VNIVERSITAT DE VALÈNCIA

FACULTAD DE ECONOMÍA

PROGRAMA DE DOCTORADO EN ECONOMÍA INDUSTRIAL

CRECIMIENTO ECONÓMICO REGIONAL
DESARROLLOS METODOLÓGICOS Y APLICACIONES PREDICTIVAS

Tesis doctoral de:

Priscila Espinosa Adamez

Dirgida por:

Dr. José Manuel Pavía Miralles

Dr. Emili Tortosa-Ausina

Enero 2024

Dedicado a
mi familia

Agradecimientos

Me gustaría comenzar este trabajo expresando mi mas sincero agradecimiento a todas las personas e instituciones que me han acompañado durante este proyecto ya que sin su apoyo incondicional este trabajo no habría sido posible.

En primer lugar, deseo expresar mi profundo agradecimiento a mis directores de tesis, a los doctores José Manuel Pavía Miralles y Emili Tortosa-Ausina. Han sido no solo guías expertos, sino también pilares sólidos y faros de luz en los momentos más difíciles de este arduo camino. Agradezco enormemente que despertaran en mi una inquietud investigadora y forjaran las alas que hoy me permiten volar. Gracias por brindarme vuestros conocimientos, por orientarme, por animarme y, sobre todo, gracias por poner todo vuestro tesón y empeño en mi persona y en mi trabajo a cualquier hora del día o de la noche. Gracias por hacerme partícipe de vuestro equipo y guiarme hacia la meta esperada. A José Manuel, por su confianza y paciencia, por su colaboración y su saber hacer; a Emili, por su amabilidad y sus palabras de ánimo para seguir adelante. Además, me gustaría agradecer especialmente a la doctora Penélope Hernández por su apoyo en los inicios de esta dura etapa y por apostar ciegamente en mi.

En segundo lugar, agradecer a la Generalitat Valenciana, concretamente a la Conselleria de Hacienda, Economía y Administración Pública (previamente Conselleria de Economía Sostenible, Sectores Productivos, Comercio y Trabajo), la cual ha sido fundamental en este viaje académico. Destacar su colaboración gracias al convenio para el fomento y desarrollo del conocimiento de la economía valenciana y sus perspectivas en el corto y medio plazo, “Desarrollo de las previsiones macroeconómicas de la Economía Valenciana” y su compromiso con la innovación y la exploración de nuevas fronteras de conocimiento que han sido una fuente constante de inspiración.

También a los proyectos AICO/2021/257 de la Conselleria d'Innovació, Universitats, Ciència i Societat Digital y CIAICO/2023-GVRTE/2023/4572860 de la Conselleria de Educación, Universidades y Empleo por su apoyo constante a la ciencia. También agradecer el respaldo del Ministerio de Ciencia, Innovación y Universidades de España a través del proyecto PID2021-128228NB-I00 y al Programa Estatal de Promoción del Talento y su Empleabilidad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020, PTA2018-015997-I, ya que sin esta formación esto no hubiese sido posible.

Finalmente, quiero agradecer esta tesis a mi familia. A Carlos, mi compañero de vida, por ayudarme a crecer y avanzar, por su apoyo inquebrantable, por su comprensión y por no cuestionar mis locuras, sin él mis retos y proyectos no serían posibles. Porque no me imagino estar en otros proyectos de vida en los que no seas partícipe. A mis hermanos, Juanma y Adrián, por ser fuente constante de inspiración y apoyo y porque han sido mi referente en todo aquello que he conseguido. Porque ellos han sido mi vía de escape y de desahogo en muchos momentos. No podría tener unos mejores escuderos. A mis padres, Juan Manuel y Rosario, por darme todas las herramientas para hoy poder estar aquí, la base de mi educación y de mi espíritu guerrero, inconformista y trabajador. Gracias por vuestras muestras de apoyo, lucha, dedicación y esfuerzo que habéis demostrado siempre. Sin vosotros no podría haber sido lo que soy hoy. Todo os lo debo a vosotros y a vuestra confianza depositada en mí. Gracias infinitas.

Y, por último, también quiero agradecer a todas las personas que han contribuido de alguna manera en este trabajo, entre ellos, a mis compañeros de trabajo y amigos.

Resumen

Uno de los desafíos fundamentales a los que se enfrentan tanto agentes públicos como privados, en general, y decisores económicos, en particular, radica en la toma de decisiones en entornos de incertidumbre. Acontecimientos recientes, como la pandemia del COVID-19, o recurrentes en forma de conflictos, como la invasión de Ucrania, intensifican la complejidad de esta tarea, creando situaciones cada vez más dinámicas y cambiantes. En este contexto, la disponibilidad de metodologías y aplicaciones flexibles para el análisis de datos económicos y la generación de predicciones se vuelve de gran interés, especialmente para las y los analistas involucrados en el estudio de la realidad regional, que tienen en cuenta las características idiosincrásicas de la región.

Esta tesis doctoral aborda tres problemas clave en el desarrollo metodológico y predictivo del crecimiento económico regional.

Desde un enfoque metodológico, la investigación trata de manera innovadora el problema de la agregación de predicciones, buscando desde el marco conceptual de la Teoría de la Información la forma óptima de combinar las predicciones. En concreto, la novedad reside en abordar por primera vez el problema utilizando técnicas de aprendizaje automático basadas en regresión con regularización empleando una función basa en el concepto de entropía. El objetivo principal es superar las limitaciones de los métodos tradicionales especialmente cuando el número de organismos que realizan predicciones supera al número de estimaciones pasadas disponibles. El algoritmo se centra en combinar algoritmos de aprendizaje automático para minimizar la distancia y maximizar la información aportada por cada organismo, evitando el sobreajuste.

La segunda de las aportaciones de esta tesis, tiene un carácter mixto. Desde la perspectiva metodológica, se centra en, por una parte, determinar cómo combinar de manera

adecuada diferentes indicadores de alta frecuencia para la construcción de un índice sintético de crecimiento económico regional y, por otro lado, medir adecuadamente su incertidumbre asociada. Desde una perspectiva aplicada, se focaliza en la creación de una herramienta semi-automática basada en el análisis dinámico de factores para la construcción de índices sintéticos. Diversos desarrollos teóricos y empíricos respaldan la idoneidad de la metodología y el enfoque para medir y predecir la evolución agregada subyacente de la economía regional.

El tercer problema aborda, desde una perspectiva más aplicada que metodológica, la evaluación del impacto sobre el crecimiento de las regiones españolas de uno de los mayores paquetes de ayuda europea implementado hasta la fecha: los Fondos Next Generation EU (NGEU). Estos fondos buscan abordar los desafíos surgidos a raíz de la crisis del COVID-19 y tienen como objetivo transformar el sistema productivo favoreciendo una economía más ecológica, digital, resiliente y mejor adaptada a los retos actuales y futuros. Dado que estos fondos no han sido distribuidos de manera homogénea en todas las regiones europeas, ya que el impacto de la COVID-19 fue heterogéneo y las situaciones de partida de cada región son también distintas, es interesante estudiar cómo impactan en términos diferenciales en las diferentes regiones. Buscando determinar qué parte del crecimiento de estas regiones se debe a la evolución natural del ciclo y qué parte es atribuible a estos fondos, se realiza un estudio sobre el impacto de los fondos NGEU en el crecimiento económico agregado regional utilizando metodologías contrafactuales y proponiendo estrategias inferenciales para una evaluación más completa del impacto.

Abstract

One of the fundamental challenges faced by both public and private entities, in general, and economic decision-makers, in particular, lies in decision-making within environments of uncertainty. Recent events, such as the COVID-19 pandemic, or recurring conflicts, like the invasion of Ukraine, heighten the complexity of this task, creating increasingly dynamic and changing situations. In this context, the availability of flexible methodologies and applications for economic data analysis and prediction becomes of great interest, especially for analysts involved in studying regional realities, who must take into account the idiosyncratic characteristics of the region.

This doctoral thesis addresses three key problems regarding methodological tools and predictive applications for measuring regional economic growth.

From a methodological perspective, this research innovatively addresses the issue of prediction aggregation, seeking, from the conceptual framework of Information Theory, the optimal way to combine predictions. Specifically, the novelty lies in tackling the problem for the first time using machine learning techniques based on regularization regression employing a function based on the concept of entropy. The main goal is to overcome the limitations of traditional methods, especially when the number of entities making predictions exceeds the number of available past estimates. The algorithm focuses on combining machine learning algorithms to minimize distance and maximize the information contributed by each entity, avoiding overfitting.

The second contribution of this thesis has a mixed nature. From the methodological perspective, it focuses on determining how to appropriately combine different non-synchronic high-frequency indicators for the construction of a synthetic index of regional economic growth and, on the other hand, accurately measuring its associated uncertainty.

From an applied perspective, it centers on the creation of a semi-automatic tool based on dynamic factor analysis for constructing synthetic indices. Various theoretical and empirical developments support the suitability of the methodology and approach for measuring and predicting the underlying aggregate evolution of the regional economy.

The third problem addresses, from a more applied than methodological perspective, the evaluation of the impact on the growth of Spanish regions of one of the largest European aid packages implemented to date: the Next Generation EU (NGEU) Funds. These funds aim to address challenges arising from the COVID-19 crisis and seek to transform the production system, promoting a more ecological, digital, resilient, and better-adapted economy to current and future challenges. Since these funds have not been uniformly distributed across all European regions, given the heterogeneous impact of COVID-19 and differing initial situations in each region, it is interesting to study how they differentially impact various regions. Seeking to determine what portion of the growth in these regions is due to the natural evolution of the cycle and what part is attributable to these funds, a study on the impact of NGEU funds on regional aggregate economic growth is conducted using counterfactual methodologies and proposing inferential strategies for a more comprehensive assessment of the impact.

Índice general

1. Introducción **27**

1.1. Crecimiento económico regional. Desarrollos metodológicos y aplicaciones predictivas 29

1.2. Contexto 31

1.3. Objetivos 34

1.4. Metodología 35

1.4.1. Aprendizaje automático para combinar predicciones de PIB 35

1.4.2. Indicador sintético y la medición de la incertidumbre 36

1.4.3. Evaluación del impacto de los fondos NGEU en el crecimiento económico 37

1.5. Avance de resultados 38

2. Un algoritmo de aprendizaje automático basado en la entropía para combinar pronósticos macroeconómicos **43**

2.1. Introducción 47

2.2. Modelo 50

2.2.1. Modelo base: Inferencia de Máxima Entropía 50

2.2.2. Inferencia en el Aprendizaje Automático 52

2.2.3. Desde la Inferencia de Entropía Máxima hasta la Inferencia de Aprendizaje Automático 53

2.3. Algoritmo 54

2.4. Análisis de datos 57

2.5. Estudio de simulación 63
2.5.1. Configuración de la simulación 63
2.5.2. Resultados de la simulación 64
2.6. Conclusiones 67

3. Automatización en la construcción de índices económicos sintéticos regionales con medición de incertidumbre 71

3.1. Introducción 75
3.2. Antecedentes 77
3.3. Métodos 80
3.3.1. Índice Sintético 80
3.3.2. Incertidumbre 86
3.4. Índice Sintético: Aplicación web 89
3.5. Aplicación Empírica 91
3.6. Medición de la incertidumbre 95
3.7. ¿Cuántas remuestras son necesarias? Un análisis de sensibilidad 97
3.8. Resumen y observaciones finales 102

4. Estimando el impacto de los Fondos Next Generation EU a nivel regional en España: una aproximación basada en el método de control Sintético 105

4.1. Introducción 109
4.2. Estímulos económicos y NGEU 113
4.2.1. Multiplicadores fiscales y *fiscal spillovers* en la UE 114
4.2.2. NGEU y su impacto: evidencia 115
4.3. Estrategia empírica 117
4.3.1. El Método del Control Sintético: Metodología 119
4.3.2. Inferencia en SCM 122
4.3.3. Datos 124
4.4. Resultados 126
4.4.1. Midiendo los impactos 126
4.4.2. Inferencia 130

4.5. Conclusiones 134

5. Conclusiones y líneas futuras de investigación **139**

5.1. Conclusiones 141

5.2. Líneas futuras de investigación 146

Índice de tablas

2.1. Pesos óptimos ω_i^* generados por el algoritmo de máxima entropía y aprendizaje automático propuesto. 61
2.2. PIB, predicciones $\hat{a}_{T+1}$ y errores cuadráticos medios de predicción (RMSE). 62
2.3. Cambios relativos del error cuadrático medio y media aritmética de las predicciones simuladas de las instituciones ("*naive2*" en la Tabla 2.2) en comparación con el algoritmo propuesto. 66
2.4. Divergencia de Kullback-Leibler entre los pesos reales y los pesos recuperados. 67

3.1. Indicadores simples considerados inicialmente para construir el índice sintético. 92
3.2. Resumen estadístico de los modelos en la selección de subconjuntos. 94
3.3. Desviaciones típicas y errores relativos en intervalos de confianza al 90 % y al 10 %. 100
3.4. Desviaciones típicas y errores relativos en intervalos de confianza al 95 % y al 5 %. 101
3.5. Desviaciones típicas y errores relativos en intervalos de confianza al 97.5 % y al 2.5 %. 101

4.1. Variables y fuente de los datos. 125
4.2. Estadísticas descriptivas, variables relevantes. 126
4.3. Medidas descriptivas de las diferencias entre las estimaciones reales y contrafactuales en los períodos pre y post-tratamiento. 128
4.4. Efectos para las regiones españolas. 129

4.5. Distribución esperada de las diferencias post-tratamiento bajo la hipótesis de tratamiento no efectivo frente a la distribución real. 132
4.6. Modelo ANOVA para las diferencias post-tratamiento. 133
4.7. Prueba de hipótesis restringida. 134

Índice de gráficos

2.1. Pseudocódigo para el algoritmo propuesto. 57

3.1. Proceso iterativo en la generación de un indicador sintético. 80
3.2. Matriz de indicadores individuales, valores observados y predichos. 85
3.3. Captura de pantalla de la aplicación web Shiny para la construcción del índice sintético. 90
3.4. Serie estimada del índice sintético y serie del PIB de la Comunitat Valenciana. 95
3.5. Representación gráfica de la incertidumbre al estimar el indicador sintético mediante *bootstrap* con B=10000. 96
3.6. Intervalos de confianza (80 %, 90 % y 95 %) en función del número de remuestras del índice sintético en determinados momentos temporales. 99

4.1. Estimaciones reales y contrafactuales de la evolución del PIB per cápita para las regiones españolas. 127

Capítulo 1

Introducción

1.1. Crecimiento económico regional. Desarrollos metodológicos y aplicaciones predictivas

La predicción económica desempeña un papel fundamental en la toma de decisiones estratégicas por parte de los agentes económicos y en la formulación de políticas públicas, siendo esencial para comprender y anticipar el desarrollo económico de una región. En particular, el Producto Interior Bruto (PIB) regional es un indicador clave que refleja la salud y el dinamismo económico de una jurisdicción. La precisión en la predicción del PIB regional no solo dota de una visión clara de la evolución económica, sino que también facilita la planificación y ejecución de intervenciones económicas pertinentes y eficientes.

En el contexto actual, donde las jurisdicciones subnacionales se enfrentan a limitaciones de recursos y capital humano en un contexto más volátil que en épocas pasadas, la capacidad de disponer de herramientas potentes capaces de realizar predicciones económicas precisas cobra una relevancia aún mayor. La implementación de políticas efectivas, especialmente en tiempos de crisis, con continuos cambios o desafíos económicos significativos, requiere una comprensión profunda de cómo ciertas medidas y ayudas pueden incidir en el PIB regional. Es en este punto en el que se destaca la importancia de evaluar y prever el impacto económico de estas intervenciones, con el objetivo de optimizar la asignación de recursos y maximizar los resultados positivos para la región en cuestión.

En particular, cuando se abordan programas de apoyo económico, como el implementado a través del Programa Next Generation EU (NGEU) y el Mecanismo de Recuperación y Resiliencia (MRR), es esencial comprender cómo estas ayudas específicas afectarán al PIB nacional y especialmente al PIB regional. La evaluación de estos impactos no solo proporciona una visión clara de la eficacia de las medidas adoptadas, sino que también sirve como guía valiosa para la toma de decisiones futuras. Conocer cómo ciertos fondos y subvenciones influyen en la economía regional permite a las autoridades tomar decisiones informadas sobre la asignación de recursos, la implementación de reformas y la orientación de inversiones hacia áreas que generen un impacto económico positivo y sostenible.

En la presente investigación, se presentan metodologías, técnicas y herramientas semiautomáticas de predicción que no solo ayudan a reducir significativamente los costes

económicos, humanos y temporales asociados a la obtención de predicciones económicas regionales, sino que también ofrecen un enfoque eficiente y preciso sobre la evolución futura regional. Se proponen métodos de agregación basados en aprendizaje automático y métodos de predicción basados en técnicas de simulación, todo ello con el fin de perfeccionar el ajuste de las predicciones, aportando así mayor confiabilidad a los resultados obtenidos.

Estos nuevos desarrollos metodológicos se complementan, desde una perspectiva más aplicada, con una evaluación del impacto sobre el PIB regional de uno de los programas de mayor alcance en Europa, los fondos NGEU, centrados en la recuperación tras la crisis económica provocada por la crisis sanitaria del COVID-19. Este análisis proporciona una perspectiva valiosa sobre las implicaciones económicas derivadas de las intervenciones, ampliando la comprensión de los factores que inciden en la economía regional.

Cabe destacar que los resultados y hallazgos derivados de este trabajo han sido presentados y discutidos en diversos congresos nacionales e internacionales, así como difundidos (o estando en evaluación) a través de publicaciones científicas, las cuales se comentan en profundidad a lo largo del presente documento, quedando disponibles para su lectura en los capítulos siguientes.

La estructura de la presente tesis doctoral prosigue de la siguiente manera. Tras esta introducción, que delinea el marco conceptual, objetivos, metodologías y principales resultados, el siguiente capítulo aborda la combinación de predicciones mediante inferencia de entropía máxima, utilizando algoritmos de aprendizaje automático para medir distancias entre predicciones reales y estimadas. En el tercer capítulo, se destaca la importancia de anticipar el futuro económico de las regiones, con la construcción de un indicador sintético de crecimiento y la aplicación de técnicas de bootstrap para medir la incertidumbre asociada. En el cuarto capítulo, evaluamos el impacto económico del programa Next Generation EU en la economía regional española mediante técnicas contrafactuales. Finalmente, el quinto capítulo presenta las principales conclusiones y sugiere posibles líneas de investigación futuras.

1.2. Contexto

En el panorama económico actual, la incertidumbre emerge como un factor omnipresente que afecta, en nuestro contexto, a la toma de decisiones de los distintos gobiernos nacionales y regionales. La planificación estratégica, la formulación de políticas y disponer de mecanismos de predicción pueden ser parte de la solución en la reducción de la incertidumbre. La capacidad de prever y comprender el futuro económico se convierte en una herramienta crucial para navegar eficazmente por los complejos matices de los mercados y las condiciones económicas. Desde el colapso de Wall Street en 1929 hasta el día de hoy, las técnicas y medidas estadísticas utilizadas han experimentado una transformación significativa (Petropoulos *et al.*, 2022; Stigler, 1986). En este contexto, la importancia radica en reducir la incertidumbre y comprender el futuro económico, destacando cómo estas capacidades impactan en la toma de decisiones a nivel macroeconómico.

La incertidumbre, intrínseca a la economía global, nacional y regional, representa un desafío significativo para el avance económico y la promoción del desarrollo sostenible en el ámbito público. En el sector gubernamental, la incertidumbre se manifiesta como un riesgo que puede impactar en la eficacia de las políticas, la asignación de recursos y la estabilidad financiera del país. Los líderes y responsables públicos y privados se encuentran en una constante encrucijada, obligados a tomar decisiones estratégicas que influyen directamente en la dirección y el bienestar de la nación, región o empresa. La habilidad para mitigar la incertidumbre se rige como un factor clave para lograr una gestión gubernamental y empresarial informada y precisa, proporcionando así una ventaja competitiva en la formulación de políticas públicas y decisiones empresariales.

En este sentido, la predicción económica se convierte en una herramienta de gran valor. Los modelos predictivos, impulsados por técnicas avanzadas como el aprendizaje automático y el análisis estadístico, permiten a los organismos y a los agentes evaluar posibles escenarios futuros y adaptar sus políticas macroeconómicas y de decisión micro en consecuencia. En el ámbito de la formulación de políticas, los gobiernos y las instituciones financieras enfrentan el desafío de diseñar medidas efectivas para promover el crecimiento económico, la estabilidad y el bienestar general de la sociedad. En este contexto, la predicción económica se convierte en un pilar esencial para desarrollar políticas que sean

proactivas y alineadas con las tendencias económicas esperadas.

En el presente análisis, por tanto, convergen tres enfoques que abordan el crecimiento económico, su predicción y la incertidumbre inherente de la misma. Desde el ámbito de la ciencia de datos y el aprendizaje automático (*Machine Learning*), el segundo capítulo (que sigue a esta introducción) introduce un enfoque innovador que fusiona la inferencia de entropía máxima (Moreno y López, 2013) con algoritmos de aprendizaje automático (*ridge, lasso* y *elastic net*). El objetivo es generar predicciones agregadas que se adapten a la realidad, incorporando, mediante la introducción de un factor de penalización, la distancia entre las predicciones reales y estimadas. Esta metodología se valida en simulaciones y, más específicamente, se aplica a predicciones del PIB español, proporcionando resultados robustos y pesos óptimos para mejorar la toma de decisiones.

La segunda aportación de este estudio resalta la importancia crítica de la predicción en el ámbito macroeconómico, especialmente en un contexto contemporáneo marcado por la aceleración de los procesos sociales, la interconexión de los mercados y la respuesta hiperreactiva por parte de los agentes a cualesquiera de las señales (Ballester *et al.*, 2023; Khan *et al.*, 2022; Sifat *et al.*, 2022; Szczygielski *et al.*, 2022; Tao, 2022).

La importancia de predicciones informadas es crucial en los ámbitos macroeconómico y empresarial, donde la rapidez con la que los impactos se propagan entre los mercados ha aumentado debido a la aceleración de los procesos sociales y la hiperreactividad de los agentes. Eventos recientes como la pandemia del COVID-19, la guerra en Ucrania o la crisis global de la cadena de suministro resaltan esta aceleración, subrayando la necesidad de herramientas efectivas para la predicción económica y la medición de la incertidumbre.

En este contexto, la investigación, realizada en el tercer capítulo presenta una herramienta semi-automática para generar un índice sintético de una economía regional, centrándose en el caso de la Comunitat Valenciana. Asimismo, también aborda la medición de la incertidumbre asociada al índice sintético, utilizando técnicas de bootstrap (Efron y Tibshirani, 1986) para simular la distribución empírica de los estadísticos de interés, teniendo en cuenta la ideosincrasia de su construcción y su naturaleza dinámica. Para acabar este capítulo se discute acerca de la cantidad mínima de simulaciones necesarias para medir adecuadamente la incertidumbre.

La tercera aportación, presente en el capítulo cuarto, se enmarca dentro de la respuesta

europea a la crisis del COVID-19. En el segundo aniversario de la puesta en marcha del Mecanismo de Recuperación y Resiliencia (MRR), llevado a cabo por la Unión Europea, surge la estrategia Next Generation EU que tiene previsto inyectar hasta un total de 750 mil millones de euros. Este mecanismo, ha sido creado en respuesta a la pandemia del COVID-19 con el objetivo de transformar el sistema productivo para impulsar la aceleración hacia la recuperación y mejora de la resiliencia de los Estados miembros de la UE. El MRR, busca impulsar una transición hacia iniciativas ambientalmente sostenibles y digitalmente impulsadas, transformando a largo plazo la economía de la UE. Sin embargo, su implementación se desarrolla en un contexto dinámico, marcado por eventos como la incursión de Rusia en Ucrania, aumentos en la inflación y la crisis energética. En este marco, el capítulo cuatro analiza el impacto que están teniendo/tendrá estos fondos en los crecimientos económicos de las diferentes comunidades autónomas españolas.

Este estudio se distingue al adoptar un enfoque de un solo país y niveles territoriales más bajos de los empleados con este propósito hasta la fecha (Barbero *et al.*, 2023) empleando modelos contrafactuales, concretamente métodos de control sintético (Abadie, 2021; Abadie *et al.*, 2010; Abadie y Gardeazábal, 2003; Eli Ben-Michael y Rothstein, 2021). Estos métodos se caracterizan por ser más flexibles y presentar altos niveles de precisión al evaluar el impacto detallado de intervenciones específicas, como los fondos NGEU. Centrado en España, el capítulo cuarto busca proporcionar una evaluación detallada y precisa del impacto heterogéneo de los fondos NGEU en sus regiones, considerando la variabilidad en la implementación y asignación de fondos en un contexto descentralizado. Además, aborda las limitaciones comunes de los modelos de control sintético al proponer estrategias inferenciales que superan las pruebas conservadoras de significación estadística, ofreciendo así una evaluación más completa del impacto de los fondos NGEU en las regiones españolas.

Al entrelazar estas tres perspectivas, se subraya la importancia de abordar el crecimiento económico y su incertidumbre desde diferentes ángulos. Desde la ciencia de datos, donde la innovación en algoritmos mejora la precisión de las predicciones, hasta la reflexión histórica sobre los métodos utilizados para desentrañar el futuro, y finalmente, la implementación de estrategias supranacionales para guiar la recuperación económica. Se destaca la necesidad de emplear herramientas sólidas y adaptativas. En un mundo

interconectado y en constante cambio, estas perspectivas convergen para subrayar la importancia de la anticipación, la adaptación y la toma de decisiones informada para afrontar los desafíos económicos y sociales.

1.3. Objetivos

El presente trabajo tiene por objetivo proponer métodos y soluciones innovadoras para la generación de predicciones de alta calidad mediante diversas técnicas estadísticas. Más allá de las metodologías desarrolladas y que las aplicaciones generadas puedan ser empleadas por otros investigadores para abordar sus propias investigaciones, el objetivo primordial de esta tesis doctoral es el análisis y medición del crecimiento económico regional con el propósito de mejorar la toma de decisiones de los principales agentes públicos.

La esencia de este trabajo reside en el desarrollo metodológico y de aplicaciones predictivas que otorguen autonomía a los decisores políticos para comprender el estado actual de la economía. Este propósito no solo se materializa a través de la aplicación creada en este estudio, sino también en las múltiples aplicaciones derivadas del convenio de colaboración llevado a cabo entre la Conselleria de Hacienda, Economía y Administración Pública y la Universitat de València para fomentar y desarrollar el conocimiento de la economía valenciana y sus perspectivas en el corto y medio plazo, mediante la investigación para el desarrollo de un marco de previsiones macroeconómicas de la economía valenciana. Específicamente, se busca establecer enfoques metodológicos que faciliten la construcción de conjuntos de predicciones de alta calidad, contribuyendo así a disminuir la incertidumbre y avanzar en la comprensión de posibles cambios en el crecimiento regional, una tarea de especial importancia.

Toda investigación y avance científico se apoya en los trabajos previos de otros investigadores, y esta tesis doctoral no es una excepción. En la misma línea de aprovechar el conocimiento acumulado por los predecesores para seguir avanzando, la contribución de este trabajo se basa en asistir a los agentes económicos en la resolución recurrente de problemas a los que se enfrentan. La tesis doctoral tiene entre sus objetivos concretos:

- Proponer un nuevo algoritmo innovador capaz de combinar predicciones de manera automática, generando mayor precisión. Dicho algoritmo introduce una penalización automática a cada predicción individual basada en su grado de acierto histórico, incorporando así principios de aprendizaje automático para mejorar la capacidad predictiva. La combinación de máxima entropía y aprendizaje automático se traduce en un método que busca optimizar la ponderación de predicciones individuales de manera adaptativa, considerando la idoneidad de cada agente (organismo) generador de predicciones.

- Proponer la construcción de herramientas y procedimientos para la obtención de un índice sintético de alta frecuencia regional, con el propósito de capturar la evolución económica en tiempo real y prever su comportamiento futuro a corto plazo. Además, se mide la incertidumbre asociada al índice mediante simulación, que incorpora la variabilidad en la medición de los indicadores elementales y en el proceso de predicción univariante y de su proceso de agregación dinámico.

- Proponer métodos contrafactuales para la estimación del impacto económico en respuesta a shocks o perturbaciones, identificando qué parte de ese impacto es atribuible a cierta variable económica y aplicando métodos de inferencia causal, adaptándolo al caso del impacto del Plan de Recuperación, Transformación y Resilencia en el crecimiento económico de las comunidades autónomas españolas.

- Exponer los resultados obtenidos detallados en los puntos anteriores, y que se han concretado en diversas publicaciones científicas.

1.4. Metodología

1.4.1. Aprendizaje automático para combinar predicciones de PIB

Desde una perspectiva metodológica, en el segundo capítulo (que precede a esta introducción) se combina teoría de la información y aprendizaje automático. En concreto, se propone un modelo novedoso para optimizar la agregación de predicciones desde la perspectiva de la Teoría de la Información (Jaynes, 1957a,b) que es resuelto utilizando

técnicas propias de aprendizaje automático basadas en regresión con regularización. Este modelo utiliza la distancia de Kullback-Leibler como criterio para determinar los pesos óptimos de agregación (Fernández-Vázquez *et al.*, 2019), priorizando la distribución de probabilidad más uniforme (maximizando la entropía (Moreno y López, 2013)) sujeta a las restricciones proporcionadas por los datos observados. En la inferencia de máxima entropía, se busca la distribución menos sesgada y consistente con los datos disponibles, aplicando restricciones a la suma de los pesos y en términos de precisión de la predicción propuesta.

El objetivo de este enfoque es superar las limitaciones de los métodos tradicionales basados en el rendimiento pasado de los organismos, especialmente cuando el número de organismos supera el número de las predicciones pasadas disponibles para evaluación. El nuevo enfoque permite superar problemas de sobreajuste y obtener una determinación dinámica de la solución.

La propuesta se centra en la combinación de algoritmos de aprendizaje automático para minimizar la distancia, maximizando la información aportada por cada organismo a la par que se evita el sobreajuste. Se incorpora una metodología que combina la distancia relativa y las ponderaciones de las restricciones en el problema de optimización, dando lugar a un problema con restricciones. Se introduce un algoritmo que utiliza conjuntos de entrenamiento, validación y prueba, eligiendo el coeficiente de penalización mediante validación cruzada (Varma y Simon, 2006).

1.4.2. Indicador sintético y la medición de la incertidumbre

El tercer capítulo de este trabajo combina numerosas metodologías estadísticas, que incluyen técnicas de series temporales, análisis factorial dinámico, métodos de remuestreo y procedimientos de selección óptima, con el objetivo de medir de manera sintética el crecimiento regional. En concreto, se centra en la creación de un indicador sintético de alta frecuencia que, siendo generalizable a cualquier región, está diseñado para capturar la evolución económica agregada de la Comunitat Valenciana. El propósito es mejorar el conocimiento instantáneo (*nowcasting*) y prever el comportamiento económico futuro a corto plazo (*short-term forecasting*) de la Comunitat Valenciana. Se hace uso de la metodo-

logía de Dagum y Cholette (2006) para desagregar la serie anual, asegurando su relación con el PIB. Además, se evalúan distintos criterios de representatividad en la selección de indicadores partiendo de una selección inicial basada en juicio experto, que es refinada mediante selección automática de subconjuntos óptimos.

En el cálculo del indicador sintético, se seleccionan, tratan y completan series con métodos de predicción univariante. La estimación del indicador se realiza mediante análisis factorial dinámico, que además sirve para alinear temporalmente las series en términos de señal cíclica.

En la última fase, se mide la incertidumbre y se determina el número mínimo de remuestras para estabilizar la estimación del error. Se utiliza la simulación por *bootstrap* para generar un intervalo de confianza. Se observa que los límites no convergen en mismo número de remuestras en todos los momentos temporales y se realizan regresiones lineales para determinar el punto de estabilización del intervalo de simulación para cualquier instante.

