CUADERNO DE EVALUACIÓN DE IMPACTO

de las políticas activas de empleo para el SNE-NL

Diciembre 2021

servicio navarro de empleo - nafar lansare

Gobierno de Navarra
Nafarroako Gobernua
Departamento de Derechos Sociales
Eskubide Sozialetako Departamentua

sarhe CONSULTORES

INDICE

TABLA DE CONTENIDOS

1. PRESENTACIÓN 4

2. CONCEPTO DE EVALUACIÓN DE IMPACTO 6
   - a. Qué se entiende por una política de empleo con impacto 7
   - b. Evaluación de resultados vs. evaluación de impacto 10
   - c. Evaluación de proceso vs. evaluación de impacto 11
   - d. Enfoques principales de las evaluaciones de impacto 12
   - e. Oportunidad y utilidad de las evaluaciones de impacto 14

3. EFECTOS BRUTOS DE NETOS DE LAS POLÍTICAS ACTIVAS DE EMPLEO 15

4. LA EVALUACIÓN EXPERIMENTAL COMO REFERENCIA 18
   - a. La lógica de la evaluación experimental 19
   - b. El desarrollo progresivo de las evaluaciones experimentales 20
   - c. Tres mantras (i): el contrafactual 22
   - d. Tres mantras (ii): el sesgo de selección 24
   - e. Tres mantras (iii): la validez interna y externa 25

5. LOS EXPERIMENTOS NATURALES 26

6. OTRAS EVALUACIONES DE IMPACTO: CUASI- EXPERIMENTOS, EVALUACIONES MACROECONÓMICAS Y METAEVALUACIONES 29
   - a. Evaluaciones cuasiexperimentales 30
   - b. Evaluaciones macroeconómicas 37
   - c. Metaevaluaciones 38

7. QUÉ PUEDE HACER LA ADMINISTRACIÓN PARA FACILITAR LAS EVALUACIONES DE IMPACTO DE POLÍTICAS DE EMPLEO 33 40
   - a. La importancia de los datos 41
   - b. La selección y la participación en políticas de empleo 43
   - c. Procedimientos de recogida de datos durante y después de la PAE 46

8. LA INTEGRACIÓN DE LA EVALUACION DE IMPACTO EN EL PROCESO DE LA POLÍTICA 48

9. BIBLIOGRAFÍA 51

1. Presentación

Conocer el impacto que tienen las políticas públicas en general, y las políticas de empleo en particular, se está convirtiendo en un factor fundamental para la gestión pública eficaz. Al igual que ocurre con otros rasgos de las políticas públicas, el conocimiento de su impacto requiere, como condición necesaria, la voluntad política de llevar a cabo evaluaciones. Como condición suficiente, es preciso que dichas evaluaciones puedan desarrollarse de forma profesional, con información de calidad, y ello aconseja que todos los actores implicados estén familiarizados con la evaluación de impacto.

De los diversos retos que afrontan las sociedades avanzadas, uno no menor es el contraste entre la creciente capacidad de tratamiento de la información a bajo coste y gran velocidad, por una parte, y las exigencias que imponen los procedimientos de las Administraciones Públicas, por otra. En este contexto, la evaluación de las políticas públicas y, más concretamente, de su impacto sobre el entorno socioeconómico, constituye un ámbito de actuación fundamental para la mejora del desempeño de la Administración.

En un sistema mixto de mercado como el que caracteriza a nuestro entorno, la intervención del Estado en la economía está justificada en razones de eficiencia, equidad y estabilidad. Para garantizar que se cumple con estas razones y que la intervención mejora el desempeño económico, las Administraciones deben evitar rigideces. La creciente disponibilidad de información debe servir para ello, facilitando la gestión y la reacción de las políticas públicas a entornos cambiantes. Pero disponer de la información no es suficiente. Es preciso aprovechar la información para generar conocimiento. Una Administración moderna requiere la toma de decisiones basada, entre otros factores, en la evidencia. Integrar toda la información disponible sobre los efectos de las políticas públicas en los procesos de toma de decisión permitirá mejorar la calidad de las políticas públicas y fortalecer el funcionamiento de las sociedades democráticas.

La toma de decisiones basada en evidencias requiere conocer el impacto de las políticas pasadas. Por esta razón se necesita poner en marcha, o mantenerlos, si ya existiesen, procedimientos que permitan conocer el impacto futuro de las medidas que se toman en el presente. De ahí que el binomio entre información y conocimiento pueda y deba completarse con un tercer elemento: la cultura de la evaluación. Solo cuando los agentes involucrados compartan una valoración positiva de la evaluación de impacto será posible impulsar y aceptar los resultados de la misma, con independencia de que éstos generen más o menos satisfacciones o incomodidades a unos y otros.

En definitiva, la tríada formada por información, conocimiento y cultura de la evaluación debe imperar sobre otra tríada, la de las "tres íes", con la que los profesores Abhijit Banerjee y Esther Duflo sintetizaban los riesgos que amenazan a las políticas públicas. Ideología, ignorancia e inercia estarían en la raíz de los fallos de las políticas, como advertían estos especialistas en evaluaciones de impacto bastantes años antes de recibir el Nobel de Economía de 2019 junto con Michael Kremer. Más recientemente, la concesión del Nobel de Economía 2021 a tres expertos en métodos de análisis de inferencia causal, los profesores David Card, Joshua Angrist y Guido Imbens, demuestra el auge que mantiene la evaluación de impacto en el panorama internacional. En esta guía se utilizarán algunos ejemplos procedentes de sus trabajos.

El Plan de Evaluación del Servicio Navarro de Empleo – Nafar Lansare 2020-2023 incluye, entre las acciones a realizar en el marco del objetivo de crear un sistema de evaluación, el diseño e implementación de un procedimiento para la evaluación del impacto. El objetivo de este Cuaderno es aportar valor añadido en este sentido. El Cuaderno se estructura en siete secciones, además de esta sección introductoria, que comprenden las cuestiones clave de la evaluación de impacto de políticas activas de empleo (en lo sucesivo, PAE).



2. Concepto de evaluación de impacto

Una evaluación de impacto consiste en un análisis de las consecuencias que tiene la participación en una política, programa o medida, sobre el colectivo al que se dirige dicha medida.

La clave de una evaluación de impacto es conseguir establecer con rigor la posible existencia de una relación causa – efecto entre el programa o medida evaluada y la evolución posterior del grupo objetivo. Para ello resulta fundamental el poder aislar los efectos de la medida en cuestión, separándolos y diferenciándolos de la influencia que pudieran tener otros factores sobre la evolución del grupo a lo largo del tiempo. De este modo, la evaluación de impacto se centra en la atribución directa de efectos observados a la medida, programa o innovación que es objeto de evaluación.

A continuación, se presentan una serie de ideas, acompañadas de algunos ejemplos, para ayudar a entender el alcance de las evaluaciones de impacto.

a. Qué se entiende por una política de empleo con impacto

Existen diversas formas de definir en qué consisten las políticas de empleo. A los efectos de este Cuaderno, se entiende como política de empleo cualquier acción cuyo propósito sea la mejora del funcionamiento del mercado de trabajo y que tenga como objetivo último la creación de empleo.

En términos generales, se pueden distinguir dos tipos de políticas de empleo: las que llevan aparejado un compromiso de gasto público, como puede ser una medida de subvenciones públicas a la contratación de un determinado colectivo; y aquellas otras que se derivan de la introducción de normas regulatorias en el mercado de trabajo, como, por ejemplo, la regulación administrativa de un nuevo tipo de contrato laboral o los cambios introducidos en contratos existentes. Ambos tipos de políticas, con programa de gasto o regulatorias, son susceptibles de generar un impacto. Por tanto, ambas son apropiadas para ser objeto de evaluaciones de impacto.

