. (There is a Spanish version of this edition with the same number).

CARLOS GONZÁLEZ PEDRAZ and ADRIAN VAN RIXTEL: The role of derivatives in market strains during the COVID-19 crisis. (There is a Spanish version of this edition with the same number).

IVÁN KATARYNIUK, JAVIER PÉREZ and FRANCESCA VIANI: (De-)Globalisation of trade and regionalisation: a survey of the facts and arguments.


CLARA I. GONZÁLEZ and SOLEDAD NÚÑEZ: Markets, financial institutions and central banks in the face of climate change: challenges and opportunities.

ISABEL GARRIDO: The International Monetary Fund’s view of social equity throughout its 75 years of existence. (There is a Spanish version of this edition with the same number).

JORGE ESCOLAR and JOSÉ RAMÓN YRIBARREN: European Central Bank and Banco de España measures against the effects of COVID-19 on the monetary policy collateral framework, and their impact on Spanish counterparties. (There is a Spanish version of this edition with the same number).

BRINDUSA ANGHEL, AITOR LACUESTA and FEDERICO TAGLIATI: 2021 Survey of Small Enterprises’ Financial Literacy: Main Results. (There is a Spanish version of this edition with the same number).

PABLO HERNÁNDEZ DE COS: Testimony before the Congress of Deputies Budget Committee on 25 October 2021 and before the Senate Budget Committee on 30 November 2021 in relation to the Draft State Budget for 2022. (There is a Spanish version of this edition with the same number).

LAURA AURIA, MARKUS BINGMER, CARLOS MATEO CAICEDO GRACIANO, CLÉMENCE CHARAVEL, SERGIO GAVILÁ, ALESSANDRA IANNAMORELLI, AVIRAM LEVY, ALFREDO MALDONADO, FLORIAN RESCH, ANNA MARIA ROSSI and STEPHAN SAUER: Overview of central banks’ in-house credit assessment systems in the euro area.

JORGE E. GALÁN: CREWS: a CAMELS-based early warning system of systemic risk in the banking sector.

ALEJANDRO FERNÁNDEZ CEREZO and JOSÉ MANUEL MONTERO: A sectoral analysis of the future challenges facing the spanish economy. (There is a Spanish version of this edition with the same number).

MANUEL A. PÉREZ ÁLVAREZ: New allocation of special drawing rights. (There is a Spanish version of this edition with the same number).

PILUCA ALVARGONZÁLEZ, MARINA GÓMEZ, CARMEN MARTÍNEZ-CARRASCAL, MYROSLAV RIDKUYKO and ERNESTO VILLANUEVA: Analysis of labor flows and consumption in Spain during COVID-19.

MATIÁS LAMAS and SARA ROMANIEGA: Designing a price index for the Spanish commercial real estate market.