



Business Administration, with International Mention (E4)

# **TOOLS TO DETECT OVERHEATING IN THE SPANISH RESIDENTIAL HOUSING MARKET: A REGIONAL ANALYSIS**

## **TRABAJO FIN DE GRADO O TRABAJO FIN DE MÁSTER**

Clave: 202120171

MADRID | Junio 2023

## Table of Contents

<i>Abstract</i> .....	3
<i>Introduction</i> .....	4
<i>Methodology</i> .....	7
Kolmogorov-Smirnov Two Sample Test .....	7
Excess of Kurtosis and Skewness .....	9
Extreme Values.....	10
Complementary Analysis.....	10
Limitations .....	10
<i>All Town Analysis</i> .....	11
Extreme Values Right Tail .....	15
Q-Q Plot Analysis .....	16
<i>Regional Analysis</i> .....	20
Madrid.....	23
Cataluña .....	28
<i>Other regional Analyses</i> .....	31
Regional Analysis Bell Curves.....	32
Murcia: .....	34
Islas Baleares: .....	35
Canarias: .....	36
Extremadura.....	38
Castilla-La Mancha.....	39
Aragón .....	40
Cantabria .....	41
Galicia.....	42
La Rioja.....	42
Valencia .....	43
Castilla Y León .....	45
Andalucía: .....	45
<i>Conclusion</i> .....	48
<i>Bibliography</i> .....	51
<i>Annex</i> .....	53

## Abstract

In this paper I search for asset bubble behaviour in the Spanish housing market. Housing assets represent an imperfect asset class whereby supply lags in response to price signals due to the asset's inherent characteristics. For such assets, I present evidence to suggest asset collateral channels and credit expansion accelerator effects cause bubble-like tendencies in the Spanish housing market.

I use static house price and income data to gauge the most accurate characterisation of demand forces in the Spanish housing market using various statistical tools and explanatory variables. The Kolmogorov-Smirnov two sample tests on regional markets are accompanied by skewness, kurtosis and extreme value theory statistics to define the relationship between income and house prices across overvalued, stable and undervalued regional markets. I find that kurtosis is a fundamental statistical tool to judge the highest price and income levels within a market and investigate the presence of asset collateral channels.

Each regional market is examined individually to appreciate the non-standard, idiosyncratic behaviour exhibited. I show that income is an important indicator of house prices. I distinguish between fundamental value and non-fundamental value drivers of price to justify house asset values.

## Introduction

Housing represents shelter, a long-term investment, and the largest asset on the household balance sheet (Causa & Woloszko, 2019). Great political importance is placed on housing due to the significant strain it can present to individuals and the grave aggregated consequences of market downturns on an economy. The Spanish housing market is a standout example of how expansionary periods can provoke strong employment and GDP growth and how downturns can produce opposite effects (Alvarez-Roman & García-Posada, 2021). Fundamentally, houses are prone to asset collateral channels whereby the asset price can deviate from its fundamental value (Goldstein & Razin, 2013). Incentives and speculative assumptions among the stakeholders can lead to extreme deviations between price and value that become unsustainable. This problem is commonly referred to as the bubble effect. The inherent characteristics of the market require intense investigation into the drivers of asset price in order to protect vulnerable stakeholders and the prosperity of the economy.

The relationship between house prices and incomes requires external financing. Much like any business project, external lenders are required to present the entrepreneur with sufficiently large upfront costs to incentivise good project choices and the avoidance of moral hazard issues (Goldstein & Razin, 2013). The entrepreneur must have a large stake in the investment or must be able to secure the loan with collateral. Spanish aspiring homeowners require 30% of the value of the house in order to comply with the mortgage's covenants and afford associated costs (La Vanguardia, 2022). This large stake ensures the search for mutually beneficial opportunities for banks and entrepreneurs. Good projects will increase the value of the asset and reduce Loan-to-Value ratio, to in turn positively affect the balance sheet of both homeowner and financial lender. The homeowner can opportunistically refinance their mortgage to obtain better interest rates; and the mortgage lender gains more confidence in its ability to cover non-performing loans and increases its borrowing capacity from other financial institutions. Both stakeholders are incentivised to invest in the most rewarding projects possible.

The incentive to invest in such asset leads to entrepreneurs investing at their maximum investment potential. The key implication is that “it would be easier to provide external financing to entrepreneurs with large assets A, since they are more likely to internalize the monetary benefit and choose the good project rather than enjoying the non-pecuniary private

benefits of the bad project” (Goldstein & Razin, 2013). Hereby, the terms to buy, build and lend are more favourable on larger assets. Larger investors have inherent advantages to receive the best returns. The marriage of entrepreneurial incentives and credit expansion causes asset collateral effects (Goldstein & Razin, 2013).

Highly financially rewarding house assets cluster where there is a high abundance of fundamental and non-fundamental factors. The existing literature popularly recognises fundamental value as a function of income, employment, population density, effort rate and housing stock (Alvarez-Roman & García-Posada, 2021). Over- or undervaluation is considered as the difference between asset price and the fundamental value. For this study the ‘difference’ is defined as non-fundamental value; it considers, but is not limited to, demographic and cultural choices, geography and landscape, size, fiscal incentives, and judicial differences. The varying magnitude of these factors lead to variations in attractiveness across the national housing market. In imperfect markets, whereby supply is slow to respond to price signals, the asset collateral channel behaviour is exhibited (Allen & Galle, 2000).

Given the supported concepts, I hypothesise the impact of asset collateral channels on market conditions.

The asset collateral concept worsens housing affordability. The search for good projects and the credit expansion phenomenon create a great deviation between house prices and the key fundamental value indicator, income.

Certain metrics justify the excessive collateral effect in attractive markets and the limited collateral effect in unattractive markets. These metrics are: favourable local interest rate on mortgages, low mortgage value to house price ratio, high transaction volume<sup>1</sup>, high foreign investment and low proportion of non-principal residencies<sup>2</sup>. In attractive markets, these variables suggest good projects that improve balance sheets, provoke credit expansions and induce inflationary effects on the asset value.

---

<sup>1</sup> Entrepreneurs want to benefit from the high returns

<sup>2</sup> This provides an indication of housing stock available to investors.

Excessive collateral effect in attractive markets can spill over into other segments in the market whereby houses have similar attributes and marginally lower prices. Increases in price in highest price segment will place upward pressure on similar assets leading to average asset values increases. Spill overs can occur infinitely given the asset attributes (non-fundamental factors) are appropriately similar. This increases average national house prices and proves the asset collateral model for this asset class.

The comparison of local market house price distribution and local market income distribution is the most important consideration to judge house market overvaluation. Similarity in right fat tail of both house price and income distributions shows proportional price to income ratios. It shows non-excessive collateral channel effects in the local market, thus exhibiting stable market behaviour.

The incentive to find good projects will cause exceptional house price outliers. The collateral channel will cause house price extreme values to be far greater than income outliers.

An excessively large price distribution fat tail suggests overheating in the local market. The right tail suggests an abundance of attractive attributes in the local market. An excessively large income distribution fat tail suggests underheating in the local market. The credit expansion is limited by the lack of highly rewards projects whereby the spill overs are limited.

Highest average price markets will present the greatest difference between house price distribution and income distribution. The local markets that present the best projects will produce the most dissimilar relationship between house prices and incomes.

Asset collateral channel effects in house assets provokes the bubble effect in housing markets.

The purpose of this study is to find evidence of the asset collateral theory present in the Spanish housing market. To explore the bubble behaviour, these hypotheses will be investigated using a variety of statistical tools. The results will lend themselves to better characterise the local regional markets (known in Spain as ‘autonomous communities’) and appreciate their idiosyncratic features.

## Methodology

The variables analysed to investigate the possibility of growing bubbles in the Spanish housing market are average town house price per square metre and average town gross income. The house price data comes from a dynamic Fotocasa database that is updated on a monthly basis (Fotocasa, 2023). The sample taken from this rich database considers house prices from February 2023 as the data was collected from 2<sup>nd</sup> March 2023 between 11am and 12pm. The average gross income data comes from the most recent Agencia Tributaria publication of 2020 (Agencia Tributaria, 2020). This dataset limits the research to 48 provinces as the Navarra, Álava, Guipúzcoa and Vizcaya are not included in the report. This dataset allows distributional analysis of both house prices and incomes across regional (autonomous community) markets and the national market. Three main statistical tools will be used to explore the similarity of distributions between average house price per square metre and average gross income. [AD HOC?]

### Kolmogorov-Smirnov Two Sample Test

For this study, the Kolmogorov-Smirnov [KS] two sample test serves as a mechanism to determine whether average prices and average incomes follow adequately similar probability density functions. The town specific variables exhibit continuous behaviour when grouped into regions. The two-sample non-parametric statistical test is optimal for this study as it does not intend to define a given probability distribution as Gaussian, uniform, exponential, etc, unlike its one sample counterpart (Young, 1977). The statistical tool recognises the idiosyncrasies that housing assets in regional markets possess given the abundance of non-fundamental value factors.

The null hypothesis,  $H_0$ , of the Kolmogorov-Smirnov two sample test proposes the house price per metre squared sample and the gross income sample come from the same continuous probability distribution function for any given test: whereby,  $f_1(x)$  is average house price per metre squared and  $f_2(x)$  is average gross income.

$$H_0: f_1(x) = f_2(x) \quad -\infty \leq x \leq +\infty$$

The alternative hypothesis,  $H_1$ , of the Kolmogorov-Smirnov two sample test is the rejection of the null hypothesis and proposes the house price per metre squared sample and the gross income sample exhibit different continuous probability distribution functions for any given test. The alternative hypothesis is defined as follows:

$$H_1: f_1(x) \neq f_2(x) \quad -\infty \leq x \leq +\infty$$

The sample probability distribution functions of the two variables, house price per metre squared and income are given by:

$$F_{n_1}(x_r) = \sum_{x=x_0}^{x_r} f_{n_1}(x)$$

$$F_{n_2}(x_r) = \sum_{x=x_0}^{x_r} f_{n_2}(x)$$

The D statistic reduces the relationship between the two probability density functions to a derived, real scalar value (Young, 1977). Graphically, this measures the greatest vertical distance between the two probability density functions. The published critical value is a function the confidence interval and the number of samples for both average house price per square metre and average gross income ( $N_1, N_2$ ). The null hypothesis is accepted when the calculated D statistic is smaller or equal to the published critical D value, given the confidence interval and number of samples. The chosen confidence interval for the study is 5%. The null hypothesis is accepted for p-values greater than 0.05 and is rejected for p-values less than 0.05. The p-values of regional market KS two sample tests will be considered to characterise the relationship between average house prices and average gross incomes within the specific market. The results will render conclusions concerning the closeness of the variable distributions and therefore, the presence of asset collateral model effects within the market.

$$D_{n,m} = \sup |F_{n_1,n}(x) - F_{n_2,n}(x)|$$

A strength of the Kolmogorov-Smirnov two sample test is the lack of importance placed on sample sizes of the two variables for any given test. Lacking data from one of the two datasets, whereby ( $N_1 \neq N_2$ ), has minimal effect on the results of the test due to the published critical values. Moreover, the literature states that  $N > 20$  is sufficiently large for the tests to produce



reliable results as the central limit theory is applicable (TDW, 2023). Consequently, Cueta and Melilla provincial markets are not considered.

The inherent mechanics of the KS two sample test allow for further analysis at given price and income subsegments in the market. By construction, as the sample sizes tend to infinity, both sample distribution curves will perfectly align,  $x = x_0$  and  $x = \infty$  (Young, 1977). Although this appears a limitation to the analysis, further KS two sample tests can investigate smaller samples at  $a < X < b$ , whereby  $a, b$  are chosen price/income normalised values. Conclusively, the results of these closer KS two sample tests will demonstrate over- or underabundance of data at given price/income points and suggest fat or thin tails. Fat house price tails will suggest excessive asset collateral channels at the high price points within the regional market.

