

Introducción

Históricamente el estudio de la evolución de los precios de la vivienda se han realizado sobre índices de precio de compraventa. A pesar de que el régimen de tenencia principal de la primera vivienda sigue siendo en propiedad, la tenencia de la vivienda en alquiler está en alza, especialmente a causa de la última gran crisis inmobiliaria en España. En paralelo, por parte de los principales actores del mercado inmobiliario y la administración, existe un interés creciente en disponer de más información del mercado inmobiliario, y conocer el impacto objetivo de fenómenos como el del alquiler vacacional o la gentrificación en los precios. A pesar de los esfuerzos de la administración, aún no se dispone de información adecuada para el análisis en detalle de los mercados inmobiliarios del alquiler residencial al nivel de granularidad temporal y espacial adecuado, por ser muy agregado o publicarse la información mucho retraso.

Objetivos generales

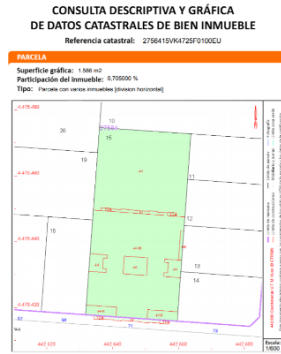
Este proyecto pretende desarrollar una metodología para la construcción de índices del precio de la vivienda con un alto nivel de desagregación funcional, por características del inmueble y por geografía. Un segundo objetivo es producir un índice lo más actualizado posible, produciendo idealmente el índice de precios mensualmente con datos del mes anterior y previos. El tercer objetivo es conseguir un índice estable en el tiempo, ajustado por calidad y sobre el que no afecten los sesgos procedentes del dato de un portal inmobiliario.

Datos

Hemos utilizado una muestra mensual detallada de anuncios del portal idealista desde 2011 a 2019, esta muestra está acotada a inmuebles de uso residencial en alquiler dentro de la Comunidad de Madrid. Junto a ella se han utilizado una explotación microdatos de alquiler a nivel de sección censal procedentes del censo de población y viviendas desarrollado por el INE en 2011. Por otro lado, se ha utilizado la información de gasto en alquiler de la encuesta de presupuestos familiares también desarrollada por el INE.

Dado que los datos del portal proceden de particulares y agencias, se ha realizado un ejercicio de eliminación de errores, valores atípicos y deduplicación (esta cuestión es importante dado que trabajaremos con inmuebles no con anuncios). Se han enriquecido de la información utilizando numerosas fuentes públicas como el catastro, Open Street Maps y el padrón continuo.

Otro aspecto clave en la investigación es la atribución de características catastrales a los anuncios del portal, a través de un proceso de cruce espacial y por dirección conseguimos imputar a cada inmueble en oferta superficie útil, año de construcción, calidad constructiva y otros aspectos morfológicos de la finca. Todos estas variables son necesarias para poder cruzar, correctamente y por características, la oferta (datos del portal) con los datos oficiales del alquiler.

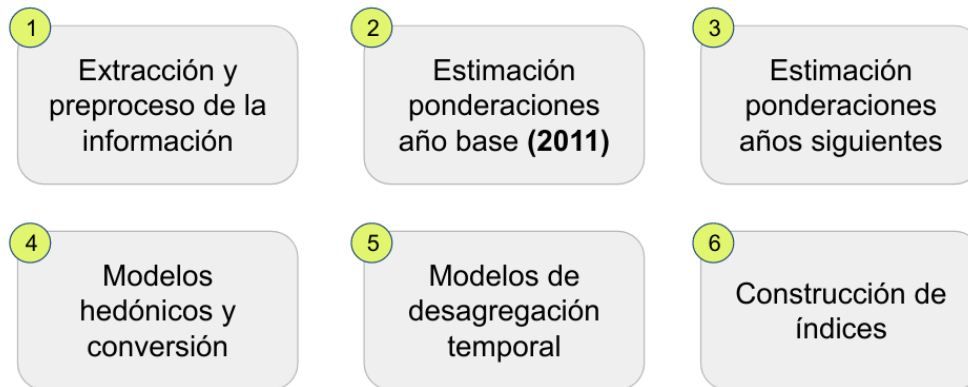


Atributos asignados:

