# Impacto del cambio climático sobre el mercado de bienes raíces: análisis contrafáctico mediante el método de control sintético, modelos no lineales y machine learning.

Bella,J. Bustos Barton,A. Fernández Bonilla,T.

Licenciatura en Economia

*Universidad Torcuato Di Tella*

Agosto 2021

# Introducción


Figura: Territorio en riesgo de inundación para el año 2050.

# Introducción

- Londres se encuentra en el tercer puesto entre las ciudades europeas más afectadas por el calentamiento global en cuanto a las posibles perdidas inmobiliarias para 2050.
- 200 millones con un aumento de 1,5$^{o}$c y 600 millones con 4$^{o}$c, pasando a ser la mas afectada de Europa
- *Neglected no more: Housing markets, mortage lending, and sea level rise.* (Keys, B. Mulder, P. 2020)
- *Does Climate Change Affect Real Estate Prices? Only If You Believe In It*. (Baldauf *et al.* 2020)

# Desarrollo Teórico

La demanda agregada de tierra es:

$$D_t = \sum_{i \in I_t} T_{i,t}^d \tag{1}$$

Donde $I_t$ representa el conjunto de demandantes en el mercado de tierras en el momento t, y $T_i$, $t_d$ es la demanda individual de tierras del individuo i en el momento t.

La utilidad de cada individuo es:

$$D_t = \sum_{t=0}^{\infty} \beta^t u_t \tag{2}$$

Donde la utilidad por unidad de tiempo es: $u_t = T_t O_t$
T representa la cantidad alquilada de unidades de tierra, y O representa la cantidad consumida de los bienes restantes.

La oferta agregada de tierras es:

$$S_t = K \tag{3}$$

Donde K es una constante.

# Desarrollo Teórico

Suponemos que el parámetro de descuento genera un ahorro nulo en equilibrio.

Normalizamos el vector de precios de los otros bienes, y suponemos un ingreso constante sin capacidad de ahorro:

$$p_t^0 = 1$$
$$\forall t \in N$$

Por lo tanto la restricción presupuestaria flujo de los individuos es:

$$alq_t T_t + O_t = M$$

Donde M es una constante, y $alq_t$ es el precio de alquiler de la tierra.

La población crece a una tasa exponencial, por lo tanto:

$$cardinal(I_t) = L^t$$

Donde L es una constante.

Para simplificar, suponemos que $\frac{M}{2K} = 1$

Se cumplen los argumentos de no arbitraje.

# Desarrollo Teórico - Equilibrio parcial

## Equilibrio Parcial

Un equilibrio competitivo es un conjunto de precios $\{alq_t | t \in N\}$ tal que $S_t = D_t \ \forall t \in N$, donde los individuos demandan maximizando su utilidad dados los precios.

Problema del agente:

$$\max_{\{T_t, O_t\}} U \ sujeto \ a \ \ alq_t T_t + O_t + ahorros_{t+1} = M + ahorros_t \ \forall \ t \in \mathbb{N}$$

Por conclusión de los supuestos antes mencionados, solo debemos maximizar la utilidad para cada momento t:

$$\max_{\{T_t, O_t\}} u_t \ sujeto \ a \quad alq_t T_t + O_t = M$$

# Desarrollo Teórico - Equilibrio parcial

Por lo tanto, dado un momento t, el lagrangiano del problema es:

$$L_t = T_t O_t - \lambda_t (alq_t T_t + O_t - M)$$

Se obtiene la siguiente demanda marshalliana de la tierra:

$$T_t^d = \frac{M}{2alq_t}$$

Como todos los individuos son iguales, la demanda agregada es:

$$D_t = \sum_{i \in I_t} T_{i,t}^d$$

$$D_t = cardinal(I_t) T_t^d$$

$$D_t = L_t \frac{M}{2alq_t}$$

## Desarrollo Teórico - Equilibrio parcial

En el equilibrio, se cumple que St=Dt, por lo tanto:

$$S_t = D_t$$
$$K = L^t \frac{M}{2alq_t}$$
$$alq_t = L^t \frac{M}{2K}$$

Como $\frac{M}{2K} = 1$, vemos que:

$$alq_t = L^t$$

# Desarrollo Teórico - Equilibrio parcial

En el equilibrio, se cumple que St=Dt, por lo tanto:

$$
\begin{aligned}
S_t &= D_t \\
K &= L^t \frac{M}{2alq_t} \\
alq_t &= L^t \frac{M}{2K}
\end{aligned}
$$

Como $\frac{M}{2K} = 1$, vemos que:

$$alq_t = L_t$$

# Desarrollo Teórico - Precio de los inmuebles

Los inmuebles pueden ser considerados como un activo que paga dividendos en forma de alquiler, por lo tanto, bajo el supuesto de no arbitraje, sabemos que el precio del bien está dado por la sumatoria de todos los alquileres futuros descontados por la tasa de interés:

$$precio_t = \sum_{j=0}^{\infty} \frac{alq_{t+j}}{(1+i)^j}$$

Donde i es la tasa de interés.

