

FEDER

Fondo Europeo de Desarrollo Regional

Una manera de hacer Europa

Evaluación del impacto de las grandes instalaciones científico-técnicas en los resultados de proyectos de investigación

Informe final



MINISTERIO
DE HACIENDA
Y FUNCIÓN PÚBLICA

SECRETARÍA GENERAL
DE FONDOS EUROPEOS

DIRECCIÓN GENERAL
DE FONDOS EUROPEOS



MINISTERIO
DE HACIENDA
Y FUNCIÓN PÚBLICA

Instituto de
Estudios Fiscales



Unión Europea



Informe final de evaluación

EVALUACIÓN DEL IMPACTO DE LAS GRANDES INSTALACIONES CIENTÍFICO-TÉCNICAS EN LOS RESULTADOS DE PROYECTOS DE INVESTIGACIÓN

**Plan de Evaluación Común FEDER 2014-2020
ES-002-Sect-O**

**Dirección General de Fondos Europeos (DGFE)
Instituto de Estudios Fiscales (IEF)
MINISTERIO DE HACIENDA Y FUNCIÓN PÚBLICA**

Evaluadora:
Bárbara López de Quintana Palacios (IEF)

Septiembre de 2023

Índice

Resumen ejecutivo	3
Executive summary	5
1. Introducción al objeto a evaluar	7
1.1. ¿Qué una gran infraestructura de investigación?	7
1.2. La participación española en grandes infraestructuras de investigación	8
1.3. La financiación de las grandes infraestructuras de investigación	9
2. El propósito de la evaluación y su metodología	10
2.1. El propósito de la evaluación	10
2.2. Los datos	10
2.3. La pregunta de evaluación	11
2.4. La metodología	11
2.5. Dificultades y limitaciones de la evaluación	16
3. Los hallazgos de la evaluación	18
3.1. Impacto en publicaciones en revistas científicas de alta calidad e influencia	19
3.2. Impacto en publicaciones en acceso abierto	24
3.3. Impacto en obtención de financiación en el Programa Marco de Investigación e Innovación de la Unión Europea	27
4. Conclusiones y recomendaciones	30
5. Referencias	32
Anexos	
1. Análisis de los supuestos para poder identificar el efecto causal	33
2. Los datos y su análisis exploratorio	35
3. Estimación del efecto causal	39
4. Análisis de sensibilidad	43

Resumen ejecutivo

Las grandes infraestructuras de investigación son instalaciones, recursos y servicios utilizados por la comunidad científica para llevar a cabo investigaciones de máxima calidad. Incluyen no solo equipamientos científicos, sino también colecciones, archivos e información científica estructurada, así como infraestructuras basadas en las tecnologías de la información y la comunicación. De acuerdo con el Consejo de la Unión Europea (UE), las infraestructuras de investigación constituyen un pilar fundamental de los ecosistemas de investigación e innovación europeos, pues son necesarias para que la comunidad científica lleve a cabo investigación de excelencia y desarrolle tecnologías de vanguardia que alimentan innovaciones, posibilitando a la UE ser más competitiva a escala global.

La participación española en grandes infraestructuras de investigación se estructura en dos niveles, el nacional y el internacional. En el nivel nacional destacan las Infraestructuras Científicas y Técnicas Singulares (ICTS), que son instalaciones de titularidad pública, únicas en su género en nuestro país, a las que pueden acceder, bajo criterios de excelencia, proyectos de investigación públicos y privados. Por su parte, el nivel internacional comprende infraestructuras científicas pertenecientes a la hoja de ruta ESFRI (Foro Estratégico Europeo sobre Infraestructuras de Investigación) y otras infraestructuras internacionales, como es el caso, por ejemplo, del CERN. Las infraestructuras de los niveles nacional e internacional colaboran entre sí y comparten recursos.

Las grandes infraestructuras de investigación requieren importantes inversiones estratégicas a lo largo de todo su ciclo de vida. Su financiación basal proviene de las Administraciones públicas titulares de las mismas; además, cuentan con los ingresos por tarifas asociadas al uso de sus instalaciones y pueden contar con ingresos competitivos o convenientes procedentes de fondos regionales, nacionales y europeos. **Los Programas Operativos españoles del Fondo Europeo de Desarrollo Regional (FEDER) del periodo 2014-2020 han destinado 532,8 millones de euros de gasto a la construcción, ampliación o mejora de las instalaciones y equipamientos de grandes infraestructuras de investigación.**

Dada la relevancia del FEDER en la financiación de las grandes infraestructuras de investigación, el Plan de Evaluación Común de los Programas Operativos del FEDER 2014-2020 incluye una evaluación del impacto que tiene el uso de dichas infraestructuras sobre los resultados de los proyectos de I+D. **La evaluación ha analizado si existe una relación causal entre hacer uso de una gran infraestructura de investigación y los siguientes resultados de los proyectos de I+D: publicaciones en revistas indexadas, en revistas indexadas en el primer cuartil, en revistas no indexadas, y en acceso abierto; tesis doctorales; patentes concedidas, licenciadas y en explotación; acuerdos de colaboración; empresas de base tecnológica creadas; y obtención de financiación en el Programa Marco de Investigación e Innovación de la UE.** Para ello, la evaluación ha aplicado la metodología de variables instrumentales a datos observacionales asociados al uso de ICTS por parte de proyectos de investigación financiados por convocatorias de ayudas de la Agencia Estatal de Investigación, dependiente del Ministerio de Ciencia e Innovación. Dichas convocatorias, que a su vez están cofinanciadas por el FEDER, constituyen la principal fuente de financiación competitiva de los proyectos de investigación desarrollados en nuestro país por universidades y demás organismos de investigación públicos y privados sin ánimo de lucro.

El análisis de impacto realizado estima que, para proyectos de una calidad científico-técnica media, usar una ICTS para llevar a cabo actividades de investigación tiene un efecto positivo sobre el número de publicaciones en revistas indexadas, en particular, el número de publicaciones en revistas indexadas que pertenecen al primer cuartil; el número de publicaciones en acceso abierto; y la obtención de financiación en el altamente competitivo Programa Marco de Investigación e Innovación de la UE.

Publicar un resultado científico en una revista indexada es un indicador de la calidad y relevancia de la investigación, especialmente si la revista se sitúa en el primer cuartil de su ranking temático. Asimismo, la publicación en revistas indexadas facilita la difusión de los resultados de investigación, lo que estimula el avance del conocimiento científico y puede sustentar, durante prolongados periodos de tiempo, numerosas e inesperadas innovaciones; así lo resalta un reciente estudio del Fondo Monetario Internacional (FMI), que pone como ejemplo el extraordinariamente rápido desarrollo de las vacunas contra el COVID-19, basadas en décadas de investigación científica previa difundida por artículos académicos. Por otro lado, el acceso abierto a publicaciones científico-técnicas, al aumentar exponencialmente la visibilidad y la accesibilidad de sus contenidos, juega un papel crucial en impulsar el impacto, la calidad, la eficiencia, la transparencia y la ética de la investigación y la innovación. Por último, obtener financiación en el Programa Marco es un indicador de que la propuesta de investigación es excelente en su concepción, viable en su implementación y relevante por sus potenciales impactos científicos, económicos y sociales.

De acuerdo con lo expuesto en el párrafo anterior, **los efectos positivos estimados sugieren que el uso de grandes infraestructuras de investigación impacta positivamente sobre la calidad de los resultados de los proyectos de investigación, su capacidad de influencia sobre futuras investigaciones e innovaciones y su capacidad de contribuir al bienestar de la ciudadanía europea y a la competitividad de los Estados miembros.**

No se ha detectado que hacer uso de una gran infraestructura de investigación tenga efectos ni positivos ni negativos sobre el resto de los resultados analizados, esto es, publicaciones en revistas no indexadas, tesis doctorales, patentes, empresas de base tecnológica creadas y acuerdos de colaboración. De todos modos, hay que señalar que es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a tesis doctorales, patentes y empresas de base tecnológica generadas por los proyectos, debido a que el tiempo que requiere producir estos resultados hace que, en muchos casos, se alcancen después del momento en que se registran los indicadores de resultado de los proyectos (normalmente, tres meses después de su finalización).

Los hallazgos de la evaluación sustentan una recomendación en línea con las conclusiones del Consejo de la UE sobre infraestructuras de investigación, adoptadas en diciembre de 2022: se recomienda a los distintos niveles administrativos (europeo, estatal y regional) que sigan colaborando para garantizar la sostenibilidad a largo plazo de las infraestructuras de investigación; en particular, **dada la relevancia del FEDER en la financiación de las grandes infraestructuras de investigación, se recomienda a la Comisión Europea y a las autoridades españolas que sigan impulsando las sinergias entre las diversas fuentes de financiación europeas, nacionales y regionales.** Asimismo, dadas las dificultades y limitaciones afrontadas por la presente evaluación causadas por la ausencia en bases de datos de información suficiente sobre el uso de grandes infraestructuras de investigación, se recomienda a las agencias financiadoras estatales y regionales y a las Administraciones titulares de las infraestructuras que implementen mecanismos de seguimiento adecuados que permitan profundizar en el conocimiento y la comprensión de los impactos que tiene el uso de estas grandes infraestructuras de investigación sobre las actividades de investigación llevadas a cabo por los sectores público y privado.

Executive summary

Large research infrastructures are facilities, resources and services used by the scientific community to carry out research of the highest quality. They include not only scientific equipment, but also collections, archives and structured scientific information, as well as infrastructures based on information and communication technologies. According to the Council of the European Union (EU), research infrastructures are a fundamental pillar of European research and innovation ecosystems, as they are necessary for the scientific community to carry out excellent research and develop cutting-edge technologies that fuel innovations, enabling the EU to be more competitive on a global scale.

Spanish participation in large research infrastructures is structured on two levels, national and international. At the national level, the Singular Scientific and Technical Infrastructures (ICTS) stand out, which are unique facilities of their kind at the national level. Their ownership is public – state, regional or mixed – and public and private research projects can access to them under criteria of excellence. For its part, the international level includes scientific infrastructures belonging to the ESFRI roadmap (European Strategic Forum on Research Infrastructures) and other international infrastructures, such as CERN. Infrastructures at the national and international levels collaborate with each other and share resources.

Large research infrastructures require significant strategic investments throughout their life cycle. Its basal financing comes from the public administrations that own them. In addition, they have the revenue from fees for the use of their facilities and may have competitive or agreed income from regional, national and European funds. **The Spanish Operational Programmes of the European Regional Development Fund (ERDF) for the period 2014-2020 allocated 532.8 million euros of expenditure to the construction, expansion or improvement of the facilities and equipment of large research infrastructures.**

Given the importance of the ERDF in the financing of large research infrastructures, the Common Evaluation Plan of the ERDF Spanish Operational Programmes 2014-2020 includes an evaluation on the impact of the use of these infrastructures on the results of R&D projects. **The evaluation has analyzed whether there is a causal relationship between making use of a large research infrastructure and the following results of R&D projects: publications in indexed journals, in journals indexed in the first quartile, in non-indexed journals, and in open access; doctoral theses; patents granted, licensed and in operation; collaboration agreements; technology-based companies created; and securing funding under the EU's Framework Programme for Research and Innovation.** To this end, the evaluation has applied the methodology of instrumental variables to observational data associated with the use of ICTS by R&D projects funded by calls for grants from the State Research Agency, under the Ministry of Science and Innovation. These calls, which in turn are co-financed by the ERDF, constitute the main source of competitive funding for research projects developed in Spain by universities and other public and private non-profit research organizations.

The impact evaluation carried out estimates that, for projects of average scientific and technical quality, using an ICTS to carry out research activities has a positive effect on the number of publications in indexed journals and, in particular, the number of publications in Q1 indexed journals; the number of open access publications; and securing funding under the EU's highly competitive Framework Programme for Research and Innovation.

Publishing a scientific result in an indexed journal is an indicator of the quality and relevance of the research, especially if the journal is in the first quartile of its thematic ranking. Likewise, publication in indexed journals facilitates the dissemination of research results, which stimulates the advancement of scientific knowledge and can sustain, over long periods of time, numerous and unexpected innovations. This is highlighted by a recent study by the International Monetary Fund (IMF), which gives as an example the extraordinarily rapid development of vaccines against COVID-19, based on decades of previous scientific research disseminated by articles published in scientific journals. On the other hand, open access to scientific-technical publications exponentially increases the visibility and accessibility of their contents, playing a crucial role in boosting the impact, quality, efficiency, transparency and ethics of research and innovation. Finally, obtaining funding under the Framework Programme is an indicator that the research proposal is excellent in its conception, viable in its implementation and relevant for its potential scientific, economic and social impacts.

In accordance with the previous paragraph, **the estimated positive effects suggest that the use of large research infrastructures has a positive impact on the quality of the results of research projects, their ability to influence future research and innovations and their ability to contribute to the well-being of European citizens and the competitiveness of the Member States.**

It has not been detected that making use of a large research infrastructure has either positive or negative effects on the rest of the results analyzed, that is, publications in non-indexed journals, doctoral theses, patents, technology-based companies created and collaboration agreements. In any case, it should be noted that it is likely that the values of the database of the State Research Agency relating to doctoral theses, patents and technology-based companies generated by the projects are not adjusted to reality, because the time required to produce these results means that, in many cases, are achieved after the time when project outcome indicators are recorded (usually three months after completion).

The findings support a recommendation in line with the EU Council conclusions on research infrastructures adopted in December 2022: it is recommended that the different administrative levels (European, state and regional) continue to collaborate to ensure the long-term sustainability of research infrastructures; in particular, **given the importance of the ERDF in the financing of large research infrastructures, it is recommended that the European Commission and the Spanish authorities continue to promote synergies between the various European, national and regional funding sources.** Likewise, given the difficulties and limitations faced by the evaluation caused by the insufficient information in databases on the use of large research infrastructures, it is recommended that state and regional funding agencies and the administrations that own the infrastructures implement adequate monitoring mechanisms on the use of large research infrastructures.

1. Introducción al objeto a evaluar

1.1. ¿Qué es una gran infraestructura de investigación?

El Consejo de la Unión Europea (UE) define las grandes infraestructuras de investigación como **las instalaciones, los recursos y los servicios utilizados por la comunidad científica para llevar a cabo investigaciones de máxima calidad** [1]. Esta definición incluye:

- equipamientos científicos como, por ejemplo, las fuentes de luz de sincrotrón, los grandes telescopios, los buques oceanográficos, los observatorios medioambientales, o las “salas blancas” para el desarrollo de nuevos materiales o de nanoelectrónica;
- colecciones, archivos e información científica estructurada como, por ejemplo, los bancos de datos sobre genómica o ciencias sociales;
- infraestructuras basadas en las tecnologías de la información y la comunicación (TIC), tales como las redes de fibra oscura para colaboración telemática o los supercomputadores para aplicaciones que requieren alta potencia de cálculo.

De acuerdo con el Consejo de la UE, **las grandes infraestructuras de investigación constituyen un pilar fundamental de los ecosistemas de investigación e innovación de la UE y contribuyen al desarrollo regional y a la formación, atracción y retención de talento**. A continuación, se resumen las principales conclusiones del Consejo sobre infraestructuras de investigación adoptadas en diciembre de 2022 [2]:

- Las grandes infraestructuras de investigación son necesarias para que el sector público y el privado lleve a cabo investigación de excelencia y desarrolle tecnologías de vanguardia que alimentan innovaciones, posibilitando a la UE ser más competitiva a escala global.
- Además, las inversiones públicas en grandes infraestructuras de investigación impactan positivamente sobre las empresas que suministran productos y servicios de alta tecnología para su construcción, equipamiento, funcionamiento y mantenimiento. Su participación no solo mejora sus resultados empresariales, sino que incrementa sus capacidades en I+D+I y les abre nuevos mercados, al tener a las infraestructuras como clientes de referencia.
- Las grandes infraestructuras de investigación contribuyen al desarrollo regional, ya que ofrecen puestos de trabajo altamente cualificados y estimulan el desarrollo de nuevas empresas, cadenas de suministro, infraestructuras civiles y servicios conexos.
- Las políticas de acceso abierto de las grandes infraestructuras de investigación promueven una formación especializada y de alta calidad para el personal de I+D+I, contribuyen a la circulación de conocimientos y talento dentro de Europa, y atraen talento de todo el mundo, mejorando así la colaboración internacional en investigación e innovación y reforzando la competitividad del ecosistema europeo de investigación e innovación.
- Los Estados miembros y la Comisión Europea han de trabajar conjuntamente para garantizar la sostenibilidad a largo plazo de las grandes infraestructuras de investigación, impulsando las sinergias entre las diversas fuentes de financiación europeas, nacionales y regionales, de modo que se garantice una planificación financiera estable, fiable y predecible.

1.2. La participación española en grandes infraestructuras de investigación

La participación española en grandes infraestructuras de investigación **se estructura en dos niveles, el nacional y el internacional.**

- **En el nivel nacional destacan las Infraestructuras Científicas y Técnicas Singulares (ICTS)**, que son instalaciones únicas en su género a nivel nacional y a las que pueden acceder, bajo criterios de excelencia, proyectos de investigación públicos y privados nacionales y extranjeros. Las ICTS son de titularidad pública (estatal, autonómica o mixta) y están distribuidas por toda la geografía española. La condición de ICTS se adquiere mediante la inclusión en el denominado “Mapa de ICTS”, instrumento creado en 2007 por la III Conferencia de Presidentes autonómicos para planificar y desarrollar la dotación de infraestructuras científico-técnicas de vanguardia de nuestro país. El Mapa es revisado y aprobado cada cuatro años por el Consejo de Política Científica, Tecnológica y de Innovación, máximo órgano de coordinación de la I+D+I en España, integrado por representantes estatales y autonómicos al más alto nivel.

El Mapa actual, aprobado para el periodo 2021-2024, está integrado por 29 ICTS que aglutinan 64 infraestructuras¹ que operan en prácticamente todas las disciplinas del conocimiento: astronomía, ciencias del mar, de la vida, de la tierra, tecnologías de la información y comunicaciones, ciencias de la salud, biotecnología, energía, ingeniería, materiales, y ciencias socio-económicas y humanidades. Al menos el 20% de la capacidad de las instalaciones esenciales de cada infraestructura debe ofrecerse en modo de acceso abierto competitivo; bajo esta modalidad, la ICTS evalúa las solicitudes de acceso y concede el mismo a aquellas de mayor calidad científico-técnica, otorgándoles además tarifas de acceso más bajas. De este modo, las ICTS proporcionan anualmente acceso competitivo a más de 5.000 proyectos de I+D que hacen uso de sus instalaciones para experimentos, simulaciones, observaciones del medio y obtención o procesamiento de datos.

