UNIVERSIDAD TORCUATO DI TELLA

IMPACTO DEL CAMBIO CLIMÁTICO SOBRE EL MERCADO DE BIENES RAÍCES: ANÁLISIS CONTRAFÁCTICO MEDIANTE EL MÉTODO DE CONTROL SINTÉTICO, MODELOS NO LINEALES Y MACHINE LEARNING

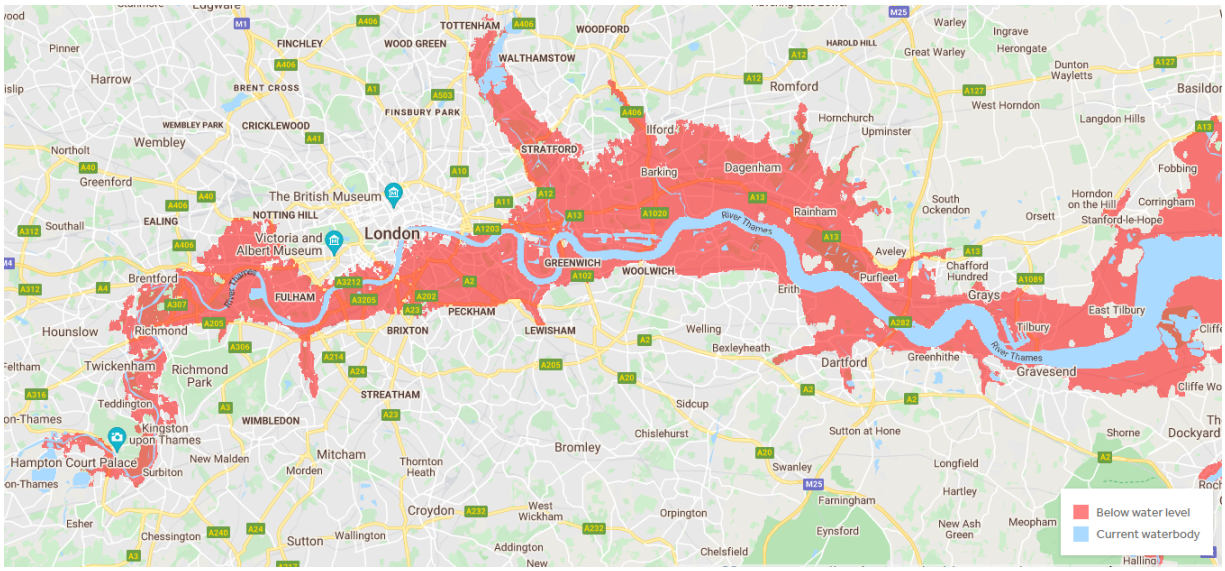
Jonatan Bella, Tomás Fernández Bonilla, Agustín Bustos Barton
Supervisor de Tesis: Emilio Espino
Departamento de Economía
Licenciatura en Economía
Agosto 2021

Índice

Resumen

Este trabajo estudia los efectos de la creencia en el cambio climático sobre el precio de los inmuebles. Un sencillo modelo de *pricing* de activos nos permite comparar la trayectoria de los precios en dos circunstancias distintas: Por un lado, tendremos la curva que describe al precio de un inmueble si los agentes son informados en un momento determinado que el activo valdrá cero en un futuro específico; por otro, la trayectoria de dicho activo si nunca recibe tal información. Esto sentará las bases para un análisis contrafáctico mediante el modelo de control sintético. Tomaremos la ciudad de Londres (rica en datos y en gran riesgo de daños considerables), para realizar un análisis econométrico. Compararemos la trayectoria de los precios en distintos territorios en situación de riesgo y territorios fuera de riesgo (es decir, entre distintos municipios londinenses llamados $boroughs$). Contaremos con datos referidos a los precios de los inmuebles según las transacciones realizadas y datos que tipifican a los distintos municipios (distribución demográfica, distribución del ingreso, etc.). A continuación, generamos una unidad sintética para realizar un análisis contrafáctico de cada municipio en situación de riesgo. Por último, utilizaremos métodos no lineales y *machine learning* para optimizar la unidad sintética. Los resultados encontrados son heterogéneos, indicando las presencia de dos subgrupos de municipios afectados. Mientras algunos responden a nuestra hipótesis, ajustando la tasa de aumento de precios de los inmuebles, otros municipios afectados no presentan cambios significantes. Si bien los resultados podrían evidenciar el reticente esceptisísmo que perdura en cierto sector de la sociedad frente a las predicciones de los expertos climáticos, se debe tener en cuenta que esta falta de reacción pueda deberse a otros tipos de información y creencias, como la posibilidad de inversión en infrastructura preventiva por parte del sector público.

1 Aspectos Generales



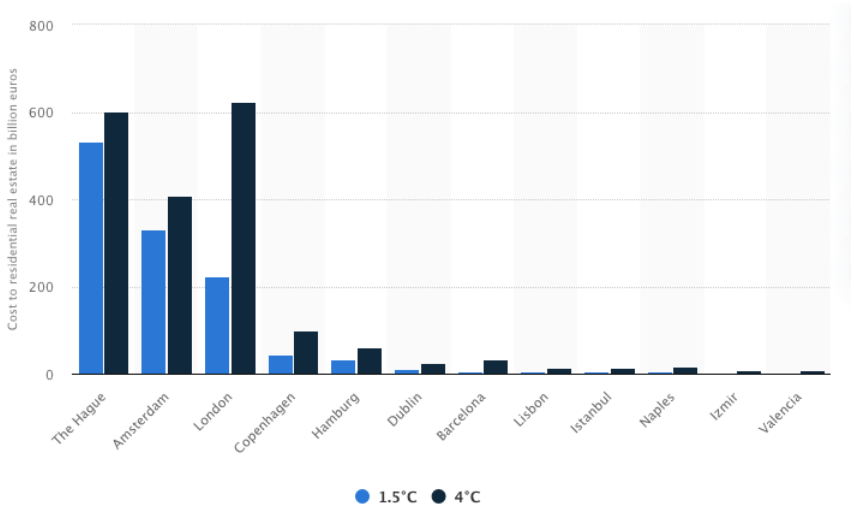
Territorio en riesgo de inundación para el año 2050 [1]

El Cambio Climático y sus consecuencias económicas parecen ser inevitables: Aún suprimiendo drásticamente las emisiones de carbono en forma inmediata, los modelos predictivos auspician un aumento de 0.5 cm del nivel del mar. De no tomarse medidas, podría alcanzar los 2 cm.[2]En el marco de un escenario semejante, reportes recientes proyectan un eventual desborde del río Támesis, poniendo en riesgo de destrucción total a miles de propiedades residenciales cuyo valor agregado de mercado alcanza los 224 billones de euros. [3]Esto posiciona a Londres como la tercera ciudad Europea cuyo mercado de bienes raíces residenciales se vería más afectado económicamente por aumentos en el nivel del mar.

[1] *Costal risk screening tool*. Costal Climate Central [En linea]: `https://coastal.climatecentral.org/map/11/-0.0516/51.5343/?theme=sea_level_rise&map_type=coastal_dem_comparison&basemap=roadmap&contiguous=true&elevation_model=coastal_dem&forecast_year=2050&pathway=rcp45&percentile=p50&refresh=true&return_level=return_level_1&slr_model=kopp_2014`

[2] *New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding.*Nature [En linea]: `https://www.nature.com/articles/s41467-019-12808-z.pdf`

[3] *Cost of rising sea-levels to coastal cities in Europe 2020, by temperature rise*. Statista [En linea]: `https://www.statista.com/statistics/1066944/estimated-cost-of-rising-sea-levels-to-select-coastal-cities-in-europe/`



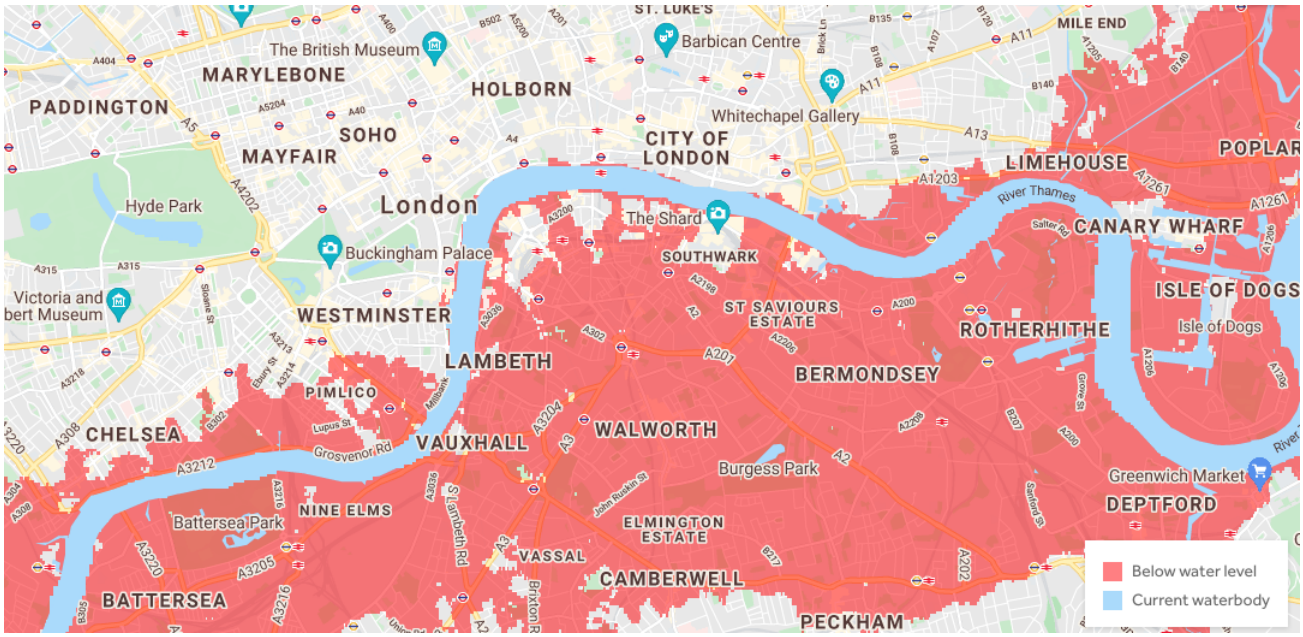
Cost of rising sea-levels to coastal cities in Europe 2020, by temperature rise. "at the current rate of warming - 0.2C per decade - global warming will reach 1.5C between 2030 and 2052" [4]

Desde 1980, las inundaciones en Londres se han cuadriplicado, y desde 2004 se han duplicado[5]. Debemos considerar que las inundaciones urbanas no sólo son provocadas por el aumento del nivel del mar y por cambios en los patrones de precipitación; se deben a un variado conjunto de razones, incluyendo el crecimiento urbano y el aumento de zonas pavimentadas que impiden un drenaje natural. [Ashley *et al.* 2005][6]. El crecimiento de la población y el aumento de la densidad urbana constituyen un aumento en tales zonas. Si las expectativas racionales de los individuos consideran al cambio climático como una amenaza real, deberíamos esperar que los precios de los inmuebles ubicados dentro del perímetro en situación de riesgo se vean afectados de manera negativa. Es decir, tomando la hipótesis de mercados eficientes de Eugene Fama, el mercado de bienes raíces debiera incorporar nueva información con respecto a la valuación de terrenos que en el futuro estarán por debajo del nivel del mar y por lo tanto tendrán una valuación nula. Nos preguntamos cómo influyeron en el mercado de bienes raíces de Londres las pronosticadas consecuencias del Cambio Climático. ¿Ha tomado en cuenta el mercado los pronósticos de los expertos?

