

La evaluación de impacto de las políticas activas de empleo en España: el deseo y la realidad

Impact evaluation of active employment policies in Spain: desire and reality

Inmaculada Cebrián López
Universidad de Alcalá

Begoña Cueto Iglesias
Francisco Javier Mato Díaz
Universidad de Oviedo

Resumen

Este artículo revisa aspectos clave relacionados con la evaluación del impacto de las políticas activas de empleo en España. Constatando que los Planes Anuales de Política de Empleo contemplan la evaluación, se advierten diversas insuficiencias para captar los efectos causales de las políticas desarrolladas. Frente a las dificultades de homogeneización de políticas e indicadores, se señala que el modelo descentralizado del Fondo Social Europeo ha potenciado las evaluaciones de impacto con enfoque contrafactual durante los últimos años. Finalmente, se analizan aspectos relacionados con los datos administrativos y las fuentes de información que podrían facilitar las evaluaciones de impacto en España.

Palabras clave: políticas activas de empleo, evaluación de impacto, contrafactual, Fondo Social Europeo, datos administrativos.

Clasificación JEL: J38, J68.

Abstract

This article reviews key aspects related to evaluating the impact of active employment policies in Spain. Noting that the Annual Employment Policy Plans contemplate evaluation, various shortcomings regarding the estimation of the causal effects of the policies are raised. As compared with the difficulties faced by homogenization of policies and indicators, it is pointed out that the decentralized model of the European Social Fund has promoted impact evaluations with a counterfactual approach in recent years. Finally, aspects related to administrative data and sources of information that could facilitate impact evaluations in Spain are analyzed.

Keywords: active employment policies, impact evaluation, counterfactual evaluation, European Social Fund, administrative data.

1. Introducción

En 1997, con la transferencia de las políticas activas de empleo a las comunidades autónomas, se pone en marcha la primera de las reformas laborales que centra su objetivo principal en la flexibilización del mercado de trabajo al tiempo que se fomenta el empleo indefinido como garantía de estabilidad laboral. Ni esta ni ninguna de las posteriores (2001, 2006, 2010, 2012) modificó sustancialmente el modelo de políticas activas, a pesar de que en todas se fueron modificando algunos elementos que tenían mucho que ver con las políticas activas de empleo, por ejemplo, el programa de incentivos económicos a la contratación, los contratos de formación y aprendizaje, las empresas de trabajo temporal (ETT) como intermediarios laborales, etc.

En España, la promoción y desarrollo de las políticas activas de empleo (PAE) tiene lugar dentro del Sistema Nacional de Empleo (SNE). Está integrado por el Servicio Público de Empleo Estatal (SEPE) y los 17 servicios públicos de empleo autonómicos (SPE). El SEPE y los SPE financian, gestionan y ejecutan las PAE, siendo el SEPE el coordinador.

El Sistema de Información de los Servicios Públicos de Empleo, conocido como SISPE, es la principal fuente de información, comunicación y coordinación del SNE. Por otro lado, el principal instrumento de colaboración de todos los SPE es la Conferencia Sectorial de Empleo y Asuntos Laborales, en la cual se determinan los servicios y programas que se van a desarrollar a nivel comunitario con cargo a los fondos distribuidos por el SEPE.

La financiación del SNE consta de cuatro niveles administrativos: la Unión Europea, la Administración General del Estado a través del SEPE, las comunidades autónomas y las entidades locales.

La Unión Europea es un agente financiador de políticas activas de empleo, principalmente a través del Fondo Social Europeo (FSE), Programas Operativos plurirregionales, regionales y de la Iniciativa de Empleo Juvenil. El SEPE desarrolla su papel como agente financiador, gestor, ejecutor y coordinador en el ámbito de las políticas activas de empleo.

Las comunidades autónomas son agentes financiadores, gestores y ejecutores de políticas activas de empleo, a través de sus respectivos SPE. Al tener las competencias en el ámbito de las políticas activas de empleo, se posicionan como agentes principales en la ejecución de esas políticas. Por otro lado, ejercen un papel significativo en su financiación, tanto en lo relativo a la financiación vinculada a programas y servicios integrados en el Plan Anual de Política de Empleo (PAPE) como en lo que respecta a la financiación de otras acciones propias no integradas en el PAPE o por la financiación que reciben de manera directa de la Unión Europea.

A nivel local, la gran mayoría de entidades locales (diputaciones, municipios, cabildos y consejos, comarcas, áreas metropolitanas y mancomunidades) juegan un papel como agentes financiadores, gestores y ejecutores de políticas activas de empleo a nivel de sus respectivos territorios.

La evaluación de las políticas activas realizada por la Autoridad Independiente de Responsabilidad Fiscal (AIReF) en 2019 puso de manifiesto las dificultades de coordinación en el Sistema Nacional de Empleo, lo que obstaculiza establecer objetivos y planificar. Así, en el sexto eje de objetivos estructurales del PAPE 2020, se indica que uno de los objetivos es la mejora de la gestión, coordinación y colaboración en el seno del Sistema Nacional de Empleo. Esta mejora contribuiría a establecer una evaluación rigurosa del gasto y del impacto de las medidas aplicadas, contribuyendo a la puesta en marcha de procedimientos que sean eficientes para destinar los fondos a las prácticas más eficaces para lograr los objetivos de las PAE.

En los apartados siguientes, se presentan una serie de análisis y reflexiones sobre el marco que regula las PAE y su evaluación, el FSE y su evaluación, así como la posibilidad que brindan los datos registrados en la Administración pública para proceder a la evaluación de impacto. Finalmente, se presentan algunas propuestas para una mejor gestión de las PAE y de su evaluación, y para favorecer la implementación de las prácticas más adecuadas.

2. La evaluación en el marco legislativo

La evaluación de las políticas activas está prevista formalmente en la mayor parte de la legislación. El artículo 39 del Real Decreto Legislativo 3/2015 por el que se aprueba el texto refundido de la Ley de Empleo establece que «el Gobierno presentará una memoria sobre el gasto, incluido el necesario para la gestión del sistema, y los resultados de las políticas activas de empleo en el conjunto del Estado».

En la Estrategia Española de Activación para el Empleo 2017-2020, se concede «especial relevancia a la planificación, evaluación y seguimiento de las políticas de activación para el empleo en base a resultados, una de las líneas de actuación prioritarias expresamente acordadas en la Conferencia de Presidentes del 17 de enero de 2017». En la misma, se establecen dos niveles para la evaluación basados en resultados:

1. Cuatro evaluaciones distintas, a realizar de forma anual o bienal:
 - a) Una evaluación estratégica del cumplimiento de objetivos.
 - b) Una autoevaluación anual del desempeño.
 - c) Una evaluación externa bienal del desempeño junto con la Red Europea de Servicios Públicos de Empleo.
 - d) Una evaluación anual de la calidad, impacto, eficacia y eficiencia del conjunto del sistema de formación profesional para el empleo.
2. Evaluaciones a realizar en cada uno de los servicios y programas concretos.

Uno de los objetivos estratégicos, el número 6.6, es el «Desarrollo e implantación de una metodología común para la evaluación del impacto de los programas de políticas de activación». De esta forma, se establecía que durante la vigencia de la Estrategia se acordaría e implantaría «una metodología para la evaluación de los

programas y medidas concretos recogidos en el PAPE». También se menciona una metodología sobre la recogida de datos y la selección de contrafactual¹. En la misma normativa, se establece que los «resultados de todas las evaluaciones se harán públicos, en cumplimiento del principio de transparencia y como factor determinante para la modernización y mejora del Sistema Nacional de Empleo».

La cuestión es hasta qué punto se ha avanzado en la consecución de estos objetivos. Ciertamente, se han cumplido parte de los objetivos al haber encargado a la AIREF la evaluación del conjunto de políticas activas (AIREF 2019; 2020). Así, en junio de 2019 se publicaba el tercer Estudio del *Spending Review* sobre Políticas Activas de Empleo² y, en octubre de 2020, la segunda fase del *Spending Review* sobre Incentivos a la contratación y al trabajo autónomo³, sobre las que hablaremos más adelante.