- Superficie útil
- Año de construcción
- Calidad constructiva
- Número de inmuebles en catastro

Metodología

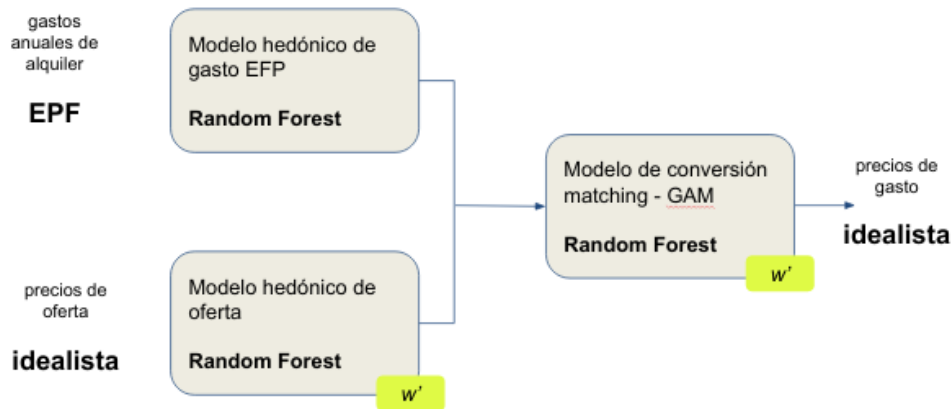
La metodología desarrollada parte del dato idealista como un dato observado sobre la población de oferta que debemos tratar para poder trasladar a la población de hogares en alquiler. Disponemos del dato del colectivo en alquiler para 2011 y el dato publicado en oferta en idealista, para poder estimar las ponderaciones de cada elemento de la muestra realizamos un proceso de calibración de los pesos originales de la muestra en base a las sumas de los distintos estratos de la población del censo.



En la estimación de las ponderaciones definimos distintas estratificaciones, en base a los atributos comunes en ambas muestras: número de habitaciones, superficie útil, tipo de municipio, antigüedad, anejos y ascensor.

Para los años 2011 a 2019 realizamos una doble calibración, la primera en base al censo y después un ajuste del censo en base a la evolución del gasto a través de la encuesta de presupuestos familiares. En este caso buscamos calibrar los pesos originales del censo usando los atributos comunes de la encuesta e idealista, que son similares a los del censo: renta media por persona de la zona, tipo de zona, nivel adquisitivo medio de los hogares, tipo de edificio, densidad de población, número de habitaciones y tamaño del municipio.

Una vez contamos con los pesos poblacionales se construyen varios modelos hedónicos, los primeros que estiman el precio del alquiler y de oferta para los atributos comunes de oferta y EFP.



Dado que las variables utilizadas para modelar son limitadas, se construirá un modelo adicional con todas las variables del inmueble que nos permita corregir el sesgo del modelo generado. Para ello construimos un modelo basado en random forest que capture tanto las interacciones de los elementos estructurales de la vivienda como los efectos de la localización. Para ellos se crea un modelo ensamblado que combina la utilidad de la localización y otro basado en atributos y características del mercado.