[4]Cost of rising sea-levels to coastal cities in Europe 2020, by temperature rise `https://www.statista.com/statistics/1066944/estimated-cost-of-rising-sea-levels-to-select-coastal-cities-in-europe/`

[5]*New data confirm increased frequency of extreme weather events: European national science academies urge further action on climate change adaptation.* ScienceDaily [En linea]: `https://www.msci.com/www/blog-posts/underwater-assets-real-estate/01593224766`

[6]R.M. Ashley*, D.J. Balmforth**, A.J. Saul* and J.D. Blanskby* *Flooding in the future – predicting climate change, risks and responses in urban areas.*



Simulación predictiva para la ciudad de Londres en 2050[7]

1.1 El Cambio Climático como información y como creencia

Una encuesta realizada en Octubre del 2020 por el Consejo de Londres ha revelado que el 82% de la población londinense dice estar preocupado por las consecuencias del cambio climático y un 52% afirma que el cambio climático afecta la forma en la que toman decisiones.[8] La problemática en cuestión ha comenzado a debatirse abiertamente en el mercado londinense.[9]

Una brevísima historia del estudio del cambio climático debe comenzar con las primeras nociones del efecto invernadero, figuradas por Joseph Fourier en 1820, quien calculó que la radiación del Sol no era lo suficientemente fuerte como para justificar las temperaturas que en aquel entonces se experimentaban cotidianamente, llegando a la conclusión de que debía tratarse de un efecto atmosférico. En 1840 el naturalista suizo Louis Aggazis publica Estudio sobre los glaciares y a partir de entonces comienza a hallar evidencia a favor de la existencia de un periodo prehistórico de bajas temperaturas al que llamó era de hielo; nace así la historia natural de los cambios climáticos. Dos décadas más tarde, John Tyndall probaría que el efecto invernadero era producido por el vapor de agua y por el dióxido de carbono; esto condujo al nobel sueco Svante Arrhenius a argumentar por primera vez que el avance de la industrialización podría afectar las temperaturas del planeta. [10] En abril de 1938, la publicación de *La producción artificial de dióxido de carbono y su influencia sobre la temperatura* por Guy

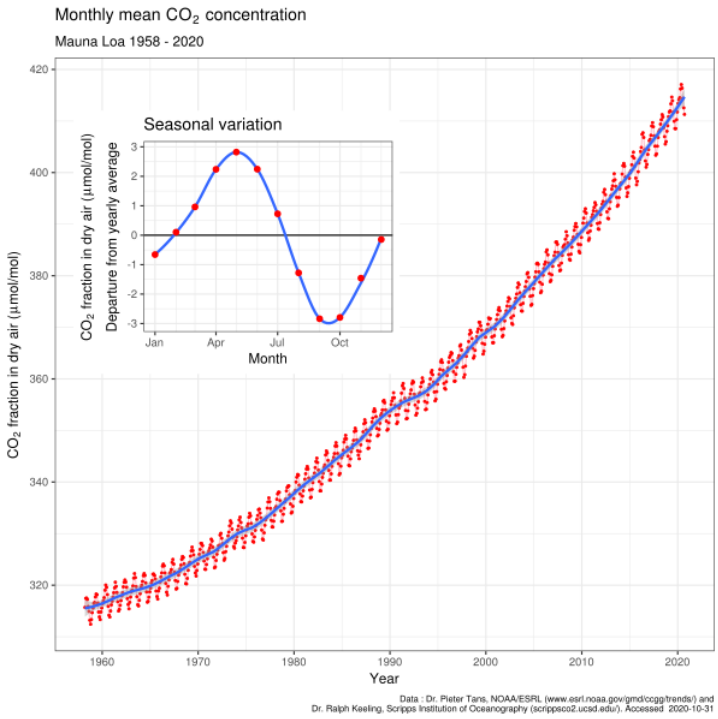
[7] *Land projected to be below annual flood level in 2050*, ibidem.

[8] *What do Londoners' think about climate change?*. Consejo de Londres [En linea]: `https://www.londoncouncils.gov.uk/our-key-themes/environment/climate-change`

[9] FALTA ESTE LINK, falta desarrollar.

[10] *Cómo descubrimos el problema del clima en*. Policy Forum [En linea]: `https://www.policyforum.net/how-we-discovered-the-climate-problem/`

Steward Callendar, daría comienzo a la historia del cambio climático producido por la actividad humana. Lo hace midiendo aumentos en la temperatura de los últimos años y midiendo el aumento en las emisiones de carbono, y el dióxido de carbono contenido en la atmósfera (Calendar, 1938)[11]. En 1961, Charles David Keeling documentó un aumento estable en la acumulación de dióxido de carbono en la atmósfera, diagramando en una curva que hoy lleva su nombre:

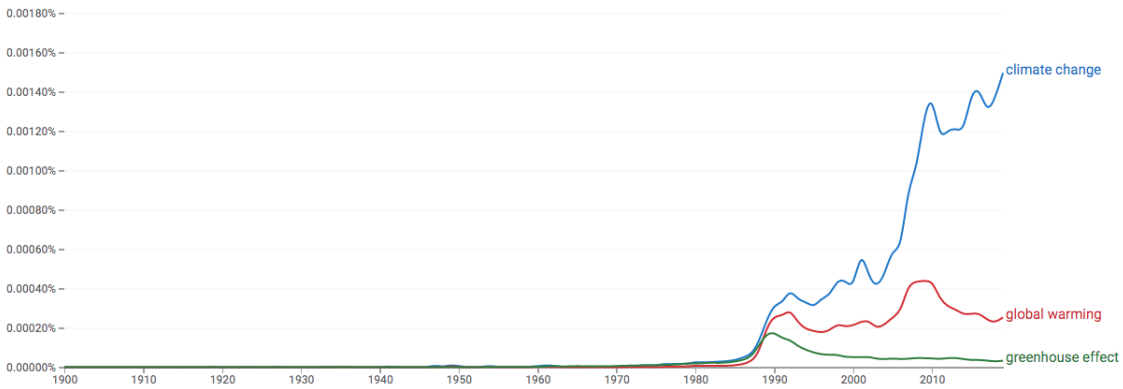


Curva de Keeling con datos del observatorio Mauna Loa en Hawaii.

En la década de los noventas, el debate en torno al agujero en la capa de ozono sitúa al cuidado del medio ambiente en un lugar central de la opinión pública; es en esta década que se acuña el término *eco-warrior* y *eco-terrorist* a para tipificar a los activistas medioambientales que comienzan a tomar posturas cada vez más radicales.[12] Sin embargo, no es hasta el estreno de la película Una Verdad Incómoda del ex vicepresidente americano Al Gore en el año 2001, que el cambio climático (hasta entonces confundido con el término calentamiento global) comienza a ocupar un lugar cada vez más dramático en la percepción popular del *mainstream*. A partir de entonces notamos un crecimiento acelerado de las menciones que se hacen en torno al cambio climático en diversas publicaciones, como podemos ver en los datos provistos por el *Ngram Viewer* de Google.

[11] Calendar, C.S. *The artificial production of carbon dioxide and its influence on temperature.*[En linea]: `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49706427503`

[12] *Eco-warrior*, Wikipedia [En linea]: `https://en.wikipedia.org/wiki/Eco-warrior`



Google Book Ngram Viewer[13]

Si bien hoy en día, la gran mayoría de las personas han oído hablar del cambio climático, la recepción de esta nueva información ha sido más bien heterogénea. Fácilmente podemos encontrarnos con alarmistas, realistas y escépticos. Últimamente se ha encontrado evidencia que sugiere que las encuestas tienden a subestimar el nivel de escepticismo climático.[14] También se ha encontrado que el público en general tiene problemas entendiendo las consecuencias del aumento del nivel del mar y su conexión con el cambio climático.[15]. Recientemente (Baldauf *et al.* 2020) [16] se ha analizado cómo el impacto en los precios de las residencias de Miami depende de las creencias de los ciudadanos; es decir, los autores estudian si los precios de los inmuebles reflejan las diversas creencias de los agentes. Con una encuesta previa que les permite distinguir sectores creyentes y sectores escépticos, encuentran que en los territorios donde los agentes confían en los expertos podemos encontrar una caída en los precios del m² de las zonas afectadas, mientras que en los territorios donde los agentes son escépticos frente a los pronósticos de los expertos, los precios no se han visto afectados. Vale subrayar que los autores encuentran una marcada correlación entre territorios que confían en el consenso científico climático y territorios donde el partido demócrata logra mayor cantidad de votos, como así los territorios con mayor cantidad de escépticos tienden a ser territorios donde el partido republicano logra una mayoría de votos. Esta variable político-cultural no se encuentra sola a la hora de afectar las creencias. El escepticismo también es heterogéneo entre las diversas generaciones. Mientras las juventudes tienden al alarmismo climático, representadas por Gretta Thumberg, sus progenitores tienden al escepticismo. En Estados Unidos, la encuestadora

[13]Google Book Ngram Viewer [En linea]:https://books.google.com/ngrams/graph?content=climate+change%2Cglobal+warming%2Cgreenhouse+effect&year_start=1900&year_end=2019&corpus=26&smoothing=0#

[14]Liam F. Beiser-McGrath,Thomas Bernauer. *Current surveys may underestimate climate change skepticism evidence from list experiments in Germany and the USA.* [En linea] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0251034

[15]Rebecca K. Priestley, Zoë Heine, Taciano L. *Milfont Public understanding of climate change-related sea-level rise*.[En linea]: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0254348

[16]Baldafur,M. Garlappi,L. Yannelis,C. *Does Climate Change Affect Real Estate Prices? Only If You Believe In It* The Review of Financial Studies, 2020. [En linea]:https://academic.oup.com/rfs/article-abstract/33/3/1256/5735306?redirectedFrom=fulltext

Gallup encontró que sólo el 53% de los mayores de 55 años estaba preocupado por el cambio climático; por otro lado, el 70% los encuestados entre 18 y 34 años de edad expresaron preocupación por la crisis venidera.[17] En los próximos años se espera la mayor transferencia de riquezas en la historia; a medida que los llamados *Millenials* (nacidos entre 1980 y 1994) hereden la riqueza de los antecesores (los *Baby Boomers* nacidos entre 1946 y 1964), es esperable que las preferencias y creencias de la demanda de inmuebles se vaya adaptando.

Existe literatura aún en desarrollo que presenta un enfoque similar al propuesto (Keys y Mulder, 2020). En él se realiza un análisis por control sintético de los precios de inmuebles en Miami. Si bien el trabajo hace hincapié en una primer caída en el volumen de transacciones que luego se ve reflejada en los precios, la conclusión de los autores anticipa un aumento en el pesimismo climático por parte de la demanda. La inclusión del mercado de hipotecas les permite constatar que no se debe a un cambio en el acceso a líneas de crédito. Por otro lado, encuentran que en los barrios más carenciados afectados por el aumento del nivel del mar, los precios de los inmuebles se ven más afectados que en aquellos barrios con mayores ingresos que también se encuentran en riesgo de inundación.[18] Por otro lado, la consultora MSCI ha desarrollado un informe en el cual señala que si bien la Barrera del Thamesís protege a la Londres Central (City of London, Holborn, Soho, etc.), otros sectores se ven muy perjudicados; como el municipio de Lewisham, que cuenta con el 35% de sus activos (según su valor capital) en situación de riesgo.[19] La Barrera del Thames constituye la mayor inversión en infraestructura de Inglaterra para hacerle frente a la crisis climática y al desborde del Támesis. Se trata de la barrera movible para prevención de inundaciones más grande del mundo y comenzará sus pruebas a partir del 11 de agosto de 2021.[20] El éxito de dicho empréstito puede alterar la percepción del público respecto a futuras inversiones en infraestructura.

[17] *Global warming age gap*. Gallup [En linea]: `https://news.gallup.com/poll/234314/global-warming-age-gap-younger-americans-worried.aspx`

[18] Benjamin J. Keys y Philip Mulder. *Neglected No More: Housing Markets, Mortgage Lending, and Sea Level Rise*

[19] *Underwater assets? Real estate exposure to flood risk* MSCI [En linea]: `https://www.msci.com/www/blog-posts/underwater-assets-real-estate/01593224766`

[20] `https://www.gov.uk/guidance/the-thames-barrier`

La Barrera del Támesis

2 Descripción del modelo teórico

Intuitivamente, al aumentar la población y al ser constante la extensión de las tierras disponibles, el precio de estas últimas debe aumentar a medida que aumenta la población. Es decir, mientras la oferta de *stock* de tierras se mantiene inalterada, la demanda de tierras aumenta a medida que aumenta la población. Consideraremos a la tierra como un activo cuyos dividendos se perciben en forma de renta (alquileres). Si bien el precio de la tierra es de una naturaleza bastante más compleja (el precio de la tierra también refleja su valor como activo colateral, su capacidad productiva, contiene un efecto de redes dependiendo de los vecinos y la infraestructura del barrio, etc.), parece ser una simplificación justificada para el estudio de nuestro caso: sencillamente consideramos una oferta constante y un aumento en la demanda agregada. En un determinado punto de corte se introduce la nueva información. Esta nueva información determina que en un momento específico del futuro, el activo será destruido. Compararemos dos trayectorias: Por un lado, la evolución del precio de la tierra donde en un determinado momento se informa que en el futuro dicho territorio sufrirá un daño irreparable y su precio caerá a cero; por otro lado, observaremos la evolución del precio de la misma tierra sin la nueva información, es decir, la trayectoria contrafáctica.

2.1 Estructura

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.

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.

2.2 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:

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

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

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

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:

$$D_t = \sum_{i \in I_t} T_{i,t}^d$$
$$D_t = cardinal(I_t) T_t^d$$
$$D_t = L_t \frac{M}{2alq_t}$$

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:

$$alq_t = L_t$$

2.3 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.