En la elaboración de los Planes Anuales de Política de Empleo, PAPE, también se destaca la evaluación del cumplimiento de objetivos. El PAPE, que se publica anualmente, establece los objetivos a alcanzar en cada año en términos de empleo y los indicadores que se utilizarán para valorar el grado de consecución. En este sentido, se indica que este Plan se configura como «instrumento de evaluación de los servicios y programas de políticas activas de empleo desarrolladas por el Servicio Público de Empleo Estatal y las comunidades autónomas».

A partir de esta breve revisión, podemos concluir que la evaluación de impacto ya está incluida en el marco legislativo de las políticas de empleo. Sin embargo, no parece que, efectivamente, se hayan puesto en marcha varias de las medidas consideradas. O, si se han puesto en marcha, no parece que sus resultados hayan sido públicos, con la excepción de las evaluaciones realizadas por la AIREF.

2.1. Indicadores en el PAPE

En cada uno de los PAPE de los últimos años, se han establecido distintos indicadores, con sus respectivos componentes, con el fin de valorar la consecución de los distintos objetivos clave, estratégicos y estructurales.

El PAPE 2019⁴ estableció un conjunto de 22 indicadores formados por 27 componentes. Esto supuso una simplificación importante respecto a años previos, aunque

¹ El enfoque contrafactual en la evaluación de políticas implica la existencia de un grupo de control que sirva de comparación para saber qué hubiera pasado al grupo de tratamiento, si éste no hubiera participado en la política a evaluar. Como se comentará más adelante, una de las claves en este tipo de evaluaciones es seleccionar un grupo de control adecuado que evite la existencia de sesgos de selección.

² <https://www.airef.es/es/politicas-activas/>

³ <https://www.airef.es/es/spending-review-fase-ii-estudio-incentivos-a-la-contratacion-y-al-trabajo-autonomo/>

⁴ Resolución de 12 de marzo de 2019, de la Secretaría de Estado de Empleo, por la que se publica el Acuerdo del Consejo de Ministros de 8 de marzo de 2019, por el que se aprueba el Plan Anual de Política de Empleo para 2019, según lo establecido en el artículo 11.2 del texto refundido de la Ley de Empleo, aprobado por el Real Decreto Legislativo 3/2015, de 23 de octubre (BOE núm. 64, de 15 de marzo de 2019).

segúan siendo muchos y muy complejos en su cálculo. En el caso del PAPE 2020⁵, se han establecido 8 indicadores para los objetivos clave, 17 para los estratégicos y 33 para los estructurales. En este caso, hay que tener en cuenta que varios de ellos se han establecido durante la situación de pandemia y su objetivo era, fundamentalmente, monitorizar en qué medida se recupera la situación de febrero de 2020.

Una de las debilidades de estos indicadores para ser utilizados como instrumento en la evaluación de las políticas de empleo es que no miden la eficacia ni la eficiencia de los distintos programas de empleo, sino que están influidos, principalmente, por la coyuntura económica del territorio en que se implementan estos programas⁶. Por ejemplo, en los indicadores de los objetivos clave se calcula: la transición del desempleo al empleo, en proporción al número de desempleados inscritos en los SPE; el número de personas que dejan de estar inscritas como desempleadas en los SPE, en proporción a los desempleados inscritos; las transiciones al empleo dentro de los 6 (12) meses siguientes a su inscripción en los SPE.

Estos indicadores tienen como objetivo medir la *contribución* a la reducción del desempleo y a su duración. Sin embargo, las cifras que se obtengan van a estar más determinadas por la estructura económica de la zona en que se llevan a cabo las políticas, y por su coyuntura, que por la eficacia de las políticas.

La mayor parte de los indicadores incluidos miden la cobertura de las medidas puestas en marcha. Así, se hace referencia a los demandantes a quienes se les proporcione atención, respecto al total de demandantes, quienes hayan recibido un itinerario, algún tipo de formación u oportunidades de empleo.

En 2019, había dos indicadores que tenían como objetivo, específicamente, evaluar el impacto de los programas. El primero trataba de medir la cobertura y el impacto que había tenido la formación profesional para el empleo impartida en los demandantes de empleo. El segundo planteaba medir la cobertura y el resultado de la atención prestada para el fomento del empleo autónomo y el emprendimiento y la eficacia en términos de inserción en el Régimen Especial de Trabajadores Autónomos (RETA).

Aunque en estos dos indicadores se establece una comparación entre los demandantes que participaron en programas de empleo con aquellos que no lo hicieron, no tienen en cuenta las características de las personas beneficiarias ni los procesos de selección que llevaron a que unos demandantes sí participaran en la política y otros no. Como consecuencia, son indicadores sesgados. Existe, además, la posibilidad de que, dado que la financiación está condicionada a estos indicadores, se opte por

⁵ Resolución de 26 de octubre de 2020, de la Secretaría de Estado de Empleo y Economía Social, por la que se publica el Acuerdo de Consejo de Ministros de 20 de octubre de 2020, por el que se aprueba el Plan Anual de Política de Empleo para 2020, según lo establecido en el artículo 11.2 del texto refundido de la Ley de Empleo, aprobado por el Real Decreto Legislativo 3/2015, de 23 de octubre (BOE núm. 289, de 2 de noviembre de 2020).

⁶ Como muestra del equívoco que pueden causar estos errores, el trabajo de Vila *et al.* (2020) señala que la evaluación contrafactual reduce en cerca de dos tercios el efecto aparente de los programas de formación y de inserción de la Comunidad Valenciana, mientras que eleva el efecto de los programas de fomento del empleo.

seleccionar los participantes que pueden tener mejor inserción laboral (independientemente de que la política les resulte útil o no, dando lugar al efecto *creaming*) o por llevar a cabo políticas en los que sea más fácil obtener un buen indicador. Por ejemplo, llevando a cabo cursos de corta duración en lugar de otros más largos; o dirigiendo las políticas a colectivos con menos dificultades, porque será más fácil obtener mayores tasas de transición al empleo que si se dirigen a grupos con graves problemas de empleabilidad.

Estos indicadores no se han incluido en 2020. En el PAPE de dicho año se indica que «el diseño del sistema de indicadores no puede ni debe limitarse a la información cuantitativa basada en los datos de la gestión de cada Servicio Público de Empleo que ofrece solamente una perspectiva de impacto de las medidas en el correspondiente ámbito territorial». Por ello, se habla de un «nuevo enfoque de evaluación» de manera que se lleve a cabo un «análisis integral de la “política de activación”» para conocer hasta qué punto se están alcanzando sus objetivos y cuáles son las medidas que se demuestran más eficientes para ello.

Este párrafo muestra que no se entiende lo que significa evaluar el impacto de un programa de empleo. Una evaluación de impacto debe aportar información que muestre en qué medida los cambios en la situación (en este caso, laboral) de las personas pueden atribuirse a la participación en un programa. La clave, por tanto, está en establecer una relación causal entre el programa y un resultado de interés.

Conocer cuántas personas se han formado o a cuántas se les ha ofrecido un itinerario de inserción es relevante. También, saber en qué medida las personas registradas en los servicios públicos de empleo transitan a la ocupación. Pero estos resultados no implican que haya impacto. Tras haber pasado por una acción formativa o por un itinerario se puede haber obtenido un empleo, pero ello no implica que *se deba a* que se ha participado en dichas políticas. Por eso son necesarias las evaluaciones de impacto, porque es necesario saber en qué medida las políticas contribuyen, en sí mismas, a que los participantes accedan al empleo. No deben ser el único factor a tener en cuenta en el diseño de las políticas, pero sí deberían considerarse.