- **El nivel internacional comprende tanto infraestructuras de investigación internacionales**, como es el caso, por ejemplo, del CERN, **como infraestructuras científicas pertenecientes a la hoja de ruta ESFRI** (Foro Estratégico Europeo sobre Infraestructuras de Investigación). La iniciativa ESFRI surgió en 2002 a instancias del Consejo de la UE con el objetivo de coordinar una estrategia europea común en materia de infraestructuras de investigación. De las 37 infraestructuras de carácter paneuropeo implementadas hasta el momento, España participa en 21, de las cuales 4 tienen su sede estatutaria en nuestro país.

Ambos niveles están interconectados, fomentándose que las ICTS colaboren activamente con las infraestructuras pertenecientes a la Hoja de Ruta ESFRI o con las grandes infraestructuras internacionales; asimismo, las infraestructuras ESFRI en las que participa España se nutren siempre que es posible de los recursos y capacidades de las ICTS.



Figura 1. El CERN (Conseil Européen pour la Recherche Nucléaire), situado en Ginebra (Suiza), es el mayor laboratorio de investigación en física de partículas elementales a nivel mundial. España tiene presencia como miembro de pleno derecho en el Consejo del CERN, ocupando el quinto lugar en cuanto a contribuciones, después de Alemania, Reino Unido, Francia e Italia. Fuente: Ministerio de Ciencia e Innovación.

¹ Hay ICTS de localización única y otras con nodos distribuidos en distintas localizaciones.

1.3. La financiación de las grandes infraestructuras de investigación

Las infraestructuras de investigación requieren importantes inversiones estratégicas a lo largo de todo su ciclo de vida. Su financiación basal proviene de las Administraciones públicas titulares de las mismas; además, cuentan con los ingresos por tarifas asociadas al uso de sus instalaciones y pueden contar con ingresos competitivos o convenientes procedentes de fondos regionales, nacionales y europeos. **Los Programas Operativos españoles del Fondo Europeo de Desarrollo Regional (FEDER) del periodo 2014-2020 han destinado 532,8 millones de euros de gasto a la construcción, ampliación o mejora de las instalaciones y equipamientos de grandes infraestructuras de investigación²**, lo que representa el 40% del gasto comprometido en el objetivo específico 1.1.2. “Fortalecimiento de las instituciones de I+D y creación, consolidación y mejora de las infraestructuras científicas y tecnológicas”.

La sostenibilidad a largo plazo de las grandes infraestructuras de investigación requiere el compromiso político de las Administraciones titulares y de la Comisión Europea. Así lo pone de manifiesto en Consejo de la UE en sus conclusiones sobre infraestructuras de investigación adoptadas en diciembre de 2022 [2] en las que:

- Invita a los Estados miembros y a la Comisión a que lleven a cabo un análisis de los tipos de apoyo financiero a las infraestructuras de investigación europeas a lo largo de su ciclo de vida, identifiquen buenas prácticas y sinergias de los diversos recursos de financiación de las infraestructuras de investigación y exploren modelos e iniciativas adecuadas para la financiación de dichas infraestructuras.
- Pide a los Estados miembros y a la Comisión que coordinen sus esfuerzos para promover e impulsar sinergias en la financiación de las infraestructuras de investigación a nivel regional, nacional, macrorregional y europeo, con el fin de aumentar la eficiencia de las inversiones en las infraestructuras de investigación.



Figura 2. ICTS Sincrotrón ALBA, radicada en Cataluña. Es un complejo de aceleradores de electrones destinado a producir luz de sincrotrón con la que visualizar la estructura y propiedades de la materia, especialmente a escala nanométrica. Fuente: Ministerio de Ciencia e Innovación.

² Para poner en contexto esta cifra respecto a la inversión pública total en grandes infraestructuras de investigación, solo se dispone de datos correspondientes a la inversión en ICTS en 2019-2021: en este periodo, prácticamente la totalidad de la ayuda pública que recibieron las ICTS para infraestructuras y equipamientos estuvo cofinanciada por el FEDER o por el Mecanismo para la Recuperación y la Resiliencia (MRR). La ayuda destinada a estos conceptos representó en torno al 25% del total de la ayuda que las Administraciones públicas españolas concedieron a las ICTS; el resto estuvo destinado a gastos de operación y mantenimiento, atracción de talento, fortalecimiento institucional y realización de proyectos de I+D y de transferencia de conocimiento.

2. El propósito de la evaluación y su metodología

2.1. El propósito de la evaluación

La presente evaluación está contemplada en el Plan de Evaluación Común de los Programas Operativos del FEDER 2014-2020. Su propósito es evaluar qué impacto tienen las grandes infraestructuras de investigación sobre los resultados de los proyectos de investigación que hacen uso de las mismas para realizar experimentos, simulaciones, observaciones, o para obtener o procesar datos.

2.2. Los datos

La manera más eficiente para estimar qué efecto tiene sobre los proyectos acceder a una gran infraestructura de investigación sería tomar un conjunto suficientemente grande de proyectos con actividades programadas susceptibles de beneficiarse del uso de las infraestructuras y asignar aleatoriamente a unos proyectos el acceso (grupo de tratamiento) y a otros no (grupo de control). De esta manera se asegura que los proyectos de los grupos de tratamiento y control tengan características idénticas; el efecto causal del uso de una gran infraestructura de investigación sería entonces la diferencia en las medias de las variables de interés en los grupos de tratamiento y control. Un diseño de este tipo —conocido como diseño experimental o experimento aleatorio— sería difícil de implementar debido tanto a los protocolos que regulan el acceso competitivo a las grandes infraestructuras de investigación como al dilema que supondría negar el acceso a proyectos de calidad que necesitan usar una gran infraestructura para alcanzar sus objetivos. Hemos de recurrir por tanto a datos observacionales, esto es, datos que no han sido obtenidos a través de experimentos aleatorios, y emplear técnicas de inferencia causal que nos permitan estimar el efecto del acceso a grandes infraestructuras de investigación.

Los datos observacionales de los que se ha dispuesto para realizar la evaluación son los asociados a los proyectos de investigación financiados por la Agencia Estatal de Investigación, dependiente del Ministerio de Ciencia e Innovación. Entre otras actuaciones, la Agencia Estatal de Investigación gestiona convocatorias de ayudas cofinanciadas por el FEDER que constituyen la principal fuente de financiación competitiva de los proyectos de investigación desarrollados en nuestro país por universidades, organismos públicos de investigación y organismos de investigación privados sin ánimo de lucro. La Agencia Estatal de Investigación evalúa la calidad de los proyectos presentados a las convocatorias y financia los mejores; la financiación va dirigida a gastos de personal, material inventariable y fungible, gastos de viajes, publicaciones científicas, patentes, y otros gastos directamente relacionados con el proyecto, como los costes de utilización de grandes infraestructuras de investigación. La cuantía concedida a cada proyecto depende de la puntuación otorgada en su evaluación, de la ayuda solicitada, de su coste financiable real y de las disponibilidades presupuestarias de cada convocatoria de ayudas. La duración de los proyectos, determinada por las correspondientes convocatorias, varía de dos a seis años. Al finalizar la ejecución del proyecto, se ha de presentar un informe final que, entre otra información, recoge una serie de indicadores directamente relacionados con los resultados del proyecto, los cuales se graban en la base de datos de la Agencia Estatal de Investigación: número de publicaciones en revistas indexadas³, en revistas indexadas en el primer cuartil Q1², en revistas no indexadas, y en acceso abierto; número de tesis doctorales leídas; número de patentes

³ Son revistas que están indexadas en determinados portales de información (como *Journal Citation Reports, JCR*). Estos portales analizan revistas científicas cuyos artículos son revisados por personas expertas con carácter previo a

concedidas, licenciadas y en explotación; número de acuerdos de colaboración; número de empresas de base tecnológica creadas; y número de proyectos financiados por el Programa Marco de Investigación e Innovación de la UE.

Para el periodo de programación FEDER 2014-2020, en la fecha de la consulta (16 de enero de 2023) la base de datos de la Agencia Estatal de Investigación disponía de 13.893 registros correspondientes a proyectos finalizados para los que había grabados tanto indicadores de resultado como información relativa a si, en el momento de la solicitud, pretendían hacer uso de alguna ICTS⁴. Un 8,5% de dichos registros (1.191) correspondían a proyectos que sí planificaban acceder a una ICTS. Se examinaron los informes finales de ejecución de dichos proyectos, encontrándose que 854 sí habían hecho uso de una gran infraestructura de investigación para realizar las actividades planificadas. Estos proyectos han constituido el grupo de tratamiento. De los proyectos restantes se han eliminado aquellos pertenecientes a áreas de conocimiento para las que ningún proyecto ha accedido a una ICTS (ciencias de la educación, ciencias sociales, derecho, economía, estudios feministas, filosofía y filología, psicología), quedando 8.157 proyectos como grupo de control. Estos son los datos observacionales de los que se ha dispuesto para realizar la evaluación. El Anexo 2 recoge una descripción de sus variables más relevantes y un análisis exploratorio de las mismas. Respecto a las variables de resultado, hay que señalar que todas presentan números valores atípicos y que seis de ellas tienen un valor nulo en más del 75% de los proyectos analizados (tercer cuartil = 0): número de patentes concedidas, licenciadas y en explotación; número de acuerdos de colaboración; número de empresas de base tecnológica creadas; y número de proyectos relacionados financiados por el Programa Marco de Investigación e Innovación de la UE.

2.3. La pregunta de evaluación

Los datos observacionales disponibles nos permiten cumplir con el propósito de la evaluación respondiendo a la siguiente pregunta:

¿El uso de instalaciones científico-técnicas singulares (ICTS) afecta a los resultados —publicaciones, instrumentos de protección y explotación de la propiedad industrial, acuerdos de colaboración, empresas de base tecnológica creadas y obtención de financiación en el Programa Marco de la UE— de los proyectos de investigación?

2.4. La metodología

La inferencia causal es el proceso que permite identificar y cuantificar relaciones causales. En nuestro caso, partiendo de datos observacionales y no de un experimento aleatorio, nos ha permitido evaluar si hacer uso de una ICTS presenta una asociación causal con diferentes resultados de los proyectos de investigación. Para ello, se ha seguido un procedimiento que consta de cuatro etapas:

su publicación (“revisión por pares”) y establecen, para cada área temática, rankings de revistas, en función de las veces que sus artículos son referenciados por otros artículos científicos. La élite de las revistas en una temática está conformada por aquellas que ocupan el primer 25% del listado (primer cuartil Q1).

⁴ Como se explica en el Anexo 2, en la fase de solicitud y para una serie de convocatorias de ayudas a proyectos, la Agencia Estatal de Investigación recopila y graba en su base de datos información relativa a si el proyecto planifica acceder a alguna ICTS; no se recopila información relativa a si se planifica acceder a una ESFRI o a otras grandes infraestructuras de investigación. Al finalizar la ejecución de los proyectos, no se graba en base de datos si el proyecto finalmente accedió a la ICTS; esta información hay que recabarla examinando las actividades reseñadas en el informe final de ejecución.

- En primer lugar, se ha utilizado el conocimiento experto sobre la materia para construir, para cada resultado de interés, una herramienta visual, denominada diagrama causal, que refleja las hipótesis de las que vamos a partir para estimar el efecto que tiene el uso de una ICTS sobre el resultado.
- En segundo lugar, se ha identificado, con el apoyo del diagrama causal, si es posible estimar a partir de los datos el efecto del uso de ICTS sobre el resultado de interés.
- En tercer lugar, se han empleado métodos estadísticos para estimar dicho efecto.
- Por último, se ha analizado la sensibilidad de las estimaciones obtenidas a incumplimientos de las hipótesis de partida reflejadas en el diagrama causal.

A) Construcción del diagrama causal

Un diagrama causal es una herramienta visual que recoge el conocimiento disponible sobre las posibles vías de asociación entre una causa o “tratamiento” y el resultado de interés. En el caso que nos ocupa, la consulta a expertos en la materia ha permitido formular las **hipótesis sobre las posibles vías de asociación entre el uso de una ICTS y los resultados de interés**. Por ejemplo, en el caso del resultado “Número de publicaciones en revistas científicas” se han formulado las siguientes hipótesis:

- El área del conocimiento en la que se encuadra el proyecto, el tipo de convocatoria de ayudas y el tipo de solicitante afectan a los objetivos del proyecto.
- Dichos objetivos determinan si se planifica hacer uso de una o varias ICTS; los objetivos también determinan la financiación solicitada para ejecutar el proyecto.
- La Agencia Estatal de Investigación evalúa la calidad científico-técnica de la propuesta presentada y otorga una puntuación, financiando aquellas propuestas con mejor puntuación hasta agotar la disponibilidad presupuestaria de la convocatoria. La cuantía concedida a cada propuesta puede ser igual o inferior a la solicitada y depende del coste financiable real estimado por la Agencia Estatal de Investigación, de la puntuación otorgada, del presupuesto de la convocatoria y del marco europeo que regula las ayudas estatales que se pueden conceder para no alterar la competencia en el mercado interior europeo.
- El presupuesto del que finalmente dispone el proyecto puede afectar al uso de la ICTS —por las posibles tarifas por uso, costes de personal, desplazamientos...— y al resto de recursos materiales y humanos a disposición del proyecto. Además del presupuesto, el que finalmente se haga uso de una determinada ICTS dependerá de la disponibilidad de los equipos o servicios de la misma en función de la demanda que ha de afrontar, de la nota otorgada por la Agencia Estatal de Investigación al proyecto si el acceso a la ICTS es en régimen competitivo⁵, y de si el grupo de investigación tiene algún tipo de ventaja para acceder a la ICTS planificada, por ejemplo, porque pertenece a la institución propietaria de la misma o a otra con algún tipo de acuerdo de colaboración para el acceso. El tener algún tipo de ventaja para acceder a una determinada ICTS puede afectar a su vez a si se planifica hacer uso de la misma.
- Finalmente, el resultado de interés “Número de publicaciones en revistas científicas”, puede verse afectado, además de por el uso de la ICTS, por la calidad científico-técnica de la propuesta, los recursos económicos, materiales y humanos a disposición del proyecto, el área de conocimiento, el tipo de solicitante y los objetivos del proyecto.

⁵ Los protocolos que regulan el acceso competitivo a una ICTS habitualmente tienen en cuenta para priorizar el acceso la puntuación otorgada por la evaluación científica-técnica de la Agencia Estatal de Investigación.

El diagrama causal de la Figura 3 recoge las hipótesis anteriores. Los nodos del diagrama representan variables⁶, las aristas entre los pares de nodos representan posibles relaciones causales entre las variables⁷ y las puntas de flecha indican el sentido del flujo de causalidad. La asociación causal a analizar está representada por la flecha verde que va del nodo de la variable de tratamiento “Uso de ICTS” al nodo de la variable de resultado “Publicaciones”. El resto de los caminos entre ambos nodos, compuestos por las aristas de color fucsia, corresponden a posibles asociaciones no causales entre el tratamiento y el resultado de interés; son asociaciones no causales porque el flujo de causalidad indicado por las flechas no fluye desde el nodo tratamiento al nodo resultado, ni de manera directa ni indirecta (a través de terceras variables). El diagrama causal es especialmente útil para identificar estas asociaciones no causales y así poder tenerlas en cuenta en la posterior estimación del efecto del tratamiento sobre el resultado, de modo que no se introduzcan sesgos.

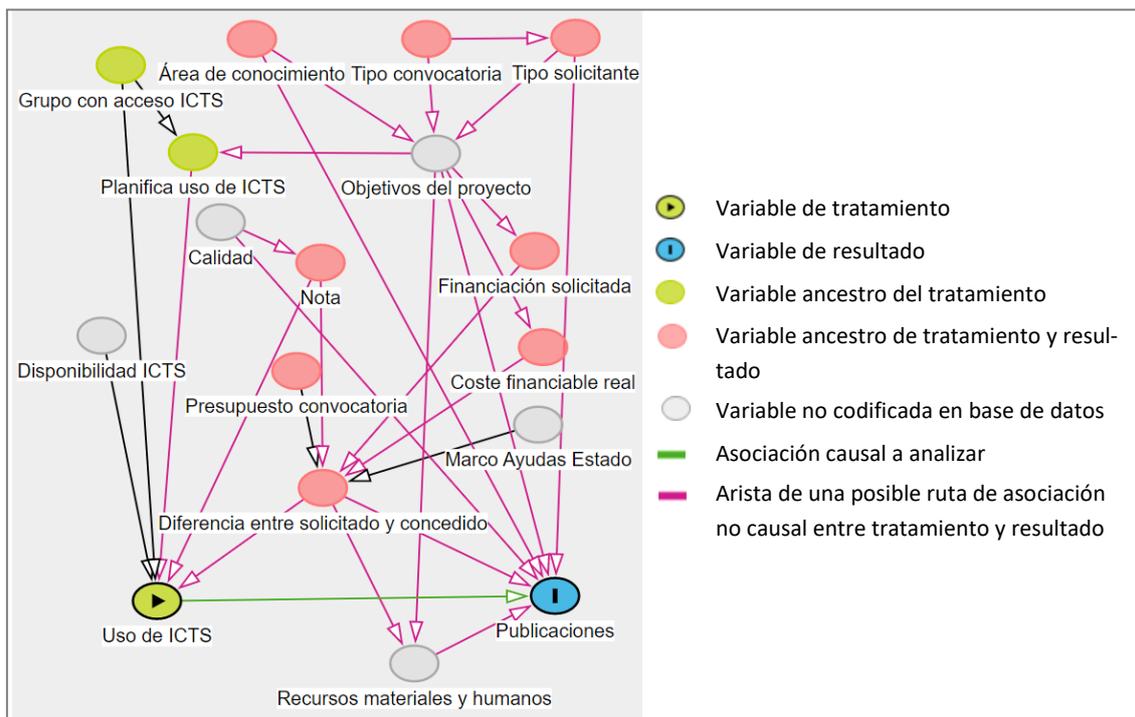


Figura 3. Diagrama causal que refleja las hipótesis de partida para estimar el efecto que tiene el uso de una ICTS sobre el número de publicaciones. Elaboración propia con el paquete de R “dagitty” [3].

⁶ No es necesario incluir todas las variables implicadas en el fenómeno a estudiar, es suficiente incluir la variable de tratamiento, la variable de resultado, las variables que son causa común tanto del tratamiento como del resultado, y las variables que son causa común de al menos dos de las variables que aparezcan en el diagrama. Así, en nuestro caso, no hubiera sido necesario incluir las variables “Disponibilidad ICTS”, “Presupuesto convocatoria” y “Marco Ayudas Estado”; se han incluido con fines de claridad en cuanto a las hipótesis de partida. La variable “Grupo con acceso ICTS”, aunque entra en esta categoría, se ha incluido porque, cómo se explicará posteriormente, es una variable “instrumental” que va a jugar un papel fundamental en la estimación del efecto causal.