2.4 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, el precio nuevo dada la mismo inmueble w, va a ser:

$$precio_t^{nueva\ info} = \sum_{j=0}^{F-t-1} \frac{alq_{t+j}}{(1+i)^j}$$

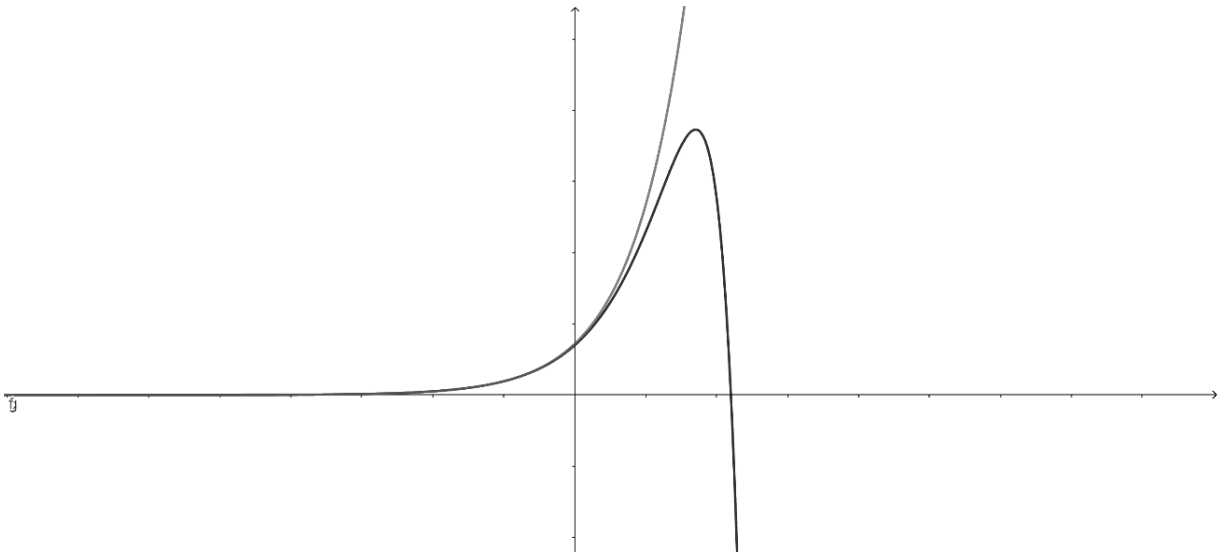
Sabemos que:

$$alq_t = L_t$$

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:



Donde el eje de abscisas representa el tiempo y el eje de ordenadas representa el precio; la curva gris es el precio de la vivienda sin nueva información, el caso placebo en el que no existe un informe climático que alerte sobre daños futuros; finalmente, la curva negra representa el precio de la vivienda en el caso de que los agentes incorporen la nueva información, siendo esta que el valor del inmueble en el punto F será nulo. Si la información se obtiene en el momento t=0, entonces el precio salta discretamente de la curva gris a la Negra en t=0, y de allí en adelante, la curva gris va a representar el contrafáctico de la casa si nunca se inundara en el futuro. 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:

$$precio\ real_t = \frac{L^t}{1 - \frac{L}{1+i}}\ si\ t < 0, cambiar\ si\ t \geq 0$$

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

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:

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

Es decir, 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.

3 Control Sintetico

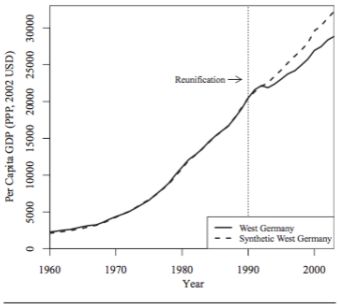
El método de control sintético es una generalización poderosa, y aún así relativamente sencilla, del método de diferencias en diferencias; permite evaluar cuantitativamente los efectos de un tratamiento específico sobre una unidad tratada mediante la creación de una unidad "sintética" contrafactual que no ha sido tratada. A grandes rasgos, los modelos de control sintético eligen un conjunto de ponderaciones óptimas que al aplicarlas al grupo correspondiente de unidades, produce una unidad contrafactual. Esta unidad sintética es comparada con la unidad tratada para capturar el efecto.

El efecto causal de un evento como el estudiado es determinado por la diferencia entre sus resultados potenciales; es decir, la diferencia entre el caso donde un municipio se ve afectado por los pronósticos de inundaciones contra el caso de la misma unidad, pero sin pronósticos. Como esto plantea que existan una misma unidad, pero bajo dos escenarios distintos al mismo tiempo, tenemos el problema de que uno de ellos no va a existir y para eso tenemos que construir su contrafactual.

3.1 Litertura

El método de control sintético fue diseñado por el economista vasco Alberto Abadie del MIT. En 2003, Abadie y Gardeazabal utilizaron dicho método contrafáctico para analizar el costo económico de los conflictos políticos internos del País Vasco. Los autores se preguntan cuál sería el ingreso per cápita de los habitantes de la comunidad autónoma si la organización separacionista E.T.A. no hubiera aterrorizado a la población local. En 2010, Abadie, Diamond y Hainmuller publican Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program donde estudian las caídas en las ventas de cigarrillos una vez implementadas políticas públicas que desincentivan su consumo.

FIGURE 2 Trends in per Capita GDP: West Germany versus Synthetic West Germany



Los autores notan una caída abrupta del consumo en comparación con el contrafáctico sintético generado a partir de estados en los que no se aplicó el programa. En 2014 los autores mencionados publican Comparative Politics and the Synthetic Control Method donde, mediante

el método de control sintético, analizan el impacto económico de la reunificación de Alemania ocupada en 1990. Los autores encuentran un menor crecimiento de Alemania Occidental tras la reunificación al compararlo con la unidad sintética. Con el tiempo, el método de control sintético ha ido ganando popularidad, al punto de que en el 2017 los economistas Athey e Imbens han escrito "the synthetic control approach developed by Abadie et al. [2010, 2015] and Abadie and Gardeazabal [2003] is arguably the most important innovation in the policy evaluation literature in the last 15 years".

3.2 Control Sintetico

Se propone una forma de construir el contrafactual con el fin de estimar el efecto, como podemos ver de la siguiente ecuación:

$$Y_{it} = \alpha_{it} D_{it} + Y_{it}^{N} = \alpha_{it} D_{it} + \theta_t Z_i + \lambda_t \mu_i + \delta_t + \epsilon_{it}$$

El efecto que se observa es el precio del inmueble Y_{it}^{N}, más el efecto del cambio climático i_t con su correspondiente variable dicotómica D_t.

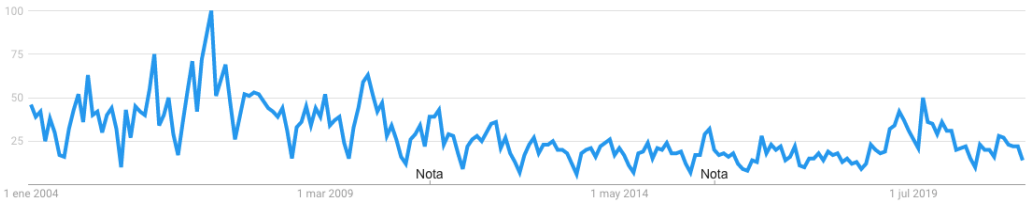
Puede uno pensar que la situación sin tratamiento es también la suma de factores observables e inobservables que están afectando a la variable de interés y estas son las que necesitamos para construir el contrafactual. Ahora, se podría utilizar un caso parecido – un match – que se acerque lo más posible a estas características para poder usarlo de contrafactual, pero también se podría usar el control sintético. Este consiste en tomar de un pool de unidades de control y promediarlas de forma tal que dicho promedio se asemeje al comportamiento de la variable de resultado de la unidad tratada previo al tratamiento. Es así como construimos los pesos w_j^*. De esta forma, si construimos promediando casos no tratados pero que, en conjunto, se asemeja a la unidad tratada podemos usarla para predecir el comportamiento del estado tratado contrafactual y, por lo tanto, estimar la diferencia entre ambos, que se puede interpretar como el efecto causal de la política.

$$\hat{\alpha_{1t}} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

3.3 Elección del año de corte: 2008

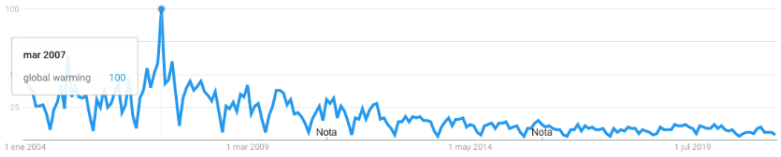
El método de control sintético requiere de un punto de corte donde comienzan a compararse la unidad tratada y la unidad placebo sintética. Hasta dicho punto, la trayectoria de la unidad tratada funciona como referencia para poder construir y fittear la unidad sintética. En el caso de un evento a estudiar (como ser la reunificación de Berlín en Abadie et al.) o el efecto de

una política pública puntual (sea el caso de un aumento en el salario mínimo como en Dube y Zipper) el punto de corte es sencillo de establecer: basta con señalar el año del evento a estudiar o la fecha en la cual se aplicó dicha política pública.



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

En nuestro caso, evaluamos respuestas del mercado inmobiliario frente a nueva información respecto a eventos futuros. Es difícil establecer cuando la población londinense tomó conciencia de que su territorio se enfrentaba a las consecuencias del cambio climático. En el año 2001, la película *Una verdad incómoda* de Al Gore llevó la discusión del clima al público mainstream. A partir de allí notamos un crecimiento acelerado de menciones del tema en todo tipo de publicaciones. Acorde a los datos de google trends, el volumen de búsquedas respecto al tema (ya sea "cambio climático" o el mal llamado "calentamiento global") alcanza su mayor magnitud el año 2007. Podemos suponer que para este año, la gran mayoría de las personas ya habían escuchado hablar del tema. Este dato concuerda con los resultados de la encuesta realizada por Yale sobre concientización de la problemática. El otro punto a considerar es la crisis de subprime que estalló en el año 2008. Al tratarse de una burbuja inmobiliaria, es de vital importancia tener en cuenta las consecuencias de dicho evento. Suponemos que al caer los precios, tras estallar la burbuja, el mercado inmobiliario debió corregir sus precios. Es cierto que dadas las circunstancias, el mercado puede realizar una sobre corrección empujando los precios por debajo de su valor al tratarse de circunstancias altamente volátiles. Consideramos, sin embargo, que tanto los datos presentados acerca de la opinión pública por un lado, y el estallido de la burbuja inmobiliaria por el otro, nos permiten realizar el corte en dicho momento.



Interés a lo largo del tiempo del Término de búsqueda "Global Warming" en Google Trends.

4 Machine Learning

El método de Machine Learning, desde el punto de vista del aprendizaje supervisado, consiste en programar computadoras para optimizar cierto criterio de performance (por ejemplo, minimizando el error cuadrático medio) 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. Actualmente dichos métodos se aplican tanto en instituciones financieras para predecir riesgo crediticio, condiciones climáticas, medicina, robótica, traducción, autos autónomos, entre otros. Cabe aclarar que este tipo de métodos requiere una abundante cantidad de datos para realizar predicciones más consistentes, ya que se dificulta encontrar patrones cuando los datos son pocos. Hal Varian, economista jefe de Google a dicho: "... my advice to grad (Economics) these days is 'go to the computer science department and take a course in machine learning'".[21]

4.1 Modelo K-vecinos

Considerando que es posible definir la variable de respuesta como $Y = f(X) + \varepsilon$

Dado una muestra Sn: $\{(X_1Y_1), ...(X_nY_n)\}$ y un parámetro $k \epsilon N$. Llamamos a X como vector de features continuos. De este modo, para cualquier $x \epsilon X$, llamamos $d_k(x, Sn) \equiv d_k(x)$ a la distancia (Euclidiana) desde x hasta su k-vecino mas cercano de entre los elementos de S_n. Si por ejemplo, $d1(x) = min_{i\,=\,1,...n} || x_i - x ||$.

Definimos $N_k(x) = \{j \epsilon \{ 1, ..., n\} \mid || x_j - x || < d_k(x)\}$, luego

$$\widehat{f(x)} = \frac{1}{k} \sum_{i \,\epsilon\, N_k(x)} yi$$

De este modo, el algoritmo estima la función en el punto X = x, haciendo un promedio local basado en los k – vecinos mas cercanos de X = x. A este método como a la mayoría de los métodos de machine learning, se le incorporan estrategias de validación cruzada para estimar el error asociado al modelo. Esto nos permite evaluar su capacidad predictiva (el fin ultimo del modelo).

4.2 Método de Kernel

Un núcleo o kernel K: $R^P x\ R^p -> R$ es una manera de computar productos internos entre observaciones proyectadas en un espacio de alta dimensión. De este modo, considerando ,$X = (x_1x_2)\ y\ X' = (x\prime_1x\prime_2)$, luego:

[21] https://people.ischool.berkeley.edu/~hal/Papers/2013/ml.pdf - Hal R. Varian - Big Data: New Tricks for Econometrics - June 2013

$$\phi^T(x)\phi(x') = x_1^2 x_1'^2 + x_2^2 x_2'^2 + 2x_1 x_1' x_2 x_2'$$
$$\phi^T(x)\phi(x') = (x_1 x_1' + x_2 x_2')^2$$
$$\phi^T(x)\phi(x') = K(x, x') \ \forall\, i, j = 1, ..n$$

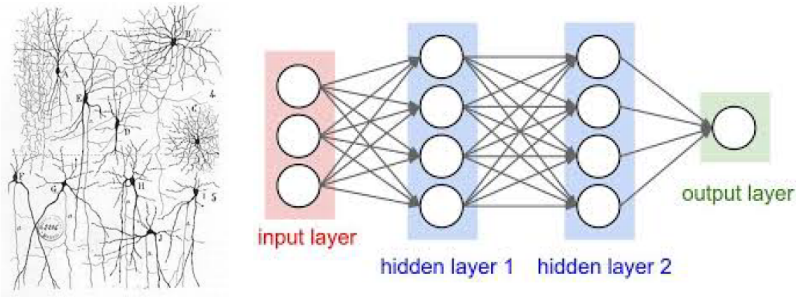
De esta forma, el método nos permite transformar problemas no lineales en formas lineales, ya que se puede demostrar que de la minimización del riesgo empírico, donde, dado φ y la muestra de entrenamiento $\{(\varphi(x_1), y_1), \ldots (\varphi(x_n), y_n)\}$,el modelo $Y = W^T\varphi(x) + \varepsilon$ es tal que minimizando la funcion de riesgo empirico:

$$RE(w, \lambda) = \frac{1}{2n}(y - X_\varphi w)^T(y - X_\varphi w) + \frac{\lambda}{2}||w||^2$$

Se obtiene que $\widehat{y_n} = \sum_{i=1}^{n} \alpha_i K(x_{new}, x_i)$

4.3 Redes Neuronales Clasicas

Los modelos de redes neuronales se caracterizan por una arquitectura formada por capas y "neuronas" en cada una de ellas. En cada una de estas neuronas residen funciones de activación que determina el output de cada neurona dado un conjunto de inputs de otras neuronas en una capa inmediatamente anterior. Los parámetros permiten a la red aprender de los datos de entrenamiento y eventualmente hacer predicciones. Por último, el algoritmo de entrenamiento nos permite ajustar los parámetros de la red para que aprenda a predecir la variable de interés.