Los PAPE se centran en indicadores de seguimiento y de resultados, pero no de impacto. Como se verá más adelante, sí que se han llevado a cabo evaluaciones de impacto, en el ámbito académico⁷ y también en algunas comunidades autónomas. Estas evaluaciones muestran con claridad que ambos conceptos –resultado e impacto– son diferentes. La evaluación de los incentivos a la contratación y al trabajo autónomo es una buena muestra de estas diferencias.

2.2. Un ejemplo de la diferencia entre seguimiento y evaluación de impacto

Tomemos, como ejemplo, la tarifa plana para trabajadores autónomos. De acuerdo con la evaluación de la AIREF, entre 2013 y 2018, se dedicaron 3.500 millones

⁷ Véase Malo y Cueto (2015) y Mato (2010).

de euros a reducciones y bonificaciones a la Seguridad Social en este programa, en torno a 600 millones cada año entre 2014 y 2018. Este gasto llegó a unos 600.000 beneficiarios cada año. ¿Quiere decir esto que hay 600.000 nuevos autónomos cada año debido a que se ha puesto en marcha la tarifa plana? En absoluto. El resultado de la política es que, efectivamente, 600.000 personas se beneficiaron cada año de una reducción o una bonificación en las cotizaciones a la Seguridad Social. En torno al 10% de los autónomos en alta en 2018 eran beneficiarios del programa, un 45% de las altas en dicho año.

Pero, en términos de impacto, es decir, en términos de aumento en la probabilidad de que los desempleados se dieran de alta como autónomos, AIReF (2020) señala que solo se encuentra un efecto positivo en personas desempleadas durante periodos muy cortos y con un empleo temporal previo. No hay ningún efecto en personas desempleadas durante más de un mes, ni en las personas desempleadas con bajo nivel educativo. Es decir, estos colectivos se hubieran dado de alta como autónomos, aunque no se hubiera puesto en marcha la tarifa plana de autoempleo. El programa no ha modificado su comportamiento, no hay relación causal entre haber recibido la tarifa plana y decidir iniciar un negocio. Por tanto, el programa presenta un elevado efecto inercia o peso muerto.

La tarifa plana ha tenido mucho éxito en términos de participación, pero su impacto sobre el empleo autónomo en España es, en el mejor de los casos, pequeño. En términos de análisis coste-beneficio, este resultado implica que la mayor parte de los 600 millones anuales mencionados en el párrafo anterior no han tenido ningún impacto sobre el mercado de trabajo y sería mejor haberlos dedicado a otro tipo de medidas que sí tuvieran efectos.

Los indicadores de los PAPE muestran que evaluar, en sentido amplio, las políticas activas de empleo, no es tarea fácil. La heterogeneidad de las medidas –incluso de aquellas que se enmarcan en un mismo concepto, como puede ser formación– es amplia en términos de colectivo al que van dirigidas, duración, requisitos de participación, entre otras características. Se pueden utilizar múltiples indicadores, cada uno de los cuales tendrá ventajas e inconvenientes. En el caso de las evaluaciones de impacto, se añade la dificultad para construir un grupo de control adecuado. Esta dificultad aumenta si, como es habitual en el caso español, las evaluaciones se plantean *a posteriori* y sin haber estado previstas en la fase de diseño y/o implementación de las políticas. La evaluación de las políticas activas que llevó a cabo la AIReF en 2018 es un buen ejemplo.

Este informe puso de manifiesto la dispersión de los datos y que, a pesar de que existe un catálogo de medidas, bajo una misma denominación pueden coexistir programas distintos. «La falta de información detallada sobre las características y requisitos de los servicios en las comunidades autónomas que se traducen a códigos comunes del SISPE» (AIReF, 2019, p. 75) fue uno de los factores que más dificultó el análisis, ya que hace más compleja la evaluación en, al menos, dos aspectos. Por una parte, si no se conocen las características de los programas a evaluar con detalle ni cuáles son los criterios a partir de los que se decide qué personas participan y

cuáles no, la construcción de un grupo de control adecuado es muy compleja. La calidad del grupo de comparación depende de que podamos establecer con claridad por qué unas personas participaron en las políticas activas y otras no, ya que esto ayuda a explicar los sesgos de selección existentes. Por otra parte, la interpretación de los resultados y las posibles recomendaciones son difíciles de hacer si no se conocen los criterios con los que se ponen en marcha los programas y cómo se desarrollan. En definitiva, para llevar a cabo una evaluación de impacto es necesaria una «descripción detallada del contenido de cada una de las actuaciones» a analizar (Felgueroso *et al.*, 2018, p. 187).

Los objetivos que se plantearon inicialmente en la evaluación de la AIReF eran muy ambiciosos, puesto que evaluar medidas que se ejecutan a nivel autonómico requiere disponer de un sistema de información con una coordinación muy eficiente. La realidad mostró que los datos disponibles no permitían llevar a cabo esta evaluación, como se verá más adelante.

Sin embargo, aunque no es posible disponer de esa información homogénea a nivel nacional, sí que hay comunidades autónomas en las que se están llevando a cabo evaluaciones con grupo de control. Casi todas ellas, en mayor o menor medida y con mayor o menor difusión, han llevado a cabo evaluaciones de impacto puntuales o las han institucionalizado en planes de evaluación, muchas de las cuales pueden consultarse en las webs de los servicios públicos de empleo autonómicos. Utilizar esas evaluaciones es un elemento imprescindible en el desarrollo de la evaluación en nuestro país.

3. La evaluación en el marco del Fondo Social Europeo

Uno de los elementos que ha impulsado la evaluación de las políticas de empleo en España durante los últimos años ha sido la percepción de fondos europeos por parte de las Administraciones públicas. La cofinanciación de la UE se ha canalizado a través de programas operativos del Fondo Social Europeo, de ámbito tanto regional como plurirregional. Durante el periodo de financiación 2014-2020, los grandes programas plurirregionales han sido los de Empleo, Formación y Educación; Inclusión Social y Economía Social; y Empleo Juvenil. A ellos se les unen 19 programas operativos regionales, correspondientes a las 17 comunidades autónomas y a Ceuta y Melilla.

Las políticas europeas traen consigo protocolos de ejecución y de evaluación. El Reglamento 1303/2013 de la Unión Europea establece la obligación de presentar a la Comisión Europea informes anuales de ejecución de los programas operativos, que son aprobados previamente por los respectivos comités de seguimiento. Respecto a la evaluación, cada programa tiene su plan de evaluación específico, que se enmarca en el Plan Estratégico Nacional de Evaluación del FSE 2014-2020, y que incorpora las directrices establecidas por la Unión Europea.

Entre las exigencias europeas para la evaluación de sus programas figura que las entidades evaluadoras sean externas e independientes, lo que ha dado lugar a contrataciones con consultoras y, en contadas ocasiones, a colaboraciones contractuales con universidades. Las evaluaciones se realizan de forma periódica, aunque no todos los años. La mecánica de las evaluaciones del FSE consiste en que cada Administración gestiona la contratación de las actividades de evaluación de los programas bajo su competencia, ofreciendo los datos necesarios y exigiendo la calidad que considere oportuna, dentro de los márgenes establecidos por el plan de evaluación, por la Unidad Administradora del FSE en España y por la normativa europea.

Por tanto, el modelo de evaluación que impulsan los programas europeos es claramente descentralizado, lo que da lugar a una variedad y diversidad de evaluaciones. La existencia de los planes de evaluación garantiza que haya una mínima comparabilidad entre las evaluaciones de los diferentes programas. El carácter descentralizado facilita que la calidad de los datos utilizados en las evaluaciones pueda ser elevada, y que los problemas de gestión y de interpretación de la información, que son normales en toda evaluación, puedan ser canalizados adecuadamente entre la administración gestora y la entidad evaluadora.