# Desarrollo Teórico - Desenlace teórico y su contrafactual

Sea un inmueble w, un bien que sigue los supuestos expresados anteriormente, su precio va a ser:

$$precio_t = \sum_{j=0}^{\infty} \frac{alq_{t+j}}{(1+i)^j}$$

Sin embargo, en t=0 obtenemos la información de que en el momento F, el inmueble se ve afectado por las inundaciones y no será de utilidad para ningún individuo, por lo tanto, el alquiler que paga el inmueble a partir de F será nulo:

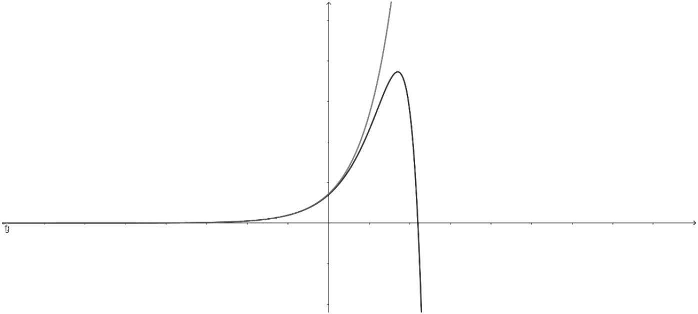
$$alq_t = 0$$
$$\forall t \geq F$$

Por lo tanto tenemos las siguientes expresiones para los precios:

$$precio_t = \frac{L^t}{1 - \frac{L}{1+i}}$$

$$precio_t^{info\ \ nueva} = (1+i)L_t \frac{1 - (\frac{L}{1+i})^{F+1-t}}{1+i-L}$$

A continuación, graficamos estas dos ecuaciones:



# Desarrollo Teórico - Desenlace teórico y su contrafactual

Nuestro trabajo va a consistir en estimar el precio contrafáctico donde el inmueble nunca se inunda (dado por la curva gris) y luego comparar su diferencia con el precio real de la casa que se informa se va a inundar (curva negra), por lo tanto, en este caso:

$$precio\ real_t = \begin{cases} \frac{L^t}{1-\frac{L}{1+i}}\ si\ \ t<0 \\ \\ (1+i)\, L^t \frac{1-\left(\frac{L}{1+i}\right)^{F+1-t}}{1+i-L}\ si\ \ t\ \ \geq 0 \end{cases}$$

$$precio\ \ contraf\acute{a}ctico_t = \frac{L^t}{1-\frac{L}{1+i}}$$

# Desarrollo Teórico - Desenlace teórico y su contrafactual

Si encontramos una diferencia significativa entre el precio contrafáctico y su precio real podemos inferir que los agentes están incorporando la nueva información y descontando del precio de adquisición la inundación de los inmuebles en el futuro. Tomando el logaritmo de la diferencia entre el precio real y el contrafáctico, obtenemos:

$$\ln\left(precio\ contrafáctico_t - precio\ real_t\right) = \ln\left(\frac{HL^{F+1}}{(1+i-L)(1+i)^F}\right) + \ln(1+i)t$$

donde:

$$\frac{M}{2K} = H$$

Esto quiere decir que si calculamos una regresión lineal del logaritmo de la diferencia contra el tiempo, el beta de la regresión va a ser:

$$\beta = \ln(1+i) \approx i$$

# Desarrollo Teórico - Desenlace teórico y su contrafactual

## Resultado

$$\ln\left(precio\ contraf\acute{a}ctico_t - precio\ real_t\right) = \ln\left(\frac{HL^{F+1}}{(1+i-L)(1+i)^F}\right) + \ln(1+i)t$$

Una vez que tengamos nuestra estimación del contrafáctico, vamos a calcular una regresión sobre el logaritmo de la diferencia, y vamos a tomar a la ordenada al origen como el salto discreto por la actualización de la información, y la pendiente como la tasa con la que los agentes descuentan la inundación.