Para describir el proceso en la cada de entrada encontramos las covariables del problema, esta a su ves son conectadas a la siguiente capa ponderadas por su respectivo peso (parámetro del modelo). Estos pesos (que podemos pensarlos como los parámetros de una simple función lineal compuesta por dichas Xs como inputs y β_s como parámetros) son el input de la neurona i-esima en la primera capa. Allí, el modelo lineal es transformado mediante una función de activación, que podemos llamar $\sigma(U),\ donde\ U = \sum_{j=0}^{p} \beta_j X_j$. Tanto el numero de capas como el de neuronas dependerá de como el analista diseñe la arquitectura de la red, y será una de las tantas variables para tener en cuenta en el fitting del modelo. El output de la

función de activiacion i-esima pasa a cada una de las siguientes neuronas en la siguiente capa ponderada por el peso correspondiente. La cantidad de estos procesos depende nuevamente de la arquitectura diseñada. Los parámetros (pesos) se ajustan para optimizar la capacidad predictiva dado un criterio seleccionado (por ejemplo, si calculamos el error mediante el error cuadrático medio, querríamos reducir este al mínimo posible, pero siempre manteniendo un conjunto de validación que nos permita ver que tanto ajusta el mismo en espacio de features sobre el que nunca recibió información), mediante el descenso del gradiente, buscamos el mínimo de la función de perdida seleccionada. Por otra parte, existen diversas funciones de activación que también forman parte del diseño de la red.

5 Control Sintético + Machine Learning

Nuestro objetivo principal es generalizar el modelo de estimación del contrafáctico al tipo de modelos predictivos usados en la disciplina del 'machine learning'. A la hora de estimar el contrafáctico, el control sintético presenta dos características que (bajo ciertas condiciones) pueden presentar mayor cantidad de desventajas que de ventajas, estas características son la extrapolación, y la linealidad. El uso de modelos predictivos puede llegar a disminuir el error de estimación lo suficiente como para realizar inferencia cuantitativa sobre los estimadores, es decir, al final del día vamos a obtener una distribución de probabilidad sobre el impacto de la información del cambio climático en el precio de las casas en Londres.

5.1 Métodos Predictivos: Principales ventajas y desventajas del control sintético.

5.1.1 Extrapolación

El control sintético presenta dos características que limitan la extrapolación de las predicciones:

a. No presenta una ordenada al origen (bias=0).

b. Los pesos no pueden ser mayores a 1 ni menores a 0

$\sum_{i \in donor\ pool} weights_i = 1\ ,\ weights_i \in [0,1]\ \forall i \in donor\ pool.$

Estas restricciones nos dan la ventaja de poder filtrar las estimaciones 'outliers' según su nivel en el output y así evitar sesgos. Por ejemplo, si tuviéramos que crear el control sintético de una casa situada en Londres, y la donor pool estuviera conformada por dos casas, una situada en Londres de precio similar y otra situada en algún barrio de bajos recursos en Estados Unidos tal que su precio es un cienavo del precio de la vivienda tratada, no sería preciso estimar el precio del contrafactico como $y_{contrafactico} = \frac{y_{Londres}+100y_{USA}}{2}$ ya que el precio de la casa de estados unidos no puede ofrecer ningún tipo de información sobre el contrafactico (estamos comparando una mansión de Londres con una choza de Estados Unidos). También funciona como método de regularización, puede verse como un modelo Lasso con una restricción adicional $(weights_i \in [0,1]\ \forall i \in donor\ pool)$, pero sin ningún hiperparametro. En este trabajo, vamos a prestar atención especial a los 'outliers' filtrados por el control sintético, pero vamos a permitir el uso de la extrapolación para aumentar la capacidad de complejidad en la función predictora. Este tipo de filtrado es una técnica heurística conocida del 'data engineering' en la cual se corren los modelos antes de trabajar con el preprocesado de datos para obtener pistas sobre posibles outliers.

5.1.2 Linealidad

Como estimación del precio contrafactico, el método del control sintético va a hacer un promedio ponderado de los precios de la donor pool:

$$\hat{y}_{contrafactico} = \sum_{i \in donor\ pool} w_i y_i$$

Esta fórmula es una extensión lineal del método de matching en la que se crea un control usando un promedio de los objetos macheados, una gran ventaja que tiene es su interpretabilidad, cuando construimos el sintético, podemos verificar el valor de cualquier otra variable contra su sintético, por ejemplo, sea hab la cantidad de habitaciones en una casa, podemos obtener la cantidad de habitaciones en el sintético usando los mismos pesos:

$$hab_{contrafactico} = \sum_{i \in donor\ pool} w_i hab_i$$

Pero el problema es el mismo que en el punto anterior, aumentar la capacidad en la complejidad del modelo puede disminuir la varianza en los estimadores.

5.2 Varianza de los estimadores y complejidad potencial

El cambio que vamos a hacer en los supuestos es muy directo, anteriormente, ya estaba demostrado[22] que, bajo ciertas condiciones, el estimador del control sintético no está sesgado, en otras palabras:

$$\hat{y}_{contrafactico} = y_{contrafactico} + \varepsilon$$

Donde $E\{\varepsilon\} = 0\ y\ Var(\varepsilon) = \sigma^2$. Por lo tanto:

$$y_{contrafactico} = \sum_{i \in donor\ pool} w_i y_i + u$$

Donde $E\{u\} = 0\ y\ Var(u) = \sigma^2$. Nuestra generalización consiste en suponer:

$$y_{contrafactico} = f(\{y_i | i \in donor\ pool\}) + v \qquad (1)$$

Donde $E\{v\} = 0\ y\ Var(v) = \sigma^2$. Es decir, suponemos que existe una función (no necesariamente lineal) que no esta sesgada. Entonces, nuestro trabajo va a consistir en buscar una estimación

[22] Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. `https://economics.mit.edu/files/11859`

de f:

$$\hat{y}_{contrafactico} = \hat{f}\left(\{y_i | i \in donor\ pool\}\right) \qquad (2)$$

Por otro lado, el parámetro a estimar es el efecto de la adquisición de información nueva, más específicamente:

$$\alpha = y_{contrafactico} - y_{real} \qquad (3)$$

Por lo tanto, el estadístico a usar es:

$$\hat{\alpha} = \hat{y}_{contrafactico} - y_{real} \qquad (4)$$

Por lo tanto, ahora el sesgo de nuestra estimación es:

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

Trabajando con (1), (2), (3) y (4), obtenemos:

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

Entonces el sesgo del estimador es el mismo que el del predictor, esto quiere decir que modelos de predicción con mucho sesgo van a transmitir su sesgo directamente hacia el estadístico, por experiencia, los modelos con mayor sesgo son los más simplistas (ejemplo, regresión lineal), esto quiere decir que el modelo de control sintético va a transmitir su sesgo directamente al estadístico en el momento en el que cambiamos a nuestros supuestos. Teniendo en cuenta que estamos trabajando con un estimador sesgado, vamos a centrar nuestros esfuerzos en disminuir el error de estimación, es decir:

$$error = E\left\{(\hat{\alpha} - \alpha)^2\right\}$$

Trabajando con (1), (2), (3), y (4), obtenemos:

$$error = \left(E\left\{\hat{f}\right\} - f\right)^2 + E\left\{\left(E\left\{\hat{f}\right\} - \hat{f}\right)^2\right\} + \sigma^2$$

Es decir, el problema de encontrar un buen estimador es análogo a resolver el bias-variance trade off:

$$error = BIAS\left\{\hat{f}\right\}^2 + Var\left\{\hat{f}\right\} + \sigma^2$$

Es por esta razón que optamos por usar modelos potencialmente más complejos.

5.3 Problemas de los modelos predictivos

Uno de los mayores problemas de este método es la replicabilidad de los resultados. Si Abadie ya estaba usando técnicas de cross validation para combatir el overfitting en un modelo lineal, es esperable que, al usar modelos no lineales, el termino de varianza sea tal $(Var\left\{\hat{f}\right\})$ que los resultados no puedan pasar los tests de robustez. Por otro lado, los modelos no lineales siempre van a tener diferentes grados de interpretabilidad, pero en general, es muy difícil sacar conclusiones sobre su funcionamiento, esta propiedad les da el nombre en la jerga popular de 'caja negra'.

5.4 Estimadores

Para poder aplicar los métodos de inferencia, vamos a definir específicamente los estimadores con los que vamos a trabajar. Primero obtenemos las predicciones del contrafactico para todo $t > T$ donde T es el punto de corte:

$$\hat{y}_t^{CF} = \hat{f}\left(\{y_{i,t} | i \in donor\ pool\}\right)$$

Donde $\hat{y}_t^{CF}$ es la estimacion del contrafactico; luego obtenemos el efecto para cada $t > T$:

$$\hat{\alpha}_t = \hat{y}_t^{CF} - y_t^{real}$$

Como vimos en la seccion teorica, vamos a trabajar con el logaritmo de las diferencias:

$$lnDif_t = ln\left(\hat{\alpha}_t\right)$$

Por lo que vimos en la seccion teorica, $lnDif_t$ tiene una forma lineal, entonces obtenemos su tendencia:

$$ln\vec{D}if_{t,i} = \gamma_i + \beta_i t$$

Como vimos, β_i es aproximadamente la tasa de descuento de la inundacion, y e^{γ_i} representa el salto discreto en t=0; sin embargo, no vamos a obtener los valores usando una regresion lineal contra el logaritmo, sino que vamos a calcular el mejor modelo del tipo $\delta_i e^{\beta_i t}$ que minimize el error cuadratico con respecto a la secuencia de $\hat{\alpha}_t$. En principio, los dos métodos parecen similares, sin embargo, cuando calculamos el error cuadrático del logaritmo de la variable, en realidad, estamos dándole mayor importancia a las diferencias de los valores mas pequeños, o, en otras palabras, con el método que vamos a usar, le damos mas importancia a los valores mas grandes. La razón de la toma de esta decisión es que, como vamos a ver con los intervalos de confianza, la varianza en la medición de $\hat{\alpha}_t$ es razonablemente constante en el tiempo, y por lo tanto, se genera heterocedasticidad cuando usamos los logaritmos.

Por último, definimos nuestro estimador:

$$\vec{v_i} = \begin{bmatrix} \delta_i \\ \beta_i \end{bmatrix}$$

6 Inferencia

En esta sección, vamos a adaptar el método de inferencia a nuestra elección de modelos de estimación de 'caja negra', siempre el objetivo en mente es obtener resultados de carácter cardinal, sin embargo, los métodos a usar van a ser una extensión de los métodos ordinales del paper Pooling Multiple Case Studies Using Synthetic Controls: An Application to Minimum Wage Policies[23]. Como estadísticos, vamos a usar el estimador del efecto $(\alpha_{t,i})$, y su tendencia (β_i), y una estimación del cambio discreto de precios (γ_i). Notar que en este caso, una unidad tratada es un municipio de Londres que se va a inundar en el futuro.

6.1 Método Principal

6.1.1

El test de hipótesis va a ser sobre la media del vector de descuento:

$$\vec{v_i} = \begin{bmatrix} \delta_i \\ \beta_i \end{bmatrix}$$

Seguimos un método de inferencia basado en el test placebo de permutaciones donde, luego de estimar los $\vec{v_i}$ de cada unidad tratada, estimamos los $\vec{v_i}$ de las unidades pertenecientes a la donor pool para crear una distribución placebo del estadístico. La diferencia con el test de rangos de Arindrajit Dube y Ben Zipperer esta en que nosotros vamos a trabajar con los desvíos estándar en vez de los rangos, y por lo tanto, en vez de usar distribuciones uniformes, bajo hipótesis nula vamos a usar la estimación kernel de la densidad del conjunto $\{\vec{v_i} | i \in donor\ pool\}$. A esta estimación la llamamos $kdensity(\delta, \beta)$, notar que usamos una estimación kernel de la densidad porque como estamos trabajando con inferencia cardinal, nos interesa extrapolar la kurtosis de la distribucion de placebos, es decir, al tener pocos datos en la donor pool, no conocemos bien el peso de las puntas de la distribucion.