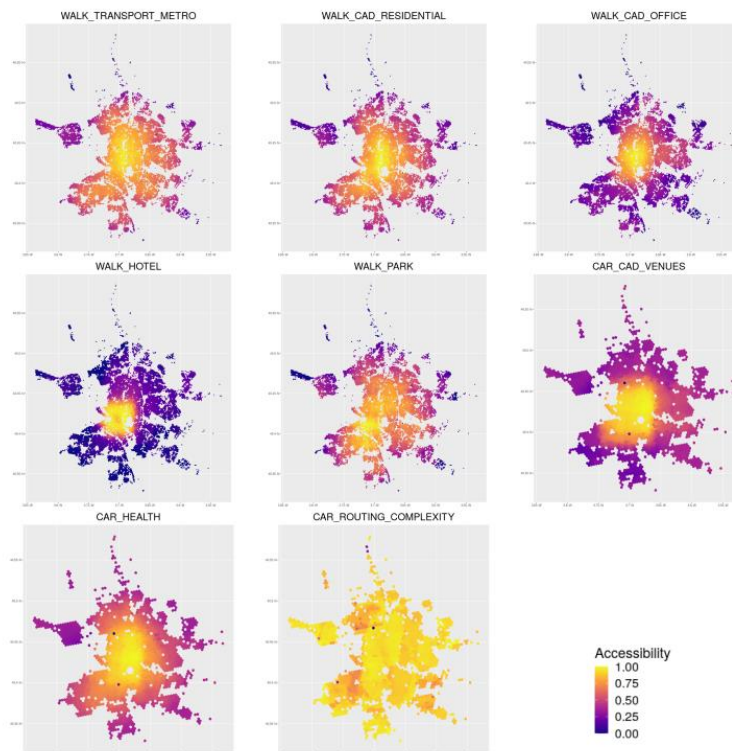
Como elemento original de la investigación introducimos en el modelo las características de la demanda de las zonas (nivel medio de visitas a la ficha del anuncio y número de mensajes enviados al propietario). Este modelo incorpora las interacciones del tiempo como dummies de tiempo:

(1)

$$p^t = \beta_0 + \sum \delta \cdot D_{nk}^t + \sum \beta \cdot S_{nk}^t + \sum \gamma \cdot Loc_{nk}^t + \sum \mu^t M_{nk} + \epsilon$$

Feature	Motivation
Structural	Structural features capture the contribution of property physical features such as square meters, room number or state of conservation
Market features	Incorporates the main offer/demand dynamics of the market where the property is located
Location	Explain the contribution of location in the land price of an asset, includes neighborhood characteristics, accessibility indices, and other geographical features
Time dummies	Captures adjustment of price over time, seasonality and trend effects

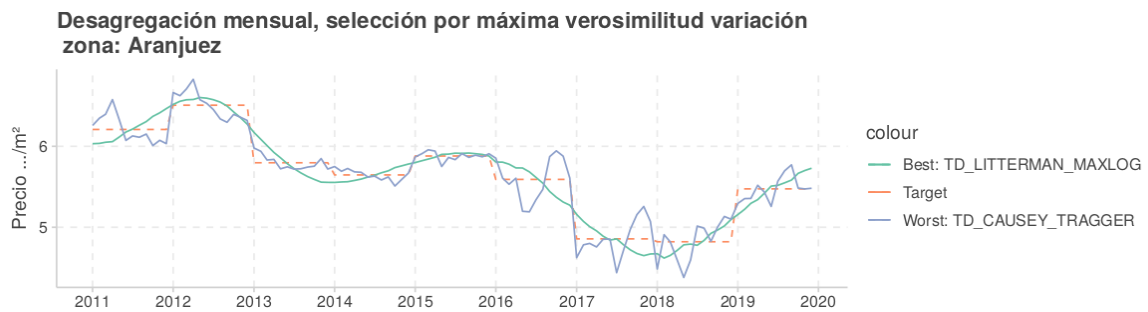
Para el modelo de utilidad se ha desarrollado un algoritmo de generación de características espaciales en base a información catastral y de openstreetmaps, usando modelos gravitatorios en base a los tiempos de desplazamiento a pie y en transporte privado.



El modelo que utilizamos para realizar la conversión se realiza sobre precios medios anuales y dado nuestro objetivo es disponer de un índice mensual, para ello realizamos un proceso de desagregación temporal seleccionando el método más adecuado para cada zona. Es necesario aplicar esta técnica puesto que la serie de alta frecuencia debe preservar las características de la de baja frecuencia, el precio medio de la serie mensual debe coincidir con el valor anual de nuestro modelo original. En la siguiente figura vemos que aplicando el modelo de conversión a un precio mensual estimado por un modelo hedónico no respeta la restricción de la media:



Utilizamos varios algoritmos de desagregación temporal, paramétricos y no paramétricos en cada uno de los estratos usando un método de selección basado en la verosimilitud de las características de la serie a ser una serie mensual. Principalmente buscamos series que no cuenten con variaciones extremas entre meses o variaciones excesivas en cada cambio de año.