The normalisation models used by the technological tools used for both variables are given as follows:

$$F_{n_1}(x_0) = F_{n_2}(x_0) = 0$$

$$F_{n_1}(\infty) = F_{n_2}(\infty) = 1$$

IBM SPSS

$$Z = \frac{x - \mu}{\sigma}$$

JAMOVI

### Excess of Kurtosis and Skewness

As the Kolmogorov-Smirnov test is limited to a judgement on the similarity of probability distributions, kurtosis and skewness tests compliment the results to characterise the regional markets. Karl Pearson's (Fiori & Zenga, 2009) kurtosis definition is the fourth moment of distribution. Kurtosis aids analysis of the distributions at the fat tails rather than central moments. High excess of kurtosis suggests collateral channel effects with bubble-like tendencies at the high price segments of the market. Equally skewness statistics indicate over or underabundant housing supply at given price segments. Overabundance at certain price levels indicates over- or underheating regional markets.

Hair et al. (2010) and Bryne (2010) argue that data is considered to be normal if skewness is between -2 to +2 and kurtosis is between -7 to +7 (Hair, Black, & Babin, 2010). The kurtosis

and skewness will be compared with the literature definitions to characterise the market distributions. Both tools supplement the results from the Kolmogorov-Smirnov two sample tests to depict the market conditions in regional markets.

### Extreme Values

To compliment the previous statistical tools, extreme values are analysed. Extreme values justify the given kurtosis results to suggest over and underheated regional markets. Extreme right tail house price values represent the most rewarding project where the assets have the most attractive attributes and credit expansion is most abundant.

### Complementary Analysis

Ad hoc regional metrics are analysed to reinforce the results from the statistical tests. The chosen metrics clarify the abundance of non-fundamental value factors and suggest the prevalence of credit expansion in given markets. The ad hoc metrics chosen are regional interest rate on mortgages, regional mortgage value to house price ratio, transaction volume, foreign investment and proportion of non-principal residencies. Markets with favourable metrics demonstrate the greatest evidence of credit expansion and the significant potential of an asset bubble.

### Limitations

A limitation of the Fotocasa data which must be considered is that when the volume of housing transactions is considered too low for the algorithm that is used to calculate the average price/m<sup>2</sup>, it applies the generic average price/m<sup>2</sup> figure for the greater region. This is significant and beneficial to not mislead the results with significant outliers; however for the purposes of this study, this has been normalised to limit its overbearing strain on the distribution curve.

The Agencia Tributaria data is limited to towns with populations greater than 1,000 inhabitants. Given the assumption that towns with less than 1,000 inhabitants have low transaction volumes, the disocclusion of these towns may inflate the results of the statistical tests within the regional and national tests.

A final consideration when considering the validity of results is the difference in sampling techniques used between the tests and the published statistics for ERI.

## All Town Analysis

This section provides an empirical characterization of the unconditional distribution of average gross income and house prices across all towns above 1,000 residents. Given the N values of both variables are sufficiently large, the results of the Kolmogorov-Smirnov two sample test are highly valid. With great confidence the null hypothesis is rejected that average house prices and average gross incomes come from the same probability density function. Therefore, the alternative hypothesis is accepted that the distributions exhibit sufficiently different distributions. The sample sizes for All Towns test are as follows:  $N_{INC} = 843$  and  $N_{HP} = 892$ . This test gives an initial indication that house prices are determined by many more fundamental and non-fundamental factors, not just income.

The additional statistical tools are essential to characterise the distributions and evaluate whether there is an over- or underabundance of data at given price/income points. The house price variable exhibits greater variability when compared to incomes to ferment the idea that there are many fundamental and non-fundamental factors that affect house prices. In combination with skewness and kurtosis values, extreme value analysis depicts where the greatest evidence of asset collateral channel effects lies in the market.

The distributions of both house price and income require further investigation after characterisation with the aforementioned statistical tools. As the KS two sample test displays the stark difference between the distributions, further analysis into the top end segments reveal the fundamental cause of the statistical differences in distributions. Given price exhibits more positively distribution, the overabundance of house price data one standard deviation above the variable mean is the root cause of the stark difference in distributions and the significantly low p-value. This demonstrates the first concrete evidence of asset collateral channels in the housing market.

### Summary of Hypothesis Test

	Null Hypothesis	Sig. <sup>a,b</sup>	Decision
1	The distribution of House Price and Income across ALLTOWNS.	<.001	Reject the null hypothesis with a 0.05 significance level

Table 1

N total		1764
Maximum Extreme Differences	Absolute	,119
	Positive	,119
	Negative	-,059
Test Statistic		2,496
Significance Level (2-sample statistic)		<,001

Table 2

	Average Gross Income	Average Price
Mean (True Value)	23,958	1,531
Standard Deviation (True Value)	5,052	861
Coefficient of Variation	21.1%	56.2%

Table 3

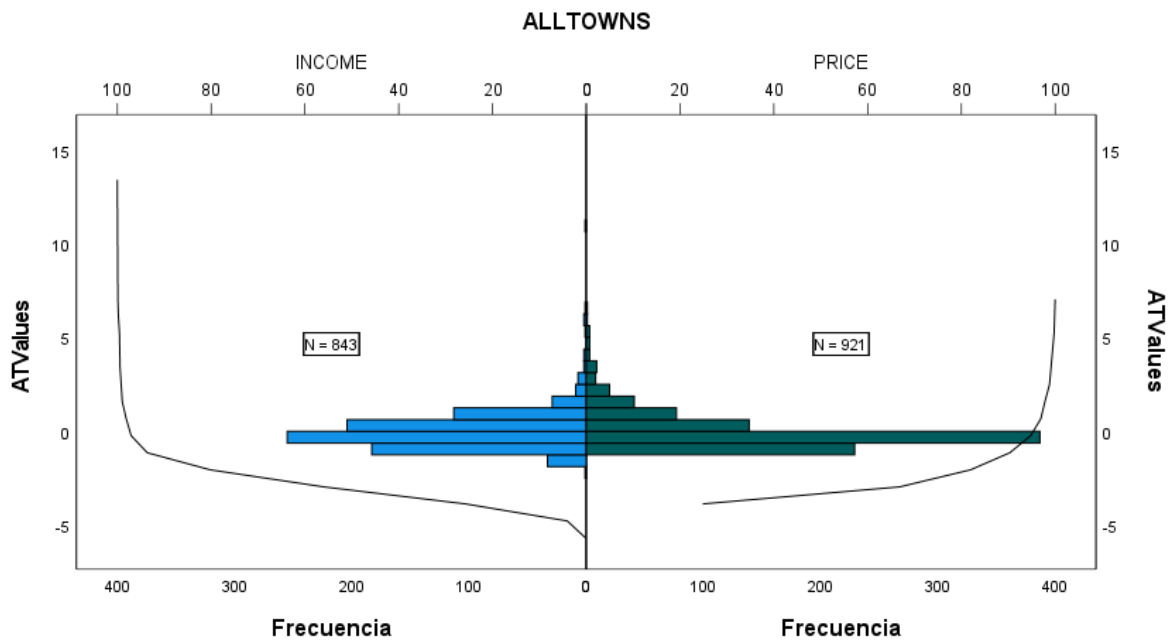


Figure 1

From the naked eye the income distribution appears more normally distributed than the price distribution, yet the statistical measures used suggest the opposite. Figure 1 shows in both cases the data is positively skewed where the median is far smaller than the mean. For income this is a common characteristic across all economies whereby there are fewer households earning considerably above the national mean. For price, the positive skew can be explained

with this same logic as higher incomes induce the ability to buy more expensive housing. Furthermore it presents evidence of collateral channel effects at the highest price segment.

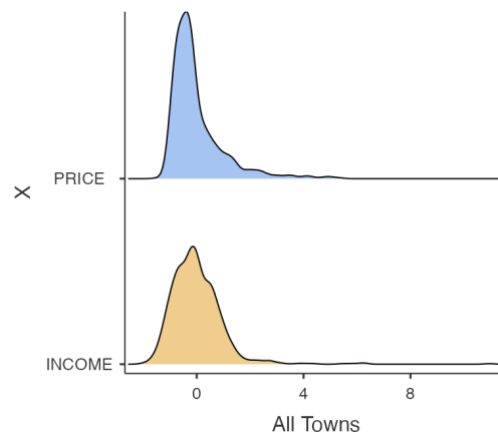


Figure 2

Although neither variable pertain truly normal distribution characteristics as defined by the existing literature, in both cases the variables exhibit behaviour close to Gaussian distribution. Figure 2 demonstrates the greater tendency house price distribution around the central location. Hair et al. and Bryne argue that data is considered to be normal if skewness is between -2 to +2 and kurtosis is between -7 to +7 (Hair, Black, & Babin, 2010). Table 4 shows the statistics for both variables using the whole dataset. Neither distribution is considered normally distributed according to the literature, although price has much closer behaviour to normal distribution than income. In the Income sample the datapoint with the largest Z score is removed from the skewness and the kurtosis statistics, displayed in Table 4, due to its overbearing effect on the tests. The datapoint is Pozuelo de Alarcón and is still considered in the rest of the tests and sections. In both cases, income distribution is much more heavily skewed and has a greater excess of kurtosis. As kurtosis, by Pearson’s method (Fiori & Zenga, 2009), is considered as the fourth moment of distribution, the following results provide initial insights to the heavily tailedness of both distributions, and alludes to income being heavier tailed.

	PRICE	INCOME
<b>SKEWNESS</b>	2.36	1.91
<b>KURTOSIS</b>	7.06	8.75

Table 4

The Skewness statistical test confirms the ruling in Figure 2. House price across all towns present a skewness statistic of 2.36 whereas income skewness is 1.91. The most positively skewed coefficient of price suggests greater variability. This is supported by the far greater coefficient of variation as well as the Q-Q plots exhibited by price when compared to income.

SD	Price	Income
X < 0	69%	56%
X > 0	31%	44%
0 < X < 1	18%	34%
1 < X < 2	7%	7%
2 < X < 3	3%	2%
X > 3	2%	1%

Table 5 Proportion of Data at various Standard Deviations

To support the findings Table 5 shows the proportional distribution of data among different Z scores. Table 5 demonstrates house price's more positive skew due to the greater abundance of data at  $SD < 0$ . Income datapoints are more abundant around the Z scores  $0 < X < 1$ . Table 5 provides sufficient evidence of a fat tail when considering the data for All Towns, as demonstrated by the proportion of data at  $SD > 1$ . The house price variable consistently produces a greater proportion of data one standard deviation above the mean even given the greater skewness coefficient. This presents evidence of asset collateral channels whereby greater asset and credit expansion is displayed at high price segments. Moreover, this gives an initial indication of the inherent financing incentive that is given to entrepreneurs to take on bigger projects as explored by (Allen & Galle, 2000). These results are also supported by McMillen's research, "Rather, bigger homes in higher-priced neighbourhood homes simply appreciated more rapidly than other homes between 1995 and 2005; i.e., the coefficients of the hedonic price functions have changed in such a way to increase the return on high-priced homes." (McMillen, 2007).

Furthermore, price distribution is more focused around its central location with an interquartile range of  $IQR_P = 0.892$  whilst  $IQR_I = 1.12$ . The proportion of data between  $1 < SD < 3$  to the right of the central location suggests variations between the two samples around 1 and 3 moments. 80.2% of the income data lies within  $-1 < X < 1$ , whereas 81.2% of the data lies

within this same range for price. 97.2% of the income data lies between  $-2 < X < 2$ , whereas 96.0% lies within this same range from price. Nevertheless, the absolute range of income exceeds that of price significantly – Income Absolute Range: 11.0 and Price Absolute Range: 5.37. All this considered, an extreme value analysis will evaluate the key differences in the distributions at the right tail.

### Extreme Values Right Tail

Table 6 shows the Z-scores of the 8 most extreme cases for price, income, and affordability ratio. Affordability Ratio represents the relationship of price to income; this concept should not be confused with overvaluation whereby the market price exceeds the intrinsic value of the asset. This section is dedicated to analysing the towns or suburbs of Spain that have the highest house prices and incomes.

The table shows that the extreme values across all three categories are overabundant in four main Autonomous Communities: The Community of Madrid, Cataluña and Islas Baleares. There are only two exceptions to this which both lie in the Affordability Ratio category: Adeje, Canarias and Tarifa, Andalucía.