# Control Sintetico

- Permite evaluar cuantitativamente los efectos de un tratamiento especıfico sobre una unidad tratada mediante la creación de una unidad "sintética" contrafactual que no ha sido tratada
- Esta unidad sintética es comparada con la unidad tratada para capturar el efecto
- $$\hat{\alpha_{1t}} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

# Control Sintetico

El método de control sintético requiere de un punto de corte donde comienzan a compararse la unidad tratada y la unidad placebo sintética. Hasta dicho punto, la trayectoria de la unidad tratada funciona como referencia para poder construir y fittear la unidad sintética.


Figura: Interés a lo largo del tiempo del 'tema' Cambio Climático en Google Trends con punto máximo en 2007.

# Control Sintetico

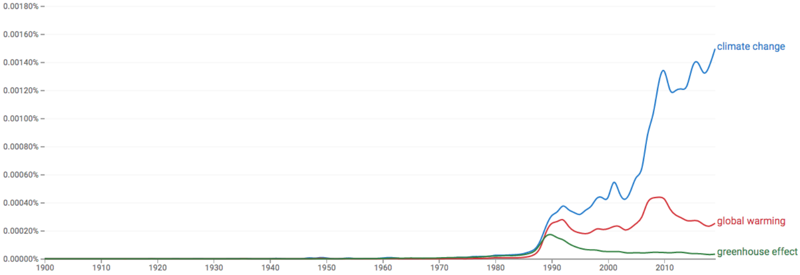



Figura: Google Book Ngram Viewer

# Machine Learning

El aprendizaje supervisado consiste en programar computadoras para optimizar cierto criterio de performance usando datos o experiencia pasada como materia prima.

De esta manera, es posible encontrar patrones y realizar predicciones que bajo métodos lineales sería imposible detectar.

Hal Varian, economista jefe de Google: *"... my advice to grad (Economics) these days is 'go to the computer science department and take a course in machine learning'"*

En nuestro caso, utilizaremos tres herramientas diferentes para construir nuestros modelos:

- Modelo K-vecinos
- Método de Kernel
- Redes Neuronales Clásicas

# Machine Learning

1. Machine Learning + Control Sintetico
   - Extrapolación
   - Linealidad
2. Complejidad Potencial
   - Bias-Variance trade off
   - Replicabilidad
3. Comparacion de Modelos

# Machine Learning + Control Sintetico

**Extrapolación**:

Desventajas

- $bias = 0$
- $weights_i \in [0, 1] \;\; \forall i \in donor \;\; pool$
- $\displaystyle\sum_{i \in donor \;\; pool} weights_i = 1$

Ventajas

- Filtrar outliers
- Regularización

**Linealidad:**

- Interpretabilidad
- $y_{contrafactico} = \displaystyle\sum_{i \in donor \;\; pool} w_i y_i + v$

Ecuaciones:

- $y_{contrafactico} = f\left(\{y_i | i \in donor\ \ pool\}\right) + v \qquad (1)$
- $\hat{y}_{contrafactico} = \hat{f}\left(\{y_i | i \in donor\ \ pool\}\right) \qquad (2)$
- $\alpha = y_{contrafactico} - y_{real} \qquad (3)$
- $\hat{\alpha} = \hat{y}_{contrafactico} - y_{real} \qquad (4)$

$$sesgo = \mathbb{E}\left\{\hat{\alpha} - \alpha\right\} = \mathbb{E}\left\{\hat{f}\right\} - f = BIAS\left\{\hat{f}\right\}$$

# Machine Learning

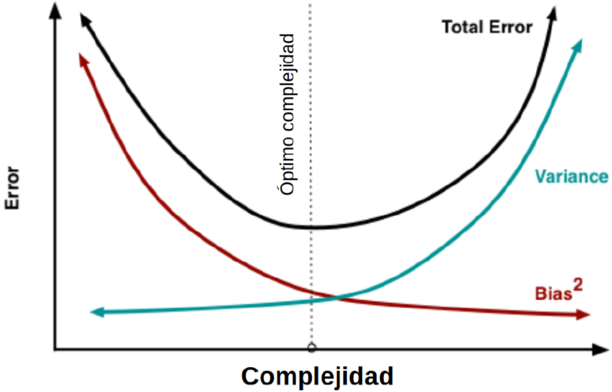
Ecuaciones:

- $y_{contrafactico} = f\left(\{y_i | i \in donor\ \ pool\}\right) + v \qquad (1)$
- $\hat{y}_{contrafactico} = \hat{f}\left(\{y_i | i \in donor\ \ pool\}\right) \qquad (2)$
- $\alpha = y_{contrafactico} - y_{real} \qquad (3)$
- $\hat{\alpha} = \hat{y}_{contrafactico} - y_{real} \qquad (4)$