1.4.3. Evaluación del impacto de los fondos NGEU en el crecimiento económico

El enfoque del cuarto capítulo es más aplicativo que metodológico. Desde la perspectiva metodológica combina métodos de control sintético (Abadie, 2021; Abadie *et al.*, 2010; Abadie y Gardeazábal, 2003; Eli Ben-Michael y Rothstein, 2021) con técnicas de contraste de hipótesis, las cuales son tratadas de manera novedosa para resolver una de las limitaciones prácticas observadas en la literatura de la identificación de impactos con el uso de técnicas contrafactuales. En concreto, el cuarto capítulo de la investigación se adentra en la evaluación del impacto de los fondos NGEU, una iniciativa diseñada para favorecer a los países más vulnerables, especialmente aquellos más afectados por la pandemia del COVID-19. A pesar de la envergadura del programa y su reciente implementación, son escasos los estudios que han abordado de manera integral su impacto económico. En el contexto español, se anticipa que estos fondos tendrán un impacto significativo.

El enfoque contrafactual, esencial para entender el impacto de intervenciones como las fomentadas por los fondos NGEU, se basa en imaginar escenarios hipotéticos. Estos

escenarios permiten comprender cómo podrían haberse desarrollado los acontecimientos si ciertos factores hubieran sido alterados, proporcionando un marco sólido para examinar las relaciones de causa y efecto.

Este enfoque metodológico permite una evaluación más detallada y específica del impacto del MRR en las regiones españolas. Los modelos contrafactuales, concretamente los métodos de control sintético imponen restricciones teóricas flexibles, lo que facilita la incorporación de la complejidad inherente a intervenciones específicas, como las relacionadas con los fondos Next Generation EU. Además, este tipo de modelos no requiere asumir un conocimiento completo de las relaciones causales en la economía, proporcionando así una evaluación más específica y detallada de cómo los fondos pueden afectar directamente variables relevantes.

Finalmente, uno de los desafíos prácticos inherentes al enfoque contrafactual se relaciona con las pruebas de significación estadística, las cuales tienden a ser relativamente conservadoras si atendemos a las que emplean comúnmente para evaluar las estimaciones (Chernozhukov *et al.*, 2013; Firpo y Possebom, 2018; Galiani y Quistorff, 2017). Estas pruebas a veces no revelan significación estadística, a pesar de que existe un claro efecto. En este contexto, se proponen estrategias de inferencia causal para abordar estas limitaciones y mejorar la evaluación de la significatividad en los modelos de control sintético.

1.5. Avance de resultados

Los resultados alcanzados de esta investigación abarcan diversas áreas, todas centradas en el análisis y la cuantificación del crecimiento regional con el objetivo de brindar información valiosa para respaldar la toma de decisiones económicas por parte de los agentes económicos.

En el segundo capítulo, cuyos resultados se validan mediante un estudio de simulación, se abordan diversos escenarios para evaluar el impacto de varios hiper-parámetros en la calidad de las predicciones. La evaluación se realiza mediante réplicas fabricadas del producto interior bruto de España que simulan series temporales objetivo. Se analiza el impacto de varios parámetros, como el número de instituciones, el tamaño de la muestra, el porcentaje de agentes clave y la desviación estándar del ruido. Los resultados alcanzados

indican que el algoritmo propuesto supera la media general del método **naive** basado en pesos constantes, mostrando mejoras significativas en el error de pronóstico. Además, la recuperación de pesos también mejora con el algoritmo propuesto, respaldando su eficacia en comparación con enfoques ponderados más complejos. Estos hallazgos concuerdan con la literatura existente sobre "la combinación de pronósticos" y establecen una conexión entre la teoría de la información y el aprendizaje automático en el contexto de la combinación de predicciones.

En el tercer capítulo, todos los avances metodológicos que se detallan se plasman en una aplicación informática propia (**app**) capaz de generar un indicador sintético que integra la información dinámica de diversos indicadores individuales mediante un análisis factorial dinámico. La creación de este índice sintético se complementa con un análisis detallado del cálculo de la incertidumbre, utilizando técnicas de simulación bootstrap. La aproximación utilizada para medir la incertidumbre se revela como una opción robusta, especialmente al combinar diversas series económicas a través del análisis factorial dinámico. Se concluye además que, en este contexto específico, es adecuado emplear 1000 remuestras para calcular de manera precisa los intervalos de confianza del índice sintético en distintos momentos temporales. Este enfoque no solo mejora la calidad de las predicciones, sino que también proporciona una herramienta valiosa para la toma de decisiones informada en el ámbito económico.

Los resultados del cuarto capítulo muestran que a pesar de las limitaciones inherentes a los modelos contrafactuales y los desafíos asociados con la construcción de las regiones donantes, el análisis exhaustivo del impacto de los fondos NGEU revela un efecto positivo generalizado en todas las regiones españolas. En término medio, las regiones habrían experimentado un 3 % más de PIB per cápita en 2022 gracias a estos fondos, y esta diferencia aumentaría en los años subsiguientes. Este impacto sostenido de crecimiento subraya la eficacia de la política en la promoción del desarrollo económico regional. En el análisis de pruebas inferenciales, se observa que antes del tratamiento no hay diferencias notables entre los valores registrados y el grupo experimental. Sin embargo, después del tratamiento, las series de valores reales muestran diferencias sistemáticamente significativas en comparación con los valores contrafactuales. Este hallazgo se confirma con análisis detallados, indicando que las diferencias son significativas a nivel individual y se amplifican a lo largo

del tiempo.

Capítulo 2

Un algoritmo de aprendizaje automático basado en la entropía para combinar pronósticos macroeconómicos

Resumen

Este estudio aplica un enfoque de aprendizaje automático con el objetivo de proporcionar una única predicción agregada de crecimiento a partir de un conjunto de predicciones individuales. Partiendo de la conocida metodología de inferencia de entropía máxima, se presenta un nuevo modelo que captura la distancia entre las predicciones agregadas reales y estimadas, planteando así un nuevo problema. Algoritmos como *ridge, lasso* o *elastic net* ayudan a desarrollar una nueva metodología para abordar y resolver este problema. Realizamos un estudio de simulación para evaluar el rendimiento de la nueva propuesta y lo aplicamos para predecir y medir la capacidad predictiva del modelo utilizando un conjunto de datos de predicciones sobre el producto interior bruto español[1].

[1]Los principales resultados que integran en este capítulo han sido publicados en el artículo "An entropy-based machine learning algorithm for combining macroeconomic forecasts", `https://doi.org/10.3390/e21101015`

2.1. Introducción

En este capítulo se aplica un enfoque de aprendizaje automático con el objetivo de proporcionar una única predicción agregada a partir de un conjunto de predicciones individuales. Partiendo de la conocida metodología de inferencia de entropía máxima, se propone un nuevo modelo a partir de la introducción de un nuevo factor que captura la distancia entre las predicciones agregadas reales y estimadas. Este nuevo modelo permite definir una nueva familia de problemas. La resolución de los problemas planteados por este factor adicional es resuelta empleando algoritmos de aprendizaje automático, como la regresión *ridge, lasso* o *elastic net*.

La contribución principal de este capítulo es un nuevo algoritmo que combina la inferencia clásica de entropía máxima con los principios de aprendizaje automático basados en regularización, mediante la aplicación de una penalización cuando la predicción agregada no coincide con el objetivo de la predicción. A través de un ejercicio de simulación, evaluamos el rendimiento del algoritmo y lo comparamos con un enfoque *naive* en el que las predicciones agregadas se construyen como media de las predicciones individuales. También aplicamos este algoritmo a un conjunto de datos de predicciones sobre el producto interior bruto español para producir pesos óptimos que luego se utilizan para realizar predicciones, cuya capacidad predictiva también se evalúa.

En la actualidad, existe un número creciente de fuentes y métodos prospectivos que ofrecen una amplia variedad de predicciones para una variable económica dada. Los métodos tradicionales de combinación de predicciones se basan en el rendimiento pasado de los organismos a través de sus pronósticos realizados. Sin embargo, el número de instituciones ha aumentado considerablemente en los últimos años, siendo los nuevos organismos tan recientes que no han dispuesto de tiempo suficiente para demostrar adecuadamente su capacidad predictiva.

La idoneidad en la combinación de resultados individuales para obtener una única predicción agregada no solo plantea problemas en el ámbito económico. En la Teoría Física, entendida como Mecánica Estadística, las obras fundamentales de Jaynes (1957a,b) establecen la conexión con la Teoría de la Información, sugiriendo un método para establecer la distribución de probabilidad con conocimiento parcial. Otra razón por la cual

un enfoque basado en la Teoría de la Información podría resultar apropiado para abordar el problema de la agregación de predicciones es de índole informativa. La expectativa racional sostiene que, eventualmente, los expertos deberían converger hacia la verdadera predicción. Después de un proceso de aprendizaje prolongado pero exitoso, se esperaría que los expertos realicen predicciones similares. Por lo tanto, una distribución uniforme sobre el conjunto de predicciones debería representar la combinación definitiva de las predicciones, maximizando así su entropía.

La literatura sobre aprendizaje automático en la combinación de predicciones es extensa e incluye, entre otros, enfoques como *bagging* (Breiman, 1996), *boosting* (Freund, 1995; Schapire, 1990) o la combinación de redes neuronales (Shnarch *et al.*, 2018). En el campo de la economía, la combinación de predicciones tiene una larga tradición y sigue siendo un área activa (ver, por ejemplo, Armstrong (2001); Chan y Pauwels (2018); Fernández-Vázquez *et al.* (2019); Timmermann (2013)). Las encuestas de perspectivas empresariales de ASA/NBER comenzaron a producir predicciones económicas agregadas en 1968, poco después de que Zarnowitz (1967) aportara evidencia sobre las ventajas de promediar varias predicciones de producción bruta (como se señala en Clemen (1989)).

Desde la teoría clásica, la combinación de resultados individuales para obtener una única predicción agregada se basa en un vector de pesos que calibra diferentes niveles de habilidad de los expertos (organismos/instituciones). Varias alternativas pueden considerarse para la combinación de predicciones que involucran diferentes grados de sofisticación. Por ejemplo, Bates y Granger (1969) considera una minimización de la varianza-covarianza; Newbold y Granger (1974) ofrece un método para calcular los pesos con el fin de minimizar la varianza del error de la combinación de predicciones. Otro método llamado "método de regresión" (Granger y Ramanathan, 1984) interpreta el vector de coeficientes de una proyección lineal como los pesos correspondientes. Esta línea de investigación tiene en cuenta el mismo problema de optimización al cambiar las condiciones de restricción. En este trabajo, presentamos nuestro modelo para el problema de optimización de la agregación de predicciones desde la perspectiva de la Teoría de la Información. Este modelo utiliza el criterio de la distancia de Kullback-Leibler para determinar los pesos de la agregación de predicciones. La naturaleza y los objetivos del problema mencionado consisten en combinar las predicciones tratando de mantener constante (uniforme) el conocimiento

proporcionado por cada una de ellas y verificar la verdadera predicción.

Bajo esta perspectiva, aplicamos un segundo enfoque, aprendizaje automático, que presenta un segundo problema de optimización. Nos inspiramos en algunos algoritmos de aprendizaje automático para sugerir una especificación que combine ambos objetivos: la expresión de distancia relativa y la parte de restricciones relacionadas con la verdadera predicción. Proponemos una nueva especificación que también introduce parámetros temporales relacionados con una estructura temporal arbitraria. Estos parámetros ponderan cada una de las divergencias entre la agregación de las predicciones y las predicciones reales. El problema de optimización resultante se asemeja al de la regresión con regularización (Hastie *et al.*, 2009), y proponemos resolverlo utilizando validación cruzada por bloques o anidada (Varma y Simon, 2006).

Las características empíricas del algoritmo propuesto se ilustran utilizando un conjunto de datos de predicciones sobre el producto interior bruto de España. El conjunto de datos utilizado en este estudio proviene de la Fundación de las Cajas de Ahorro, FUNCAS. Se trata de un conjunto de datos completo con un número suficiente de instituciones que realizan predicciones, lo que permite el uso del algoritmo propuesto. Utilizando este conjunto de datos, el algoritmo propuesto genera pesos óptimos que luego se utilizan para producir tanto predicciones como para evaluar la capacidad predictiva del mismo. Aunque el conjunto de datos concreto no nos permite desentrañar diferencias claras entre el algoritmo propuesto y un pronóstico *naive* (esto no es así en general como se muestra en el estudio de simulación), el algoritmo es robusto en el sentido de que la selección de predicciones realizadas en julio o diciembre conduce a resultados e interpretaciones similares.

Las diferencias entre el algoritmo propuesto y la predicción *naive* se exploran más a fondo en un estudio de simulación. Este estudio revela que, a medida que aumenta la longitud de la serie temporal objetivo, disminuye el número de instituciones de predicción y las instituciones con predicciones más afinadas que el resto se reducen y se distancian más del resto, el algoritmo propuesto se vuelve más adecuado que la simple media *naive*.

El resto del capítulo se organiza de la siguiente manera. En la Sección 2.2 presentamos el modelo. En la Sección 2.3 introducimos el algoritmo de aprendizaje automático aplicado al problema de inferencia de entropía máxima. En la Sección 2.4, el algoritmo se aplica a

un conjunto de datos de predicciones sobre el producto interior bruto de España, y en la Sección 2.5 el algoritmo es evaluado mediante un ejercicio de simulación. En la Sección 2.6 se presentan las observaciones finales.

2.2. Modelo

En esta sección se presenta, en primer lugar, el modelo para el problema de optimización de la agregación de predicciones desde la perspectiva de la Teoría de la Información. Este modelo activa el criterio de la distancia de Kullback-Leibler para determinar los pesos óptimos de la agregación de predicciones. Posteriormente se introduce un segundo enfoque, basado en aprendizaje automático, que permite definir un segundo modelo. Ambos modelos son integrados, y finalmente se describe la relación entre ambos enfoques.

2.2.1. Modelo base: Inferencia de Máxima Entropía

Dado un conjunto de agentes I, sea $\{y_{i,t}\}_{i\in I, t\geq 0}$ las predicciones realizadas para una variable económica en el periodo t realizadas en un momento anterior. Consideramos la combinación de los resultados individuales, ponderados por un vector de parámetros para cada posible pronóstico, denotado por ω_i. Los pesos ω_i se interpretan como los grados de experiencia/calidad de cada agente $i \in I$. Al asumir una distribución no degenerada de pesos, la predicción verdadera en el periodo t se denota como a_t, que verifica $\sum_{i\in I} \omega_i y_{i,t} = a_t$. El primer problema que abordamos es determinar los pesos ω_i de manera que la predicción verdadera se ajuste a la agregación de pronósticos.

Un problema paralelo que consideramos es la maximización de la entropía de la distribución de $\{\omega_i\}_{i\in I}$ sujeta a que el valor verdadero coincida con la agregación de predicciones para todos los posibles horizontes temporales t. Este problema de optimización se expresa de la siguiente manera:

$$\max_{\omega_i} \sum_{i \in I} \omega_i \log \omega_i^{-1} \tag{2.1}$$

$$\begin{aligned} \textit{sujeto a} \qquad & \textstyle\sum_{i \in I} \omega_i = 1 \qquad \omega_i \geq 0, \\ & \textstyle\sum_{i \in I} \omega_i y_{i,t} = a_t \quad \textit{para } t \geq 0 \end{aligned}$$

Esta metodología, conocida como inferencia de entropía máxima, es equivalente al problema de encontrar una distribución no negativa de pesos $\{\omega_i\}_{i \in I}$ que minimice la distancia de Kullback-Leibler entre dicha distribución y la distribución uniforme sobre el conjunto de los agentes, es decir, $\frac{1}{|I|}$. La distancia de Kullback-Leibler entre dos distribuciones p y q se define como $K(p,q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$. Cabe destacar que la distancia de Kullback-Leibler es siempre no negativa, pero no es una distancia propiamente dicha, ya que no verifica las propiedades simétricas ni triangulares. Este enfoque asegura un resultado no sesgado sobre el conjunto de agentes.

Formalmente,

$$\min_{\omega_i} \sum_{i \in I} \frac{1}{|I|} \log(\omega_i |I|)^{-1} \tag{2.2}$$

$$\begin{aligned} \textit{sujeto a} \qquad & \textstyle\sum_{i \in I} \omega_i = 1 \qquad \omega_i \geq 0, \\ & \textstyle\sum_{i \in I} \omega_i y_{i,t} = a_t \quad \textit{para } t \geq 0 \end{aligned}$$

Para resolver el sistema anterior, comenzamos con el Lagrangiano del problema. Cabe destacar que la cardinalidad de las variables en comparación con el conjunto de restricciones puede no ser suficiente para garantizar una solución única, o incluso la existencia de una solución. A pesar de poder caracterizar un conjunto de pesos posibles que minimizan la distancia relativa, es posible que no cumpla con la condición de predicción verdadera.

Este problema es bien reconocido en la literatura, ya que la complejidad de encontrar una solución adecuada aumenta con la cardinalidad de los parámetros y las condiciones; por lo tanto, es necesario utilizar algoritmos numéricos para determinar (si existe) el conjunto de candidatos de solución.

Con el fin de reducir la complejidad del problema, consideramos una nueva parametrización de $\{\omega_i\}_{i\in I}$. En particular, podemos reparametrizar $\omega_i = \frac{e^{x_i}}{\sum e^{x_i}}$ para $x_i \in \mathbf{R}$. Esto garantiza que $\sum_i \omega_i = 1$ con $\omega_i > 0$, al mismo tiempo que simplifica el problema de optimización al reducir el número de restricciones de tres a una.

Otra manera de abordar el problema original consiste en permitir un equilibrio entre las dos restricciones presentes en el Lagrangiano. Por un lado, la distribución de pesos debería minimizar la entropía relativa; por otro lado, este sistema debería generar una agregación de predicciones lo más cercana posible a la predicción verdadera.

2.2.2. Inferencia en el Aprendizaje Automático

La naturaleza y los objetivos del problema (2.2) radican en: (i) combinar las predicciones de un conjunto de instituciones, (ii) tratar de mantener constante (uniforme) el conocimiento (la información) proporcionado por cada una de ellas y (iii) verificar (acercarse tanto como sea posible) a un conjunto de restricciones. Desde esta perspectiva, nos inspiramos en algunos algoritmos de aprendizaje automático como *ridge, lasso* o *elastic net* (en los cuales el objetivo es minimizar una distancia, controlando el número de parámetros del modelo para evitar el sobreajuste; todo ello se regula mediante un parámetro que permite reescalar o determinar la importancia relativa de cada función de error de origen) para proponer una especificación que combine ambos objetivos: la expresión de distancia relativa y la parte de restricciones relacionada con las predicciones verdaderas. Presentamos una nueva especificación que introduce los parámetros δ_t relacionados con una estructura temporal arbitraria (por ejemplo, cada uno de estos parámetros puede depender de la distancia temporal de la restricción al período predicho o de la certeza disponible sobre el valor correspondiente de la restricción), con un parámetro λ que pondera las restricciones impuestas por la distancia entre la agregación de las predicciones y las predicciones reales. Dado que el problema está en $\mathbf{R}^n$ y nos encontramos en un espacio vectorial normado, es posible considerar una familia de normas.

Al reunir ambas expresiones y tener en cuenta que los parámetros $\{\omega_i\}$ están parametrizados, formalizamos el problema de minimización desde una perspectiva de aprendizaje automático de la siguiente manera:

$$\min_{\omega_i} \sum_{i \in I} \frac{1}{|I|} \log(\omega_i |I|)^{-1} + \lambda \sum_t \delta_t || \sum_{i \in I} \omega_i y_{i,t} - a_t || \tag{2.3}$$

La relación de la especificación propuesta con la literatura de aprendizaje automático se origina en la forma de la función objetivo (ecuación 2.3) y sus dos sumandos. El primero hace referencia a la divergencia de Kullback-Leibler: $\sum_{i \in I} \frac{1}{|I|} \log(\omega_i |I|)^{-1}$. El segundo corresponde (se asemeja) a un término de regularización flexible: $\lambda \sum_t \delta_t || \sum_{i \in I} \omega_i y_{i,t} - a_t ||$.

El parámetro de penalización λ se utiliza para seleccionar los pesos que minimicen la divergencia de Kullback-Leibler hacia una distribución uniforme y penalizar la magnitud de la desviación de la predicción ponderada del valor observado. Por un lado, cuando λ es igual a 0, no hay penalización por predicciones pasadas y el resultado es equivalente al modelo clásico sin restricciones temporales. Por otro lado, a medida que λ aumenta, el incumplimiento de las restricciones temporales gana peso y domina la ecuación (2.3). En este último caso, el problema puede entenderse como un problema de regresión ponderada, pero con coeficientes restringidos al ser positivos y sumar uno, y sin mostrar las limitaciones de los procedimientos tradicionales cuando el número de organismos es mayor que el número de restricciones temporales.

Los parámetros delta, δ, son una medida de mejora para la magnitud de la importancia que λ asigna al incumplimiento de las restricciones. En otras palabras, δ pondera la importancia relativa de la restricción de un año a otro.

2.2.3. Desde la Inferencia de Entropía Máxima hasta la Inferencia de Aprendizaje Automático

El problema (2.2) comparte en realidad la misma esencia que la minimización del problema (2.3). El primer problema es un problema de optimización con restricciones, y el segundo incorpora esta restricción a la función objetivo. La metodología para resolver el problema (2.2) se basa en el método de los multiplicadores de Lagrange. Específicamente, el problema con restricciones se transforma en una forma estructural donde tanto el objetivo como las condiciones restringidas se multiplican por parámetros que dependen del conjunto de restricciones. Al resolver las condiciones de primer orden de la función

lagrangiana, se obtiene el óptimo. La lagrangiana para el problema (2.2) se formula de la siguiente manera:

$$\mathscr{L} = \sum_{i \in I} \frac{1}{|I|} \log(\omega_i |I|)^{-1} + \sum_t \lambda_t (\sum_{i \in I} \omega_i y_{i,t} - a_t) \tag{2.4}$$

Cabe destacar que la solución $\{(\omega_i^*, \lambda_t^*)\}_{(i \in I, t)}$, si existe, reduce la segunda parte del Lagrangiano a 0, ya que las restricciones deben cumplirse y, además, minimizan la distancia relativa.

Ahora, asumamos una familia de problemas denominada $\mathscr{P}(\lambda)$. Al fijar λ, tenemos el siguiente problema de minimización:

$$\min_{x_i} \sum_{i \in I} \frac{1}{|I|} \log(\omega_i |I|)^{-1} + \lambda \sum_t \delta_t || \sum_{i \in I} \omega_i y_{i,t} - a_t || \tag{2.5}$$

Cuando la norma es la distancia absoluta y $\lambda_t = \lambda \delta_t$, ambos problemas, (2.5) y (2.2), coinciden. Si existe una solución para el primer problema (2.5), entonces dicha solución es un candidato para el segundo problema (2.2) para el λ_t específico. Solo las restricciones pueden no cumplirse en el problema (2.5) si esta distorsión permite reducir (si es posible) la entropía relativa con respecto a la distribución uniforme. Por lo tanto, bajo el supuesto de existencia de una solución, ambos problemas ofrecerán la misma clase de solución.

La consideración en el óptimo nos permite abordar el problema (2.5) desde otra perspectiva cuando, de hecho, el problema (2.2) no tiene solución o es demasiada compleja de encontrar. Los algoritmos y formas estructurales tomados del aprendizaje automático podrían ser una manera de abordar la solución desde un marco de aprendizaje automático.

2.3. Algoritmo

Esta sección propone un algoritmo para abordar el problema (2.5). El algoritmo propuesto encuentra una solución para el problema (2.5) de manera análoga a cómo lo hacen los algoritmos de regularización y aprendizaje automático. Los pasos principales del algoritmo incluyen la división de los datos en conjuntos de entrenamiento, validación y prueba, y la elección del coeficiente de penalización λ, mediante la validación cruzada

en conjuntos de validación. Los parámetros δ_t son exógenos. A continuación, se puede calcular el error de predicción del algoritmo en los conjuntos de prueba, y se estiman los valores de x_i utilizando todo el conjunto después de abordar los parámetros λ y δ, teniendo en cuenta que los parámetros $\{\omega_i\}$ están parametrizados. Los valores estimados se utilizan finalmente como pesos para combinar las predicciones individuales.

En cuanto a la validación cruzada, seguimos la literatura de aprendizaje automático de series temporales y proponemos el uso de la evaluación de origen en movimiento (Tashman, 2000), también conocida como evaluación de recalibración de origen en movimiento (Bergmeir y Benítez, 2012). Estas son formas de validación cruzada anidadas, o por bloques, que deberían proporcionar una estimación casi imparcial del error (Varma y Simon, 2006). Una vez seleccionado el número de instituciones (organismos) que se pueden usar para definir adecuadamente los conjuntos de entrenamiento, validación y prueba, podemos comenzar a resolver el problema de optimización. Como ya hemos notado, en nuestra versión actual del algoritmo, las instituciones deben ser las mismas en los conjuntos de entrenamiento, prueba y validación. Si esta condición no se cumple, el problema no estará bien definido. Para abordar este problema, en nuestra aplicación (ver Sección 2.4), la dimensionalidad de la base de datos inicial se redujo de 21 a alrededor de 10 instituciones que cumplen con la condición de tener datos disponibles para las tres fases. Esto nos proporciona tres conjuntos de muestreo de datos con alrededor de 10 instituciones para cada fase.

Como especificación potencial, seleccionamos una de las opciones disponibles: consideramos la norma cuadrática, una regresión de *ridge*, en la función objetivo, y agregamos un parámetro λ y los δque caracterizan la holgura del proceso. Por simplicidad, utilizamos $\delta_t = 1$, donde otorgamos igual importancia a todas las restricciones. Probamos diferentes valores de λ, tomados de una cuadrícula de valores, para encontrar el óptimo que minimiza la divergencia y penaliza la predicción combinatoria en comparación con el valor observado.

Los pasos del algoritmo propuesto se describen detalladamente en la Figura 2.1. La salida del algoritmo es una predicción para el período $T+1$ denotada por $\hat{a}_{T+1}$. Los requisitos para aplicar este algoritmo son: (i) las tres divisiones de conjuntos de datos mencionadas anteriormente (entrenamiento, validación y prueba), (ii) un conjunto de

valores discretos de λ entre 0 e infinito, (iii) un conjunto de valores de descuento δ, que ponderan la importancia relativa de la restricción de un año a otro y que enfatizan el parámetro λ, y (iv) una función de error de predicción. El algoritmo resuelve el problema de optimización en el subconjunto de entrenamiento para cada uno de los diferentes valores de λ y δ. Una vez que se resuelve el problema de optimización, obtenemos un conjunto de errores de predicción en el conjunto de validación, tantos como valores de λ. Posteriormente, mediante validación cruzada, seleccionamos el λ que minimiza este error de predicción. Gracias a esta selección, es posible obtener la mejor penalización en términos de error de predicción. Una vez obtenido el mejor λ, aplicamos el algoritmo en el conjunto de prueba y evaluamos su rendimiento. Obtenemos los ω_i que minimizan la función objetivo y una medida de su error de predicción.

Figura 2.1: Pseudocódigo para el algoritmo propuesto.

Algoritmo: Entropía basada en aprendizaje automático

Input:

Datos de pronóstico realizados por la institución i para el año n, $\{y_{i,n};\ i \text{ en } 1:I,\ n \text{ en } 1:N+1\}$ $(N \geq 2)$

Valores realizados, $a_{1:N}$

Conjunto de coeficientes de penalización, $\{\lambda_j,\ j \text{ en } 1:J\}$

Conjunto de coeficientes de descuento $\{\delta_{t,T},\ t \text{ en } 1:(N-1),\ T \text{ en } 2:N\}$

Función de error de pronóstico f

Output:

Predicción $\hat{a}_{N+1}$

Pseudocódigo:

Para n en $2:N$

Para j en $1:J$

Resolver para los pesos utilizando el subconjunto de entrenamiento $y_1,\ldots,y_{n-1}$:

Conjunto $\omega_{i,n,j} = \text{argmin}_{\{\omega_i\}} \sum_{i\in I} \frac{1}{|I|} \log(\omega_i |I|)^{-1} + \lambda_j \sum_{t=1}^{n-1} \delta_{t,n} || \sum_{i\in I} \omega_i y_{i,t} - a_t ||$

Determinar el error de pronóstico utilizando el conjunto de validación y_n:

Conjunto $e_{n,j} = f(a_n, \sum_{i\in I} \omega_{i,n,j} y_{i,n})$

Fin del bucle

Fin del bucle

Conjunto $j^* = \text{argmin}_j (N-1)^{-1} \sum_{t=2}^{N} e_{t,j}$

Conjunto $\lambda^* = \lambda_{j^*}$

Resolver para los pesos utilizando λ^* y el conjunto de datos completo:

Conjunto $\omega_i^* = \text{argmin}_{\{\omega_i\}} \sum_{i\in I} \frac{1}{|I|} \log(\omega_i |I|)^{-1} + \lambda^* \sum_{t=1}^{N} \delta_{t,N} || \sum_{i\in I} \omega_i y_{i,t} - a_t ||$

Conjunto $\hat{a}_{N+1} = \sum_{i\in I} \omega_i^* y_{i,N+1}$

2.4. Análisis de datos

Se utiliza un conjunto de datos de predicciones sobre el producto interior bruto español para ilustrar las características empíricas del algoritmo propuesto. El algoritmo produce pesos óptimos ω_i^* (Tabla 2.1) que se utilizan para generar pronósticos $\hat{a}_{T+1}$ (Tabla 2.2), cuya capacidad predictiva se puede evaluar. La capacidad predictiva del algoritmo propuesto

para este conjunto de datos para el último período es similar a la de otros algoritmos de predicción *naive*, en concordancia con el ejercicio de simulación de la Tabla 2.3.

El conjunto de datos utilizado en este estudio proviene de la Fundación de las Cajas de Ahorro, FUNCAS. La muestra abarca las predicciones económicas de diferentes instituciones desde 2000 hasta 2018. La muestra seleccionada contiene un total de 21 instituciones: Analistas financieros, Asesor, Bankia, BBVA, Caixabank, Cámara de Comercio de España, CatalunyaCaixa, CEEM-URJC, Cemex, CEOE, CEPREDE-UAM, ESADE, Funcas, ICAE-UCM, IEE, Instituto de Macroeconomía y Finanzas (Universidad CJC), Instituto Flores de Lemus, Intermoney, Repsol, Santander, Solchaga Recio & asociados). Cada organismo realiza dos predicciones al año, en julio y diciembre, tanto para el año en curso como para el siguiente. Por lo tanto, cada organismo predice cada año hasta 4 veces. Los paneles de predicción de FUNCAS son muy conocidos, con una destacada experiencia en investigación económica y por su trabajo exhaustivo en la recopilación de pronósticos a nivel regional y nacional. Además, FUNCAS proporciona esta información de forma gratuita (consultar `www.funcas.es`).

En este análisis de datos, se ha empleado una función de error cuadrático de predicción $f(x,y) = (x-y)^2$, junto con los siguientes valores para los parámetros del algoritmo: $\lambda \in \{1\times 10^{-4}, 2\times 10^{-4}, \ldots, 8\times 10^{15}, 9\times 10^{15}\}$ y $\delta_{t,T} = 1$ para todo t, T. Los problemas de optimización se han resuelto utilizando el software libre R versión 3.6.1 (R Core Team, 2018) y los algoritmos de optimización disponibles en la biblioteca nloptr, que sirve como interfaz para la biblioteca NLOPT (González-Rodríguez, 2019). Los algoritmos de NLOPT pueden ser globales o locales, basados en derivadas o sin gradiente, e incluyen el algoritmo del Lagrangiano aumentado, que utiliza algoritmos locales de optimización subsidiarios. Todas las optimizaciones se iniciaron con un punto de inicio uniforme.

Para ilustrar la aplicación del algoritmo, las Tablas 2.1 y 2.2 se centran en el subconjunto del conjunto de datos completo que incluye solo predicciones para cada año en julio del mismo año. Se podrían aplicar restricciones alternativas al conjunto de datos completo, como predicciones para cada año en diciembre del mismo año o predicciones para cada año en julio del año anterior. Estas restricciones alternativas conducen a características clave similares en términos de capacidad predictiva y pesos óptimos, como se describe a continuación.