The extreme value Price Z-scores are much closer together than the Income Z-scores. Considering only these top eight values, the price Z-score range is 2.13 whereas the income Z-score range is 7.16. This is evidently caused by the significant Z-score of Pozuelo de Alarcón which has an average gross income of €79,327. This notable outlier is responsible for the considerably larger excess of kurtosis score and absolute range exhibited by income over price. However, other than the outlier, the values of both income and price tend to align at the most extreme right tail of the distributions. Furthermore, the income column is entirely dominated by Cataluña and Madrid, showing a concentration of wealth in these Autonomous Communities. The high income statistics in these regions is a direct cause of the absence of towns from these two regions in the affordability ratio column.

	Price			Income			Affordability Ratio		
	AC <sup>3</sup>	Town	Z	AC	Town	Z	AC	Town	Z
1	<i>Madrid</i>	Chamberí	<b>5.67</b>	<i>Madrid</i>	Pozuelo de Alarcón	<b>10.9</b>	<i>Illes Balears</i>	Santa Eulalia	<b>9.08</b>
2	<i>Illes Balears</i>	Santa Eulalia	<b>5.41</b>	<i>Madrid</i>	Alcobendas	<b>6.28</b>	<i>Illes Balears</i>	Eivissa	<b>5.39</b>
3	<i>Cataluña</i>	Sarrià – Sant Gervasi	<b>5.30</b>	<i>Cataluña</i>	Sant Cugat	<b>6.19</b>	<i>Illes Balears</i>	Andratx	<b>5.39</b>
4	<i>Illes Balears</i>	Eivissa	<b>5.16</b>	<i>Madrid</i>	Majadahonda	<b>5.84</b>	<i>Illes Balears</i>	Calvià	<b>4.38</b>
5	<i>Madrid</i>	Chamartín	<b>4.93</b>	<i>Madrid</i>	Las Rozas	<b>5.47</b>	<i>Illes Balears</i>	Sant Josep	<b>4.15</b>
6	<i>Cataluña</i>	Eixample	<b>4.48</b>	<i>Cataluña</i>	Sitges	<b>4.36</b>	<i>Canarias</i>	Adeje	<b>4.15</b>
7	<i>Illes Balears</i>	Andratx	<b>4.49</b>	<i>Madrid</i>	Villaviciosa de Odón	<b>4.03</b>	<i>Andalucía</i>	Tarifa	<b>4.02</b>
8	<i>Madrid</i>	Retiro	<b>4.46</b>	<i>Cataluña</i>	Castelldefels	<b>3.74</b>	<i>Illes Balears</i>	Campos	<b>3.97</b>

TABLE 6

### Q-Q Plot Analysis

To compliment the previous analysis, Q-Q plots aid in characterising the behaviour of the distributions at various segments. The steep curvature exhibited by the house price Q-Q plot curve demonstrates how house prices are more strongly skewed than incomes (a contradiction to the skewness test in Table 4). Around  $-1 < SD < 1.5$ , the income curve correlates very closely to the theoretical quantiles whereas the price curve is exponential. This suggests that the house prices exhibit less normal distribution than incomes, more variance and the heavier tails.

Furthermore, the gradient of the line is much steeper for price than it is for incomes. This proposes the idea that house prices exhibit greater variability than income and confirms the notion that house prices are not merely driven by the fundamental price driver of income. It suggests that there may be potential market imbalance across various price segments. The affordability ratio Q-Q plot visualises the almost perfect dissection affordability creates between the price and income curves. This is relevant as it considers the ratio between price

---

<sup>3</sup> Autonomous Community



and incomes of specific locations. It suggests that income is still a very strong indicator of house prices.

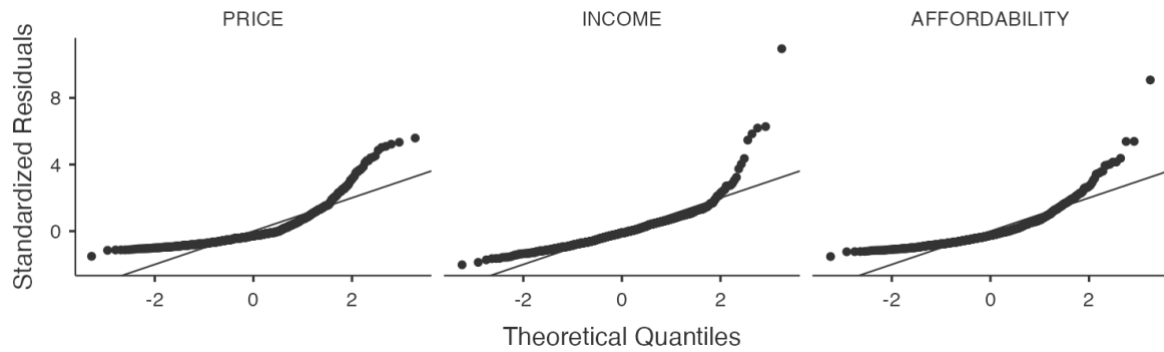


Figure 3

Given all the analysis above, an investigation into the overabundance of house price within the region  $1 < SD$  with respect to income distribution is necessary to further analyse the situation. Although the extreme values appear similar, the differences in fat tails demonstrated by the Q-Q plots suggest an imbalance. Building on this hypothesis I use two Kolmogorov-Smirnov 2 sample tests to investigate the difference between the two distributions between two observationally chosen ranges,  $1 < SD$  in order to test this theory.

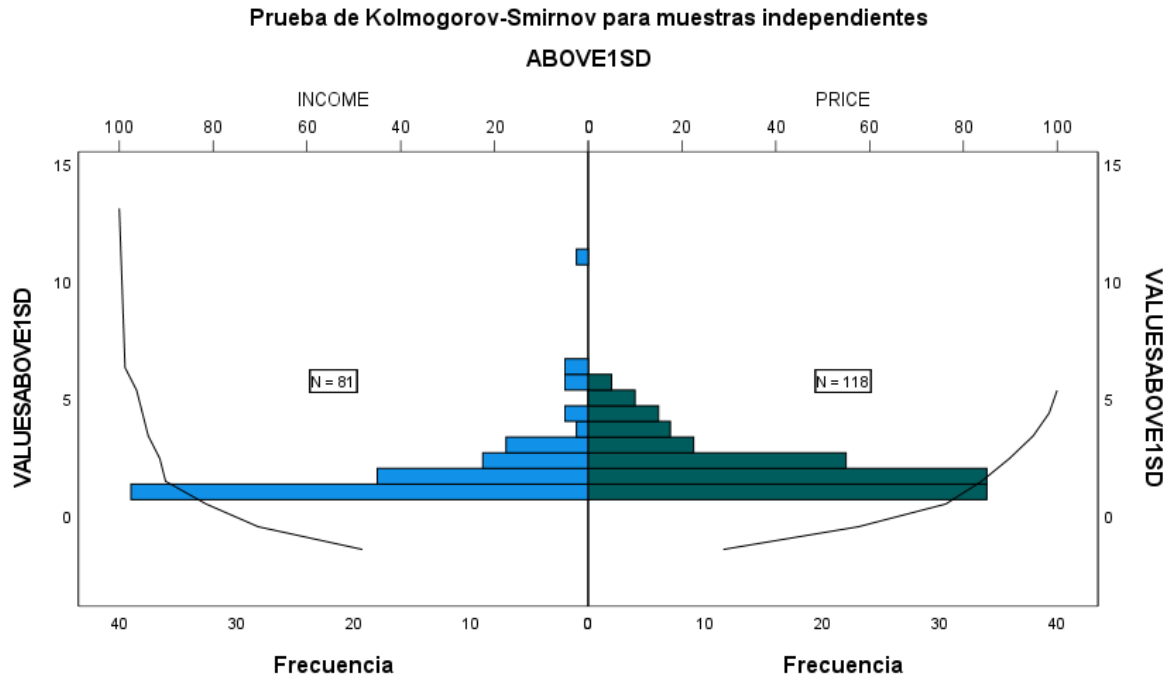


Figure 4

Figure 4 shows the significantly larger abundance of datapoints one standard deviation above the mean for the house price variable in comparison to the income variable. The null hypothesis is again rejected for this test. This provides sufficient evidence of asset collateral effects within the market. This market behaviour is driven by the market price dynamics of the best projects where the average house price per square metre is maximised.

For house price and income variables across the whole of Spain, the null hypothesis is rejected; this demonstrates that these variables do not come from the same distribution. It implies that the price of housing varies across different markets and price segments. Price demonstrates greater variability, and the tests confirm that there are more factors than just income that effect house price valuations. At the highest price and income levels, the KS tests suggest that incomes and house prices align whereby we accept the null hypothesis that the samples come from the same probability density function. The greatest differences between the two distributions occurs where the normalised samples range between the standard distributions  $1 < SD$ , where the null hypothesis is again rejected.

Both distributions exhibit heavy tails but price has a significantly larger proportion of the samples in  $SD > 1$  than income. There is initial evidence to suggest that in Cataluña, Madrid

and Islas Baleares have strong asset collateral effects due to the high extreme values within these markets. A local analysis will take place to understand whether income distribution justifies the price distribution or not. This now requires investigation into specific markets to understand whether these price and income imbalances are justified in their local markets or not (from a distributional perspective).

## Regional Analysis

Regional housing markets are assumed heterogenous across the whole of Spain due to their inherent fundamental and non-fundamental attributes. Kolmogorov-Smirnov two sample test is a fundamental statistical tool for this analysis as it appreciates the idiosyncrasies of market forces in individual regional markets. Overvalued towns in All Town analysis can be completely justified when considered the specific regional or provincial market. These tests benefit from using regional averages as opposed to national averages from comparison, hence will be more closely related and produce closer Z-scores.

The regions considered are: Andalucía, Aragón, Asturias, Canarias, Cantabria, Castilla y León, Castilla-La Manche, Cataluña, Extremadura, Galicia, Illes Balears, La Rioja, Madrid, Murcia and Valencia. Madrid and Cataluña are analysed first as they provide a benchmark analysis for comparison of the smaller regions. The regional markets vary greatly in geographic size, average price and income, Kurtosis and Skewness, mortgage value, proportion of non-principal residencies, market turnover, and attractiveness to foreign investment. These metrics aid in the characterisation of the relationship between incomes and house prices to, therefore, reach conclusions on the prevalence of credit expansion and asset collateral effects within the regional markets.

This section comments on the over- or “underheatedness” of regional markets to suggest the formation of asset bubbles. The conclusions are based upon the asset collateral theory concept and the provocation of credit expansions. Significant suggestions of said theories imply accelerator effects which cause over expansion in both asset and credit markets. The markets are judged as undervalued, stable, or overvalued. This section stresses the importance of using statistical tools and several metrics in order to gage the most complete picture of demand forces within the housing market. Although the average national house price statistic gives a suggestion of market conditions, it does not give any indication of bubble tendencies. This paper suggests that bubbles can form in local markets given that average national house prices are still 6.2% below peak (INE, 2023). The entrepreneur’s incentive to undertaken good projects that increase in value can inflate stakeholder perception of valuation. Asset collateral channels can cluster in certain segment of the market and lead to overvaluations when national peak is still low.

This has two major consequences: positive shocks in the higher segments generate positive sentiment which can lead to upward pressure on houses with similar attributes but lower prices, as both segments benefit from higher collateral. The second consequence is the contagion effect in neighbouring markets. Positive shocks at the highest segment may signal to markets in close towns or provinces with similar attributes upward price pressure for the market as a whole.

The overvaluation report by (Alvarez-Roman & García-Posada, 2021) is used to compliment and scrutinise the analysis. In some cases, fundamental analysis solidifies the findings; in other cases, the weakness of multiple linear regression analysis to identify is highlighted.

KS tests are used to characterise the distributions of income and price, check for fat tailedness and test the study's hypotheses. This section groups the KS tests by separation into observationally chosen p-value ranges:  $0.05 < p \leq 0.40$ ,  $0.40 < p \leq 0.60$ ,  $0.60 < p \leq 0.80$ , and  $0.80 < p$ .

### Summary of Findings

The regional markets produced p-values from the KS two sample tests varied from  $p = 0.118$  to  $p = 0.996$ . The similarity of house price and income distributions varied immensely across the complex markets.

Excess of kurtosis statistic is a significant determinant of p-value. Specifically, the p-value of regional tests tended to be much lower when price excess of kurtosis greatly exceeded income excess of kurtosis. Highly positive kurtosis statistics suggest the presence of fat tails in the market that create spill overs across various price segments. Negative kurtosis statistics suggest a lack of fat tails did not provide evidence for asset collateral channel effects. Skewness appeared a less relevant factor in the regional analysis. In tests whereby price kurtosis is severely negative, markets are unattractive, and asset and credit contraction are suggested. These theories were confirmed by regional interest rates and mortgage value to house price ratios.