La pregunta lógica es: ¿impacto sobre qué, exactamente? Pues bien, si las subvenciones a la contratación se centran en un colectivo determinado, por las razones que sean, como puede ser que se trate de un grupo desaventajado en el mercado de trabajo, entonces parece razonable que la evaluación de impacto trate de descubrir si las subvenciones han conseguido impulsar el empleo de dicho grupo. De igual manera, cabe suponer que la regulación gubernamental responda a objetivos concretos, cuya consecución centre los esfuerzos de atribución del impacto de la norma aprobada.



Sin embargo, no hay por qué descartar que las políticas de empleo produzcan impactos en ámbitos adicionales o diferentes a los inicialmente deseados. Por ejemplo, la subvención de la contratación temporal, o la regulación de determinado contrato temporal, podría estar guiada por el esfuerzo en reducir el desempleo. Entonces, ¿se debería evaluar su impacto sobre el desempleo exclusivamente, o cabe plantear hipótesis adicionales? Por ejemplo, con independencia del impacto sobre el desempleo, dichas medidas podrían aumentar la tasa de temporalidad en el empleo. O podrían afectar a otras políticas de empleo, como la formación, pues es sabido que la temporalidad no ayuda precisamente a la formación para el empleo, pues la incertidumbre sobre la duración del contrato a medio y largo plazo desincentiva tanto al trabajador como a la empresa ante la asunción de costes y/o esfuerzos de la formación.

En términos generales, pueden establecerse tres tipos de impactos sobre los que evaluar las políticas de empleo:

- Impacto sobre aspectos de cantidad. Incluyen, entre otros, cambios en el volumen de empleo propiamente dicho, pero también en las horas trabajadas, en el número de contratos firmados, en el volumen de desempleo o en su composición. Distintos tipos de empleo –por ejemplo, indefinido o temporal- y de paro –por ejemplo, de larga duración- estarían incluidos en esta categoría, junto con indicadores similares relativos a la intensidad del empleo o del desempleo. Lógicamente, estos indicadores pueden ser generales, o bien centrarse en un colectivo determinado. También pueden medirse en niveles o en tasas de variación. Las evaluaciones disponibles, como se verá más adelante, suelen centrarse en impactos de este tipo.

- Impacto sobre aspectos de precio. En esta categoría cabe incluir, entre otros, los impactos de las políticas de empleo sobre los costes laborales de las empresas, de tipo tanto salarial como no salarial, y los impactos sobre los ingresos de los trabajadores o de las personas desempleadas.

- Impacto sobre otros aspectos, como, entre otros muchos, los patrones de búsqueda de empleo, las condiciones laborales, la jornada laboral, la seguridad laboral, los accidentes de trabajo, la conflictividad, la salud laboral o incluso la salud mental de los trabajadores y sus familias. En esta categoría encajaría el ejemplo citado anteriormente sobre el posible impacto de una medida que fomenta el empleo temporal sobre la formación.

Los numerosos impactos potenciales de las políticas no implican que toda política tenga impacto. Con frecuencia se evalúan medidas o programas con impactos esperados muy claros que, sin embargo, no llegan a poder demostrarse. Cuando esto ocurra, puede resultar extraño o puede parecer inexplicable, especialmente a las personas más involucradas en el día a día de la medida concreta. No obstante, cabe recordar que la imposibilidad de demostrar atribución de impacto no exime a la medida de haber supuesto una cierta contribución a los logros obtenidos.

En todo caso, un informe de evaluación de impacto deberá recoger las estimaciones específicas del impacto atribuido a la medida evaluada, estimaciones que acompañará de información sobre su significación estadística. Por ejemplo, se señalarán los puntos porcentuales de variación del empleo como consecuencia de la medida evaluada, y se indicará si esos puntos estimados son estadísticamente significativos y en qué grado.



b. Evaluación de resultados vs. evaluación de impacto

La delimitación precisa de la evaluación de impacto frente a otro tipo de evaluaciones, como la evaluación de resultados o la evaluación de efectos, no es tarea sencilla. Con frecuencia estos términos se utilizan indistintamente. Sin embargo, lo que define a la evaluación de impacto es la capacidad de establecer con rigor una relación causa – efecto entre la medida analizada y la realidad observada.

Por ejemplo, un programa de subvenciones a la contratación por parte de empresas privadas puede tener como resultado la contratación de un número concreto de personas durante el tiempo que duren las subvenciones. ¿Podría argumentarse que esto es un impacto? A salvo de los detalles de cada caso, lo normal es que se considere un resultado del programa, y que su impacto sea evaluado mediante el análisis riguroso de las personas que continúen trabajando una vez que terminen las subvenciones.

De forma similar, un curso de formación para personas desempleadas puede contribuir al logro de certificados de profesionalidad por parte de los asistentes. Cabe argumentar que dicho logro sea más un resultado que un impacto, que vendría dado, en su caso, por la mejora del empleo a más largo plazo del grupo de personas formadas y cualificadas.

Por otra parte, la evaluación de impacto no sustituye a otros tipos de evaluación, sino que las complementa. Idealmente, la toma de decisiones basadas en la evidencia considerará como algo valioso el disponer de evaluaciones diversas, como las de implementación, de costes, de satisfacción, de resultados o de impacto. En el ejemplo anterior de las subvenciones a la contratación, parece razonable que exista un paralelismo entre resultados (personas contratadas) e impacto posterior. Conocer ambos extremos, al igual que los que produzcan otras evaluaciones, contribuirá a mejorar las decisiones futuras sobre programas de este tipo.

c. Evaluación de proceso vs. evaluación de impacto

La llamada evaluación de proceso es también diferente, pero complementaria, a la evaluación de impacto. Este tipo de evaluación trata de comprender el funcionamiento de un programa mediante el análisis conjunto de su diseño, su implementación y la obtención de resultados. El objetivo es que dicho análisis contribuya a mejorar el programa en el futuro.

La evaluación de proceso, también llamada evaluación de seguimiento, atenderá a aspectos de una medida o programa tales como, primero, su relevancia de cara a la consecución de los objetivos últimos de la política. En segundo lugar, este enfoque abordará la coherencia entre las diversas medidas impulsadas en un ámbito determinado, entre las medidas y los recursos movilizados, y entre estos elementos y los objetivos iniciales.

Una evaluación de proceso de un programa de formación, por ejemplo, no dejará de atender a otras PAE que coincidan en el tiempo sobre el mismo colectivo. Si se diera el caso de que las personas inmersas en un plan formativo reciban ofertas para incorporarse a un programa de empleo subvencionado, es evidente que esta coincidencia afectará al resultado de ambas.

Una diferencia básica entre evaluaciones de impacto y de proceso es que las técnicas utilizadas para evaluar el impacto son predominantemente estadísticas, mientras que la evaluación de proceso suele basarse en técnicas de tipo cualitativo, como entrevistas, grupos de discusión o estudios de caso.

En todo caso, disponer de una evaluación de proceso contribuirá a comprender mejor los resultados de una evaluación de impacto, especialmente si éstos se alejan del objetivo fijado inicialmente.



d. Enfoques principales de las evaluaciones de impacto

Una tipología de interés, dentro de las evaluaciones de impacto, es la que distingue entre tres orientaciones: de eficacia o efectividad, de eficiencia y de utilidad.

La evaluación de eficacia o efectividad se centra en analizar los efectos de la medida en cuestión. El impacto generado por una PAE, en los términos descritos anteriormente, relacionados con aspectos de cantidad, de precio u otros aspectos, encaja en este tipo de análisis.

La evaluación orientada a la eficiencia parte de la anterior, añadiendo el análisis de los costes. Su intención es valorar la forma más económica de alcanzar determinado impacto.

Por ejemplo, una evaluación de eficacia o efectividad podría determinar que el programa evaluado tiene un impacto sobre el empleo del grupo objetivo del 10 por ciento. Y una evaluación de eficiencia trataría de contrastar dicho programa con otros cuyos efectos sean comparables. Esta evaluación permitiría conocer cuál de los programas permite incrementar en un determinado tanto por ciento el empleo del grupo de la forma más económica.

