



# Estimating regional house price levels: Methodology and results from a pilot project with Spain

Pierre-Alain PIONNIER and Johannes SCHUFFELS  
(OECD – Statistics and Data Directorate)

Recent trends in the real estate market and its analysis –  
2021 edition (5 Nov. 2021)



## A broader OECD project on Housing...

---

- Our recent statistical projects are building blocks of a [broader OECD project on Housing](#), looking at several issues:
  - Housing, economic resilience and performance (e.g. sustainability of house price developments, origin and effect of house price bubbles on the functioning of the economy)
  - Housing and labour mobility (e.g. differences in housing affordability across regions acting as a barrier to labour mobility)
  - Housing and inequality (e.g. housing contributing to wealth and income inequality across households)
  - Housing and environmental efficiency
- For all these issues, understanding challenges and improving economic policies requires to **go local**.



## ... with a strong statistical and subnational component

---

Phase 1 (completed): Collecting and analysing information on regional house price *developments over time*

- [OECD database on national and regional house price indices](#) (only contains information from official statistical sources)
- [Statistical Insights](#): Location, location, location – House price developments across and within OECD countries
- Recently relaunched [web application](#) allowing to visualise regional house price developments with maps



# OECD Housing project

## Estimating regional house price *levels*

---

### Phase 2 (ongoing): Focus on regional **house price levels**:

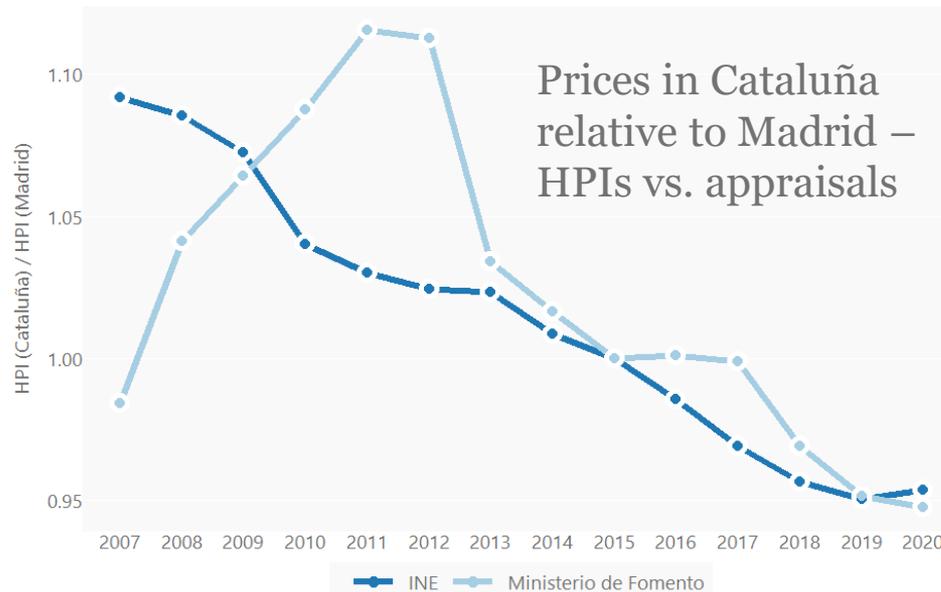
- House price *indices* allow tracking house price developments over time.
- House price *levels* are needed to assess **housing affordability**, potential **barriers to labour mobility** across regions, and for the design of **housing policies**

Work with **Spain** as a pilot country to estimate house price levels that are **representative of the stock of dwellings**, with a method that is **scalable** to other countries.



# Motivation: House price levels rarely available from official sources

- In the absence of official statistics on house price levels, public or private data providers may fill the gap.
- **Alternative data providers** may rely on different data sources (e.g. asking vs. transaction prices) and/or not adequately control for quality effects.
  - > **Risk of inconsistency with HPIs.**
  - > **Need for NSOs to coordinate or take the lead.**





# Relationship between house price *levels* and *indices*

---

- The quality of properties in the stock of dwellings can be considered stable from one year to the next: mostly second-hand dwellings, investments
- Therefore, the evolution of **representative house price levels** should correspond to pure price changes only
- Where they don't correspond to the evolution of quality-adjusted HPIs, quality and price changes are mixed



# Overview of the OECD method

---

- **Two-stage method:**
  - First estimate stratified house price levels, using transaction prices of second-hand dwellings and stock weights (i.e. the number of dwellings in the stock) to aggregate strata.
  - Then improve upon these first-stage estimates by using the information on house price changes provided by quality-adjusted HPIs.
- **Benefits of the OECD method:**
  - Estimated house price levels have the same evolutions as the corresponding HPIs provided by NSOs.
  - They use the available information on house price levels at all dates without having to choose a reference period arbitrarily, thus leading to only minor revisions over the past when new data come in.