## Resultado

$$\mathbb{E}\left\{(\hat{\alpha} - \alpha)^2\right\} = BIAS\left\{\hat{f}\right\}^2 + Var\left\{\hat{f}\right\} + \sigma^2$$

Replicabilidad:



Comparación de Modelos:

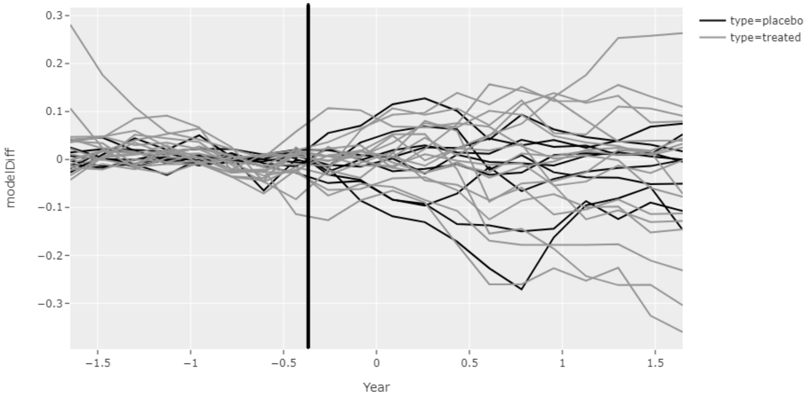



Figura: Control Sintetico - Error: 0.005157038599079725

# Machine Learning

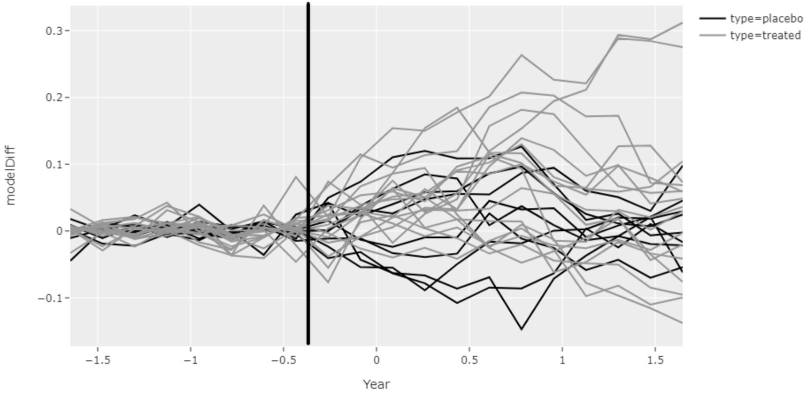



Figura: Lazy Ridge - Error: 0.0029926219842486335

# Machine Learning

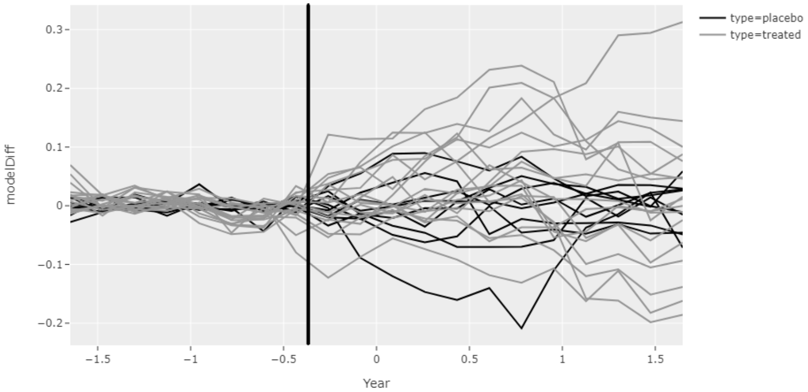
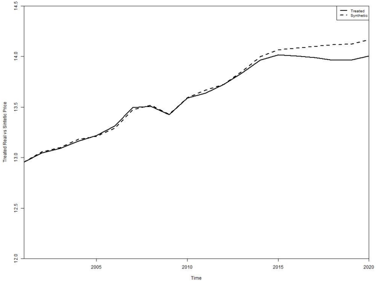
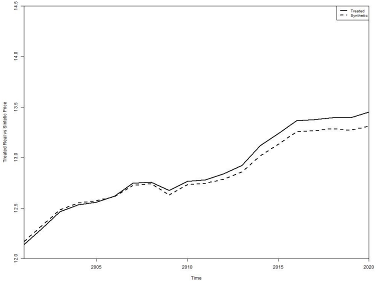