Algunas características clave de los pesos óptimos ω_i^* generados por el algoritmo propuesto e incluidos en la Tabla 2.1 es la variación de pesos a lo largo de los años y entre las instituciones. Estas variaciones pueden ser sustanciales, pero también revelan algunas consistencias. Los años para los cuales la Tabla 2.1 informa sobre los pesos óptimos abarcan desde 2002 hasta 2018. En los primeros dos años de predicción, todos los pesos son insignificantes excepto uno, con esa única institución clave representando aproximadamente el 10 % del total de instituciones. Para los quince años siguientes, los pesos se distribuyen, produciendo entre el 20 % y el 60 % de instituciones clave. Las instituciones varían desde aquellas que reciben pesos óptimos significativos (por ejemplo, CatalunyaCaixa con un 100 % en 2002-2003 o IEE e ICO con aproximadamente un 75 % en 2004 y 2006, respectivamente) hasta aquellas que reciben pesos insignificantes. Algunas instituciones no se consideran en ciertos años. De las 21 instituciones iniciales en el conjunto de datos completo, solo 13 realizaron predicciones desde 2000, de las cuales solo 9 continuaron haciendo pronósticos al final de la muestra. Considerando años e instituciones conjuntamente, se obtienen dos grupos de instituciones: aquellas con pesos significativos (usando el 25 % como umbral, como CatalunyaCaixa, IEE e ICO) y el resto de las instituciones.

Al evaluar la capacidad predictiva de las predicciones $\hat{a}_{T+1}$ generadas mediante el algoritmo propuesto y presentado en la Tabla 2.2, como parámetros en el estudio de simulación, se han modificado algunos factores clave como parámetros en el estudio de simulación. Este estudio considera diversas combinaciones de diferentes parámetros (Tabla 2.3). Una combinación de parámetros que se asemeja a las características en los datos podría ser: (i) un 40 % de agentes clave, dado que aproximadamente la mitad de los pesos óptimos estimados en la Tabla 2.1 no son despreciables, es decir, ≥ 4 % (cabe señalar que la fracción de pesos no despreciables aumenta sustancialmente con el tiempo en los datos, aunque permanece constante en el estudio de simulación); (ii) 10 agentes de predicción (organismos), dado que el número de agentes disminuye de 13 en 2000 a 9 en 2018; y (iii) un tamaño de muestra de $T = 20$ años, con los datos que cubren diecinueve años (2000-2018). Según el estudio de simulación, esta combinación de parámetros parece tener el potencial de favorecer tanto al algoritmo *naive* como al propuesto, dependiendo del grado de variabilidad entre las predicciones. Una variabilidad de $SD = 0.2$ podría considerarse razonable para los datos, ya que los pronósticos para un año se realizan en

julio de ese mismo año. Esta cantidad de variabilidad produce un aumento medio del 3.21 % en el error cuadrático medio de predicción en comparación con el algoritmo *naive* en el estudio de simulación. Esto concuerda con las diferencias informadas en la Tabla 2.3 para los datos.

Tabla 2.1: Pesos óptimos ω_i^* generados por el algoritmo de máxima entropía y aprendizaje automático propuesto.

Institución	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Analistas Financieros	0.00	0.00	0.01	0.01	0.01	0.04	0.04	0.01	0.01	0.05	0.08	0.09	0.09	0.10	0.10	0.16	0.16
Bankia	0.00	0.00	0.01	0.01	0.01	0.04	0.04	0.01	0.01	0.05	0.08	0.08	0.08	0.07	0.08	0.11	0.11
BBVA	0.00	0.00	0.01	0.01	0.01	0.04	0.04	0.01	0.01	0.04	0.08	0.07	0.07	0.04	0.04	0.06	0.06
Caixabank	0.00	0.00	0.01	0.01	0.01	0.05	0.05	0.01	0.01	0.05	0.08	0.09	0.09	0.11	0.09	0.13	0.14
CatalunyaCaixa	1.00	1.00	0.03	0.04	0.06	0.17	0.17	0.36	0.37	0.33	0.08	0.10	0.10				
CEPREDE-UAM	0.00	0.00	0.01	0.01	0.01	0.05	0.05	0.01	0.01	0.04	0.08	0.07	0.07	0.04	0.04	0.05	0.05
Funcas	0.00	0.00	0.01	0.01	0.01	0.04	0.04	0.01	0.01	0.04	0.08	0.08	0.08	0.06	0.07	0.10	0.10
ICAE-UCM	0.00	0.00	0.01	0.01	0.01	0.03	0.03	0.01	0.01	0.04	0.08	0.08	0.08	0.06	0.05	0.07	0.07
ICO	0.00	0.00	0.13	0.64	0.74	0.26	0.26										
IEE	0.00	0.00	0.75	0.22	0.11	0.13	0.13	0.52	0.51	0.13	0.08	0.10	0.10	0.25	0.27		
Instituto Flores de Lemus	0.00	0.00	0.01	0.01	0.01	0.04	0.04	0.01	0.01	0.04	0.08	0.08	0.08	0.06	0.06		
Intermoney	0.00	0.00	0.01	0.01	0.01	0.04	0.04	0.01	0.01	0.08	0.08	0.08	0.08	0.08	0.09	0.12	0.12
Santander	0.00	0.00	0.02	0.02	0.02	0.07	0.07	0.03	0.03	0.12	0.08	0.09	0.09	0.12	0.12	0.19	0.19

Tabla 2.2: PIB, predicciones $\hat{a}_{T+1}$ y errores cuadráticos medios de predicción (RMSE).

Año	PIB	*Naive*	*Naive2*	Aprendizaje Automático
2000	5.20	4.01		
2001	4.00	3.00		
2002	2.90	2.09	2.01	2.30
2003	3.20	2.28	2.22	2.10
2004	3.20	2.81	2.77	2.74
2005	3.70	3.28	3.29	3.28
2006	4.20	3.37	3.39	3.30
2007	3.80	3.85	3.88	3.84
2008	1.10	1.74	1.69	1.76
2009	-3.60	-3.64	-3.57	-3.65
2010	0.00	-0.59	-0.52	-0.72
2011	-1.00	0.79	0.80	0.86
2012	-2.90	-1.69	-1.56	-1.60
2013	-1.70	-1.49	-1.48	-1.50
2014	1.40	1.19	1.20	1.18
2015	3.60	3.05	3.03	3.09
2016	3.20	2.85	2.83	2.88
2017	3.00	3.15	3.16	3.16
2018	2.60	2.79	2.82	2.84
Muestra RMSE		0.76	0.73	0.74

Nota: Los métodos utilizados fueron: la media aritmética de las predicciones realizadas por todas las instituciones (*naive*); el algoritmo propuesto (aprendizaje automático); y la media aritmética del subconjunto de predicciones utilizado para realizar los predicciones con el algoritmo propuesto (*naive2*).

2.5. Estudio de simulación

2.5.1. Configuración de la simulación

El estudio de simulación abarca una amplia gama de escenarios, cada uno evaluado mediante 30 réplicas. Cada réplica se construye utilizando el siguiente proceso. Inicialmente, se utiliza la información real del PIB de España que se va a prever en nuestro análisis de datos para obtener estimaciones de parámetros ($\hat{\mu}$, $\hat{\phi}$ y $\hat{\sigma}$) para un proceso autorregresivo estándar de orden uno, $y_t = \mu + \phi y_{t-1} + \sigma_\epsilon \epsilon_t$. Estas estimaciones se utilizan en cada réplica para generar una serie temporal objetivo simulada preliminar, $\{\tilde{y}_t^*\}$. Luego, este valor preliminar se utiliza para generar predicciones simuladas para cada institución, $\{\tilde{y}_{i,t}\}$, mediante la adición de ruido (se consideran todos los ruidos, es decir, ϵ_t, η_t y ε_t), son ruidos gaussianos estándar independientes con diferentes intensidades parametrizadas por su desviación estándar, es decir, $\tilde{y}_{i,t} = \tilde{y}_t^* + \sigma_\eta \eta_t$. Estos pronósticos simulados se agregan utilizando pesos simulados, $\{\tilde{\omega}_i\}$. Los pesos simulados dependen del número de agentes clave (instituciones) considerados. Para un 100 % de agentes clave, los pesos simulados se establecen como pesos iguales. Para un 40 % y un 10 % de agentes clave, ese porcentaje del total de instituciones se selecciona al azar y se les asignan pesos uniformes entre 0.5 y 1 de forma aleatoria. A las demás instituciones se les asigna un peso despreciable y todos los pesos se vuelven a escalar para sumar uno. Estos pesos simulados se utilizan para producir la serie temporal objetivo final simulada, $\tilde{y}_t = \sum_i \tilde{\omega}_i \tilde{y}_{i,t} + \sigma_\varepsilon \varepsilon_t$ (con σ_ε fijado en 0.1 para introducir cierta pero no mucha desviación de la agregación directa). El rendimiento del algoritmo para diferentes series temporales objetivo simuladas se analiza al variar los siguientes parámetros: el número de instituciones, el tamaño de la muestra, el porcentaje de agentes clave y la desviación estándar del ruido.

El número de instituciones o agentes toma valores de 10, 20 y 40. Los dos primeros valores son ligeramente inferiores y ligeramente superiores al número de instituciones en nuestro análisis de datos (Sección 2.4, Tabla 2.1). El tercer valor corresponde a un número ideal de instituciones. El tamaño de la muestra (T) toma valores de 6, 10 y 20. El primer valor coincide con las observaciones disponibles en nuestro análisis de datos, y los otros valores consideran horizontes razonables y deseables, respectivamente. Los porcentajes de agentes clave considerados son 10 %, 40 % y 100 %, siendo este último correspondiente

a todas las instituciones ponderando de manera igual en la generación de la serie temporal objetivo. La desviación estándar del ruido (SD), σ_η, toma valores de 0.1, 0.2 y 0.3. Mientras que los dos primeros valores son apropiados para predicciones a corto plazo (por ejemplo, predicciones para un año dado realizados en diciembre de ese mismo año), el último valor corresponde a predicciones más lejanas en el tiempo (por ejemplo, para un año dado, predicciones realizados en julio del año anterior).

2.5.2. Resultados de la simulación

Los resultados del estudio de simulación están en línea con lo esperado (Tablas 2.3 y 2.4). El algoritmo propuesto se vuelve más favorable que la media general del método *naive* a medida que aumenta la longitud de la serie temporal objetivo y disminuye el número tanto de instituciones como de instituciones clave. El estudio de simulación revela que el error cuadrático medio puede más que duplicarse al usar el algoritmo *naive* en lugar del propuesto. Además, aunque los resultados muestran un buen número de mejoras en el error relativo por encima del 20 %, los resultados negativos parecen detenerse en torno al 12 %.

Los resultados en términos de recuperación de pesos se presentan en la Tabla 2.4. Evaluamos la recuperación de pesos a través de la divergencia de Kullback–Leibler entre los pesos reales y los recuperados. Una pequeña divergencia de Kullback–Leibler entre estos pesos se relaciona con las mejoras identificadas por el estudio de simulación en el error de predicción resultante de la aplicación del algoritmo propuesto. Los resultados de la Tabla 2.4 concuerdan con los de la Tabla 2.3.

El denominado "puzzle de la combinación de predicciones" radica en la observación de que combinaciones simples de predicciones puntuales superan a combinaciones ponderadas más elaboradas en aplicaciones empíricas repetidas (Smith y Wallis, 2009). En esta línea, Smith y Wallis (2009) señalaron errores en la estimación de pesos en muestras finitas como la causa más probable. Más recientemente, Genre *et al.* (2013) establecen que "no concluiríamos que existe un fuerte argumento para considerar combinaciones distintas al peso igual como un medio para resumir mejor la información recopilada como parte de las rondas trimestrales regulares" de la Encuesta de Organismos Predictores. Nuestros

hallazgos están en sintonía con esta literatura. La concordancia se observa tanto desde la perspectiva empírica como en el estudio de simulación. Este acuerdo complementa la principal contribución de este capítulo al conectar la literatura de teoría de la información con la literatura de aprendizaje automático en el contexto de la combinación de predicciones. El éxito de la ponderación igual en la combinación de predicciones también puede atribuirse al hecho de que las instituciones de previsión tienden a formar un consenso bien informado, que se beneficia simultáneamente de un efecto de manada (Pons-Novell, 2003) y un efecto de sabiduría de las masas (Fleenor, 2006).

Tabla 2.3: Cambios relativos del error cuadrático medio y media aritmética de las predicciones simuladas de las instituciones ("*naive2*" en la Tabla 2.2) en comparación con el algoritmo propuesto.

Agentes clave		**Ruido** $SD = 0.1$ Tamaño muestra (T)			**Ruido** $SD = 0.2$ Tamaño muestra (T)			**Ruido** $SD = 0.3$ Tamaño muestra (T)		
	Agentes	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$
	10	9.054	4.412	18.990	30.949	38.425	54.184	47.125	104.228	103.081
10 %	20	-10.597	-4.592	-1.951	0.189	6.892	9.293	11.791	21.052	34.796
	40	-2.586	-5.273	-3.923	-2.806	-1.578	-0.214	-1.774	0.725	2.796
	Agentes	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$
	10	-7.664	-7.808	-3.638	-1.210	-3.582	-3.214	2.117	0.619	4.131
40 %	20	-9.622	-0.997	-4.204	-10.501	-7.674	-6.759	-8.960	-7.862	-4.530
	40	-6.705	-3.98	-6.479	3.712	-1.271	-6.835	-4.716	-5.989	-3.499
	Agentes	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$
	10	-5.369	-5.737	-6.963	1.667	-10.27	-6.739	-9.857	-9.534	-8.856
100 %	20	-11.972	-8.850	-3.873	-9.304	-12.018	-7.438	-11.344	-8.470	-6.108
	40	-10.179	-9.636	-5.620	-11.407	-8.597	-3.475	-9.157	-7.703	-7.479

Nota: Los cambios relativos se encuentran representados en porcentajes. El error cuadrático medio se obtiene como la media de errores a lo largo de los años. Los parámetros incluyen el número de instituciones o agentes, el tamaño de la muestra T (dimensiones internas de la subtabla), agentes clave y la desviación estándar del ruido (dimensiones externas)

Tabla 2.4: Divergencia de Kullback-Leibler entre los pesos reales y los pesos recuperados.

Agentes clave		**Ruido** $SD = 0.1$ Tamaño muestra (T)			**Ruido** $SD = 0.2$ Tamaño muestra (T)			**Ruido** $SD = 0.3$ Tamaño muestra (T)		
	Agentes	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$
	10	2.653	1.197	0.362	0.815	0.376	0.154	0.686	0.166	0.116
10%	20	2.964	2.378	1.438	1.958	1.316	0.574	1.663	0.930	0.286
	40	3.118	2.623	2.253	2.483	2.314	1.832	2.122	1.960	1.295
	Agentes	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$
	10	2.495	1.613	1.229	1.429	1.015	0.702	1.229	0.734	0.438
40%	20	2.638	2.232	1.143	1.497	1.083	0.960	1.208	0.986	0.871
	40	2.053	1.462	1.178	1.281	1.220	1.022	1.039	1.163	0.960
	Agentes	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$	$T = 6$	$T = 10$	$T = 20$
	10	1.647	1.346	0.357	0.639	0.356	0.068	0.371	0.253	0.035
100%	20	1.173	0.860	0.613	0.637	0.259	0.098	0.222	0.135	0.092
	40	0.856	0.731	0.392	0.423	0.304	0.325	0.294	0.294	0.155

2.6. Conclusiones

De acuerdo con las teorías de predicción y muestreo, los errores y varianzas de predicciones individuales pueden reducirse mediante la combinación de predicciones individuales. Los métodos tradicionales para combinar pronósticos se basan en evaluar el rendimiento pasado relativo de las instituciones que se van a combinar. Sin embargo, el

problema se vuelve indeterminado tan pronto como el número de instituciones es mayor que el número de resultados pasados. Para superar este problema, una alternativa es asumir un conjunto de pesos a priori y aplicar el principio de máxima entropía para obtener un conjunto de pesos a posteriori, sujeto a la restricción de que las predicciones combinadas sean iguales a los valores realizados. Desafortunadamente, este es un problema complejo que crece con la cardinalidad de las variables y no se garantiza encontrar una solución.

Con el fin de llegar a una solución dentro del marco de la teoría de la información, proponemos un enfoque nuevo para el problema e, inspirados en la literatura de aprendizaje automático, sugerimos una nueva especificación basada en regresión con regularización y un algoritmo para resolverlo. El nuevo enfoque siempre produce una solución y, además, es bastante flexible. Permite el uso de diferentes normas para medir las discrepancias entre las predicciones combinadas y los valores realizados para ponderar la importancia relativa de las discrepancias. Nuestro enfoque de regularización también tiene la ventaja de producir, como subproducto, los pesos asignados a las diferentes instituciones. Estos pesos podrían entenderse como una medida de la capacidad de los organismos y utilizarse como herramienta para decidir qué metodologías merecen más crédito.

Se podría introducir mayor flexibilidad en nuestro modelo. Por ejemplo, sustituyendo en la ecuación (2.3) los valores individuales de previsión por funciones de predicción (por ejemplo, ecuaciones de regresión). En este caso, los parámetros de tales funciones de predicción se estimarían simultáneamente durante el paso de validación cruzada. Esto nos permitiría aplicar nuestra propuesta en un solo paso cuando, por ejemplo, intentamos obtener una predicción para una economía regional a partir de un conjunto de previsiones nacionales, donde las estimaciones individuales no están disponibles. Podríamos sustituir las estimaciones regionales individuales (no disponibles) por una función parametrizada (por ejemplo, una ecuación de regresión dinámica) de los valores nacionales.

En nuestro algoritmo, hemos considerado una norma cuadrática (una penalización *ridge*) y una evaluación de origen móvil como estrategia de validación cruzada. Por supuesto, también son posibles otras penalizaciones (por ejemplo, *lasso* o *elastic net*), y del mismo modo, hay margen para implementar otros métodos de validación cruzada. Por ejemplo, podríamos omitir explícitamente el orden temporal de los datos en los conjuntos de entrenamiento y realizar una validación cruzada de dejar uno fuera. Al final, la impor-

tancia relativa de las predicciones más recientes puede incorporarse de manera implícita en nuestra especificación a través de los coeficientes δ.

En cuanto a nuestro estudio, como es práctica común, hemos utilizado las últimas cifras disponibles y confiables del producto interior bruto, todos los países elaboran varias versiones del PIB. Sin embargo, las cuentas nacionales se revisan regularmente a medida que se amplía la información estadística. Por ejemplo, en el caso de España, las estimaciones de cada año sufren tres revisiones hasta que se consideran definitivas (Cabrer-Borrás *et al.*, 2017) como valores realizados, a_t. Sin embargo, en nuestra opinión, esta no es la mejor estrategia para un "combinador" de predicciones macroeconómicos, pensamos que deberían usarse estimaciones preliminares. Las estimaciones preliminares (flash, aunque son las cifras más provisionales y menos fiables) son las más atractivas, reciben una gran atención. Por un lado, ocupan las portadas de los medios de comunicación y son las más analizadas, debatidas y comentadas. Por otro lado, y más importante aún, las estimaciones preliminares sirven como marco para la toma de decisiones por parte de los agentes económicos. Decisiones que pueden dar lugar a derechos y obligaciones: compromisos de estabilidad presupuestaria en la UE, límites de gasto del gobierno general, tamaño del déficit o deuda permitidos. Esto puede tener consecuencias significativas en los pesos que recibe cada institución.

La contribución fundamental de este capítulo consiste en establecer una conexión entre la metodología de inferencia de entropía máxima de la literatura de la teoría de la información y la regularización de la literatura de aprendizaje automático, con el objetivo último de combinar predicciones. Aunque se podría concebir la posibilidad de conectar algoritmos de combinación de predicciones que no empleen regularización (como el *boosting* o *bagging*) con la literatura de teoría de la información, no resulta inmediatamente claro cómo lograrlo. Esta inmediatez parece ser una de las ventajas de la regularización sobre otros algoritmos alternativos al momento de establecer la conexión entre la literatura de aprendizaje automático y la teoría de la información.

Capítulo 3

Automatización en la construcción de índices económicos sintéticos regionales con medición de incertidumbre

Resumen

Las jurisdicciones subnacionales, en comparación con las maquinarias de los países y las grandes instituciones, tienen menos recursos y capital humano disponible para llevar a cabo un seguimiento agregado actualizado de la economía (*nowcasting*) y para generar predicciones económicas (*forecasting*). Este capítulo presenta los resultados de nuestra investigación con el objetivo de facilitar la toma de decisiones económicas de los agentes públicos regionales. Por un lado, presentamos una *app* interactiva que, basada en el análisis dinámico de factores, simplifica y automatiza la construcción de indicadores económicos sintéticos y, por otro lado, evaluamos cómo medir la incertidumbre asociada al indicador sintético que construye. Los desarrollos teóricos y empíricos muestran la idoneidad de la metodología y el enfoque para medir y predecir la evolución agregada subyacente de la economía y, dada la complejidad asociada con la metodología de análisis dinámico de factores, la oportunidad de utilizar técnicas de *bootstrap* para medir el error de estimación asociado. También mostramos que cuando combinamos diferentes series económicas mediante análisis dinámico de factores, alrededor de 1000 remuestras son suficientes para calcular adecuadamente los intervalos de confianza del índice sintético en los diferentes momentos temporales[1].

[1]Los principales resultados que integran en este capítulo han sido publicados en el artículo "Automation in regional economic synthetic index construction with uncertainty measurement", https://doi.org/10.3390/forecast5020023.

3.1. Introducción

Desde tiempos inmemoriales, los humanos han buscado reducir la incertidumbre sobre lo que depara el futuro o incluso predecir directamente el futuro. La incertidumbre genera ansiedad y puede intensificar la sensación de amenaza potencial causada por una situación. No sorprende, por lo tanto, que nuestros antepasados hayan buscado prever el futuro en los sueños, el cielo o incluso las entrañas de los animales. Aunque estos esfuerzos por predecir el futuro no han disminuido, la forma en que buscamos respuestas a este desafío ha cambiado. Desde el colapso de Wall Street en 1929 hasta el día de hoy, las técnicas y medidas estadísticas utilizadas han experimentado una transformación significativa (Petropoulos *et al.*, 2022; Stigler, 1986).

La necesidad de predicciones informadas es especialmente intensa en los campos macroeconómico y empresarial, donde la toma de decisiones es fundamental para elegir una opción entre varias alternativas. La aceleración de los procesos sociales, la creciente interconexión de los mercados y la hipersensibilidad de los agentes ante todo tipo de señales han provocado un crecimiento sustancial en la velocidad con la que los impactos se transfieren entre los mercados (Ballester *et al.*, 2023; Khan *et al.*, 2022; Sifat *et al.*, 2022; Szczygielski *et al.*, 2022; Tao, 2022), como lo demuestran eventos recientes como la pandemia del COVID-19, la guerra en Ucrania o la ruptura global en la cadena de suministros.

Esto ha creado una necesidad aún mayor de contar con herramientas adecuadas para la predicción económica y la medición de la incertidumbre. No sorprende, por lo tanto, que en las últimas décadas haya habido un aumento notable de la conciencia acerca de la importancia de predecir el futuro económico, con el objetivo de poder reaccionar adecuadamente, tomar decisiones acertadas y hacerlo rápidamente (Chernis *et al.*, 2020; Chow *et al.*, 2023; Hall *et al.*, 2023).

La toma de decisiones incorrectas o tardías por parte de los agentes públicos y gubernamentales podría tener consecuencias graves para una región o país, con efectos en la sostenibilidad de la deuda o la continuidad del sistema de seguridad social (Antipa *et al.*, 2012). Como consecuencia de una predicción económica errónea o tardía, pueden surgir problemas en las finanzas públicas debido a defectos en el equilibrio fiscal, por ejemplo, provocando un aumento del gasto público en un momento del ciclo en el que sería óp-

timo hacer lo contrario. Una toma de decisiones deficiente podría generar la quiebra de los países (Hakura, 2020). En estos entornos, los responsables de la toma de decisiones públicas necesitan reaccionar rápidamente ante los cambios en el ciclo económico, y esto solo es posible si se comprende plenamente la realidad económica de la región o el país (Andreini *et al.*, 2023; Bitetto *et al.*, 2023; Kuck y Schweikert, 2021).

Desafortunadamente, en comparación con la maquinaria de los países y las grandes instituciones, que son capaces de desarrollar y gestionar métodos complejos, incluidos enfoques bayesianos y/o de aprendizaje automático (Dauphin *et al.*, 2022; Gil Martín *et al.*, 2019), las jurisdicciones subnacionales cuentan con menos recursos y capital humano para llevar a cabo un seguimiento actualizado y ajustado a corto plazo de la realidad económica (Chernis *et al.*, 2020). Dentro de este contexto, el objetivo de este estudio fue buscar una manera de facilitar la toma de decisiones económicas de los agentes públicos regionales.

Este capítulo presenta una herramienta (semi)automática que permite generar un índice sintético de la economía regional de la Comunitat Valenciana mediante la combinación dinámica de indicadores económicos individuales. También evalúa cómo medir la incertidumbre que inducen las predicciones de las series individuales en el proceso de construcción del índice sintético.

El índice sintético se obtiene a través de una aplicación web interactiva diseñada especialmente para ser utilizada con el conocimiento teórico mínimo y con su metodología de cálculo adaptada a las características idiosincráticas de la Comunitat Valenciana. Sin embargo, aunque la herramienta se haya diseñado pensando en la Comunitat Valenciana, su flexibilidad permite que otros analistas la utilicen para construir sus propios índices sintéticos.

Con la ayuda de la herramienta, los responsables de la política económica regional pueden actualizar fácil y autónomamente el indicador sintético tantas veces como sea necesario, utilizándolo como proxy del crecimiento económico regional para el momento actual y el futuro inmediato, para reducir su incertidumbre al tomar decisiones. Asimismo, la estimación de sus errores y sus intervalos de confianza, como medidas de incertidumbre, ayudan a limitar el rango de valores entre los cuales puede encontrarse la trayectoria de la evolución económica y el conjunto de escenarios posibles. Esto facilita una mejor evaluación de la situación y aumento de la capacidad de reacción de los agentes en la toma

de decisiones.

El resto del capítulo está estructurado de la siguiente manera. La Sección 3.2 esboza el problema y sus antecedentes. La Sección 3.3 detalla la metodología, especificando cómo se construye el índice sintético y la estrategia que seguimos para medir la incertidumbre. La Sección 3.4 presenta la aplicación web, mientras que la Sección 3.5 desarrolla un ejemplo. La Sección 3.6 está dedicada a medir la incertidumbre y la Sección 3.7 a responder a la pregunta de cuántas réplicas son necesarias. El capítulo se cierra con una discusión de las principales conclusiones.

3.2. Antecedentes

La estimación de la actividad económica de España y de sus regiones la realiza actualmente el Instituto Nacional de Estadística (INE). El INE ofrece una visión agregada de la evolución de la economía española de forma anual y trimestral a través de la contabilidad nacional. El INE también recopila la contabilidad regional de España para medir y describir la actividad económica regional. Sin embargo, esta última información se publica con un retraso significativo y solo de manera anual. Esto significa que los agentes económicos regionales se ven privados de una herramienta fundamental para el *nowcasting* y *short-term forecasting*.

Afortunadamente, a nivel regional existe un conjunto de series económicas (indicadores simples) de alta frecuencia (mensual y trimestral) que ofrecen una visión parcial de la evolución económica regional (Chernis *et al.*, 2020). Esta información puede aprovecharse para un análisis económico coyuntural (a corto plazo) mediante la construcción de un indicador/síntesis económico a través de la "agregación" de varios indicadores simples. Este índice puede utilizarse como una herramienta por sí misma o como un componente intermedio para derivar nuevas series sintéticas; por ejemplo, para estimar/predecir en "tiempo real" una macromagnitud económica específica, como el producto interior bruto trimestral. El PIB es el indicador macroeconómico sintético por excelencia, siendo un buen medidor de crecimiento que proporciona información sobre el nivel de desarrollo y bienestar de un país (Kuznets, 1955).

Esta magnitud macroeconómica, sin embargo, no está disponible en la naturaleza y

requiere la creación de varios componentes que, a su vez, no son fácilmente accesibles. Todo esto dificulta la tarea de identificar la situación económica de una región o país. Entre los estudios de investigación sobre este tema se encuentran los realizados por Camacho y Pérez-Quirós (2010), Cuevas *et al.* (2017) y Dauphin *et al.* (2022). Camacho y Pérez-Quirós (2010) propusieron, entre otras contribuciones, un modelo de predicción económica a corto plazo para prever el crecimiento de la Eurozona en tiempo real. Cuevas *et al.* (2017) sugirieron una metodología para estimar el PIB en tiempo real mediante modelos factoriales dinámicos que utilizan indicadores de actividad económica. Además, Dauphin *et al.* (2022) aplicaron modelos factoriales dinámicos y varios algoritmos de aprendizaje automático para realizar pronósticos inmediatos del crecimiento del PIB en varias economías europeas, combinando variables tradicionales como no tradicionales (como la calidad del aire y las búsquedas en Google).

Además, descendiendo al nivel regional, podemos mencionar a Gil Martín *et al.* (2019), Chernis *et al.* (2020) Kuck y Schweikert (2021). Gil Martín *et al.* (2019) propusieron un modelo para generar previsiones instantáneas del crecimiento del PIB de las regiones españolas mediante modelos bayesianos de factores dinámicos. Chernis *et al.* (2020) desarrollaron un modelo de factores dinámicos de tres frecuencias para la previsión inmediata del crecimiento del PIB de las provincias canadienses. Además, Kuck y Schweikert (2021) evaluaron, entre otros modelos, un modelo dinámico de factores de frecuencia mixta para prever el crecimiento económico en Baden-Württemberg (Alemania).

En la aplicación que desarrollamos, detallada en la siguiente sección, utilizamos un modelo de factores dinámicos como instrumento metodológico. Esto nos permite desarrollar tasas sintéticas que resumen las características comunes que contiene un conjunto de indicadores, combinando las señales de variables económicas adelantadas, retardadas y coincidentes.

El uso y la construcción de índices sintéticos como fuente de conocimiento de la evolución económica no es nuevo y se remonta al trabajo realizado durante la primera parte del siglo pasado por el National Bureau of Economic Research de los Estados Unidos (Burns y Mitchell, 1946). Por ejemplo, Bitetto *et al.* (2023) emplea factores dinámicos para desarrollar un índice capaz de identificar crisis bancarias y de deuda tanto en economías fuertes como en desarrollo, mientras que Liang *et al.* (2021) utilizó un modelo de factores

dinámicos para prever la inflación en China. En una línea similar, Anesti *et al.* (2022) y Mumtaz y Musso (2021) fueron un paso más allá al proponer, respectivamente, un modelo de factores dinámicos con parámetros variables del tiempo y la volatilidad estocástica para estudiar el impacto de la volatilidad tanto regional como global en la economía mundial, y un modelo de factores dinámicos mejorado mediante la incorporación de nuevas variables para perfeccionar el modelo estándar de factores dinámicos al realizar predicciones trimestrales del PIB con un trimestre de anticipación.

En este estudio, junto con la elaboración de un índice sintético, también abordamos la medición de su incertidumbre. Cuantificar la incertidumbre de una estimación/predicción la hace más flexible y, de alguna manera, sirve como medida de su precisión/corrección/robustez frente a la imposibilidad práctica de conocer los verdaderos valores objetivos. La incertidumbre, según la entendemos en este estudio, es una medida asociada a la dispersión existente alrededor de los valores predichos/estimados, la cual será aproximada, dada la complejidad de la metodología utilizada, mediante técnicas de *bootstrap*.

El método de simulación *bootstrap* proviene de Efron (1979), quien afirmó que, dada una muestra lo suficientemente grande, es posible generar muestras de reemplazo de la original para aproximar la distribución empírica de cualquier estadístico de interés. La idea es que dado que estas muestras se generan a partir de la muestra original comparten su función de distribución empírica, lo que permite derivar varianzas o intervalos de confianza del parámetro de interés con la precisión deseada.

En la literatura existen innumerables trabajos que aplican la metodología de *bootstrap* para la medición de la incertidumbre. Por ejemplo, Meyer *et al.* (1986) aplican técnicas de *jackknife* y *bootstrap* para estimar sesgo, errores estándar y distribuciones muestrales de tasas de crecimiento per cápita; Hasni *et al.* (2019) llevan a cabo una revisión y crítica de la literatura sobre los métodos de *bootstrap* disponibles; y Fresoli (2022), en el contexto de los modelos VAR (modelos autorregresivos vectoriales), analizan los impactos del modelo, los parámetros estimados y la distribución del error en las coberturas empíricas de los intervalos multidimensionales de Bonferrini pronosticados, obteniendo mayores rendimientos al considerar las incertidumbres de los parámetros y el orden del número de retardos.