Por último, la evaluación orientada a la utilidad plantea un enfoque basado en la valoración no solamente del coste de una política, sino también de los cambios en el bienestar de la población como consecuencia de la misma. Enlazando con el ejemplo anteriormente citado, este enfoque requeriría el poder disponer de estimaciones sobre la utilidad o bienestar que la sociedad asigna al aumento de un determinado tanto por ciento en el empleo del grupo objetivo del programa.

Como punto final de este apartado, el cuadro siguiente recoge ejemplos de algunos impactos relacionados con los diferentes tipos de PAE, permitiendo, además, incluir otros aspectos de las políticas, como costes, factores, productos y resultados.

TIPOS DE POLÍTICAS	COSTES	FACTORES	PRODUCTOS	RESULTADOS	IMPACTO
Intermediación y orientación laboral	-Salarios del personal técnico y orientador	- Personal técnico y orientador	- Nº de servicios prestados, por tipos	- Tasa de registro de vacantes - Tasas de éxito de registro - Tasas de penetración	- Reducción posterior del desempleo en el grupo objetivo - Variación de los resultados de la intermediación a medio y largo plazo
Formación	- Salarios profesorado - Alquiler aulas	- Aulas - Profesorado - Consumibles	- Nº de cursos - Nº de horas - Nº de personas formadas	- Nº de personas que terminan la formación - Nº de certificados de profesionalidad	- Reducción posterior del desempleo en el grupo objetivo - Mantenimiento del empleo en el grupo objetivo con posterioridad a la formación
Subvenciones a la contratación	- Subvenciones otorgadas	- Personal de gestión	- Nº de contratos subvencionados - Nº de personas contratadas	- Experiencia laboral de las personas contratadas	- Nº de personas que continúan contratadas al terminar la subvención - Nº de personas que están trabajando con posterioridad al programa
Subvenciones al autoempleo	- Salarios personal de apoyo - Subvenciones otorgadas	- Asesores de planes de negocio	- Nº de personas perceptoras	- Nº de negocios establecidos	- Supervivencia de los negocios al cabo de un tiempo - Contrataciones adicionales creadas por las personas autoempleadas
Programas de empleo público	- Salarios personal de gestión	- Personal de gestión	- Nº de personas contratadas	- Experiencia laboral de las personas contratadas	- Acceso al empleo ordinario en el grupo objetivo - Mejora de las perspectivas de jubilación de las personas contratadas

Fuente.: Elaboración propia.



e. Oportunidad y utilidad de las evaluaciones de impacto

En la presentación se ha argumentado la conveniencia de combinar información, conocimiento y cultura de la evaluación para favorecer la toma de decisiones basadas en evidencias y evitar efectos perversos causados por la inercia, la ideología o la ignorancia. A todo ello cabe añadir un aspecto más que apunta a la utilidad de las evaluaciones de impacto de políticas de empleo.

La progresiva disponibilidad de evaluaciones de impacto va ofreciendo un conjunto de evidencias creciente, que permitirá obtener un panorama sobre qué medidas o programas funcionan mejor para qué grupos de población. Esto constituye una oportunidad que no se puede desaprovechar. Cuanto mayor sea el número de evidencias sobre impactos de PAE, será más sencillo determinar qué medidas son, a priori, más apropiadas para abordar determinadas necesidades de políticas de empleo, cuáles son de impacto dudoso, y qué otras medidas tienen, con alta probabilidad, impactos nulos o incluso negativos.

Como ejemplo, a lo largo de este Cuaderno se irán presentando varios recuadros que sintetizan casos concretos de PAE, las evaluaciones realizadas y los impactos obtenidos.



3. Efectos brutos y netos de las políticas activas del empleo

Una distinción clave para comprender plenamente la filosofía que subyace tras las evaluaciones de impacto es la que separa los efectos brutos y los efectos netos de las políticas evaluadas. Esta separación está íntimamente asociada a la posibilidad de que se produzca una falacia, muy conocida en ciencias sociales, resumida mediante la expresión latina "post hoc ergo propter hoc". La idea tras esta falacia, también conocida como la de la causa falsa, es que cuando un suceso viene seguido de otro a continuación, se interpreta que el segundo viene causado por el primero.

Los efectos brutos de una PAE vienen dados por la observación de lo que acontece al grupo objetivo de la política, en relación con los objetivos de la misma, con posterioridad a su realización. Los efectos brutos se obtienen a partir de un análisis descriptivo.

Por ejemplo, si la tasa de empleo de un grupo de personas desempleadas al que se dirigía la PAE evaluada, observada al cabo del tiempo establecido como referencia temporal adecuada, asciende al 70 por ciento del colectivo, esta cifra representará el efecto bruto de la medida en cuestión.

Los efectos netos, sin embargo, resultan bastante más complejos de calcular. Estos se definen como los efectos cuya atribución causal a la política evaluada ha quedado demostrada a partir de la evaluación de impacto.

En el ejemplo, es bastante improbable que la totalidad del 70 por ciento de las personas que trabajan lo haga como consecuencia directa de la PAE. Tras esta cifra se encontrarán todo tipo de casos, aunque cabe una clasificación en cuatro patrones básicos:

- 1. Habrá personas que habrían encontrado trabajo igualmente, aunque no hubiesen sido incluidas en la PAE, porque la manera de acceder al empleo y las causas de su reclutamiento hayan sido ajenas a su participación en la medida.

- 2. También puede haber personas desaventajadas para quienes la medida no sea suficiente como para mejorar sus perspectivas de empleo de manera efectiva.

- 3. El caso ideal vendrá dado por las personas a quienes la PAE consiga situar en una nueva perspectiva, influyendo realmente en su paso del desempleo al empleo.

- 4. Un último caso menos intuitivo, pero también posible, es el de personas a quienes la participación en la medida haya perjudicado. No hay que descartar que participar en la PAE haya distraído a estas personas de la búsqueda de empleo mediante alguna fórmula más adecuada a sus necesidades.

Los patrones 1 y 2 conforman, en la evaluación de impacto, el llamado efecto de inercia. Se trata de personas para quienes la PAE no supone ningún cambio, pues sus condiciones de otro tipo predominan sobre la posible influencia de la medida evaluada.

El efecto neto de la PAE vendrá dado por la diferencia entre los patrones 3 y 4. Es decir, la influencia causal de la medida será el resultado de comparar el grupo influido positivamente por la misma con aquel que haya resultado perjudicado, si lo hubiera.

En el ejemplo anterior, la tasa de empleo observada del 70 por ciento (efecto bruto) es compatible, con que un porcentaje inferior haya sido la consecuencia atribuible a la medida. Por ejemplo, un patrón 3 que ascienda al 20 por ciento y un patrón 4 del 3 por ciento daría como resultado un efecto neto del 17 por ciento.

El recuadro 1 recoge un caso real donde se especifican estos efectos.



Recuadro 1. Efectos brutos y netos de la tarifa plana para trabajadores autónomos

En el año 2013, el Gobierno de España estableció una medida orientada a fomentar el empleo autónomo, consistente en una "tarifa plana" de cotización a la Seguridad Social de 50 euros al mes (Real Decreto-ley 4/2013 de 22 de febrero, de medidas de apoyo al emprendedor y de estímulo del crecimiento y de la creación de empleo). La medida supuso una reducción significativa de la cuota previamente existente durante los primeros seis meses de actividad (un 80% menos en la base mínima de cotización) y una reducción menor durante los dos años siguientes. Aplicada inicialmente a la población joven, bajo el objetivo de reducir el paro juvenil, y a personas con discapacidad, esta medida se amplió posteriormente a toda persona que se diese de alta como autónoma por vez primera en cinco años.