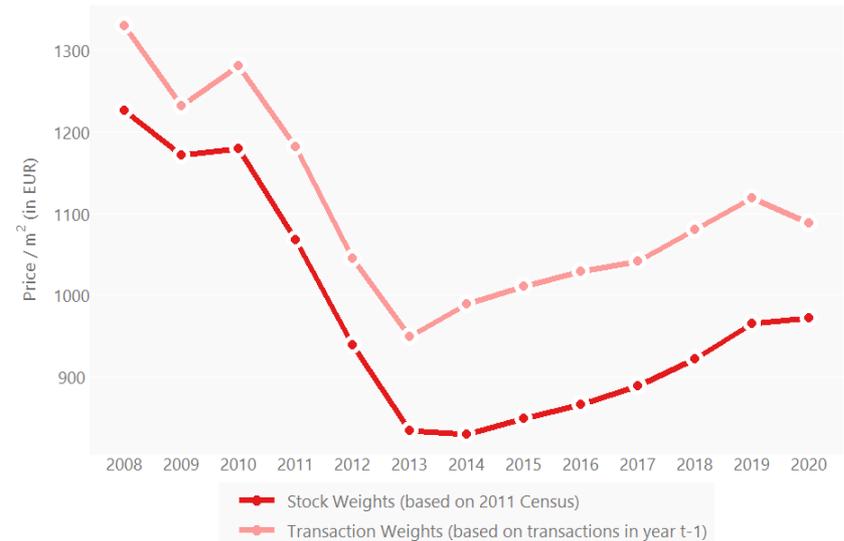
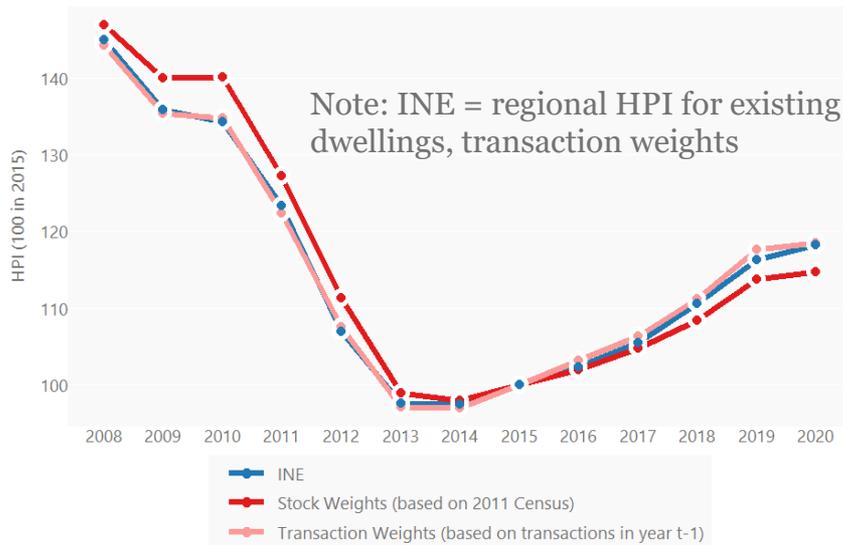


# Stratified house price levels: Sensitivity analysis (1/2)

## Sensitivity to the type of weights (stock vs. transaction weights):

- **Weights matter:** Using stock weights increases the weight of rural areas and reduces regional house price levels by up to 10-20% for some Spanish regions (e.g. Andalucía).
- Indices are usually less sensitive to the type of geographical weights than price levels.

### House price levels in Andalucía with a stratification at municipality level and either stock or transaction weights to aggregate strata (2008-20)