En el contexto de la elaboración de índices sintéticos, nuestra investigación aborda

también la pregunta sobre el número mínimo de simulaciones *bootstrap* necesario para medir adecuadamente la incertidumbre asociada con cada uno de los valores de una serie sintética obtenida mediante análisis factorial dinámico. Según Efron y Tibshirani (1994), 50 o incluso 25 remuestras serían suficientes para obtener un buen estimador del error estándar. Sin embargo, Hesterberg (2011) afirma en su artículo "*Bootstrap*"que en poblaciones sesgadas, los intervalos de confianza generados con un pequeño número de remuestras por las técnicas de *bootstrap* son imprecisos, siendo necesario un número de remuestras *bootstrap* mayor de lo normalmente considerado óptimo. Dado que consideramos que la solución puede depender del nivel de confianza preestablecido para el intervalo, estudiamos empíricamente considerando hasta 10000 remuestras en qué punto los límites de los intervalos de confianza se estabilizan en función del número de remuestras.

3.3. Métodos

3.3.1. Índice Sintético

La construcción de un índice sintético depende de su propósito de creación y de los indicadores simples de los que está compuesto, así como de la combinación de estos; todos estos elementos interactúan dinámicamente para producir una solución (ver Figura 3.1).

Figura 3.1: Proceso iterativo en la generación de un indicador sintético.

En nuestro caso, el objetivo es construir un índice sintético de alta frecuencia (mensual) capaz de capturar la evolución económica agregada de la Comunitat Valenciana. La finalidad es incrementar, de manera inmediata y sintética, nuestro conocimiento acerca de la situación actual por la que está atravesando la economía regional (*nowcasting*) y prever su comportamiento futuro(predicción a corto plazo). Nuestros índices de referencia (es decir, sus componentes) deben estar relacionados con el PIB de la Comunitat Valenciana. Desafortunadamente, como se mencionó en la sección anterior, las series de PIB regional solo están disponibles anualmente y, como suele ser el caso en muchos países (Chernis *et al.*, 2020), con un retraso temporal significativo. Por lo tanto, teniendo en cuenta nuestro objetivo de desarrollar un índice mensual, el primer problema a abordar es cómo comparar/relacionar los posibles indicadores elementales con la serie de referencia. De las dos soluciones disponibles —agregar temporalmente las series mensuales y estudiar las relaciones de frecuencia anual o desagregar las series anuales y estudiar las relaciones de frecuencia mensual— hemos optado por la segunda solución. Las posibilidades de análisis son más amplias y los tamaños de muestra son mayores. Específicamente, utilizamos la metodología de desagregación temporal de alta frecuencia de Dagum y Cholette (2006), que interpola mensualmente la serie anual sujeta a la restricción de agregación anual.

La nueva serie mensual de PIB se utiliza en la parte empírica de este estudio para cuantificar/objetivar los criterios de representatividad y significancia en la selección de indicadores simples. Esto es menos restrictivo de lo que podría parecer a primera vista. Desde una perspectiva metodológica, no presupone nada sobre qué indicadores iniciales deben considerarse o cómo deben combinarse los indicadores seleccionados. Desde una perspectiva procedimental, la aplicación presentada en este estudio puede ser utilizada por cualquier investigador/analista para construir su propio índice sintético, utilizando sus propios indicadores elegidos según sus propios criterios. Y, desde una perspectiva sustantiva, los indicadores simples no se eligen para replicar/predecir la serie artificial de PIB construida, sino porque proporcionan información relevante sobre la evolución económica general. La introducción de los indicadores en el índice sintético elaborado es lo suficientemente flexible como para permitir que los movimientos de alta frecuencia de los indicadores simples informen sobre los cambios que están ocurriendo en el futuro económico. La serie mensual de PIB construida no contiene esta información.

Junto con el criterio de representatividad y alta significancia, en nuestra aplicación empírica, la selección inicial de indicadores refleja, a partir de una selección previa basada en el conocimiento experto, los criterios de disponibilidad, frecuencia y retraso. Los indicadores deben estar disponibles con poco retraso, el coste de obtenerlos debe ser bajo y su frecuencia de actualización debe ser alta, preferiblemente mensual.

Una vez elegido un conjunto inicial de indicadores, este puede ser refinado mediante métodos de selección de variables, dos de los cuales son los métodos de *"subset selection"* (incluyendo los métodos hacia adelante y hacia atrás) y los métodos de regularización (Hothorn *et al.*, 2006; Tibshirani, 1996; Zou y Hastie, 2005). Los métodos de *"subset selection"* se basan en la identificación del mejor grupo de variables que conjuntamente mejor predicen/explican la variable de respuesta, o al menos la identificación de un buen subconjunto. Los métodos de regularización *ridge, lasso* o *elastic net* seleccionan las características de manera indirecta, ajustando un modelo con todos los predictores incorporando una penalización que castiga los coeficientes (acercándolos a cero) asociados con las variables que aportan poca capacidad predictiva. En nuestra aplicación empírica, utilizamos un método completo de *"subset selection"*.

Una vez seleccionado el conjunto final de indicadores simples, comienza la fase de construcción del índice sintético. Existen múltiples métodos de agregación de indicadores (Cabrer Borrás, 2001), desde los más simples basados en una estructura de ponderación fija, determinada de manera exógena o según algún criterio univariante, hasta métodos más complejos basados en técnicas de componentes principales y análisis de factores. Sin embargo, muchos de estos métodos enfrentan el problema de la asincronía que presentan las series económicas en términos de su información cíclica. No todas las series se pueden considerar sincrónicas con el ciclo económico; encontramos series adelantadas, retardadas y coincidentes. Por lo tanto, es necesario fechar y retrasar/adelantar las series para sincronizarlas (Cuevas *et al.*, 2015). Algunos de los estudios que abordan este problema, en el marco de la construcción de índices sintéticos, son los de Mondéjar-Jiménez y Vargas-Vargas (2008), Domínguez Serrano *et al.* (2011) y Cuevas y Quilis (2012). Todos ellos proponen técnicas de agregación de indicadores a través de componentes principales y análisis de factores, ofreciendo una revisión de la literatura sobre los métodos de agregación existentes. Consulte también Doz y Fuleky (2020).

Una forma de evitar el problema del fechado (sincronización temporal) es utilizar métodos de análisis factorial dinámico. Las técnicas de análisis factorial dinámico incorporan la dinámica de las series en su búsqueda del conjunto de señales temporales comunes a todas las series. Específicamente, nuestra aplicación toma como referencia la investigación de Cuevas *et al.* (2015) y utiliza el análisis factorial dinámico para reducir la dimensionalidad del problema y descubrir un único factor latente (el primer factor dinámico) que representa las variables originales con la menor pérdida de información. El proceso completo se muestra en la Figura 3.1, que incluye la fase de tratamiento y predicción univariante de las series, abordada a continuación. El modelo implementado permite incorporar nuevas observaciones de indicadores simples según la disponibilidad de los datos.

Siguiendo el procedimiento descrito en la Figura 3.1, para el cálculo del indicador sintético, dado un conjunto inicial de series temporales, se selecciona un conjunto de N series temporales utilizando el método de *"subset selection"*. Estas series se completan hasta el mismo momento temporal mediante métodos de predicción univariante. Se corrigen previamente de efectos calendario y se ajustan estacionalmente. Varios análisis previos (no presentados en este documento) revelan que la predicción univariante de los indicadores simples ofrece mejores resultados y flexibilidad que la extensión predictiva del propio factor dinámico, ya sea realizada de manera exógena (predicción después de modelar univariantemente) o endógena, durante el proceso de su cálculo a través del análisis factorial dinámico.

En nuestra investigación, el tratamiento y corrección de efectos estacionales, valores atípicos y efectos de calendario de los indicadores simples se lleva a cabo utilizando la metodología descrita en JDemetra+ (Grudkowska, 2017). Específicamente, como complemento de la aplicación del indicador sintético, existe una aplicación alternativa que permite procesar cada serie utilizando la metodología TRAMO-SEATS (Maravall *et al.*, 2015) o X-13ARIMA-SEATS (US Census Bureau, 2017), basada en la identificación (definida por el usuario o de forma automática) del modelo autoregresivo integrado de media móvil (ARIMA) de la serie a tratar. Además, una ventaja de esta aplicación alternativa es la simplicidad de su uso, es que, además de los componentes habituales de corrección de calendario (efecto móvil de Semana Santa, ciclo semanal o año bisiesto), permite la incorporación de otros componentes construidos a partir del calendario festivo y comercial de la Comunitat

Valenciana (IVE, 2016). Al igual que la aplicación principal (para la construcción del indicador sintético), esta aplicación `https://apps.uv.es/app_extraccion_componentes/` (consultada el 17 de abril de 2023) también puede resultar de interés para otros analistas.

Es común que los indicadores elementales no estén todos disponibles en un momento específico. Difícilmente estarán todos disponibles hasta el momento actual (*nowcasting*) y ciertamente no lo estarán hasta el momento el cual queremos realizar una predicción (*short-term forecasting*). Por esta razón, imponemos una restricción con respecto al número máximo de meses que cada serie puede ser predicha (establecido en un máximo de 12 meses con respecto a la serie más retrasada). Cada serie se predice hasta que todas ellas se ubiquen en el mismo instante en el tiempo (ver Figura 3.2).

La Figura 3.2 muestra esquemáticamente la estructura de datos observados y predichos para cada serie; $x_{t,i}$ representa el valor observado del indicador i durante el instante t y $\hat{x}_{T-f_i+k,i}$ es la predicción asociada al indicador i correspondiente al instante $T-f_i+k$ (con $f_i \geq k$), siendo T el total de períodos (horizonte temporal), N el número de indicadores elementales y $f_i (\leq 12)$ el número de períodos no observados (meses) (a predecir) hasta que la serie i se completa hasta el instante T.

En nuestra aplicación, cada serie/indicador se completa/predice utilizando modelos ARIMA univariantes identificados automáticamente. El proceso de identificación (y predicción) del modelo se realiza utilizando la metodología de Box–Jenkins (ver, por ejemplo, Box *et al.* (2015)). La mayor ventaja proporcionada por este tipo de modelos, en comparación con otros más sofisticados, suele residir en su capacidad para generar predicciones óptimas para el futuro inmediato (Wan Ahmad y Ahmad, 2013). Por supuesto, el analista puede optar por utilizar un procedimiento alternativo de predicción trabajando con series ya completas/predichas antes de la construcción del índice sintético.

Una vez que se dispone de una matriz de datos completa, como la que se muestra en la Figura 3.2, el índice sintético (IS) se estima utilizando análisis factorial dinámico, como en Cuevas *et al.* (2015), después de aplicar una serie de transformaciones a las series para eliminar las diferencias de escala y asegurar que sean estacionarias con una distribución normal aproximada. Específicamente, la *app* realiza en cada serie de indicadores: (i) una transformación logarítmica para garantizar la normalidad y simetría de los datos; (ii) una diferenciación para eliminar los componentes de tendencia de las series; y (iii) una

estandarización de las variables para homogeneizar sus escalas y varianzas.

Figura 3.2: Matriz de indicadores individuales, valores observados y predichos.

	Indicador					
Tiempo	1	2	...	*i*	...	N
1	$x_{1,1}$	$x_{1,2}$	⋮	$x_{1,i}$	⋮	$x_{1,N}$
2	$x_{2,1}$	$x_{2,2}$	⋮	$x_{2,i}$	⋮	$x_{2,N}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
t	$x_{t,1}$	$x_{t,2}$	⋮	$x_{t,i}$	⋮	$x_{t,N}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	$x_{T-f_1,1}$	⋮	⋮	$x_{T-f_i,i}$	⋮	⋮
⋮	$\hat{x}_{T-f_1+1,1}$	⋮	⋮	$\hat{x}_{T-f_i+1,i}$	⋮	$x_{T-f_N,N}$
⋮	⋮	$x_{T-f_2,2}$	⋮	$\hat{x}_{T-f_i+2,i}$	⋮	$\hat{x}_{T-f_N+1,N}$
⋮	$\hat{x}_{T-1,1}$	$\hat{x}_{T-f_2+1,2}$	⋮	⋮	⋮	⋮
T	$\hat{x}_{T,1}$	$\hat{x}_{T,2}$	⋮	$\hat{x}_{T,i}$	⋮	$\hat{x}_{T,N}$

Nota: Los valores observados de cada serie están sombreados en gris. Los valores predichos, alcanzados para completar todas las series en el mismo mes, están sombreados en rojo oscuro.

El análisis factorial dinámico permite expresar cada serie como una función lineal de un conjunto $r \ll N$ de factores comunes, que capturan los patrones de movimientos conjuntos compartidos por todas las series, y algunos efectos idiosincráticos específicos de cada serie. En nuestra aproximación $r = 1$. Matemáticamente, las relaciones se pueden expresar para $t = 1, \ldots, T$ a través de las ecuaciones (3.1) y (3.2).

$$\boldsymbol{x_t} = \boldsymbol{\lambda}(\boldsymbol{L})\boldsymbol{f_t} + \boldsymbol{\varepsilon_t} \tag{3.1}$$

$$f_t = \Psi(L) f_{t-1} + \eta_t \tag{3.2}$$

donde $\boldsymbol{x_t}$ y $\boldsymbol{\varepsilon_t}$ son vectores de dimensión $N \times 1$ (siendo N el número de indicadores y $\boldsymbol{\varepsilon_t}$ el vector de componentes específicas o idiosincráticas de los indicadores en el tiempo t), f_t es el vector factorial de r dimensiones que captura los valores del factor en el instante t, η_t es un vector ruido blanco de r dimensiones y $\lambda(L)$ y $\Psi(L)$ son polinomios de retardo matriciales de dimensiones $N \times r$ y $r \times r$, respectivamente, donde L es el operador de retardos.

Por un lado, el polinomio de retardo $\lambda_i(L)$ de la i-ésima fila de $\lambda(L)$, se denomina carga del factor para la i-ésima serie y $\lambda_i(L)\ f_t$ es el componente común de la i-ésima serie. Los polinomios de carga $\lambda_i(L)$ miden la sensibilidad de la señal de crecimiento del i-ésimo indicador ante cambios en el factor. Por otro lado, la matriz de retardo polinómico $\Psi(L)$ captura la estructura dinámica de los factores comunes.

El sistema anterior se estima bajo las asunciones clásicas de estacionariedad para las ecuaciones (3.1) y (3.2), asumiendo que las componentes idiosincráticas no están correlacionadas con todos los factores en todos los retardos, es decir, $E\left[\varepsilon_{t,i}\eta'_{t-k,j}\right] = 0\ \forall i,j,k$. Una vez que el sistema ha sido estimado, el modelo factorial dinámico nos proporciona la comunalidad correspondiente al primer factor; a partir de la cual derivamos mediante la ecuación (3.3) el índice sintético en base uno, donde $IS_1 = 1$.

$$IS_{t+1} = IS_t(1 + f_t) \tag{3.3}$$

Va

3.3.2. Incertidumbre

El segundo de nuestros objetivos es generar medidas de precisión del índice sintético para medir su incertidumbre. Desde una perspectiva frecuentista clásica, la incertidumbre de un estadístico se mide teniendo en cuenta la incertidumbre asociada a la muestra utilizada para su construcción. Sin embargo, en muchas ocasiones, como en macroeconomía, no es posible tener diferentes muestras de la misma variable o repetir un estudio en un gran número de ocasiones. En estas circunstancias, la incertidumbre se mide en función de las suposiciones del modelo que se utiliza para resolver el problema o después de asumir una cierta estructura generadora para los datos, como ocurre en el contexto del modelado de series temporales ARIMA.

La construcción de un índice sintético utilizando análisis factorial dinámico, sin embargo, implica operaciones matemáticas muy complejas que son difíciles de rastrear, lo que impide transferir analíticamente las estructuras de error a través del sistema. En estos contextos, así como en general, es posible utilizar técnicas de simulación para resolver el problema: técnicas de remuestreo de las cuales podemos obtener estimaciones de errores

estándar, intervalos de confianza o resolver pruebas de hipótesis. En esta investigación, utilizamos la metodología de bootstrap (ver, por ejemplo, Efron y Tibshirani (1986)).

Según Efron y Tibshirani (1994), la relación entre la muestra original y el *bootstrap* se explica mediante la conexión entre dos mundos, el mundo real y el mundo *bootstrap*. En el mundo real, la distribución de probabilidad multivariada conjunta, F, del fenómeno en estudio es desconocida y se dispone de una muestra aleatoria, $\boldsymbol{x} = (x_1, \ldots, x_n)$, que se utiliza para estimar una estadística de interés $\hat{\theta} = s(\boldsymbol{x})$. En el mundo *bootstrap*, se obtiene la distribución empírica $\hat{F}$ de los datos remuestreando con reemplazo la muestra observada (o los errores estimados) del mundo real, $\boldsymbol{x}^* = (x_1^*, \ldots, x_n^*)$, utilizando esta estimación de la distribución empírica para obtener parámetros de interés que serían imposibles de obtener en el mundo real.

En nuestro caso, el análisis factorial dinámico es un proceso determinista: dados los mismos inputs, siempre se obtiene la misma salida. En otras palabras, condicionado a los valores observados de los indicadores elementales, el índice sintético es invariable: no está sujeto a ninguna incertidumbre. Sin embargo, hay dos fuentes de incertidumbre que podrían introducirse en el proceso. Por un lado, los indicadores elementales son aproximaciones a la realidad y están sujetos a errores de medición. Por otro lado, como se mencionó anteriormente, los objetivos de *nowcasting* y *forecasting* impelan el uso de predicciones univariantes para completar las series de indicadores (ver Figura 3.2). Este es un proceso sujeto a incertidumbre, que también aumenta a medida que aumenta el horizonte temporal de predicción.

Aunque en esta investigación no incluimos la incertidumbre asociada a la primera fuente, ya que según un enfoque basado en el modelo de Valliant *et al.* (2000) se consideran fijos los valores observados, para aquellos investigadores interesados podemos señalar al menos dos estrategias para su inclusión. Por un lado, las series observadas podrían perturbarse localmente, manteniendo la estructura de dependencia de los datos, utilizando el bootstrap por bloques (ver, por ejemplo Veres-Ferrer y Pavía (2021)). Por otro lado, los errores asociados al modelo ARIMA de los indicadores podrían utilizarse para generar nuevas series como composición de los valores observados y errores remuestreados.

La segunda fuente de incertidumbre (aquella asociada con la predicción univariante de cada indicador) se incorpora al proceso y, dada la estructura dinámica del factor construido,

termina afectando a todos los instantes temporales del índice sintético, no solo a los instantes en los que se ha realizado una predicción. La incertidumbre asociada con las predicciones específicas de cada serie de indicadores se incorpora construyendo nuevas series como una concatenación de los valores observados de la serie y réplicas de las predicciones obtenidas mediante muestreo en sus intervalos de predicción asociados.

Matemáticamente, continuando con la notación introducida en la subsección anterior, el proceso parte de los valores observados $(x_{1,i}, x_{2,i}, \ldots, x_{t,i}, \ldots, x_{T-f_i,i})$ y varía las predicciones $\hat{x}_{T-f_i+k,i}, k = 1 \ldots f_i$ de cada serie en cada remuestreo. Para hacer esto, denominando $\hat{x}^-_{T-f_i+k,i}$ y $\hat{x}^+_{T-f_i+k,i}$ a los valores mínimo y máximo (obtenidos, por ejemplo, de un intervalo de predicción con una confianza dada) entre los cuales se encontraría razonablemente la predicción para el instante $T - f_i + k$ de la i-ésima serie, extraemos aleatoriamente un valor $x^*_{T-f_i+k,i}$ del intervalo $[\hat{x}^-_{T-f_i+k,i}, \hat{x}^+_{T-f_i+k,i}]$, para $k = 1, \ldots, f_i$, y construimos una nueva serie de la siguiente manera: $x_{1,i}, x_{2,i}, \ldots, x_{t,i}, \ldots, x_{T-f_i,i}, \ldots, x^*_{T-f_i+k,i}, \ldots x^*_{T,i}$. Este proceso se repite B veces para cada una de las N series. Específicamente, para $b = 1, \ldots, B$ e $i = 1, \ldots, N$, $\boldsymbol{x}_i^{(b)} = (x_{1,i}, x_{2,i}, \ldots, x_{t,i}, \ldots, x_{T-f_i,i}, \ldots, x^{*(b)}_{T-f_i+1,i}, \ldots, x^{*(b)}_{T,i})$ denota el b-ésimo remuestreo de la i-ésima serie de indicadores y $\boldsymbol{X^{(b)}} = \left[\boldsymbol{x_1^{(b)}, x_2^{(b)}, \ldots, x_i^{(b)}, \ldots, x_N^{(b)}}\right]$ denota la matriz de indicadores de orden $T \times N$ correspondiente al b-ésimo remuestreo.

En este punto, tenemos B matrices, $\boldsymbol{X^{(b)}}$ $b = 1, \ldots, B$, de indicadores simples. El procedimiento de construcción del índice sintético descrito en la subsección anterior se aplica a cada una de estas matrices, obteniendo B series de indicadores sintéticos: $IS^{(1)}, IS^{(2)}, \ldots, IS^{(b)}, \ldots, IS^{(B)}$. Todas estas comparten la misma distribución de probabilidad, por lo que se pueden considerar como una muestra aleatoria simple, $\boldsymbol{IS}^* = (IS^{(1)}, \ldots, IS^{(b)}, \ldots, IS^{(B)}) \overset{\text{iid}}{\sim} \hat{F}_{IS}$, de las cuales podemos extraer las estimaciones del índice sintético medio, el error estándar asociado a cada instante de temporal o construir intervalos de confianza para el valor del índice sintético, $t = 1, \ldots, T$.

En particular, denotando $IS_t^{(b)}$ al valor del índice sintético obtenido para el instante t con la b-ésima remuestra, se obtienen el estimador del índice sintético medio, cuyo valor esperado coincidirá con el índice IS obtenido en la sección anterior, y sus desviaciones típicas asociadas en cada instante, siguiendo Efron y Tibshirani (1994) y utilizando las ecuaciones (3.4), (3.5) y (3.6):

$$\bar{IS} = \left(\bar{IS}_1, \bar{IS}_2, \ldots, \bar{IS}_t, \ldots, \bar{IS}_T\right) \tag{3.4}$$

donde

$$\bar{IS}_t = \sum_{b=1}^{B} \frac{IS_t^{(b)}}{B} \tag{3.5}$$

y

$$\hat{\sigma}_{\bar{IS}_t} = +\sqrt{\frac{\sum_{b=1}^{B}\left(IS_t^{(b)} - \bar{IS}_t\right)^2}{B-1}} \tag{3.6}$$

Hay dos posibilidades para construir intervalos de confianza. Una es asumir una distribución normal para los valores del índice sintético en cada instante, y la otra es utilizar directamente los intervalos de confianza del *bootstrap*. La primera aproximación implica calcular los intervalos por $\left[\bar{IS}_t \pm z_{\frac{\alpha}{2}}\hat{\sigma}_{\bar{IS}_t}\right]$, donde $z_{\frac{\alpha}{2}}$ es el percentil $1-\alpha/2$ de una distribución normal estándar. La segunda estrategia estima el intervalo de confianza, con confianza $1-\alpha$, para el valor del índice en el instante t a partir de los percentiles del *bootstrap*, ecuación (3.7).

$$\left[IS_t^{\alpha/2}, IS_t^{1-\alpha/2}\right] \tag{3.7}$$

donde $IS_t^{1-\alpha/2}$ denota el percentil $1-\frac{\alpha}{2}$ del índice sintético remuestreado en el periodo t. Estos percentiles se obtienen como los valores del conjunto $\left\{IS_t^{(1)}, IS_t^{(2)}, \ldots, IS_t^{(b)}, \ldots, IS_t^{(B)}\right\}$ por debajo de los cuales se encuentran $(\alpha/2)$ % y $(1-\alpha/2)$ % de los valores, respectivamente.

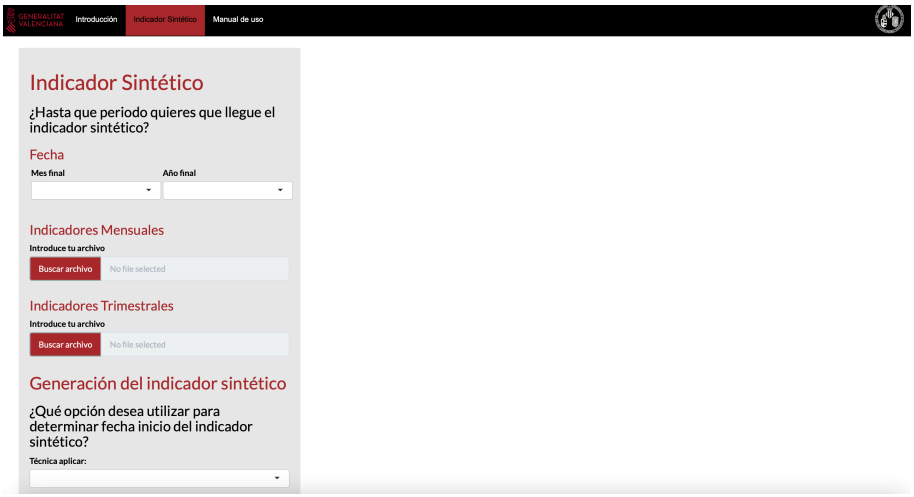
3.4. Índice Sintético: Aplicación web

La metodología de construcción del índice sintético es compleja e implica una serie de operaciones laboriosas que, afortunadamente, pueden automatizarse. La automatización de estos procesos facilita la toma de decisiones económicas, especialmente para los agentes públicos regionales cuyas habilidades estadísticas y computacionales suelen ser

limitadas. Por lo tanto, de acuerdo con uno de los objetivos de esta investigación, desarrollamos una aplicación interactiva de fácil uso para la construcción de un índice sintético, dado un conjunto de indicadores simples, utilizando la metodología de análisis factorial dinámico descrita en la sección anterior. La aplicación se desarrolló utilizando el lenguaje de programación R (R Core Team, 2022), común en áreas de investigación para el análisis estadístico, y emplea el paquete Shiny (Chang *et al.*, 2023), que simplifica la creación de aplicaciones interactivas.

La aplicación "Indicador Sintético" desarrollada con Shiny (ver Figura 3.3) se compone de dos componentes funcionales: una función de interfaz de usuario (UI) y una función de servidor (SERVER). La función UI se encarga de generar la parte visible de la aplicación web (*app*) interactiva y es la que contiene la estructura visual. La función SERVER se encarga de la parte mecánica y ejecutiva interna de la aplicación web interactiva y contiene las instrucciones para la aplicación.

Figura 3.3: Captura de pantalla de la aplicación web Shiny para la construcción del índice sintético.



La *app* consta de tres pestañas. La primera, llamada "Introducción", explica brevemente el objetivo de la aplicación y sus principales resultados. La segunda pestaña es el generador del indicador sintético de crecimiento económico. En esta segunda pestaña, la aplicación web interactiva solicita al usuario varios datos de entrada, como la fecha (mes y año) hasta la cual desea que llegue el indicador sintético, así como uno (o dos) archivo(s)

de Microsoft Excel con los indicadores mensuales y/o trimestrales que el usuario desea combinar para construir el indicador sintético. La primera columna de estos archivos contiene las fechas en formato YYYYMMM (por ejemplo, 2022M12) para series mensuales y en formato YYYYTQ (por ejemplo, 2023T2) para series trimestrales. Los datos, con un encabezado de nombre para cada serie, se encuentran en las columnas que siguen. Dado que el indicador se construye mensualmente, las series trimestrales (si se introducen) se interpolan mensualmente utilizando Dagum y Cholette (2006). Finalmente, la app pregunta al usuario qué opción utilizar para combinar los indicadores en términos de fecha de inicio: "Predeterminado: series observadas" o *backcasting*. La opción de *backcasting* realiza una estimación puntual hacia atrás de todos los indicadores económicos hasta completarlos hasta la fecha del indicador para el cual se dispone de la fecha más antigua. La opción predeterminada inicia todas las series tomando como fecha de inicio la fecha más antigua para la cual hay observaciones conocidas para todas las series.

Una vez que el usuario introduce estos datos, la app muestra los valores del indicador sintético generado como resultado y muestra una representación gráfica en la pantalla. El usuario puede descargar los datos en un archivo en formato .xlsx.

La aplicación `https://apps.uv.es/app_indicador_sintetico/` (acceso el 17 Abril de 2023) también cuenta con una tercera pestaña que contiene el manual de usuario, que explica paso a paso en español cómo crear un indicador sintético de crecimiento económico regional.

3.5. Aplicación Empírica

El uso de indicadores económicos simples para monitorizar el ciclo económico es muy útil para medir la efectividad de las políticas públicas que se están implementando (Cabrer Borrás, 2001; Chernis *et al.*, 2020). Cada indicador proporciona información sobre el pasado y la situación del fenómeno que captura, proyectando su evolución en el presente y futuro. Sin embargo, la información proporcionada por los indicadores es parcial y a veces contradictoria. Distintos indicadores pueden encontrarse en diferentes momentos del ciclo (por ejemplo, la tasa de desempleo tiende a mejorar tiempo después de que ha ocurrido una mejora en el ciclo, mientras que un aumento en la demanda de financiación

empresarial suele anticipar una mejora en el ciclo). Alternativamente, los fenómenos subyacentes que los indicadores miden pueden presentar evoluciones diferentes. Para superar esta limitación, debemos recurrir a métodos de agregación que nos permitan medir las variables latentes no observables del conjunto de indicadores.

Con el objetivo de construir un indicador altamente correlacionado con el PIB de la Comunitat Valenciana y prestando especial atención a criterios como la disponibilidad de la variable y su alta significatividad estadística, en esta investigación seleccionamos inicialmente 14 indicadores para construir el índice sintético, como se detalla en la Tabla 3.1. La fuente oficial que produce todos los indicadores con los que se ha trabajado en esta investigación es el Instituto Nacional de Estadística (INE).

Tabla 3.1: Indicadores simples considerados inicialmente para construir el índice sintético.

Código	**Descripción**
AFSST	Total afiliados al sistema de seguridad social
AFSSC	Total afiliados al sistema de seguridad social en el sector de la construcción
CPPT	Consumo de productos petrolíferos
CVV	Ventas de propiedades
EXPORT	Exportaciones
GTOTUR	Gasto turístico
IASS	Indicador de actividad del sector de servicios
ICMG	Índice de volumen de ventas minoristas
IMPORT	Importaciones
IPI	Índice de producción industrial
MATTUR	Matriculaciones de vehículos
MATVC	Matriculaciones de vehículos pesados
PHT	Pernoctaciones totales en establecimientos hoteleros
VET	Visados totales de construcción

Aunque los 14 indicadores individuales podrían combinarse en un único índice sintético, basándonos en el principio de parsimonia y en nuestro objetivo de representatividad del índice en términos de PIB, como se indica en la sección de metodología, realizamos una

selección previa de indicadores utilizando *"subset selection"*. Utilizando el criterio de AIC, identificando el subconjunto de indicadores que juntos proporcionan un AIC más bajo en relación con el PIB. La principal ventaja de realizar una preselección radica en la reducción de la complejidad y los costes computacionales. Antes de aplicar el método de *"subset selection"*, la serie anual de PIB se transforma en una mensual aplicando la metodología de Dagum y Cholette (2006) y las series mensuales de indicadores elementales se estandarizan para evitar efectos de escala.

En este caso específico, dado que tenemos 14 variables/predictores independientes, se ajustan $2^{14} - 1 = 16383$ modelos diferentes, después de considerar todas las combinaciones posibles con al menos un predictor. Se presenta un extracto con un resumen estadístico con tres de los modelos ajustados en la Tabla 3.2. La última columna de la Tabla 3.2 muestra los valores para el coeficiente $AIC = 2k - 2ln(\hat{L})$ obtenidos para cada modelo, donde k es el número de parámetros estimados (como indicador de complejidad) del modelo y $\hat{L}$ es el valor de la función de verosimilitud del modelo asociado. Según el criterio del AIC, el modelo seleccionado sería el etiquetado con el número 9908 que contiene 8 indicadores. Como se puede observar en la tabla, la mayoría de los indicadores seleccionados son indicadores de demanda y del sector de servicios.