En 2020, la Autoridad Independiente de Responsabilidad Fiscal (AIReF) llevó a cabo una evaluación de esta medida en la que puso de manifiesto que durante los años 2014 a 2018 se había asumido un coste en términos de bonificaciones a la Seguridad Social cercano a los 600 millones de euros al año, y se habían beneficiado de la tarifa plana cerca de 600.000 personas al año.

De la evaluación de la AIReF se desprende que los efectos brutos de la tarifa plana habrían consistido en la salida desde el desempleo hacia el empleo autónomo de unas 600.000 personas al año.

Sin embargo, el análisis causal realizado en dicha evaluación especifica que el efecto neto de la medida, es decir, el incremento de la probabilidad de que las personas desempleadas iniciasen una actividad como autónomas, atribuible directamente a este programa, es mucho menor. Concretamente, dicho impacto se limita a personas que estuvieron desempleadas durante periodos muy cortos y con un empleo temporal previo, lo que la evaluación considera "transiciones directas de un empleo a otro". No se encuentra evidencia de ese efecto positivo entre episodios de desempleo de duración superior al mes ni tampoco entre personas desempleadas con bajo nivel educativo.

Es decir, la gran mayoría de personas beneficiadas por el programa se hubieran dado de alta como autónomos, aunque no existiese la tarifa plana. Por tanto, el programa presenta un elevado efecto inercia y un efecto neto muy reducido. Esto quiere decir que la mayor parte de los 600 millones anuales que ha costado esta medida no han tenido ningún impacto sobre la dinámica de reducción del desempleo y, como mucho, han contribuido marginalmente a las transiciones directas desde el empleo asalariado hacia el autoempleo. La reducción del paro juvenil, uno de los objetivos iniciales, ha quedado totalmente exenta del impacto de la tarifa plana.

4. La evaluación experimental como referencia

Dado que la evaluación de impacto tiene como elemento central la atribución de causalidad, una de sus referencias metodológicas más importantes, si no la que más, es la llamada evaluación experimental, a la que se dedica esta sección. En el primer apartado se explica su lógica, que se acompaña posteriormente del relato del aumento creciente de este tipo de evaluaciones y de un repaso de lo que podemos considerar los tres “mantras” de los evaluadores experimentales: el contrafactual, el sesgo de selección y la validez interna y externa.

a. La lógica de la evaluación experimental

Una evaluación experimental consiste, en síntesis, en la atribución de un efecto causal a partir de la comparación de un grupo afectado por un determinado tratamiento y un grupo de control. Las unidades que forman los grupos pueden ser, entre otras y dependiendo del tipo de tratamiento, las siguientes: personas, animales, empresas, escuelas, hospitales, y entidades geográficas de diversos niveles como localidades, regiones o países.

Las evaluaciones experimentales más conocidas proceden del ámbito de la salud. La efectividad de los tratamientos médicos de todo tipo ha de ser demostrada rigurosamente mediante experimentos, como ha puesto de manifiesto la reciente oleada de análisis de vacunas frente al COVID.

La diversidad de tratamientos que son objeto de evaluación experimental no impide, sino que, más bien al contrario, anima a exponer la lógica común a este tipo de análisis.

La razón fundamental de llevar a cabo tratamientos experimentales es poder extraer conclusiones sobre su impacto. Esta, y no otra, es la razón de que dichos tratamientos afecten a un grupo y excluyan a otro.

La lógica de la evaluación experimental es el poder tomar como única diferencia entre los dos grupos de personas o entidades el tratamiento que ha recibido uno de ellos y la ausencia de tratamiento del otro. Para que la comparación sea válida se precisa que el tratamiento se asigne aleatoriamente, de modo que constituya la única diferencia existente entre ambos grupos.



b. El desarrollo progresivo de las evaluaciones experimentales

Los tratamientos experimentales y su evaluación proceden de la medicina y de las ciencias de la salud en general. Los experimentos son habituales en los laboratorios científicos, requieren protocolos muy detallados y se necesita confirmar los impactos antes de aceptar la existencia de evidencia científica.

Puede parecer iluso, atrevido o incluso una pérdida de tiempo y recursos el intentar adoptar métodos propios de las ciencias de la salud por parte de las ciencias sociales. Sin embargo, durante las últimas décadas eso es exactamente lo que ha venido ocurriendo. En el ámbito de las políticas públicas, la necesidad de obtener evidencia sobre el funcionamiento de programas y medidas (las ventajas) se va imponiendo progresivamente a los problemas inherentes que presenta la evaluación experimental fuera del laboratorio (los inconvenientes).

Concretamente, en el campo de la cooperación al desarrollo, han proliferado los programas experimentales basados en seleccionar al azar grupos de personas, escuelas, empresas, poblados o regiones que se beneficiarían de determinadas medidas. Simultáneamente, los correspondientes grupos de control quedarían excluidos del alcance de dichas medidas. Los problemas éticos derivados de estas exclusiones quedarían superados por las ventajas de disponer de evidencia sobre el impacto de cada medida. A su vez, esta evidencia justificaría el escalamiento de las medidas exitosas hacia el conjunto de la población.

Las políticas de desarrollo y, más específicamente, las medidas contra la pobreza, han experimentado cuatro procesos paralelos durante las últimas dos décadas: un incremento sustancial de los programas experimentales, un aumento equivalente de la obtención de evidencia sobre el impacto de diversas medidas, un impulso de organizaciones y programas destinados a escalar las medidas exitosas, y un reconocimiento público creciente. Este último queda personificado en los investigadores galardonados con el Premio de Ciencias Económicas del Banco de Suecia en Memoria de Alfred Nobel en 2019, Abhijit Banerjee, Esther Duflo y Michael Kremer. Estos autores, entre otros, han trabajado sobre políticas de desarrollo en muy diversos campos, ayudando a superar prejuicios sobre el funcionamiento de la ayuda al desarrollo con la obtención de evidencia causal.

En el ámbito de las PAE las evaluaciones experimentales habían empezado anteriormente, en las últimas décadas del siglo XX, sobre todo en los EEUU. Uno de los programas pioneros, el Job Training Partnership Act, produjo evidencia sobre el impacto positivo de la formación para el empleo entre mujeres adultas de grupos desaventajados, a la vez que se obtuvo un impacto mínimo en el caso de los hombres, e incluso negativo entre la población joven. Desde entonces se han llevado a cabo innumerables evaluaciones de este tipo sobre políticas de empleo y formación, más en Norteamérica que en Europa. El Recuadro 2 recoge un ejemplo reciente de evaluación experimental realizado en los Países Bajos.

Recuadro 2. Evaluación experimental de la intermediación laboral mediante agencias de trabajo temporal

En los Países Bajos, las oficinas del Servicio Público de Empleo organizan eventos en los que reúnen a personas perceptoras de prestaciones por desempleo con representantes de empresas de trabajo temporal. El objetivo de estos encuentros, iniciados en 2011, es el intercambio de información respecto a las características de los demandantes de empleo y de los puestos vacantes disponibles, así como, en ocasiones, asesoramiento laboral y revisión del currículum vitae.

Los eventos suelen ir dirigidos a personas que llevan menos de tres meses desempleadas, a quienes se invita a participar desde el SPE. No hay obligación, pero la asistencia cuenta como una solicitud de empleo, y en ese país los perceptores de prestaciones deben llevar a cabo al menos una solicitud a la semana. Los eventos se han popularizado porque su organización resulta poco costosa y las empresas de trabajo temporal, que no son remuneradas, los consideran útiles para recabar personal.

Entre 2014 y 2016, investigadores de la Universidad Libre de Ámsterdam, a iniciativa del SPE, realizaron una evaluación experimental de este programa. Para ello se seleccionó de forma aleatoria entre las personas desempleadas a un grupo que recibiría las invitaciones para asistir, dejando al resto sin invitación. Posteriormente se compararon datos de contrataciones y también se envió un cuestionario a ambos grupos, sin hacer mención explícita a la evaluación.