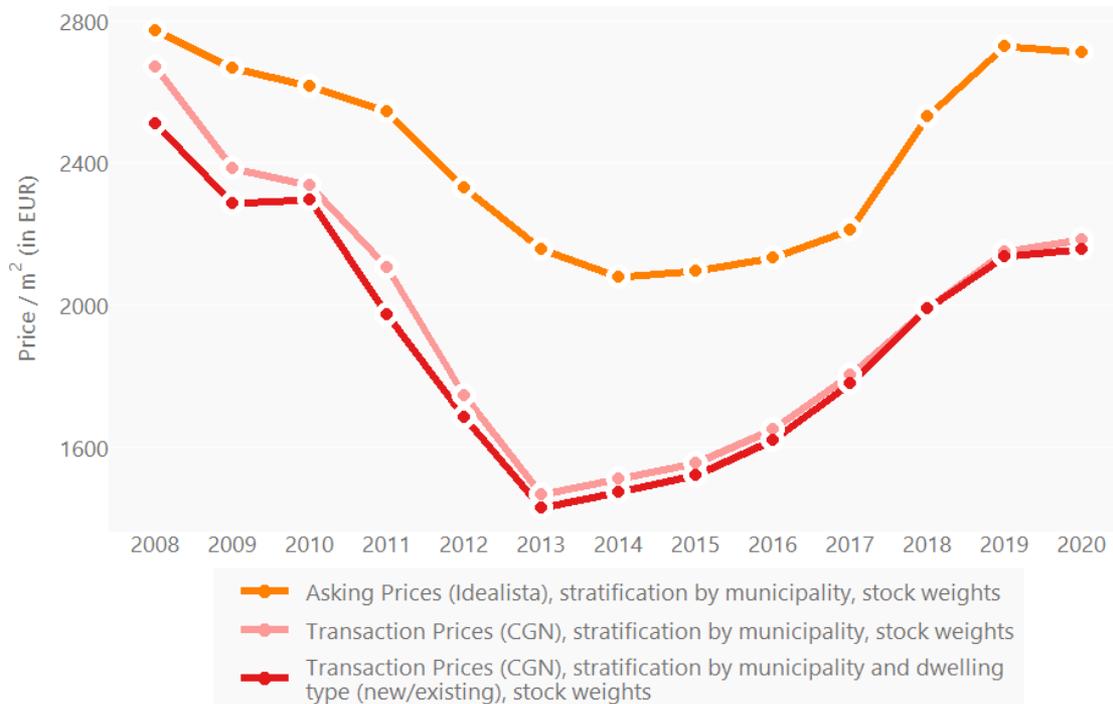


# Stratified house price levels: Sensitivity analysis (2/2)

Sensitivity to the type of input prices (asking vs. transaction prices):

- **Differences are region specific, time varying and potentially very large.** ([more details](#))
- Reasons: selection bias; price cut dependent on business cycle

**House price levels in Madrid based on asking or transaction prices (2008-20)**



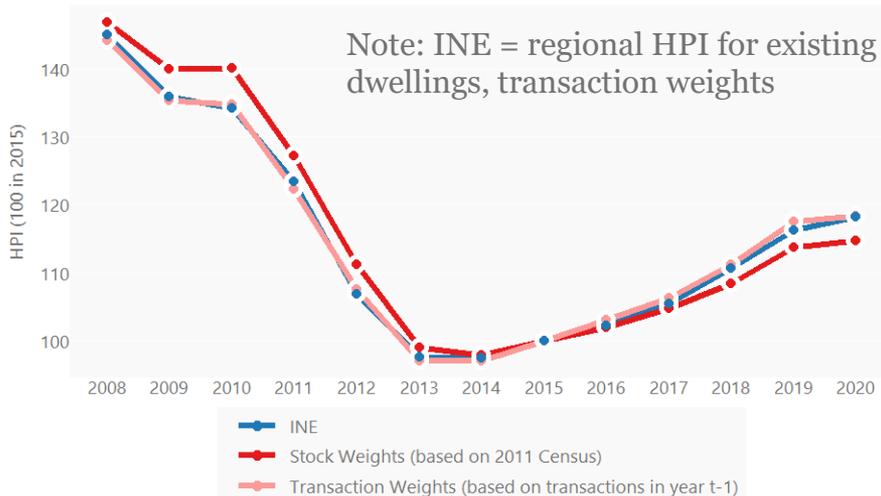
Source: Idealista, *Consejo General del Notariado (CGN)*, OECD calculations



# Discrepancies between the evolutions of stratified house price levels and HPIs

- Even when relying on transaction prices at municipality level, **stratification cannot fully control for quality effects in all regions.**
- Discrepancies appear especially in the years after the financial crisis of 2008-09.

## Andalucía



## Comunitat Valenciana



Source: INE, Consejo General del Notariado (CGN), OECD calculations



# Using HPIs to improve the accuracy of stratified house price levels

## Second stage of the OECD method:

- Consider stratified house price levels as noisy estimates of the true, but unobserved, representative house price levels (noise due to unaccounted quality effects).
- Since the evolutions of representative house price levels should correspond to pure price changes, quality-adjusted HPIs can provide additional information to estimate house price levels.
- State-space models provide the adequate statistical framework to combine stratified house price levels with HPIs.
- Similar in spirit to [Rao et al. \(2010\)](#) who estimate PPP time series based on **benchmark PPPs** and **GDP deflators**.