The most attractive markets had high house price extreme values and fat tails. When the statistical tests were accompanied by the complementary metrics, attractive markets presented strong evidence for asset collateral channels and credit expansion. Attractive markets contain house assets that are considered as good projects that will increase in value. Attractive markets have abundant fundamental and non-fundamental value drivers. Markets deemed as unattractive and/or undervalued consistently present low excess of kurtosis.

There is a strong connection in the regional analysis between foreign investment and asset collateral effects. This is fermented by the difference in average transaction value by foreign entrepreneurs and domestic entrepreneurs. Average transaction value by foreign entrepreneurs is 12% higher than that of domestic entrepreneurs (Ministerio de Transportes, Movilidad y Agenda Urbana, 2023). Higher average transaction price causes greater fat tails in certain market. The tests provide sufficient evidence that attractive markets have fat tails whereby the highest price town markets create signals to entrepreneurs regarding markets with similar attributes and ultimately cause spill over effects.

In the majority of tests, both house price and income distributions complied with the literature definition of kurtosis and skewness ( $7 < X < 7$  excess of kurtosis test and  $-2 < X < 2$  skewness test) (Hair, Black, & Babin, 2010). This is a significant finding when compared to the all-town analysis whereby these statistics were much greater.

## Madrid

Madrid is the second largest regional market in Spain with 6.7m inhabitants across the region (INE, 2023). The population density and housing stock density vary greatly across the region with the highest concentration of inhabitants and houses within the city (Carpio-Pinedo, 2020). The exponentially decreasing housing unit density relative to closeness of the city demonstrates how the regional market can be considered as a relatively closed system and a lack of spillovers in neighbouring regional markets. As suggested by the literature (Alvarez-Roman & García-Posada, 2021), population density is a significant fundamental value factor and causes the relatively normal distribution of house prices within the market. The strong presence of fundamental factors suggests a small difference between asset price and fundamental value and hint towards the Madrid Region representing a relatively stable market.

	<b>HOUSE PRICE</b>	<b>GROSS INCOME</b>
<b>AVERAGE</b>	2,524 €/m <sup>2</sup>	€ 31,796
<b>STANDARD DEVIATION</b>	1,203 €/m <sup>2</sup>	€ 10,101
<b>COV<sup>4</sup></b>	0.48	0.32
<b>SKEWNESS</b>	1.42	2.06
<b>KURTOSIS</b>	2.84	5.67

Table 11

<b>CC.AA</b>	<b>GROWTH RATE IN 22</b>	<b>MORTGAGE VALUE /</b>	<b>INTEREST RATE</b>	<b>% NON-PRINCIPLE</b>	<b>NUMBER OF TRANSACTIONS</b>	<b>FOREIGN INVESTMENT</b>
Madrid	9.6%	69.2%	2.21%	30.3%	83,388	5.0%

Figure 12

The sample results state that the region of Madrid has the second highest house prices and the highest average incomes, as demonstrated by the table above. Average house prices for the region stand at €2,524/m<sup>2</sup> whereas average gross income stands at €31,796. These statistics are considerable when considering the size of the market in terms of inhabitants and quantity of housing stock (Carpio-Pinedo, 2020). The statistical tests suggest relative stability in the

---

<sup>4</sup> Coefficient of Variation

market in terms of similarity of income and house price distributions, presence of right-end fat tails, and extreme values.

The combination of statistical results and complimentary metrics strongly suggest asset collateral channel effects within the market but not explicit signs of overheating. Madrid housing projects represent the largest projects whereby the entrepreneur’s motivation to internalise benefit aligns both bank and entrepreneur incentives. The regional interest rate is attractive given the abundance of fundamental and non-fundamental drivers of value and mortgage value to house price ratio is significantly lower than market average. While there is evidence to suggest that affordability is low, the alignment of house price and income distributions, alongside extreme value analysis, suggest house prices are relatively stable.

### Summary of Hypothesis Test

	Null Hypothesis	Sig. <sup>a,b</sup>	Decision
1	The distribution House Price and Income across Madrid.	0.815	Accept the null hypothesis with a 0.05 significance level

Table 7

For the Madrid region KS two sample test, the null hypothesis is accepted with a significant p-value of 0.815. There is a high certainty that house prices and incomes come from the same probability density function. The Caixabank supply shortage report suggests that Madrid experiences a shortage of housing in the city centre as well as in the surrounding peripheral towns (Garriga, 2023). The KS two sample test opposes this sentiment as at given price levels there are sufficient house listings<sup>5</sup> within the region when compared to relative income levels. Figure 5 suggests greater normality of distribution when compared to All Towns analysis previously discussed whereby the distributions correlate greatly.

---

<sup>5</sup> Implied by Fotocasa average prices



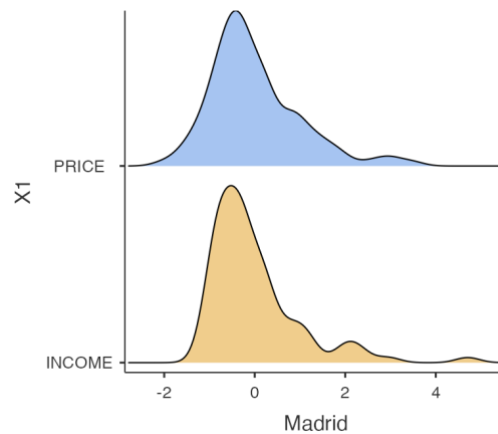


Figure 5

The Madrid statistical test results suggest very opposing market behaviour when compared to All Town analysis. Figure 5 demonstrates that the Madrid house price variable has a greater proportion of data to the right of the regional mean than the Madrid income variable. Madrid region house price variable shows a much greater abundance of data points above the mean when compared to (a) Madrid regional income and (b) national price sample. This alludes to an attractive market whereby good projects experience asset collateral channels effects at the middle-to-top price segments.

The skewness statistic hints towards above average inequitable distribution of income in the region of Madrid. Skewness of income exceeds that of price whereas the opposite is true for All Towns. As the Z score tends to -2, the frequency of income data drops drastically presenting floor-like tendencies. A floor effect in incomes is intuitively anticipated to be the product of government equality-seeking intervention to protect the vulnerable stakeholders within the aggressive regional market. Although the skewness and kurtosis statistics align with the literature definitions for normality, the heavily skewed income data when compared to house prices effects affordability at the bottom segments for income and price (Hair, Black, & Babin, 2010). This is confirmed by the significantly high coefficient of variation produced by the Madrid sample in comparison to the national sample (Madrid: 31.7%, Spain: 21.1%).

The difference in Z-scores of the extreme values for the Madrid test in comparison to the national test demonstrates the normality of the house price values given the specific market conditions. Three of the eight most extreme house price values in All Towns analysis were from the Madrid region: Chamberí, Chamartín, and Retiro. The Z-scores in the regional KS

two sample test are much closer to the mean given the fundamental and non-fundamental value drivers across the aggregated regional market. These local markets cause the regional market to have a right fat tail due to the understanding of the greatest projects due to the abundance of attractive factors.

	All Towns Z-score	Madrid Region Z-score
<i>Chamberí</i>	3.47	5.67
<i>Chamartín</i>	2.92	4.93
<i>Retiro</i>	2.55	4.46

Table 8

### Q-Q plot

The Madrid regional Q-Q plots support the idea that house price distribution is highly Gaussian and the top end market is influencing house prices in the rest of the regional market. Figure 6 shows how the standardised residuals of highest price segment ( $Z > 2$ ) tend further away from the theoretical quantiles than the standardised residuals of income. This suggests a larger fat tail in house prices than in incomes in the Madrid Region market. These findings lend themselves to the central theory regarding the attractiveness of large projects for financing reasons discussed in the “3 crises” academic research (Goldstein & Razin, 2013). Given the normality of the rest of the market, this fat tail appears excessive. Given the largest projects for entrepreneurs present themselves in Madrid, favourable interest rates are offered due to the significantly large nominal upfront costs presented by the market. Credit expansion and asset accelerator effects are visible in this market.

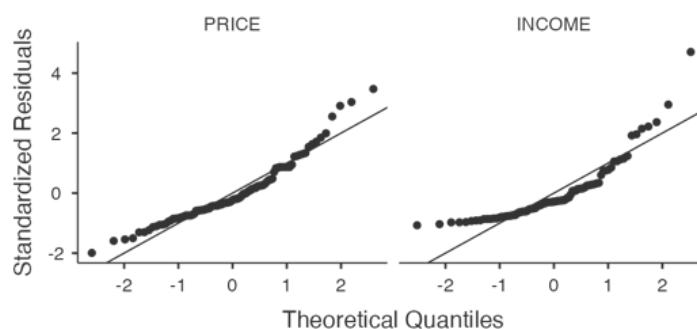


Figure 6

Madrid consistently demonstrates significant evidence of the asset collateral channel concept in housing. The house price fat tail is concentrated around the city due to the greatness of opportunity for entrepreneurs. Although there is a fat tail, the market adjusts to compensate for the high prices, and house prices across the region maintain highly Gaussian distribution. The ERI reports display the greater confidence in Madrid asset and credit markets through the consistently lower interest rates on mortgages, lower loan-to-value ratios and higher variable-to-fixed mortgage rates over the last decade (ERI, 2023). Entrepreneurial incentives alongside collateral channels cause highly inflationary effects in the market. Further analysis is required to investigate the lack of spill overs to neighbouring markets with fundamental value analyses considering population and housing stock densities.

## Cataluña

Cataluña is the largest regional market by population with 7.7m inhabitants (INE, 2023). Dissimilarly to the uni-provincial Madrid market, the Cataluña market comprises of four provinces with highly variable fundamental and non-fundamental factors of price. These markets are: Barcelona, Girona, Tarragona, Lleida. The regional market is divided by economic prosperity and demographic differences into the provincial markets. The Barcelona province presents 73% of the Catalan population and 74% of the economy (OECD, 2010). The relatively higher asset value of houses in Barcelona is driven by the abundance of fundamental factors present with respect to the surrounding provinces. Barcelona represents the top end provincial housing market whereby houses are much higher than surrounding provinces. The economic nature of Cataluña limits the significance of the statistical tests as the regional market comprises of smaller markets with different attributes. Upon consideration of the Catalan market in its entirety, the evidence suggests that asset collateral channels do not produce the same high level of credit expansion recognised in the Madrid market and there is considered more stable.

	<b>HOUSE PRICE</b>	<b>GROSS INCOME</b>
<b>AVERAGE</b>	2,236 €/m <sup>2</sup>	€ 28,140
<b>STANDARD DEVIATION</b>	1,243 €/m <sup>2</sup>	€ 5,248
<b>COV<sup>6</sup></b>	0.56	0.18
<b>SKEWNESS</b>	1.11	2.73
<b>KURTOSIS</b>	1.05	10.5

Table 9

<b>CC.AA</b>	<b>GROWTH RATE IN 22</b>	<b>MORTGAGE VALUE /</b>	<b>INTEREST RATE</b>	<b>% NON-PRINCIPLE</b>	<b>NUMBER OF TRANSACTIONS</b>	<b>FOREIGN INVESTMENT</b>
Cataluña	6.7%	73.3%	2.33%	40.0%	102,231	13.6%

Figure 7

The study sample results states that the Catalan region has the third highest house prices and the second highest average incomes of all of Spain. Average house prices for the region stand at €2,236/m<sup>2</sup> whereas average gross income standards at €28,140. The regional average

<sup>6</sup> Coefficient of Variance

variables are notably high given the greater variability of house prices and incomes across the provinces within in the Catalan economy.