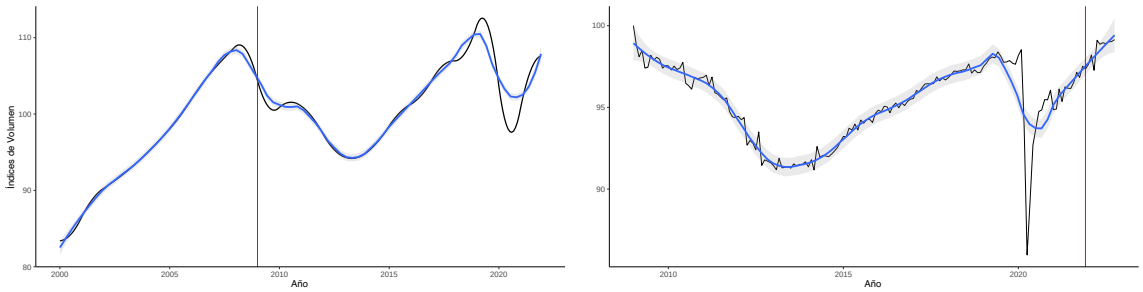
Como muestra la Tabla 3.2, en términos de ajuste no hay diferencias significativas entre el modelo más parsimonioso seleccionado y el modelo más complejo construido mediante combinación lineal de 14 indicadores preseleccionados. Esto se debe a la selección previa de los indicadores iniciales, donde se utilizó el conocimiento experto para su selección. De hecho, los resultados de esta aplicación pueden considerarse robustos, ya que se obtiene casi el mismo índice sintético si se hubiesen utilizado los 14 indicadores iniciales seleccionados.

Tabla 3.2: Resumen estadístico de los modelos en la selección de subconjuntos.

Modelo	**N**	**Predictores**	R^2	**Adj.** R^2	**AIC**
9908	8	AFSST GTOTUR IASS ICMG IMPORT MATTUR MATVC PHT	0.94	0.94	**15.31**
…	…	…	…	…	…
16383	14	AFSST AFSSC CPPT CVV EXPORT EXPORT GTOTUR IASS ICMG IMPORT IPI MATTUR MATVC PHT VET	0.94	0.94	23.12
…	…	…	…	…	…
14	1	CPPT	0.13	0.12	361.72
…	…	…	…	…	…

El resultado del cálculo del índice sintético utilizando la opción “datos observados” de la aplicación con los 8 indicadores seleccionados en el modelo 9908 se muestra en el panel izquierdo de la Figura 3.4 para el período de enero de 2009 a octubre de 2022. Como elemento de comparación, en el panel derecho de la Figura 3.4 se representa gráficamente la serie mensual del PIB de la Comunitat Valenciana, disponible hasta 2021. En ambos paneles se ha incluido una curva suavizada LOESS para facilitar la observación de sus evoluciones subyacentes. Asimismo, se utilizan líneas verticales en color rojo en ambos paneles para indicar los períodos durante los cuales la otra serie está (o no está) disponible

Figura 3.4: Serie estimada del índice sintético y serie del PIB de la Comunitat Valenciana.



Nota: Las unidades de medidas de la serie estimada del índice sintético (panel izquierdo) son índices de volumen y la serie de PIB (panel izquierdo), se encuentran en índices de volumen encadenados. En ambos paneles se ha incluido una curva suavizada de loess. La parte más a la derecha en el panel izquierdo, a la derecha de la línea roja vertical, corresponde a meses en los que el PIB aún no está disponible. La parte más a la izquierda del panel derecho, a la izquierda de la línea roja vertical, corresponde a años en los que algunos de los indicadores no están disponibles. Como se puede observar, en el período de intersección en el que ambas series están disponibles, los movimientos de ambas curvas son bastante similares.

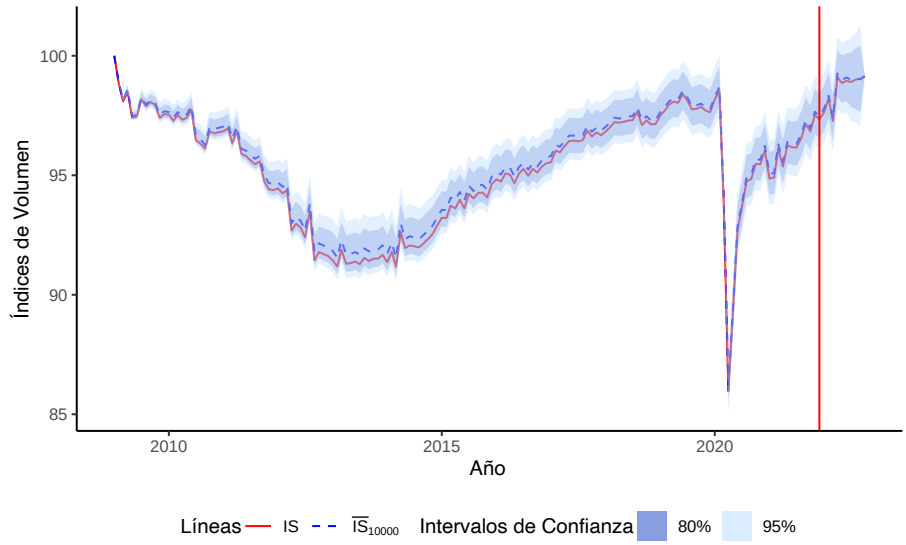
Como se puede observar al comparar ambos paneles de la Figura 3.4, durante el periodo de intersección (cuando ambas series están disponibles), los movimientos de ambas curvas son bastante similares, lo que indica la utilidad del índice sintético, para el cual tenemos estimaciones de *nowcasting* y *forecasting*. En ambas curvas, son evidentes las consecuencias económicas del confinamiento implementado para mitigar la pandemia del COVID-19. Es notable que la brusca caída en la actividad económica causada por el confinamiento es más evidente en el índice sintético que en la serie mensual del PIB. Por un lado, hubo un efecto de suavizado estadístico causado por el método de desagregación mensual y, por otro lado, la importante ayuda financiera que acompañó al confinamiento permitió mantener un grado significativo de consumo a pesar de la reducción en la actividad.

3.6. Medición de la incertidumbre

Una predicción o estimación no está completa sin una medición de su error de predicción/estimación, es decir, del nivel de incertidumbre asociado. El cálculo de la incertidumbre no siempre es sencillo, como es el caso de la construcción de nuestro índice sintético que emplea el análisis factorial dinámico. Afortunadamente, como se discute en la sección

de metodología, las técnicas de simulación pueden utilizarse (casi siempre) para generar estimaciones de la incertidumbre. En esta sección, mostramos (como se puede observar en la Figura 3.5) los resultados de medir la incertidumbre asociada con la construcción del índice sintético utilizando $B = 10000$, según el procedimiento descrito en la subsección 3.3.2.

Figura 3.5: Representación gráfica de la incertidumbre al estimar el indicador sintético mediante *bootstrap* con B=10000.



Nota: Se ha representado tanto el índice sintético de referencia (línea roja) como el índice sintético medio de *bootstrap* (línea discontinua roja). También se muestran los intervalos de confianza *bootstrap* al 80 % y al 95 % de confianza.

La Figura 3.5 muestra gráficamente el índice sintético junto con la estimación de sus intervalos de confianza de *bootstrap* al 80 % y al 95 % de confianza. Al analizar la figura, se destacan varios hechos. En primer lugar, las bandas de confianza de la incertidumbre no son homogéneas a lo largo de todo el período. Durante los primeros años y en el momento del confinamiento pandémico, los niveles de incertidumbre son más bajos. Los niveles más altos de incertidumbre se observan durante el período de predicción/extrapolación. En segundo lugar, los intervalos *bootstrap* no son simétricos con respecto al índice base, IS, ni con respecto al índice medio, $\bar{IS}$, siendo los intervalos significativamente más asimétricos con respecto al primero. En general, el rango del intervalo hacia el extremo superior es mayor que hacia el extremo inferior. En tercer lugar, como regla general, las series de índices IS y $\bar{IS}$ son casi coincidentes, trazando caminos muy similares. En resumen,

la metodología de *bootstrap* se revela adecuada para aproximar la incertidumbre en la estimación del índice sintético de actividad y predicción económica de la Comunitat Valenciana.

3.7. ¿Cuántas remuestras son necesarias? Un análisis de sensibilidad

La sección anterior concluye que la simulación *bootstrap* es una buena alternativa para medir la incertidumbre cuando combinamos diferentes series económicas mediante el análisis factorial dinámico. Sin embargo, en la aplicación anterior, utilizamos 10000 simulaciones, lo que conllevó un alto coste computacional (tiempos y ciclos de procesamiento elevados), incluso al utilizar sistemas de procesamiento paralelo que reducen los tiempos de cálculo. La pregunta natural que surge es si es necesario realizar tantas simulaciones. En otras palabras, ¿cuál sería el número mínimo de remuestras necesario para garantizar la convergencia de los intervalos de confianza?. Conocer el número mínimo de escenarios simulados necesarios para lograrla estabilidad en la estimación de los errores estándar e intervalos de confianza es relevante porque, por un lado, nos permitiría estar seguros de las estimaciones obtenidas y, por otro lado, evitaría desperdiciar recursos al incurrir en un coste computacional excesivo con cálculos innecesarios. Brindar una respuesta a esta pregunta es la segunda de las contribuciones de este capítulo.

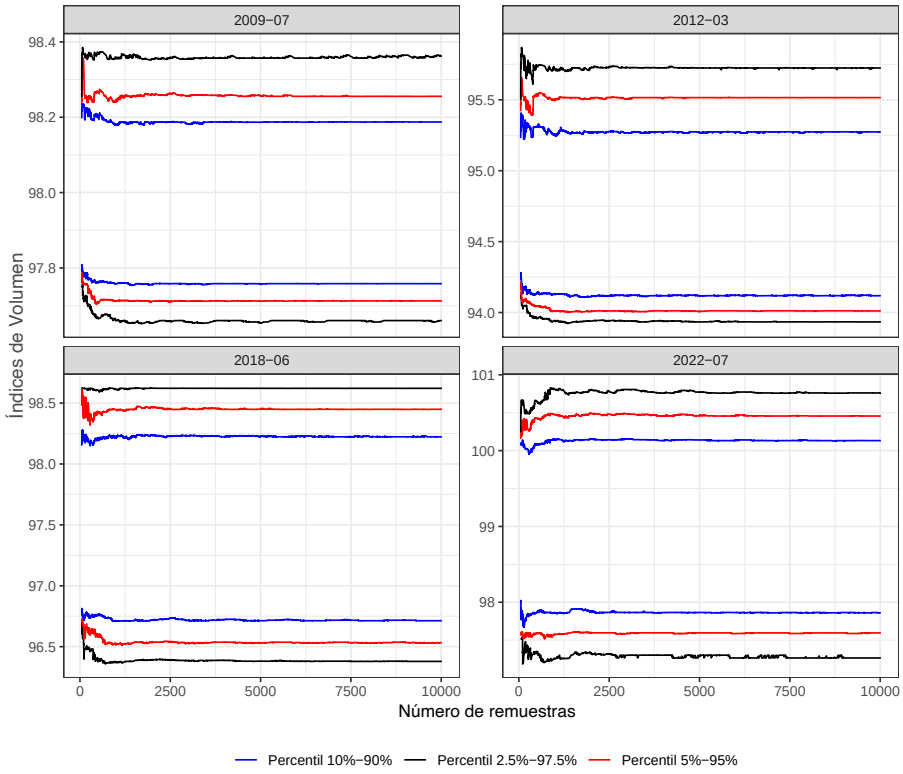
Sin embargo, en la literatura, no hay consenso sobre cuál debería ser el número mínimo de remuestras para lograr estabilidad en las estimaciones. Efron y Tibshirani (1994) afirman que con un número relativamente bajo, como 50 o incluso 25 remuestras, podríamos lograr estabilidad en la estimación del error estándar, aunque el número de remuestras necesario para estimar los intervalos de confianza sería algo mayor. Otros autores como Wilcox (2010) recomiendan utilizar un mínimo de 599 simulaciones como regla general, mientras que Davidson y MacKinnon (2000) defiende que el número óptimo mínimo para B debe ser razonablemente grande. En la práctica, Davidson y MacKinnon (2000) indican que usar 399 simulaciones sería suficiente, aunque recomiendan utilizar un número mayor si deseamos aumentar la potencia y precisión en la estimación del error de estimación y los

intervalos de confianza con $1-\alpha = 0.95$.

Otros autores son más conservadores (Hesterberg, 2011) consideran que se necesitan 1000 remuestras para lograr estabilidad en la estimación del error y recomienda el uso de 10000 remuestras cuando el objetivo es estimar intervalos de confianza con cobertura suficiente. Finalmente, Chernick (2011) afirma que no hay un número fijo de remuestras que garantice la convergencia media.

Para responder a las preguntas planteadas en el contexto de esta investigación, simulamos $B = 10000$ remuestras, un número superior al recomendado por Wilcox (2010) y Davidson y MacKinnon (2000) e igual al recomendado por Hesterberg (2011), y (i) estudiamos, como en función del número de remuestras b, cuándo se alcanza la estabilidad y (ii) comparamos las diferencias observadas entre los límites de los intervalos obtenidos para diferentes números de remuestreos, $b = \{25, 50, 399, 599, 1000, 10000\}$. Los análisis también se llevaron a cabo considerando el posible impacto de la cobertura $(1-\alpha)$ del intervalo utilizado en los resultados. De hecho, como se muestra en la Figura 3.6, donde se presentan los límites de los intervalos de confianza obtenidos para el índice sintético en cuatro momentos diferentes en función del número de remuestreos b, no todos los límites del intervalo tienden a estabilizarse con el mismo número de remuestras.

Figura 3.6: Intervalos de confianza (80 %, 90 % y 95 %) en función del número de remuestras del índice sintético en determinados momentos temporales.



Para estudiar cuándo se alcanza la estabilidad en los límites de los intervalos de confianza, utilizamos regresiones lineales con variable independiente el número de remuestras y variables dependientes los límites de los intervalos: $\ell_t^{(b)} = \alpha + \beta b$ para $b \geq b_0$, donde $\ell_t^{(b)}$ representa, para el instante t, cualquiera de los límites del intervalo de confianza del índice obtenido utilizando las primeras b remuestras.

Dado que, lógicamente, los percentiles *bootstrap* tienden a estabilizarse a medida que aumenta el número de remuestras, el cálculo de los p-valores asociados con los coeficientes β para $b \geq b_0$ indica a partir de qué número de remuestras b_0 el coeficiente β ya no es estadísticamente significativo. A partir de este número, podemos decir que el conjunto correspondiente de límites se ha estabilizado, ya que la línea de regresión se ajusta bien mediante una constante.

Adoptando un criterio conservador, según el cual necesitamos al menos tantas re-

muestras como el número máximo de remuestras necesario para alcanzar la estabilidad en todos los puntos temporales, se deduce que, para los niveles de confianza habituales ($1-\alpha = 0.80, 0.90, 0.95$), necesitaríamos simular casi 10000 remuestras. Este resultado estadístico no es muy útil desde un punto de vista práctico porque, como se muestra en la Figura 3.6, generalmente la estabilidad práctica se alcanza mucho antes.

Para responder a la pregunta desde una perspectiva práctica, comparamos los límites obtenidos para diferentes combinaciones remuestras y analizamos los errores y desviaciones típicas asociadas con sus diferencias. Para cada combinación posible del conjunto $\Omega = \{25, 50, 399, 599, 1000, 10000\}$, calculamos a lo largo de los períodos las estadísticas $sd_t\left(\ell_t^{(b_1)} - \ell_t^{(b_2)}\right)$ y $sd_t\left(\ell_t^{(b_1)} - \ell_t^{(b_2)}\right) / mean_t\left(\ell_t^{(b_2)}\right)$, donde $b_1, b_2 \in \Omega$ y $b_1 < b_2$. Las Tablas 3.3, 3.4 y 3.5 muestran los valores obtenidos para dichas estadísticas asociadas con los intervalos de confianza al 80 %, 90 % y 95 %, respectivamente.

Tabla 3.3: Desviaciones típicas y errores relativos en intervalos de confianza al 90 % y al 10 %.

	25	**50**	**399**	**599**	**1000**	**10000**	**25**	**50**	**399**	**599**	**1000**	**10000**
25		0.1018	0.0991	0.1001	0.1034	0.0985		0.1058	0.1029	0.1040	0.1074	0.1023
50	0.1171		0.0352	0.0439	0.0629	0.0590	0.1232		0.0365	0.0456	0.0653	0.0612
399	0.0969	0.0688		0.0195	0.0460	0.0405	0.1020	0.0724		0.0203	0.0477	0.0420
599	0.0954	0.0683	0.0098		0.0312	0.0260	0.1005	0.0719	0.0103		0.0324	0.0270
1000	0.1124	0.0682	0.0254	0.0247		0.0114	0.1183	0.0718	0.0267	0.0260		0.0118
10000	0.1079	0.0704	0.0181	0.0177	0.0106		0.1136	0.0742	0.0191	0.0186	0.0112	

Nota: Desviaciones típica (panel izquierdo) y errores relativos en porcentaje (panel derecho) a lo largo del tiempo entre los límites de los intervalos de confianza al 90 % (triángulos superiores) y 10 % (triángulos inferiores) para un número seleccionado de remuestras.

Tabla 3.4: Desviaciones típicas y errores relativos en intervalos de confianza al 95 % y al 5 %.

	25	50	399	599	1000	10000	25	50	399	599	1000	10000
25		0.0858	0.1149	0.1373	0.1270	0.1260		0.0890	0.1190	0.1422	0.1315	0.1305
50	0.0623		0.0701	0.0904	0.0951	0.0963	0.0656		0.0726	0.0936	0.0985	0.0997
399	0.0855	0.0868		0.0396	0.0518	0.0508	0.0901	0.0915		0.0410	0.0537	0.0526
599	0.0892	0.0837	0.0205		0.0595	0.0578	0.0940	0.0883	0.0217		0.0616	0.0598
1000	0.1006	0.0861	0.0438	0.0305		0.0119	0.1061	0.0908	0.0463	0.0322		0.0123
10000	0.0999	0.0887	0.0345	0.0232	0.0140		0.1053	0.0936	0.0364	0.0245	0.0148	

Nota: Desviaciones típica (panel izquierdo) y errores relativos en porcentaje (panel derecho) a lo largo del tiempo entre los límites de los intervalos de confianza al 95 % (triángulos superiores) y 5 % (triángulos inferiores) para un número seleccionado de remuestras.

Tabla 3.5: Desviaciones típicas y errores relativos en intervalos de confianza al 97.5 % y al 2.5 %.

	25	50	399	599	1000	10000	25	50	399	599	1000	10000
25		0.1178	0.1112	0.1160	0.1437	0.1483		0.1219	0.1150	0.1199	0.1485	0.1532
50	0.1031		0.1263	0.1241	0.1280	0.1310	0.1085		0.1305	0.1282	0.1323	0.1354
399	0.0911	0.0859		0.0238	0.0558	0.0618	0.0961	0.0906		0.0246	0.0577	0.0638
599	0.0970	0.0874	0.0191		0.0970	0.0874	0.1024	0.0922	0.0202		0.0487	0.0531
1000	0.1029	0.0926	0.0245	0.0181		0.0142	0.1087	0.0978	0.0258	0.0191		0.0147
10000	0.1059	0.0913	0.0286	0.0212	0.0126		0.1119	0.0964	0.0302	0.0224	0.0133	

Nota: Desviaciones típica (panel izquierdo) y errores relativos en porcentaje (panel derecho) a lo largo del tiempo entre los límites de los intervalos de confianza al 97.5 % (triángulos superiores) y 2.5 % (triángulos inferiores) para un número seleccionado de remuestras.

El análisis de los datos de las tablas muestran que los límites inferiores de los intervalos presentan una mayor estabilidad que los límites superiores y, además, su error relativo alcanza un valor cercano al 1 % bastante pronto. Las distancias medias a lo largo de los períodos entre los límites inferiores son prácticamente equivalentes a partir de 399 remuestras. Sin embargo, para los límites superiores, necesitamos alrededor de 1000 remuestras para obtener límites similares a los obtenidos con 10000 simulaciones. A la luz de estos resultados, consideramos que, en este contexto, es más que suficiente realizar 1000 remuestras para calcular los intervalos de confianza para el índice sintético en los diferentes momentos temporales.

3.8. Resumen y observaciones finales

Los agentes económicos, tanto públicos como privados, tienen la responsabilidad de tomar decisiones a diario que pueden tener importantes consecuencias sociales y financieras. Estas decisiones se toman en función de las condiciones actuales y futuras de la región o país. Por lo tanto, contar con herramientas adecuadas para el análisis coyuntural, la predicción económica y la medición de la incertidumbre constituye un instrumento poderoso que puede mejorar la calidad de las decisiones.

Este documento muestra una parte de la investigación llevada a cabo con el objetivo de facilitar la toma de decisiones económicas de los agentes públicos regionales. En este trabajo, presentamos una aplicación web interactiva (`https://apps.uv.es/app_indicador_sintetico/`, consultada el 17 de abril de 2023) que permite generar fácilmente un indicador sintético mediante el análisis de factorial dinámico a través de la combinación dinámica de indicadores económicos individuales, y analizamos cómo medir su incertidumbre asociada. Nuestra investigación también responde a la pregunta sobre el número mínimo de simulaciones de *bootstrap* necesarias para medir adecuadamente la incertidumbre de una serie sintética obtenida mediante el análisis dinámico de factores.

Este artículo muestra una parte de las investigaciones que se están llevando a cabo los autores con el objetivo de facilitar la toma de decisiones económicas de los agentes públicos regionales. En este trabajo presentamos una aplicación web interactiva (`https://apps.uv.es/app_indicador_sintetico/`) que permite generar fácilmente un indicador sintético utilizando análisis factorial dinámico a través de la combinación dinámica de indicadores económicos individuales, y analizamos cómo medir su incertidumbre asociada. Nuestra investigación también responde a la pregunta sobre el número mínimo de simulaciones bootstrap necesario para medir adecuadamente la incertidumbre de una serie sintética obtenida mediante análisis de factores dinámicos.

A lo largo de nuestro estudio, se verificó que la metodología propuesta por Cuevas *et al.* (2015) para obtener un indicador sintético basado en factores latentes permite capturar la señal conjunta común de un grupo de indicadores simples, estimando adecuadamente la evolución económica agregada de la Comunitat Valenciana. Esto permite a los tomadores de decisiones públicas mejorar sintéticamente su comprensión sin retrasos y, gracias a la

aplicación, de manera autónoma, permite conocer la situación actual de la economía regional (*nowcasting*) y también prever su evolución inmediata futura (*short-term forecasting*). A la luz de estos hallazgos, como recomendación de política, consideramos que los agentes económicos de la Comunitat Valenciana pueden confiar en esta nueva herramienta para la toma de decisiones.

La construcción del índice sintético se complementa con el estudio del cálculo de su incertidumbre, para lo cual, dada la complejidad asociada con la metodología, utilizamos técnicas de *bootstrap*. La aproximación elegida se revela como una buena alternativa para medir la incertidumbre cuando combinamos diferentes series económicas a través del análisis factorial dinámico, demostrando que, en este contexto, 1000 simulaciones son suficientes para calcular adecuadamente los intervalos de confianza del índice sintético en diferentes momentos temporales.

Sin embargo, nuestra herramienta y resultados no están exentos de limitaciones. La aplicación que desarrollamos solo considera el primer factor dinámico e impone un período máximo de predicción de 12 meses basado en la serie más retardada. Además, predice cada indicador utilizando un modelo ARIMA univariante identificado automáticamente. Todos estos problemas limitan la flexibilidad de la herramienta y pueden no conducir a las estimaciones más precisas en ciertas circunstancias. Sin embargo, creemos que asumirlos representa un compromiso razonable dado el usuario objetivo de nuestra aplicación.

Asimismo, vale la pena señalar que, si las series mensuales en nuestra aplicación se hubieran agregado anualmente para estudiar su relación con el PIB, si se hubiera utilizado un enfoque diferente para la selección de variables para elegir los indicadores individuales, y/o si se hubiera tenido en cuenta el error de medición de los indicadores al estimar la incertidumbre del índice sintético, se podrían haber alcanzado soluciones ligeramente diferentes.

Finalmente, los lectores podrían estar interesados en saber que, aunque este índice es valioso como herramienta independiente, actualmente estamos trabajando en una nueva aplicación en la cual el código de construcción del indicador sintético y su metodología asociada se utilizan como un componente intermedio para estimar las series trimestrales regionales del PIB (ver, por ejemplo, Pavía-Miralles y Cabrer-Borrás (2007)).

Capítulo 4

Estimando el impacto de los Fondos Next Generation EU a nivel regional en España: una aproximación basada en el método de control Sintético

Resumen

El programa Next Generation de la Unión Europea (NGEU) y su implementación a través del Fondo de Recuperación y Resiliencia (FRR) han sido concebidos con la premisa de promover una respuesta fiscal coordinada dentro de la Unión Europea para abordar los desafíos surgidos a raíz de la crisis del COVID-19 y con el objetivo de transformar el sistema productivo hacia una economía más verde, sostenible y digital. El programa proporciona a los Estados miembros acceso a subvenciones y préstamos destinados a respaldar sus planes de recuperación y resiliencia, los cuales deben incorporar paquetes coherentes de reformas e inversiones. En este capítulo evaluamos el impacto económico regional del programa NGEU en España, uno de los países europeos más afectados por la pandemia y, en consecuencia, uno de los principales beneficiarios del programa. Para ello, empleamos técnicas contrafactuales, que son particularmente útiles al evaluar escenarios alternativos, como la existencia o ausencia de fondos NGEU. De acuerdo con nuestros resultados, los fondos tienen un impacto económico mesurable en términos de aumento del PIB per cápita durante 2022, una perspectiva que se proyecta en 2023 y 2024, en comparación con un escenario sin fondos NGEU. Durante la evaluación de los resultados, abordamos también una de las limitaciones actuales de los modelos de control sintético que, debido al conservadurismo de sus tests estadísticos asociados, encuentran dificultades para detectar efectos aún estando estos presentes. En concreto, proponemos estrategias inferenciales que superan las pruebas conservadoras de significación estadística, ofreciendo así una evaluación más completa del impacto de los fondos NGEU en las regiones españolas[1].

[1] Los principales resultados que integran en este capítulo han sido publicados en el artículo, en evaluación "On the Impact of Next Generation EU Funds: A Regional Synthetic Control Method Approach", disponible como working paper en https://ideas.repec.org/p/jau/wpaper/2023-07.html.

4.1. Introducción

El segundo aniversario del establecimiento del Mecanismo de Recuperación y Resiliencia (en adelante, MRR) brinda una valiosa ocasión para el análisis retrospectivo de los logros y desafíos inherentes a este mecanismo excepcional. Posicionado como el centro de la estrategia de recuperación Next Generation EU de 750 000 millones de euros para la Unión Europea (también conocidos como "fondos NGEU"). El MRR fue instituido en el contexto de la pandemia del COVID-19 con el objetivo principal de acelerar el proceso de recuperación entre los Estados miembros y mejorar su resiliencia.

Este programa ha sido concebido como un instrumento potente a nivel supranacional para facilitar una transición rápida y ambiciosa hacia iniciativas ambientalmente sostenibles y digitalmente impulsadas, mejorando la transformación a largo plazo de la economía de la UE. Sin embargo, la ejecución del MRR se desarrolla en un contexto permanentemente dinámico, caracterizado por la incursión de Rusia en Ucrania, que ha acelerado el aumento de los niveles de inflación y ha desatado una crisis energética.

Aunque se espera que este enorme paquete de estímulos tenga un fuerte impacto positivo en las economías de la UE, la evidencia empírica actualmente disponible sigue siendo limitada y generalmente se basa en modelos de equilibrio general estocástico dinámico (DSGE) como EAGLE o el modelo semi-estructural multi-país ECB-MC (Bańkowski *et al.*, 2022, 2021). Según estos estudios, los fondos NGEU podrían impulsar el PIB de la zona del euro en un 0.5 % durante 2022-23 (1.5 % a medio plazo), contribuyendo aumentar la producción potencial entre un 1.0 % y un 1.4 % a largo plazo (hasta un 1.5 % para 2026). Sin embargo, tanto la magnitud como la persistencia de este efecto positivo dependerán de la capacidad productiva global de cada economía.

Otros estudios que también encuentran efectos positivos incluyen Pfeiffer *et al.* (2023), quienes consideran el modelo QUEST de la Comisión Europea (Picek, 2020) y utilizan, de manera similar a Fernández-Cerezo *et al.* (2023), que se basan en la matriz de input-output (World Input Database), un modelo de input-output multirregional.

Una característica común a todos los estudios anteriores es que se concentran en el ámbito nacional. Hasta donde llega nuestro conocimiento, solo Barbero *et al.* (2023) considera niveles territoriales más bajos (230 regiones NUTS2 europeas). Sin embargo, y

desde un punto de vista metodológico, también adopta modelos de equilibrio general.

En este capítulo, contribuimos a esta incipiente literatura desde un punto de vista diferente. Por un lado, adoptamos una perspectiva de un solo país que, aunque menos general en alcance geográfico, puede proporcionar un análisis más detallado de la efectividad del mecanismo del MRR al considerar niveles territoriales más bajos, en línea con el estudio de Barbero *et al.* (2023). Por otro lado, desde una perspectiva metodológica, utilizamos métodos de control sintético (Abadie, 2021; Abadie *et al.*, 2010; Abadie y Gardeazábal, 2003; Eli Ben-Michael y Rothstein, 2021), que se han vuelto relativamente populares pero que hasta la fecha no se han considerado para evaluar los efectos de los fondos NGEU en el PIB de la zona del euro.

El método de control sintético (SCM) utiliza datos reales para desarrollar un modelo que evita las suposiciones teóricas impuestas por los modelos DSGE, proporcionando un tipo alternativo de análisis. Los SCM han demostrado ser una herramienta poderosa, especialmente al tratar con tamaños de muestra pequeños, como en el caso de unidades agregadas expuestas a un tratamiento o evento. Este método es particularmente adecuado para evaluar el impacto de shocks agregados, como políticas públicas, desastres naturales y ataques terroristas, entre otros. En nuestro caso, basándomos en datos observados, el SCM simplifica la evaluación del panorama económico de España construyendo un homólogo sintético de cada región de estudio. Esto se logra utilizando regiones de control (no intervenidas) con características pre-intervención similares a las de las regiones tratadas.

A diferencia de los modelos DSGE, el SCM impone restricciones teóricas más flexibles, facilitando la incorporación de la complejidad inherente a intervenciones específicas, como las relacionadas con los fondos NGEU, y eliminando la necesidad de asumir un conocimiento completo de las relaciones causales en la economía. Al medir el impacto directo de la intervención, estos modelos proporcionan una evaluación más específica y detallada de cómo los fondos NGEU podrían afectar directamente a variables relevantes, lo que es esencial para la toma de decisiones informada. Además, los modelos de control sintético tienen una mayor capacidad para realizar análisis a niveles más detallados, como sectores o regiones, en los que nos enfocamos, creando un homólogo sintético de la región en estudio a través de regiones de control comparables. Esto podría permitir una comprensión más precisa de cómo ciertas políticas afectan a diversos segmentos de la economía.

Por lo tanto, este método puede contribuir a una comparación y comprensión precisa del impacto heterogéneo de los fondos NGEU en las regiones españolas, sin tener que imponer fuertes suposiciones sobre sus dinámicas económicas específicas. Además, estos modelos también son especialmente adecuados para abordar asimetrías y no linealidades en los datos.

Este enfoque se puede vincular con la extensa literatura sobre multiplicadores fiscales, que está estrechamente relacionada con el tema que analizamos aquí y, aunque muchas contribuciones relevantes también consideran modelos DGSE (por ejemplo, Brueckner *et al.*, 2023), otros aplican estrategias empíricas como el diseño de regresión de discontinuidad (RDD) (ver Corbi *et al.*, 2019). En este sentido, nuestros objetivos y enfoque están más cerca de estudios recientes como Alloza y Sanz (2021),[2] quienes evaluaron los efectos sobre el desempleo del llamado *Plan E*, un gran programa de estímulo fiscal implementado en España en el período 2009-2011. Estos autores utilizan un enfoque de diferencias en diferencias aprovechando la variación en el momento de ejecución de proyectos entre municipios; es decir, también consideraron unidades subnacionales de gobierno como lo hacemos nosotros.

Coincidentemente, nuestro contexto empírico también es España, que experimentó un confinamiento draconiano y, como consecuencia, fue testigo de una de las caídas más pronunciadas en el PIB (casi el 12 %) en Europa. Por lo tanto, la expectativa es que, junto con Italia, el Fondo de Recuperación y Resiliencia debería tener un efecto particularmente fuerte en comparación con sus homólogos europeos (Bańkowski *et al.*, 2022). De hecho, el NGEU fue diseñado para favorecer a los países vulnerables y de bajos ingresos y aquellos especialmente afectados por la pandemia (Pfeiffer *et al.*, 2023). España tiene ventajas adicionales como laboratorio, ya que su estructura territorial cuasi-federal implica que, en la práctica, la cantidad de fondos asignados a cada región, la forma en que se ha implementado el mecanismo y los fondos invertidos varían considerablemente entre regiones. Esto puede proporcionar más información sobre la efectividad del MRR, dependiendo del grado de descentralización de cada país de la UE.

Nuestros resultados, señalan un impacto económico relevante del programa NGEU en

[2]Ver también Carozzi y Repetto (2019) y Montolio (2018), que analizan otros aspectos relacionados con este mecanismo específico.

las regiones españolas y demuestran el potencial significativo de los modelos contrafactuales como herramientas efectivas para medir el impacto de eventos o perturbaciones disruptivas en un contexto económico dado. Sin embargo, es importante reconocer que, a pesar de su utilidad, estos modelos pueden tener ciertas limitaciones que deben tenerse en cuenta al interpretar los resultados.

Específicamente, uno de los inconvenientes prácticos del enfoque contrafactual reside en las pruebas relativamente conservadoras que generalmente se emplean para evaluar la significación estadística de sus estimaciones (Chernozhukov *et al.*, 2021; Firpo y Possebom, 2018; Galiani y Quistorff, 2017). Estas pruebas pueden no revelar significación estadística a pesar de informar claramente un efecto gráficamente. Nuestro artículo contribuye a la literatura actual describiendo cómo se puede superar esta limitación. Específicamente, mostramos que la evidencia acumulativa proporcionada por los efectos de una intervención en un conjunto de unidades constitutivas de una unidad global puede aprovecharse para superar las limitaciones de las pruebas actualmente disponibles. Utilizamos diferentes estrategias inferenciales para evaluar conjuntamente la significación de un conjunto de estimaciones contrafactuales, incluso cuando ninguna estimación individual se identifica como estadísticamente significativa.

En este contexto, realizamos pruebas unilaterales de significación de medias, pruebas de diferencia de medias antes y después de la intervención, y pruebas de medias restringidas, todas las cuales son elementos cruciales para evaluar las diferencias entre la tendencia y magnitud del impacto de los fondos en los períodos posteriores al tratamiento en la región analizada y su homólogo sintético. Es importante destacar que, después de estas pruebas, este efecto, que observamos tanto gráficamente como ahora numéricamente, no se debe simplemente al azar, sino que refleja una influencia sustancial. Especialmente, este patrón de tendencia permanece consistente en todas las regiones españolas analizadas, un hallazgo que aumenta la confianza en la efectividad de los fondos como impulsores del crecimiento económico en el contexto específico examinado. Las pruebas estadísticas realizadas en este estudio ofrecen una opción a las alternativas conservadoras utilizadas y disponibles en el software para implementar SCM y evaluar la existencia de impacto.

El resto del capítulo está organizado de la siguiente manera. Después de esta introducción, la siguiente sección revisa (muy brevemente y parcialmente) la literatura sobre

multiplicadores fiscales y *fiscal spillovers* en la Unión Europea (subsección 4.2.1), y sobre el impacto económico de los fondos Next Generation (subsección 4.2.2). La estrategia empírica se presenta en la Sección 4.3, tanto la metodología (subsección 4.3.1) como los datos utilizados en el estudio (subsección 4.3.3). Por último, los resultados se informan en la Sección 4.4 y la Sección 4.5 proporciona algunas conclusiones.

4.2. Estímulos económicos y NGEU

En respuesta a la gravedad de las consecuencias económicas del virus COVID-19, la Comisión Europea diseñó el paquete Next Generation EU como un instrumento para redistribuir y estabilizar la economía. El paquete se implementa a través de préstamos obtenidos por la Comisión Europea durante el período presupuestario 2021-2027 (además del presupuesto regular). Los reembolsos se realizarán mediante las contribuciones de los Estados miembros con pagos de intereses bajos, así como cantidades significativas de capital (Picek, 2020).

El paquete NGEU se divide en varios mecanismos, como el Fondo de Recuperación y Resiliencia Europeo, un programa de reconstrucción post-COVID-19 implementado a través de subvenciones a los Estados miembros, que totaliza 312.5 mil millones de euros. Una cantidad similar (360 mil millones de euros) corresponde a préstamos, generando ahorros en tasas de interés para los Estados miembros. Finalmente, una proporción mucho menor de los fondos NGEU se asigna a través de otros programas existentes de la UE (71.9 mil millones de euros). El paquete se completa con una pequeña suma de garantías.

En cuanto al momento, la mayor parte de los pagos reales (tres cuartas partes de los pagos del MRR) se han distribuido a lo largo del 2023 o más tarde, lo que ha puesto en duda la efectividad del mecanismo en sus primeras etapas. Sin embargo, los reembolsos se llevarán a cabo mucho más tarde: las subvenciones financiadas por un préstamo se pagarán a partir de 2028 y hasta 2058.

4.2.1. Multiplicadores fiscales y *fiscal spillovers* en la UE

El debate sobre la efectividad de las políticas fiscales para estimular la economía ha atraído considerable atención teórica y empírica. Académicos y responsables de políticas han abordado persistentemente preguntas de si una política fiscal expansiva, una reducción de impuestos o un aumento del gasto público pueden impulsar la producción y el consumo o multiplicar los empleos, y en qué medida (Hebous, 2011; Hebous y Zimmermann, 2013). Las severas crisis que han afectado a la mayoría de los países de todo el mundo desde el comienzo del siglo, especialmente la Gran Recesión de 2007/08 y la recesión por el COVID-19, también han contribuido significativamente al impulso de esta literatura.

Este debate ha sido más intenso en ciertos ámbitos, como la Unión Europea, no solo debido a consideraciones sobre el papel óptimo de las políticas fiscales en una unión monetaria, sino también debido al efecto severo que la Gran Recesión tuvo en sus estados miembros, especialmente en los países de la zona euro (Coelho, 2019). Como resultado, hay una parte en la literatura empírica especialmente centrada en este contexto, que comprende dos corrientes principales: la literatura de multiplicadores transversales macroeconómicos (por ejemplo, Amendola *et al.*, 2020; Brueckner *et al.*, 2023), y la literatura de intervenciones políticas basadas en la localización (por ejemplo, Bartik, 2020; Brachert *et al.*, 2019; Ehrlich y Overman, 2020).

En el caso de Europa, también ha surgido una literatura paralela que intenta medir cómo los fondos estructurales y de cohesión europeos pueden afectar al crecimiento y al empleo a nivel regional. Contribuciones relevantes a esta corriente de la literatura incluyen, entre otras a Becker *et al.* (2010); Mohl y Hagen (2010) y Becker *et al.* (2012, 2018). Esta línea de investigación está estrechamente relacionada con la literatura sobre las ganancias de bienestar de una unión fiscal redistributiva en Europa (Bargain *et al.*, 2013), políticas que han recibido, en general, un respaldo empírico positivo. El impacto en el crecimiento del PIB tiende a ser positivo[3] (Beugelsdijk y Eijffinger, 2005; Ederveen *et al.*, 2006), aunque los resultados son menos concluyentes para el empleo. Los resultados también varían considerablemente entre regiones (por ejemplo, Becker *et al.*, 2013).

[3]Excepto por Sala-i-Martin (1996).

En cuanto al caso específico de los *fiscal spillovers*, se han utilizado ampliamente modelos de input-output para identificar los efectos de un shock de demanda final a través de matrices multirregionales de flujos intra-industriales de bienes intermedios (Picek y Schröder, 2018). Esta literatura ha prestado especial atención al tamaño de los *spillovers* para diferentes países y regiones europeas (Picek, 2020), y una extensa corriente ha considerado autorregresiones vectoriales (Beetsma y Giuliodori, 2011; Beetsma *et al.*, 2006), cuyas contribuciones más relevantes son revisadas por Hebous (2011).

La imposición de medidas de austeridad durante la crisis de la deuda soberana de 2010–12 generó tensiones repetidas entre la Comisión Europea y los Estados miembros, algunos de los cuales solicitaron un mayor uso de estímulos fiscales (Coelho, 2019). En este sentido, la actual postura fiscal expansiva debido a la pandemia del COVID-19 da un giro al debate de austeridad de la última década (Pfeiffer *et al.*, 2023). Por lo tanto, nuestro análisis contribuye a la investigación sobre *fiscal spillovers* y multiplicadores fiscales en la UE, extendiendo, a la luz de las políticas de la NGEU, la corriente de la literatura que mide el efecto de los fondos de cohesión en el crecimiento y el empleo regionales (Di Caro y Fratesi, 2022).

4.2.2. NGEU y su impacto: evidencia

A pesar de la magnitud del programa, debido a su reciente implementación, la literatura que evalúa el impacto económico de los fondos Next Generation en los países de la Unión Europea aún es escasa. Los estudios más explícitos son los de Bańkowski *et al.* (2021) y su extensión más reciente en Bańkowski *et al.* (2022). Este último es un estudio más completo que utiliza un modelo semi-estructural considerando los tres canales económicos a través de los cuales NGEU afectará a las economías de la zona euro desde una perspectiva macroeconómica, como: (i) el canal de la prima de riesgo; (ii) el canal de estímulo fiscal; y (iii) el canal de reforma estructural. A través de estos tres mecanismos, que operan de manera bastante diferente, se espera que NGEU aumente el PIB en la zona euro hasta un 1.5 % para 2026. Aunque este aumento podría parecer modesto *a priori*, los efectos variarán considerablemente de un país a otro, con países como España e Italia siendo los más beneficiados. Sin embargo, los resultados dependen de la suposición de que las

medidas de reforma e inversión se implementarán completamente y estén sujetas a una gran cantidad de incertidumbre y limitaciones.

Algunas de estas limitaciones se refieren a los efectos positivos (probables) provenientes del resto de la UE debido a la atención centrada en la zona del euro. Estas limitaciones son abordadas en parte por las propuestas de Pfeiffer *et al.* (2023). Estos autores consideran explícitamente los efectos secundarios del comercio y amplían el modelo QUEST de la Comisión Europea para capturar de manera más detallada la dinámica de las inversiones públicas, incorporándolo en una estructura multi-país diseñada para el análisis de efectos secundarios y vínculos comerciales. Según sus resultados, en un escenario de 6 años, el nivel del PIB real en la UE-27 podría ser más del 1.2 % más alto en 2026 en comparación con un escenario sin cambios en las políticas. Sin embargo, factores como la reacción de la política monetaria, los efectos mejoradores de la productividad al estímulo de inversión y la velocidad a la que se distribuyen las ayudas determinarán los efectos macroeconómicos finales de NGEU.

Así, los efectos para cada país específico se miden teniendo en cuenta los efectos de secundarios que, en el caso de España, por ejemplo, se sumarían a un efecto medio anual para el horizonte de 6 años (hasta 2026) del 2.09 %, que no está lejos del 2.4 % encontrado por Picek (2020). Picek y Schröder (2017, 2018) utilizan un modelo multirregional de input-output basado en la Base de Datos del World Input-Output Database (WIOD, por sus siglas en inglés) y también encontraron que Grecia sería el país más beneficiado.

Aunque los estudios a nivel de un solo país son escasos, existen estudios específicos para estos contextos, como Malliaropulos *et al.* (2021) para Grecia, y Fernández-Cerezo *et al.* (2023) para España. Este último estudio presenta un modelo de red de producción para la economía española, calibrado con una matriz de I-O de la Base de Datos del World Output Database (Timmer *et al.*, 2015), que se aplica al impacto de los fondos NGEU. Sus resultados difieren de los de Pfeiffer *et al.* (2023) y Picek (2020), ya que encontraron un impacto más modesto, entre el 1.15 % y el 1.75 %. Por lo tanto, dado que los resultados no coinciden completamente entre los estudios, se agrega más evidencia, especialmente teniendo en cuenta la dimensión regional de los datos.

Barbero *et al.* (2023) proporcionan esa evidencia, contribuyendo al debate sobre el impacto macroeconómico del MRR al considerar el nivel regional (NUTS2). Para ello,

adopta un modelo de equilibrio general espacial que, al aplicarse a 230 regiones europeas, pronostica un aumento promedio entre el 0.85 % y el 1.36 % (dependiendo de si se incluyen o no préstamos).

Mientras que toda esta evidencia (relativamente limitada) ha considerado los métodos mencionados por la literatura anterior, como se revisa en la subsección 4.2.1, hay evidencia limitada con respecto a la aplicación de técnicas de inferencia causal (particularmente contrafactuales), como las empleadas en Alloza y Sanz (2021) y Carozzi y Repetto (2019), para evaluar el impacto del *Plan E* en España. Hasta donde alcanza nuestro conocimiento, nuestro estudio representa el primer uso de técnicas de inferencia causal para evaluar los efectos de NGEU.

Nuestros métodos también son particularmente adecuados para nuestro caso, ya que pueden acomodar efectos espaciales más fácilmente (similar a Alloza y Sanz, 2021), aunque dejamos esto para investigaciones futuras.

4.3. Estrategia empírica

Como se afirma y argumenta en la introducción de este capítulo y se desarrolla en esta sección, en este análisis utilizaremos técnicas contrafactuales (Carrillo y Rothbaum, 2016; Chernozhukov *et al.*, 2013). Los contrafactuales implican imaginar escenarios hipotéticos en los que se crea una realidad o línea de tiempo alternativa basada en circunstancias diferentes. Estos escenarios se utilizan comúnmente en filosofía y ciencias sociales para explorar las posibles consecuencias de diferentes elecciones y acciones. Al considerar los contrafactuales obtenemos ideas sobre cómo podrían haberse desarrollado los eventos de manera diferente si se hubieran alterado ciertos factores, lo que nos permite examinar las relaciones de causa y efecto. Básicamente, los contrafactuales proporcionan un marco para comprender los resultados potenciales que podrían haber ocurrido si se hubieran tomado decisiones o circunstancias diferentes, manteniendo constantes otros factores, es decir, *cæteris paribus*.

En este estudio, estamos interesados en explorar el escenario hipotético (el contrafactual) de que las regiones españolas (NUTS2) no hubieran recibido los fondos NGEU. Para abordar esta pregunta, enfrentamos dos desafíos principales. El primer desafío es

encontrar grupos de control naturales que se asemejen a las regiones españolas pero que no hayan recibido fondos NGEU. Como se indicó anteriormente, aplicamos el método del control sintético (SCM) para superar esta limitación. Si bien los modelos de equilibrio general estocástico dinámico (DSGE) son valiosos por sus profundas capacidades analíticas, a menudo requieren supuestos teóricos detallados (Fernández-Villaverde, 2010). Esto puede agregar capas de complejidad y restar claridad a la atribución de los resultados de las políticas (Chari *et al.*, 2009). El SCM, por otro lado, adopta un enfoque basado en datos, aprovechando la información disponible para crear un análogo sintético de las regiones tratadas. Este método es especialmente adecuado para analizar el impacto de algunas iniciativas a nivel regional, donde cada unidad tiene su propio perfil económico (por ejemplo, Castillo *et al.*, 2017; Mora-Sanguinetti y Spruk, 2023; Pinotti, 2015). Al adoptar el SCM, buscamos fomentar un proceso analítico más transparente y empíricamente fundamentado que mejore nuestra comprensión del efecto de este enorme paquete de estímulos, en las regiones españolas, sin la modelización intrínseca de supuestos económicos específicos de cada región (Stiglitz, 2018).

El segundo desafío surge debido a que la mayoría de las regiones europeas que pueden actuar como donantes para nuestras regiones sintéticas españolas han recibido en mayor o menor medida fondos NGEU, lo que hace imposible obtener un efecto neto del tratamiento de los fondos. Sin embargo, debido a que el nivel de fondos recibidos varía significativamente entre los países de la UE, podemos obtener al menos una imagen parcial del efecto de los fondos.

España es uno de los principales beneficiarios de los fondos, en comparación con otros países, y como tal, su posición destacada en la distribución de la asignación de fondos en relación con su PIB total permite una comparación con aquellas regiones que han recibido considerablemente menos fondos en relación con su PIB total. A partir de esta comparación, podemos abordar la siguiente pregunta: *¿Cuál habría sido el desarrollo económico de las regiones españolas si hubieran recibido un nivel de fondos, en relación con el PIB, equivalente al percentil 25 o inferior de la distribución total de los países de la UE?*

Para responder a esta pregunta, primero seleccionamos las regiones europeas en países que han recibido un volumen de fondos por debajo del percentil 25 en relación con su

PIB total.[4] Estas regiones actuarán como donantes en nuestro análisis. Una vez que se ha identificado este grupo de países donantes[5], aplicaremos el SCM para crear regiones sintéticas españolas utilizando información socioeconómica y características de las regiones de control NUTS2 en estos países, en línea con la literatura sobre crecimiento económico (Arribas *et al.*, 2020; Sala-i-Martin, 1996; Tortosa-Ausina *et al.*, 2005).

El método del control sintético nos permitirá construir regiones que se asemejen estrechamente a las regiones españolas en términos de variables relevantes (por ejemplo, inversión, nivel educativo, crecimiento de la población, etc.) pero que no hayan recibido los mismos fondos NGEU. Al utilizar esta técnica, podemos comparar el desarrollo económico de las regiones españolas con un escenario hipotético en el que sí se hubieran recibido fondos equivalentes al percentil 25 o inferior de la distribución total de los países de la UE.

4.3.1. El Método del Control Sintético: Metodología

En esencia, y como se mencionó anteriormente, el método del control sintético se basa en la idea de que una combinación ponderada de unidades no tratadas puede proporcionar un grupo de control adecuado cuando el número de unidades tratadas es pequeño. La intuición del método es que construimos nuestras regiones españolas "sintéticas" basándonos en una combinación ponderada de las demás regiones que actuarán como donantes, es decir, aquellas regiones que caen por debajo del percentil 25 en cuanto a la proporción total de fondos NGEU con respecto al PIB. Al hacer esto, pretendemos trazar la trayectoria que nuestra variable de interés (PIB per cápita) habría seguido si el nivel de fondos recibidos, para las regiones españolas, hubiera permanecido por debajo del percentil 25.

Formalmente, utilizando la notación de Abadie (2021), asumiendo que observamos $J+1$ regiones: $j = 1, 2, \ldots, J+1$, y luego suponiendo que la primera región es la única expuesta al evento $j = 1$, es decir, la unidad tratada. De esta manera, las J regiones restantes corresponden al grupo "donante". De manera similar, consideramos que la unidad tratada $j = 1$ está continuamente expuesta al tratamiento. Supongamos que nuestro conjunto de

[4] La información sobre la proporción por país de los fondos NGEU totales respecto al PIB total se obtuvo de la Comisión Europea.

[5] Estos países son: Dinamarca, Alemania, Irlanda, Luxemburgo, Países Bajos, Finlandia y Suecia.

datos abarca T períodos, con T_0 siendo los períodos antes del inicio del evento; por lo tanto, tenemos $1 \leq T_0 < T$. Sea Y_{jt} el resultado de interés para cada unidad j y período t. Siguiendo la misma notación, definimos Y_{jt}^N como el resultado potencial sin intervención para la región j y el período t. En consecuencia, caracterizamos Y_{jt}^I como el resultado potencial bajo intervención. Este último resultado se aplica solo a la unidad afectada por el tratamiento $j = 1$ durante el período posterior a la intervención $t > T_0$. Por último, si buscamos evaluar el efecto del evento en la unidad tratada, derivamos la siguiente ecuación:

$$\tau_{1t} = Y_{1t}^I - Y_{1t}^N \tag{4.1}$$

donde τ_{1t} es el efecto del tratamiento para la unidad afectada en el período t, satisfaciendo $(t > T_0)$. Nótese que en la Ecuación (4.1), el efecto del shock puede cambiar con el tiempo, dando lugar a diferentes valores para cada período t. En este escenario, solo observamos la evolución del PIB per cápita de las regiones españolas en su estado actual, es decir, Y_{1t}^I. En consecuencia, necesitamos desarrollar una unidad de comparación adecuada para ver cuál habría sido el resultado en ausencia de este shock (el shock, en este caso, corresponde al nivel real de fondos recibidos por las regiones españolas). Esencialmente, necesitamos estimar Y_{1t}^N, que representaría una "región española sintética" que ha recibido un nivel de fondos NGEU por debajo del percentil 25. Repetimos este proceso para cada una de las 17 comunidades autónomas españolas, eliminando cada vez las otras 16 regiones de la muestra.

Así, Y_{1t}^N se obtiene como media ponderada de las unidades no tratadas que mejor replican las características de la unidad tratada antes del período de intervención. Matemáticamente, nuestras "regiones españolas sintéticas" se definen por un vector $\mathbf{W} = (w_2, \ldots, w_{J+1})$ de pesos de tamaño $J \times 1$, donde $\mathbf{W}$ son combinaciones de regiones que caen por debajo del percentil 25 de la distribución. Con esta selección, los sintéticos potenciales para cada región española se representan por:

$$\widehat{Y}_{1t}^N = \sum_{j=2}^{J+1} w_j Y_{jt} \tag{4.2}$$

y, por lo tanto, el estimador del efecto del tratamiento mostrado en la ecuación (4.1) es:

$$\widehat{\tau}_{1t} = Y_{1t}^{I} - \sum_{j=2}^{J+1} w_j Y_{jt} \tag{4.3}$$

Además, necesitamos un conjunto de k posibles predictores para la variable dependiente para aplicar el método de control sintético. En este sentido, por un lado, tenemos $\mathbf{X}_1$, que es un vector de tamaño $k \times 1$ que contiene los valores de covariables (predictores) antes de la intervención en la unidad tratada. Por otro lado, tenemos $\mathbf{X}_0$, que es una matriz de tamaño $k \times J$ que contiene los valores de las mismas covariables para las regiones no tratadas. Típicamente, estos predictores incluyen valores retardadas de la variable respuesta, así como determinantes conocidos de la variable respuesta.

El desafío fundamental de esta metodología radica en seleccionar la combinación óptima de pesos $\mathbf{W} = (w_2, \ldots, w_{J+1})$. Para este escenario, seguimos las recomendaciones de Abadie y Gardeazábal (2003) y Abadie *et al.* (2010). Por lo tanto, seleccionamos el vector de pesos óptimo $\mathbf{W}^*$ que minimiza la siguiente expresión:

$$|\mathbf{X}_1 - \mathbf{X}_0\mathbf{W}| = \left(\sum_{h=1}^{k} v h \left(X_{h1} - w_2 X_{h2} - \ldots - w_{J+1} X_{hJ+1} \right)^2 \right)^{1/2} \tag{4.4}$$

Dado que los pesos óptimos $\mathbf{W}^*$ que minimizan la Ecuación (4.4) dependen de $\mathbf{V} = (v_1, \ldots, v_k)$, la selección de $\mathbf{V}$ implica decidir la importancia relativa asignada a cada covariable (predictor) para minimizar la Ecuación (4.4), midiendo en última instancia la discrepancia entre $\mathbf{X}_1$ y $\mathbf{X}_0\mathbf{W}$. En este sentido, el $\mathbf{V}$ óptimo se elige siguiendo las pautas de Abadie (2021), seleccionando aquél que mejor reproduce la trayectoria de la variable de la unidad tratada durante el período pre-intervención. Posteriormente, se selecciona $\mathbf{W}(\mathbf{V})$ para minimizar el error cuadrático medio de la predicción (RMSPE), que mide el grado de ajuste entre la trayectoria de la variable de interés (PIB per cápita) en la región tratada y su contrafactual. De hecho, este es el objetivo principal del método, ya que un RMSPE más bajo antes de la intervención indica un mejor ajuste de nuestra región sintética y, por lo tanto, una mayor confianza del efecto potencial mostrado durante el período post-tratamiento.

4.3.2. Inferencia en SCM

En el contexto de los métodos de control sintético, la inferencia desempeña un papel crucial para abordar la significación estadística de los efectos del tratamiento en una unidad específica. Este enfoque, que tiene como objetivo evaluar el impacto de una intervención al comparar la unidad tratada con una combinación ponderada de unidades no tratadas, se basa en estrategias clave para mejorar su robustez inferencial.

Una estrategia común en este sentido es el análisis de placebo (Galiani y Quistorff, 2017), una técnica donde la metodología se aplica a unidades que no han estado sujetas a la intervención. Este enfoque proporciona una línea base crítica para evaluar la validez de los efectos observados en la unidad tratada. Al utilizar estas "unidades de placebo", podemos distinguir entre cambios reales asociados con la intervención y variaciones aleatorias, mejorando así la capacidad del método para ofrecer inferencias válidas y fiables.

Otra herramienta para evaluar la precisión de estos modelos es la comparación del error cuadrático medio de predicción (RMSPE) antes y después de la intervención (Abadie, 2021). El RMSPE proporciona evidencia cuantitativa del impacto de la intervención al medir la precisión predictiva antes y después del evento. Sin embargo, estas estrategias inferenciales presuponen la existencia de unidades de comparación capaces de reflejar con precisión las características preintervención de la unidad tratada. Si dichas unidades de comparación son subóptimas pero representan las mejores opciones disponibles, surge un desafío importante para una inferencia adecuada. Un desajuste en la trayectoria preintervención entre las unidades tratadas y sintéticas puede llevar a inferencias engañosas (Ferman y Pinto, 2019), advirtiendo que el efecto del tratamiento puede confundirse con diferencias preexistentes. En estas circunstancias, los investigadores deben reconocer las limitaciones de estas herramientas de inferencia y considerar procedimientos alternativos (Cattaneo *et al.*, 2021; Chernozhukov *et al.*, 2021; Fratesi y Wishlade, 2017).

La selección adecuada de las unidades de placebo es un aspecto crucial de la inferencia en los métodos de control sintético. Estas unidades se seleccionaron según criterios sólidos de similitud con la unidad tratada, asegurando que las características observables y no observables relevantes sean comparables. En este sentido, se analizan cuidadosamente las tendencias pretratamiento en las unidades de placebo para evaluar si reflejan de manera

consistente el comportamiento de la unidad tratada antes de la intervención.

También es necesario considerar la estabilidad temporal de las unidades de placebo, asegurando que mantengan una evolución temporal similar a la unidad tratada. Otro aspecto interesante es la relevancia contextual de las unidades de placebo, teniendo en cuenta las características específicas del contexto y la intervención. Además, contar con un número suficiente de unidades de placebo es esencial para mejorar la validez de la comparación. Para evaluar el impacto del SCM, hemos aplicado pruebas de hipótesis como parte fundamental del proceso inferencial. Estas pruebas desempeñan un papel crucial en la determinación de la significación estadística de las estimaciones y proporcionan una base rigurosa para la toma de decisiones informada.

En nuestro contexto específico, las pruebas de hipótesis de los modelos contrafactuales, específicamente aquellos basados en métodos de control sintético, se revelan como insuficientes para capturar con precisión el efecto subyacente que los fondos del NGEU tienen en España, a pesar de la asignación significativa de recursos que el país ha recibido. Debido en parte al conservadurismo de las pruebas actualmente disponibles y también como consecuencia de la dificultad para imitar las trayectorias pretratamiento de las regiones españolas con los donantes disponibles, no se puede concluir un impacto significativo cuando se observa cada serie de manera individual. Sin embargo, cuando se contemplan en su conjunto, la evidencia de un impacto es abrumadora. En nuestros análisis, operacionalizamos esto mediante el uso de pruebas unilaterales de significación de medias, pruebas de diferencia de medias antes y después, y pruebas de medias con restricciones. Por lo tanto, realizamos diversas pruebas paramétricas, que aprovechan el carácter multiunitario de nuestras estimaciones, para probar la supuesta limitación de estos modelos y validar la existencia de un efecto.

La idea no es considerar el impacto estimado de cada región de manera aislada, sino como parte de un conjunto de estimaciones interrelacionadas. En última instancia, todos los impactos estimados son consecuencia de la misma intervención. Esto otorga más robustez a las pruebas de hipótesis individuales y nos permite evaluar si existe evidencia acumulada de un impacto real. Nos preguntamos si los resultados observados en su conjunto podrían haberse obtenido por azar en ausencia de un impacto, concluyendo, en nuestro estudio, que este no puede ser el caso. Estas pruebas destacan de manera

concluyente el impacto positivo que los fondos del NGEU están generando.

En este sentido, una de las pruebas más comúnmente utilizadas es la significación de medias. Al evaluar la diferencia entre las medias de las variables de interés antes y después de la intervención, esta prueba proporciona información sobre la magnitud y dirección del impacto. Un resultado estadísticamente significativo respalda la existencia de un impacto real y medible.

También hemos utilizado pruebas de hipótesis con restricciones en parámetros específicos del modelo. Estas pruebas permiten a los investigadores evaluar hipótesis específicas relacionadas con el impacto de la intervención en momentos particulares o en subconjuntos específicos de la población, proporcionando una comprensión más detallada del fenómeno estudiado. Cuando se utilizan en conjunto, estas pruebas de hipótesis ayudan a fortalecer la validez de la inferencia realizada a través de los modelos de control sintético. Al abordar diversas dimensiones del impacto y considerar diferentes escenarios, estas pruebas ofrecen un enfoque integral para evaluar rigurosamente los efectos de las intervenciones estudiadas.

4.3.3. Datos

Nuestra investigación utiliza un conjunto de datos que contiene datos de panel a nivel regional para las regiones europeas NUTS2, abarcando los años 2017 a 2024. Los años 2022, 2023 y 2024 corresponden a proyecciones realizadas por la Comisión Europea. Las previsiones regionales son proporcionadas por AMECO, la base de datos macroeconómica anual de la Dirección General de Asuntos Económicos y Financieros de la Comisión Europea, a nivel de país, y se desarrollan bajo la supervisión tanto del Centro Común de Investigación como de la Dirección General de Política Regional y Urbana. Estas proyecciones aseguran uniformidad en la construcción del modelo contrafactual, ya que cada una de las proyecciones de las regiones donantes fue preparada por la misma entidad y con los mismos estándares de calidad.

En nuestro análisis, consideramos el PIB per cápita como la variable dependiente, con el objetivo de proporcionar evidencia empírica adicional sobre los efectos del programa Next Generation EU (NGEU). Basándonos en investigaciones previas (Bańkowski *et al.*,

2022, 2021; Barbero *et al.*, 2023; Picek, 2020), buscamos explorar cómo la iniciativa NGEU ha impactado en este indicador macroeconómico.

Para asegurar la solidez de nuestros hallazgos y mitigar la influencia de factores externos confusos, incorporamos los principales determinantes del crecimiento del PIB, como se describe en (Mankiw *et al.*, 1992). Este enfoque nos permite desentrañar los efectos específicos del programa NGEU de las tendencias macroeconómicas más amplias e influencias en el PIB per cápita. Estos controles incluyen: la formación bruta de capital fijo ($GFCP$), que refleja la inversión en activos físicos; la proporción de la población con educación terciaria ($EDUC$), capturando el impacto de la educación en la productividad y la innovación de acuerdo con Barro (2001); y el crecimiento de la población (POP_GROWTH), que puede impulsar la producción económica a través de una fuerza laboral más grande o potencialmente tensionar los recursos cuando supera las tasas de crecimiento económico. También se incluye la densidad de población (POP_DENS), que controla el tamaño de las regiones. La información sobre la definición específica de las variables, así como sus estadísticas descriptivas, se puede encontrar en las Tablas 4.1 y 4.2.

Tabla 4.1: Variables y fuente de los datos.