Here playing the same role as the noisy house price level estimates obtained with stratification methods

Here playing the same role as quality-adjusted HPIs



# State-Space Model: Specification

- **Measurement Equation:**

$$\log P_t^A = \log \alpha_t + \epsilon_t$$

- $P_t^A$ : stratified house price level at date  $t$  in region  $A$ , using stock weights to aggregate strata. (*In the case of Spain, we focus on existing dwellings, representing more than 95% of the dwelling stock in all regions, at all dates.*)
- $\alpha_t$ : true but unobserved house price level at date  $t$
- $\epsilon_t$ : estimation error (e.g. due to unaccounted quality effects by stratification).

- **State Equations:**

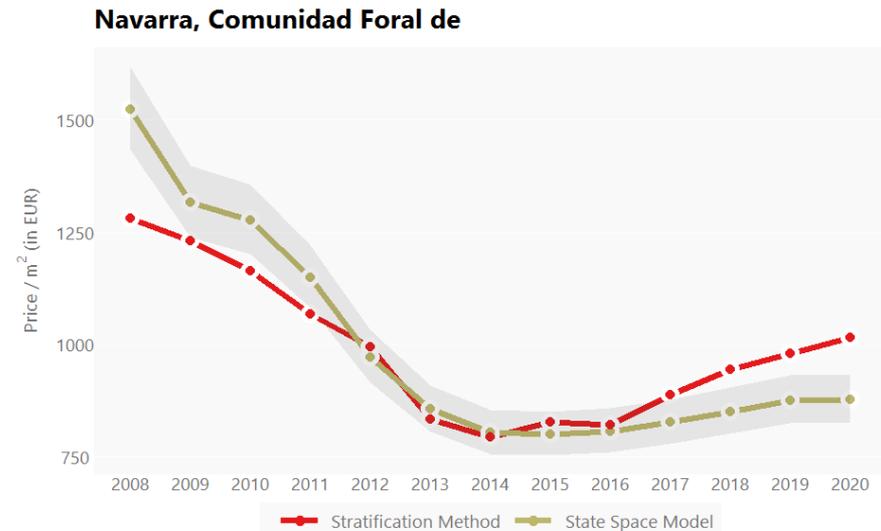
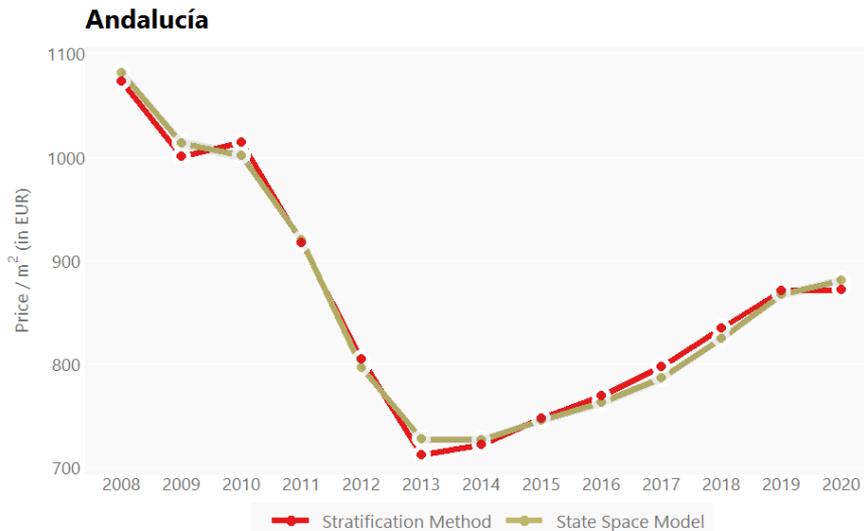
$$\left\{ \begin{array}{l} \log \alpha_t = \log \alpha_{t-1} + \Delta \log(HPI_t^A) \\ \epsilon_t = \sum_{i=1}^p \phi_i \epsilon_{t-i} + \eta_t \end{array} \right.$$

- $HPI_t^A$ : quality-adjusted house price index for region  $A$ , at date  $t$ . (*In the case of Spain, we use transaction-weighted HPIs for existing dwellings as a proxy for stock-weighted HPIs.*)
- Region-specific lag length  $p$



# State-Space Model: Results (1/2)

- The autoregressive parameters  $\phi_i$  and the variance of the Gaussian process  $\eta_t$  are estimated by maximum likelihood.  $\alpha_t$  and the corresponding confidence intervals are estimated with a Kalman Smoother.
- Depending on the ability of stratification methods to control for quality-effects,  $\alpha_t$  may deviate (significantly) from  $P_t^A$  or be close to it.