Las directrices europeas respecto a la evaluación de actividades del FSE, además de establecer unos mínimos que permitan la ya mencionada comparabilidad, incluyen otras orientaciones que, sin ser de obligado cumplimiento, marcan pautas deseables para la evaluación. Así, durante los últimos periodos de financiación, la Comisión Europea ha ido incluyendo progresivamente orientaciones para que la parte de las evaluaciones dedicada al impacto incorpore enfoques contrafactuales (Comisión Europea, 2013; 2015). En este sentido, el enfoque descentralizado de la evaluación facilita que cada Administración decida la conveniencia de incluir este enfoque para alguna de las medidas realizadas, en función del tipo de medida, de la disponibilidad de datos, del horizonte temporal de la evaluación y de sus propios objetivos, entre los cuales tampoco cabe excluir la oportunidad política.

Por consiguiente, el hecho de que la Unión Europea establezca los contenidos mínimos, pero no los máximos, para la evaluación de las políticas de empleo que contribuye a financiar, significa una de las grandes ventajas del modelo de evaluación descentralizada: dar cabida a la iniciativa de las Administraciones gestoras en materia de evaluación. Además, como queda dicho, la Comisión valora las evaluaciones de impacto con enfoque contrafactual⁸.

Este planteamiento contribuye a explicar la aparición, durante los últimos años, de un número creciente de evaluaciones de tipo contrafactual referidas a programas europeos en España. Por ejemplo, a escala estatal, en 2018 se llevó a cabo la segunda evaluación de la Iniciativa de Empleo Juvenil, que incluyó por vez primera un análisis del impacto de las medidas y un enfoque contrafactual (Ministerio de Trabajo, Migraciones y Seguridad Social, 2018).

⁸ Buena prueba del interés de la Comisión Europea en intensificar las evaluaciones de impacto es el Informe publicado en julio de 2019, que analiza la posibilidad de realizar este tipo de estudios en distintos países, entre ellos, España (Comisión Europea, 2019).

Independientemente de la evaluación que el FSE requiere, por lo que respecta a las comunidades autónomas, la proliferación de evaluaciones con contrafactual es notable, aunque no alcanza a todas ellas, como se observa en el Cuadro 1. Dos comunidades autónomas (Cataluña y Navarra) tienen planes propios de evaluación de las políticas activas, en los que se contempla expresamente la evaluación con grupo de control. Además, esta proliferación coincide con el hecho de que, en comunidades como Cataluña, la evaluación de impacto de las políticas de empleo es especialmente intensa. También en Madrid, anualmente, se evalúa el impacto de las políticas activas. En otros casos, las iniciativas no tienen una periodicidad establecida, pero sí se ha realizado alguna evaluación de impacto. A su vez, esta intensidad parece independiente de la existencia de planes autonómicos de evaluación diferentes a los ya citados, que vienen impuestos por el FSE (Cuadro 1).

En definitiva, la evaluación de políticas de empleo en el marco del FSE, al ser descentralizada, remite al hecho de que algunas Administraciones van más allá de los mínimos establecidos por la UE, desarrollando sus propias evaluaciones de impacto, con o sin plan regional de evaluación. A su vez, la evaluación descentralizada de programas del FSE bajo un paraguas común, pero flexible, establece el ámbito de la evaluación al mismo nivel del ámbito de la gestión, mostrando, por tanto, un elemento de contraste muy claro con el enfoque del PAPE que se presentó en la sección anterior. Finalmente, este marco facilita la iniciativa de los gestores, lo que a su vez podría favorecer la difusión de buenas prácticas de evaluación.

4. Datos y fuentes de información para la evaluación

No hay duda de que todos los estamentos implicados en la elaboración y ejecución de las PAE coinciden en considerar que la evaluación de estas políticas es imprescindible para determinar cuál es su impacto y el grado de cumplimiento de los objetivos que persiguen, siendo el mejor camino para medir su eficiencia y eficacia.

En España, la práctica de la evaluación no se ha institucionalizado globalmente, de manera que, aunque recientemente se han llevado a cabo diversos estudios de investigación, sobre todo de carácter académico y con técnicas cuasiexperimentales, solo se dispone de conclusiones parciales y dispersas, que no ayudan a la toma de decisiones.

El modelo ideal para poder realizar una evaluación es aquel que permite identificar qué parte de los resultados obtenidos se deben única y exclusivamente al hecho de haber recibido algún «tratamiento» por parte de una política activa, es decir, que permite establecer una relación de causalidad entre el programa y sus resultados⁹.

⁹ Un magnífico ejemplo de la relevancia de los análisis basados en las relaciones de causalidad entre sucesos es la reciente concesión del premio Nobel de Economía a David Card, Joshua Angrist y Guido Imbens por su contribución al desarrollo de instrumentos analíticos para identificar qué programas pueden considerarse causantes, y en qué grado, de los resultados observados.

**CUADRO 1
PANORAMA DE LAS EVALUACIONES DE POLÍTICAS ACTIVAS DE EMPLEO EN LAS CC. AA.**

Comunidad autónoma	Plan propio de evaluación*	Alguna evaluación de impacto (2017-2021)	Alguna evaluación de impacto (2017-2021) con enfoque contrafactual
Andalucía	No	Sí	Evaluación específica sobre la Iniciativa de Cooperación Social y Comunitaria [Programa Empleo 30+]
Aragón	No		
Principado de Asturias	No	Sí	Informe de evaluación. PO FSE Asturias 2014-2020 (2019) Informe de evaluación. PO FSE Asturias 2014-2020 (2017)
Baleares	No	Sí	Evaluación de los objetivos/resultados del PO FSE de Baleares (2019) Evaluación de impacto del Programa VISIBLES dirigido a desempleados maduros (2018) Evaluación de políticas activas de ocupación: los programas MIXTS de formación y ocupación (2018)
Comunidad Valenciana	No	Sí	Informe de evaluación de los objetivos/resultados del Programa Operativo FSE 2014-2020 de la Comunidad Valenciana (2019)
Canarias	No	Sí	Evaluación de los objetivos/resultados del PO FSE 2014-2020 de Canarias (2019)
Cantabria	No	No	
Castilla-La Mancha	No	No	
Castilla y León	No	No	

NOTA: * Excluyendo los planes de evaluación específicos de los Programas Operativos del FSE.

FUENTE: Elaboración propia.

CUADRO 1 (Continuación)
PANORAMA DE LAS EVALUACIONES DE POLÍTICAS ACTIVAS DE EMPLEO EN LAS CC. AA.

Comunidad autónoma	Plan propio de evaluación*	Alguna evaluación de impacto (2017-2021)	Alguna evaluación de impacto (2017-2021) con enfoque contrafactual
Cataluña	Sí	Sí	Evaluación de impacto del programa Nuevas Oportunidades de Empleo 2015-2017 (2020) Evaluación del programa de formación en Áreas prioritarias (FOAP), ejecución 2016-2019 (2020) Evaluación de diseño, implementación e impacto del programa de formación de los CIFO, ejecución 2016-2019 (2020) Informe comparativo de los programas de formación destinados a personas en situación de paro, ejecución 2016-2019 (2020) Evaluación de impacto del programa Trabajo y Formación – Línea parados de larga duración. Convocatoria 2016 (2020) Evaluación de impacto del programa Trabajo y Formación (convocatoria 2015) en el marco del PO FSE Cataluña 2014-2020 (2017) Evaluación de implementación e impacto del programa TLN Mobilicat, convocatoria 2015 (2018)
Extremadura	No	Sí	Evaluación de los objetivos/resultados del PO FSE 2014-2020 de la comunidad autónoma de Extremadura (2019)
Galicia	No	Sí	Evaluación de los objetivos/resultados de las prioridades del PO FSE Galicia 2014-2020 (2019)
La Rioja	No	No	
Madrid	No	Sí	Evaluación del Impacto Individual de Políticas Activas de Empleo 2018 (2019)
Murcia	No	Sí	Informe de evaluación de los objetivos/resultados del Programa Operativo FSE 2014-2020 Región de Murcia (2019)
Navarra	Sí	No	
País Vasco	No	Sí	Informe de evaluación de los objetivos/resultados del Programa Operativo FSE 2014- 2020 del País Vasco (2019)

NOTA: * Excluyendo los planes de evaluación específicos de los Programas Operativos del FSE.