El test de hipótesis es:

$$H_0 : \vec{v}_{tratados} = \vec{v}_{placebos}$$
$$H_1 : \vec{v}_{tratados} > \vec{v}_{placebos}$$

Donde:

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

[23] http://ftp.iza.org/dp8944.pdf

En este caso, un vector es mayor a otro cuando todos sus valores internos son mayores.

6.1.2 Distribucion de las medias

La distribución del estadístico en cualquier estudio de control sintético es delicada, no tenemos tan pocos datos como para usar probabilidad bayesiana, pero tampoco tenemos tantos datos como para asumir una convergencia normal, para resolver este problema, Dube y Zipperer usan una distribución teórica de la media de los rangos (Irwin-Hall distribution), y eso es una ventaja de la inferencia ordinal, es mucho mas robusta; pero si tomamos a la variable aleatoria $\vec{V_i}$ como el descuento del estado tratado 'i' en el caso en el que no exista ningún efecto sobre este, entonces, bajo hipótesis nula, $\vec{V_i}$ se distribuye como $kdensity(\delta, \beta)$. $kdensity(\delta, \beta)$ se centra en captar la incertidumbre que se deriva de la ignorancia sobre la capacidad del grupo de control para reproducir un contrafactico [24].

Entonces, sabemos que, bajo hipótesis nula, $\vec{v_i}$ es una realización de $\vec{V_i}$, por lo tanto, la distribución de $\vec{v}_{tratados}$ bajo hipótesis nula sigue la misma distribución que $\bar{V}$ donde:

$$\bar{V} = \frac{1}{cardinalidad(tratados)} \sum_{j \in tratados} \vec{V_j}$$

Notar que, bajo hipótesis nula, la adquisición de información no genera un efecto, por lo tanto, las distribuciones $\vec{V_j}$ son independientes, así que, si asumimos homogeneidad en las distribuciones placebo, la distribución de $\bar{V}$ convergería a una distribucion normal bivariada. Sin embargo, como solo tenemos 17 distritos tratados, vamos a usar una simulación Montecarlo para aproximar la distribucion de $\bar{V}$.

6.2 Estimador de la diferencia y su inversión

Para obtener una distribución de probabilidad sobre los efectos del tratamiento, debemos recurrir a la probabilidad bayesiana.

Aplicando los mismos conceptos, podemos obtener la distribución placebo de cada efecto $(\alpha_{j,t})$ par cada momento t:

$$alpha\_acumulado_{j,t}(x) = \frac{1}{O} \sum_{i \in donor\ pool} \mathbf{1}_{\alpha_i \leq x}$$

Donde O es un numero tal que `alpha_acumulado` es una distribucion de probabilidad y $\mathbf{1}_{\alpha_i \leq x}$ es una función indicadora. Llamamos $kalpha\_acumulado_{j,t}(x)$ a la estimación kernel de $alpha\_acumulado_{j,t}(x)$, sin embargo, en este caso, la interpretación de esta función acumulativa es diferente a la anterior, en este caso, sea $\widetilde{\alpha}_{j,t}$ la variable aleatoria que se distribuye como

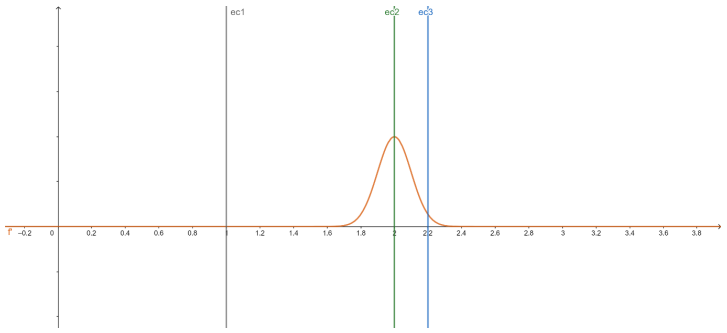
[24] Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program

$kalpha\_acumulado_{j,t}$, entonces, vamos a ver que la variable $\hat{\alpha}_{j,t} + \widetilde{\alpha}_{j,t}$ va a representar la distribucion de probabilidad de los efectos:

a. Como sabemos, la distribucion empirica $kalpha\_acumulado_{j,t}(x)$ capta la incertidumbre sobre la capacidad del conjunto del donor pool de reproducir un contrafactico, es decir, dado un contrafactico real y^{CF}, la distribución de probabilidad de la estimación del contrafactico es:

$$density\left(\hat{y}^{CF} - y^{CF} \middle| y^{CF}\right) = density(alpha\_acumulado_{j,t}\left(\hat{y}^{CF} - y^{CF}\right))$$

El grafico de esta ecuacion seria:



En este caso, la linea negra vertical ubicada en x=1, representa el valor real de la unidad tratada (y^{real}), la linea vertical verde representa el contrafactico real de la unidad tratada (y^{CF}), la línea vertical azul representa la estimación del contrafactico obtenida por el modelo $(\hat{y}^{CF})$. Obviamente, si pudiéramos conocer y^{CF}, entonces el efecto real quedaría $\alpha = y^{CF} - y^{real}$, pero los unicos datos que tenemos son y^{real}, $\hat{y}^{CF}, ydensity\left(\hat{y}^{CF} - y^{CF} \middle| y^{CF}\right)$. Notar que la densidad $(density(kalpha\_acumulado_{j,t}\left(\hat{y}^{CF} - y^{CF}\right)))$ es la función de verosimilitud, por lo tanto, podemos aplicar la formula bayesiana:

$$density\left(\alpha|\hat{\alpha}\right) = \frac{density\left(\hat{\alpha}|\alpha\right) priori(\alpha)}{constante}$$

Ahora, nosotros conocemos $density\left(\hat{\alpha}|\alpha\right)$, ya que, si conocemos y^{CF}, entonces conocemos $\alpha = y^{CF} - y^{real}$, y por lo tanto, $density\left(\hat{\alpha}|\alpha\right) = density\left(\hat{y}^{CF} - y^{CF} \middle| y^{CF}\right)$. Por otro lado, la constante es un numero que se encarga de normalizar la distribución. En último lugar, la distribución a priori es un tema que puede generar discusión, es decir, lo que vamos a usar nosotros es $priori\left(\alpha\right) = 1$ (notar que es una distribución impropia) porque a priori, creemos que el efecto puede estar en cualquier punto de la recta con probabilidad uniforme, sin embargo, proponemos una priori que puede ponderar diferentes grados de creencia intuitiva sobre el problema inicial.

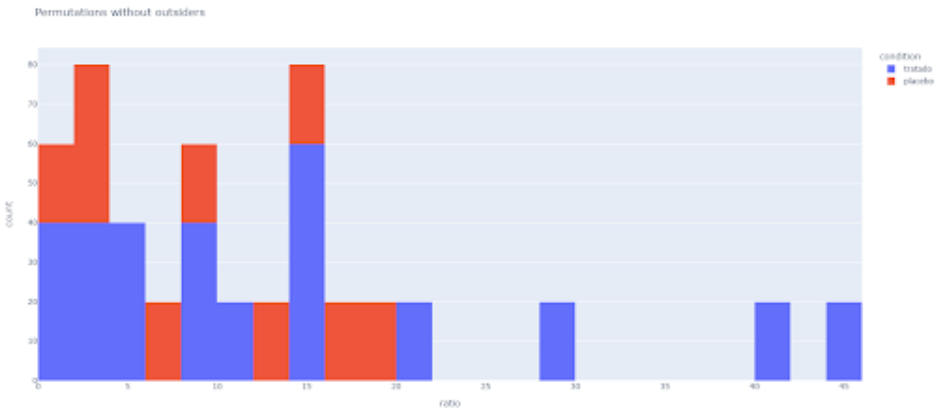
6.3 Test Final

Por ultimo vamos a realizar los tests de medias, pero usando los $\vec{v}_i$ ajustados, y siguiendo una interpretacion bayesiana sobre los resultados, es decir, en este caso la variable $\vec{v}_{tratados} + \bar{V}$ va a representar la distribucion de probabilidad de la media de $\vec{v}_i$.

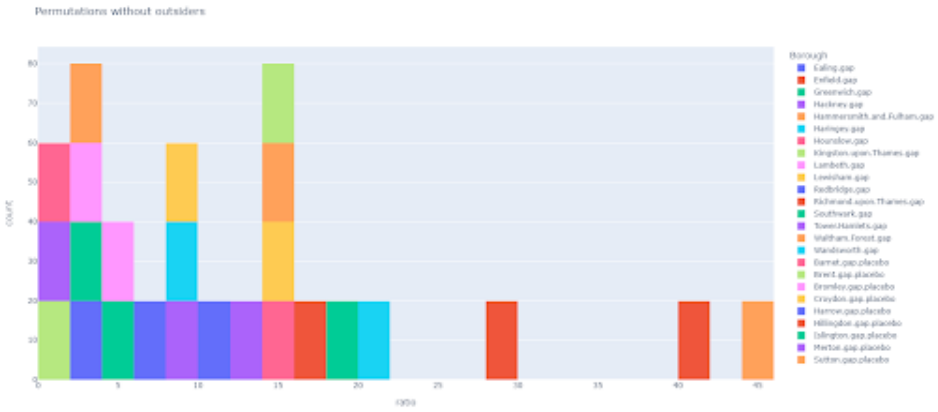
7 Resultados

7.1 Control Sintetico

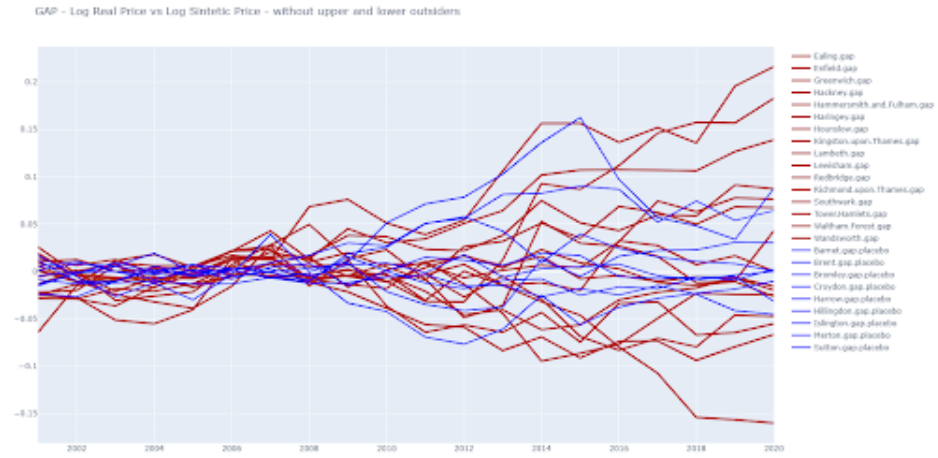
En primer lugar, luego de quitar aquellos municipios que mediante control sintético no pueden replicarse de forma confiable, dada las restricciones que impone el mismo con respecto a la extrapolación, realizamos un histograma con la relación entre los errores post vs pre tratamiento. Esto nos permite visualizar cuáles de los barrios testeados presentan cambios significativos luego del año de corte como se esperaría, entendiendo significativos como aquellos que se alejan ampliamente de la distribución de los placebos (que no deberían presentar un gran cambio en su variabilidad). Como vemos, el caso emblemático es el de Hammersmith and Fulham, cuya ratio post/pre se encuentra en 45, y por lo tanto a 4.8 veces la media de la distribución de los placebos (media de placebos: 9.377, desvío: 6.124). Esto otorga un alto grado de confianza en la predicción del barrio que se encontraría más afectado por la inundación y cuya dirección inferida mediante control sintético respalda. Por otra parte, Richmond Upon Thames se encuentra a 4.42 veces la media de los placebos, Enfield a 3.17 veces. Todos ellos, se encuentran a más de 2 desvíos de la media de los placebos. Además, dentro de este grupo, todos presentan la dirección esperada, lo que brinda fortaleza a la hipótesis de que la expectativa de inundación está siendo incorporada por los agentes. Sin embargo, del grupo de tratados, tan sólo contamos con 3 de los 16 barrios, dado este nivel de robustez. Por otra lado, tan solo 8 de los 16 presentan una diferencia a la esperada por dicha hipótesis.