### Summary of Hypothesis Test

	Null Hypothesis	Sig. <sup>a,b</sup>	Decision
1	The distribution of House Price and Income across Cataluña.	.126	Accept the null hypothesis with a 0.05 significance level

Table 10

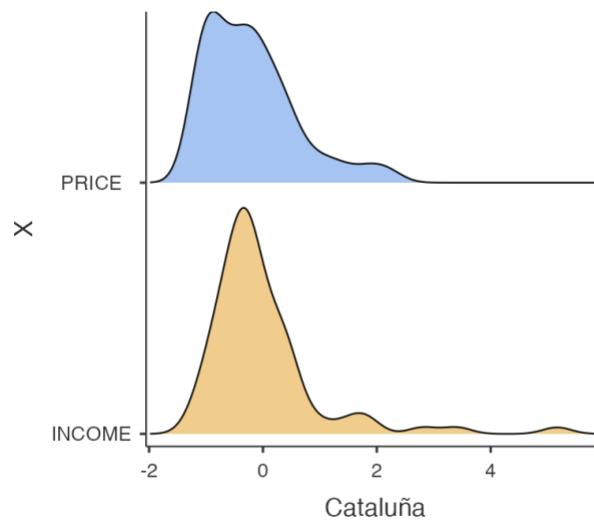


Figure 8

The null hypothesis for the Cataluña Kolmogorov-Smirnov two sample test is accepted and signifies that house prices and incomes exhibit the same probability density function with a 95% degree of confidence. Although the null hypothesis is accepted, the produced p-value is among the lowest scores of all the regional tests. This underscores the high variability of market conditions within the regional market and highlights the importance of individual market analysis.

The Catalan regional market shares many common statistical attributes to the Madrid regional market related to the central location and below. Where the regional markets differ most significantly is in the coefficients of variation and relationship between house prices and

incomes at the highest segments. Figure 8 demonstrates the uniformity that income distribution exhibits with a lacking fat tail. In the Catalan market, the price fat tail is not as significant as the income fat tail. There is less evidence of credit expansion due to the lacking fat tail in the market. The low abundance at the highest end does not allow for asset collateral channels exhibited in the Madrid market and limits growth at the highest end.

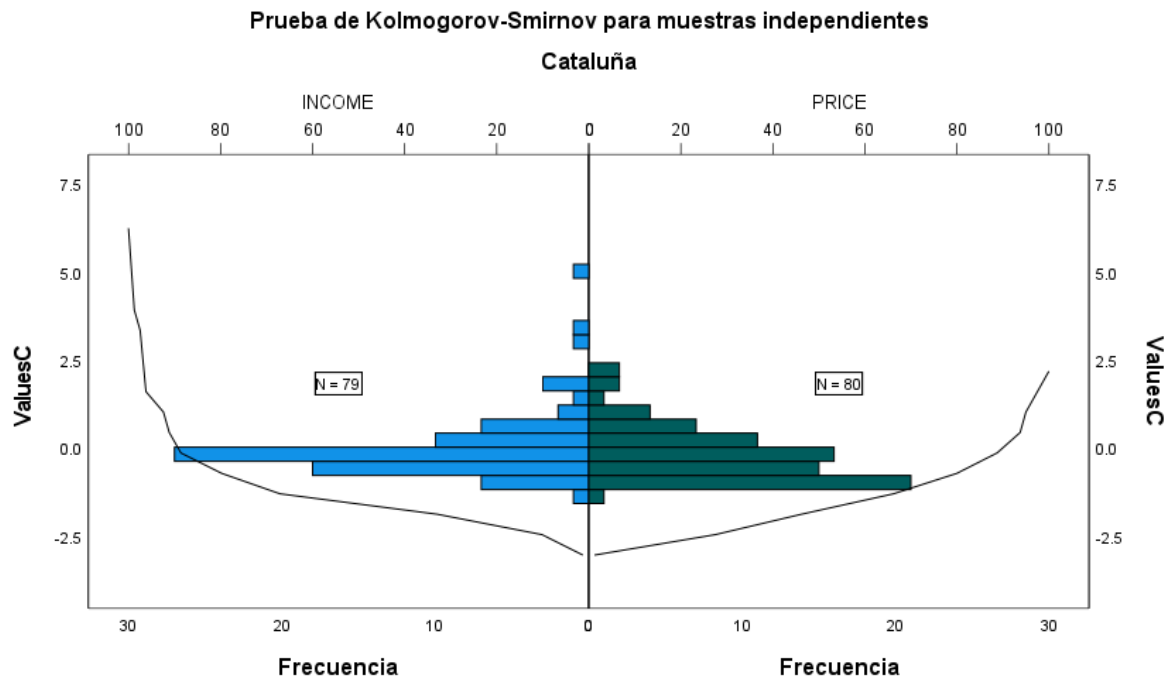


Figure 9

SD	Price	Income
$X < 0$	59%	66%
$X > 0$	41%	46%
$0 < X < 1$	26%	29%
$1 < X < 2$	9%	10%
$2 < X < 3$	6%	6%
$X > 3$	1%	1%

Table 11

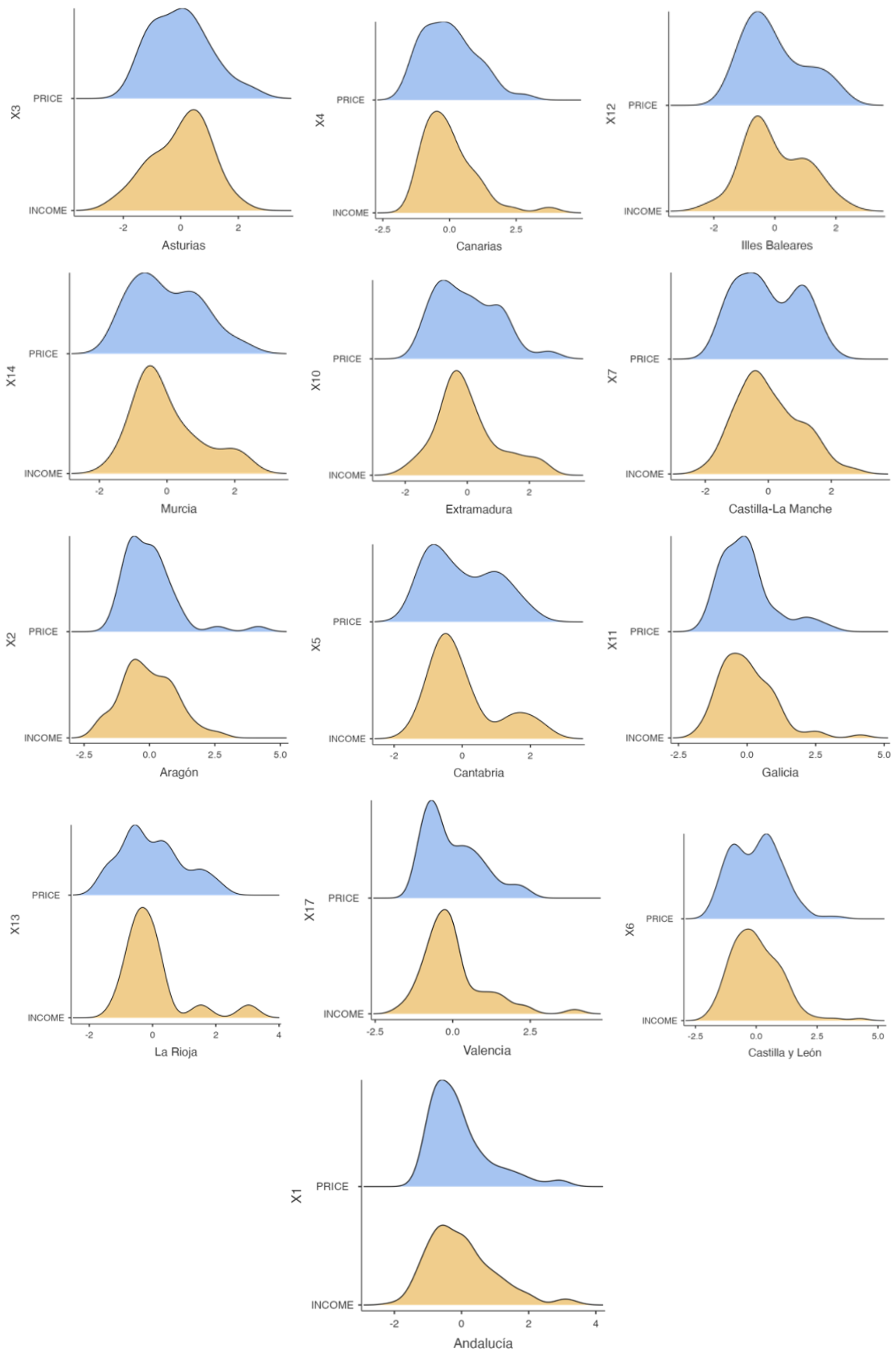
Ultimately, the Cataluña housing market exhibits very different market conditions to the Madrid market due to the difference in inherent characteristics, fundamental and non-fundamental drivers of price. Cataluña does not suggest the same degree of asset collateral channels and credit expansion due to the prevalence of spill overs across the provincial markets.

## Other regional Analyses

CC.AA	P-VALUE	MEAN AC		KURTOSIS		GROWTH RATE IN 22	MORTGAGE VALUE / PRICE	MEAN I RATE	% NON- PRINCIPAL	NUMBER OF TRANSACTIONS	FOREIGN INVESTMENT	X > 3 (AC PRICE)
		PRICE	INCOME	PRICE	INCOME							
a) Asturias, Principado de	<b>0.996</b>	1,379	25,410	-0.02	-0.21	1.5%	84.6%	2.37%	29.1%	11,394	3.6%	0
Murcia, Región de	<b>0.984</b>	1,050	22,537	-0.62	0.25	6.0%	75.2%	2.36%	10.6%	22,425	19.1%	0
Balears, Illes	<b>0.978</b>	3,439	25,497	-0.64	-0.38	12.3%	60.4%	2.36%	23.9%	17,826	34.4%	0
Canarias	<b>0.913</b>	1,897	22,631	0.08	3.92	12.7%	57.9%	2.46%	23.2%	28,173	28.6%	0
b) Extremadura	<b>0.836</b>	921	19,928	-0.06	0.72	9.5%	94.4%	2.43%	34.7%	1,034	1.2%	0
Castilla - La Mancha	<b>0.809</b>	1,043	22,409	-1.18	-0.27	1.9%	93.4%	2.35%	30.4%	267	4.3%	0
c) Aragón	<b>0.582</b>	1,378	25,049	6.44	-0.06	5.4%	77.4%	2.29%	24.1%	16,669	6.4%	1
Cantabria	<b>0.560</b>	1,628	25,616	-1.15	0.44	5.2%	71.4%	2.17%	16.3%	8,595	3.2%	0
Galicia	<b>0.527</b>	1,365	23,300	1.30	3.00	4.8%	75.4%	2.35%	32.9%	21,708	1.2%	0
Rioja, La	<b>0.512</b>	1,245	24,379	-0.47	5.52	6.7%	78.9%	2.24%	12.6%	4,966	6.9%	0
Comunitat Valenciana	<b>0.493</b>	1,707	23,885	-0.06	3.35	6.5%	69.8%	2.35%	18.7%	99,869	26.8%	0
d) Castilla y León	<b>0.359</b>	1,137	23,757	-0.08	1.93	6.6%	81.7%	2.22%	34.4%	28,014	2.4%	1
Andalucía	<b>0.118</b>	1,411	22,078	2.07	1.02	9.4%	69.1%	2.30%	24.8%	133,481	15.0%	2

Table 10

Figure 11: Regional Analysis Bell Curves





a) KS tests with  $0.8 < p\text{-score} < 1.0$

Regarding fundamental drivers of price these regions exhibit vastly different characteristics, they produce the greatest p-score demonstrating the closeness between distributions of price and income. Average price among these regions range from 1,050 €/m<sup>2</sup> to 3,439 €/m<sup>2</sup>, which alludes to the idea that average price isn't overly significant when analysed solely. Affordability ratio also varies greatly between these regions. An analysis of each region is necessary to depict whether the market is overheating relative to incomes in the given market.

Asturias

Asturias produces the highest p-value at  $p = 0.996$  and exhibits the behaviour most attributed to normal distribution. Both price and income, of which are normalised using Asturian means and standard deviations, produce relatively tight absolute ranges, of which  $AR_P = 3.67$  and  $AR_I = 3.95$ . The data does not produce significantly varied data for either variable, and shows an absence of fat tails. The standardised residuals correlate very closely to the theoretical quantiles exhibiting highly Gaussian behaviour.

Where Asturias differs from the rest of the tests is in Skewness. Asturias represents the only regional market whereby income skewness produces a negative value: Skewness of Price: 0.576; Skewness of Income: -0.436. The negatively skewed distribution (-0.03) suggests a higher proportion of data earns greater than mean income. This represents a limitation to the results principally caused by the equal weighting given to all towns irrespective of population size.

Kurtosis for both price and income distribution are low and negative demonstrating the lack of fat tails in the market. Kurtosis of price: -0.02; Kurtosis of income: -0.21. This signifies that the distribution is lighter at the tails than normal distribution. The stability of price and income across the distribution demonstrates a fixed correlation between house prices and incomes; this is also referred to as stable affordability ratio across the region. The KS test suggests that income is the main factor of price in Asturias whereby other fundamental and non-fundamental factors are negligible.

Although this analysis may suggest market stability, the complete absence of fat tails may suggest the market's lack of confidence in selecting good projects within Asturias. The similarity in behaviour between house prices and incomes may represent an anomaly in the results, whereby other local markets exhibit characteristics whereby price is more negatively skewed than income. The ERI report suggests similar findings; Asturias had the lowest price growth rate in 2022, low transaction volume and low foreign interest. The overvaluation study demonstrates concurrent results whereby the region was considered c.20% undervalued in 2018 according to the overvaluation model.

#### Murcia:

The Murcia KS two sample test demonstrates the high degree of similarity between the house price and income distribution, with a p-value of 0.984. Murcia differs from the other KS tests in  $0.8 < p < 1.0$  category due to its significantly lower prices (€1,050) and greater affordability ratio. Murcia's negative kurtosis statistic of -0.62 presents a platykurtic house price variable distribution whereby data tends towards the central location at there is a significant lack of fat tails. The lack of fat tails implies an absence of very good projects within the region. This scarcity of fat tails in Murcia is a unique characteristic that sets Murcia apart from the other coastal regions.

The complementary metrics suggest favourable market conditions that would hint towards asset collateral channels and credit expansion. The surrounding coastal markets with similar attributes have fat house price tails whereby entrepreneurs with lower investment capacity should be incentivised to invest in Murcia, given the favourable published metrics. Murcia housing market's high proportion of primary residency homes coupled with substantial market turnover rates are important non-fundamental drivers of house prices. The magnitude of the outlier, Manga del Mar Menor, demonstrates the relevance non-fundamental value factor "closeness to the sea" has on house prices. The similarity in non-fundamental value drivers would suggest upward pressure should be placed on Murcia due to its neighbouring markets, Andalusian and Valencian. This has not been found.

The similarity of house price and income distributions, negative kurtosis of price and low average price suggest the Murcian housing market is undervalued with a lack of asset collateral

channels leading to a bubble effect. The projects are not considered sufficiently adequate to decrease loan-to-value ratios and therefore the credit market remains stagnant.

#### Islas Baleares:

There is compelling evidence to suggest that the house price and income variables come from the same probability density function in Islas Baleares. Islas Baleares house prices are among the highest in all of Spain. Z-scores of Islas Baleares towns in All Towns analysis give the impression of overvaluation however the p-score signifies that the distributions almost perfectly align. The Islas Baleares KS two sample test produces a p-value of 0.978. Regarding the provincial data, Islas Baleares incomes are 0.39 standard deviations to the right of the national mean, at €25,497, whereas prices are 3.10 standard deviations greater, at €3,439/m<sup>2</sup>. Affordability ratio would suggest that the prices unexplainably high but p-value = 0.978 suggests almost perfect symmetry between distributions.

The market with the highest average price and near perfect correlation between the distributions presents a phenomenon. House prices are more positively skewed than income meaning there is a smaller proportion of house data above the mean than income. Furthermore, both house price and income kurtosis are negative showing an absence of fat tails (HP: -0.64, I: -0.38). The statistical results show the standardised residuals correlate closely to the theoretical quantiles. The distributions correlate very strongly.

The house prices in Islas Baleares exceed their fundamental values considerably due to the abundance of non-fundamental factors (Alvarez-Roman & García-Posada, 2021). This regional market benefits from the same intrinsic singular quality that Canarias does, which indicates investors cannot find opportunities with truly similar attributes. According to the statistics in Table 10, Islas Baleares and Canarias are very similar markets. With almost identical proportion of primary residencies, mortgages value/house price statistics and growth rates. Where the two autonomous communities differ is the significantly largest foreign investment that Islas Baleares receives (34.4%). This region attracts more capital-rich investors. The high proportion of primary residencies also limits the housing stock available for foreign investment which will put upward pressure on price.

The extreme low house price values (left tail) emphasise the significance of non-fundamental drivers of price in the autonomous community. The lowest priced towns, Inca (€1,807) and Alaoir (€1,808), are 1.8 standard deviations below the Islas Baleares. Geographically these are the most central points on the islands. This demonstrates the significance that “closeness to the sea” has on house prices.

The statistical tests suggest, although highly correlated, that Islas Baleares is more stable than Canarias. The negative kurtosis indicates the lack of fat tails. There is evidence of asset collateral accumulation across the entire Islas Baleares regional market and lead to a significantly large average house price statistics. The closeness of data and correlation blurs market segmentation in Islas Baleares making it a unified market. Although these markets are considered less susceptible to top price segments leading to overvaluation and “bubble-like behaviour”, it must be duly noted that these markets are highly susceptible to international risks such as Covid-19. The exceptionally high growth rates may be consequential to the reopening of international borders.

#### Canarias:

The Canarias KS two sample test exhibits a very high similarity of distribution between house price and incomes, with a p-score of 0.913. The null hypothesis is accepted with great confidence that the two variables come from the same probability density function. The regional market represents a special case due to its abundant non-fundamental factors and attraction as a holiday destination and its insular configuration. Average prices (1,897€/m<sup>2</sup>) are significantly higher than the national average, whereas average incomes (€22,631) are significantly lower. Although affordability ratio appears very high and unsustainable, the distributions come from the same probability density function. On the contrary, the high p-value accompanied by very acceptable levels of skewness and kurtosis suggest that house prices are very coordinated with incomes.

House prices demonstrate a negligible presence of fat-tailedness, with a significantly low excess of kurtosis (at 0.08) and standardised residuals correlating very closely to theoretical quantiles. On the other hand, income does suggest slight fat tailedness at  $SD > 2$ , however this can be expected when average income is so low. Moreover, income is more heavily skewed than house prices, which indicates a higher proportion of data above the mean for house prices

than income. This alludes to potential collateral channel effects occurring at the top segment in the market and price signals causing upward pressure on middle-to-high price segments.

The ERI reports and government statistics indicates that the abundance of non-fundamental factors may cause upward pressure on house prices. Canarias has a below average proportion non-principal residency housing stock. Under the assumption that residents with primary residencies in Canarias will stay, below average secondary residence housing coupled with high turnover and foreign investment may suggest upward pressure. It can be posited that foreign investors deem the returns property in Canarias as sufficiently rewarding due to limited stock available. The islands attract wealthier investors considering the average price of the market coupled with the lowest mortgage value to transaction price in Spain. Investors require lots of capital and therefore interest in the market, which require high returns. The high growth rate in 2022 (12.7%) is consolidated by inherent lack of spill-overs into other autonomous markets. Investors with lower capital capacity or risk-reward positions cannot encounter opportunities with similar attributes but lower price due to the absence of opportunities with truly similar attributes. The abundance of non-fundamental factors makes the opportunity singular.

Although the KS test shows a strong correlation between house price and incomes distributions, other indicators suggest possible accelerator effects, especially caused by its insular configuration.

b)  $0.60 < p < 0.80$

### Extremadura

Extremadura is the autonomous community with the lowest price and income averages, with values of 921€/m<sup>2</sup> and €19,928. The high p-value of 0.836 demonstrates the closeness in distributions. The combination of lowest price and income, and high similarity of distributions suggest low asset and credit activity in the regional market.

House prices demonstrates a platykurtic distribution with a negative kurtosis of -0.06. There is a significant lack of activity at the highest price segment. Bajadoz represents the most extreme value with a Z-score of 2.30. The affordability ratio for Bajadoz is the same as the national mean. The combination of lowest regional price and lack of extreme values ferment the lack of fat tails in the market to create upward price pressure within the market.

Extremadura has the highest mortgage value to transaction price value of all the autonomous communities whereby the average mortgage is 94.4% of the house price. Alongside this it represents the AC with the lowest number of transactions in 2022 (1,034) and the lowest foreign investment (1.2%). Extremadura provide evidence opposite evidence to the asset collateral model. The projects are considered bad and whereby the loan-to-value ratios are likely to increase as opposed to favourably decreasing. Entrepreneurs are not required to take a large stake in the investment and moral hazard issues may arise. This market is considered highly unattractive. The Extremadura regional markets face the risk of opposite collateral effects and negative accelerator effects. This marked is defined as underheated.

Purely based on the statistical evidence, a hypothesis can be made regarding the type of buyer and the type of investment. Extremadura has the highest proportion of non-primary residencies at 34.7%. The troublesome market outlook suggests that investors searching for good projects and healthy returns will not consider the market. Therefore, a deduction can be made that the majority of this non-principal residency stock is not for investment purposes but a leisure residency by Spanish citizens. Holiday home buyers are not obligated to rent or consider rental yields. These houses exhibit characteristics closer to consumption goods than investment assets. Spanish citizens who are priced out of the more demanded holiday destinations have the option to buy in Extremadura. This hypothesis is based off of the following key variables:

low price, low frequency of data 2 standard deviations above the AC mean, high mortgage value/house price ratio, low transactions volume (as holiday homes tend to be kept for a long time) and the high portion of non-primary residencies. Reasons to invest or buy should be considered a non-fundamental driver of price which has largely not been considered in previous literature.

### Castilla-La Mancha

Castilla-La Mancha exhibits almost identical behaviour to Extremadura according to the statistical results. The KS test value of p-value of 0.809 demonstrates the high likelihood that both variables exhibit the same probability density function. Castilla-La Mancha has the second lowest prices in Spain at €1,043. Although the prices are 13% higher than Extremadura, affordability ratios are identical for both of the regional markets. The statistical tests for Castilla-La Mancha suggest the same opposite collateral channels and potential negative accelerator effects as Extremadura.

The thin tailedness at the highest price segment in Castilla-La Mancha is highlighted by its significantly negative excess of kurtosis (-1.18). This housing market's most expensive segment uniquely has the lowest extreme values, whereby the standardised residuals drift down from the theoretical quantiles one standard deviation above the mean. The lack of suitably good projects to undertake causes a platykurtic distribution of price. The abnormally low frequency of house prices at the higher intervals is indicative of the hypothesis applied to Extremadura, whereby entrepreneurs are willing to undertake poor projects as holiday homes. At the higher price levels, entrepreneurs seek better holiday locations closer to the sea. This market is considerably large (INE, 2023) however only experienced 267 transactions in 2022 (ERI, 2023). The large mortgage value coupled with low prices allows holiday homebuyers to keep their home for long periods of time without the intention of selling.

$$0.40 < p < 0.60$$

### Aragón

Aragón price and income samples are likely come to from the same probability density function, with a p-score of 0.582. Both house price and income variables are below the national mean at 1,378€/m<sup>2</sup> and €25,049. Figure 11 displays the significantly more positively skewed nature that price exhibits than income (P: 2.06, I: 0.34) and shows presents the negative fat tail of exhibited by income but not price. Price tends to exert more variability than income in these regional specific tests however in this case price has a higher probability of X lying closer to the central location. This is confirmed by the differences in interquartile range: P: 0.93, I:1.40. The results allude to a theoretical lower bound or floor in the Aragón house market, as shown by the stark increase in frequency from  $SD > -0.5$ .

Aragón exhibits the same price floor behaviour as Cataluña.

Aragón houses price tends to be more tightly distributed around the central tendency than income – a phenomenon for this test.

The Aragón regional test present evidence for spill overs across regions. The left tail of house price distribution exhibits price floor behaviour whereby there is a theoretical lower boundary do not reduce to. This characteristic is exhibited by the closest neighbouring region, Cataluña, as previously discussed. A hypothesis can be made that there are abundant non-fundamental factors that warrant Aragón property to be considered as a good second home project. The test suggests that Aragón is considered a moderately attractive project when compared a neighbouring market with high prices and similar attributes.

Aragón exhibits an extraordinary excess of kurtosis unique from the other regional tests. Due to two significant outliers in the high price segment, excess of kurtosis for price is 6.44 giving the distribution a significant heavy tail. The extreme values are attractive tourist spots due to their significantly abundant non-fundamental value drivers.: Sallent de Gállego (3,176€/m<sup>2</sup>) and Biescas (2,502€/m<sup>2</sup>). House prices are very high due to the unique and irreproducible offering given by these skiing towns.



The suggestion of potential spill-overs from adjacent markets and the high variation in incomes relative to house prices does indicate the potential for house prices to grow. The tightness house prices exhibit to the mean compared to income is unique to Aragón. On the contrary, excluding the extreme values mentioned due to their inherent characteristics, there is little evidence of a fat tail by which asset collateral channels can generate positive market sentiment and accelerator effects.

The overvaluation report by the Bank of Spain suggests that Teruel, a province within Aragón, was the only mainland market to be considered overvalued in 2018 when compared to its fundamental factors. The results of this study would suggest that considering distribution of house prices and income, there is little evidence at the present moment of a bubble forming in Aragón's housing market.

### Cantabria

The Cantabria regional KS two sample test produced a p-value of 0.560 thus exhibiting moderate similarity between house price and income distributions. This is caused by the significant lack of extreme house price values and the negative kurtosis statistic of -1.15. The lack of extreme values induces a platykurtic distribution where the highest price local markets are not considered sufficiently good projects. The lack of confidence in the most attractive markets does not send upward pressure signals into lower priced markets with similar attributes.

The Cantabrian market importantly implies the limiting power of the credit market in affecting asset valuations. The statistical analyses suggest the Cantabria market to be undervalued and unattractive due to its similarities with Aragón, however Cantabrian regional interest rates are the lowest in Spain, 2.17%. This highlights the limitation of cheap credit to incentivise investment when a project is deemed insufficiently good due to its fundamental and non-fundamental value drivers. Cantabria ferments the hypothesis that credit expansion leads to accelerator and decelerator effects rather than directly incentivising decision making. The relatively high price coupled with low rates shows the banks willingness to expand credit markets but the lack of fat tails does not lead to asset collateral accelerator effects at the highest price segments.

## Galicia

The Galician KS two sample tests produces a moderate p-value, of 0.527, suggesting moderate confidence that both variables come from the same probability density function. Although both prices and income are below the national average at 1,365€/m<sup>2</sup> and €23,330, there is evidence of fat tail in both variables. Galicia represents a special case whereby excess of kurtosis for price is sufficiently large (1.20). Whilst kurtosis of income exhibits relatively predictable behaviour considering the market, the presence of a fat tail hints towards asset collateral effects in the market.

Whilst skewness is almost identical for both variables, the Galician house price and income variables distributions suggest a divide into two markets, a high price segment and a non-high price segment. At  $SD > 1.2$ , the standardised residuals exceed theoretical residuals considerably. There is a significantly low frequency of data between  $0.5 < SD < 2$ , however there are important extreme values of price at  $SD > 2$ . These are Sansenxo, Nigrán, A Coruña and Santiago de Compostela. There is an abundance of non-fundamental factors in these towns. Affordability ratios greatly exceed the average Galician affordability ratio. Considering these extreme values, alongside the high non-principal residency proportion, high transaction volume and the low foreign investments, a hypothesis can be made of the general interest in second property for Spanish investors. Galician extreme house price values present good projects in the extreme house price values. There is some evidence to suggest that asset collateral channels are putting upward pressure on the extreme values and the surrounding local markets do not have adequately similar attributes to be considered equally good projects.

## La Rioja

La Rioja KS two sample tests produces a moderate p-value, of 0.512, suggesting moderate confidence that both variables come from the same probability density function. Average regional prices are 1,245€/m<sup>2</sup> and average regional incomes at €24,379. The lack of significantly extreme house price values suggest low accelerator effects in the market. Moreover, the relationship kurtosis between the two variables has proven to suggest unattractive markets – whereby kurtosis of income greatly exceed that of price (I: 5.52; HP: -0.47). The lack of fundamental and non-fundamental drivers of value fail to present asset collateral channels for intense expansion. This statistical analysis coupled with low non-

principal residencies, low market turnover and low foreign interest, suggest a stagnant market with market imbalances. The overvaluation report shares this sentiment placing La Rioja in the most undervalued segment.

### Valencia

The Valencian region KS two sample test produces a p-value of 0.493 whereby we accept the null hypothesis that the variables come the same distributions. The average Valencian affordability is relatively due to its high average prices and below national average income (1,707€/m<sup>2</sup>; €23,885). Excess Kurtosis of price is negative (-0.06) which indicates that price in this AC exhibits platykurtic distribution whereas income has a highly positive excess of kurtosis (3.35) showing a significant fat tail.

The Valencian region (as opposed to the provincial area) attracts high interest from foreign investors and exhibits high market turnover due to its abundant non-fundamental factors (ERI, 2023). The extreme price values are all concentrated around the sea and big cities, which demonstrates the importance the non-fundamental factor, “closeness to the sea”, oasset values in this regional market. The magnitude of this factor is portrayed by the only landlocked province, Albacete, only containing one datapoint in the top twenty most expensive towns in the Valencian region. The market does not appear to have a fat tail but there is some suggestion of a growing separation of the provincial markets within the regional market. Foreign investment constitutes a considerable 26.8% of the 99,869 total transactions in 2022. The variance in goodness of projects causes collateral channel effects in varying magnitudes across the region. Valencia exhibits similar characteristics to Galicia.

The Valencian statistical tests demonstrate the significance of foreign investment in spurring asset collateral channel effects. In previous tests foreign investment ratios in this range have not occurred in markets with such high turnover, showing the abundance of non-fundamental factors in this large market. The ERI report highlights the stark difference in foreign investment between the three provinces. The portion of transactions by foreigners are: Alicante 42.9%, Valencia 16.7%, Albacete 2.57%. Alicante receives the most foreign investment of all the provinces in Spain. Alicante also exhibits the greatest difference in house price and income distributions. Foreign investors tend to invest in larger projects than citizens; average transaction price by foreign entrepreneurs is 12% higher than by domestic entrepreneurs

(Ministerio de Transportes, Movilidad y Agenda Urbana, 2023). Investment in larger projects will put upward pressure on the fat tails to drive prices up further. Valencia demonstrates perfect evidence of high foreign investment provoking lower regional interest rates to cause credit expansion.

$$0.05 < p < 0.40$$

### Castilla Y León

With respect to p-values the Castilla y León test stands alone with a p-value of 0.357. House prices and incomes are significantly different, but the null hypothesis is accepted that the distributions came from the same probability density function. Castilla y León exhibits characteristics very similar to Castilla La Manche across all complementary metrics in Table 10. Prices and incomes are very low compared to the national average – P: 1,137€/m<sup>2</sup> and I: €23,757 – and the price distribution exhibits a slightly negative kurtosis of price but very positive kurtosis of income. The distribution of price is highly normal with little to no evidence of fat tails.

The affordability ratio is useful in this test to explain the abundance of affordable local markets in Castilla y León. Only 20% of the sample data exceeds the national average affordability ratio which is inherently positively skewed. This overvaluation results are concurrent and a predominant factor in their analysis would be its very low population density; this factor produced a cointegration vector of 1.032 (the highest of the fundamental drivers of price). The market does not exhibit asset collateral channel effects and the market is deemed unattractive due to lack of fat tails to signal upward pressure in the market.

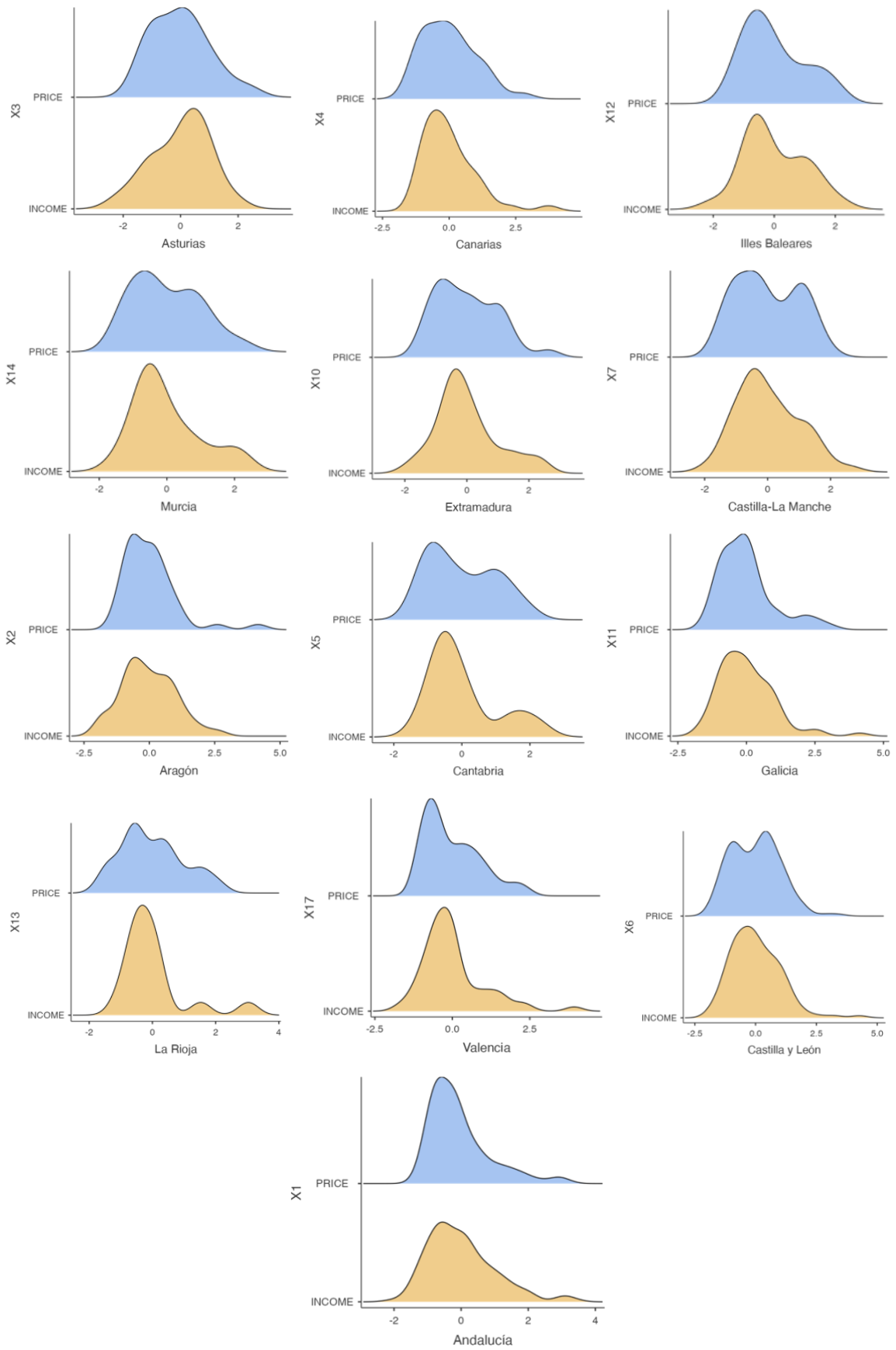
### Andalucía:

The Andalusian KS two sample test produces the lowest p-value of all the regional markets, p-value = .118. Andalusian housing distribution and income distribution are the least like variables of all the test to follow the same distribution. The regional mean is very close to the national mean at 1,411 €/m<sup>2</sup>.