Los efectos del programa se resumen, en aspectos de cantidad en un incremento del 7 por ciento en el acceso al empleo a corto plazo, sin impacto significativo a medio y largo plazo. En aspectos de precio, el programa causa una sustitución de ingresos de prestaciones por ingresos del trabajo. Adicionalmente, se concluyó que el comportamiento de búsqueda de empleo del grupo tratado se modificó, mostrando un mayor esfuerzo de búsqueda y cierta moderación de sus expectativas salariales, en comparación con el grupo de control.

Las posibles reticencias que pueda haber para llevar a cabo programas experimentales contrastan con otras situaciones en que la asignación al azar de participantes se convierte en una solución, o simplemente encaja con el marco de la política. Piénsese, por ejemplo, en los sorteos para asignar viviendas sociales, que han sido utilizados en nuestro país como medida ante el exceso de solicitantes. Otra posibilidad es utilizar la selección aleatoria al inicio de un programa, seleccionando los participantes en la fase piloto, lo que no impedirá la entrada de participantes adicionales más adelante.

c. Tres mantras (i): el contrafactual

En la evaluación experimental pueden reconocerse tres mantras cuya explicación es necesaria, pues los tres se utilizan posteriormente en el conjunto de evaluaciones de impacto, constituyendo conceptos clave en esta literatura.

El primero de ellos es el llamado contrafactual. Su importancia es tal que con frecuencia las evaluaciones de impacto con grupo de control se han dado en llamar evaluaciones con enfoque contrafactual.

Se denomina el contrafactual a lo que le habría ocurrido al grupo objeto de la política o programa evaluado en el caso de que no hubiese existido tal programa. Es evidente que el interés se encuentra en la evolución de dicho grupo en relación al objetivo del programa. Si el objetivo es la reducción del desempleo, interesa conocer el contrafactual respecto al paro del grupo al que se dirige el programa.

En los estudios de laboratorio el contrafactual se obtiene mediante la observación del grupo de control. En una evaluación experimental ocurre lo mismo. Las unidades que conforman el grupo de control reflejan la ausencia de tratamiento.

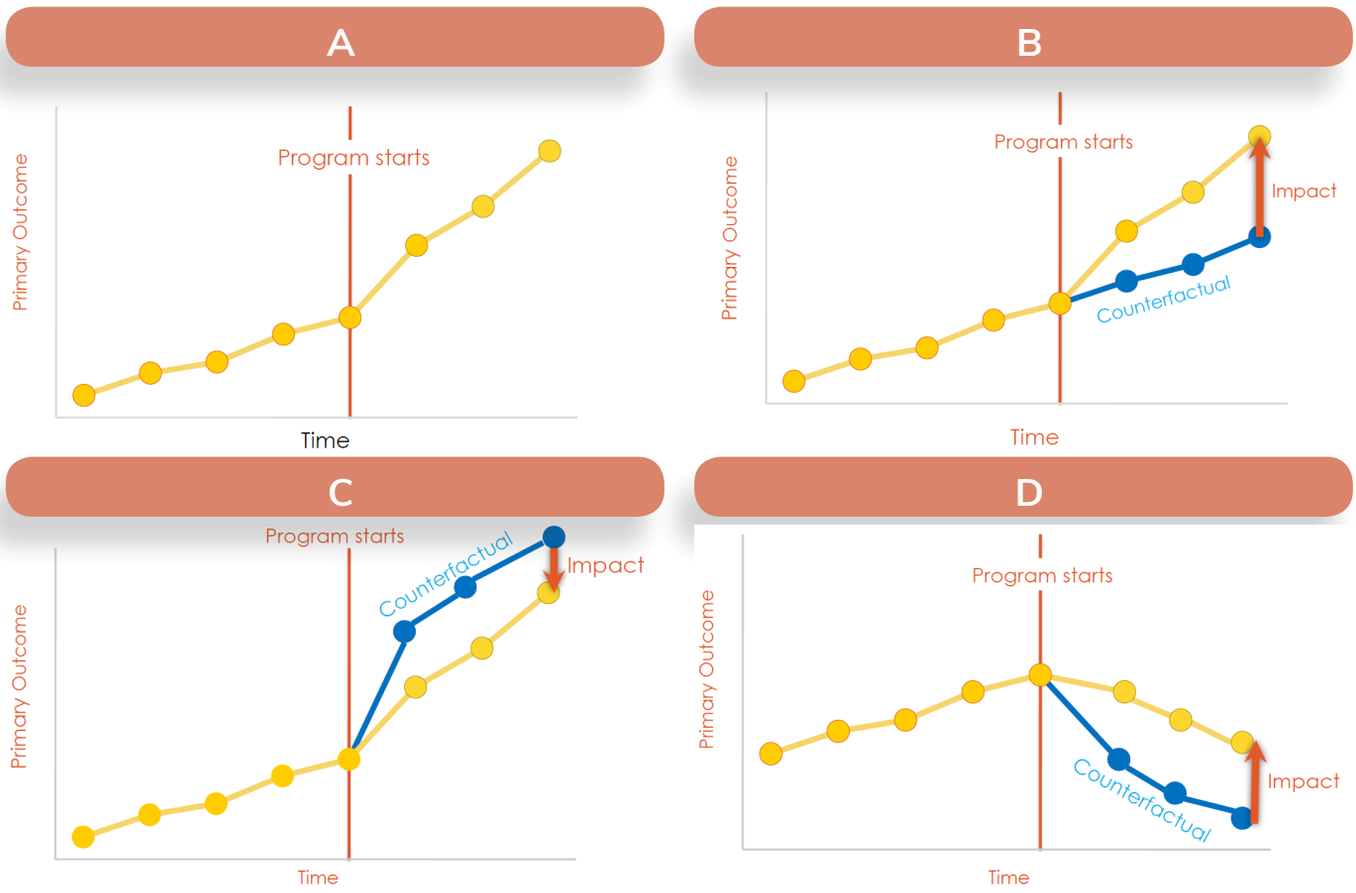
Sin embargo, la evaluación de impacto de PAE, aunque sea experimental, no se produce en las circunstancias de aislamiento de un laboratorio. El desempeño del grupo de control en relación con el objetivo del programa estará sometido a la influencia de factores externos. Aunque no hayan recibido el tratamiento, los miembros del grupo de control evolucionan, de la manera en que lo hagan, bajo la influencia, por ejemplo, del ciclo económico o del comportamiento del mercado de trabajo.

Por tanto, el dinamismo inherente a que detrás de las políticas de empleo se encuentren grupos de personas dificulta la evaluación de impacto, incluso cuando se dispone de un programa experimental.

La figura 1 recoge gráficamente cuatro posibles impactos de una PAE, señalando la evolución del grupo en cuanto al objetivo del programa (eje vertical) y distinguiendo con claridad el momento a partir del cual podrían ponerse de manifiesto los efectos de la medida (eje horizontal). En el gráfico A da la impresión de que el programa beneficia al colectivo, pues mejora su desempeño posterior. Pero los otros gráficos muestran la importancia del contrafactual, señalando que la misma evolución mostrada en A puede contener un impacto positivo (B) o negativo (C). El cuarto gráfico (D) también refleja que el impacto puede ser positivo incluso cuando los registros del grupo en relación con el objetivo del programa hayan empeorado.

Por ejemplo, una crisis económica explicaría que lo que le habría ocurrido al grupo objetivo del programa en ausencia de éste (es decir, el contrafactual) fuese peor de lo que realmente ocurrió.

Figura 1. El contrafactual en la evaluación de impacto



Fuente. : Latif Jameel Poverty Action Lab.

d. Tres mantras (ii): el sesgo de selección

El segundo mantra de la literatura sobre evaluación de impacto es el sesgo de selección. Este sesgo, que debe evitarse a toda costa, se produce cuando el grupo tratado y el grupo de control difieren no solamente por el tratamiento, sino por alguna otra variable que es susceptible de afectar al desempeño posterior.