Histograma de ratios de errores pre vs post tratamiento por grupo tratado vs control



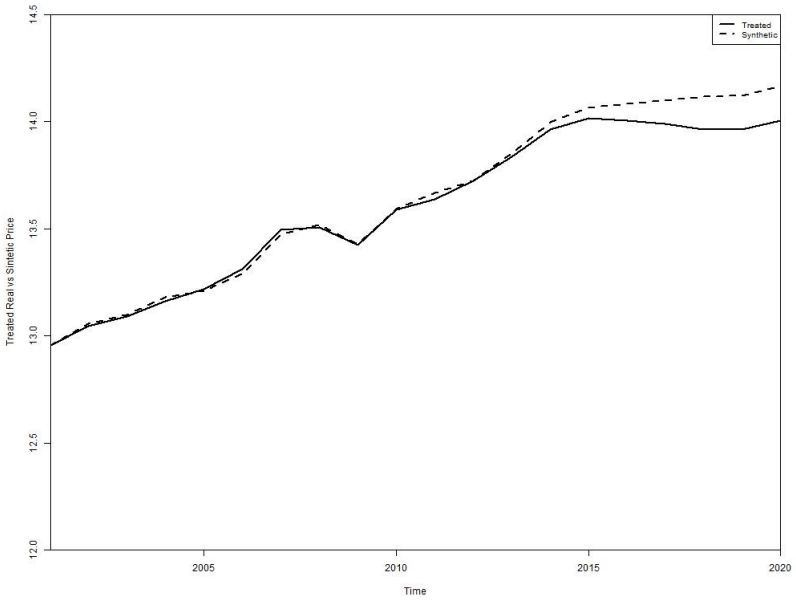
Histograma de ratios de errores pre vs post tratamiento de cada barrio

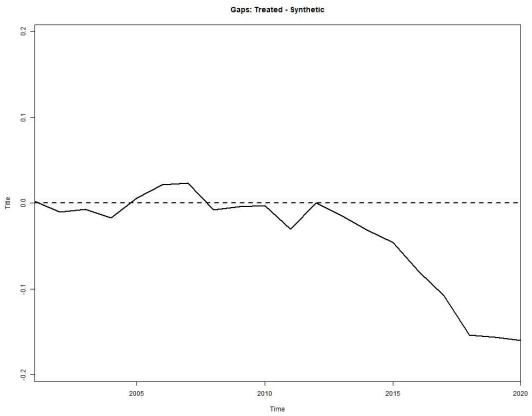


Gap entre el sintético y la serie real de la muestra de barrios sin outsiders, donde aquellos tratados se presentan de color rojo, mientras que el grupo de control de color azul.

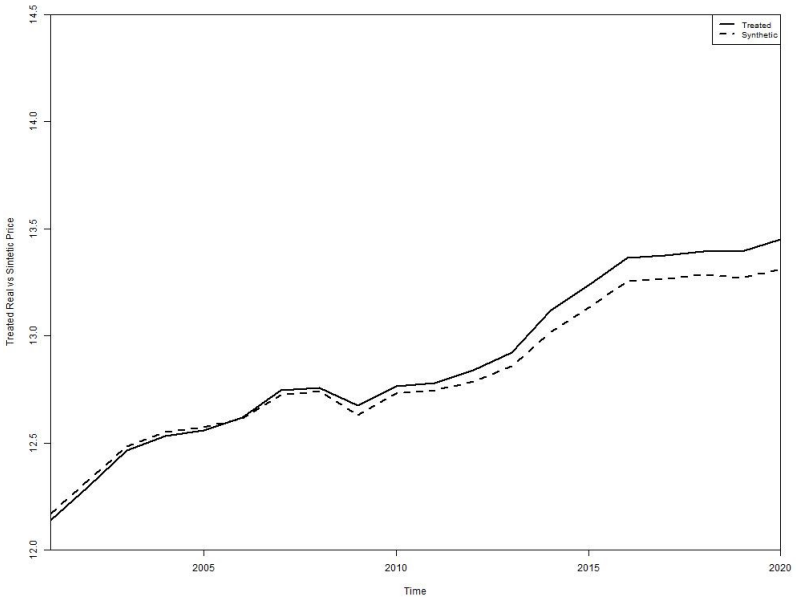
Para cada municipio se presentan dos gráficos que ilustran los resultados encontrados. Por un lado, comparamos las trayectorias de los precios de los inmuebles en relación a la unidad sintética placebo. El segundo gráfico consiste en la diferencia entre la unidad tratada (en nuestro caso, en riesgo de inundación) y la unidad sintética.

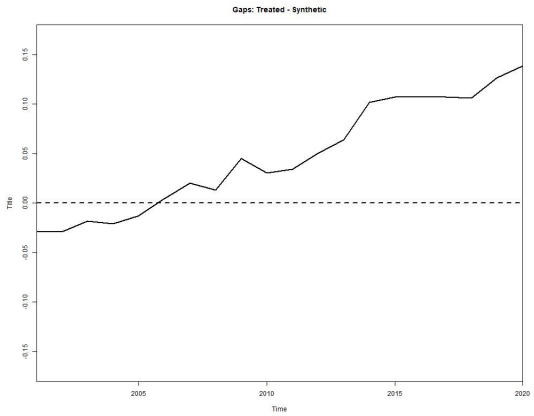
7.2 Caso Hammersmith and Fulham



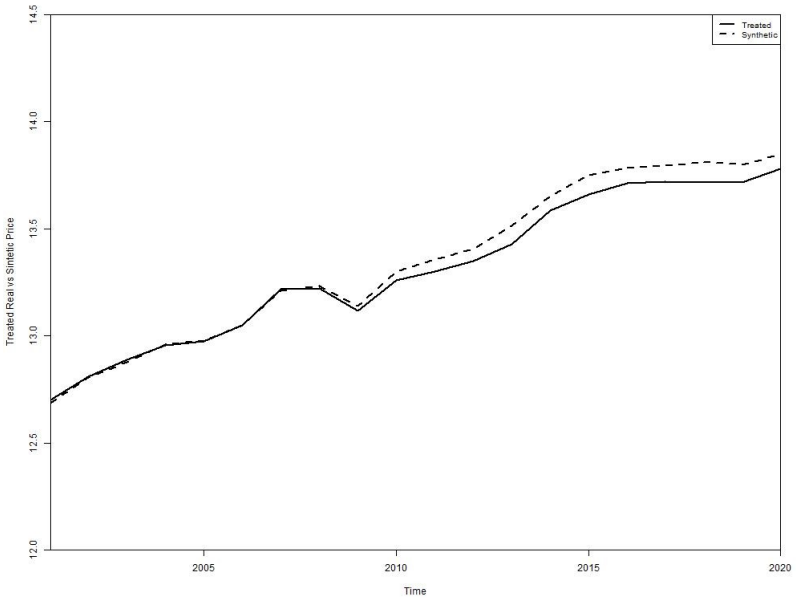


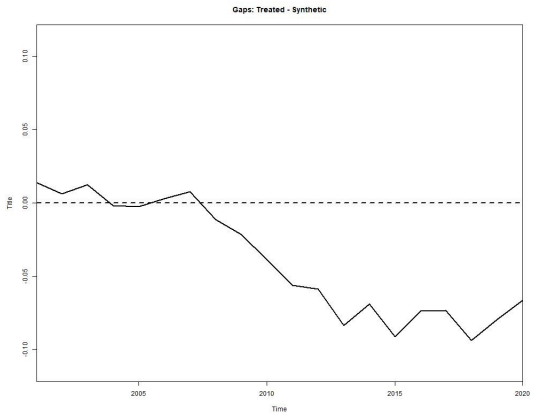
7.3 Caso Lewisham





7.4 Caso Richmond upon Thames





8 Resultados con Machine Learning

8.1 Comparación de modelos

8.1.1 Control Sintético

Tal como se acostumbra en todo trabajo de machine learning, usamos el error del test set como punto de comparación entre los modelos. En este caso, el test set es el conjunto de placebos luego del punto de corte. Pero antes de hacer la comparación, vamos a filtrar los municipios outliers usando las predicciones del control sintético:

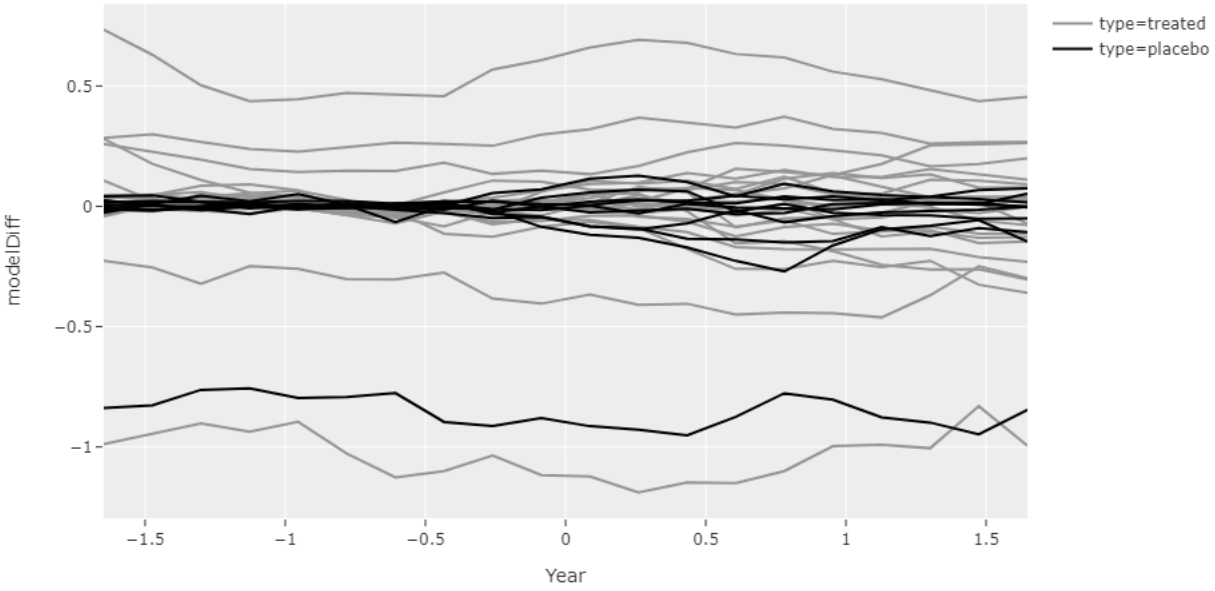
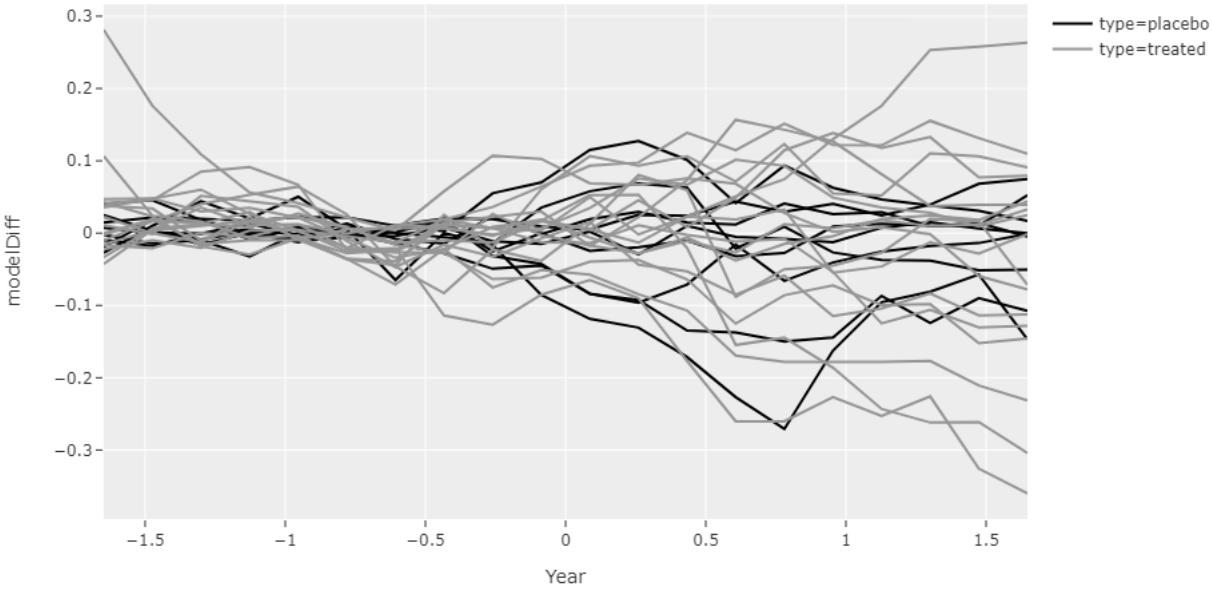


Figure 1: En este caso, los outliers son los municipios donde los precios de los inmuebles son mucho mayores (o mucho menores) a la media. error: 0.08311114586964855

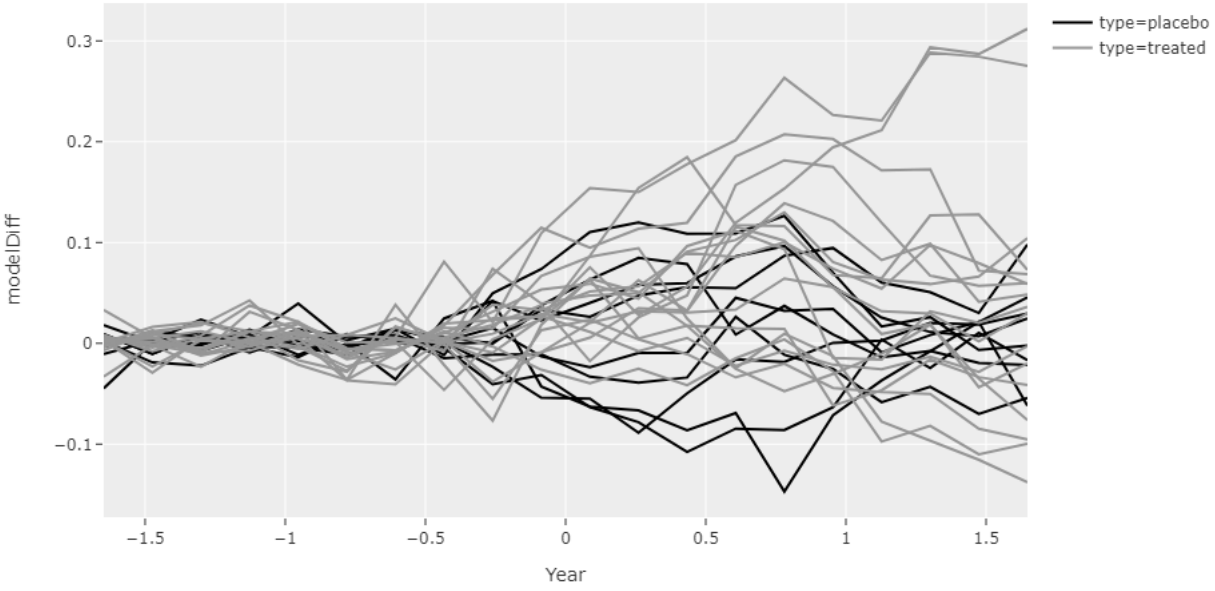
A continuación, al deshacernos de los outliers señalados, compararemos los distintos modelos.



error: 0.005157038599079725

este gráfico es el análogo a los gráficos en los trabajos de control sintético. Las variables están normalizadas, y por lo tanto, el punto de corte es en -0.42. Notar que este es el mismo gráfico que el presentado en la introducción, pero en este caso usamos $\alpha = y^{CF} - y^{real}$ mientras que en ese caso usamos $\alpha = y^{CF} - y^{real}$ (y también se encuentra en otra escala de normalización).