Figura: Kernel - Error: 0.0025659922177516

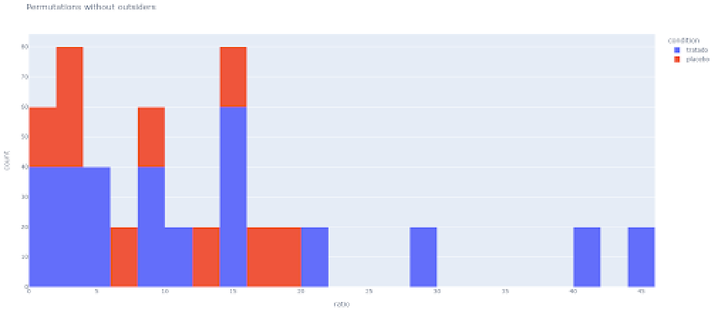
Resultados Hammersmith and Fulham y Lewisham:



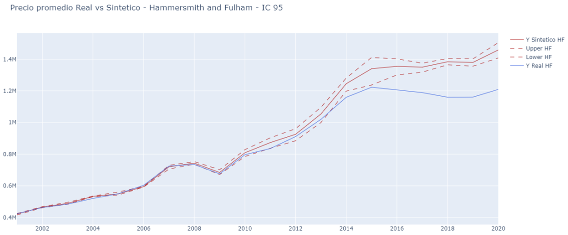
# Resultados

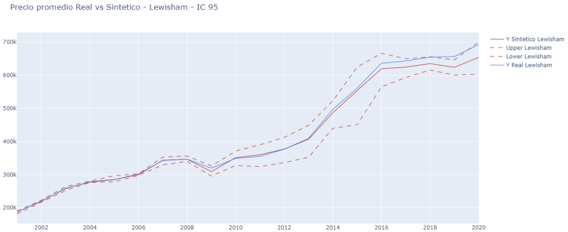
Interpretación estandar:



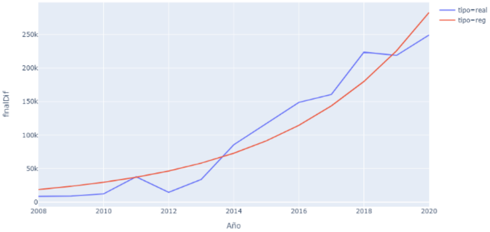
# Resultados

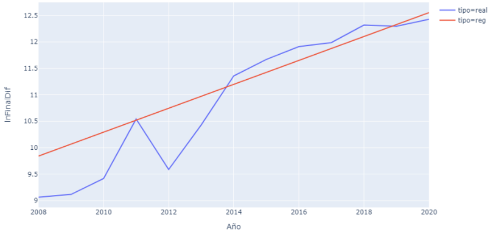
## Intervalo de confianza:



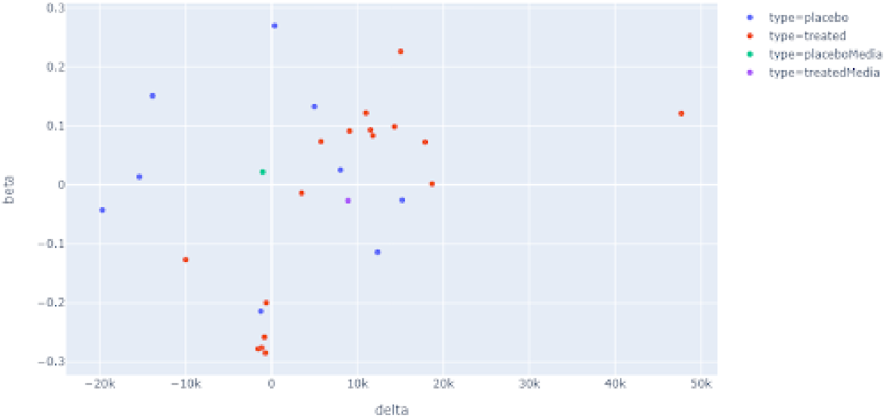
Heterogeneidad y tendencia:



# Resultados

Distribución de tendencias:



Test de Hipótesis:

$$H_0 : \vec{v}_{tratados} = \vec{v}_{placebos}$$
$$H_1 : \neg(\vec{v}_{tratados} < \vec{v}_{placebos})$$

Donde:

$$\vec{v}_{tratados} = \frac{1}{cardinal(tratados)} \sum_{i \in tratados} \vec{v}_i$$
$$\vec{v}_{placebos} = \frac{1}{cardinal(donor\ pool)} \sum_{i \in donor\ pool} \vec{v}_i$$

# Resultados

Distribución bajo hipótesis nula, p-valor menor a 0.01: