

## **Memoria Proyecto EDA**

### **Situación económica del Estado español: Una perspectiva de deuda**

**ADRIAN GARCIA HERNANDEZ**

Para poder analizar la situación económica del Estado español desde una perspectiva de deuda se deben considerar diferentes variables. Empezando por la evolución del PIB en términos absolutos, así como su crecimiento real. Estos datos ya indican la tendencia que toman los ingresos del estado, ingresos que como veremos, provienen principalmente de tres impuestos, el IRPF, las cotizaciones a la Seguridad Social y el IVA.

El IRPF, así como las cotizaciones a la Seguridad Social están completamente ligados al desempeño del mercado laboral, por lo que este también es tomado en cuenta.

Una vez detallados los ingresos del Estado, se analizará la evolución del gasto público y su tamaño respecto al PIB. Durante este análisis es primordial tener en cuenta la deuda, ya que en caso contrario las cuentas públicas no estarían equilibradas, los ingresos serían mucho menores que los gastos. La emisión de deuda por parte del Estado (préstamos a corto, medio y largo plazo), viene a solucionar este desajuste. Por una parte, se aumentan los ingresos, y por otra se incluye la deuda; de ahí la necesidad de tener datos sobre la deuda acumulada y el déficit (o superávit) del Estado de los últimos años.

Para poder calcular el ratio de deuda/PIB es primordial tener en cuenta la inflación.

Esta no suele afectar negativamente mucho los ingresos del Estado ya que la mayoría de los impuestos están basados en porcentajes, por lo que conforme la inflación aumenta, los ingresos del estado aumentan (prácticamente) de manera proporcional. No obstante, la inflación afecta mucho al valor real de la deuda acumulada, haciéndole perder valor, reduciendo su peso mientras el PIB en términos absolutos aumenta.

Cabe recalcar que este trabajo lleva a cabo un análisis descriptivo y comparativo. Esto supone que las expectativas extraídas tras analizar la tendencia de los indicadores macroeconómicos no necesariamente se deben de cumplir, no se han calculado probabilidades ni construido un modelo de previsión con estimaciones fiables. Así mismo, es necesario analizar las correlaciones entre todas las variables para determinar con precisión la verdadera relación entre ellas, como por ejemplo la correlación entre el PIB y la recaudación por IVA.

Los datos usados para el análisis provienen en su mayoría del portal estadístico EUROSTAT. Esta plataforma ha sido de una gran ayuda a la hora de obtener los datos, ya que posibilita la selección de los datos de nuestro interés dentro de un gran dataset, incluyendo el periodo, el tipo de unidad de medida de los valores, así como diferentes tipos de formato de presentación de los datos, entre otros. Además, hace una previsualización de los datos seleccionados y los muestra en diferentes gráficos y mapas. La descarga de los datos puede ser hecha en diferentes formatos como, por ejemplo, xlsx, SDMX-CSV o TSV.

No obstante, al descargar el fichero en cualquiera de los formatos disponibles, el archivo incluye información complementaria. Esta información es importante para la correcta comprensión de los datos, pero dificulta la importación de estos a un dataset en Python. Por ello el procedimiento usado ha sido descargar el archivo en xlsx, abrirlo con Excel y crear una nueva hoja de cálculo donde pegar únicamente los datos de la tabla (esto ya supone una primera parte de limpieza de datos). Posteriormente exportar esta hoja de cálculo a un archivo CSV para poderlo importar con pandas en Python usando el Visual Studio Code.

Parte de los datos han sido obtenidos a través de la página web del Fondo Monetario Internacional, en esta el procedimiento ha sido muy similar al proceso anteriormente detallado sobre EUROSTAT. En este caso la descarga de los datos solo se podía hacer en xlsx, por lo que se ha abierto en Excel y posteriormente han sido exportados a CSV.

Por último, el resto de los datos han sido obtenidos a través de informes del Ministerio de Hacienda, todos ellos almacenados en documentos xlsx. El procedimiento ha sido idéntico al anteriormente descrito.

En algunos pocos casos también se han modificado internamente los CSV siendo abiertos con un editor de texto, pero sólo para modificaciones muy puntuales.

Una vez los datos han sido exportados usando la librería de pandas en Python dentro de un Jupiter Notebook usando Visual Studio Code, empieza la segunda parte de limpieza de los datos. En esta fase principalmente se han hecho estos cambios:

- Cambiar comas por puntos
- Eliminar puntos que indican los miles
- Sustitución de símbolos que implican falta de datos por NaN
- Eliminación de ciertas filas por existencia de demasiados NaN
- Eliminación de decimales y/o redondear valores
- Conversión de los valores de tipo object a tipo int o float
- Ajustar la columna de la leyenda como índice
- Ajustes de los nombres de los países
- Eliminación de múltiples columnas vacías o sin interés para nuestro análisis

Debido a la naturaleza de los datos, principalmente por tratarse todos de tipo *panel data* (cada uno mide conceptos diferentes), no se han unificado en un único dataframe. Por cada dataset se ha creado un notebook con el procedimiento de limpieza de este y una vez acabada la limpieza se ha exportado a un archivo nuevo csv con el dataset limpio.

Posteriormente se ha creado otro archivo ipynb para cada uno de los datasets que contiene la creación de los gráficos, así como los cálculos necesarios para, por ejemplo, calcular la media de los países mostrados en el gráfico.

Durante todo este trabajo ha primado la búsqueda de la consistencia; si han faltado algunos datos, estos no se han añadido al dataset original con datos provenientes de otras fuentes. En el ámbito de la economía y las finanzas, existen muchos criterios diferentes para calcular los indicadores macroeconómicos, por tanto, los criterios para el cálculo de indicadores por parte de EUROSTAT no tienen por qué ser los mismos que los del FMI; por esta razón se han obtenido la mayor cantidad de datos posible únicamente de EUROSTAT.

Para la presentación de los resultados también ha primado la consistencia. Al principio se hicieron diferentes pruebas en búsqueda de un estilo de gráfico que fuera capaz de representar la mayor cantidad posible de variables (países) con la mayor claridad posible. Una vez creado el código base del gráfico, se ha utilizado para cada uno de los datasets, haciendo los ajustes necesarios para su correcto funcionamiento. Siguiendo las directrices para una buena consistencia y claridad, se han creado una paleta de colores fija con colores con bastante contraste. En todos los gráficos España está representada con un color rosa fucsia y, en el caso de haber una variable que describe una media, esta viene representada por el negro. Con el resto de países se ha querido mantener el mismo color para cada uno de ellos. Por esta razón los países analizados en la mayor parte de los casos son siempre los mismos y se han ploteado en el gráfico siempre de manera ordenada. Únicamente existen dos gráficos (ambos tasa de desempleo) con un estilo muy similar pero no completamente igual debido a un problema a la hora de mostrar la leyenda del eje x.

En un principio se planteó la posibilidad de crear mapas para mostrar los indicadores macroeconómicos para toda Europa. No obstante, finalmente esta idea se desechó por dos razones. En primer lugar, debido a los problemas encontrados a la hora de plotear el mapa (problemas con el PATH del Mac). En segundo lugar, por la pérdida de información que supone dibujar mapas que únicamente muestran los valores para un año en concreto sin poder ver la evolución, consecuentemente no pudiendo explicar las variaciones de este a lo largo de los años.

La elección de los datos también ha sido en parte complicada debido a, como mencionado anteriormente, la existencia de múltiples criterios a la hora de calcular los indicadores económicos. Dos importantes decisiones han sido tomadas, por una parte para los datos sobre el empleo se ha elegido el concepto “doméstico”; este se ajusta mejor a nuestro análisis, ya que toma en cuenta ajustes geográficos en términos de empleo. Por ejemplo, toma en cuenta los trabajadores en España que no son residentes en el país y resta aquellas personas que pese a ser oficialmente residentes en el país, trabajan en otro (no cotizando en España).

Por otra parte, también hay diferentes criterios a la hora de calcular la deuda de cada estado. Existen varios tipos de niveles a la hora de calcular la deuda, pero los dos más superiores son el “General Government” y el “Central Government”. La existencia de

estos dos es causada por la variedad en la organización administrativa de los estados en el mundo. Para nuestro análisis se ha elegido (en casi todos los casos) el “General Government” ya que es el término que abarca más en términos de deuda. No obstante, los datos del ratio deuda/PIB se ha elegido el “Central Government” debido a la imposibilidad de encontrar datos con el otro concepto. No obstante, se han comprobado los datos de los países representados con otras fuentes de datos, así como comparado con otros indicadores mostrados en el trabajo y no existen grandes diferencias. La elección en este caso de ese concepto no modifica el análisis que podemos extraer de las gráficas creadas y concuerda con el resto del análisis del trabajo.

Fuentes de los datos:

PIB en términos absolutos (2002-2022) (Precios corrientes y constantes)

[https://ec.europa.eu/eurostat/databrowser/view/NAMA\\_10\\_GDP\\_custom\\_5452592/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_GDP_custom_5452592/default/table?lang=en)

Crecimiento real del PIB (2011 - 2022)

[https://ec.europa.eu/eurostat/databrowser/view/TEC00115/default/table?lang=en&category=na10.nama10.nama\\_10\\_ma](https://ec.europa.eu/eurostat/databrowser/view/TEC00115/default/table?lang=en&category=na10.nama10.nama_10_ma)

Población empleada (2002-2022)

[https://ec.europa.eu/eurostat/databrowser/view/NAMA\\_10\\_PE\\_custom\\_5444527/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_PE_custom_5444527/default/table?lang=en)

Tasa de desempleo (2011-2022)

<https://ec.europa.eu/eurostat/databrowser/view/TPS00203/default/table?lang=en&category=labour.employ.lfsi.une>

Deuda acumulada (2002-2023) (en formato cuatrimestral)

[https://ec.europa.eu/eurostat/databrowser/view/GOV\\_10Q\\_GGDEBT\\_custom\\_5788872/default/table](https://ec.europa.eu/eurostat/databrowser/view/GOV_10Q_GGDEBT_custom_5788872/default/table)

Ratio deuda respecto al PIB (2002-2022)

[https://www.imf.org/external/datamapper/GG\\_DEBT\\_GDP@GDD/CHN/FRA/DEU/ITA/JPN/GBR/USA/RUS/ESP](https://www.imf.org/external/datamapper/GG_DEBT_GDP@GDD/CHN/FRA/DEU/ITA/JPN/GBR/USA/RUS/ESP)

Tamaño del gasto del sector público respecto al PIB (2010-2022)

[https://ec.europa.eu/eurostat/databrowser/view/TEC00023/default/table?lang=en&category=gov.govgfs10.gov\\_10a](https://ec.europa.eu/eurostat/databrowser/view/TEC00023/default/table?lang=en&category=gov.govgfs10.gov_10a)

Ingresos fiscales por categoría, estructura porcentual de los ingresos fiscales y presión fiscal tanto en España como en los países de la OCDE así como la media.

[https://www.hacienda.gob.es/es-ES/Areas%20Tematicas/Impuestos/Direccion%20General%20de%20Tributos/Paginas/Estadisticas\\_Recaudacion.aspx](https://www.hacienda.gob.es/es-ES/Areas%20Tematicas/Impuestos/Direccion%20General%20de%20Tributos/Paginas/Estadisticas_Recaudacion.aspx)

Deficit - Superávit del “General government” (2010-2021)

<https://ec.europa.eu/eurostat/databrowser/view/TEC00127/default/table>

Los resultados del trabajo vienen resumidos junto con la verificación de las hipótesis iniciadas planteadas:

*“El ratio de deuda/PIB de España actualmente se sigue manteniendo por encima de la época pre-pandemia pero este se ha empezado a reducir a una mayor velocidad que antes de la pandemia.”*

Hemos demostrado que efectivamente el ratio deuda/PIB en España se sitúa por encima de la época prepandemia y todo parece indicar que la reducción de este ha sido con mayor velocidad que anteriormente. No obstante, para poder realmente confirmar esta tendencia se necesitan más datos.

*“El aumento del endeudamiento español ha sido igual o menor al de los países de la Unión Europea.”*

Se ha demostrado que esta hipótesis es falsa, en todo caso el Estado Español se ha endeudado igual o más que el resto de países de la UE.

*“La presión fiscal en España es relativamente menor que en otros países, haciendo más necesario la emisión de deuda para financiar gastos.”*

Dependiendo de sobre qué países lo comparemos, la presión fiscal en España es menor o mayor. Actualmente se sitúa en la media de los países de la OCDE. Si comparamos con países como Chile, Estados Unidos o Irlanda, la presión fiscal es mucho mayor en España. Si, por el contrario comparamos con Suecia, Dinamarca o Italia, nuestra presión fiscal es mucho menor.

*“La pandemia afectó especialmente al PIB español y la recuperación de este ha sido más lenta que en el resto de países de la UE.”*

Totalmente verdadera, España fue el país (entre los de mayor PIB) con mayor afección a su PIB debido a la pandemia. La recuperación así mismo también ha sido relativamente más lenta, no obstante, se prevé que los próximos años España experimente unos niveles de crecimiento mayores que la mayoría de economías europeas.

*“El mercado laboral está mejorando sus datos, lo que hace prever un aumento de los ingresos por parte del Estado y, por tanto, una posible mayor reducción de deuda.”*

Verdadero. La buena evolución del mercado laboral no sólo aumenta los ingresos (más cotizaciones sociales e IRPF) sino que también reduce los gastos (menos gastos por prestaciones por desempleo). Todo aquello que mejore los ingresos y baje los gastos ayuda a una posible reducción de deuda.