8.1.2 Lazy ridge



Error: 0.0029926219842486335

En primer lugar, queremos aclarar que este es un modelo lineal, el problema que estamos atacando en este caso, es el de la extrapolación; ahora pasamos a explicar el modelo $ridge$ y luego el modelo $lazy\ ridge$: Como argumentamos en secciones anteriores, un potencial problema del control sintético es su imposibilidad para extrapolar resultados, lo que sucede es que como ya filtramos los outliers antes de comparar los modelos, entonces nos podemos permitir usar una ordenada al origen y pesos menores a 0 o mayores a 1, así que este modelo es una simple regresión lineal usando una variación del modelo $ridge$ como regularización. Con respecto a la regularización, el control sintético ya aplica regularización lasso cuando impone la restricción:

$$\sum_{i \in placebos} weights_i = 1$$

La diferencia con el modelo $lasso$ estándar es que no usa ningún hiperparámetro, es decir, generalmente, se hace cross validation para encontrar el hiperparámetro donde:

$$\sum_{i \in placebos} |weights_i| = \lambda$$

Pero en el caso del control sintético, se usa $\lambda = 1$, y también el módulo no es necesario porque ya sabemos que los pesos son mayores a 0. Por otro lado, nosotros decidimos usar el modelo $ridge$. La idea principal es permitir mayor extrapolación, es decir, por el uso de

módulos, el modelo lasso muchas veces genera resultados de esquina en los weights, pero como solo tenemos 10 placebos disponibles para crear el sintético, no podemos permitirnos perder información; si que en el caso del modelo $ridge$, la restricción sobre los pesos es:

$$\sum_{i \in placebos} {weights_i}^2 = \lambda$$

Por último, el método típico es hacer una malla de hiperparametros sobre λ, y chequear cual es el que minimiza el error de validación. Nosotros vamos a intentar evitar la creación de mallas de hiperparametros, ya que la cantidad de chequeos sobre el conjunto de validación crece exponencialmente con cada hiperparámetro nuevo que se agrega al modelo; por lo tanto inventamos un tipo de regularización que llamamos $lazy\ ridge$ [25].

En primer lugar definimos la regresión como si fuera una red neuronal sin hidden layer y sin activation function, y luego usamos $stochastic\ gradient\ descent$ para obtener los pesos imponiendo un λ=0 (es decir, sin regularización), luego usamos SGD con un λ alto, ejemplo $\lambda = 0.01$, y por último volvemos a usar SGD con $\lambda = 0$, sin embargo, en el último entrenamiento, usamos 'early stopping' para elegir el mejor modelo:

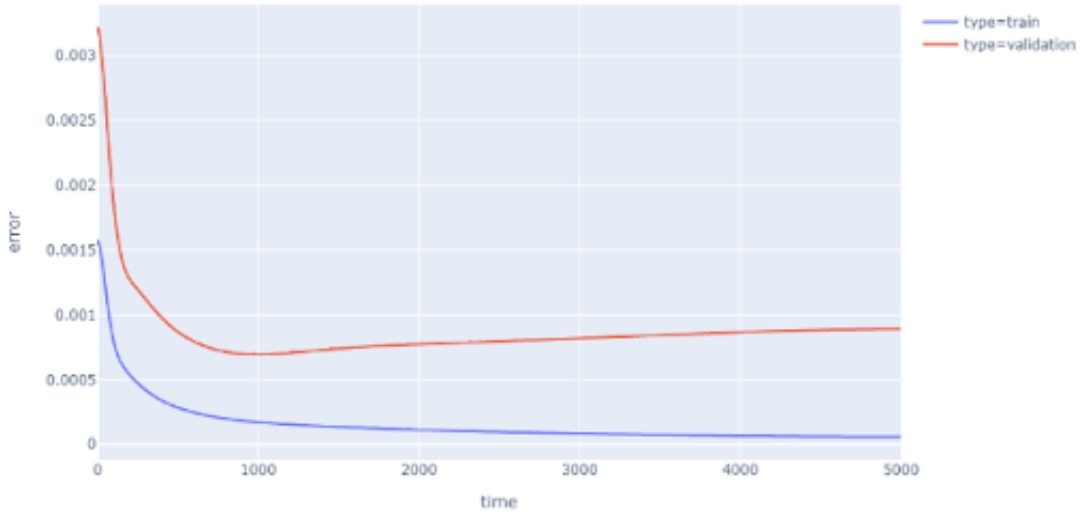
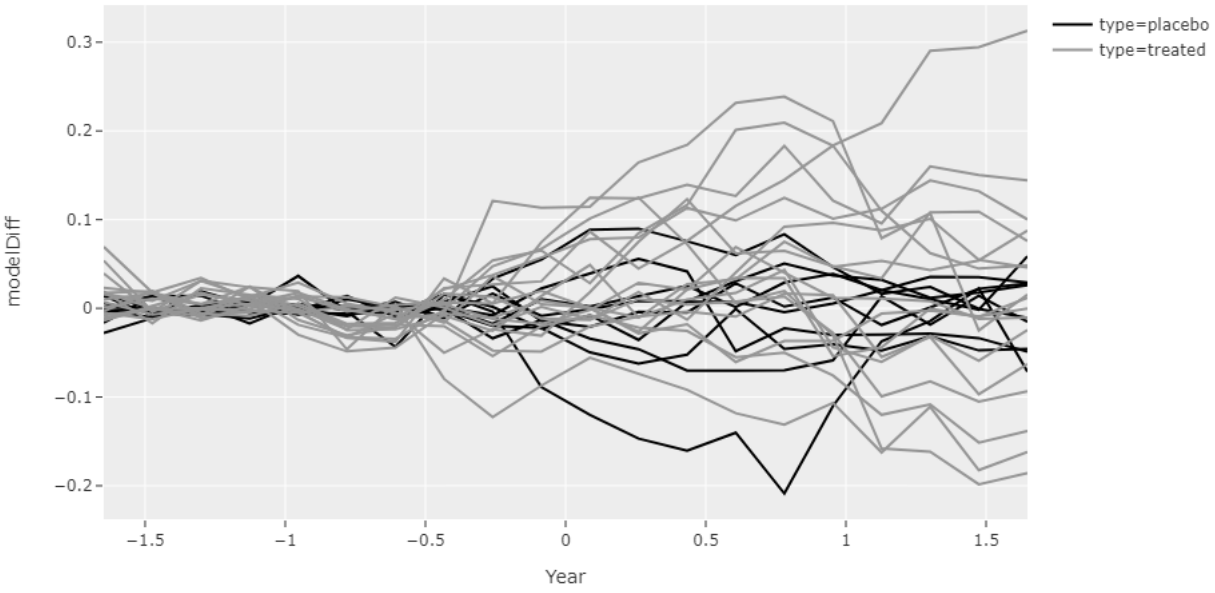


Figure 2: En nuestro caso, elegimos el modelo que se encuentra cerca de la iteración número 1000

La idea es usar el concepto de triangulación para replicar el sendero de expansión de los pesos generado por λ; en vez de ir paso a paso por el sendero de expansión, primero vamos a un punto donde sepamos que el modelo hace $underfitting$, y luego usamos SGD para llegar al punto donde se hace $overfitting$, en este caso, se recrea un sendero de expansión artificial.

[25] inspirada en el paper Cyclical Learning Rates for Training Neural Networks de Leslie N. Smith

8.1.3 Kernel Ridge



Error: 0.0025659922177516

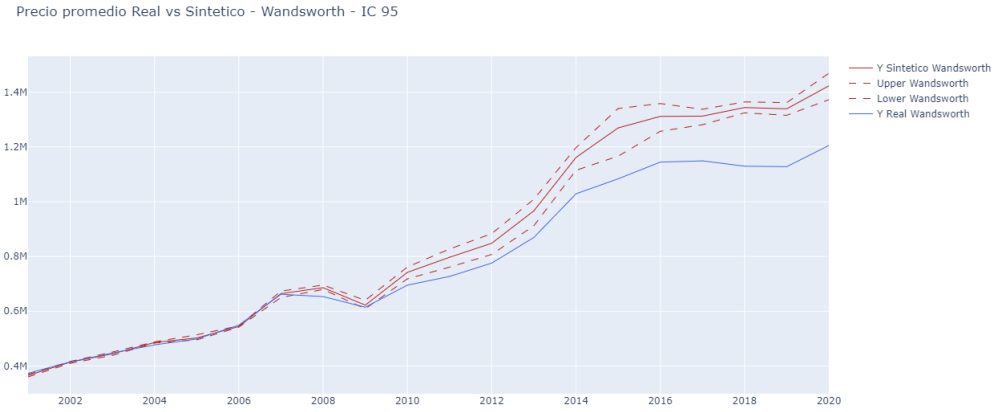
Con este modelo, intentamos solucionar el problema de la linealidad, es decir, utilizamos un modelo que usa una regresión lineal $ridge$ pero en un espacio nuevo de features usando el $kernel\ trick$ para generar un modelo no lineal. Todo este análisis hace posible disminuir la incertidumbre sobre la capacidad que tiene nuestro modelo para generar contrafácticos precisos sobre los distritos tratados.

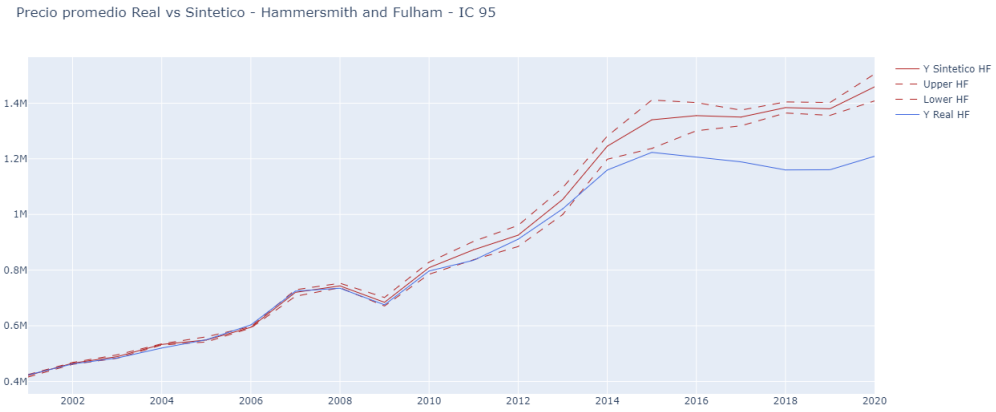
Para hacer el test de hipótesis, vamos a usar los resultados del modelo $lazy\ ridge$ ya que el cambio del error con respecto al modelo $Kernel\ ridge$ es muy bajo, pero a la vez, (como es un modelo lineal) no presenta tantos problemas de varianza como el modelo más potente (si es que en un futuro se busca replicar los resultados).