FUENTE: Elaboración propia.

Este tipo de procedimiento no es sencillo, pero cabe la posibilidad de que se diseñe una determinada política de manera que se pueda poner en marcha un experimento previo. Por ejemplo, mediante experiencias piloto que permitan, de manera previa al lanzamiento de un programa, medir la eficacia de dichas medidas. En este sentido, los experimentos en cuestiones relacionadas con la pobreza y el desarrollo, de Duflo y Banerjee constituyen una referencia (Banerjee & Duflo, 2009; Duflo, 2017), si bien estas aplicaciones estuvieron precedidas de otras muchas en el ámbito de la economía laboral y, específicamente, en el de las políticas activas de mercado de trabajo (véase Heckman, Lalonde & Smith, 1999). Desafortunadamente, en Europa, las evaluaciones experimentales en este campo son escasas y se concentran principalmente en los países nórdicos¹⁰.

La dificultad de realizar este tipo de evaluaciones, que debe ser prevista en el momento de diseño de la política, puede ser solventada en ocasiones gracias a la existencia de metodologías estadísticas conocidas como cuasiexperimentales, que utilizan los datos disponibles, como son los procedentes de registros administrativos, para extraer información, que permiten interpretar los resultados como si fuesen fruto de la comparación de los beneficiarios de una política con un grupo de control definido mediante un experimento. No obstante, poder desarrollar estas metodologías adecuadamente requiere de un conocimiento detallado del significado y recolección de los datos para que los resultados del análisis no estén invalidados debido a los sesgos introducidos por otras variables que escapan al control del investigador.

En España, la ausencia de evidencia experimental es una grave limitación, porque la realización de experimentos permite conocer el impacto de las políticas activas contando con información de calidad recogida explícitamente para ello. Las evaluaciones cuasiexperimentales están concentradas en muy pocas comunidades autónomas, siendo conveniente que se extiendan a todos los niveles, lo que podría ser impulsado desde el SEPE, quien además podría coordinarlas a través de la organización administrativa existente de los SPE. El coste sería reducido, pues los datos ya existen y lo único que se necesita es una extracción correcta de las bases de datos disponibles y un análisis que, hoy en día, es mucho más fácil de acometer gracias a diferentes paquetes estadísticos (Malo & Cueto, 2015). Los investigadores d'Hombres y Santangelo (2019) presentan las ventajas del uso de los datos administrativos para hacer evaluación de impacto de manera sistemática, porque proporcionan gran cantidad de datos y de información con un coste de utilización muy bajo, en comparación con el coste derivado de la realización de encuestas. Destaca, sin embargo, que España no aparece en el grupo de países que tienen publicados análisis de evaluación utilizando modelos contrafactuales.

Ante esta evidente escasez de procesos de evaluación adecuados y completos, no podemos dejar de llamar la atención del que podría ser el principal motivo que lo explica.

¹⁰ Véase, por ejemplo, el metaanálisis realizado por Card, Kluve y Weber (2018).

La posibilidad de poner en práctica estos procesos de evaluación en España está condicionada por la disponibilidad de datos adecuados para tal fin. No se puede afirmar que no haya información, sino que es complicado acceder a ella y, en ocasiones, es incompleta. Además, incluso cuando existe y está disponible, resulta difícil manejarla para analizar lo que se está buscando.

Partiendo del supuesto de que las PAE dejan claramente definidos, estructurados y concretados los objetivos que persiguen, las fuentes de datos utilizadas en la evaluación han de contener información suficiente para establecer cuál es la variable que permite medir el resultado, así como para aislar qué efectos se han debido exclusivamente al tratamiento recibido por la PAE, de aquellos que se alcanzan con independencia de ella. Como ya se ha mencionado anteriormente, resulta imprescindible llegar a identificar a la población que recibe el tratamiento de la PAE y diseñar un grupo de control que tenga idénticas características que el grupo beneficiario, pero que no ha recibido ningún tipo de tratamiento.

Por otra parte, una vez que estén bien definidos y validados los elementos y cuestiones que han de ser analizados, hay que identificar los criterios para la estandarización de conceptos y los indicadores que permitan hacer el seguimiento y la posterior evaluación. En este proceso, es importante identificar a los interlocutores y su intervención en la oferta y puesta en marcha de los servicios y programas. El conocimiento referido al detalle de la implementación y el diseño de la evaluación requieren de una estrategia que ayude a recopilar toda la información necesaria para su realización. No solo es necesario definir la idoneidad de la estrategia propuesta en relación con las necesidades del mercado de trabajo, sino el diseño del procedimiento que se va a utilizar en la posterior evaluación.

A modo de ejemplo, se puede considerar que una variable resultado puede ser en qué medida mejora la empleabilidad o el desempeño laboral de un individuo tras ser tratado por una PAE. Ahora bien, medir el grado de empleabilidad o el desempeño laboral no es tarea fácil¹¹, hay que poder determinar qué se entiende y cómo se mide, porque no solo depende de que se logre salir del desempleo, sino del tiempo que se tarda en lograrlo, de la proporción de tiempo que el individuo está empleado, el nivel de ganancias acumuladas, el tiempo transcurrido hasta el primer empleo, del tipo de contrato con el que se accede, de la duración de dicho contrato, del nivel de cualificación y su adecuación con el empleo, entre otras.

4.1. El encaje temporal de la evaluación

Si la evaluación está prevista desde la fase de diseño de la política, debe decidirse la metodología a emplear y, con ello, también el protocolo de recogida de información en la que se basará la evaluación. Como ya se ha mencionado, una vez

¹¹ Toharia *et al.* (2006) analizan con detalle de qué manera hay que tratar los datos registrados en las demandas de empleo en el SEPE para poder establecer alguna medida de la empleabilidad.

que estén bien definidos y validados los elementos y las cuestiones que han de ser analizadas, hay que identificar los criterios para la estandarización de conceptos y los indicadores que permitan hacer el seguimiento y la posterior evaluación. En este proceso es importante identificar a los interlocutores y su intervención en la oferta y puesta en marcha de los servicios y programas. El conocimiento referido al detalle de la implementación y el diseño de la evaluación requieren de una estrategia que ayude a recopilar toda la información necesaria para su realización. No solo es necesario definir la idoneidad de la estrategia propuesta en relación con las necesidades del mercado de trabajo, sino el diseño del procedimiento que se va a utilizar en la posterior evaluación. En este sentido, el papel de los agentes participantes, de las oficinas de empleo y, por ejemplo, de los centros de formación resulta muy relevante para poder tener el control de las medidas de eficiencia y eficacia de los servicios y programas prestados.

Lamentablemente, la decisión sobre evaluar PAE se toma, en muchas ocasiones, cuando la política ya se ha implementado y, así, gran parte de las dificultades para evaluar reside en que la recogida de datos no ha sido adecuada a la necesidad del programa a evaluar. Estas dificultades se pusieron claramente de manifiesto en la primera evaluación que intentó hacer la AIReF, donde se indica que «los datos suministrados han sido insuficientes y de una calidad mejorable. La evaluación completa de la eficacia de las políticas activas por comunidad autónoma no ha sido posible» (AIReF, 2019, p. 6).

Los objetivos que se planteaban inicialmente eran muy ambiciosos. Sin embargo, finalmente, se evaluaron únicamente dos programas: medidas de diagnóstico individualizado en Aragón y el PREPARA, cuya información era recogida por el SEPE a nivel nacional. Emplearemos el plan PREPARA para mostrar la dificultad de configurar un grupo de control adecuado, en la línea de lo planteado anteriormente.