Variable	**Definición**	**Fuente**
$GDPpc$	PIB per cápita a precios corrientes (NUTS2)	ARDECO
POP_DENS	Población total (NUTS2)/Km^2 (NUTS2)	ARDECO
$GFCF$	Formación bruta de capital fijo (NUTS2) / PIB (NUTS2)	ARDECO
$EDUC$	Proporción de la población total con educación terciaria (NUTS2)	ARDECO
POP_GROWTH	Tasa de crecimiento de la población total (NUTS2)	ARDECO

Tabla 4.2: Estadísticas descriptivas, variables relevantes.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
$GDPpc$	712	42,745.41	16,594.01	17,198.95	135,837.20
POP_DENS	534	33.59	569.75	3.40	4.339
$GFCF$	712	0.219	0.057	0.130	0.944
$EDUC$	534	35.61	8.19	20.60	57.50
POP_GROWTH	712	0.471	0.485	–1.36	2.46

Nota: Todas las variables se miden a nivel regional. $GDPpc$, $GFCF$ y POP_GROWTH contienen información para el período 2017-2024. POP_DENS y $EDUC$ contienen información para el período 2017-2022.

4.4. Resultados

Esta sección presenta, analiza y compara la evolución de la serie (esperada) del PIB per cápita de las regiones españolas en el escenario real, en el cual se implementan los fondos Next Generation (NGEU), y en el escenario simulado, aquellas regiones europeas que recibieron menos financiación (por debajo del percentil 25). El objetivo es medir cuánto del crecimiento económico (esperado) de las regiones españolas se puede atribuir a los fondos del NGEU y determinar si, a pesar del inherente conservadurismo de las pruebas de hipótesis del método de control sintético (SCM) (Galiani y Quistorff, 2017), este tratamiento (los fondos) podría tener un impacto significativo en el desarrollo económico de las regiones españolas. Para abordar estos problemas, realizamos un análisis comparativo a lo largo del tiempo, considerando el año 2021 como referencia para dividir los períodos pre y post-tratamiento.

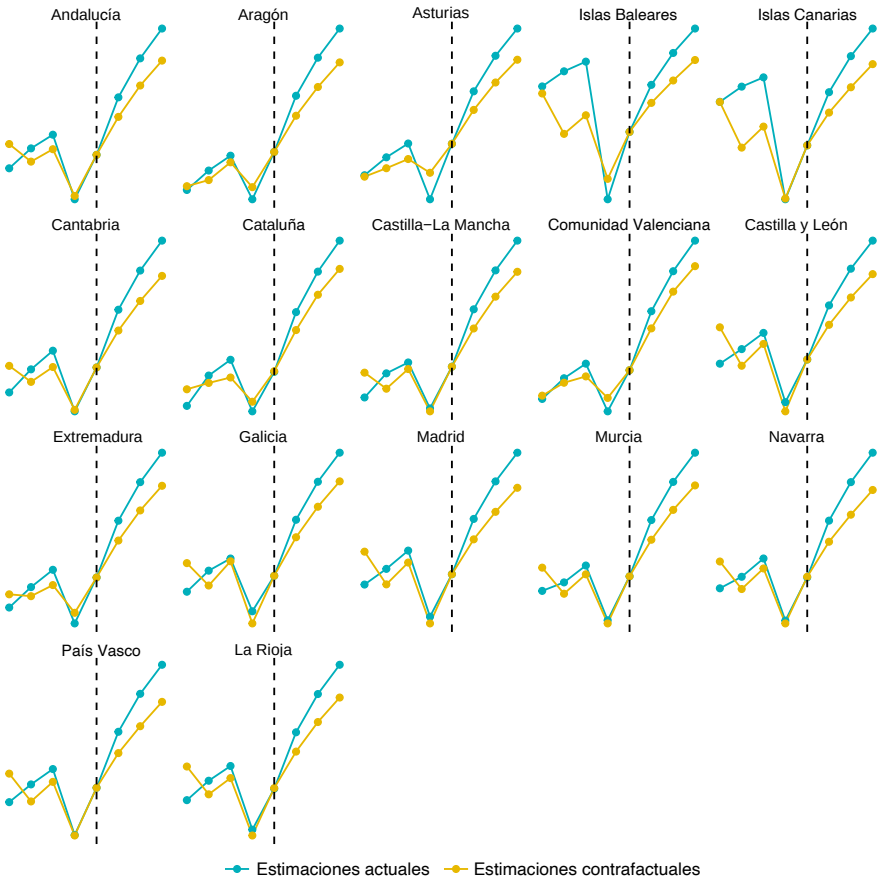
4.4.1. Midiendo los impactos

Después de la crisis del COVID-19, se espera que la implementación de medidas de política económica, como el despliegue de los fondos del NGEU por parte de Europa, tenga un impacto positivo tanto en el crecimiento nacional como regional, especialmente en países como España que sufrieron graves repercusiones a raíz de la pandemia.

La Figura 4.1 muestra estimaciones reales y contrafactuales de la evolución del PIB per

cápita para las regiones españolas a precios corrientes, demostrando que el impacto de la crisis del COVID-19 fue bastante desigual entre las regiones. Estos impactos regionales dispares pueden explicarse parcialmente por diversos factores, incluidas las estructuras productivas regionales o la alta dependencia del turismo en algunas regiones. De hecho, las regiones que sufrieron más fueron aquellas que dependen más intensamente del turismo (Duro *et al.*, 2021).

Figura 4.1: Estimaciones reales y contrafactuales de la evolución del PIB per cápita para las regiones españolas.



Nota: representación gráfica del PIB per cápita para cada una de las regiones españolas en los grupos experimental y sintético durante el período 2017-2024.

Los resultados en la Figura 4.1 se presentan divididos en dos periodos temporales: (i) el período pretratamiento (2017-2020), que se refiere al marco temporal previo a la

asignación de fondos; y (ii) el período post-tratamiento (2022-2024). Una comparación de las relaciones entre la evolución de las estimaciones correspondientes a los periodos pre y post-tratamiento revela una mayor cercanía y diversidad entre las curvas reales y contrafactuales durante el período pretratamiento.

Por un lado, observamos que la distancia media entre las dos curvas es menor durante el período pretratamiento que durante el post-tratamiento, y al mismo tiempo presenta más variabilidad (ver Tabla 4.3). Por otro lado, observamos una relación consistente y sistemática entre ambos conjuntos de líneas en el periodo post-tratamiento. Las líneas reales son consistentemente más altas que las contrafactuales y, además, presentan una tendencia en la que la distancia entre ellas aumenta con el tiempo. De hecho, el análisis de la Tabla 4.3, que informa sobre las medidas descriptivas resumidas de las diferencias entre las estimaciones reales y contrafactuales en los periodos pre y post-tratamiento, revela que las discrepancias medias en el período pretratamiento son aproximadamente diez veces más pequeñas que en el período post-tratamiento. Este hallazgo sugiere una diferencia significativa en las dinámicas de crecimiento entre los dos períodos y es coherente con la hipótesis de que los fondos tienen un impacto positivo en el crecimiento económico regional en España.

Tabla 4.3: Medidas descriptivas de las diferencias entre las estimaciones reales y contrafactuales en los períodos pre y post-tratamiento.

Tratamiento	**Media**	**Desv. Típica**	**Total casos**
Pre-tratamiento	130.7573	892.0527	68
Post-tratamiento	1129.9890	387.2677	51
Total	558.9994	872.9100	119

El análisis detallado del impacto de los fondos muestra un impacto generalmente positivo para todas las regiones españolas, sin excepciones, como se muestra en la Tabla 4.4. Un examen más minucioso de esta tabla revela que, en media, las regiones españolas habrían tenido un 3 % menos de PIB per cápita (a precios corrientes) en 2022, si no hubieran recibido los fondos del NGEU. Esta diferencia aumenta a un 4.1 % en 2023 y luego a un 4.5 % en 2024. Nuevamente, esto significa que, si España no hubiera recibido los fondos,

el PIB per cápita medio en las regiones habría sido un 3 %, 4.1 % y 4.5 % menor para 2022, 2023 y 2024, respectivamente. Como se puede observar, este impacto sostenido y creciente subraya la efectividad continua de esta política para fomentar el desarrollo económico regional.

Tabla 4.4: Efectos para las regiones españolas.

Región	**Year**		
	2022	**2023**	**2024**
Andalucía	2.9	3.8	4.3
Aragón	3.0	4.2	4.7
Asturias	3.0	4.1	4.6
Cantabria	3.1	4.2	4.7
Castilla-La Mancha	2.8	3.7	4.2
Castilla León	2.6	3.1	3.7
Cataluña	3.1	4.3	4.8
Comunidad Valenciana	2.9	4.0	4.5
Extremadura	2.5	2.8	3.3
Galicia	3.0	4.0	4.5
Islas Baleares	3.3	4.7	5.2
Islas Canarias	3.3	4.7	5.2
La Rioja	2.7	3.6	4.0
País Vasco	3.2	4.7	5.1
Madrid	3.1	4.4	4.9
Murcia	3.0	4.0	4.5
Navarra	3.2	4.6	5.1
Media	3.0	4.1	4.5

Nota: Porcentaje del producto interior bruto per cápita efectivo que se puede atribuir al tratamiento.

Deberíamos enfatizar que este impacto positivo observado se atribuye a la mayor proporción total de fondos del Next Generation (NGEU) con respecto al PIB de España en

comparación con las regiones del grupo de control. Como se mencionó anteriormente, el grupo de control en este análisis está compuesto por regiones por debajo del percentil 25 de la distribución de la proporción total de fondos NGEU/PIB. Por lo tanto, los efectos que observamos son probablemente conservadores, y es muy probable que, con un grupo de control "puro" (es decir, un grupo sin fondos NGEU), el impacto real sería aún más sustancial.

Profundizando más en detalle, un aspecto particularmente notable es la heterogeneidad en el impacto entre las regiones. Por ejemplo, se espera que las Islas Baleares y Canarias experimenten un efecto sustancial, equivalente al 5.2 %. En el extremo opuesto, aunque todas las regiones experimentarán un crecimiento positivo, algunas de ellas están ligeramente por debajo de la media nacional. Extremadura y Castilla y León, por ejemplo, muestran impactos del 3.3 % y 3.7 %, respectivamente. Sin embargo, también debemos tener en cuenta que los factores que contribuyen a estos efectos regionales heterogéneos pueden ser intrínsecos y van más allá del alcance de esta investigación. A pesar de esta complejidad, se debe reconocer la contribución sustancial que estos fondos hacen al PIB per cápita regional y el hecho de que todas las regiones parecen beneficiarse de ellos.

4.4.2. Inferencia

En el contexto de los modelos contrafactuales, se anticipa que el impacto generado por la intervención de una política específica será significativo si dicha intervención ha tenido un efecto tangible y real. Sin embargo, a pesar de observar una tendencia y magnitud notable de este impacto, en algunas ocasiones, las pruebas de hipótesis asociadas con el modelo contrafactual pueden no ser completamente capaces de capturar de manera concluyente este efecto significativo. Esto es exactamente lo que ocurre con los resultados obtenidos cuando son evaluados mediante las pruebas predeterminadas (Abadie *et al.*, 2010). Para entender este resultado paradójico, debemos tener en cuenta la naturaleza conservadora de estas pruebas, que pueden no ser completamente sensibles para detectar tal efecto (Galiani y Quistorff, 2017).

No rechazar la hipótesis nula con respecto al efecto de los fondos lleva a lo que se conoce en estadística como un "Error Tipo II". Este tipo de error ocurre cuando la prueba

no tiene suficiente potencia, lo que lleva a una conclusión incorrecta. Aunque esto es bastante obvio para cualquier persona familiarizada con estadísticas a nivel universitario, debemos prestar especial atención a esto en nuestro caso específico, ya que la prueba predeterminada afirma que la medida económica no tiene efecto, lo que puede llevar a interpretaciones erróneas del estudio. Esta observación subraya la importancia de no solo considerar los resultados estadísticos obtenidos mediante modelos contrafactuales, sino también examinar la tendencia general y la magnitud del impacto.

Una manera efectiva de mitigar este tipo de error es garantizar un tamaño de muestra suficiente (Akobeng, 2016). Sin embargo, en nuestro caso, esto no es factible debido a la implementación reciente de los fondos en el verano de 2021 y al retraso en la disponibilidad de información regional de las instituciones públicas. A pesar de esto, en el período post-tratamiento, se observa un crecimiento más pronunciado en todas las regiones del grupo experimental en comparación con el grupo sintético. En nuestra opinión, estos resultados no pueden atribuirse simplemente al azar, sino que señalan la existencia de un efecto significativo generado por la intervención. Por lo tanto, a continuación, implementamos un conjunto de pruebas alternativas para decidir si estos resultados sistemáticos, observados como conjunto, pueden atribuirse al azar o no.

El contraste de hipótesis relativa a la diferencia de medias entre los valores observados de las regiones reales y las contrafactuales ($H_0 : \mu_d^b = 0$) revelan resultados no estadísticamente significativos ($p-$valor : 0.2310). Es decir, antes de la implementación del tratamiento, no hubo diferencias sustanciales entre los valores registrados para el grupo experimental y el grupo sintético. Por el contrario, una prueba sobre las diferencias medias después del tratamiento ($H_1 : \mu_d^a > 0$) concluye resultados estadísticamente significativos ($p-$valor < 0.0001). Además, si consideramos la distribución de las diferencias post-tratamiento y las comparamos con la distribución teórica de las diferencias pre-tratamiento, podemos concluir sin lugar a dudas que la serie de valores reales es sistemáticamente más alta ($p-$valor < 0.0001) que la serie de valores contrafactuales (ver Tabla 4.5, donde S_d denota la desviación típica estimada de las diferencias pre-tratamiento).

Tabla 4.5: Distribución esperada de las diferencias post-tratamiento bajo la hipótesis de tratamiento no efectivo frente a la distribución real.

	$(-\infty, -S_d)$	$(-S_d, 0)$	$(0, S_d)$	$(S_d, +\infty)$
Esperado	8.09	17.41	17.41	8.09
Actual	0	0	17	34

Nota: Número de regiones esperado si la distribución observada en el período pre-tratamiento se mantuviera igual en el período post-tratamiento. S_d hace referencia a la desviación típica

El contenido de la Tabla 4.5 corrobora la existencia de diferencias sistemáticas a nivel individual entre la evolución de las series antes y después del tratamiento. Para medir la magnitud de estas diferencias a nivel agregado, hemos implementado una prueba adicional con el propósito de determinar cuántas desviaciones por encima de la media pre-tratamiento se encuentra la media de las diferencias post-tratamiento:

$$\begin{cases} H_0: & \mu_d^a \leqslant \mu_d^b + K\left(\frac{S_d^b}{\sqrt{67}}\right) \\ H_1: & \mu_d^a > \mu_d^b + K\left(\frac{S_d^b}{\sqrt{67}}\right) \end{cases} \tag{4.5}$$

Cuando se exploraron diversos valores de K en el conjunto $1, 2, 3, \ldots, 9, 10$, observamos que, para todos los casos hasta $K = 9$, la prueba arroja resultados estadísticamente significativos ($p-\text{valor}: 0.0028$, para $K = 9$).

El análisis descriptivo no solo revela que las diferencias son mayores en el período post-tratamiento, sino que también señala un ensanchamiento sistemático de estas diferencias con el tiempo. Por lo tanto, la fase final del análisis aplica pruebas de hipótesis restringidas (Silvapulle y Sen, 2005). Este enfoque nos permite evaluar hipótesis en las cuales ciertos parámetros o relaciones están sujetos a restricciones específicas. En nuestro estudio, aplicamos pruebas de hipótesis restringidas para evaluar la hipótesis de que las diferencias medias entre los años post-tratamiento experimentaron un aumento anual significativo.

$$\begin{cases} H_0: & \mu_{d,2022} < \mu_{d,2023} < \mu_{d,2024} \\ H_1: & No\ H_0 \end{cases} \tag{4.6}$$

Para llevar a cabo este análisis, hemos implementado un enfoque de dos fases. La primera fase consiste en un análisis de varianza (ANOVA) en el cual la variable dependiente está definida por las diferencias y como variable independiente un factor que indica el año. Los resultados del ajuste del modelo, detallados en la Tabla 4.6, indican que todas las categorías son estadísticamente significativas, con un ajuste notable, R^2 igual a 0.935.

Tabla 4.6: Modelo ANOVA para las diferencias post-tratamiento[a,b].

	Año	**Media**	**Std. Error**	**t-value**	**Pr(>|t|)**
Grupo	2022	818.920	75.976	10.779	$< 0.0001^{***}$
	2023	1,188.349	75.976	15.641	$<0.0001^{***}$
	2024	1,382.698	75.976	18.199	$<0.0001^{***}$
RSE			313.26		
R^2			0.935		

[a] $^{*}\ p < 0.1$; $^{**}\ p < 0.05$; $^{***}\ p < 0.01$.

[b] Análisis de varianza, utilizando las diferencias en el PIB per cápita entre los grupos sintético y experimental como la variable dependiente, y cada uno de los años en el período post-tratamiento como la variable independiente.

Posteriormente, se ha aplicado la prueba de hipótesis restringida utilizando el paquete `restriktor` (Vanbrabant y Kuiper, 2023) de `R` (R Core Team, 2022). Los resultados de la prueba (ver Tabla 4.7) indican claramente que hay suficiente evidencia estadística para considerar que la magnitud del efecto tiende a aumentar con el tiempo. Este resultado es coherente con la hipótesis de un efecto de tratamiento significativo. Por lo tanto, a la luz de estos resultados, podemos confirmar que se espera que la intervención tenga un impacto significativo y positivo en la evolución de la serie española del PIB per cápita.

Tabla 4.7: Prueba de hipótesis restringida.

Contraste	H_0	H_1	$\bar{F}$**-test statistic**	**p-value**
Hipótesis	Todas las restricciones se mantienen en la población	Al menos una restricción se incumple	0.0000	1.0000

[a] * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

[b] Prueba de hipótesis para el aumento en las diferencias medias entre 2022 y 2024.

4.5. Conclusiones

Las consecuencias económicas de la pandemia del COVID-19 persisten y se han agravado debido a las recientes guerras en Ucrania e Israel-Palestina. La pandemia ha afectado a los países independientemente de su nivel de desarrollo, aunque de manera diferente, y, como tal, las reacciones políticas y las medidas de contención también difirieron (Giménez *et al.*, 2023). En el caso específico de la Unión Europea, cuyos miembros fueron gravemente afectados por la crisis sanitaria (especialmente España e Italia), la Comisión Europea diseñó el programa conocido como Next Generation EU, implementado a través del Mecanismo de Recuperación y Resiliencia (MRR), que representa la mayor parte del NGEU.

Dado lo reciente de esta iniciativa, la evidencia disponible sobre la efectividad del mecanismo, especialmente en términos de impactos económicos, sigue siendo limitada, y la mayoría de los estudios se centran ya sea a nivel de la UE (Bańkowski *et al.*, 2022, 2021; Pfeiffer *et al.*, 2023; Picek, 2020) o a nivel de país (Fernández-Cerezo *et al.*, 2023; Malliaropulos *et al.*, 2021), con evidencia muy limitada a nivel regional (Barbero *et al.*, 2023). Contribuimos a esta creciente literatura desde un punto de vista empírico y metodológico. Por un lado, y en contraste con la mayoría de los estudios anteriores sobre el tema específico, nuestro análisis se centra a nivel regional dentro de un país específico, España, que es uno de los países europeos más afectados por la pandemia. Por otro lado, mientras que la mayoría de los estudios anteriores han considerado modelos de equilibrio general dinámico y similares, nosotros utilizamos modelos contrafactuales, herramienta muy valiosa en investigación, que buscan ofrecer una visión de lo que podría haber sucedido en ausencia del shock o perturbación.

En este sentido, los resultados deben abordarse con una rigurosa cautela, dada la meticulosidad inherente en la evaluación de la significación de estos modelos. Este rigor metodológico surge de la naturaleza de los datos en los que se basan, extraídos de registros históricos de las variables de interés y suposiciones sobre las relaciones causales subyacentes a un evento perturbador, lo cual también es una limitación de los métodos que empleamos. Este rigor puede llevar a resultados que no solo ilustran lo que "podría haber sucedido si...", sino también lo que "realmente ha sucedido", ya que a menudo reflejan la realidad observada en lugar de una simulada.

Múltiples factores contribuyen a la percepción de los modelos contrafactuales como conservadores. En primer lugar, su fuerte dependencia de datos históricos implica que la calibración y predicción de estos modelos están fuertemente influenciadas por eventos anteriores al evento perturbador. Además, la simplificación de las relaciones causales puede llevar a una representación inexacta de la complejidad de la realidad por parte del contrafactual. A su vez, la disponibilidad limitada de información a veces implica la no obtención de todos los datos necesarios para construir el contrafactual, lo que puede resultar en predicciones de variables individuales que solo se obtiene información hasta el momento del shock o perturbación. Además, los modelos contrafactuales tienden a evitar predicciones o escenarios que se desvíen demasiado de la realidad, manteniendo coherencia con los datos históricos. Finalmente, la validación empírica del modelo a menudo no es factible hasta que haya pasado un tiempo significativo, y en otros casos, no es posible observar lo que sucedió porque es una situación simulada.

Nuestro estudio ha puesto de manifiesto el inherente conservadurismo de este tipo de modelos. También ha destacado cómo una prueba que no arroja resultados estadísticamente significativos puede no estar reflejando la complejidad y riqueza de los eventos en cuestión. Este reconocimiento de la sutileza y limitaciones de los modelos contrafactuales contribuye a una interpretación más matizada y precisa de sus resultados.

Los análisis presentados proporcionan, por lo tanto, una evaluación relativamente precisa del impacto de los fondos. A pesar de las limitaciones como el tamaño de la muestra o la información limitada disponible hasta ahora, nuestros resultados respaldan la existencia de un impacto positivo de la intervención. La atención cuidadosa a los errores de Tipo II y el uso de pruebas alternativas fortalecen la validez de los resultados, que sugieren

que el impacto no es el resultado de la aleatoriedad estadística, tanto a nivel individual como agregado. Además, las pruebas de hipótesis restringidas indican que la magnitud del efecto tiende a aumentar con el tiempo. Así, nuestros análisis y hallazgos ofrecen no solo una alternativa a al conservadurismo de las pruebas actualmenye asociadas a modelos SCM para probar la existencia de un impacto, sino también una base sólida para afirmar que la intervención del NGEU tiene un impacto positivo y significativo en el PIB regional español, y que este impacto continuará en el futuro inmediato.

Capítulo 5

Conclusiones y líneas futuras de investigación

5.1. Conclusiones

La generación de predicciones económicas juega un papel fundamental en la toma de decisiones estratégicas de los agentes económicos y en la formulación de las políticas públicas. Tomar decisiones macroeconómicas puede ser contemplado como una actividad de riesgo dada la multitud de ramificaciones que conlleva y las situaciones tan dinámicas y cambiantes en las que deben realizarse. La situación es incluso más compleja en el escenario económico actual, en el que las fuentes de incertidumbre crecen, lo que constituye un desafío constante, que afecta a las decisiones de gobiernos nacionales y regionales.

La capacidad de anticipar y comprender el futuro económico se convierte, por tanto, en una herramienta crucial para navegar eficazmente por los matices de los mercados y las condiciones económicas. La incertidumbre, omnipresente en la economía global, representa un riesgo significativo para el avance económico y la promoción del desarrollo sostenible en el ámbito público. En este sentido, la generación de predicciones de crecimiento regional emerge como una herramienta valiosa para anticipar nuevos escenarios, adaptarse a un entorno económico cambiante y diseñar políticas que impulsen un crecimiento sostenible y resiliente.

La obtención de predicciones precisas genera innumerables beneficios, entre los que destacan la mejora de la toma de decisiones, la optimización de recursos, el avance en estabilidad y resiliencia, la mejor planificación a largo de los gobiernos y el incremento de la eficiencia en las políticas públicas. La reducción de la incertidumbre en la predicción económica permite a los tomadores de decisiones, ya sean gobiernos, empresas o individuos, contar con información más precisa y confiable. Esto facilita la formulación de estrategias más efectivas y adaptativas, al estar basadas en una comprensión más nítida de los posibles escenarios. La incertidumbre económica puede dar lugar a asignaciones ineficientes de recursos. Al reducir la incertidumbre en el futuro económico, se mejora la capacidad en la planificación y se realiza una asignación de recursos más precisa. Una predicción económica menos incierta contribuye a la estabilidad económica y fomenta la resiliencia frente a posibles perturbaciones. Los agentes económicos, como gobiernos y empresas, pueden anticipar y prepararse para eventos adversos, mitigando así su impacto y fortaleciendo la capacidad de recuperación frente a crisis o cambios inesperados. Tanto

los gobiernos como las empresas necesitan planificar a medio y largo plazo para lograr el crecimiento sostenible. La reducción de la incertidumbre en las predicciones económicas brinda la confianza necesaria para formular estrategias a largo plazo, lo que es esencial para el desarrollo sostenible y la implementación exitosa de políticas y proyectos. Los gobiernos pueden diseñar políticas públicas más efectivas cuando tienen una comprensión clara de cómo ciertas medidas afectarán a la economía. Por ello, la predicción económica precisa y menos incierta permite una evaluación más veraz del impacto de las políticas, facilitando una toma de decisiones gubernamentales más informadas y eficientes.

Por ello, la generación de predicciones económicas nacionales y regionales de PIB permite a las autoridades gubernamentales tomar decisiones informadas sobre la asignación de recursos, la implementación de reformas y la orientación de inversiones. Al conocer cómo ciertos fondos y subvenciones influirán en la economía regional, se pueden diseñar estrategias más efectivas y adaptativas. Esto se vuelve aún más relevante en un contexto donde los recursos son limitados y la eficiencia en su utilización es crucial. Bajo esta premisa, en los capítulos anteriores de esta investigación hemos desarrollado metodologías y aplicaciones predictivas valiosas para la generación de estimaciones del crecimiento regional, la gestión de la incertidumbre y la toma de decisiones informada, destacando enfoques innovadores y procedimientos robustos.

El primer estudio combina teoría de información y técnicas de muestreo para abordar el problema de la combinación de predicciones individuales. La combinación de predicciones constituye una estrategia relevante para mitigar los errores inherentes a las predicciones económicas individuales. Los métodos tradicionales de combinación de predicciones se centran en la evaluación retrospectiva del desempeño de las instituciones que se van a combinar. Esta aproximación se torna indeterminada cuando el número de instituciones es superior al número de predicciones pasadas de que se dispone. Con el objetivo de intentar dar solución a este problema, se propone un enfoque basado en la teoría de la información y en el uso del concepto de entropía. Se introduce una especificación innovadora de regresión con regularización y un algoritmo eficaz para resolver el problema. Este enfoque no solo asegura alcanzar una solución, sino que también brinda flexibilidad al permitir diversas normas para medir discrepancias y ponderar su importancia.

El algoritmo, que se desarrolla en el segundo capítulo, considera una norma cuadrática

y la evaluación de origen móvil como estrategia de validación cruzada. La metodología que se propone, se aplica al producto interior bruto (PIB), utilizándose estimaciones finales del PIB nacional en lugar de previsiones preliminares, consideradas, por algunos autores, más atractivas y relevantes para la toma de decisiones económicas. La contribución clave de este primer estudio se ha centrado en establecer una conexión entre la inferencia de entropía máxima y la regularización para resolver el problema de la combinación de predicciones. La conexión de la teoría de la información con la regularización del aprendizaje automático crea un marco capaz de superar las limitaciones que presentan otros algoritmos de combinación, y abre un camino para estudiar el impacto de otros enfoques también basados en aprendizaje automático, como *boosting* o *bagging*. Este enfoque proporciona una solución garantizada y una conexión más inmediata entre la literatura de aprendizaje automático y la teoría de la información.

En el segundo estudio, se hace mención a los actores económicos del sector público, en especial a la encrucijada a la que se enfrentan diariamente en la toma de decisiones y las repercusiones que pueden tener estas en el ámbito social y económico. Estas decisiones, moldeadas por las condiciones presentes y futuras de la región o país, demandan el respaldo de herramientas adecuadas para llevar a cabo análisis coyunturales, predicciones económicas y evaluaciones de la incertidumbre asociada. La disponibilidad de tales instrumentos se erige como una herramienta poderosa capaz de elevar la calidad de las decisiones tomadas.

En este contexto, se presenta una aplicación web interactiva (`https://apps.uv.es/app_indicador_sintetico/`) diseñada para generar de manera sencilla un indicador sintético mediante análisis de factorial dinámico. La metodología combina de manera activa indicadores económicos individuales y explora el método más adecuado para medir la incertidumbre asociada. Asimismo, se aborda la pregunta acerca del número mínimo de simulaciones *bootstrap* necesario para realizar una evaluación precisa de la incertidumbre de una serie sintética obtenida mediante el análisis dinámico de factores.

En esta investigación, se ha verificado que la metodología propuesta por Cuevas *et al.* (2015) para la obtención de un indicador sintético basado en factores latentes permite capturar de manera precisa la señal común de un conjunto de indicadores simples. Este método permite a los tomadores de decisiones públicos mejorar sintéticamente su

comprensión de la evolución económica agregada de una región, y particularmente de la Comunitat Valenciana. A través de la aplicación, los usuarios pueden obtener de forma autónoma una visión de la situación económica actual de la región (*nowcasting*) y anticipar su desarrollo inmediato futuro (*short-term forecasting*). Basándonos en estos resultados, recomendamos que los actores económicos de la Comunitat Valenciana confíen en esta herramienta innovadora para la toma de decisiones.

La construcción del índice sintético se complementa con la exploración del cálculo de su incertidumbre, donde, debido a la complejidad inherente a la metodología, basada en análisis factorial dinámico, recurrimos a técnicas de *bootstrap*. Mostramos que, en este contexto, 1000 simulaciones son suficientes para calcular de manera adecuada los intervalos de confianza del índice sintético en distintos momentos temporales.

Finalmente, y en conexión con esta investigación, consideramos relevante informar que, aunque la herramienta y el índice representan contribuciones valiosas por sí misma, constituyen elementos intermedios para la generación de otros componentes valiosos en los que actualmente estamos trabajando para que estén disponibles a través de una nueva aplicación. En esta nueva aplicación, el código de construcción del indicador sintético y su metodología asociada, se emplearán como un componente intermedio para estimar las series trimestrales regionales del PIB, como se discute en (Pavía-Miralles y Cabrer-Borrás, 2007).

Por último, el tercer estudio se ha centrado en medir el impacto en España de la respuesta de la Unión Europea (UE) para hacer frente a las secuelas económicas de la pandemia del COVID-19, las cuáles se han agravado debido a los recientes conflictos internacionales, como la guerra en Ucrania o la reactivación del conflicto Israel-Palestina. La influencia de la pandemia ha afectado de manera diferente a los diversos países, independientemente de su nivel de desarrollo, generando respuestas políticas y medidas de contención diversas (Giménez *et al.*, 2023). En el contexto específico de la UE, especialmente afectada por la crisis sanitaria, ha surgido el programa Next Generation EU, implementado a través del Mecanismo de Recuperación y Resiliencia (MRR).

Dada la novedad de esta iniciativa, la evidencia sobre la efectividad del mecanismo, especialmente en términos de impactos económicos, sigue siendo limitada. La mayoría de los estudios se centran en la UE (Bańkowski *et al.*, 2022, 2021; Pfeiffer *et al.*, 2023; Picek,

2020) o a nivel país (Fernández-Cerezo *et al.*, 2023; Malliaropulos *et al.*, 2021), con escasa evidencia a nivel regional (Barbero *et al.*, 2023). Nuestro análisis se ha centrado a nivel regional en España, uno de los países más afectados por la pandemia y beneficiario de las ayudas económicas. A diferencia de estudios anteriores, que utilizan modelos de equilibrio general dinámico, en nuestra investigación se ha optado por modelos contrafactuales, los cuáles son valiosos ya que ofrecen una visión de lo que podría haber sucedido sin una determinada perturbación.