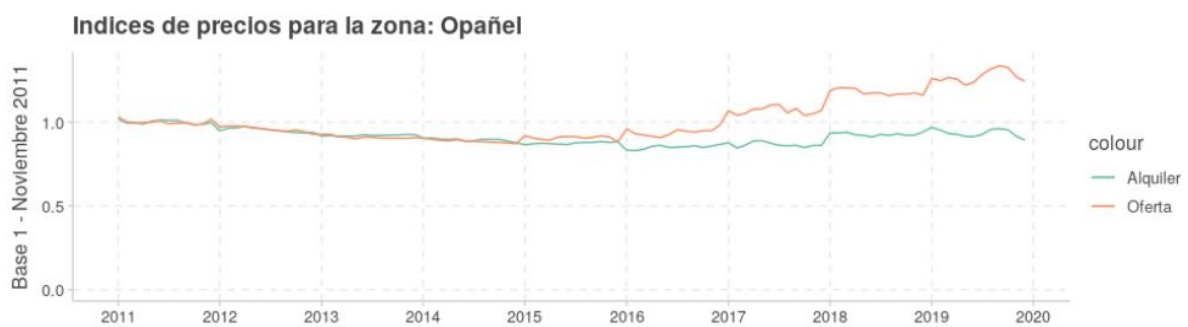


Finalmente construimos una serie de índices de precios encadenados de Laspeyres desagregable por varias dimensiones: Madrid o provincia, tamaño del municipio, número de habitaciones, barrio o municipio.

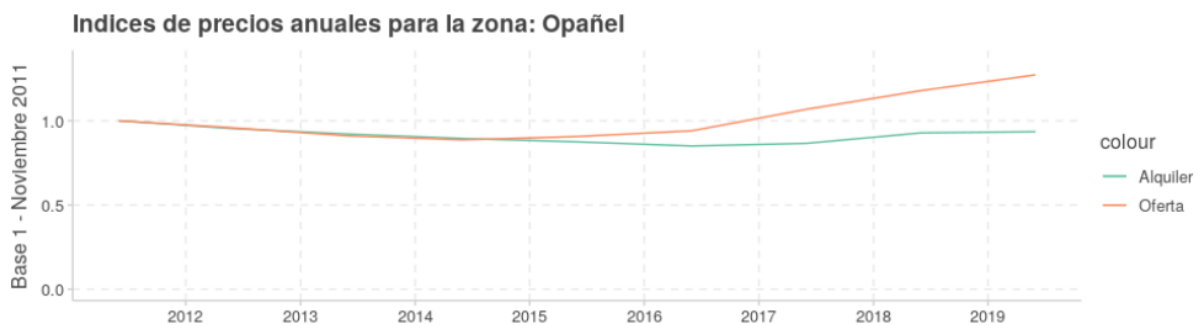
Conclusiones y futuros trabajos

Este método proporciona un índice de gasto del alquiler que respeta los datos recogidos por los organismos estadísticos pero que nos permite un alto nivel de desagregación, también permite generar datos de alta frecuencia y recientes en base a datos del portal. Reduce los sesgos habituales de la oferta y ofrece una herramienta analítica para estudiar en el tiempo la evolución y la relación de las magnitudes de oferta y alquiler.

Las series obtenidas guardan coherencia con las series oficiales y permiten analizar de una forma mínima sesgada el comportamiento de los mercados de compra y de alquiler. Tanto en su versión anual como su versión desagregada, vemos cómo el índice de precios de oferta actúa como indicador adelantado de la del alquiler. Hasta 2014 con una ligera caída, mientras que a partir de mitad de 2016 con incrementos en el precio del alquiler.



En los análisis preliminares hemos podido relacionar de forma positiva la velocidad de incorporación de los cambios de oferta en el alquiler con indicadores de demanda como el número de contactos al mes en la web por inmueble.



Como continuación a este trabajo, pretendemos continuar con un indicador a nivel nacional, enriqueciendo las fuentes de información de la oferta. Se profundizará en la desagregación infra municipal y en la estimación de la relación de los precios de oferta y de gasto, y su relación con el ciclo económico y las características de cada zona. Esto dará lugar a la posibilidad de segmentar el espacio en submercados de la vivienda de alquiler por criterios morfológicos y de mercado