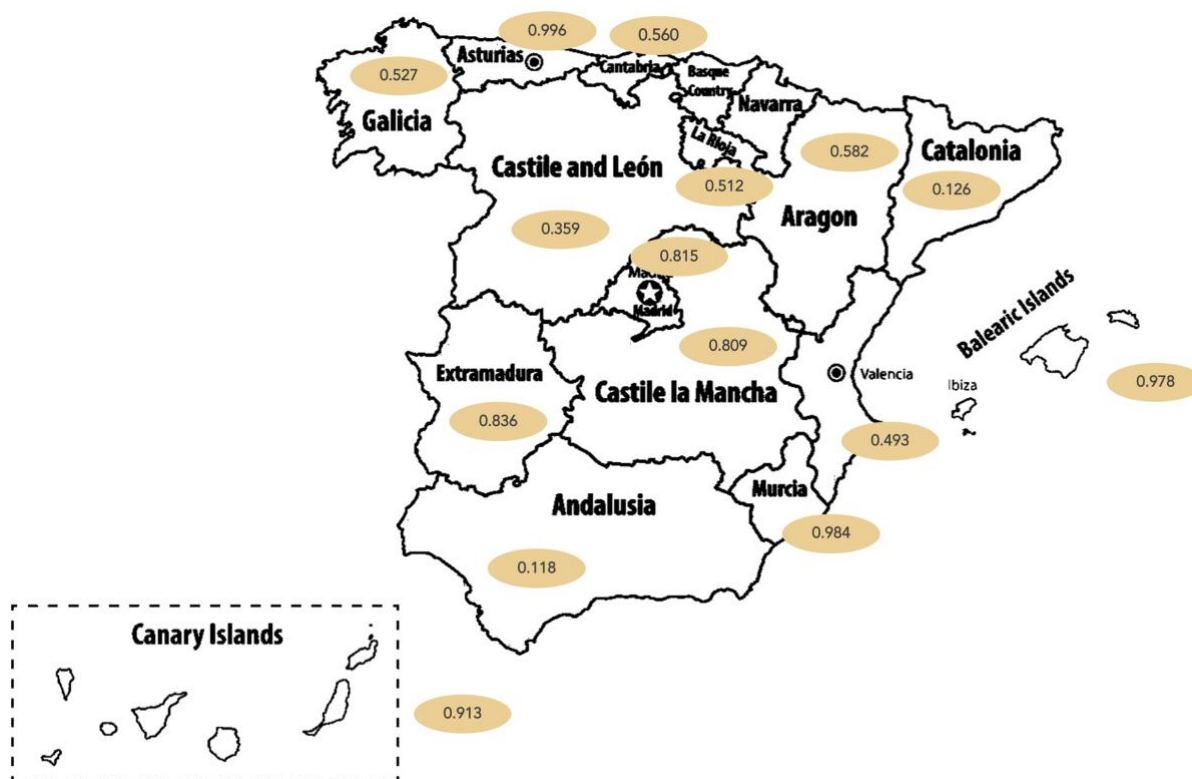
The greatest difference between the Andalusian KS test and the rest of the tests whereby the null hypothesis was accepted is the excess of kurtosis in price. The test produced a price excess of kurtosis of 2.07 which is significantly greater than the rest of the KS two sample tests. This result coupled with the Q-Q plot suggests that Andalucía has the greatest price fat tail of these KS tests. There is evidence for house unaffordability at the highest price segments; 10 local

markets produce affordability z-scores two standard deviations above the national mean. This fat tail distinguishes the Andalusian market from other Spanish regional markets and fundamentally causes the substantially lower p-value. These local markets are concentrated near the sea in Málaga and Cadiz, confirming the importance of non-fundamental drivers of price in Costa del Sol. The growth rates specifically seen in these markets create signals to markets with similar attributes to grow. There is evidence to suggest that the houses represent good projects with high returns. There is sufficient evidence for asset collateral channels at the most extreme house price values,

This autonomous community market shows evidence of the highest price segments driving price in the rest of the market. It also suggests the importance of foreign investment in the highest segment to affect market sentiment. The overvaluation report deems Andalucía the least undervalued market in 2018. Although the null hypothesis is accepted in this test, the considerably lower p-value alongside the structural and cyclical factors analysed alongside the test, suggests bubble forming behaviour whereby the fat tails drive average price.



## Conclusion



There is real evidence in the Spanish housing market that the imperfect asset class, housing, demonstrates asset collateral channel and credit expansionary effects due to supply lags. To assess the degree of overheating with a housing market, it is fundamental to examine the degree of the asset collateral effect and credit expansion. The entrepreneur's incentive to encounter the best housing project possible to benefit for refinancing and the financial lender's incentive to increase confidence in cover non-performing loans gears the market conditions to grow in unregulated housing markets. It is the intrinsic large stake in the investment that the entrepreneur must take to incentivize good decisions that causes asset bubbles in the Spanish housing market.

Asset collateral cycles worsen housing affordability across a regional market. The incentive to search for good projects leads to distributional fat tails at the highest price segments. These statistical anomalies send signals to other markets with similar attributes to adjust price in order to also benefit from lower loan-to-value ratios and increase household equity. Spill over effects take many forms; through price segments, across local markets and across regional markets.



The joint pursuit of higher asset values increases the difference between fundamental value and asset values, causing bubble-like effects. The inherent high variance of price compared to income causes asset collateral channels to worsen housing affordability across the market.

Statistical tools accompanied by explanatory variables are essential to judge the “overheatedness” of a housing market. One statistical measure is not sufficient to capture the market behaviour in its entirety. Nor is a multiple linear regression analysis considering fundamental value of the asset, as it does not evaluate bubble-like tendencies. Kolmogorov-Smirnov two sample tests, kurtosis, skewness, extreme value analysis and complementary explanatory variables are fundamental to understand demand forces in housing markets. Markets need to be sufficiently small to uncover the idiosyncratic market conditions specific to said market. Individual markets can present bubble tendencies when the national market is considered in recovery.

Kurtosis is significantly revealing when comparing house price and income distributions. Overheated markets are suggested by good projects. The presence of good projects in statistical terms are shown by high price kurtosis relative to income kurtosis, very high extreme house price values, lower interest rates on mortgages, low mortgage value to house price ratio and high foreign interest. Underheated markets have abundant bad projects which are statistically suggested by low price kurtosis relative to income kurtosis, platykurtic distribution of house prices, high interest rates on mortgages, high mortgage value to house price ratio and low foreign interest.

Income is a powerful variable to asset values, but it is not sufficient solely due to the abundance of fundamental and non-fundamental factors that drive asset price. The entrepreneur’s incentive to undertake good projects that increase in value inflates stakeholder perception of valuation and leads to credit expansion and collateral cycles.

The Madrid regional housing market is the purest form of asset collateral channels and credit expansion due to its unique attributes, and its abundant fundamental and non-fundamental factors of value. Although asset values appear very high, the statistical tests suggest that the market is stable. The Madrid market also provides evidence that stakeholders are incentivised to take on larger projects due to external financing benefits.

Future studies should use dynamic data whereby the house price and income distributions can be compared over a time series. This will demonstrate the dynamic results whereby size and severity of housing asset bubble can be judged.

## Bibliography

- Atlas; Gesvalt. (2021). *Vivienda asequible en España*. Estudio de situación y desglose de planes actuales de vivienda.
- Romero, T. (2022). *Average annual wages in Spain from 2000 to 2021*. Retrieved from Statista: <https://www.statista.com/statistics/419513/average-annual-wages-spain-y-on-y-in-euros/#:~:text=The%20average%20annual%20wage%20stood,in%20France%2C%20Ireland%20and%20Germany.>
- ESRB. (2019). *Vulnerabilities in the residential real estate sectors of the EEA countries*.
- Alvarez-Roman, L., & García-Posada, M. (2021). Are house prices overvalued in Spain? A regional approach. *Economic Modelling* 99 (2021) 105499, 1.
- Ouasbaa, G., & Viladecans Marsal, E. (2021). *Spain's once-substantial rental market is now one of the smallest in Europe*. Retrieved from Brookings: <https://www.brookings.edu/essay/spain-rental-housing-markets/>
- INE. (2023). *Índice de Precios de Vivienda*. Retrieved from INE: <https://www.ine.es/jaxiT3/Datos.htm?t=25171#!tabs-grafico>
- Aquila Capital. (2022). *SPAIN'S RESIDENTIAL MARKET IS DIFFERENT A EUROPEAN OPPORTUNITY*.
- Talenom. (2023). *Attracting Foreign Investment*. Retrieved from <https://avalanding.com/spain-attracts-foreign-investment/>
- Ministerio de Transportes, Movilidad y Agenda Urbana. (2023). *Transacciones inmobiliarias (compraventa)*. Retrieved from Actividades y Servicios: <https://apps.fomento.gob.es/BoletinOnline2/?nivel=2&orden=34000000>
- ERI. (2023). *4 trimestre 2022*. Colegio de Registradores de la Propiedad, Bienes Muebles y Mercantiles de España.
- World Bank. (2020). *International tourism, number of arrivals - Spain*. Retrieved from The World Bank Data: <https://data.worldbank.org/indicator/ST.INT.ARVL?locations=ES>
- Goldstein, I., & Razin, R. (2013). THREE BRANCHES OF THEORIES OF FINANCIAL CRISES. *NATIONAL BUREAU OF ECONOMIC RESEARCH*.
- La Vanguardia. (2022, February 14). *Requisitos para contratar una hipoteca en 2022* . Retrieved from <https://www.lavanguardia.com/economia/finanzas-personales/20220114/7984940/requisitos-contratar-hipoteca-2022.html>
- Allen, F., & Galle, D. (2000). Bubbles and Crises. *Royal Economic Society*, 110, 236-255.
- McMillen, D. (2007). Changes in the Distribution of House Prices over Time: Structural Characteristics, Neighborhood, or Coefficients?
- Agencia Tributaria. (2020). *Posicionamiento de los municipios mayores de 1.000 habitantes por Renta bruta media*. Retrieved from [https://sede.agenciatributaria.gob.es/AEAT/Contenidos\\_Comunes/La\\_Agencia\\_Tributaria/Estadisticas/Publicaciones/sites/irpfmunicipios/2020/jrubikf7024f7e68a6d19091b823f761818f14de77d7950.html](https://sede.agenciatributaria.gob.es/AEAT/Contenidos_Comunes/La_Agencia_Tributaria/Estadisticas/Publicaciones/sites/irpfmunicipios/2020/jrubikf7024f7e68a6d19091b823f761818f14de77d7950.html)
- Fotocasa. (2023, March 2). *Índice Inmobiliario de precios Fotocasa*. Retrieved from <https://www.fotocasa.es/indice-precio-vivienda>
- Hair, J., Black, W., & Babin. (2010). *Multivariate Data Analysis*. 7.
- Fiori, A., & Zenga, M. (2009). Karl Pearson and the Origin of Kurtosis. *International Statistics Review*.
- Eurostat. (2022). *Housing in Europe*. Retrieved from Eurostat: <https://ec.europa.eu/eurostat/web/products-eurostat-news/w/DDN-20230216-1>
- Causa, O., & Woloszko, N. (2019). *HOUSING, WEALTH ACCUMULATION AND WEALTH DISTRIBUTION: EVIDENCE AND STYLIZED FACTS* . OECD.

- Young, I. (1977). PROOF WITHOUT PREJUDICE: USE OF THE KOLMOGOROV-SMIRNOV TEST FOR THE ANALYSIS OF HISTOGRAMS FROM FLOW SYSTEMS AND OTHER SOURCES. *THE JOURNAL OF HISTOCHEMISTRY AND CYTOCHEMISTRY*.
- TDW. (2023). *Is 20 A Good Sample Size For Quantitative Research?* Retrieved from <https://www.thedonutwhole.com/is-20-a-good-sample-size-for-quantitative-research/#:~:text=In%20general%2C%20sample%20sizes%20of,to%20the%20Central%20Limit%20Theorem>.
- Pinterest. (2023). Retrieved from <https://www.pinterest.com/pin/a-guide-to-the-regions-of-spain--440438038548707674/>
- INE. (2023). *Population and Housing Census*. Retrieved from [https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica\\_C&cid=1254736176992&menu=ultiDatos&idp=1254735572981](https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176992&menu=ultiDatos&idp=1254735572981)
- Carpio-Pinedo, J. (2020). Consumption and symbolic capital in the metropolitan space: Integrating ‘old’ retail data sources with social big data.
- Garriga, J. (2023). *What big data reveal about Spain’s supply of new housing and demographic trends* . Retrieved from Caixabank Research: <https://www.caixabankresearch.com/en/sector-analysis/real-estate/what-big-data-reveal-about-spains-supply-new-housing-and-demographic>
- OECD. (2010). *Summary of Hypothesis Test Null Hypothesis Sig.a,b Decision 1 La distribución de House Price and Income across ALLTOWNS. <.001 Reject the null hypothesis with a 0.05 significance level* . Barcelona.

# Annex