El plan PREPARA entró en vigor en febrero de 2011¹², con el fin de apoyar la cualificación profesional de las personas que agotaban su protección por desempleo, mediante acciones de políticas activas de empleo y la percepción de una ayuda económica de apoyo. Los requisitos para recibir la ayuda variaron a lo largo del tiempo¹³, lo que ya es una dificultad en sí misma, dado que los grupos de tratamiento y de control deben tener en cuenta los cambios realizados. Por otra parte, se disponía de muy poca información sobre los servicios ofrecidos a los beneficiarios. En el informe de la AIReF se indica que no se incluye en el análisis la duración de dichos servicios por la mala calidad del dato. Dada la naturaleza del programa y su apuesta por combinar la ayuda con acciones de políticas activas, estas últimas eran un aspecto clave en la evaluación. Es más, la comparación entre grupos que recibieran

¹² Real Decreto-ley 1/2011, de 11 de febrero, de medidas urgentes para promover la transición al empleo estable y la recualificación profesional de las personas desempleadas.

¹³ Real Decreto-ley 14/2017, de 6 de octubre, por el que se aprueba la reactivación extraordinaria y por tiempo limitado del programa de recualificación profesional de las personas que agoten su protección por desempleo.

distintas medidas (orientación, formación, u otras) podría haber aportado resultados muy interesantes sobre qué tipo de programas podrían ser más útiles para la inserción laboral de los beneficiarios.

En el análisis del PREPARA se incide en que genera desincentivos a incorporarse al mercado laboral en los seis meses que dura la prestación. Aquí, podemos señalar la importancia de los aspectos subjetivos de valoración de resultados. Dado que el colectivo destinatario está formado por personas desempleadas de larga duración, principalmente con baja cualificación, ¿podemos esperar resultados positivos a corto plazo en una coyuntura económica de crisis? Por otra parte, imaginemos que las personas están participando en actividades formativas y que, dado que reciben una ayuda, prefieren seguir formándose a aceptar un empleo que, probablemente, sea de muy corta duración. ¿Este desincentivo lo calificamos como «negativo»? Por supuesto, lo que estamos planteando son hipótesis que, con datos de calidad, podrían contrastarse.

Por último, hay que señalar las diferencias que se encontraron entre comunidades autónomas que aportan información para el aprendizaje mutuo de gran interés. Los resultados transcurridos dos años mostraban que, en tres casos, el impacto era negativo, en dos casos positivo y, en el resto, nulo. Además de vincular esta información con la diferente situación de los mercados de trabajo regionales, la explicación también podría estar vinculada con la distinta supervisión de la búsqueda activa de empleo o las distintas medidas en las que se participaba. Desafortunadamente, la información para seguir ahondando en los resultados del programa no estaba disponible.

Respecto a las medidas de diagnóstico individualizado en Aragón, son un buen ejemplo de qué se puede hacer si se cuenta con información adecuada. El informe de la AIReF disponía de información sobre qué personas habían recibido servicios tras el diagnóstico individualizado. Los datos permitían caracterizar a las personas beneficiarias en términos de trayectoria laboral y acciones del servicio público de empleo en las que habían participado. En definitiva, la calidad de la información sobre ese programa permitió llevar a cabo la evaluación de impacto.

4.2. El potencial de los registros administrativos

Los datos administrativos procedentes de los registros de la Seguridad Social o del Servicio Público de Empleo son una importante fuente de información con un coste relativamente bajo que, además de contener una cantidad elevada de información, permiten llegar a unidades de análisis territoriales pequeñas. Su relevancia es tal que estos datos administrativos podrían ser elevados a la consideración de «bien público». Si fuese fácil y sencilla su disponibilidad para las personas residentes en un país, habría un mayor y mejor conocimiento de la sociedad, además de proporcionar la tan demandada e imprescindible transparencia en favor de la democracia.

Estas bases de datos administrativas han sido utilizadas previamente por investigadores que han evaluado políticas de empleo, principalmente incentivos a la

contratación, y que han puesto de manifiesto sus fortalezas y debilidades como fuente de información¹⁴.

Desde 2005, España dispone del Sistema de Información de los Servicios Públicos de Empleo (SISPE), en el cual se recoge importante información sobre los demandantes de empleo que están inscritos en las oficinas públicas de empleo. Asimismo, se incorpora información sobre los Servicios y Programas comunes prestados por el Estado y propios de las comunidades autónomas. Y también integra el historial de servicios recibidos por las personas inscritas en algún momento y se dispone de información referida a determinados programas, como el PREPARA. Por otro lado, el sistema incluye toda la información recogida en el registro de contratos, o en los sistemas que inscriben a los perceptores de prestaciones por desempleo.

La evaluación de la eficacia de las políticas activas de empleo parte de la selección de las personas inscritas en el SEPE que son demandantes de empleo que realmente quieren trabajar, porque hay demandantes inscritos por otras razones relacionadas con los posibles derechos que el registro puede reportarles tanto en el corto, como en el medio y largo plazo. De ahí que dentro del SISPE se clasifique a los colectivos inscritos en función de que realmente estén buscando un puesto de trabajo porque quieren trabajar. El problema es que, seguramente, puede haber personas clasificadas como paradas que en realidad no se encuentren disponibles para ocupar un empleo. Además, la disponibilidad para trabajar es una cuestión que no puede ser observada, que depende de las personas, sus actitudes y sus condiciones, lo que sitúa a la disponibilidad en un plano subjetivo de muy difícil medición. De ahí que haya ciertos problemas con este tipo de datos, que sin lugar a dudas es inmensamente rico e interesante, que levantan preocupaciones y obligan a ser cautos tanto en su uso como explotación e interpretación.

Además, hay que añadir las numerosas reticencias y temores por parte de los agentes gestores de los registros ante la necesidad de guardar la confidencialidad que, en ocasiones, se debe más a un exceso de celo que a un problema real, pues podrían resolverse, y de hecho así se hace en otras fuentes, con la debida anonimización de las bases de datos y la agregación de determinadas categorías.

Por otra parte, el proceso de solicitud es complicado, y se complica más cuando resulta imprescindible la unión de diversas bases de datos administrativos, puesto que hay que implicar a diversos gestores que, incluso perteneciendo a la misma institución, utilizan usos y costumbres dispares.

Otra magnífica fuente de información que pone a disposición de los investigadores datos de la Tesorería General de la Seguridad Social es la Muestra Continua de Vidas Laborales (MCVL). Esta base de datos es resultado de una operación estadística emprendida por la Dirección General de Ordenación de la Seguridad Social, con el objetivo de explotar la enorme riqueza de información que existe en los registros

¹⁴ Véase, por ejemplo, Cebrián *et al.* (2011), trabajo en el que colabora Luis Toharia, investigador con una participación decisiva para la puesta en marcha de la renovación del sistema de registros del Servicio Público de Empleo, así como pionero en enlazar diferentes fuentes de registros administrativos.

informatizados de la Seguridad Social, no solo desde el punto de vista de la afiliación durante los periodos de empleo, sino también en lo referente a las prestaciones sociales recibidas¹⁵.

El primer año de su publicación fue 2004, pero contiene el número de episodios de afiliación iniciados en cada año, ya sea por empleo por cuenta ajena o propia, por una prestación por desempleo o por alguna pensión. Además, proporciona todo el historial de la vida laboral que haya registrado para cada una de las personas que forman parte de la muestra, lo que supone una de sus principales ventajas al permitir obtener datos longitudinales para el seguimiento de la trayectoria laboral de las personas de la muestra y así tener un panel con los historiales de todos los participantes. Ahora bien, recientemente se ha incorporado a la solicitud de dicha base el requisito de que los datos sean destruidos anualmente, lo que, si no se modifica, hará inviable su uso para el análisis longitudinal.