⁷ La conexión entre cualquier par de nodos no indica que exista una relación causal entre las variables que representan, sino simplemente que puede existir una relación causal. En cambio, el que dos nodos no estén conectados implica que se parte de la hipótesis de que las variables que representan no están relacionadas de ninguna manera.

B) Identificación de si es posible estimar el efecto causal

Para identificar si es posible estimar el efecto causal, una aproximación comúnmente adoptada es la denominada *backdoor criterion* o “criterio de puerta trasera”, que consiste en comprobar, con ayuda del diagrama causal, si se pueden cerrar todos los caminos de asociación no causal entre el tratamiento (“Uso de ICTS”) y el resultado (por ejemplo, “Número de publicaciones en revistas científicas”). En nuestro caso, se pueden cerrar controlando por las variables “Nota”, “Diferencia entre solicitado y concedido”, “Grupo con acceso ICTS” y “Planifica uso de ICTS”. Entonces, es posible estimar el efecto causal empleando diversos tipos de estimadores siempre que se cumplan cuatro supuestos: el modelo causal representado por el diagrama causal contiene todas las variables que afectan a la asociación entre tratamiento y resultado (*supuesto de selección en observables*); para cada proyecto, no existen diferentes versiones del tratamiento que conduzcan a diferentes resultados potenciales⁸ (*supuesto de consistencia*); los resultados potenciales de cada proyecto no se ven afectados por la exposición de otros proyectos al tratamiento (*supuesto de no interferencia*); y, para cada combinación de valores del conjunto de variables de control —“Nota”, “Diferencia entre solicitado y concedido”, “Grupo con acceso ICTS” y “Planifica uso de ICTS”—, podemos encontrar proyectos del grupo de tratamiento (hacen uso de una ICTS) y del grupo de control (no hacen uso) (*supuesto de positividad*). Aunque se acepte el cumplimiento de los tres primeros supuestos, que son inverificables empíricamente, sí que se puede comprobar que el cuarto no se cumple: hay una violación estructural del mismo porque todos los proyectos que no planifican acceder a una ICTS (valor de “Planifica uso ICTS” = 0) pertenecen al grupo de control. Por tanto, al violarse el cuarto supuesto, no es posible estimar de manera confiable el efecto del tratamiento empleando esta aproximación, ni otras que exigen positividad, como el emparejamiento o *matching*⁹.

La violación estructural del supuesto de positividad nos obliga a emplear otra aproximación: intentar encontrar una variable o “instrumento” que sea causa del tratamiento y analizar si cumple una serie de supuestos, los cuales, excepto uno, son inverificables empíricamente y solo se pueden aceptar o rechazar empleando conocimiento experto. En nuestro caso, como se puede observar en el diagrama causal, una posible causa de la variable de tratamiento es la variable “Grupo con acceso ICTS”, que representa si el grupo de investigación que lleva a cabo el proyecto tiene algún tipo de ventaja para acceder a alguna ICTS, por ejemplo, porque pertenece a la institución propietaria de la misma o a otra con algún tipo de acuerdo de colaboración para el acceso. Para ser considerada un instrumento válido, tienen que cumplirse los siguientes supuestos: el instrumento está asociado con el tratamiento (*supuesto de relevancia*); el instrumento solo afecta al resultado a través del tratamiento (*supuesto de exclusión*); el resultado y el instrumento no comparten causas no observadas (*supuesto de selección en observables para el caso instrumental*); para cada combinación de valores del conjunto de variables de control necesarias para que se cumplan los dos supuestos anteriores, hay proyectos que reciben el instrumento y otros que no (*supuesto de positividad*); para cada proyecto, no existen diferentes versiones del instrumento que conduzcan a diferentes tratamientos potenciales¹⁰ ni tampoco existen diferentes versiones del tratamiento que conduzcan a diferentes resultados potenciales (*supuesto de*

⁸ Cada proyecto tiene dos resultados potenciales: el que obtendría con tratamiento (hace uso de una ICTS) y el que obtendría sin tratamiento (no hace uso de la ICTS). Como cada proyecto, o bien hace uso de una ICTS, o bien no lo hace, solo uno de los resultados potenciales es observado; el resultado potencial no observado se suele denominar contrafactual.

⁹ Si, para subsanar la violación estructural de positividad, se restringiera el análisis a los proyectos que planeaban inicialmente el acceso a una gran infraestructura de investigación, seguiría existiendo un limitado soporte común por el reducido grupo de control que quedaría (de los 1.191 proyectos que planificaban acceder, 854 recibieron tratamiento, por lo solo quedarían 337 proyectos como grupo de control).

¹⁰ Cada proyecto tiene dos tratamientos potenciales: el que obtendría con instrumento (pertenencia a un grupo de investigación con ventaja para acceder a una ICTS) y el que obtendría sin instrumento (no pertenencia). Como cada

consistencia); los tratamientos potenciales de cada proyecto no se ven afectados por la exposición de otros proyectos al instrumento, ni tampoco los resultados potenciales de cada proyecto se ven afectados por la exposición de otros proyectos al tratamiento (*supuesto de no interferencia*); y el instrumento nunca puede desincentivar el tratamiento, aunque puede fallar en incentivarlo (*supuesto de monotonicidad*). En el Anexo 1 se analizan los supuestos anteriores, concluyéndose que es posible aceptar su cumplimiento, aunque hay que señalar que la información insuficiente con la que se ha contado para definir la variable de tratamiento ha obligado a hacerlo de una manera demasiado amplia o poco concreta (uso o no uso de una gran infraestructura de investigación), lo que compromete el cumplimiento del supuesto de consistencia y puede inducir un sesgo positivo en la magnitud de las estimaciones [5, pág. 14].

La aceptación de los supuestos requeridos por la técnica de variables instrumentales permite estimar el efecto causal a partir de los datos mediante diversos estimadores estadísticos propuestos en la literatura. La amplia mayoría de estos no permiten estimar el efecto promedio del tratamiento¹¹ entre la población de interés (por ejemplo, el número medio de publicaciones que generan los proyectos de investigación financiados por la Agencia Estatal de Investigación que hacen uso de ICTS para sus actividades de investigación respecto al número medio de publicaciones que obtendrían si no hicieran uso). Bajo el supuesto de monotonicidad, lo que estiman son medias ponderadas del efecto promedio local del tratamiento (también conocido como LATE por las siglas de su denominación en inglés, *Local Average Treatment Effect*), que es el efecto medio del tratamiento para aquellos individuos de la población cuyo valor del tratamiento cambiaría cuando cambiase el valor del instrumento (los individuos denominados “cumplidores”). El tipo de ponderación depende del estimador empleado.

C) Estimación del efecto causal

Se ha optado por emplear el estimador de variables instrumentales más habitual en la literatura, el estimador de mínimos cuadrados ordinarios en dos etapas¹². Como se explica en el Anexo 3, en el caso que nos ocupa —con una variable de tratamiento binaria, un único instrumento también binario que presenta una asociación suficientemente fuerte con el tratamiento (estadístico $F = 1.226,47$), una monotonicidad fuerte y una variable de control para garantizar que el instrumento solo está asociado al resultado a través del tratamiento—, el efecto causal estimado se interpreta como una media ponderada del LATE para cada área de conocimiento, siendo el LATE para cada área el promedio del efecto del tratamiento para el subconjunto de proyectos que hace uso de una ICTS si y solo si el grupo de investigación tiene ventaja para acceder a una ICTS. Es un subconjunto de proyectos interesante de estudiar porque, a priori, serán proyectos de una calidad científico-técnica media, suficientemente buena como para recibir financiación de la Agencia Estatal de Investigación, pero no lo suficientemente excelente para tener una nota¹³ o una financiación que les incentive a acceder a la ICTS en concurrencia competitiva con otros proyectos. Por tanto, el posible impacto del uso de una ICTS en sus resultados será una estimación conservadora del posible impacto en toda la población de proyectos de investigación financiados por la Agencia Estatal de Investigación.

El Anexo 3 detalla el análisis realizado para estimar el efecto causal.

proyecto, o bien es ejecutado por un grupo con ventaja para acceder a una ICTS, o bien por un grupo que no la tiene, solo uno de los tratamientos potenciales es observado.

¹¹ También conocido por su denominación y sus siglas en inglés, *Average Treatment Effect*, ATE.

¹² Siguiendo las recomendaciones de Angrist y Pischke, no se ha tenido en cuenta la no linealidad asociada al tratamiento binario y a algunos resultados binarios (véase <https://www.mostlyharmlesseconometrics.com/2009/07/is-2sls-really-ok/>).

¹³ Véase la Nota 5.

D) Análisis de la sensibilidad de las estimaciones a violaciones de los supuestos

Cuando se aplica la técnica de variables instrumentales a datos observacionales, es probable que se violen los supuestos mencionados en el apartado 2.4.B), en particular, los supuestos de selección en observables (el resultado y el instrumento no comparten causas no observadas) y de exclusión (el instrumento solo afecta al resultado a través del tratamiento). **Los análisis de sensibilidad permiten evaluar en qué medida deben violarse los supuestos para cambiar las conclusiones respecto al efecto del tratamiento.** En la presente evaluación se ha empleado el análisis de sensibilidad propuesto en [4], que evalúa la sensibilidad en términos de cuánta varianza residual del instrumento y del resultado pueda ser predicha por las posibles variables omitidas en la estimación por mínimos cuadrados en dos etapas sin que varíen las conclusiones respecto al efecto del tratamiento.

El Anexo 4 recoge los resultados y la interpretación del análisis de sensibilidad realizado.

2.5. Dificultades y limitaciones de la evaluación

La principal **dificultad** ha estado relacionada con la ausencia en bases de datos de información suficiente sobre el uso de grandes infraestructuras de investigación. La Agencia Estatal de Investigación solo registra si el proyecto, en la fase de solicitud de financiación, planea hacer uso de infraestructuras pertenecientes al Mapa de ICTS, pero, al finalizar los proyectos, no vuelve a recabar información específica. Ha sido necesario examinar los informes finales de justificación científico-técnica de todos los proyectos que planeaban hacer uso de una ICTS para determinar si el proyecto finalmente hizo uso o no de una ICTS para sus actividades de investigación.

Dicho seguimiento insuficiente del uso de grandes infraestructuras de investigación es el responsable de una serie de **limitaciones** de la evaluación:

- La revisión de los informes finales para la construcción de la variable de tratamiento binaria “Uso de ICTS” es un proceso que inevitablemente ha introducido ruido en los datos, pues no siempre se desprende claramente del informe del final si se accedió o no a una gran infraestructura de investigación.
- La información insuficiente recogida en los informes finales ha obligado a definir el tratamiento de manera demasiado amplia o poco concreta, lo que compromete el cumplimiento del supuesto de consistencia y puede inducir un sesgo positivo en la magnitud de las estimaciones mediante variables instrumentales [5, pág. 14].
- Para la construcción de la variable instrumento binaria “Grupo con acceso ICTS”¹⁴ solo se ha dispuesto de información relativa a 2021. Esto también ha podido introducir ruido en los datos, pues se han utilizado estos valores de la variable instrumento para todos los proyectos de la muestra (proyectos solicitados desde el 1 de enero de 2014 y que en enero de 2023 habían finalizado y presentado el informe final de justificación científico-técnica).
- La Agencia Estatal de Investigación no registra información relativa a si el proyecto planea acceder a otras grandes infraestructuras de investigación europeas o internacionales. Dado su volumen, no se analizaron los informes finales de ejecución de los proyectos que inicialmente declararon que no planificaban acceder a una ICTS, por lo que puede darse el caso de que algunos accedieran a una ESFRI o a otra gran infraestructura de investigación, o incluso a una ICTS. Por esta razón hay que tener en cuenta que los posibles impactos hallados

¹⁴ Como se ha explicado anteriormente, representa si el grupo de investigación que lleva a cabo el proyecto tiene algún tipo de ventaja para acceder a alguna ICTS, por ejemplo, porque pertenece a la institución propietaria de la misma o a otra con algún tipo de acuerdo de colaboración para el acceso.

por la evaluación pueden ser menores a los reales, ya que en el grupo de control puede haber proyectos que sí hicieron uso de una gran infraestructura de investigación.

- La Agencia Estatal de Investigación registra información relativa a si el proyecto planea acceder a infraestructuras del mapa de ICTS solo para determinadas convocatorias, básicamente para todas las convocatorias de proyectos de I+D dirigidas a organismos públicos de investigación, universidades, centros de investigación, centros tecnológicos, institutos de investigación y demás instituciones de I+D+I públicas o privadas sin ánimo de lucro. Para las convocatorias de colaboración pública-privada, no se registra ninguna información relativa a grandes infraestructuras de investigación. Tampoco el Centro para el Desarrollo Tecnológico Industrial (CDTI), que es la Agencia Estatal de Innovación, hace un seguimiento específico. Por tanto, la evaluación no ha podido analizar el impacto del uso de grandes infraestructuras de investigación sobre los resultados de proyectos de investigación desarrollados por empresas.

Por último, las características de los datos también han impuesto limitaciones:

- Es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a tesis doctorales, patentes y empresas de base tecnológica generadas por los proyectos, debido a que el tiempo que requiere producir estos resultados¹⁵ hace que, en muchos casos, se alcancen después del momento en que se registran los indicadores de resultado de los proyectos (normalmente, tres meses después de su finalización).
- La violación estructural del principio de positividad imposibilita emplear estimadores que hacen uso del *propensity score*¹⁶ e impide que sean confiables las estimaciones realizadas mediante estandarización o regresión lineal múltiple pues, en ausencia de positividad, están basadas en gran medida en la extrapolación.
- La presencia de números valores atípicos en todas las variables de resultado compromete la precisión de las estimaciones mediante variables instrumentales, dada la sensibilidad de la técnica a la presencia de atípicos.



Figura 4. ICTS Observatorios de Canarias. Está compuesta por el Observatorio del Roque de los Muchachos, en la isla de La Palma, y el Observatorio del Teide, en Tenerife. Alberga telescopios e instrumentos pertenecientes a instituciones de casi 40 países, constituyendo el grupo de instalaciones más importante de la UE para astrofísica nocturna, solar, visible e infrarroja. Fuente: Ministerio de Ciencia e Innovación.

¹⁵ Respecto a las tesis, según datos del Ministerio de Educación, en el periodo 2015-2021 el número medio de años desde el inicio de la formación investigadora hasta la lectura de la tesis doctoral fue de 4,9 años. Respecto a las patentes, el tiempo medio de concesión en la Oficina Española de Patentes y Marcas es de 2 años y, en la Oficina Europea de Patentes, de 3 a 5 años.

¹⁶ El *propensity score* es la probabilidad que tiene un proyecto de recibir el tratamiento. En nuestra muestra, todos los proyectos que inicialmente no planificaban acceder a una ICTS tienen un *propensity score* nulo, lo que impide el uso de técnicas como el *propensity score matching* o de estimadores que ponderan por la inversa del *propensity score*.

3. Los hallazgos de la evaluación

Para los proyectos de investigación financiados por las convocatorias de la Agencia Estatal de Investigación, **se estima que acceder a una gran infraestructura de investigación para llevar a cabo determinadas actividades del proyecto tiene un efecto positivo en una serie de resultados relacionados con la calidad de la investigación, la difusión e influencia de sus resultados y el impulso de la competitividad y el crecimiento de la UE.** Así, se han estimado efectos positivos sobre:

- el número de publicaciones en revistas científicas de alta calidad e influencia,
- el número de publicaciones en acceso abierto y
- la obtención de financiación en el altamente competitivo Programa Marco de Investigación e Innovación de la UE.

No se ha detectado que hacer uso de una gran infraestructura de investigación tengo efectos ni positivos ni negativos sobre el resto de los resultados analizados, esto es, publicaciones en revistas no indexadas, tesis doctorales, patentes, empresas de base tecnológica creadas y acuerdos de colaboración. De todos modos, como ya se ha señalado, es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a tesis doctorales, patentes y empresas de base tecnológica generadas por los proyectos, debido a que el tiempo que requiere producir estos resultados hace que, en muchos casos, se alcancen después del momento en que se registran los indicadores de resultado de los proyectos (normalmente, tres meses después de su finalización).



Figura 5. Buque Hespérides, perteneciente a la ICTS distribuida Flota Oceanográfica Española. La Flota está compuesta por diez buques oceanográficos que poseen instrumental para realizar tareas de investigación científica en mares y océanos. Tienen su sede en Galicia, Illes Balears y Región de Murcia Fuente: Ministerio de Ciencia e Innovación.



Figura 6. ICTS Reserva Biológica de Doñana, en Andalucía. Cuenta con casi 10.000 hectáreas que acogen cuatro ecosistemas (playa, dunas, monte mediterráneo y marisma), constituyendo una plataforma ideal para la experimentación en campo y el intercambio de experiencias asociadas al impacto ambiental del cambio global. Fuente: Ministerio de Ciencia e Innovación.

3.1. Impacto en publicaciones en revistas científicas de alta calidad e influencia

A) ¿Qué son las publicaciones en revistas indexadas y por qué son relevantes?