8.2 Test de hipotesis

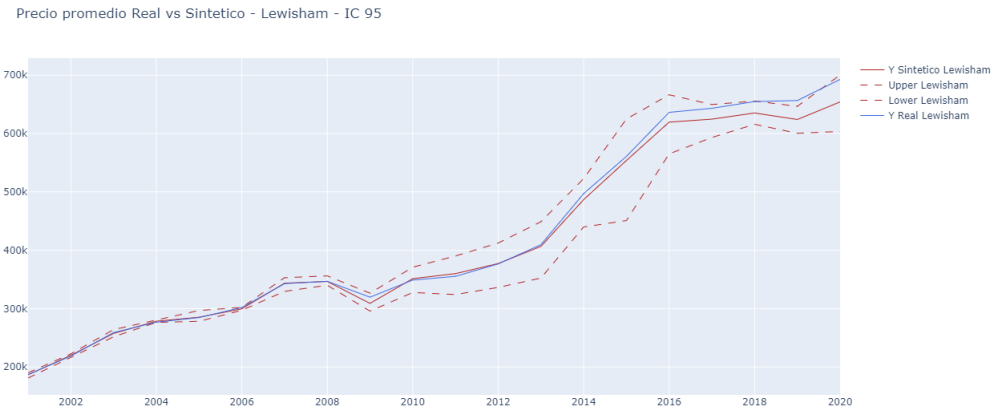
8.2.1 Selección del modelo exponencial

Siguiendo los argumentos de la sección de inferencia, obtuvimos un intervalo de confianza sobre las predicciones del contrafáctico, los barrios tratados más significativos presentan un gran cambio:





Por otro lado, los barrios que presentan un cambio negativo no tienen un efecto significativo:



Como se puede ver en estos gráficos, la varianza de la predicción a partir de 2008 se mantiene constante en el tiempo, esto quiere decir, que si tomamos estos resultados según la interpretación bayesiana (como una distribución de probabilidad), entonces el logaritmo de

las diferencias va a presentar problemas de heterocedasticidad. Es por esto que directamente aproximamos un modelo exponencial sobre las diferencias:

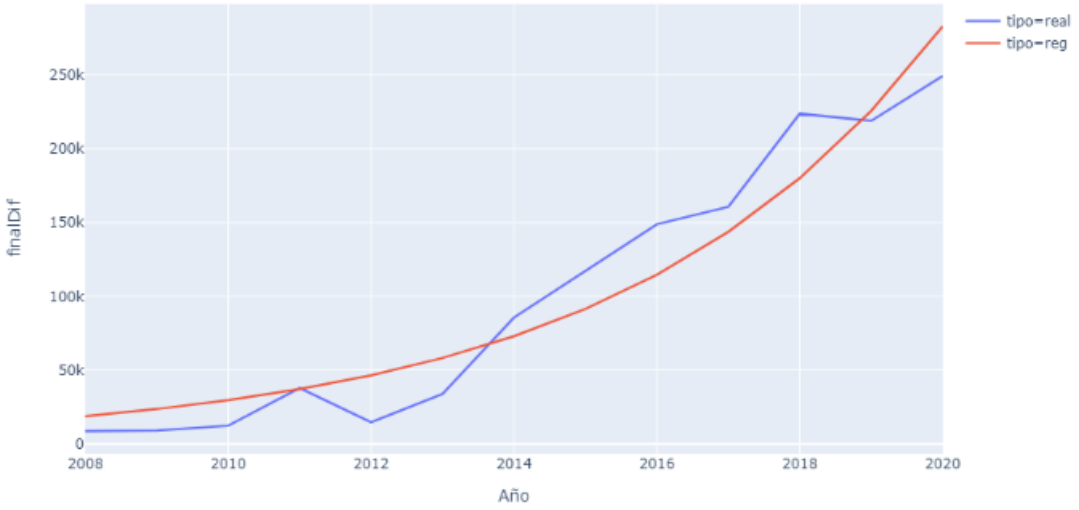
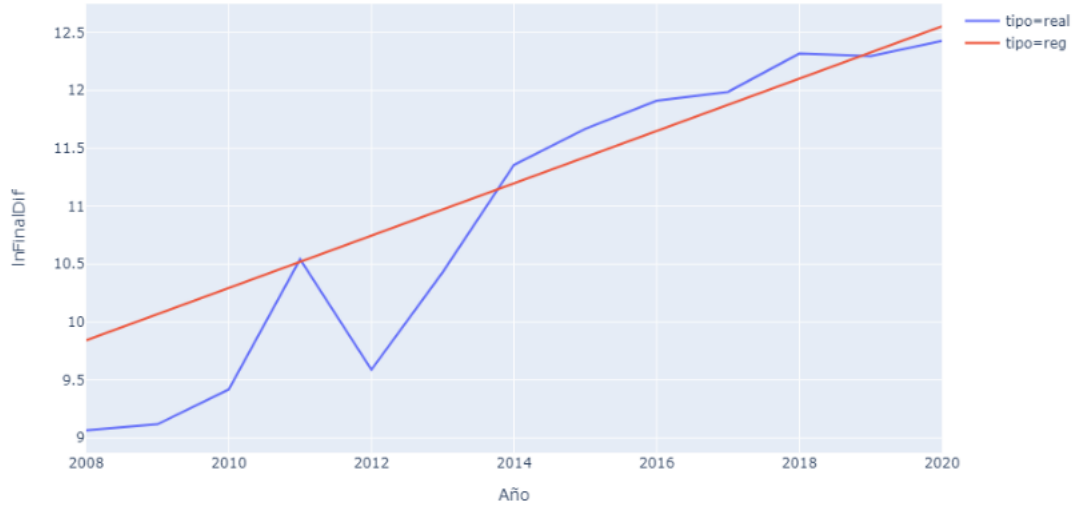


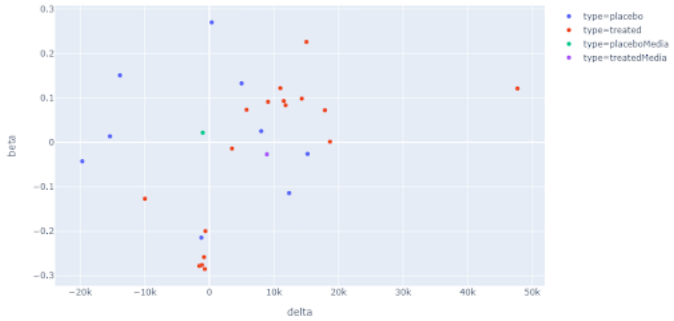
Figure 3: Aproximación exponencial de la tendencia en Hammersmith and Fulham

La ventaja de este método es darle menos importancia al error cuadrático de las diferencias logarítmicas de los años donde el efecto no es apreciable, por lo tanto, el logaritmo del grafico anterior queda:



8.2.2 Distribución de los parámetros tendenciales calculados

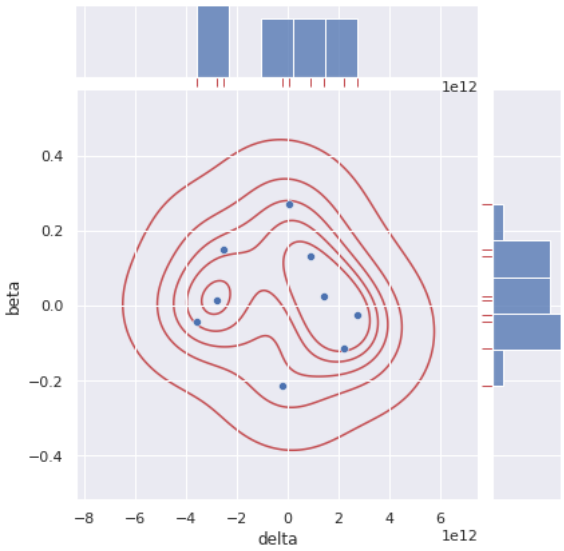
Una vez que calculamos las tendencias exponenciales de cada diferencia, obtenemos:



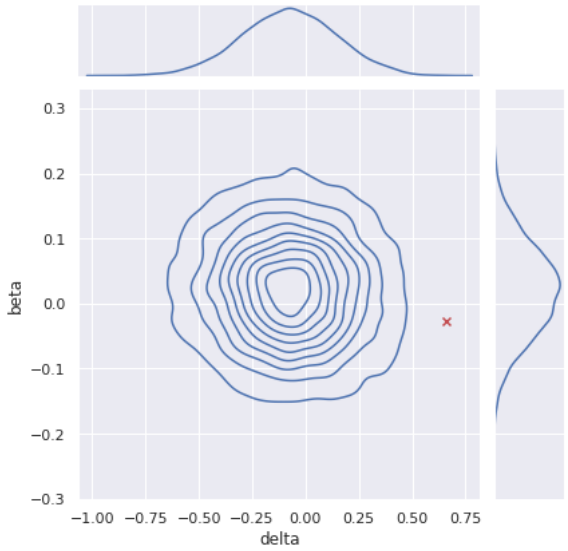
Es pertinente aclarar que multiplicamos por -1 los betas de los vectores donde delta es menor a cero, con esto logramos generar una pendiente negativa cuando los vectores se encuentran debajo del eje de las abscisas, y una pendiente positiva en el otro caso; notar que sigue siendo posible usar la densidad kernel porque a pesar de la transformación, la función sigue siendo continua.

Antes de hacer el análisis formal, podemos observar que, como esperábamos, los valores de los placebos se encuentran cerca del eje. También podemos detectar la formación de 2 grupos diferentes dentro de los tratados, el primero tiene betas y deltas positivos y al parecer, significativos, y el segundo no tiene deltas significativos, pero tiene betas negativos y significativos.

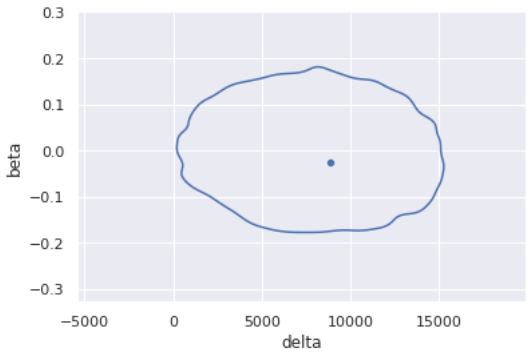
Como argumentamos en la sección de inferencia, vamos a obtener una estimación kernel de la distribución placebo de $\vec{v}_i$, la selección del hiperparámetro 'bandwidth' la hicimos usando 'cross validation':



Por último, para obtener el p-valor, usamos un método Montecarlo para estimar la distribución de la media de los distritos tratados bajo hipótesis nula, es decir, bajo hipótesis nula, los v_i de los distritos tratados se distribuyen según la densidad del gráfico anterior, entonces simulamos 1000000 de medias y obtenemos un p-valor igual a 0.00016:



Como podemos ver, el delta de los tratados es mayor, pero el beta es menor, esto se debe a la creación de estos dos grupos observada anteriormente donde un grupo tiene betas negativos y significativos; el hecho de que exista heterogeneidad en los resultados del análisis del control sintético es común, sin embargo, en nuestro caso, pudimos identificar estos dos grupos diferentes. Usando esta distribución sobre las medias, podemos obtener la distribución de probabilidad bayesiana sobre nuestra estimación de la media de los placebos:



Hicimos un corte al 95% sobre la distribución de la variable aleatoria $v_{\vec{tratados}} + \vec{V}$ para identificar nuestra seguridad sobre la estimación de la media. Podemos concluir que captamos un efecto generalizado en el aumento de las diferencias en los precios del corto plazo (delta) entre tratados y control, pero existen efectos heterogéneos en el largo plazo (beta).

9 Bibliografía

Climate Central, Land projected to be below annual flood level in 2050, online [07/07/2021], https://coastal.climatecentral.org/map/11/-0.0516/51.5343/?theme=sea_level_rise&map_type=coastal_dem_comparison&basemap=roadmap&contiguous=true&elevation_model=coastal_dem&forecast_year=2050&pathway=rcp45&percentile=p50&refresh=true&return_level=return_level_1&slr_model=kopp_2014

Cyclical Learning Rates for Training Neural Networks de Leslie N. Smith

Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program

Pooling Multiple Case Studies Using Synthetic Controls: An Application to Minimum Wage Policies http://ftp.iza.org/dp8944.pdf

Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. https://economics.mit.edu/files/11859

https://people.ischool.berkeley.edu/~hal/Papers/2013/ml.pdf - Hal R. Varian - Big Data: New Tricks for Econometrics - June 2013

Underwater assets? Real estate exposure to flood risk https://www.msci.com/www/blog-posts/underwater-assets-real-estate/01593224766

Benjamin J. Keys y Philip Mulder Neglected No More: Housing Markets, Mortgage Lending, and Sea Level Rise

Neglected No More: Housing Markets, Mortgage Lending, and Sea Level Rise de Benjamin J. Keys y Philip Mulder

Gallup https://news.gallup.com/poll/234314/global-warming-age-gap-younger-americans-worried.aspx

Does Climate Change Affect Real Estate Prices? Only If You Believe In It [Baldauf et al.] https://academic.oup.com/rfs/article-abstract/33/3/1256/5735306?redirectedFrom=fulltext

Rebecca K. Priestley, Zoë Heine, Taciano L. Milfont Public understanding of climate change-related sea-level rise. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0254348

Calendar, C.S. The artificial production of carbon dioxide and its influence on temperature. https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49706427503

Flooding in the future – predicting climate change, risks and responses in urban areas. R.M. Ashley*, D.J. Balmforth**, A.J. Saul* and J.D. Blanskby*

New data confirm increased frequency of extreme weather events: European national science academies urge further action on climate change adaptation." ScienceDaily, March 21, 2018. Citado por https://www.msci.com/www/blog-posts/underwater-assets-real-estate/01593224766

Liam F. Beiser-McGrath,Thomas Bernauer. Current surveys may underestimate climate change skepticism evidence from list experiments in Germany and the USA. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0251034