En algunas ocasiones, ha sido necesario solicitar el enlace de diferentes fuentes administrativas, como las de los registros del SEPE y los registros de la Seguridad Social de la MCVL¹⁶. En todas ellas el procedimiento ha sido largo y tedioso y, a veces, casi imposible, aunque también es cierto que ha ido mejorando con la práctica y experiencia, hasta tal punto que, en ocasiones, se oyen voces comentando el interés de las partes en que el enlace se convierta en algo habitual. Esto evitaría que una vez finalizado el estudio, los datos hayan de ser destruidos impidiendo que se vuelva a utilizar dicho enlace.

Un buen ejemplo de enlace de datos administrativos lo podemos encontrar en la unión del fichero de contratos registrados en el SEPE y la MCVL. El resultado que se obtiene es uno de los muchos ejemplos que confirman lo conveniente que es la colaboración entre Administraciones para poder llevar a cabo cualquier análisis de evaluación. La principal ventaja de este enlace es que permite complementar la información de ambas bases de información. Por un lado, el fichero de contratos registrados no permite identificar la duración efectiva del contrato, solo el momento en el que se inicia y, en todo caso, se podría conocer si el individuo tuvo con anterioridad o posterioridad otro contrato, de modo que se podría calcular el total de contratos firmados a lo largo de un periodo, pero en ningún caso se podrá saber el tiempo transcurrido en cada tipo de contrato ni entre la finalización de uno y el comienzo del otro. En cambio, esta información es posible conocerla con los datos aportados por la MCVL, ya que figura el momento del inicio de la relación laboral, el momento de finalización, si es que lo hay, e incluso se puede identificar si el individuo ha cambiado de contrato dentro de la misma relación laboral, pudiendo deducirse el tiempo transcurrido desde el inicio de dicha relación y el tiempo pasado en los diferentes contratos que se hayan ido encadenando a lo largo de ella, e incluso el tiempo fuera del sistema

¹⁵ No se encuentran incluidos los demandantes de empleo cuando no reciben prestaciones y los inactivos (distintos de los pensionistas), además de los trabajadores que tienen un sistema de previsión social distinto de la Seguridad Social (los funcionarios de Clases Pasivas) o no tienen ninguno (como quienes trabajan en la economía informal).

¹⁶ Cebrián, Moreno y Toharia (2011a); Cebrián, Moreno y Toharia (2011b); Cebrián y Moreno (2012).

de afiliación. Sin embargo, la MCVL, a pesar de incorporar datos procedentes del padrón aportados por el INE, proporciona una información sobre el nivel de estudios que no tiene por qué corresponder con la del individuo a la firma de cada contrato que haya podido tener a lo largo de su vida laboral. En cambio, dicha información sí se encuentra disponible en el registro de contratos. Lo mismo ocurre con el tamaño de la plantilla, ya que dicha información se registra en la MCVL según consta en la Seguridad Social en el momento en el que se extrae la muestra de referencia, no cuando se registra el contrato. Además, el fichero de contratos contiene información detallada sobre el tipo de ocupación a la que se refiere el contrato, en cambio, este detalle no se recoge en la Seguridad Social. Por ello, gracias al enlace se combina toda la información de la Seguridad Social con la aportada por el registro de contratos. Por otro lado, el registro de contratos aporta mucho detalle sobre la normativa que regula la contratación, si el contrato está sujeto a algún tipo de bonificación o deducción financiada por el Servicio Público de Empleo, así como el tipo de indemnización por despido improcedente que, antes de la reforma laboral de 2012, podía ser de 33 días por año trabajado con un máximo de 24 mensualidades o de 45 días por año trabajado con un máximo de 42 mensualidades, en función de las características del contrato.

El principal inconveniente es el propio proceso de enlace. Ya que requiere que las Administraciones competentes acuerden cuál será el criterio y procedimiento para realizar la unión y la anonimización del identificador de la persona física y del código de la cuenta de cotización, que son las dos variables básicas para proceder al enlace. Sirva como otro ejemplo de la complejidad para llevar a cabo este enlace que el identificador de la persona física según el criterio que aplica la Seguridad Social y el utilizado por el fichero de contratos son diferentes, de manera que resulta prácticamente imposible identificar adecuadamente a los trabajadores de nacionalidad extranjera por ser diferente el criterio de identificación seguido por la Seguridad Social y por el SEPE. Esta circunstancia impide que se identifiquen adecuadamente aquellos contratos de la población extranjera que forma parte de la MCVL. Por ello, en ocasiones, el estudio se centra exclusivamente en aquellos individuos que en el momento de firmar un contrato poseen la nacionalidad española, con independencia de cuál fuese su país de nacimiento. Evidentemente, esto puede dar lugar a que se elimine casi a un quinto del total de contratos y afiliaciones.

En resumen, desde el punto de vista del análisis y la investigación, resulta absolutamente necesario poder utilizar los datos administrativos con la imprescindible colaboración e implicación directa de sus gestores y propietarios, de manera que se garantice el acceso e interpretación adecuada de los conceptos y registros, elementos base para tener un buen conocimiento y entendimiento de los sistemas de información.

Para evitar retrasos y errores sería muy conveniente que hubiese unidades dedicadas en la planificación y coordinación de la evaluación de las políticas activas en las que estuvieran implicadas y colaboraran las diferentes unidades gestoras de las fuentes, así como las entidades encargadas de realizar las evaluaciones.

5. Conclusiones

En este trabajo se han repasado una serie de aspectos clave en relación con la evaluación del impacto de las políticas activas de empleo en España. A partir de la experiencia de los autores y de un análisis documental de la bibliografía disponible, se han puesto de manifiesto varias cuestiones, comenzando por la constatación de que la normativa en materia de PAE incluye a la evaluación como parte integrante de los Planes Anuales de Política de Empleo, si bien la concreción de los indicadores que abarcan los PAPE resultan insuficientes para captar los verdaderos efectos de las políticas que se llevan a cabo. En este sentido, se han expuesto las carencias de los PAPE en comparación con la necesidad de evaluar el impacto causado por las medidas, disociándolo con claridad de la influencia de otros factores, como el ciclo económico o las diferencias entre diversos mercados de trabajo regionales. Se han puesto de manifiesto, como ejemplo, los significativos efectos de inercia de la tarifa plana para trabajadores autónomos.

A su vez, se ha analizado la evaluación de las políticas activas llevada a cabo por la AIREF en 2019, y los problemas destapados a partir de la misma en relación con las diferencias en las pautas de registro de información por parte de las comunidades autónomas.

Los problemas derivados del marco normativo estatal y de la falta de homogeneidad de la información se han contrastado con el mecanismo de evaluación que utiliza el Fondo Social Europeo, que hace recaer sobre las Administraciones gestoras, ya sean nacionales o regionales, las tareas de evaluación, sometidas a ciertos principios comunes. Se han señalado las ventajas e inconvenientes de este modelo descentralizado, aportando evidencia del incremento creciente del número de evaluaciones de impacto con enfoque contrafactual durante los últimos años.

Adicionalmente, en el artículo se ha abordado el contraste entre la falta de una institucionalización de la evaluación y la existencia de numerosos trabajos académicos que aportan evidencias específicas sobre el impacto de las PAE. No obstante, también cabe señalar que no se ha realizado ninguna evaluación de tipo experimental que, a pesar de tener un coste elevado, presenta ventajas derivadas de poder tener acceso a información ajustada al objetivo de la evaluación mediante encuestas *ad hoc*. Por el contrario, los datos administrativos, aunque su objetivo no es facilitar la evaluación, constituyen una magnífica fuente de información, cuyo coste es más reducido y permiten acceder a unidades de análisis muy desagregadas, incluso a la propia población tratada.