Las denominadas revistas indexadas son aquellas que están indexadas en determinados portales de información que se dedican a medir el impacto de lo que se publica. Los portales más conocidos a nivel internacional son *Journal Citation Reports (JCR)* y *SCIMAGO Journal and Country Rank (SJR)*. Ambos portales analizan revistas de literatura científica cuyo contenido es revisado por personas expertas con carácter previo a su publicación (la denominada “revisión por pares”) y, para cada revista y año, establecen un factor de impacto en función del número de veces que sus artículos son referenciados por otros artículos científicos. Dichos factores de impacto permiten establecer, por áreas temáticas, rankings de revistas. Los rankings se dividen en cuartiles, denominándose Q1 al grupo que ocupa el primer cuartil. Las revistas Q1 conforman la élite en su temática; las revistas que ocupan los siguientes cuartiles tienen una repercusión decreciente. Por tanto, **publicar un resultado científico en una revista indexada es un indicio de la calidad y relevancia de la investigación, especialmente si la revista se sitúa en el primer cuartil de su ranking temático.**

Las publicaciones en revistas indexadas son relevantes porque facilitan la difusión de los resultados de investigación, estimulando el avance del conocimiento científico y sustentando las innovaciones; además, los hallazgos científicos divulgados por las publicaciones pueden tener impactos económicos durante periodos de tiempo prolongados. Así lo señala un reciente estudio del Fondo Monetario Internacional (FMI) [6, pág. 65], del cual se ha extraído la Figura 7 relativa a los documentos que las patentes referencian como antecedentes de sus invenciones¹⁷: como se observa, a partir de mediados de los años 80, el número medio de referencias a artículos científicos (línea azul) supera al número medio de referencias a otras patentes (línea roja), lo que refleja que **las innovaciones se sustentan en gran medida sobre hallazgos científicos difundidos por publicaciones académicas.**

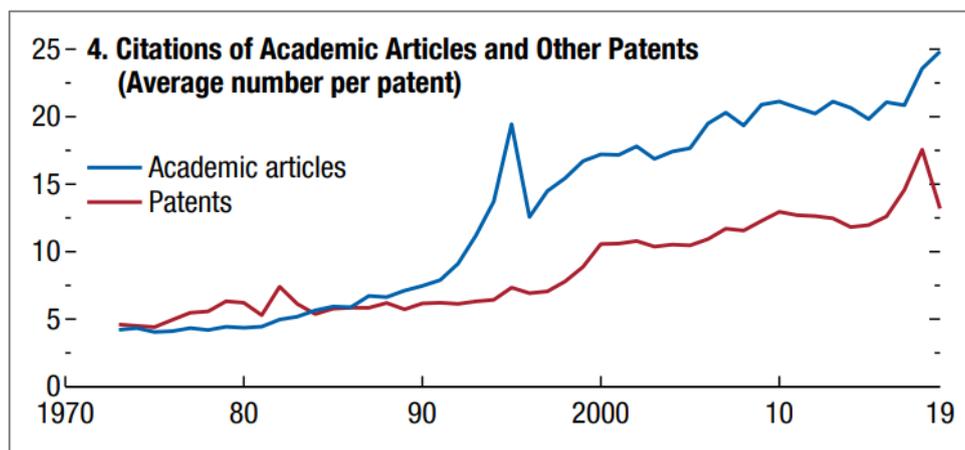


Figura 7. Evolución del número medio de artículos científicos y de patentes citados por documentos de patente. Fuente: FMI [6, figura 3.1].

¹⁷ La normativa armonizada a nivel mundial sobre propiedad industrial exige que los documentos de patente recojan los antecedentes científico-técnicos sobre los que se sustenta la invención que reivindican. Para cumplir con este requisito, normalmente los documentos de patente citan otras patentes y artículos científicos.

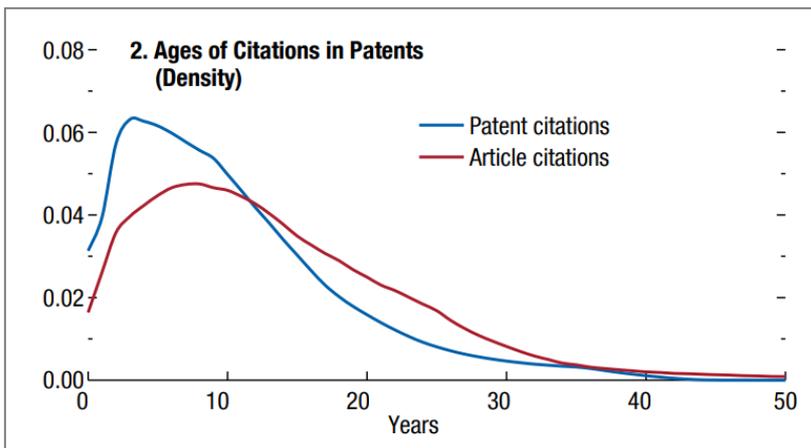


Figura 8. Antigüedad de las referencias citadas en los documentos de patente publicados en el periodo 2010-2019. Fuente: FMI [6, figura 3.4].

roja) tiene una antigüedad en torno a los ocho años, frente a los tres para las patentes referenciadas (línea azul). De acuerdo con el FMI, esta evidencia sugiere que **los hallazgos científicos difundidos por publicaciones académicas pueden tener impactos económicos durante largos periodo de tiempo** [6, pág. 69].

Además, el conocimiento divulgado por los artículos científicos presenta una influencia más duradera en el tiempo que el conocimiento divulgado por las patentes. La Figura 8 recoge la densidad de la antigüedad de las referencias citadas en los documentos de patente: la mayor parte de los artículos científicos referenciados (línea

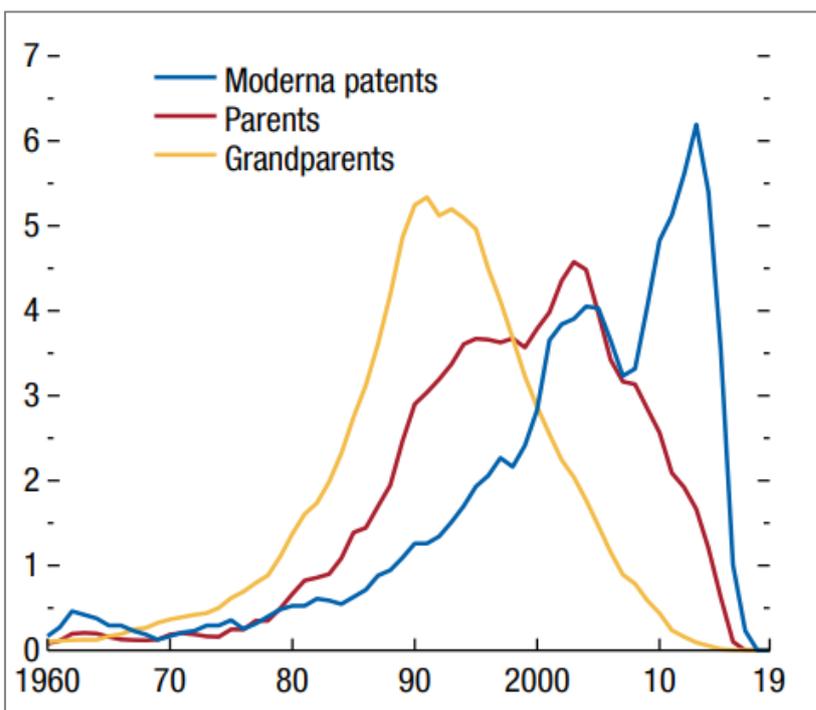


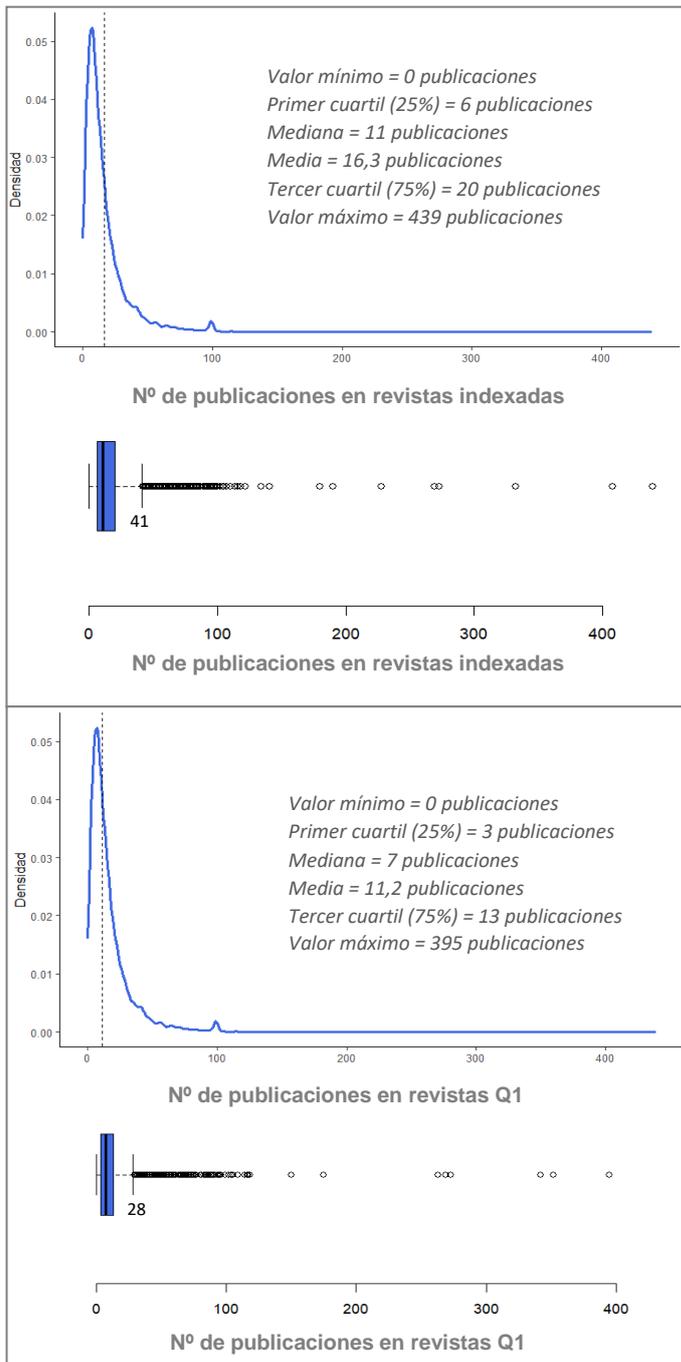
Figura 9. Artículos científicos referenciados por las patentes de Moderna y sus ancestros por año de publicación. Fuente: FMI [6, figura 3.1.1].

enorme: probablemente ha acertado la pandemia en años, y parece que va a revolucionar los tratamientos médicos en los próximos años. Esta tecnología está basada en sucesivas oleadas de descubrimientos científicos anteriores divulgados por artículos científicos, como muestra la Figura 9 para el caso de la vacuna de Moderna. La distribución de las fechas de publicación de los artículos científicos citados por las patentes de la vacuna de Moderna (línea azul), captura la

El informe del FMI recoge diversos casos de estudio para ilustrar cómo los avances científicos difundidos por publicaciones científicas influyen en los pilares de las tecnologías futuras y generan beneficios económicos durante prolongados periodos de tiempo. Uno de estos casos estudia el desarrollo de las vacunas contra el COVID-19 que utilizan la nueva tecnología de ARN mensajero (ARNm). El impacto social y económico de la tecnología de ARNm ha sido

dependencia directa de descubrimientos anteriores, principalmente de los avances en la comprensión de la función del ARNm de comienzos de la década de 2010. Por su parte, la dependencia indirecta de descubrimientos anteriores es capturada por las distribuciones de las fechas de publicación de los artículos citados por las patentes “madres” (patentes referenciadas por las patentes originales) y las patentes “abuelas” (referenciadas por las patentes “madres”): la primera (línea roja), realiza un seguimiento de los descubrimientos en el ámbito de la edición de códigos genéticos, siendo los de especial influencia los de comienzos de la década de 2000; la segunda (línea amarilla) hace un seguimiento de los avances anteriores en la lectura de códigos genéticos, siendo los de mayor influencia los de comienzos de la década de 1990.

B) ¿Cuántas publicaciones en revistas indexadas generan los proyectos de investigación financiados por la Agencia Estatal de Investigación?



El 50% de los proyectos de investigación analizados produjeron entre 6 y 20 publicaciones en revistas indexadas, con una mediana de 11 publicaciones; como se observa en los gráficos de la Figura 10, la distribución presenta numerosos valores atípicos: casi el 7% de los proyectos han generado más de 41 publicaciones, con un valor máximo de 439.

Figura 10. Gráficos de densidad y de caja del número de publicaciones en revistas indexadas generados por los proyectos de investigación de la muestra analizada.

Respecto a las publicaciones en revistas indexadas pertenecientes al primer cuartil, el 50% de los proyectos de investigación analizados produjeron entre 3 y 13 publicaciones, con una mediana de 7 publicaciones; como refleja la Figura 11, también presenta numerosos valores atípicos: casi el 7% de los proyectos han generado más de 28 publicaciones, con un valor máximo de 395 publicaciones.

Figura 11. Gráficos de densidad y de caja del número de publicaciones en revistas Q1 generados por los proyectos de investigación de la muestra analizada.

C) ¿El uso de grandes infraestructuras de investigación impacta en el número de publicaciones en revistas indexadas que generan los proyectos?

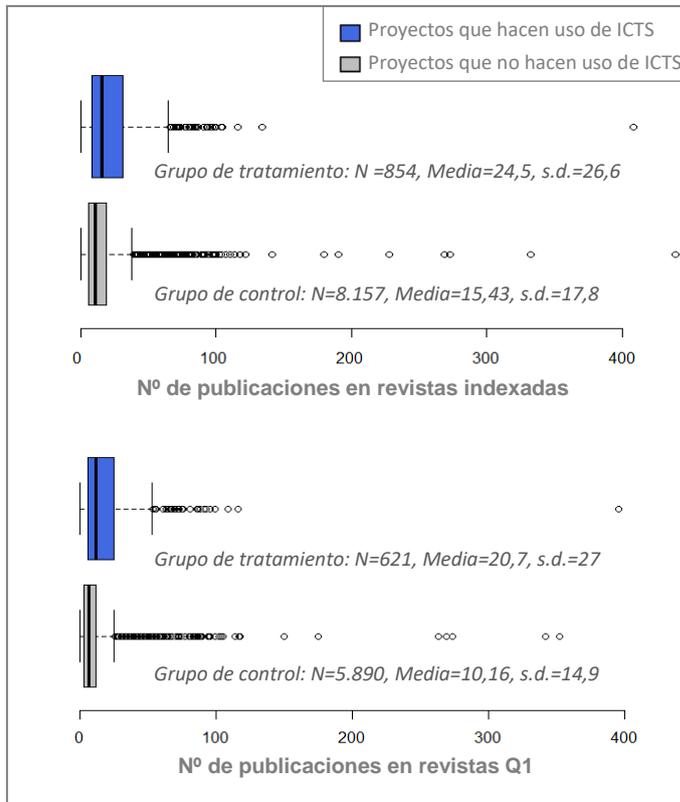


Figura 12. Gráficos de caja del número de publicaciones en revistas indexadas para los grupos de tratamiento y control.

una calidad científico-técnica media¹⁹, usar una ICTS para llevar a cabo actividades de investigación tiene un efecto positivo sobre el número de publicaciones en revistas indexadas y, en particular, sobre el número de publicaciones en revistas indexadas Q1. En cuanto a la cuantificación de dichos efectos, el uso de las ICTS está asociado con un incremento de 5,1 publicaciones en revistas indexadas (intervalo de confianza al 95% de 1,4 – 8,9 publicaciones) y con un incremento de 5,6 publicaciones en revistas indexadas que pertenecen al primer cuartil (con un intervalo de confianza al 95% de 1,5 – 9,6 publicaciones). Esto es, el efecto estimado es mayor en el caso de publicaciones Q1. De todos modos, como se explica en el Anexo 3, es probable que la magnitud de las estimaciones esté sesgada hacia arriba debido a la definición demasiado amplia de la variable de tratamiento; además, la presencia de numerosos valores atípicos en ambas variables también afecta a la exactitud de las estimaciones. En cualquier caso, aunque las estimaciones puedan ser imprecisas, sí que está constatada la robustez de los hallazgos respecto a los efectos positivos sobre las publicaciones.

De acuerdo con lo expuesto en el apartado 3.1.A), **los efectos positivos estimados sugieren que el uso de grandes infraestructuras de investigación impacta positivamente en la calidad de los resultados de los proyectos de investigación y en su capacidad de influencia sobre futuras investigaciones e innovaciones.**

¹⁸ La muestra empleada para las publicaciones en revistas indexadas Q1 (N = 6.511 proyectos) es menor debido a que existen numerosos valores perdidos en los datos extraídos de la base de datos de la Agencia Estatal de Investigación.

¹⁹ Véase explicación del apartado 2.4.C) sobre el subconjunto de proyectos al que aplica el efecto causal estimado.



Figura 13. La ICTS distribuida Bases Antárticas Españolas tiene dos bases en la Antártida, las cuales permanecen operativas durante el verano austral. Están situadas en las islas Shetland del Sur: la base Juan Carlos I en la isla de isla Livingston y la base Gabriel de Castilla en la isla Decepción. En estas bases se realizan valiosos estudios científicos, tanto biológicos, como geológicos, climatológicos, etc. Fuente: Ministerio de Ciencia e Innovación.

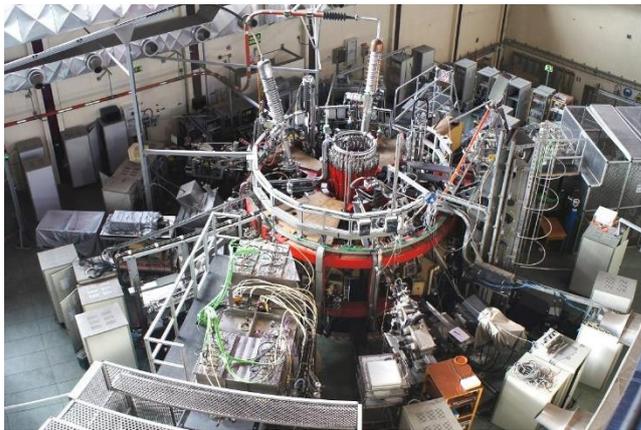


Figura 14. ICTS Laboratorio Nacional de Fusión. Radicado en Madrid, está dedicado a la investigación y desarrollo de tecnologías y materiales para construir y operar los reactores de fusión del futuro. Su principal instalación es el TJ-II, un dispositivo de confinamiento magnético de tipo stellarator; cuenta además con instrumentación para la caracterización y modificación por radiación de materiales. Fuente: Ministerio de Ciencia e Innovación.



Figura 15. ICTS Plataforma Solar de Almería, en Andalucía. Está dedicada a la experimentación de sistemas de energía solar. Dispone de grandes instalaciones de ensayo, como sistemas de receptor central, campos de captadores de cilindros parabólicos, hornos solares y plantas piloto de desalación, fotoquímica y fotocatalisis. Fuente: Ministerio de Ciencia e Innovación.



Figura 16. La ICTS distribuida Red de Salas Blancas de Micro y Nano Fabricación ofrece servicios en los campos de la micro y nano fabricación y la fotónica, tales como técnicas de procesado, crecimiento de muestras, encapsulado, integración y caracterización de dispositivos. Consta de tres infraestructuras radicadas en Cataluña, Comunitat Valenciana y Madrid. Fuente: Ministerio de Ciencia e Innovación.

3.2. Impacto en publicaciones en acceso abierto

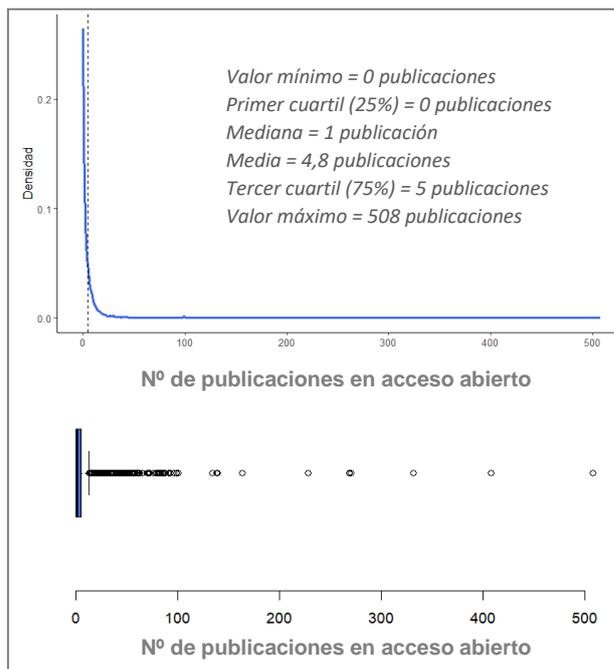
A) ¿Qué son las publicaciones en acceso abierto y por qué son relevantes?

El acceso abierto (*open access*) se refiere a información en línea de libre acceso. **Las publicaciones científico-técnicas en acceso abierto son gratuitas y sus términos de copyright y licencias de uso son menos restrictivos que los de las publicaciones tradicionales.** El acceso abierto puede materializarse a través de diferentes vías, que se pueden agrupar en dos:

- La vía verde, que implica el depósito en repositorios de libre acceso. Estos repositorios pueden tener carácter institucional, como DIGITAL.CSIC, temático, como arXiv, o generalista, como Zenodo.
- La vía dorada, que implica la publicación en acceso abierto en revistas revisadas por pares; dichas revistas pueden tener todos sus artículos en abierto o bien ser híbridas, con una mezcla de artículos en abierto y artículos de acceso restringido. Al no mediar pago para acceder al artículo, las revistas cubren los gastos operativos asociados a su publicación mediante otras fuentes, como subvenciones, apoyo en especie, o cuotas que han de asumir las propias personas autoras o sus instituciones o agencias financiadoras (denominadas comúnmente APC - *Article Processing Charges*).

La indexación automatizada y masiva por parte de un amplio abanico de motores de búsqueda gratuitos, portales temáticos y agregadores de resultados de investigación en acceso abierto, hace que el acceso abierto ejerza un efecto multiplicador sobre la diseminación de los resultados de investigación [7]. De acuerdo con las conclusiones del Consejo de la UE sobre la evaluación de la investigación y la aplicación de la ciencia abierta [8], **al aumentar exponencialmente la visibilidad y la accesibilidad de los contenidos de las publicaciones científico-técnicas, el acceso abierto juega un papel crucial en impulsar el impacto, la calidad, la eficiencia, la transparencia y la ética de la investigación y la innovación.**

B) ¿Cuántas publicaciones en acceso abierto generan los proyectos de investigación financiados por la Agencia Estatal de Investigación?



Como se observa en los gráficos de la Figura 17, el 25% de los proyectos de investigación analizados no generó ninguna publicación en abierto, el 50% generó entre 1 y 5 publicaciones y el 25% restante generó más de 5, con una distribución muy dispersa, con numerosos valores atípicos.

Figura 17. Gráficos de densidad y de caja del número de publicaciones en acceso abierto generados por los proyectos de investigación de la muestra analizada.

C) ¿El uso de grandes infraestructuras de investigación impacta en el número de publicaciones en acceso abierto que generan los proyectos?

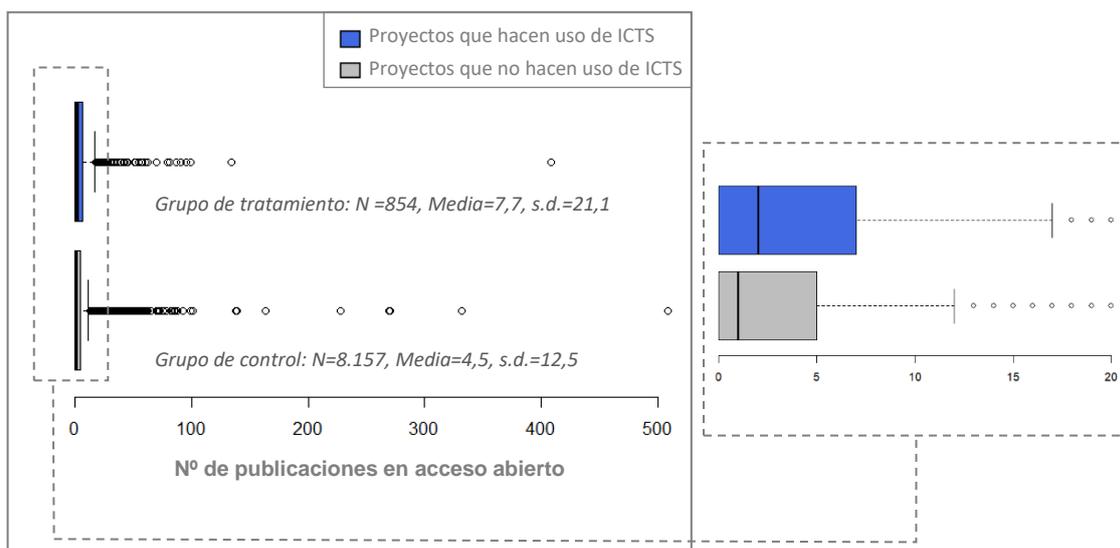


Figura 18. Gráficos de caja del número de publicaciones en acceso abierto para los grupos de tratamiento y control.

Si se agrupan las publicaciones en acceso abierto en función del uso o no de grandes infraestructuras de investigación, se observa que ambos grupos presentan distintas distribuciones, tal y como refleja la Figura 18. El grupo de proyectos que ha hecho uso de ICTS para actividades de investigación (grupo de tratamiento, arriba en azul) presenta mayores cuartiles segundo, tercero y cuarto, mayor media y mayor desviación típica.

Como se está trabajando con datos observacionales, las diferencias encontradas no tienen por qué responder a una asociación causal entre hacer uso de una ICTS para actividades de investigación y el número de publicaciones en acceso abierto. Para evaluar si existe dicha relación causal se ha llevado a cabo el análisis de impacto detallado en el Anexo 3. Se ha hallado que, **para proyectos de una calidad científico-técnica media²⁰, usar una ICTS para llevar a cabo actividades de investigación tiene un efecto estimado positivo sobre el número de publicaciones en acceso abierto.** En cuanto a la cuantificación de dicho efecto, el uso de las ICTS está asociado con un incremento de 2,5 publicaciones en acceso abierto (con un intervalo de confianza al 95% de 0,8 – 4,3 publicaciones en acceso abierto). De todos modos, como se explica en el Anexo 3, es probable que la magnitud de la estimación esté sesgada hacia arriba debido a la definición demasiado amplia de la variable de tratamiento; además, la presencia de numerosos valores atípicos también afecta a la exactitud de las estimación. En cualquier caso, aunque la estimación pueda ser imprecisa, sí que está constatada la robustez del hallazgo respecto al efecto positivo sobre las publicaciones en acceso abierto.

De acuerdo con lo expuesto en el apartado 3.2.A), **este hallazgo sugiere que el uso de grandes infraestructuras de investigación impacta positivamente sobre la difusión, la calidad, la eficiencia, la transparencia y la ética de los resultados de los proyectos de investigación.**

²⁰ Véase explicación del apartado 2.4.C) sobre el subconjunto de proyectos al que aplica el efecto causal estimado.



Figura 19. La ICTS Red Académica y de Investigación Española (RedIRIS) es una red de comunicaciones avanzadas a la que se conectan, bien directamente o mediante redes autonómicas, más de 500 instituciones académicas y científicas españolas. RedIRIS opera en sincronía con otras redes, como GÉANT, que agrupa en Europa las principales redes académicas y de investigación nacionales. Fuente: Ministerio de Ciencia e Innovación.



Figura 20. Supercomputador MareNostrum, ubicado en Cataluña. Pertenece a la ICTS Red Española de Supercomputación y Datos (RES). La RES es una infraestructura de supercomputadores ubicados en catorce localizaciones cuyo objetivo es proporcionar a la comunidad científica servicios de supercomputación y de almacenamiento masivo de datos. Fuente: Ministerio de Ciencia e Innovación.



Figura 21. La ICTS Red Distribuida de Imagen Biomédica da servicio en el campo de la imagen molecular y funcional, así como en imagen avanzada y de alto rendimiento. Consta de cuatro infraestructuras radicadas en la Comunitat Valenciana, Madrid y País Vasco. Fuente: Ministerio de Ciencia e Innovación.



Figura 22. Gran Tanque de Ingeniería Marítima de Cantabria. Junto con otras cuatro instalaciones radicadas en Canarias, Cataluña, Madrid y País Vasco, forma parte de la ICTS distribuida MARHIS (Maritime Aggregated Hydraulic Infrastructures), que ofrece instalaciones numéricas, experimentales y de campo en ingeniería marítima. Fuente: Ministerio de Ciencia e Innovación.

3.3. Impacto en obtención de financiación en el Programa Marco de Investigación e Innovación de la Unión Europea

A) ¿Qué es el Programa Marco de Investigación e Innovación de la Unión Europea?

El Programa Marco es la principal iniciativa de la UE de apoyo a la investigación y a la innovación. Tiene como objetivos últimos mejorar el bienestar de la ciudadanía, hacer frente a desafíos globales, como el cambio climático, e impulsar la competitividad industrial europea. Para ello, financia la creación y la difusión de conocimientos y tecnologías de gran calidad, moviliza y atrae talento científico-tecnológico, y fomenta la colaboración científico-técnica entre los países y regiones de la UE y también con terceros países.

Las propuestas presentadas para ser financiadas por el Programa Marco se evalúan basándose en tres criterios: la excelencia del concepto, de los objetivos y del enfoque; el impacto de los resultados potenciales; y la calidad y la eficiencia del plan de trabajo. La mayor parte de las propuestas se desarrollan en consorcio, siendo los socios entidades de distintos países. Dada la alta competencia por los fondos del Programa, solo el 12,2% de las propuestas presentadas en el periodo 2014-2020 obtuvieron financiación. Por tanto, **obtener financiación en el Programa Marco es un indicador de que la propuesta de investigación es excelente en su concepción, viable en su implementación y relevante por sus potenciales impactos científicos, económicos y sociales.**

B) ¿Cuántos proyectos de investigación financiados por la Agencia Estatal de Investigación han dado lugar a proyectos financiados por el Programa Marco?

Como se observa en el gráfico de la Figura 23, el 21% de los proyectos analizados ha dado lugar a proyectos relacionados que han recibido financiación del Programa Marco de Investigación e Innovación de la UE.

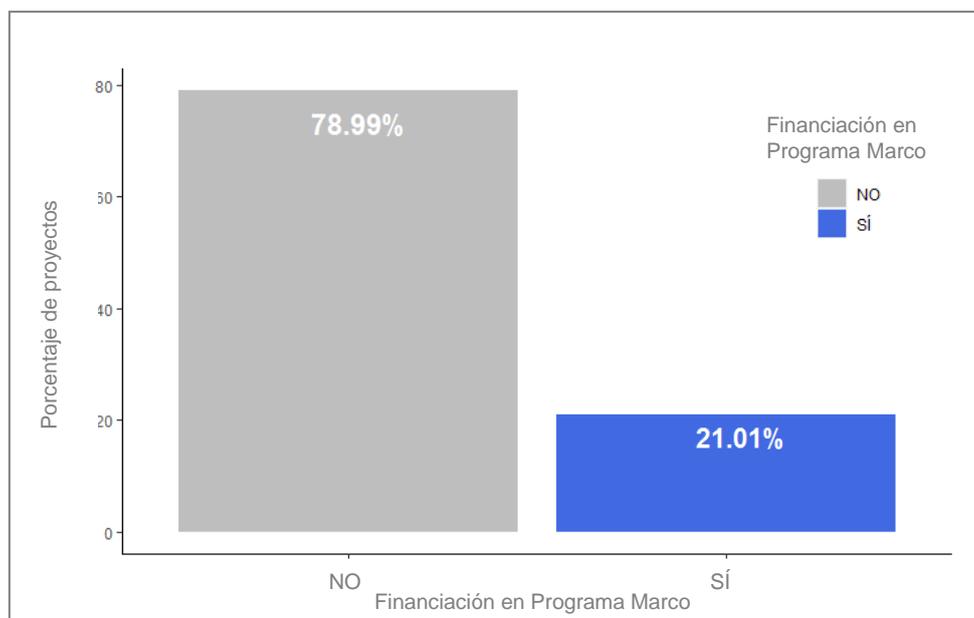


Figura 23. Gráfico de barras del porcentaje de proyectos que ha generado propuestas financiadas por el Programa Marco.

C) ¿El uso de grandes infraestructuras de investigación impacta en la generación de proyectos de investigación financiados por el Programa Marco?

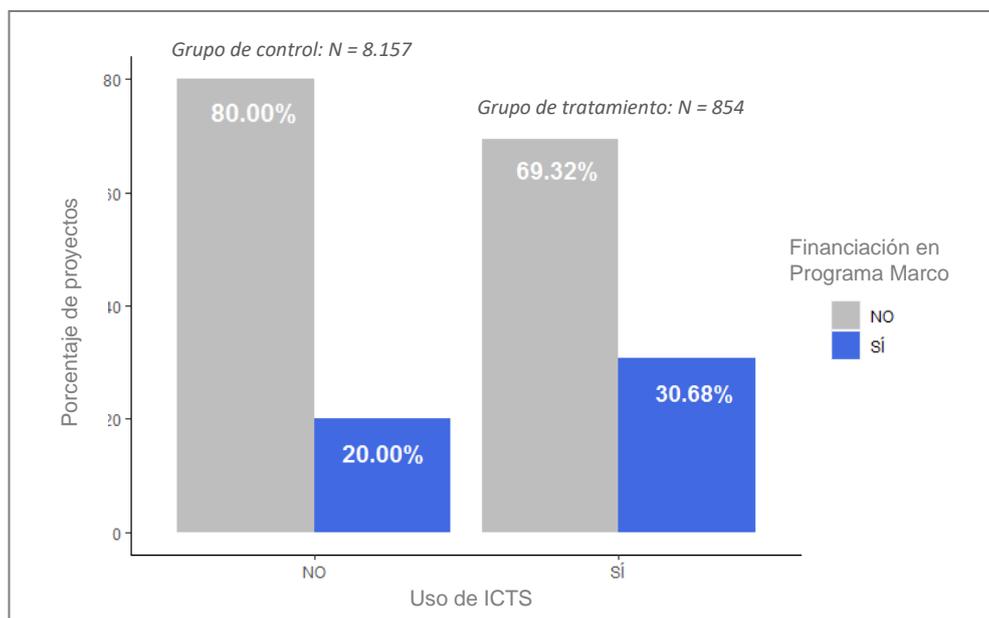


Figura 24. Gráficos de barras del porcentaje de proyectos que ha generado propuestas financiadas por el Programa Marco para los grupos de tratamiento y control.

El grupo que ha hecho uso de grandes infraestructuras de investigación presenta un porcentaje mayor de proyectos que ha dado lugar a propuestas financiadas por el Programa Marco (30,68% vs. 20%), tal y como refleja la Figura 24.

Como se está trabajando con datos observacionales, las diferencias observadas no tienen por qué responder a una asociación causal entre hacer uso de una ICTS para actividades de investigación y la obtención de financiación en el Programa Marco. Para evaluar si existe dicha relación causal se ha llevado a cabo el análisis de impacto detallado en el Anexo 3. Se ha hallado que, **para proyectos de una calidad científico-técnica media²¹, usar una ICTS para llevar a cabo actividades de investigación tiene un efecto estimado positivo sobre la probabilidad de dar lugar a un proyecto relacionado que obtenga financiación en el Programa Marco.** En cuanto a la cuantificación de dicho efecto, se estima que el uso de las ICTS incrementa en 17 puntos porcentuales la probabilidad de dar lugar a un proyecto financiado por el altamente competitivo Programa Marco (con un intervalo de confianza al 95% de 8 - 25 puntos porcentuales). De todos modos, como se explica en el Anexo 3, es probable que la magnitud de la estimación esté sesgada hacia arriba debido a la definición demasiado amplia de la variable de tratamiento; además, la presencia de numerosos valores atípicos también afecta a la exactitud de las estimación. En cualquier caso, aunque la estimación pueda ser imprecisa, sí que está constatada la robustez del hallazgo respecto al efecto positivo sobre la generación de proyectos de investigación financiados por el Programa Marco.

De acuerdo con lo expuesto en el apartado 3.3.A), **este hallazgo sugiere que el uso de grandes infraestructuras de investigación impacta positivamente en la calidad de los resultados de los proyectos de investigación y en su capacidad de contribuir al bienestar de la ciudadanía europea y a la competitividad de los Estados miembros.**

²¹ Véase explicación del apartado 2.4.C) sobre el subconjunto de proyectos al que aplica el efecto causal estimado.



Figura 25. La ICTS NANBIOSIS proporciona soluciones integradas a medida para los desafíos que encuentra la investigación en nanomedicina, diagnóstico médico e ingeniería de tejidos y dispositivos de medicina regenerativa. Está constituida por 27 unidades complementarias y coordinadas, ubicadas en diferentes centros de Andalucía, Aragón, Cataluña, Extremadura, Madrid y País Vasco. Fuente: Ministerio de Ciencia e Innovación.



Figura 26. La ICTS Laboratorio Subterráneo de Canfranc, en Aragón, está dedicada a la física de astropartículas y a la investigación en física subterránea. Por su extensión y características, es el segundo laboratorio subterráneo europeo tras el laboratorio del Gran Sasso, en Italia. Fuente: Ministerio de Ciencia e Innovación.



Figura 27. La ICTS Centro Nacional de Investigación sobre la Evolución Humana (CENIEH), ubicada en Castilla y León, gestiona y conserva las colecciones paleontológicas y arqueológicas procedentes de las excavaciones de Atapuerca y de otros yacimientos similares. Posee laboratorios de sedimentología y caracterización de materiales, geocronología, y conservación y restauración. Fuente: Ministerio de Ciencia e Innovación.



Figura 28. La ICTS Observatorio de Yebes, ubicada en Castilla-La Mancha, está dedicada al desarrollo y construcción de instrumentación en el campo de la radioastronomía, así como a la realización de observaciones astronómicas tanto de interés astronómico como geodésico o geofísico. Su radiotelescopio es uno de los nodos más importantes de la Red Europea de Interferometría de Muy Larga Base. Fuente: Ministerio de Ciencia e Innovación.

4. Conclusiones y recomendaciones

De acuerdo con el Consejo de la UE, las grandes infraestructuras de investigación son necesarias para que la comunidad científica lleve a cabo investigación de excelencia y desarrolle tecnologías de vanguardia que alimentan innovaciones, posibilitando a la UE ser más competitiva a escala global. Empleando datos observacionales asociados a proyectos de investigación que han accedido a grandes infraestructuras pertenecientes al Mapa de Infraestructuras Científicas y Técnicas Singulares (ICTS), **la presente evaluación ha encontrado evidencias de que efectivamente el uso de grandes infraestructuras de investigación impacta positivamente sobre la calidad de los resultados de los proyectos de investigación, su capacidad de influencia sobre futuras investigaciones e innovaciones y su capacidad de contribuir al bienestar de la ciudadanía europea y a la competitividad de los Estados miembros.**

- El análisis de impacto realizado estima que, para proyectos de una calidad científico-técnica media, usar una ICTS para llevar a cabo actividades de investigación tiene un efecto positivo sobre el número de publicaciones en revistas indexadas y, en particular, sobre el número de publicaciones en revistas indexadas que pertenecen al primer cuartil.

Publicar un resultado científico en una revista indexada es un indicador de la calidad y relevancia de la investigación, especialmente si la revista se sitúa en el primer cuartil de su ranking temático. Asimismo, la publicación en revistas indexadas facilita la difusión de los resultados de investigación, lo que estimula el avance del conocimiento científico y puede sustentar, durante prolongados periodos de tiempo, numerosas e inesperadas innovaciones. Así lo resalta un reciente estudio del FMI [6], que pone como ejemplo el extraordinariamente rápido desarrollo de las vacunas contra el COVID-19, basadas en décadas de investigación científica previa difundida por artículos publicados en revistas científicas.

Por tanto, el efecto estimado sugiere que el uso de grandes infraestructuras de investigación impacta positivamente sobre la calidad de los resultados de los proyectos de investigación y sobre su capacidad de influencia sobre futuras investigaciones e innovaciones.

- Asimismo, el análisis estima que, para proyectos de una calidad científico-técnica media, usar una ICTS para llevar a cabo actividades de investigación tiene un efecto positivo sobre el número de publicaciones en acceso abierto.

El acceso abierto a publicaciones científico-técnicas, al aumentar exponencialmente la visibilidad y la accesibilidad de sus contenidos, juega un papel crucial en impulsar el impacto, la calidad, la eficiencia, la transparencia y la ética de la investigación y la innovación.

Por tanto, este segundo efecto estimado también sugiere que el uso de grandes infraestructuras de investigación impacta positivamente sobre la calidad de los resultados de los proyectos de investigación y su capacidad de influencia sobre futuras investigaciones e innovaciones.

- Por último, el análisis estima que, para proyectos de una calidad científico-técnica media, usar una ICTS para llevar a cabo actividades de investigación tiene un efecto positivo sobre la probabilidad de dar lugar a un proyecto financiado por el altamente competitivo Programa Marco de Investigación e Innovación de la UE.

Obtener financiación en el Programa Marco es un indicador de que la propuesta de investigación es excelente en su concepción, viable en su implementación y relevante por sus potenciales impactos científicos, económicos y sociales.

Por tanto, el tercer efecto estimado sugiere de nuevo que el uso de grandes infraestructuras de investigación impacta positivamente sobre la calidad de los resultados de los proyectos de investigación y sugiere asimismo que impacta positivamente sobre su capacidad de contribuir al bienestar de la ciudadanía europea y a la competitividad de los Estados miembros.

No se ha detectado que hacer uso de una gran infraestructura de investigación tenga efectos ni positivos ni negativos sobre el resto de los resultados analizados, esto es, publicaciones en revistas no indexadas, tesis doctorales, patentes, empresas de base tecnológica creadas y acuerdos de colaboración. De todos modos, hay que señalar que es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a tesis doctorales, patentes y empresas de base tecnológica generadas por los proyectos, debido a que el tiempo que requiere producir estos resultados hace que, en muchos casos, se alcancen después del momento en que se registran los indicadores de resultado de los proyectos (normalmente, tres meses después de su finalización).

Las evidencias encontradas de los efectos positivos sobre la calidad de los resultados de los proyectos de investigación, su capacidad de influencia sobre futuras investigaciones e innovaciones y su capacidad de contribuir al bienestar de la ciudadanía europea y a la competitividad de los Estados miembros, sustentan las siguientes **recomendaciones**:

- En línea con las conclusiones del Consejo de la UE sobre infraestructuras de investigación adoptadas en diciembre de 2022 [2], se recomienda a los distintos niveles administrativos (europeo, estatal y regional) que sigan colaborando para garantizar la sostenibilidad a largo plazo de las infraestructuras de investigación. En particular, dada la relevancia del FEDER en la financiación de las grandes infraestructuras de investigación, se recomienda a la Comisión Europea y a las autoridades españolas que sigan impulsando las sinergias entre las diversas fuentes de financiación europeas, nacionales y regionales.
- Dadas las dificultades y limitaciones afrontadas por la presente evaluación debido a la ausencia en bases de datos de información suficiente sobre el uso de grandes infraestructuras de investigación, se recomienda a las agencias financiadoras estatales y regionales y a las Administraciones titulares de las infraestructuras que implementen mecanismos de seguimiento adecuados que permitan profundizar en el conocimiento y comprensión de los impactos del uso de grandes infraestructuras de investigación sobre las actividades de investigación llevadas a cabo por los sectores público y privado.

5. Referencias

- [1] Consejo de la Unión Europea (2009). *Reglamento (CE) nº 723/2009 del Consejo, de 25 de junio de 2009, relativo al marco jurídico comunitario aplicable a los Consorcios de Infraestructuras de Investigación Europeas (ERIC)*. Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32009R0723>
- [2] Consejo de la Unión Europea (2022). *Council conclusions on Research Infrastructures*. Disponible en: <https://data.consilium.europa.eu/doc/document/ST-15429-2022-INIT/en/pdf>
- [3] Textor, J., van der Zander, B., Gilthorpe, M.S., Liskiewicz, M. y Ellison, GTH. (2016). *Robust causal inference using directed acyclic graphs: the R package 'dagitty'*. *International Journal of Epidemiology*, 45 (6): 1887–1894. Disponible en: <https://doi.org/10.1093/ije/dyw341>
- [4] Cinelli, C. y Hazlett, C. (2022). *An omitted variable bias framework for sensitivity analysis of instrumental variables*. Disponible en: [https://carloscinelli.com/files/Cinelli%20and%20Hazlett%20\(2020\)%20-%20OV%20for%20IV.pdf](https://carloscinelli.com/files/Cinelli%20and%20Hazlett%20(2020)%20-%20OV%20for%20IV.pdf)
- [5] Felton, C. y Stewart, B. M. (2022). *Handle with Care: A Sociologist's Guide to Causal Inference with Instrumental Variables*. Disponible en: <https://doi.org/10.31235/osf.io/3ua7q>
- [6] Fondo Monetario Internacional (2021). *Research and Innovation: Fighting the Pandemic and Boosting Long-Term Growth*. Disponible en: <https://www.imf.org/-/media/Files/Publications/WEO/2021/October/English/ch3.ashx>
- [7] Centro Superior de Investigaciones Científicas (s.f.). *Publicar en acceso abierto*. Disponible en: <https://digital.csic.es/dc/accesoAbierto.jsp> [Última consulta: 04-06-2023]
- [8] Consejo de la Unión Europea (2022). *Council conclusions on Research assessment and implementation of Open Science*. Disponible en: <https://www.consilium.europa.eu/media/56958/st10126-en22.pdf>
- [9] Angrist, J. y Krueger, A. (2001). *Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments*. *Journal of Economic Perspectives*, 15 (4): 69-85. Disponible en: <https://www.aeaweb.org/articles?id=10.1257/jep.15.4.69>
- [10] Marshall, J. (2016). *Coarsening bias: How coarse treatment measurement upwardly biases instrumental variable estimates*. *Political Analysis*, 24 (2): 157–171. Disponible en: <https://doi.org/10.1093/pan/mpw007>
- [11] Słoczyński, T. (2021). *When Should We (Not) Interpret Linear IV Estimands as LATE?* Institute of Labor Economics. Disponible en: <https://docs.iza.org/dp14349.pdf>

Anexo 1: Análisis de los supuestos para poder identificar el efecto causal

Como se ha explicado en el apartado 2.4 de metodología, para que la variable “Grupo con acceso ICTS” —que representa si el grupo de investigación que lleva a cabo el proyecto tiene algún tipo de ventaja para acceder a alguna ICTS, por ejemplo, porque pertenece a la institución propietaria de la misma o a otra con algún tipo de acuerdo de colaboración para el acceso— pueda ser considerada un instrumento válido, tiene que cumplir una serie de supuestos, los cuales, excepto uno, son inverificables empíricamente y solo se pueden aceptar o rechazar empleando conocimiento experto. A continuación, se analizan dichos supuestos y, de acuerdo con las recomendaciones de la literatura [5], para cada supuesto se discuten sus violaciones plausibles:

1. *Supuesto de selección en observables para el caso instrumental*: el resultado y el instrumento no comparten causas no observadas y, si comparten una causa observada, es posible cerrar el flujo de asociación entre el instrumento y el resultado controlando por dicha causa. Una variable no observada es aquella que, o bien no está considerada en el modelo causal y por tanto no aparece reflejada en el diagrama causal, o bien sí que está reflejada pero no está codificada en la base de datos. En nuestro caso, se incumpliría este supuesto si el diagrama causal estuviera omitiendo alguna variable que afecte tanto al resultado como a la pertenencia a un grupo de investigación con ventaja para acceder a una ICTS. También se incumpliría si el diagrama estuviera omitiendo alguna relación entre variables que conduzca a que resultado e instrumento compartan una causa. Observando el diagrama causal, se puede identificar una posible causa del resultado que puede ser a su vez causa de que el grupo de investigación tenga ventaja para acceder a una ICTS: el área de conocimiento del proyecto, si se diera el caso de que determinadas áreas indujeran más acuerdos de acceso. El análisis de los datos revela que esta variable explica el 7,8% del instrumento; respecto a las variables de resultado, los porcentajes explicados varían: por ejemplo, explica el 27,16% del resultado número de publicaciones en revistas indexadas, el 30,28% del resultado número de publicaciones en revistas Q1, el 10,43% del resultado número de publicaciones en abierto y el 4,7% del resultado obtención de financiación en el Programa Marco. Por tanto, se ha incluido como covariable en el análisis para evitar una violación del supuesto.
2. *Supuesto de exclusión*: el instrumento solo afecta al resultado a través del tratamiento, esto es, pertenecer a un grupo de investigación con ventaja para acceder a una ICTS solo afecta al resultado de interés —por ejemplo, número de publicaciones en revistas indexadas— a través del uso de las instalaciones, recursos o servicios de grandes infraestructuras de investigación para las actividades del proyecto. De nuevo se incumpliría este supuesto si el diagrama causal estuviera omitiendo alguna variable que medie entre el instrumento y el resultado. También se incumpliría si el diagrama estuviera omitiendo alguna relación causal entre el instrumento y una variable que sea causa del resultado. Observando el diagrama causal, se concluye que solo una de las posibles causas del resultado, la calidad científico-técnica del proyecto, podría verse afectada por la pertenencia a un grupo de investigación con ventaja para acceder a una ICTS —si se diera el caso de que estos fueran grupos con una excelencia científica-técnica distinta a los grupos que no tienen ventajas para el acceso—. Pero las personas expertas consultadas han descartado dicha relación causal; es más: el análisis de los datos ni siquiera muestra una asociación significativa entre la variable “Grupo con acceso ICTS” y la variable “Nota” otorgada al proyecto (que podemos usar como un proxy de “Calidad”, que es una variable no observada), obteniéndose un coeficiente de correlación biserial de -0,09.

3. *Supuesto de positividad*: para cada combinación de valores del conjunto de variables de control necesarias para que se cumplan los supuestos anteriores, hay proyectos que reciben el instrumento y otros que no. En nuestro caso, la única covariable es el área de conocimiento del proyecto y en todas las áreas hay proyectos que reciben el instrumento y otros que no.
4. *Supuesto de relevancia*: el instrumento está asociado con el tratamiento. Este supuesto es el único verificable empíricamente, siendo práctica habitual emplear el valor del estadístico F parcial para analizar la fortaleza de la asociación. En nuestro caso, como se recoge en el Anexo 3, asciende a 1.226,47, lo que indica una relación fuerte entre ambas variables.
5. *Supuesto de consistencia*: para cada proyecto, no existen diferentes versiones del instrumento que conduzcan a diferentes tratamientos potenciales²² ni tampoco existen diferentes versiones del tratamiento que conduzcan a diferentes resultados potenciales²³. Este supuesto se incumpliría en el caso de que existieran versiones ocultas del instrumento o del tratamiento. En nuestro caso, tanto el instrumento como el tratamiento están definidos por variables binarias:
 - Si el proyecto es ejecutado por un grupo de investigación con alguna ventaja para acceder a una ICTS, “Grupo con acceso ICTS” = 1; en caso contrario, “Grupo con acceso ICTS” = 0.
 - Si el proyecto hace uso de una ICTS para cumplir con sus objetivos, “Uso de ICTS” = 1; en caso contrario, “Uso de ICTS” = 0.

Pudiera ser, por ejemplo, que, en épocas de alta demanda, si el “Grupo con acceso ICTS” = 1 es un grupo que pertenece a la institución propietaria de una ICTS, se observe “Uso de ICTS” = 1, pero que si el “Grupo con acceso ICTS” = 1 pertenece a otra institución, se observe “Uso de ICTS” = 0. Por tanto, existirían diferentes versiones del instrumento dentro de la categoría “Grupo con acceso ICTS” = 1 que conducirían a diferentes tratamientos, violándose el supuesto. En el caso del tratamiento, el problema es que la insuficiente información de la que se ha dispuesto ha obligado a definirlo de una manera demasiado amplia (“uso o no uso de una ICTS”), lo que puede estar ocultando versiones del tratamiento como, por ejemplo, proyectos que sí han accedido a una gran infraestructura, pero el tiempo de uso concedido no ha sido suficiente para realizar todas las actividades proyectadas inicialmente, lo que puede afectar a sus resultados. La definición demasiado amplia de los tratamientos puede conducir a sesgos positivos en las estimaciones por variables instrumentales²⁴.

6. *Supuesto de no interferencia*: los tratamientos potenciales de cada proyecto no se ven afectados por la exposición de otros proyectos al instrumento; asimismo, los resultados potenciales de cada proyecto no se ven afectados por la exposición de otros proyectos al tratamiento. Aunque es razonable aceptar este supuesto, hay que señalar que puede incumplirse si la disponibilidad de los equipos o servicios de la ICTS no fuera capaz de satisfacer todas las solicitudes de acceso bajo demanda: en este caso, el que un proyecto de un grupo de investigación con ventaja para acceder a la ICTS haga uso de la misma sí que podría verse afectado por otros proyectos con “Grupo con acceso ICTS” = 1.
7. *Supuesto de monotonicidad*: el instrumento nunca puede desincentivar el tratamiento, aunque puede fallar en incentivarlo; en nuestro caso se traduce en que pertenecer a un grupo de investigación con ventaja para acceder a alguna ICTS no puede influir en que el grupo decida que su proyecto no haga uso de una ICTS. Parece razonable aceptar este supuesto.

²² Véase Nota 10.

²³ Véase Nota 8.

²⁴ Para una discusión formal sobre porqué los tratamientos definidos de manera amplia a menudo introducen un sesgo positivo en la estimaciones del efecto, véase [10], citado en [5, pág. 14].

Anexo 2: Los datos y su análisis exploratorio

Como se explica en el apartado 2.2 del cuerpo del informe, los datos observacionales de los que se ha dispuesto para realizar la evaluación son los asociados a los proyectos de investigación financiados por la Agencia Estatal de Investigación, dependiente del Ministerio de Ciencia e Innovación. Entre otras actuaciones, la Agencia Estatal de Investigación gestiona convocatorias de ayudas cofinanciadas por el FEDER que constituyen la principal fuente de financiación competitiva de los proyectos de investigación desarrollados en nuestro país por universidades, organismos públicos de investigación y otros organismos de investigación. La Agencia Estatal de Investigación evalúa la calidad de los proyectos presentados a las convocatorias y financia los mejores; la financiación va dirigida a gastos de personal, material inventariable y fungible, gastos de viajes, publicaciones científicas, patentes, y otros gastos directamente relacionados con el proyecto, como los costes de utilización de grandes infraestructuras de investigación. La cuantía concedida a cada proyecto depende de la ayuda solicitada, de su coste financiable real y de las disponibilidades presupuestarias de cada convocatoria de ayudas. La duración de los proyectos, determinada por las correspondientes convocatorias, varía de dos a seis años.

- Al finalizar la ejecución del proyecto se ha de presentar un informe final que, entre otra información, recoge una serie de indicadores directamente relacionados con los resultados del proyecto, los cuales se graban en la base de datos de la Agencia Estatal de Investigación: número de publicaciones en revistas indexadas, en revistas indexadas en el primer cuartil Q1, en revistas no indexadas, y en acceso abierto; número de tesis doctorales leídas; número de patentes concedidas, licenciadas y en explotación; número de acuerdos de colaboración; número de empresas de base tecnológica creadas; número de proyectos financiados por el Programa Marco de Investigación e Innovación de la UE.
- Para una serie de convocatorias de ayudas a proyectos²⁵, en la fase de solicitud, la Agencia Estatal de Investigación recopila y graba en su base de datos información relativa a si el proyecto planifica acceder a alguna ICTS; no se recopila información relativa a si se planifica acceder a una ESFRI o a otras grandes infraestructuras de investigación. Al finalizar la ejecución de los proyectos, no se graba en base de datos si el proyecto finalmente accedió a la ICTS²⁶; esta información hay que recabarla examinando las actividades reseñadas en el informe final de ejecución.

Para el periodo de programación FEDER 2014-2020 en la fecha de la consulta (16 de enero de 2023) la base de datos de la Agencia Estatal de Investigación disponía de 13.893 registros correspondientes a proyectos finalizados para los que había grabados tanto indicadores de resultado como información relativa a si, en el momento de la solicitud, pretendían hacer uso de alguna ICTS. Un 8,5% de dichos registros (1.191) correspondían a proyectos que sí planificaban acceder a una ICTS. Se examinaron los informes finales de ejecución de dichos proyectos, encontrándose que 854 sí habían hecho uso de una gran infraestructura de investigación para realizar las actividades planificadas. Estos proyectos han constituido el grupo de tratamiento. De los proyectos restantes se han eliminado aquellos pertenecientes a áreas de conocimiento para

²⁵ Básicamente, para todas las convocatorias de proyectos de I+D dirigidas a organismos públicos de investigación, universidades, centros de investigación, centros tecnológicos, institutos de investigación y demás instituciones de I+D+I públicas o privadas sin ánimo de lucro.

²⁶ Hay que tener en cuenta que, como se ha señalado en apartados anteriores, un porcentaje importante de accesos a las ICTS es en régimen competitivo, por lo que la solicitud de acceso puede ser denegada.

las que ningún proyecto ha accedido a una ICTS (ciencias de la educación, ciencias sociales, derecho, economía, estudios feministas, filosofía y filología, psicología), quedando 8.157 proyectos como grupo de control. Para cada proyecto, existe información sobre las siguientes variables (solo se especifican las relevantes para el análisis):

- Variable de tratamiento “Uso de ICTS”: variable binaria; si el proyecto hace uso de ICTS para cumplir con sus objetivos, “Uso de ICTS” = 1; en caso contrario, “Uso de ICTS” = 0.
- Variable instrumento “Grupo con acceso ICTS”: variable binaria que representa si el grupo de investigación que lleva a cabo el proyecto tiene algún tipo de ventaja para acceder a alguna ICTS, por ejemplo, porque pertenece a la institución propietaria de la misma o a otra con algún tipo de acuerdo de colaboración para el acceso. Si el proyecto es ejecutado por un grupo de investigación con alguna ventaja para acceder a alguna ICTS, “Grupo con acceso ICTS” = 1; en caso contrario, “Grupo con acceso ICTS” = 0. Hay que señalar que para la construcción de esta variable solo se ha dispuesto de información relativa a 2021; esto ha podido introducir ruido en los datos, pues se han utilizado estos valores de la variable instrumento para todos los proyectos de la muestra (proyectos solicitados desde el 1 de enero de 2014 y que en enero de 2023 habían finalizado y presentado el informe final de justificación científico-técnica). La Tabla 1 recoge la frecuencia relativa de uso de ICTS entre los proyectos que no reciben el instrumento y entre los que sí lo reciben: la frecuencia relativa de uso es mayor entre el grupo que sí recibe el instrumento (43% vs. 5,7%).

		Uso de ICTS	
		No	Sí
Grupo con ventaja para acceder a alguna ICTS	No	94,3%	5,7%
	Sí	57%	43%

Tabla 1. Frecuencia relativa de uso de ICTS en función de la recepción o no del instrumento.

- Covariable “Planifica uso de ICTS”: variable binaria; si el proyecto planifica hacer uso de ICTS en el momento de la solicitud de la financiación a la Agencia Estatal de Investigación “Planifica uso de ICTS” = 1; en caso contrario, “Planifica uso de ICTS” = 0.
- Covariable “Nota”: variable numérica continua que adopta valores entre 3 y 5 y que corresponde a la calificación otorgada al proyecto por la evaluación científico-técnica realizada por la Agencia Estatal de Investigación.
- Covariable “Diferencia entre solicitada y concedido”: variable numérica continua que corresponde a la diferencia en euros entre la ayuda solicitada por el proyecto y la finalmente concedida por la Agencia Estatal de Investigación.
- Covariable “Área de conocimiento del proyecto”: variable categórica con 21 niveles correspondientes a las áreas de conocimiento en las que la Agencia Estatal de Investigación clasifica los proyectos.
- Variable de resultado “Número de publicaciones en revistas indexadas”: variable numérica discreta que corresponde al número de publicaciones en revistas indexadas directamente relacionadas con el proyecto.
- Variable de resultado “Número de publicaciones en revistas indexadas en el primer cuartil”: variable numérica discreta que corresponde al número de publicaciones en revistas Q1 directamente relacionadas con el proyecto. De los 9.011 registros, 2.500 presentan valores perdidos para esta variable.
- Variable de resultado “Número de publicaciones en revistas no indexadas”: variable numérica discreta que corresponde al número de publicaciones en revistas no indexadas directamente relacionadas con el proyecto.

- Variable de resultado “Número de publicaciones en acceso abierto”: variable numérica discreta que corresponde al número de publicaciones en acceso abierto directamente relacionadas con el proyecto.
- Variable de resultado “Número de tesis doctorales leídas”: variable numérica discreta que corresponde al número de tesis finalizadas y leídas directamente relacionadas con el proyecto. Hay que señalar que, según datos del Ministerio de Educación del periodo 2015-2021, el número medio de años desde el inicio de la formación investigadora hasta la lectura de la tesis doctoral fue de 4,9 años. Por tanto, es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a esta variable, debido a que el tiempo que requiere generar este resultado hace que, en muchos casos, se alcance después del momento en que se registran los indicadores de resultado de los proyectos (normalmente, tres meses después de su finalización).
- Variable de resultado “Número de patentes concedidas”: variable numérica discreta que corresponde al número de patentes concedidas directamente relacionadas con el proyecto. Hay que señalar que, de acuerdo con sus respectivos sitios web, el tiempo medio de concesión en la Oficina Española de Patentes y Marcas es de 2 años y, en la Oficina Europea de Patentes, de 3 a 5 años. Por tanto, es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a esta variable, debido a que el tiempo que requiere generar este resultado hace que, en muchos casos, se alcance después del momento en que se registran los indicadores de resultado de los proyectos.
- Variable de resultado “Número de patentes licenciadas”: variable numérica discreta que corresponde al número de patentes directamente relacionadas con el proyecto que han sido licenciadas a terceros para su explotación. Aplica el mismo comentario de la variable de resultado “Número de patentes concedidas”: es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a esta variable, ya que, en la mayor parte de los casos, los tiempos necesarios para conceder una patente y luego licenciarla no permitirán que el informe final presentado al acabar el proyecto recoja este resultado.
- Variable de resultado “Número de patentes en explotación”: variable numérica discreta que corresponde al número de patentes directamente relacionadas con el proyecto que están en explotación. Aplica el mismo comentario de la Variable de resultado “Número de patentes concedidas”: es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a esta variable, ya que, en la mayor parte de los casos, los tiempos necesarios para conceder una patente y explotarla no permitirán que el informe final presentado al acabar el proyecto recoja este resultado.
- Variable de resultado “Número de acuerdos de colaboración”: variable numérica discreta que corresponde al número de acuerdos de I+D+I suscritos con terceros que tengan que ver con el objeto del proyecto subvencionado. En sentido estricto, un acuerdo de colaboración supone la firma de un documento en el que consta la aportación de las partes que lo suscriben, con un fin concreto, y e implica una compartición del resultado final de la colaboración en términos de propiedad industrial. Pero a veces se asimilan también como acuerdos de colaboración los acuerdos de investigación bajo contrato entre dos partes, donde una solicita a otra un desarrollo a cambio de una contraprestación económica.
- Variable de resultado “Número de empresas de base tecnológica creadas”: variable numérica discreta que corresponde al número de empresas creadas que basan su actividad en la aplicación de los conocimientos científicos y/o tecnológicos desarrollados por el proyecto. Como otras variables de resultado anteriores, es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a esta variable, ya

que, en muchos casos, los tiempos necesarios para poner en marcha una empresa que explote los desarrollos del proyecto no permitirán que el informe final presentado al acabar el proyecto recoja este resultado.

- Variable de resultado “Número de proyectos financiados por el Programa Marco de Investigación e Innovación de la UE”: variable numérica discreta que corresponde al número de proyectos directamente relacionados con el proyecto subvencionado que han recibido financiación del altamente competitivo Programa Marco.

El análisis exploratorio de las variables numéricas anteriores evidencia que todas las variables están sesgadas hacia la derecha, con una media mayor que la mediana. Todas las variables de resultado presentan numerosos valores atípicos y seis de ellas tienen un valor nulo en más del 75% de los proyectos analizados (tercer cuartil = 0): número de patentes concedidas, licenciadas y en explotación; número de acuerdos de colaboración; número de empresas de base tecnológica creadas; y número de proyectos relacionados financiados por el Programa Marco de Investigación e Innovación de la UE. La Tabla 2 recoge los principales estadísticos descriptivos de las variables numéricas relevantes para el análisis y el número de valores atípicos extremos²⁷.

Variable	Nº observaciones	Mínimo	Primer cuartil	Mediana	Media	Tercer cuartil	Máximo	Desviación típica	Nº valores atípicos extremos
Uso de ICTS	9.011	0,00	0,00	0,00	0,09	0,00	1,00	0,29	854
Grupo con acceso ICTS	9.011	0,00	0,00	0,00	0,10	0,00	1,00	0,30	894
Nota	9.011	3,00	3,91	4,11	4,15	4,33	5,00	0,29	20
Diferencia solicitado y concedido	9.011	-242	28.563	69.348	115.442	135.140	18.168.755	330.581,5	472
Nº publicaciones indexadas	9.011	0,00	6,00	11,00	16,29	20,00	439,00	19,04	258
Nº publicaciones indexadas Q1	6.511	0,00	3,00	7,00	11,16	13,00	395,00	16,72	206
Nº publicaciones no indexadas	9.011	0,00	0,00	1,00	2,87	3,00	140,00	6,63	493
Nº publicaciones acceso abierto	9.011	0,00	0,00	1,00	4,83	5,00	508,00	13,58	401
Nº tesis doctorales leídas	9.011	0,00	0,00	1,00	1,86	3,00	69,00	2,21	29
Nº patentes concedidas	9.011	0,00	0,00	0,00	0,31	0,00	17,00	0,90	1.596
Nº patentes licenciadas	9.011	0,00	0,00	0,00	0,09	0,00	12,00	0,45	541
Nº patentes en explotación	9.011	0,00	0,00	0,00	0,05	0,00	12	0,34	297
Nº acuerdos colaboración	9.011	0,00	0,00	0,00	0,36	0,00	16	1,13	1.498
Nº empresas base tecnológica	9.011	0,00	0,00	0,00	0,07	0,00	10	0,36	503
Nº proyectos Programa Marco	9.011	0,00	0,00	0,00	0,34	0,00	45	1,03	1.893

Tabla 2. Análisis descriptivo de las variables relevantes para el análisis.

²⁷ Se ha considerado que un valor es atípico extremo cuando es mayor que el tercer cuartil + 3 veces el rango intercuartílico (tercer cuartil menos el primer cuartil)

Anexo 3: Estimación del efecto causal

Siguiendo la metodología explicada en el apartado 2.4 del cuerpo del informe, se ha construido para cada variable de resultado su correspondiente diagrama causal, que ha resultado ser básicamente el mismo en todos los casos. A continuación, dado el diagrama causal y los datos observacionales disponibles, se ha identificado que es posible estimar el efecto causal empleando la técnica de variables instrumentales, puesto que para todas las variables de resultado existe un instrumento para el que es posible aceptar el cumplimiento de los supuestos necesarios, tal y como se ha argumentado en el Anexo 1. Dicho instrumento es la variable binaria “Grupo con acceso a ICTS”, que representa si el grupo de investigación que lleva a cabo el proyecto tiene algún tipo de ventaja para acceder a alguna ICTS, por ejemplo, porque pertenece a la institución propietaria de la misma o a otra con algún tipo de acuerdo de colaboración para el acceso. El tratamiento es la variable binaria “Uso de la ICTS” y los resultados son las siguientes variables numéricas: número de publicaciones en revistas indexadas, en revistas indexadas Q1, en revistas no indexadas, y en acceso abierto; número de tesis doctorales; número de patentes concedidas, licenciadas y en explotación; número de acuerdos de colaboración; número de empresas de base tecnológica creadas; y número de proyectos relacionados financiados por el Programa Marco de Investigación e Innovación de la UE.

Como la técnica de variables instrumentales es sensible a la presencia de valores atípicos y el análisis exploratorio de los datos recogido en el Anexo 2 ha evidenciado que todas las variables de resultado presentan numerosos valores atípicos, se ha procedido del siguiente modo:

- Se han binarizado las variables de resultado cuyo valor es nulo en más del 75% de los proyectos analizados (tercer cuartil = 0): número de patentes concedidas, licenciadas y en explotación; número de acuerdos de colaboración; número de empresas de base tecnológica creadas; y número de proyectos relacionados financiados por el Programa Marco de Investigación e Innovación de la UE.
- Para el resto de variables (número de publicaciones en revistas indexadas, en revistas indexadas Q1, en revistas no indexadas, en acceso abierto; número de tesis doctorales), se han winsorizado²⁸ los valores atípicos al 99%, 97%, 95% y 93% para comprobar la robustez de la estimación.

El efecto del tratamiento se ha estimado empleando el estimador de mínimos cuadrados ordinarios en dos etapas. Este método emplea dos regresiones: la primera, conocida habitualmente como “primera etapa”, predice la variable de tratamiento usando el instrumento y otras posibles covariables; la segunda regresión predice la variable de resultado usando las predicciones de la primera etapa y las posibles covariables. Como se ha expuesto en el Anexo 1, para evitar una violación del supuesto de selección en observables, se ha incluido en el análisis la covariable “Área de conocimiento” del proyecto. Esta variable explica el 7,8% del instrumento y, respecto a las variables de resultado, los porcentajes varían. Por ejemplo, explica el 27,16% del resultado número de publicaciones en revistas indexadas, el 30,28% del resultado número de publicaciones en revistas Q1, el 10,43% del resultado número de publicaciones en abierto y el 4,7% del resultado obtención de financiación en el Programa Marco.

²⁸ La winsorización que se ha realizado consiste en sustituir los valores atípicos por encima del percentil especificado por el valor de dicho percentil. Así, la winsorización al 95% ha consistido en asignar a los registros cuyo valor de la variable de resultado está por encima del percentil 95 el valor de dicho percentil 95.

En este escenario, con una covariable categórica, un solo instrumento binario, un tratamiento también binario y un cumplimiento fuerte del supuesto de monotonicidad²⁹, el efecto estimado mediante mínimos cuadrados en dos etapas se interpreta como una media ponderada³⁰ del efecto promedio local del tratamiento (LATE) para cada área de conocimiento. El LATE para cada área es el promedio del efecto del tratamiento para el subconjunto de proyectos de dicha área que hace uso de una ICTS si y solo si el grupo de investigación tiene ventaja para acceder a una ICTS. Es un subconjunto de proyectos interesante de estudiar porque, a priori, serán proyectos de una calidad científico-técnica media, suficientemente buena como para recibir financiación de la Agencia Estatal de Investigación, pero no lo suficientemente excelente para tener una nota o una financiación que les incentive a acceder a la ICTS en concurrencia competitiva con otros proyectos. Por tanto, el posible impacto del uso de una ICTS en sus resultados será una estimación conservadora del posible impacto en toda la población de proyectos de investigación financiados por la Agencia Estatal de Investigación.

La Tabla 3 muestra los resultados de la estimación.

- Para cada variable de resultado se recoge el efecto del tratamiento, estimado mediante mínimos cuadrados en dos etapas, y tres intervalos de confianza al 95%: intervalo robusto a la heterocedasticidad, intervalo *bootstrap* (10.000 réplicas) e intervalo de Anderson-Rubin. En el caso de las variables no binarizadas, también se recogen las estimaciones con las diferentes winsorizaciones con sus intervalos de confianza robustos al 95%.
- Asimismo, se recogen los efectos estimados y los correspondientes intervalos de confianza robustos de la primera etapa y de la denominada “forma reducida”, que es la regresión de la variable de resultado sobre el instrumento y las posibles variables de control³¹.
- Por último, se muestran los efectos estimados mediante mínimos cuadrados ordinarios controlando por las variables necesarias para satisfacer el criterio de puerta trasera y por sus interacciones (hay que tener en cuenta que los efectos estimados mediante mínimos cuadrados ordinarios son efectos promedios del tratamiento, no efectos promedios locales).



Figura 29. El ESRF (European Synchrotron Radiation Facility) es un “súper microscopio” que permite desentrañar los detalles más íntimos de la materia. España es socio fundador de esta instalación europea, con una participación del 4% en el capital de la sociedad. Gracias a ello, los científicos españoles tienen derecho a tiempo de experimentación en sus 52 líneas de haz. Fuente: Ministerio de Ciencia e Innovación.

²⁹ Véase [11, pág. 27].

³⁰ Los pesos son una función de cuánta varianza induce el instrumento en el tratamiento en cada nivel de la covariable [5, pág. 55].

³¹ El efecto estimado de la primera etapa representa cuan fuertemente el instrumento afecta al tratamiento; cuanto más pequeño sea, mayor es la posibilidad de sesgos en la estimación del efecto del tratamiento. Por su parte, el efecto estimado de la forma reducida es proporcional al efecto estimado del tratamiento; si no es significativamente distinto de cero, se debe presumir que el efecto del tratamiento es nulo o bien que la asociación entre instrumento y tratamiento es demasiado débil para detectarlo [9, pág. 80].

Variable de resultado	Estimador	Estimación puntual	Intervalo robusto 95%	Intervalo bootstrap 95%	Intervalo Anderson-Rubin 95%	p-valor
	Primera etapa Estadístico F = 1226,47	0,33	[0,30, 0,36]			
Nº publicaciones indexadas	Mín. cuadrados 2 etapas	5,12	[1,56, 8,68]	[1,56, 8,81]	[1,40, 8,85]	0,005
	Forma reducida	1,69	[0,51, 2,87]			0,007
Winsor. 99%	Mín. cuadrados 2 etapas	5,63	[2,18, 9,09]		[2,45, 8,82]	0,001
Winsor. 97%	Mín. cuadrados 2 etapas	4,88	[1,97, 7,78]		[2,20, 7,56]	0,001
Winsor. 95%	Mín. cuadrados 2 etapas	4,25	[1,66, 6,85]		[1,83, 6,85]	0,001
Winsor. 93%	Mín. cuadrados 2 etapas	3,68	[1,32, 6,04]		[1,45, 5,91]	0,002
	Min. cuadrados (ATE)	5,01	[1,80, 8,26]			0,002
Nº publicaciones indexadas Q1	Mín. cuadrados 2 etapas	5,56	[1,65, 9,47]	[1,73, 9,61]	[1,52, 9,62]	0,005
	Forma reducida	1,72	[0,50, 2,93]			0,007
Winsor. 99%	Mín. cuadrados 2 etapas	6,02	[2,39, 9,64]		[2,89, 9,16]	0,001
Winsor. 97%	Mín. cuadrados 2 etapas	4,40	[1,64, 7,16]		[1,94, 6,86]	0,002
Winsor. 95%	Mín. cuadrados 2 etapas	3,57	[1,21, 5,94]		[1,42, 5,73]	0,003
Winsor. 93%	Mín. cuadrados 2 etapas	3,05	[0,92, 5,17]		[1,09, 5,01]	0,005
	Min. cuadrados (ATE)	3,76	[0,21, 7,31]			0,038
Nº publicaciones no indexadas	Mín. cuadrados 2 etapas	0,35	[-1,04, 1,75]	[-1,01, 1,82]	[-1,03, 1,74]	0,617
	Forma reducida	0,12	[-0,40, 0,71]			0,585
Winsor. 99%	Mín. cuadrados 2 etapas	0,10	[-0,89, 1,10]		[-1,00, 1,21]	0,839
Winsor. 97%	Mín. cuadrados 2 etapas	0,11	[-0,69, 0,90]		[-0,77, 0,98]	0,794
Winsor. 95%	Mín. cuadrados 2 etapas	0,10	[-0,58, 0,79]		[-0,63, 0,84]	0,767
Winsor. 93%	Mín. cuadrados 2 etapas	0,11	[-0,49, 0,72]		[-0,52, 0,75]	0,711
	Min. cuadrados (ATE)	1,98	[1,14, 2,81]			0,000
Nº publicaciones en acceso abierto	Mín. cuadrados 2 etapas	2,06	[-0,39, 4,51]	[-0,36, 4,57]	[-0,77, 4,89]	0,010
	Forma reducida	0,68	[-0,13, 1,49]			0,154
Winsor. 99%	Mín. cuadrados 2 etapas	2,52	[0,61, 4,43]	[0,61, 4,49]	[0,76, 4,29]	0,001
Winsor. 97%	Mín. cuadrados 2 etapas	2,05	[0,63, 3,47]		[0,70, 3,40]	0,005
Winsor. 95%	Mín. cuadrados 2 etapas	1,87	[0,67, 3,06]		[0,74, 3,01]	0,002
Winsor. 93%	Mín. cuadrados 2 etapas	1,80	[0,74, 2,86]		[0,81, 2,80]	0,001
	Min. cuadrados (ATE)	0,20	[-2,88, 3,29]			0,897
Nº tesis doctorales leídas	Mín. cuadrados 2 etapas	0,18	[-0,23, 0,65]	[-0,21, 0,60]	[-0,29, 0,65]	0,383
	Forma reducida	0,06	[-0,08, 0,20]			0,445
Winsor. 99%	Mín. cuadrados 2 etapas	0,14	[-0,24, 0,53]		[-0,26, 0,55]	0,458
Winsor. 97%	Mín. cuadrados 2 etapas	0,12	[-0,24, 0,48]		[-0,26, 0,51]	0,501
Winsor. 95%	Mín. cuadrados 2 etapas	0,12	[-0,22, 0,46]		[-0,24, 0,48]	0,491
Winsor. 93%	Mín. cuadrados 2 etapas	0,13	[-0,19, 0,46]		[-0,20, 0,47]	0,412
	Min. cuadrados (ATE)	0,03	[-0,30, 0,35]			0,859
Patentes (binaria)	Mín. cuadrados 2 etapas	-0,07	[-0,15, 0,01]	[-0,15, 0,01]		0,010
	Forma reducida	-0,02	[-0,05, 0,004]			0,096
	Min. cuadrados (ATE)	-0,05	[-0,12, 0,02]			0,197
Patentes licenciadas (binaria)	Mín. cuadrados 2 etapas	0,004	[-0,05, 0,06]	[-0,05, 0,06]	[-0,05, 0,05]	0,895
	Forma reducida	0,001	[-0,02, 0,02]			0,891
	Min. cuadrados (ATE)	-0,02	[-0,06, 0,03]			0,521
Patentes en explotación (binaria)	Mín. cuadrados 2 etapas	-0,01	[-0,05, 0,02]	[-0,05, 0,02]		0,484
	Forma reducida	-0,004	[-0,02, 0,01]			0,506
	Min. cuadrados (ATE)	0,002	[-0,03, 0,03]			0,907
Acuerdos colaboración (binaria)	Mín. cuadrados 2 etapas	0,06	[-0,02, 0,15]	[-0,02, 0,15]		0,157
	Forma reducida	0,02	[-0,01, 0,05]			0,134
	Min. cuadrados (ATE)	-0,07	[-0,15, 0,004]			0,063
Empresas base tecnológica (binaria)	Mín. cuadrados 2 etapas	-0,01	[-0,06, 0,04]	[-0,06, 0,04]	[-0,06, 0,04]	0,679
	Forma reducida	-0,003	[-0,02, 0,01]			0,679
	Min. cuadrados (ATE)	0,02	[-0,02, 0,06]			0,335
Proyectos Programa Marco UE (binaria)	Mín. cuadrados 2 etapas	0,17	[0,07, 0,26]	[0,07, 0,26]	[0,08, 0,25]	0,001
	Forma reducida	0,05	[0,02, 0,09]			0,0002
	Min. cuadrados (ATE)	0,01	[-0,07, 0,09]			0,778

Tabla 3. Estimaciones de los efectos del tratamiento sobre las variables de resultado.

A continuación, se comentan los resultados de la Tabla 3:

- La primera etapa, idéntica en todos los análisis puesto que las variables instrumento y covariable son comunes a todas las variables de resultado, evidencia que el instrumento afecta al tratamiento y que la asociación entre ambas variables es suficientemente fuerte, con un estadístico F parcial de 1.226,47. El test F robusto a heterocedasticidad y *clustering* arroja como resultado que el coeficiente del instrumento es significativamente distinto de cero (p -valor $< 2.2e-16$).
- Como era esperable, los efectos promedios del tratamiento estimados mediante mínimos cuadrados ordinarios están sesgados hacia los efectos en el grupo de no tratados o control. Esto es debido a la violación estructural del supuesto de positividad: todos los proyectos que no planifican acceder a una ICTS (valor de “Planifica uso ICTS” = 0) pertenecen al grupo de control, por lo que las estimaciones para el efecto del tratamiento están basadas en gran medida en la extrapolación de los resultados obtenidos por el grupo de control.
- Respecto a la estimación por mínimos cuadrados en dos etapas, **tres variables de resultado presentan una estimación positiva y significativa del efecto del tratamiento: el número de publicaciones en revistas indexadas, el número de publicaciones en revistas indexadas Q1 y la existencia de proyectos relacionados que han obtenido financiación en el Programa Marco de Investigación e Innovación de la UE**. Para las tres variables, la estimación del coeficiente del instrumento en la forma reducida es estadísticamente significativa y es coherente en signo y magnitud con el efecto esperado sobre el resultado. En el caso de las dos variables numéricas, las sucesivas winsorizaciones al 99%, 97%, 95% y 93% siguen mostrando estimaciones positivas.
- La variable de resultado número de publicaciones en acceso abierto presenta una estimación positiva pero no es estadísticamente distinta de cero (los correspondientes intervalos de confianza al 95% contienen al cero). Sin embargo, las sucesivas winsorizaciones al 99%, 97%, 95% y 93% sí que son estadísticamente significativas. Por tanto, se considera que **hay suficiente evidencia que soporta la estimación positiva y significativa del efecto del tratamiento para la variable número de publicaciones en acceso abierto**.
- En cualquier caso, **hay que resaltar que es probable que las magnitudes de las estimaciones obtenidas tengan un sesgo positivo** debido a que el tratamiento “uso de las instalaciones, recursos o servicios de grandes infraestructuras de investigación para las actividades del proyecto” está definido de una manera demasiado amplia³². Además, los numerosos valores atípicos afectan a la exactitud de las estimaciones en el caso de las variables no binarizadas.
- Para el resto de los resultados, la estimación del efecto del tratamiento no es estadísticamente distinta de cero (los correspondientes intervalos de confianza al 95% contienen al cero). De todos modos, como se ha indicado en el Anexo 2, hay que señalar que es probable que no estén ajustados a la realidad los valores de la base de datos de la Agencia Estatal de Investigación relativos a tesis doctorales, patentes y empresas de base tecnológica generadas por los proyectos, debido a que el tiempo que requiere producir estos resultados hace que, en muchos casos, se alcancen después del momento en que se registran los indicadores de resultado de los proyectos (normalmente, tres meses después de su finalización).

³² Véase Nota 24.

Anexo 4: Análisis de sensibilidad

Como se explica en el apartado 2.4.D) del cuerpo del informe, cuando se aplica la técnica de variables instrumentales a datos observacionales es probable que se violen los supuestos analizados en el Anexo 1, en particular, los supuestos de selección en observables (el resultado y el instrumento no comparten causas no observadas) y de exclusión (el instrumento solo afecta al resultado a través del tratamiento). Los análisis de sensibilidad permiten evaluar en qué medida deben violarse los supuestos para cambiar las conclusiones respecto al efecto del tratamiento. En la presente evaluación se ha empleado el análisis de sensibilidad propuesto en [4], que emplea dos estadísticos principales que se calculan respecto a la forma reducida del estimador de mínimos cuadrados en dos etapas:

- El valor de robustez (*robustness value RV*): si existen variables omitidas en la estimación por mínimos cuadrados en dos etapas que afectan tanto al instrumento como al resultado, el RV indica la mínima varianza residual de las variables instrumento y resultado que tiene que ser explicada por dichas variables omitidas para que la estimación del efecto del tratamiento no sea estadísticamente distinta de cero a un nivel de significación α (esto es, para que el intervalo de confianza $1-\alpha$ contenga al cero).
- El valor de robustez extremo (*extreme robustness value XRV*): si existen variables omitidas en la estimación por mínimos cuadrados en dos etapas que afectan tanto al instrumento como al resultado y no se imponen restricciones a su fortaleza de asociación con el resultado, el XRV indica la mínima varianza residual del instrumento que tiene que ser explicada por dichas variables omitidas para que la estimación del efecto del tratamiento no sea estadísticamente distinta de cero a un nivel de significación α (esto es, para que el intervalo de confianza $1-\alpha$ contenga al cero).

Para las variables con estimaciones del efecto del tratamiento estadísticamente distintas de cero, la Tabla 4 recoge los resultados del análisis de sensibilidad a la violación de los supuestos de exclusión y de selección en observables para el caso instrumental. La estimación menos sensible es la correspondiente a la obtención de financiación en el Programa Marco de Investigación e Innovación de la UE y la más sensible es la del efecto del tratamiento sobre el número de publicaciones en revistas indexadas.

Variable resultado	Estadísticos ($\alpha=0,05$)	Interpretación ³³
Nº publicaciones indexadas	Valor robusto	0,0280
	Valor robusto extremo	0,0077

Las posibles variables no observadas necesitan explicar al menos el 2,8% de la varianza residual de las variables instrumento y resultado para que el intervalo de confianza al 95% de la estimación del efecto del tratamiento contenga al cero (o el 0,77% de la varianza residual del instrumento en el caso extremo en el que las variables no observadas expliquen el 100% de la varianza residual del resultado).

Tabla 4. Análisis de sensibilidad a la violación de los supuestos de exclusión y de selección en observables para el caso instrumental.

³³ Los resultados del análisis de sensibilidad son exactos para una sola variable no observada y conservadores para múltiples variables no observadas, posiblemente actuando de forma no lineal.

Variable resultado	Estadísticos ($\alpha=0,05$)	Interpretación	
Nº publicaciones indexadas Q1	Valor robusto	0,0329	Las posibles variables no observadas necesitan explicar al menos el 3,29% de la varianza residual de las variables instrumento y resultado para que el intervalo de confianza al 95% de la estimación del efecto del tratamiento contenga al cero (o el 0,91% de la varianza residual del instrumento en el caso extremo en el que las variables no observadas expliquen el 100% de la varianza residual del resultado).
	Valor robusto extremo	0,0091	
Nº publicaciones en acceso abierto	Valor robusto	0,0359	Las posibles variables no observadas necesitan explicar al menos el 3,59% de la varianza residual de las variables instrumento y resultado para que el intervalo de confianza al 95% de la estimación del efecto del tratamiento contenga al cero (o el 1,57% de la varianza residual del instrumento en el caso extremo en el que las variables no observadas expliquen el 100% de la varianza residual del resultado).
	Valor robusto extremo	0,0157	
Proyectos financiados Programa Marco (binaria)	Valor robusto	0,0386	Las posibles variables no observadas necesitan explicar al menos el 3,86% de la varianza residual de las variables instrumento y resultado para que el intervalo de confianza al 95% de la estimación del efecto del tratamiento contenga al cero (o el 1,86% de la varianza residual del instrumento en el caso extremo en el que las variables no observadas expliquen el 100% de la varianza residual del resultado).
	Valor robusto extremo	0,0186	

Tabla 4 (continuación). *Análisis de sensibilidad a la violación de los supuestos de exclusión y de selección en observables para el caso instrumental.*

Para poder interpretar adecuadamente los resultados mostrados en la Tabla 4, se pueden establecer comparaciones con las varianzas residuales explicadas por la covariable observada, el “Área de conocimiento” del proyecto. Esta variable explica el 7,8% del instrumento, el 27,16% del resultado número de publicaciones en revistas indexadas, el 30,28% del resultado número de publicaciones en revistas Q1, el 10,43% del resultado número de publicaciones en abierto y el 4,7% del resultado obtención de financiación en el Programa Marco. Por tanto, si las variables no observadas fueran capaces de explicar la misma variabilidad explicada por la variable área de conocimiento del proyecto, los efectos estimados no serían estadísticamente distintos de cero. Lo anterior se puede visualizar mediante gráficos de sensibilidad como los de la Figura 30 que representan, en el eje horizontal, la parte residual de la variación del instrumento que hipotéticamente se explica por las variables no observadas y, en el eje vertical, la parte residual de la variación del resultado que hipotéticamente se explica por las variables no observadas. Las curvas de contorno que aparecen en los gráficos muestran qué estimación de la forma reducida³⁴ se habría obtenido incluyendo en la regresión las variables no observadas con dichas fortalezas hipotéticas. La línea roja discontinua corresponde al límite de significación estadística. Los rombos rojos indican la asociación con el instrumento y con el resultado de la covariable área de conocimiento del proyecto: si una o múltiples variables no observadas tuvieran una asociación tan fuerte, los efectos del tratamiento estimados estarían en la región de no significancia estadística.

Las personas expertas consultadas contemplan como poco probable que exista una variable no observada con una fortaleza de asociación análoga, aunque, como se observa en los gráficos, variables no observadas con menores fortalezas de asociación también podrían hacer estadísticamente no significativos los efectos estimados.

³⁴ Se emplea la forma reducida porque, como se ha explicado en la Nota 31, el efecto estimado de la forma reducida es proporcional al efecto estimado del tratamiento; si no es significativamente distinto de cero, se debe presumir que el efecto del tratamiento es nulo.

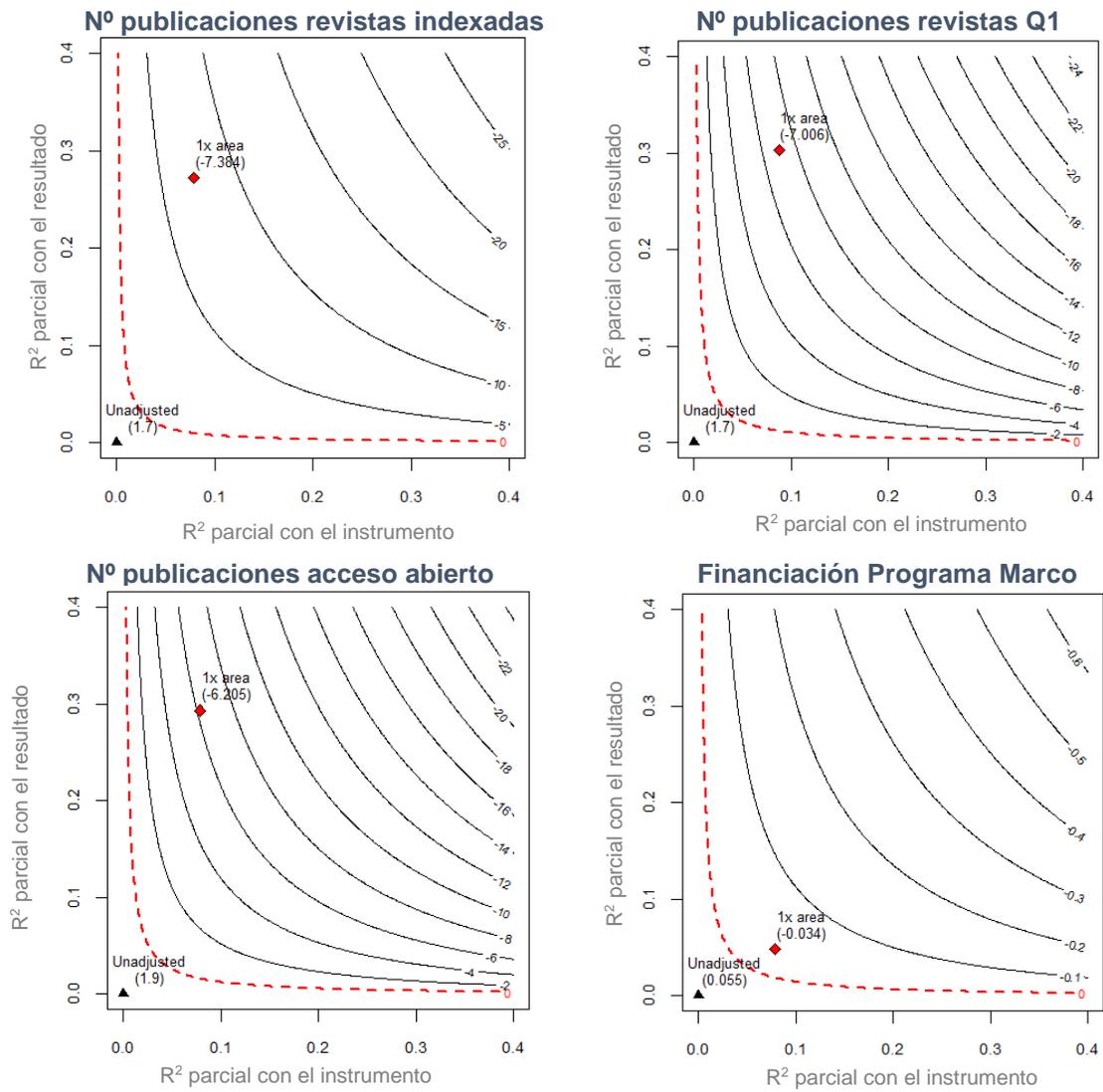


Figura 30. Gráficos de sensibilidad de las estimaciones de la forma reducida para las cuatro variables de resultado con estimaciones positivas del efecto de tratamiento.

FEDER

Fondo Europeo de Desarrollo Regional

Una manera de hacer Europa