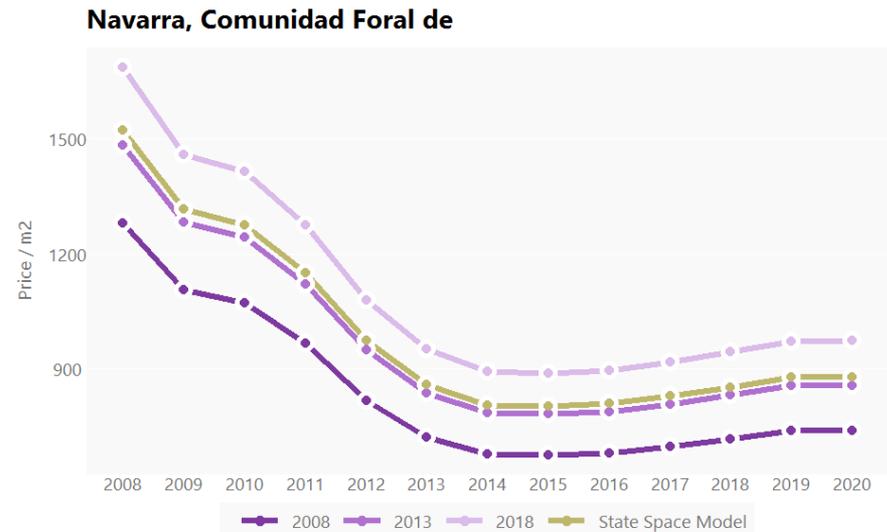
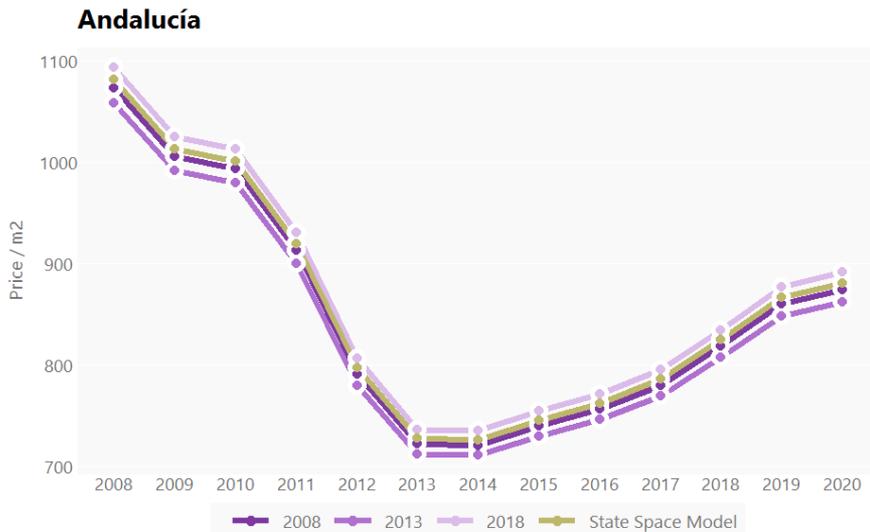


Source: INE, Consejo General del Notariado (CGN), OECD calculations



# State-Space Model: Results (2/2)

- The state-space model uses as input the stratified house price levels that are available for all years.
- This avoids picking up an arbitrary reference year and limits the size of revisions when new data come in ([more details](#)).



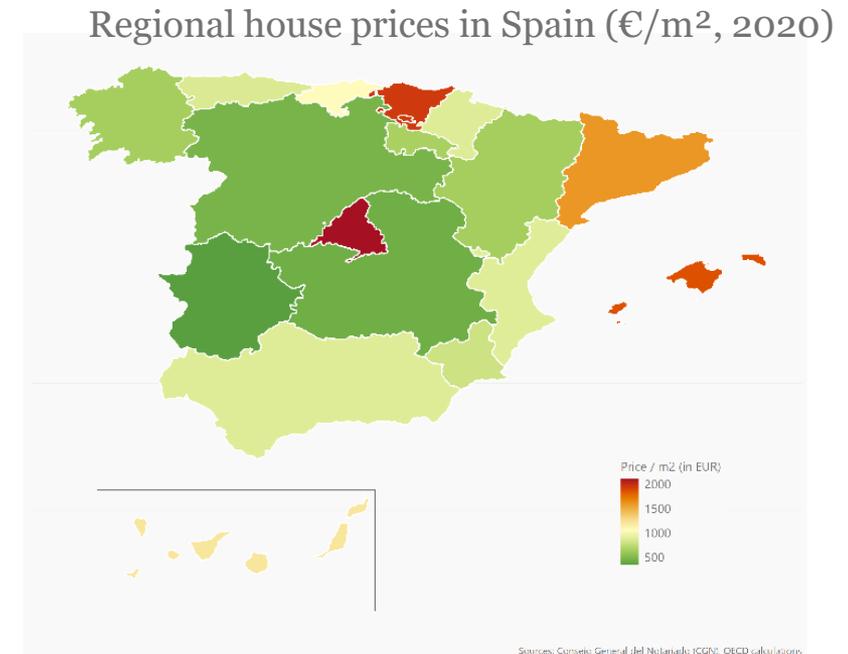
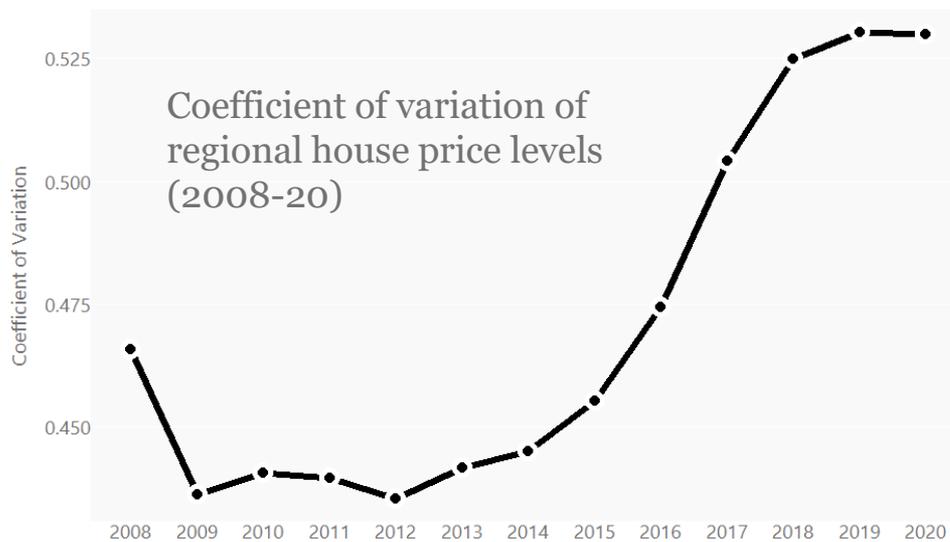
Source: INE, Consejo General del Notariado (CGN), OECD calculations



# Brief economic analysis of the results:

## Price dispersion

- The decline in house prices after the 2008-09 crisis reduced the **heterogeneity in house prices across Spanish regions** but this heterogeneity has increased again since 2012.

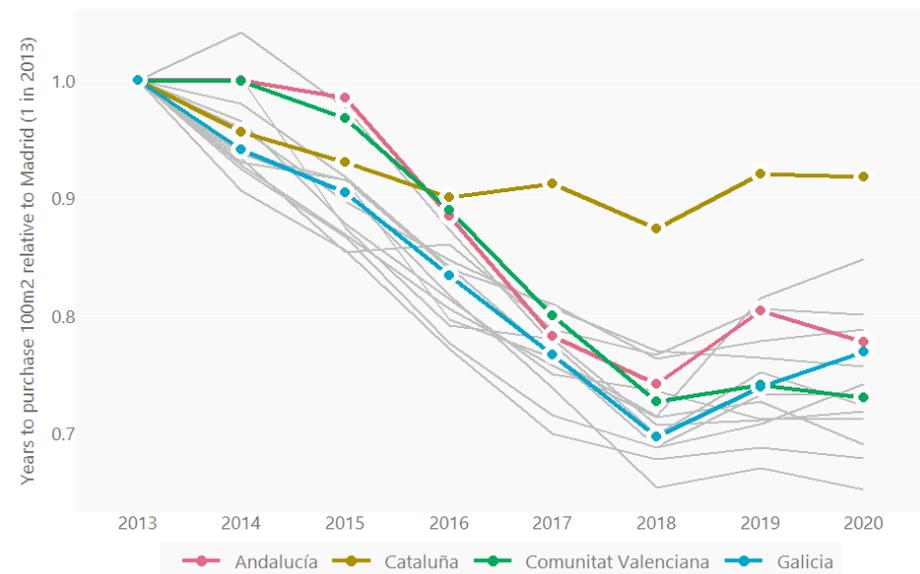
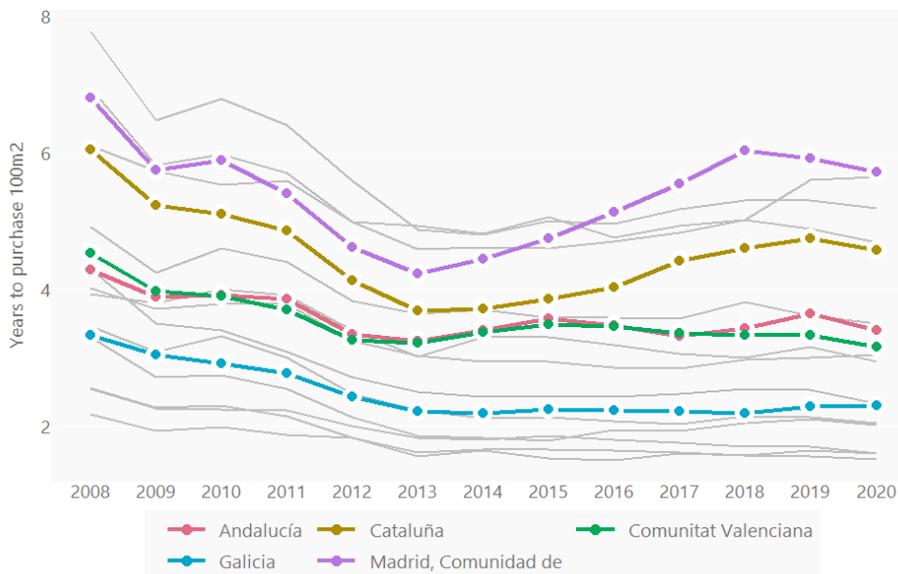


Source: INE, Consejo General del Notariado (CGN), OECD calculations



# Brief economic analysis of the results: Housing affordability

- Compared to the rest of the country, **housing affordability in Madrid** (measured by the number of years of average regional income to purchase a dwelling of 100 m<sup>2</sup>) has deteriorated strongly since 2013.
- High relevance of such statistics in the wake of the **COVID-19** crisis.





# Conclusions

---

Recommendations for the estimation of regional house price levels:

- Improve upon stratified house price levels by using information on evolutions from quality-adjusted HPIs
- Ensure that new dwellings do not distort the results
- Use the same data source on house prices for the compilation of house price levels and indices (i.e. transaction prices or appraisals)



---

Thank you for your attention.

[johannes.schuffels@oecd.org](mailto:johannes.schuffels@oecd.org)

[pierre-alain.pionnier@oecd.org](mailto:pierre-alain.pionnier@oecd.org)

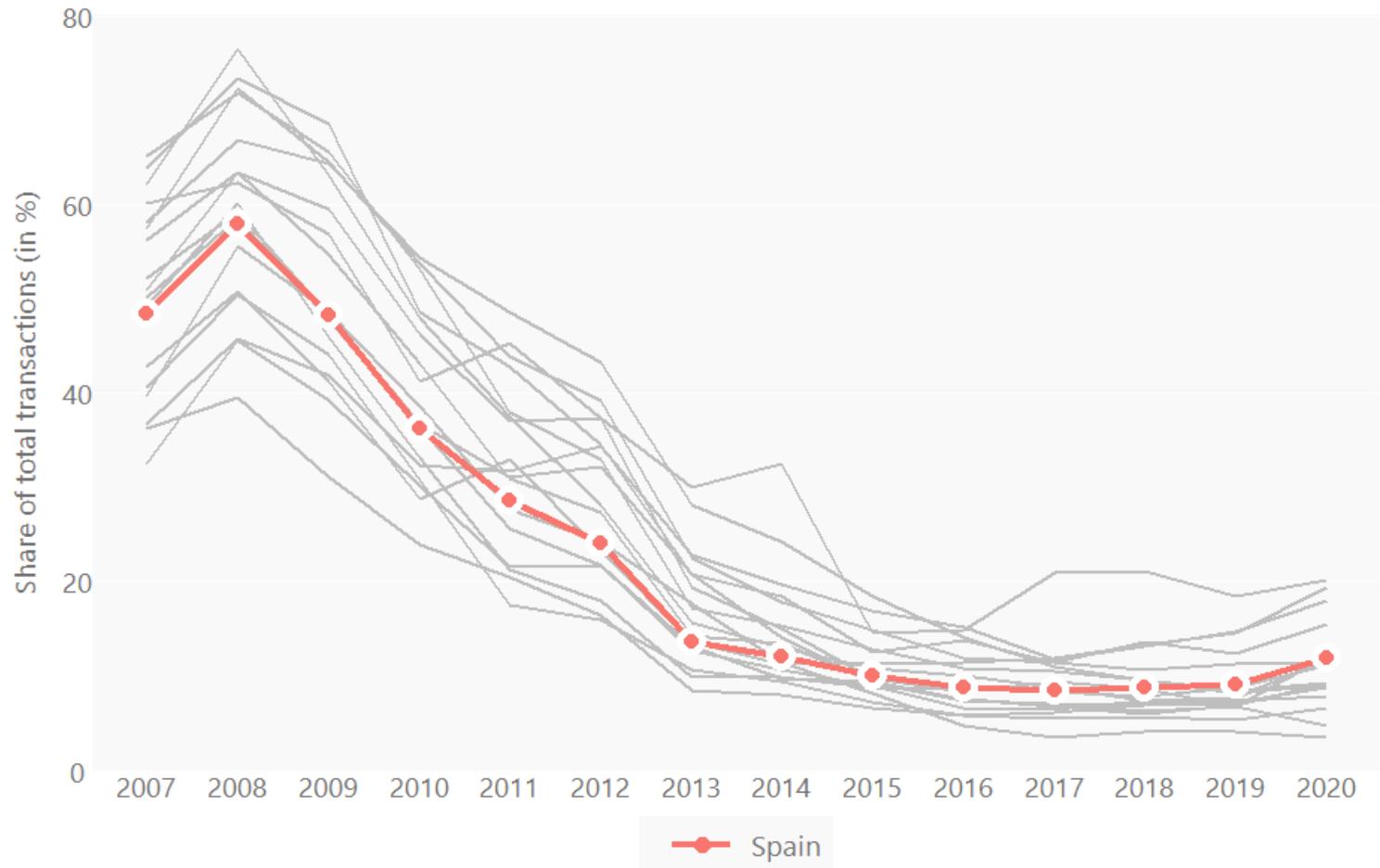
[SDD.Prices@oecd.org](mailto:SDD.Prices@oecd.org)



# SUPPORTING MATERIAL



# Share of new dwellings in total transactions in Spain





# Percent difference between asking and transaction prices

Autonomous Communities	2008	2013	2020
Andalucía	17.55%	44.72%	26.22%
Aragón	16.84%	45.65%	27.26%
Asturias, Principado de	7.34%	47.37%	27.26%
Balears, Illes	10.10%	40.02%	35.64%
Canarias	1.96%	29.56%	16.67%
Cantabria	-1.02%	35.66%	19.95%
Castilla - La Mancha	17.61%	56.88%	33.95%
Castilla y León	28.02%	65.38%	42.77%
Cataluña	2.35%	50.10%	16.73%
Comunitat Valenciana	1.33%	47.31%	24.38%
Extremadura	61.00%	72.82%	55.64%
Galicia	15.48%	50.14%	31.23%
Madrid, Comunidad de	3.92%	46.93%	23.83%
Murcia, Región de	-5.57%	41.61%	22.18%
Navarra, Comunidad Foral de	5.20%	23.36%	9.61%
País Vasco	5.38%	31.52%	10.64%
Rioja, La	-5.32%	45.39%	33.89%

[back](#)

*Source: Consejo General del Notariado (CGN), Idealista, authors' calculations. Positive values mean that asking prices are higher than transaction prices.*



# Extrapolation of individual reference year vs. OECD method

*Table 5: Average (root mean squared) revision in price level estimates over the sample due to a reference year update (backcasting/extrapolation method) or the addition of an additional year of data (state-space model) over 2008-2013*

Autonomous Communities	Backcasting/Extrapolation Method	State-Space Model
Andalucía	1.876%	0.003%
Aragón	3.239%	0.086%
Asturias, Principado de	4.018%	0.117%
Balears, Illes	1.914%	0.129%
Canarias	1.585%	0.054%
Cantabria	1.559%	0.088%
Castilla - La Mancha	1.537%	0.004%
Castilla y León	2.607%	0.122%
Cataluña	2.134%	0.025%
Comunitat Valenciana	2.504%	0.306%
Extremadura	3.085%	0.148%
Galicia	3.039%	0.084%
Madrid, Comunidad de	2.199%	0.008%
Murcia, Región de	2.001%	0.028%
Navarra, Comunidad Foral de	4.030%	0.106%
País Vasco	2.543%	0.119%
Rioja, La	5.968%	0.052%

[back](#)