A este respecto, la disponibilidad de datos de calidad resulta crucial, lo que remite a los procedimientos de recogida y tratamiento de la información. Se ha tomado como muestra de estos problemas el hecho de que la evaluación de eficacia de las PAE hecha por la AIREF y citada anteriormente, que ambicionaba un ámbito estatal, hubiese de conformarse en cuanto a análisis de impacto causal con el análisis de un solo programa estatal (PREPARA) y otro de una comunidad autónoma. En contraste, se han presentado las importantes ventajas de la Muestra Continua de Vidas Laborales, de la que también se han señalado algunos inconvenientes específicos.

El conjunto de análisis realizado permite establecer tres comentarios a modo de conclusión, a los que se unen algunas propuestas de mejora de la evaluación de las PAE. Entre los comentarios, primero, cabe subrayar el problema derivado de la falta de colaboración entre las Administraciones que son propietarias de la información y las entidades que podrían llevar a cabo evaluaciones de impacto. A este problema se le suma, en segundo lugar, que muchos datos que podrían publicarse no se publican. En tercer lugar, algo que quizá ayude a entender los dos puntos anteriores, cabe subrayar la falta de homogeneidad de los sistemas de recogida de información de las Administraciones públicas con competencias en la gestión de las PAE.

Entre las propuestas de mejora de la evaluación de políticas de empleo se pueden incluir, sin ánimo de exhaustividad, las siguientes: primero, dados los decepcionantes resultados de las iniciativas de evaluación de PAE a escala estatal, cabría avanzar hacia un modelo descentralizado inspirado en el que utiliza el FSE, donde la responsabilidad de la evaluación recaiga sobre las Administraciones responsables de la gestión de las medidas. Este modelo, a pesar de los problemas de comparabilidad de las evaluaciones, tiene indudables ventajas, visibles ya al aparecer cada vez más evaluaciones de impacto que avanzan conclusiones útiles sobre la causalidad de las medidas de empleo.

En segundo lugar, cabe defender que se avance también hacia un mayor compromiso por parte de las Administraciones públicas que deberían protagonizar o facilitar evaluaciones a escala estatal, como la Seguridad Social, el SEPE y los ministerios con competencias en PAE y en registro de información necesaria para la evaluación.

Por último, una propuesta muy específica que facilitaría la evaluación de las PAE es la posibilidad de vincular las diversas fuentes de datos generados por las Administraciones públicas (al menos Seguridad Social y Servicio Público de Empleo, pero también, Agencia Tributaria, INE), como ya se hace cuando la MCVL se enlaza con datos del INE y de la Agencia Tributaria. Su explotación por parte de las diversas Administraciones y la comunidad científica y académica, gracias a la digitalización y las cada vez más avanzadas tecnologías disponibles, proporcionarán la generación de un mayor y mejor conocimiento sobre el impacto de las PAE.

Por último, cabe mencionar que la aplicación estricta, y a veces muy restrictiva, tanto de la ley de protección de datos¹⁷, como del Reglamento General de Protección de Datos europeo para garantizar la seguridad de los datos de carácter personal que se incluyen en las fuentes de información, limitan notablemente la explotación y uso de estas fuentes por parte de investigadores externos a la Administración pública, lo que impide que se puedan llevar a cabo estudios de evaluación de las políticas públicas o replicar las evaluaciones ya realizadas como garantía de la propia evaluación.

Algunos países europeos con la misma legislación han arbitrado mecanismos de acceso a los datos para facilitar que personal investigador externo a la propia Administración pública pueda proceder a la realización de evaluaciones. Este podría ser

¹⁷ Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

el ejemplo de la iniciativa desarrollada por la Seguridad Social, el INE, la Agencia Tributaria y el Banco de España para poner en marcha un sistema de acceso a sus bases de datos con fines científicos. Se han puesto a disposición del personal investigador solamente dos salas de investigación en Madrid y en Barcelona, lo que ayuda a garantizar el secreto estadístico, la confidencialidad de los datos y la protección de los datos personales, pero quizá limite un acceso amplio.

Referencias bibliográficas

- AIReF, Autoridad Independiente de Responsabilidad Fiscal. (2019). *Evaluación Del Gasto Público 2018. Estudio Programa Políticas Activas de Empleo*.
- AIReF, Autoridad Independiente de Responsabilidad Fiscal. (2020). *Evaluación Del Gasto Público 2019. Estudio Incentivos a La Contratación*.
- Banerjee, A. V., & Duflo, E. (2009). The Experimental Approach to Development Economics. *Annual Review of Economics*, 1, 151-178.
<https://doi.org/10.1146/annurev.economics.050708.143235>
- Card, D., Kluve, J., & Weber, A. (2018). What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. *Journal of the European Economic Association*, 16(3), 894-931.
<https://doi.org/10.1093/jea/jvx028>
- Cebrián, I., Moreno, G., & Toharia, L. (2011a). *Evaluación microeconómica de los programas de fomento del empleo indefinido a partir de datos de los servicios públicos de empleo y de la Muestra Continua de Vidas Laborales*. Ministerio de Trabajo e Inmigración. ISBN 9788484173847.
- Cebrián, I., Moreno, G., & Toharia, L. (2011b). La estabilidad laboral y los programas de fomento de la contratación indefinida. *Hacienda Pública Española / Revista de Economía Pública*, 198, 103-127. ISSN 0210-1173.
- Cebrián, I., & Moreno, G. (2012). La estabilidad de los nuevos contratos indefinidos durante la crisis económica. *Estudios de Economía Aplicada*, 30(1), 183-208. ISSN 1133-3197.
- Comisión Europea. (2013). *Design and commissioning of counterfactual impact evaluations*.
- Comisión Europea. (2015). *Programming period 2014-2020. Monitoring and Evaluation of European Cohesion Policy. European Social Fund. Guidance document*.
- Comisión Europea. (2019). *Pilot and feasibility study on the sustainability and effectiveness of results for European Social Fund participants using counterfactual impact evaluations. Final Report*.
<https://op.europa.eu/es/publication-detail/-/publication/84cc9eb9-b33d-11e9-9d01-01aa75ed71a1>
- d'Hombres B., & Santangelo, G. (2019). Use of Administrative Data for Counterfactual Impact Evaluation of Active Labour Market Policies in Europe: Country and Time Comparisons. In N. Crato, & P. Paruolo (eds.), *Data-Driven Policy Impact Evaluation* (pp. 271-287). Springer. https://doi.org/10.1007/978-3-319-78461-8_17
- Duflo, E. (2017). The Economist as Plumber. *American Economic Review*, 107(5), 1-26.
<https://doi.org/10.1257/aer.p20171153>
- Felgueroso, F., García-Pérez, J. I., & Jiménez-Martín, S. (2018). *Perfilado Estadístico: Un Método Para Diseñar Políticas Activas de Empleo*. Fundación Ramón Areces. FEDEA.

- Heckman, J. J., Lalonde, R., & Smith, J. (1999). Chapter 31. The economics and econometrics of Active Labour Market Policies. *Handbook of Labor Economics*, 3(A), 1865-2097.
- Malo, M. A., & Cueto, B. (2015). El impacto de las políticas activas de mercado de trabajo en España. *Documentación social*, 178, 105-120.
- Mato, F. J. (2010). La formación en España desde una perspectiva comparada. Balance y propuestas de mejora. *Papeles de Economía Española*, 124, 266-280.
- Ministerio de Trabajo, Migraciones y Seguridad Social. (2018). *II Evaluación de la Iniciativa de Empleo Juvenil. Informe de Evaluación*. Subdirección General de Programación y Evaluación del Fondo Social Europeo.
- Toharia, L., Pérez Infante, J. I., & Prudencio, C. A. (2006). *La ocupabilidad de los parados registrados y la tasa de cobertura de las prestaciones por desempleo. Informe para el Servicio Público de Empleo Estatal*.
- Vila, J. E., Gómez, Y., Martínez, J., & Ferri, J. L. C. (2020). Evaluación de políticas públicas con métodos cuasi-experimentales: una aplicación a las políticas activas de empleo en la Comunidad Valenciana. *Cuadernos económicos de ICE*, 99, 11-24.