Los resultados que hemos obtenido deben ser interpretados con cautela debido a las características metodológicas de la aproximación y las limitaciones que imponen la naturaleza de los datos disponibles. A pesar de que los modelos contrafactuales ilustran no solo lo que "podría haber sucedido si...", sino también lo que "realmente ha sucedido", reflejando la realidad observada en lugar de simulada, varios factores contribuyen a la percepción de los modelos contrafactuales como conservadores para el objetivo de detección de impactos. Su fuerte dependencia de datos históricos implica que la calibración y predicción están influenciadas por eventos anteriores al perturbador. Además, la simplificación de las relaciones causales puede conducir a una representación inexacta de la complejidad de la realidad por parte del contrafactual. La disponibilidad limitada de información, en muchas ocasiones proveniente de la falta de datos, implica resultados que solo incorporan información hasta el momento del shock. Además, tienden a evitar predicciones que se desvíen demasiado de la realidad, manteniendo coherencia con los datos históricos. La validación empírica a menudo no es factible hasta pasado un tiempo significativo.

Nuestro estudio resalta el inherente conservadurismo de estos modelos y cómo una prueba sin resultados aparentemente estadísticamente significativos puede no reflejar la complejidad de los eventos. Este reconocimiento contribuye a una interpretación más matizada y precisa de los resultados. Sin embargo, los análisis proporcionan una evaluación precisa del impacto de los fondos. A pesar de las limitaciones como el tamaño de la muestra o la información limitada, nuestros resultados respaldan un impacto positivo de la intervención. La atención a los errores de Tipo II y el uso de pruebas alternativas fortalecen la validez, sugiriendo que el impacto no es aleatorio, ni a nivel individual ni a nivel agregado. Las pruebas de hipótesis restringidas han permitido concluir la existencia de efecto y, además que la magnitud del efecto tenderá a aumentar con el tiempo.

En conjunto, estos tres estudios reflejan la diversidad de enfoques y metodologías existentes para abordar la generación de predicciones, la gestión de la incertidumbre y la mejora de la toma de decisiones. Las consideramos valiosas contribuciones que enriquecen la comprensión de fenómenos económicos complejos y proporcionan perspectivas clave para la investigación futura en campos tan dinámicos como la macroeconomía.

5.2. Líneas futuras de investigación

La investigación desarrollada en la presente tesis ha marcado importantes líneas de investigación en el ámbito de la economía regional, especialmente en el análisis del crecimiento económico, la gestión de la incertidumbre y la toma de decisiones fundamentada. No obstante, nuestro trabajo no acaba aquí, de hecho, recién acaba de empezar. A continuación, se exponen algunas ideas que estamos desarrollando en estos momentos, así como algunos proyectos que tenemos intención de llevar a cabo a corto plazo.

En el primer estudio, los futuros trabajos de investigación se enfocarán, desde una perspectiva metodológica, en mejorar la adaptabilidad y robustez del algoritmo. Se propone avanzar en el desarrollo de una versión capaz de gestionar paneles de datos desbalanceados, lo que ampliaría su aplicabilidad a escenarios de datos más diversos y desafiantes, incluidos escenarios con datos missing. Además, se explorarán nuevas funciones de penalización, como *lasso* o *elastic net*, con el objetivo de optimizar la capacidad de generalización del modelo.

Otra dirección alentadora implica la investigación y desarrollo del impacto de diferentes métodos de validación cruzada. Este enfoque busca perfeccionar la evaluación del rendimiento, asegurando una validación más rigurosa y evitando posibles sesgos en la obtención e interpretación de los resultados.

Desde una perspectiva numérica, una futura línea de trabajo implicaría implementar estimaciones flash en lugar de estimaciones permanentes. Esta estrategia promete proporcionar resultados preliminares de manera más rápida, permitiendo una adaptación más ágil a entornos dinámicos y decisiones informadas.

Además, valoramos explorar otras propuestas dentro del enfoque de aprendizaje automático, aunque abandonando el marco conceptual entrópico. Esto incluiría métodos

de combinación de predicciones como *bagging*, *boosting* o *stacking*, entre otros, con el objetivo de aprovechar la diversidad de modelos para mejorar la precisión y estabilidad de las predicciones. Estas líneas de investigación sugieren un camino interesante hacia la evolución y perfeccionamiento continuo del algoritmo en cuestión.

En el segundo estudio, las perspectivas de investigación futuras en el ámbito de la generación de indicadores sintéticos y la medición de la incertidumbre plantean una serie de desafíos que buscan enriquecer y perfeccionar el proceso. En este sentido, una posible línea futura consistiría en realizar un análisis sistemático y exhaustivo del impacto que tendrá anualizar los indicadores mensuales y trimestrales dentro del proceso de selección de variables para la construcción del indicador sintético. La diferente consideración de esta dimensión temporal no solo podría influir en la precisión del indicador, sino que también podría revelar patrones y tendencias más sólidas en la información recopilada.

Otra dirección de investigación relevante implica examinar el impacto de utilizar métodos alternativos de predicción para los meses faltantes en las series individuales, desafiando la actual dependencia de la función `auto.arima` de librería `forecast` (Hyndman *et al.*, 2023). La propuesta consiste en diversificar los métodos de predicción, evaluando su eficacia y confiabilidad en comparación con el enfoque actual. Esta diversificación podría proporcionar nuevas perspectivas sobre la predicción precisa de datos faltantes y mejorar la robustez del indicador sintético.

Un aspecto fundamental a explorar es la viabilidad y eficacia del método de *backcasting* para la construcción del indicador sintético. En esta línea, la idea sería realizar un estudio comparativo a lo largo de un extenso periodo que permita evaluar la idoneidad de esta técnica, especialmente en situaciones en las que todas las series económicas poseen datos faltantes. Este análisis debe extenderse a un ejercicio de simulación para identificar posibles escenarios en los cuales el método de *backcasting* podría destacar, proporcionando una comprensión más completa de sus ventajas y limitaciones.

Otra línea potencial en la que estamos comenzando a trabajar consiste en valorar comparativamente y de manera exhaustiva como funcionarían otros métodos de cálculo de indicadores sintéticos, como el método de contabilidad nacional input-output, el modelado econométrico o el método del valor bruto de producción. Este enfoque busca determinar las bondades relativas de cada aproximación, considerando sus fortalezas

y limitaciones en diversos contextos económicos. La comparación detallada de estos métodos proporcionaría una visión más completa de las opciones disponibles y permitiría elegir la metodología más apropiada según las características específicas del análisis.

Finalmente, a partir del último estudio, centrados en la investigación futura sobre modelos contrafactuales, específicamente en el análisis del impacto de los fondos NGEU desde su implementación, consideramos que se abre un horizonte prometedor tanto desde la perspectiva de la aplicabilidad como desde el punto de vista metodológico.

En términos de aplicación, un estudio interesante implicaría realizar un análisis detallado sobre cómo podría afectar el empleo de países donantes externos a la UE. La inclusión de estos países en el análisis podría mejorar el proceso del *match* en el pretratamiento, especialmente al considerar que no reciben fondos NGEU y disponer de más variedad en términos de estructuras productivas. Este enfoque podría proporcionar una visión más completa y robusta de los efectos al tener en cuenta una gama más amplia de contextos y condiciones.

Otra línea, que también consideramos prometedora, consistiría en investigar cómo la inclusión de un mayor número de variables en la determinación de los contrafactuales podría afectar a los resultados. La ampliación de las variables podría influir significativamente en la calidad del *match* en el pretratamiento, mejorando así la precisión y validez de los resultados. Este enfoque podría revelar patrones más sutiles y complejos, permitiendo una comprensión más completa del impacto de los fondos NGEU.

Una dirección futura relevante en cuanto a la replicación del estudio sería repetirlo unos años más tarde. La disponibilidad de datos a lo largo de varios años permitiría una mejor evaluación a largo plazo de los efectos de los fondos NGEU, proporcionando una perspectiva temporal más sólida y una comprensión más profunda de su impacto a lo largo del tiempo.

Desde el punto de vista metodológico, se plantea la necesidad de estudiar la sensibilidad del método para la detección de shocks en función de todas las variables consideradas. Este análisis aportaría claridad sobre la robustez del método utilizado, ayudando a comprender mejor cómo diferentes variables pueden afectar a la capacidad del modelo para identificar y cuantificar el impacto de los fondos NGEU.

En conjunto, estas líneas de investigación ofrecen una perspectiva integral para fortale-

cer la validez y aplicabilidad de los modelos contrafactuales en el análisis del impacto de los fondos NGEU, abriendo oportunidades para mejorar la comprensión de las dinámicas económicas y políticas asociadas con esta iniciativa.

Asimismo, dentro del marco de estudio de esta tesis, se identifican diversas líneas de investigación pertinentes que podrían enriquecer y ampliar de manera significativa el conocimiento en áreas clave. Estas perspectivas futuras abordan aspectos fundamentales en el ámbito de la investigación económica y regional, entre los cuales destacamos la contabilidad trimestral regional, la contabilidad sectorial, el Big Data y los modelos contrafactuales.

La implementación de la contabilidad trimestral regional se presenta como una dirección esencial para profundizar en el análisis económico a nivel local. Este enfoque posibilitaría la captura más precisa de variaciones estacionales y tendencias a corto plazo que influyen en la dinámica económica de regiones específicas. Explorar la contabilidad trimestral regional ofrecería una perspectiva dinámica y detallada de los cambios en el crecimiento económico regional. Esto no solo permitiría una mejor comprensión de las tendencias a corto plazo, sino que también facilitaría una capacidad de reacción más rápida y precisa ante cambios en las condiciones económicas del entorno. La contabilidad trimestral a nivel regional se configura, por tanto, como una herramienta particularmente útil para la toma de decisiones informada y estratégica.

En consonancia con este enfoque, surge otro enfoque complementario como es la descomposición de la contabilidad regional desde el punto de vista sectorial, derivada de la necesidad de avanzar en la predicción económica regional tanto en términos anuales como trimestrales. La contabilidad sectorial regional se centra en el análisis económico desde una perspectiva sectorial, descomponiendo la actividad económica según su aportación al crecimiento económico de la región. Este conocimiento facilitaría una comprensión más profunda de la estructura económica regional y permitiría a los gobiernos diseñar estrategias específicas de desarrollo económico, identificando las áreas de fortaleza y debilidad de cada sector.

La importancia de las predicciones recae en su capacidad para proporcionar información valiosa sobre el comportamiento económico futuro, también a nivel regional. Estas dos herramientas propuestas (contabilidad trimestral regional y contabilidad sectorial re-

gional) son esenciales para la planificación estratégica de políticas y la asignación eficiente de recursos.

Finalmente, en los últimos años hemos asistido a una integración creciente y al incremento de herramientas de análisis de conjuntos de datos masivos, conocido como Big Data, por lo que consideramos avanzar en su incorporación como una herramienta poderosa más para el estudio económico y la generación de predicciones regionales. Explorar esta línea permitiría aprovechar la abundancia de información disponible, desde transacciones comerciales hasta datos sociales, para obtener perspectivas más ricas y detalladas. El uso adecuado de Big Data podría revelar patrones ocultos, mejorar la capacidad predictiva y proporcionar *insights* valiosos para la toma de decisiones económicas a nivel regional.

Por último, la evaluación de modelos contrafactuales sigue siendo una piedra angular en la investigación económica. Estos modelos desempeñan un papel crucial en la evaluación de impactos derivados de intervenciones económicas y políticas. Aunque esta investigación ha abordado ciertos matices en la aplicación de modelos contrafactuales para España y sus regiones, una dirección futura interesante sería replicar este tipo de modelos y la inferencia causal propuesta en otro tipo de contexto. La aplicación de esta metodología en un entorno estable y sin grandes shocks, permitiría simular un shock exógeno aplicando el modelo contrafactual y evaluar la calidad inferencial propuesta por la aproximación.

Bibliografía

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425.

Abadie, A., Diamond, A., y Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A. y Gardeazábal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1):113–132.

Akobeng, A. K. (2016). Understanding type I and type II errors, statistical power and sample size. *Acta Paediatrica*, 105(6):605–609.

Alloza, M. y Sanz, C. (2021). Jobs multipliers: Evidence from a large fiscal stimulus in Spain. *The Scandinavian Journal of Economics*, 123(3):751–779.

Amendola, A., Di Serio, M., Fragetta, M., y Melina, G. (2020). The euro-area government spending multiplier at the effective lower bound. *European Economic Review*, 127:103480.

Andreini, P., Hasenzagl, T., Reichlin, L., Senftleben-König, C., y Strohsal, T. (2023). Nowcasting german gdp: Foreign factors, financial markets, and model averaging. *International Journal of Forecasting*, 39(1):298–313.

Anesti, N., Galvão, A. B., y Miranda-Agrippino, S. (2022). Uncertain Kingdom: Nowcasting Gross Domestic Product and its revisions. *Journal of Applied Econometrics*, 37(1):42–62.

Antipa, P., Barhoumi, K., Brunhes-Lesage, V., y Darné, O. (2012). Nowcasting German GDP: A comparison of bridge and factor models. *Journal of Policy Modeling*, 34(6):864–878.

Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners*, volumen 30. Springer.

Arribas, I., Peiró-Palomino, J., y Tortosa-Ausina, E. (2020). Is full banking integration desirable? *Journal of Banking & Finance*, 112:105185.

Ballester, L., López, J., y Pavía, J. M. (2023). European systemic credit risk transmission using dynamic bayesian networks. *Research in International Business and Finance*, 65:101914.

Bańkowski, K., Bouabdallah, O., Semeano, J. D., Dorrucci, E., Freier, M., Jacquinot, P., Modery, W., Rodríguez Vives, M., Valenta, V., y Zorell, N. (2022). The economic impact of Next Generation EU: a euro area perspective. Occasional Paper 291, ECB.

Bańkowski, K., Ferdinandusse, M., Hauptmeier, S., Jacquinot, P., y Valenta, V. (2021). The macroeconomic impact of the Next Generation EU instrument on the euro area. Occasional Paper 255, ECB.

Barbero, J., Conte, A., Crucitti, F., Lazarou, N.-J., Sakkas, S., y Salotti, S. (2023). The impact of the recovery fund on EU regions: a spatial general equilibrium analysis. *Regional Studies*, forthcoming.

Bargain, O., Dolls, M., Fuest, C., Neumann, D., Peichl, A., Pestel, N., y Siegloch, S. (2013). Fiscal union in Europe? Redistributive and stabilizing effects of a European tax-benefit system and fiscal equalization mechanism. *Economic Policy*, 28(75):375–422.

Barro, R. J. (2001). Human capital and growth. *American Economic Review*, 91(2):12–17.

Bartik, T. J. (2020). Using place-based jobs policies to help distressed communities. *Journal of Economic Perspectives*, 34(3):99–127.

Bates, J. M. y Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.

Becker, S. O., Egger, P. H., y Von Ehrlich, M. (2010). Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics*, 94(9-10):578–590.

Becker, S. O., Egger, P. H., y Von Ehrlich, M. (2012). Too much of a good thing? On the growth effects of the EU's regional policy. *European Economic Review*, 56(4):648–668.

Becker, S. O., Egger, P. H., y Von Ehrlich, M. (2013). Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy*, 5(4):29–77.

Becker, S. O., Egger, P. H., y Von Ehrlich, M. (2018). Effects of EU regional policy: 1989-2013. *Regional Science and Urban Economics*, 69:143–152.

Beetsma, R. y Giuliodori, M. (2011). The effects of government purchases shocks: review and estimates for the EU. *The Economic Journal*, 121(550):F4–F32.

Beetsma, R., Giuliodori, M., y Klaassen, F. (2006). Trade spill-overs of fiscal policy in the European Union: a panel analysis. *Economic policy*, 21(48):640–687.

Bergmeir, C. y Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191:192–213.

Beugelsdijk, M. y Eijffinger, S. C. (2005). The effectiveness of structural policy in the European Union: An empirical analysis for the EU-15 in 1995–2001. *JCMS: Journal of Common Market Studies*, 43(1):37–51.

Bitetto, A., Cerchiello, P., y Mertzanis, C. (2023). On the efficient synthesis of short financial time series: A Dynamic Factor Model approach. *Finance Research Letters*, 53:103678.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., y Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

Brachert, M., Dettmann, E., y Titze, M. (2019). The regional effects of a place-based policy–Causal evidence from Germany. *Regional Science and Urban Economics*, 79:103483.

Breiman, L. (1996). Bagging predictors. *Machine learning*, 24:123–140.

Brueckner, M., Pappa, E., y Valentinyi, A. (2023). Geographic Cross-Sectional Fiscal Spending Multipliers and the Role of Local Autonomy: Evidence from European Regions. *Journal of Money, Credit and Banking*, forthcoming.

Burns, A. F. y Mitchell, W. C. (1946). *Measuring Business Cycles*. National Bureau of Economic Research.

Cabrer-Borrás, B., Serrano, G., y Pavía, J. M. (2017). Evaluación del sesgo en las estimaciones de contabilidad nacional trimestral: Estudio de las añadas en españa. *Studies of Applied Economics*, 35(2):271–298.

Cabrer Borrás, B. (2001). Indicadores económicos y su problemática: Una visión de síntesis. En Cabrer Borrás, B., editor, *Análisis Regional: El Proyecto Hispalink*, pp. 259–275. Mundi Prensa Libros, Madrid.

Camacho, M. y Pérez-Quirós, G. (2010). Introducing the euro-sting: Short-term indicator of euro area growth. *Journal of Applied Econometrics*, 25(4):663–694.

Carozzi, F. y Repetto, L. (2019). Distributive politics inside the city? The political economy of Spain's Plan E. *Regional Science and Urban Economics*, 75:85–106.

Carrillo, P. E. y Rothbaum, J. L. (2016). Counterfactual spatial distributions. *Journal of Regional Science*, 56(5):868–894.

Castillo, V., Garone, L. F., Maffioli, A., y Salazar, L. (2017). The causal effects of regional industrial policies on employment: A synthetic control approach. *Regional Science and Urban Economics*, 67:25–41.

Cattaneo, M. D., Feng, Y., y Titiunik, R. (2021). Prediction intervals for synthetic control methods. *Journal of the American Statistical Association*, 116(536):1865–1880.

Chan, F. y Pauwels, L. L. (2018). Some theoretical results on forecast combinations. *International Journal of Forecasting*, 34(1):64–74.

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., y Borges, B. (2023). *Shiny: Web Application Framework for R*. R package version 1.7.4.9002.

Chari, V. V., Kehoe, P. J., y McGrattan, E. R. (2009). New keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, 1(1):242–266.

Chernick, M. R. (2011). *Bootstrap Methods: A Guide for Practitioners and Researchers*. John Wiley & Sons.

Chernis, T., Cheung, C., y Velasco, G. (2020). A three-frequency dynamic factor model for nowcasting Canadian provincial GDP growth. *International Journal of Forecasting*, 36(3):851–872.

Chernozhukov, V., Fernández-Val, I., y Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268.

Chernozhukov, V., Wüthrich, K., y Zhu, Y. (2021). An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls. *Journal of the American Statistical Association*, 116(536):1849–1864.

Chow, H. K., Fei, Y., y Han, D. (2023). Forecasting GDP with many predictors in a small open economy: forecast or information pooling? *Empirical Economics*.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International journal of forecasting*, 5(4):559–583.

Coelho, M. (2019). Fiscal stimulus in a monetary union: Evidence from eurozone regions. *IMF Economic Review*, 67:573–617.

Corbi, R., Papaioannou, E., y Surico, P. (2019). Regional transfer multipliers. *The Review of Economic Studies*, 86(5):1901–1934.

Cuevas, A., Pérez-Quirós, G., y Quilis, E. M. (2017). Integrated model of short-term forecasting of the spanish economy (mipred model). *Revista de Economía Aplicada*, 25(74):5–25.

Cuevas, A. y Quilis, E. M. (2012). A factor analysis for the Spanish economy. *SERIEs*, 3(3):311–338.

Cuevas, A., Quilis, E. M., y Espasa, A. (2015). Quarterly Regional GDP Flash Estimates by Means of Benchmarking and Chain Linking. *Journal of Official Statistics*, 31(4):627–647.

Dagum, E. B. y Cholette, P. A. (2006). The cholette-dagum regression-based benchmarking method — the additive model. En *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*, Lecture Notes in Statistics, pp. 51–84. Springer, New York.

Dauphin, J. F., Dybczak, K., Maneely, M., Sanjani, M. T., Suphaphiphat, N., Wang, Y., y Zhang, H. (2022). *Nowcasting GDP - A Scalable Approach Using DFM, Machine Learning and Novel Data, Applied to European Economies*. International Monetary Fund.

Davidson, R. y MacKinnon, J. G. (2000). Bootstrap tests: How many bootstraps? *Econometric Reviews*, 19(1):55–68.

Di Caro, P. y Fratesi, U. (2022). One policy, different effects: Estimating the region-specific impacts of EU cohesion policy. *Journal of Regional Science*, 62(1):307–330.

Domínguez Serrano, M., Blancas Peral, F. J., Guerrero Casas, F. M., y González Lozano, M. (2011). Una revisión crítica para la construcción de indicadores sintéticos. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 11:41–70.

Doz, C. y Fuleky, P. (2020). Dynamic factor models. *Macroeconomic Forecasting in the Era of Big Data: Theory and Practice*, pp. 27–64.

Duro, J. A., Perez-Laborda, A., Turrion-Prats, J., y Fernández-Fernández, M. (2021). Covid-19 and tourism vulnerability. *Tourism Management Perspectives*, 38:100819.

Ederveen, S., De Groot, H. L., y Nahuis, R. (2006). Fertile soil for structural funds? A panel data analysis of the conditional effectiveness of European cohesion policy. *Kyklos*, 59(1):17–42.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26.

Efron, B. y Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence lntervals, and other measures of statistical accuracy. *Statistical Science*, 1(1):54–77.

Efron, B. y Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC Press.

Ehrlich, M. v. y Overman, H. G. (2020). Place-based policies and spatial disparities across European cities. *Journal of Economic Perspectives*, 34(3):128–149.

Eli Ben-Michael, A. F. y Rothstein, J. (2021). The Augmented Synthetic Control Method. *Journal of the American Statistical Association*, 116(536):1789–1803.

Ferman, B. y Pinto, C. (2019). Inference in differences-in-differences with few treated groups and heteroskedasticity. *Review of Economics and Statistics*, 101(3):452–467.

Fernández-Cerezo, A., Moral-Benito, E., y Quintana, J. (2023). A production network model for the Spanish economy with an application to the impact of NGEU funds. Documentos de Trabajo 5, Banco de España.

Fernández-Vázquez, E., Moreno, B., y Hewings, G. J. (2019). A data-weighted prior estimator for forecast combination. *Entropy*, 21(4):429.

Fernández-Villaverde, J. (2010). The econometrics of dsge models. *SERIEs*, 1(1-2):3–49.

Firpo, S. y Possebom, V. (2018). Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets. *Journal of Causal Inference*, 6(2):20160026.

Fleenor, J. W. (2006). The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economics, societies and nations. *Personnel Psychology*, 59(4):982.

Fratesi, U. y Wishlade, F. G. (2017). The impact of European Cohesion Policy in different contexts.

Fresoli, D. (2022). Bootstrap VAR forecasts: The effect of model uncertainties. *Journal of Forecasting*, 41(2):279–293.

Freund, Y. (1995). Boosting a weak learning algorithm by majority. *Information and computation*, 121(2):256–285.

Galiani, S. y Quistorff, B. (2017). The synth_runner package: Utilities to automate synthetic control estimation using synth. *The Stata Journal*, 17(4):834–849.

Genre, V., Kenny, G., Meyler, A., y Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, 29(1):108–121.

Gil Martín, M., Leiva León, D., Pérez García, J. J., y Urtasun Amann, A. (2019). An application of dynamic factor models to nowcast regional economic activity in spain. *Documentos Ocasionales/Banco de España, 1904*, (1904).

Giménez, V., Prior, D., Thieme, C., y Tortosa-Ausina, E. (2023). International comparisons of covid-19 pandemic management: What can be learned from activity analysis techniques? *Omega*, p. 102966.

González-Rodríguez, B. (2019). *nlopt: Call Optimization Solvers with .nl Files*. R package version 0.1.1.

Granger, C. W. y Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3(2):197–204.

Grudkowska, S. D. (2017). *JDemetra+ Reference Manual version 2.1*. Narodowy Bank Polski Education.

Hakura, D. (2020). What is debt sustainability? *International Monetary Fund*.

Hall, S. G., Tavlas, G. S., y Wang, Y. (2023). Forecasting inflation: The use of dynamic factor analysis and nonlinear combinations. *Journal of Forecasting*, 42(3):514–529.

Hasni, M., Aguir, M., Babai, M., y Jemai, Z. (2019). Spare parts demand forecasting: A review on bootstrapping methods. *International Journal of Production Research*, 57(15-16):4791–4804.

Hastie, T., Tibshirani, R., Friedman, J. H., y Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volumen 2. Springer.

Hebous, S. (2011). The effects of discretionary fiscal policy on macroeconomic aggregates: A reappraisal. *Journal of Economic Surveys*, 25(4):674–707.

Hebous, S. y Zimmermann, T. (2013). Estimating the effects of coordinated fiscal actions in the euro area. *European Economic Review*, 58:110–121.

Hesterberg, T. (2011). Bootstrap. *WIREs Computational Statistics*, 3(6):497–526.

Hothorn, T., Hornik, K., y Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, 15(3):651–674.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., y Yasmeen, F. (2023). *forecast: Forecasting functions for time series and linear models*. R package version 8.20.

IVE (2016). *Estándar del SEEDS de la Generalitat Valenciana para la Corrección de Efectos Estacionales y de Calendario en las Series Coyunturales*. Ed. Generalitat Valenciana.

Jaynes, E. T. (1957a). Information theory and statistical mechanics. *Physical review*, 106(4):620.

Jaynes, E. T. (1957b). Information theory and statistical mechanics. ii. *Physical review*, 108(2):171.

Khan, S. A. R., Razzaq, A., Yu, Z., Shah, A., Sharif, A., y Janjua, L. (2022). Disruption in food supply chain and undernourishment challenges: An empirical study in the context of Asian countries. *Socio-Economic Planning Sciences*, 82:101033.

Kuck, K. y Schweikert, K. (2021). Forecasting Baden-Württemberg's GDP growth: MIDAS regressions versus dynamic mixed-frequency factor models. *Journal of Forecasting*, 40(5):861–882.

Kuznets, S. (1955). Economic Growth and Income Inequality. *The American Economic Review*, 45(1):1–28.

Liang, R., Wang, F., y Xu, J. (2021). Nowcasting china's ppi inflation using low-frequency and mixed-frequency dynamic factor models. *J. Financ. Res*, 494:22–41.

Malliaropulos, D., Papageorgiou, D., Vasardani, M., y Vourvachaki, E. (2021). The impact of the recovery and resilience facility on the Greek economy. *Bank of Greece Economic Bulletin*, (53):7–28.

Mankiw, N. G., Romer, D., y Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 107(2):407–437.

Maravall, A., Gómez, V., y Caporello, G. (2015). Statistical and econometrics software: Tramo and seats. *Statistical and Econometrics Software*. Banco de España.

Meyer, J. S., Ingersoll, C. G., McDonald, L. L., y Boyce, M. S. (1986). Estimating Uncertainty in Population Growth Rates: Jackknife vs. Bootstrap Techniques. *Ecology*, 67(5):1156–1166.

Mohl, P. y Hagen, T. (2010). Do EU structural funds promote regional growth? New evidence from various panel data approaches. *Regional Science and Urban Economics*, 40(5):353–365.

Mondéjar-Jiménez, J. y Vargas-Vargas, M. (2008). Indicadores sintéticos: Una revisión de los métodos de agregación. *Economía, Sociedad y Territorio*, 8(27):565–585.

Montolio, D. (2018). The effects of local infrastructure investment on crime. *Labour Economics*, 52:210–230.

Mora-Sanguinetti, J. S. y Spruk, R. (2023). Economic effects of recent experiences of federalism: Analysis of the regionalization process in spain. *Journal of Regional Science*, 63(1):30–63.

Moreno, B. y López, A. J. (2013). Combining economic forecasts by using a maximum entropy econometric approach. *Journal of forecasting*, 32(2):124–136.

Mumtaz, H. y Musso, A. (2021). The Evolving Impact of Global, Region-Specific, and Country-Specific Uncertainty. *Journal of Business & Economic Statistics*, 39(2):466–481.

Newbold, P. y Granger, C. W. (1974). Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society: Series A (General)*, 137(2):131–146.

Pavía-Miralles, J. M. y Cabrer-Borrás, B. (2007). On estimating contemporaneous quarterly regional GDP. *Journal of Forecasting*, 26(3):155–170.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., y Babai (2022). Forecasting: Theory and practice. *International Journal of Forecasting*, 38(3):705–871.

Pfeiffer, P., Varga, J., y in 't Veld, J. (2023). Quantifying spillovers of coordinated investment stimulus in the EU. *Macroeconomic Dynamics*, 27(7):1843–1865.

Picek, O. (2020). Spillover effects from next generation EU. *Intereconomics*, 55(5):325–331.

Picek, O. y Schröder, E. (2017). Euro area imbalances: How much could an expansion in the North help the South? Working Paper 180, IMK, Düsseldorf.

Picek, O. y Schröder, E. (2018). Spillover effects of Germany's final demand on Southern Europe. *The World Economy*, 41(8):2216–2242.

Pinotti, P. (2015). The economic costs of organised crime: Evidence from southern italy. *The Economic Journal*, 125(586):F203–F232.

Pons-Novell, J. (2003). Strategic bias, herding behaviour and economic forecasts. *Journal of Forecasting*, 22(1):67–77.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Sala-i-Martin, X. (1996). Regional cohesion: evidence and theories of regional growth and convergence. *European Economic Review*, 40(6):1325–1352.

Schapire, R. E. (1990). The strength of weak learnability. *Machine learning*, 5:197–227.

Shnarch, E., Alzate, C., Dankin, L., Gleize, M., Hou, Y., Choshen, L., Aharonov, R., y Slonim, N. (2018). Will it blend? blending weak and strong labeled data in a neural network for argumentation mining. pp. 599–605.

Sifat, I., Zarei, A., Hosseini, S., y Bouri, E. (2022). Interbank liquidity risk transmission to large emerging markets in crisis periods. *International Review of Financial Analysis*, 82:102200.

Silvapulle, M. J. y Sen, P. K. (2005). *Constrained statistical inference: Inequality, order and shape restrictions*. John Wiley & Sons.

Smith, J. y Wallis, K. F. (2009). A simple explanation of the forecast combination puzzle. *Oxford bulletin of economics and statistics*, 71(3):331–355.

Stigler, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900*. Harvard University Press.

Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy*, 34(1-2):70–106.

Szczygielski, J. J., Brzeszczyński, J., Charteris, A., y Bwanya, P. R. (2022). The COVID-19 storm and the energy sector: The impact and role of uncertainty. *Energy Economics*, 109:105258.

Tao, Y. (2022). Research on the impact of trade uncertainty on national grain supply and risk cost control. *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science*, 72(1):92–104.

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. *International journal of forecasting*, 16(4):437–450.

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 58(1):267–288.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., y De Vries, G. J. (2015). An illustrated user guide to the world input-output database: the case of global automotive production. *Review of International Economics*, 23(3):575–605.

Timmermann, A. (2013). *Handbook of economic forecasting*. Elsevier/North Holland.

Tortosa-Ausina, E., Pérez, F., Mas, M., y Goerlich, F. J. (2005). Growth and convergence profiles in the spanish provinces (1965–1997). *Journal of Regional Science*, 45(1):147–182.

US Census Bureau (2017). *X-13ARIMA-SEATS Reference Manual*. US Census Bureau, Washington DC.

Valliant, R., Dorfman, A., y Royall, R. (2000). *Finite Population Sampling and Inference: A Prediction Approach*. Número 4.

Vanbrabant, L. y Kuiper, R. (2023). *restriktor: Restricted Statistical Estimation and Inference for Linear Models*. R package version 0.5-30.

Varma, S. y Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC bioinformatics*, 7(1):1–8.

Veres-Ferrer, E. J. y Pavía, J. M. (2021). Elasticity as a measure for online determination of remission points in ongoing epidemics. *Statistics in Medicine*, 40(4):865–884.

Wan Ahmad, W. K. A. y Ahmad, S. (2013). Arima model and exponential smoothing method: A comparison. *AIP Conference Proceedings*, 1522(1):1312–1321.

Wilcox, R. R. (2010). The Bootstrap. En Wilcox, R. R., editor, *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*, pp. 87–108. Springer, New York.

Zarnowitz, V. (1967). *An Appraisal of Short-Term Economic Forecast*. National Bureau of Economic Research.

Zou, H. y Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.