Imaginemos que un programa de orientación laboral se ofrece prioritariamente a personas desempleadas residentes en las comarcas con mayores niveles de paro de una comunidad. Esta posibilidad abre la puerta al sesgo de selección, porque si finalmente se configura un grupo de tratamiento con más miembros de comarcas en crisis y un grupo de control con más miembros de comarcas prósperas, podría ocurrir que la comparación posterior ofreciese un impacto sesgado del programa. Es probable que el impacto estimado sea negativo y que ello pueda deberse a las mayores dificultades de los miembros del grupo participante a la hora de poner en práctica los contenidos de la orientación.

Sesgos de este tipo podrían producirse si en los grupos de tratamiento y de control hay proporciones distintas de personas con determinado dinamismo de búsqueda de empleo, con ciertos niveles de estudios, con mayor o menor experiencia laboral, con cierto grado de desventaja ante la búsqueda de empleo o con otros factores que pueden influir sobre el acceso al empleo.

Al lector avezado no se le escapará que la mejor forma de evitar el sesgo de selección es seleccionar aleatoriamente a los miembros del grupo de tratamiento y de control, a partir del colectivo potencialmente tratado. De este modo, los factores coadyuvantes al logro del objetivo del programa estarán repartidos de forma similar en ambos grupos.

e. Tres mantras (iii): la validez interna y externa

El tercer mantra tiene que ver con la validez de la evaluación experimental. Cuando se habla de validez interna se quiere decir que el experimento se ha realizado correctamente, evitando sesgos y ofreciendo indicadores de impacto rigurosos y representativos. La validez interna depende, por tanto, del cuidado que pongan todos los agentes involucrados en la evaluación, y también de las personas encargadas de la gestión, como se verá más adelante.

La validez externa apela a otro rasgo deseable de la evaluación de impacto, pero que es más difícil de probar. Se trata de que el impacto estimado en unas circunstancias concretas sea extrapolable a otras diferentes, externas a las primeras. ¿Se producirían resultados similares a los obtenidos en los Países Bajos si se llevase a cabo un experimento como el de los eventos, que recoge el Recuadro 2, en la Comunidad de Navarra?.

Es evidente que la validez externa resulta muy difícil de probar. Frente a los experimentos de laboratorio, el análisis socioeconómico está expuesto a diversidad de influencias adicionales a los tratamientos que se busca evaluar. Lo que produce impacto en un entorno o en un momento determinado no tiene porqué producirlo en otras circunstancias. Sin embargo, si se dispone de evaluaciones reiteradas sobre políticas de contenido similar, se observan patrones comunes que pueden resultar muy útiles a la hora de tomar decisiones sobre políticas de empleo. A este respecto, resultan muy relevantes las metaevaluaciones de PAE, que se explicarán más adelante.

En definitiva, una evaluación experimental constituye la mejor oportunidad para obtener impactos atribuibles a la política evaluada, haciendo uso del contrafactual, obtenido éste mediante selección aleatoria de las unidades participantes en la medida y de las que conforman el grupo de control, y alcanzando validez interna. No obstante, las dificultades de diverso tipo para llevar a cabo evaluaciones experimentales ayudan a explicar que todavía sean escasas. En el ámbito de la Unión Europea, caracterizado por un peso significativo de las políticas y los derechos sociales, con frecuencia se presta más atención a la privación de participar en una medida a los miembros del grupo de control que al interés en conocer el impacto de dicha medida. Así, podría darse la paradoja de que nos oponemos a evaluar de esta forma basándonos en que no se debe privar a nadie de servicios cuyo impacto nos es desconocido.

5. Los experimentos naturales

Los llamados experimentos naturales constituyen una alternativa muy interesante ante las evidentes dificultades de las políticas públicas para llevar a cabo selección aleatoria de perfiles y grupos beneficiarios de las mismas. Un experimento natural surge cuando, por alguna circunstancia, se produce un cambio significativo que afecta a una parte de la población o del entorno, pero no a su totalidad. Si este cambio afecta a una variable con impacto potencial sobre alguno de los aspectos de cantidad, de precio u otros, podrían darse las circunstancias para evaluar el experimento natural.

Para comprender mejor la idea del experimento natural pueden plantearse circunstancias que podrían dar lugar al mismo. Por ejemplo, un cambio legislativo que, por la razón que sea, afecte a unas regiones, pero no a la totalidad de un país. O un retraso en la implementación de un programa, motivado por circunstancias ajenas al mismo, que genere que haya zonas o grupos tratados y zonas o grupos no tratados. Si se da la circunstancia de que sea posible comparar a las regiones o a los grupos afectados por los cambios con aquellos que no han sufrido modificaciones, en relación con alguna variable de interés, la lógica de la evaluación de impacto es parecida a la descrita anteriormente para los experimentos.

Por ejemplo, si un programa de formación para el empleo fuese implementado en unas comarcas, pero no lo fuese en otras, o sufriese retrasos significativos, la variabilidad entre comarcas permitiría una comparación de interés en la línea de un experimento natural. Adicionalmente, si con posterioridad el programa fuese implementado en las comarcas donde no lo hubo al principio, podría repetirse la evaluación de impacto. Entonces, las comarcas de control del primer momento pasarían a ser comarcas de tratamiento, y viceversa.

Existen numerosos casos de experimentos naturales, de interés para el análisis del mercado de trabajo, con origen en movimientos migratorios y de refugiados. Cuando estos movimientos son rápidos y vienen causados por factores no laborales, como conflictos bélicos o crisis políticas, ello ha sido aprovechado por numerosos investigadores para estudiar el impacto de las migraciones sobre los mercados de trabajo de destino. El Recuadro 3 toma un estudio de caso de experimento natural diferente, relativo al impacto del salario mínimo.

Recuadro 3. Evaluación mediante un experimento natural del incremento del salario mínimo

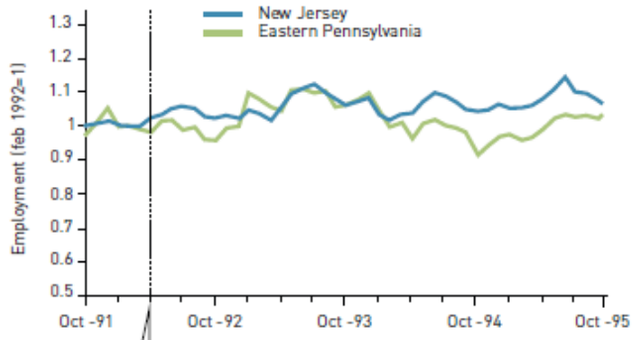
Probablemente el ejemplo más conocido de experimento natural en el ámbito de las políticas de mercado de trabajo se debe a uno de los profesores galardonados con el Nobel de Economía en 2021, David Card. Junto con el tristemente fallecido Alan Krueger, Card observó que el incremento del salario mínimo aprobado en el año 1992 en Nueva Jersey era susceptible de ser evaluado mediante comparación con el vecino estado de Pensilvania, donde el salario mínimo no sufrió cambios. El hecho de que ambos estados compartan una frontera bastante irregular, que en gran parte cruza zonas urbanas y residenciales, permitió esta comparación.

Card y Krueger analizaron el impacto del incremento del salario mínimo seleccionando para ello un sector de actividad como los restaurantes de comida rápida, cuyas plantillas en su mayor parte son remuneradas con el mínimo legal. Su trabajo tuvo gran relevancia, primero, por el esfuerzo de análisis empírico que supuso. Pero fueron sus conclusiones las que más contribuyeron a que la evaluación saltara a los libros de texto. En contra de lo esperado, el nivel de empleo en los establecimientos del sector en Nueva Jersey evolucionó por encima del nivel de Pensilvania, demostrando que el incremento del salario mínimo no generó destrucción de empleo.

(Fuente del gráfico: Fundación Nobel).

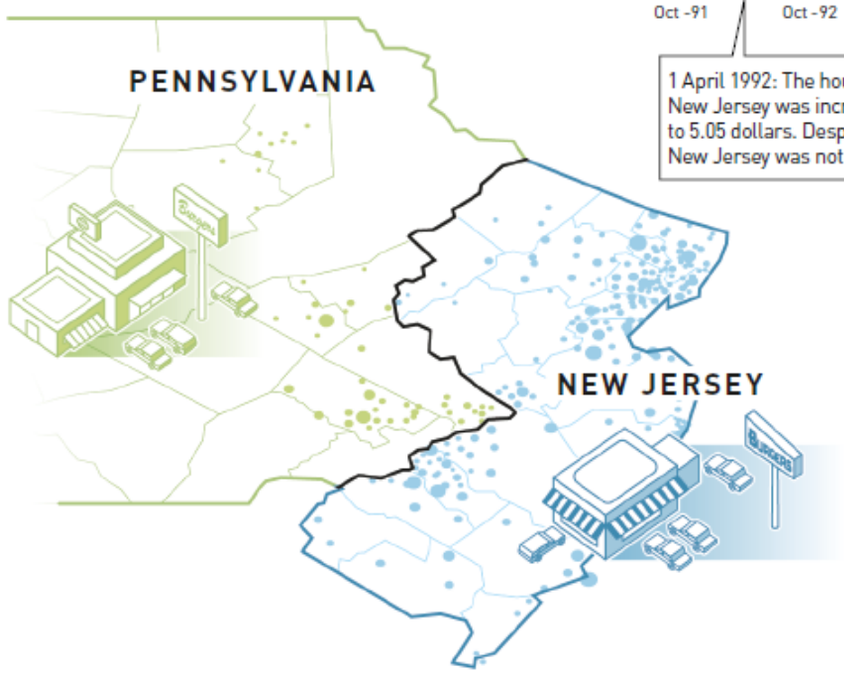
Card and Krueger used a natural experiment to study how increasing the minimum wage affects employment.

The researchers identified a treatment group (restaurants in New Jersey) and a control group (restaurants in eastern Pennsylvania) to measure the effect of increasing the minimum wage.



1 April 1992: The hourly minimum wage in New Jersey was increased from 4.25 dollars to 5.05 dollars. Despite this, employment in New Jersey was not affected.

CONTROL GROUP TREATMENT GROUP



6. Otras evaluaciones de impacto: cuasi-experimentos, evaluaciones macroeconómicas y metaevaluaciones

En esta sección se repasan algunos métodos de evaluación de impacto, sin ánimo de exhaustividad técnica, con el objetivo de dar a conocer las ideas básicas que recoge cada uno de ellos a personas que no necesariamente serán expertos en evaluación de este tipo.

El denominador común a estos métodos es la búsqueda de atribución causal entre una política y su impacto posterior, en ausencia de evaluación experimental. Se trata, por tanto, de técnicas que pueden ser útiles para evaluar el impacto de las PAE en entornos donde es infrecuente llevar a cabo políticas o programas con selección aleatoria.

Cada método es diferente y su utilidad dependerá de cada caso. Se han seleccionado dos de ellos, los de tipo cuasi-experimental y la metaevaluación. La razón para seleccionar estos dos métodos es el interés que pueden tener para las personas involucradas en la gestión de las PAE de la Comunidad Foral de Navarra.

a. Evaluaciones cuasiexperimentales

La inclusión de los cuasiexperimentos en este Cuaderno se debe a que, dadas las dificultades para evaluar PAE de forma experimental, resulta más probable que la evaluación de impacto en el marco de una comunidad autónoma pueda adoptar la forma de cuasiexperimentos.

Las evaluaciones cuasiexperimentales tratan de acercarse lo más posible al modelo descrito anteriormente para los experimentos, en ausencia de programas con selección aleatoria de participantes. Para ello utilizan técnicas matemáticas y econométricas relativamente sofisticadas, cuya explicación excede los objetivos de este Cuaderno. No obstante, es aconsejable presentar algunos de sus rasgos básicos, que ayuden a acercar la distancia que separa a la documentación y explicaciones más técnicas de los textos y contenidos de programas y medidas de las PAE.

i. Análisis de emparejamiento

Una referencia útil para entender los cuasiexperimentos es la técnica del emparejamiento o "matching", que de algún modo se encuentra en los orígenes de los métodos más avanzados. La idea es sencilla: en ausencia de selección aleatoria de participantes en PAE, la posible comparación entre éstos y un grupo de control podrá tener sesgos de selección. Entonces, el emparejamiento trata de segmentar cada uno de los grupos y, posteriormente, llevar a cabo comparaciones por segmentos.

Por ejemplo, en el grupo tratado en la PAE puede haber "muchos" jóvenes universitarios y "pocos" adultos con estudios primarios, y en el grupo de comparación puede ser al revés: "pocos" universitarios y "muchos" adultos con primaria. El método del emparejamiento segmentaría cada grupo por niveles de estudios y compararía el impacto de la PAE en cada nivel, calculando posteriormente el impacto medio ponderado. Con ello se evitaría sobrevalorar el impacto si, como es probable, la mejor inserción laboral de los universitarios tratados tienda a inflar los efectos aparentes (brutos) de la política.

El emparejamiento resulta sencillo en casos como el del ejemplo, donde hay una variable afectada, el nivel educativo, con unas pocas categorías. El problema aparece cuando las diferencias entre el grupo tratado y el posible grupo de comparación afectan a un gran número de características. Para resolver estas situaciones hay métodos algo más sofisticados, como el Propensity Score, entre otros, que permiten estimar impactos con emparejamiento de un elevado número de variables. De este modo, factores que influyen sobre la inserción laboral, como rasgos sociodemográficos o de historia laboral, pueden ser controlados para conseguir aislar el impacto de la PAE de dichos factores. La ventaja de poder emparejar por características numerosas es que puede llegar a evitarse el sesgo de selección explicado en la sección cuarta, obteniendo una estimación aceptable del impacto de la política.

El hecho de que la mayoría de las evaluaciones de PAE disponibles en nuestro entorno hayan utilizado técnicas de emparejamiento explica que los recuadros 4, 5 y 6 recojan ejemplos de este tipo. En ellos se resumen, respectivamente, un análisis del impacto de cursos de formación en Cataluña, una evaluación de iniciativas locales de empleo realizada en Andalucía, y un estudio sobre prevención del abandono escolar temprano realizado en Asturias.



Recuadro 4. Evaluación de impacto de la formación para el empleo

Entre los numerosos estudios sobre el impacto de la formación para el empleo, un buen ejemplo es el del trabajo publicado en 2010 por Begoña Cueto, Luis Toharia, Carlos García Serrano y Joan A. Alujas sobre la formación ocupacional en Cataluña. En él se utiliza el método del emparejamiento y se analiza a personas registradas como desempleadas en el servicio público de empleo catalán. Se establecen comparaciones entre el grupo de parados que asistieron a cursos de formación con parados registrados no formados. Se analizó el impacto de la formación sobre la probabilidad de empleo a lo largo de los dos años siguientes a la formación.

Parte del interés de este trabajo es que, al disponer de datos de registro, permitió utilizar bases de datos de tamaño suficiente para distinguir entre distintos programas formativos, como la formación dirigida a nuevas cualificaciones o la formación en empresas con compromiso de contratación. El gran número de datos también permitió obtener impactos por sexo, edad, duración del desempleo y percepción de prestaciones.

Los resultados de este estudio mostraron que la formación incrementa la probabilidad de empleo en cerca de 10 puntos porcentuales. Pero el impacto varía mucho según la duración de la formación. Así, mientras los cursos de corta duración tienen un impacto muy reducido (2 puntos porcentuales), los cursos de duración media (251 – 500 horas) presentan un impacto de 6 puntos porcentuales y los cursos largos (más de 500 horas) muestran un impacto en el orden de 7 a 9 puntos porcentuales.

Por otra parte, este estudio muestra que la formación impactó en mayor medida sobre el empleo de las mujeres que sobre los hombres, y también señala que el impacto crece cuando se considera como grupo de tratamiento exclusivamente a quienes finalizaron la formación con aprovechamiento, es decir, con una evaluación positiva por parte de los formadores.

Recuadro 5. Evaluación de iniciativas locales de empleo

Las iniciativas o programas locales de empleo son políticas dirigidas a personas desempleadas y, con frecuencia, a parados especialmente desaventajados en el mercado de trabajo. Estos programas tienen como objetivo el favorecer la empleabilidad de dichos colectivos, para lo que ofrecen contratos de trabajo en el ámbito de las administraciones públicas (normalmente, en ayuntamientos). Los contratos son de duración temporal y los programas, aunque de contenido muy similar, reciben denominaciones diferentes dependiendo de la comunidad autónoma, diputación o municipio que las gestione.

En un estudio publicado en 2021, Yolanda Rebollo-Sanz y José I. García Pérez recogen la evaluación del impacto de estos programas, que en Andalucía reciben el nombre de Proyectos de Cooperación Social y Comunitaria. Como es frecuente en otras comunidades autónomas, el programa consiste en financiar a los ayuntamientos que, a su vez, son quienes proceden a contratar a las personas desempleadas que cumplan los requisitos. En este caso, los contratos puestos en marcha tuvieron una duración de entre 3 y 6 meses y la contratación se realizó sobre la base de la vulnerabilidad social, priorizando a personas con menor renta y mayor duración del desempleo, entre los residentes inscritos como demandantes de empleo.

El estudio analiza el impacto por separado para los programas dirigidos a menores y mayores de 30 años, obteniendo que en ambos casos la participación en las iniciativas de empleo reduce la probabilidad de estar trabajando al cabo de seis meses. El impacto es casi idéntico, con un empeoramiento de 1,2 puntos porcentuales en la probabilidad de empleo de ambos grupos. Sin embargo, el impacto desagregado por colectivos más específicos muestra que las personas mayores de 30 años que llevaban desempleadas más de dos años o que carecían de experiencia laboral ven mejorar su empleabilidad entre 2,1 y 3,1 puntos porcentuales.

Recuadro 6. Evaluación de programas de prevención de abandono escolar temprano

Es sabido que una parte importante de los problemas de inserción de los jóvenes en el mercado de trabajo viene influida por el abandono escolar temprano. De ahí que programas de la Unión Europea, financiados por el Fondo Social Europeo, y especialmente a raíz del establecimiento de la Garantía Juvenil, aborden este problema social.

Una evaluación de impacto con técnicas de emparejamiento, publicada en 2021 por Javier Mato, Ángel Alonso, Rosario González e Israel Escudero, aborda un programa de prevención del abandono, el llamado Programa de Mejora del Aprendizaje y del Rendimiento (PMAR), con datos de Asturias. Se trata de un programa que atiende a estudiantes de la ESO en riesgo de abandono durante los cursos intermedios (2º y 3º). Los estudiantes retornan al grupo ordinario en el 4º y último curso de la ESO.

Esta evaluación contrasta al alumnado de la ESO que participó en el programa con un grupo de comparación formado por alumnado de los mismos centros educativos, pero que no participó. Las dos variables de impacto consideradas son la permanencia en el sistema educativo al cabo de dos años de entrar en el PMAR y la tasa de titulación en la ESO. En ambos casos se obtuvieron impactos positivos. La participación en el PMAR incrementa la permanencia en el sistema educativo en cerca de 12 puntos porcentuales al cabo de dos años. En el caso de la finalización de la ESO con titulación, el impacto varía entre 8 y 13 puntos porcentuales, dependiendo del horizonte temporal considerado.

Uno de los rasgos más señalados en esta evaluación es el impacto significativamente mayor del programa sobre las mujeres que sobre los hombres. Las mujeres ven incrementarse la probabilidad de terminar la ESO con éxito entre 10 y 22 puntos porcentuales, dependiendo del horizonte temporal considerado.

ii. Análisis de regresión en discontinuidad

Una técnica crecientemente utilizada para la evaluación cuasiexperimental es la llamada regresión en discontinuidad. En síntesis, se trata de un análisis que aprovecha la existencia de un criterio de selección a la PAE que genera un punto de corte o efecto frontera. Por ejemplo, si el acceso a un determinado servicio para personas desempleadas quedase determinado por la antigüedad como parado registrado, de tal manera que se dirija a personas con más de 25 semanas en alta, este tipo de evaluación utilizará como grupo de tratamiento al subgrupo de los participantes que se encuentren más próximos a ese punto de corte, como las personas en altas durante 25, 26 y 27 semanas. Y utilizará como grupo de comparación, para evaluar el impacto del servicio en cuestión, a una submuestra de desempleados registrados lo más cercanos posible al mismo punto, como personas con 22, 23 y 24 semanas en alta. La idea subyacente es que ambos grupos serán muy similares. Es más, si existen suficientes casos para realizar un análisis estadístico, la comparación podría restringirse a personas con 24 semanas en paro (no tratadas por la PAE) frente a las de 25 semanas (tratadas).

A este ejemplo pueden añadirse otros: existen PAE a la que se accede con límite de edad, que actuaría como punto de corte. Un caso concreto son las subvenciones a la contratación. Otro ejemplo son los servicios que restringen el acceso en función de los ingresos, como las prestaciones por desempleo de tipo no contributivo. En el ámbito de las políticas dirigidas a jóvenes también hay casos con puntos de corte dados por las notas, como ocurre cuando, por ejemplo, un programa de prevención del abandono escolar se dirige a quienes tengan un número concreto de asignaturas suspensas y/o de cursos repetidos.

La regresión en discontinuidad, por tanto, requiere que exista un punto de corte. Pero requiere también que la gestión de la PAE haya sido estricta en la aplicación del criterio que genera dicho punto. Por ejemplo, si durante la implementación de la PAE se hubiera relajado el requisito en cuestión, la aplicación de este método puede peligrar.

iii. Análisis de diferencias en diferencias

Una técnica adicional para los cuasiexperimentos es la llamada doble diferencia o de diferencias en diferencias. La idea es muy sencilla: si se dispone de datos históricos sobre la variable objetivo del colectivo al que se dirige la PAE, pueden aprovecharse para evaluar su impacto haciendo acopio de los mismos datos con posterioridad al tratamiento. El caso ideal es aquel en el que los datos históricos muestran un paralelismo entre quienes accedieron a la PAE y el grupo de comparación. Estudiando los datos posteriores a la aplicación de la medida, si se produce una desviación entre ambos grupos, será evidencia del impacto de la política.

El nombre del método procede de un caso más general, que puede ejemplificarse con datos inventados. Supongamos que la diferencia en la tasa media de empleo (histórica) entre quienes han accedido a la PAE y el grupo de comparación era del 10%, favorable a los tratados. Si, una vez finalizada la medida y transcurrido un plazo razonable, la diferencia asciende al 25%, favorable también a los tratados, entonces la evaluación por doble diferencia nos va a decir que la PAE ha tenido un impacto del 15% sobre la tasa de empleo (es decir, la resta de 25% menos 10%). Evidentemente, para validar este resultado, por intuitivo que resulte, es preciso un análisis riguroso con datos de calidad.

Los tres métodos presentados en esta sección comparten una condición que es preciso poner de manifiesto: las variables o características que se analicen, bien porque se utilicen para el emparejamiento, porque den lugar al punto de corte de la regresión en discontinuidad, o porque sirvan como base para la doble diferencia, deben ser observables. Dicho de otro modo, estos métodos se basan en que las condiciones de la evaluación de impacto son medibles, como ocurre con la antigüedad en el paro registrado, la edad, la nota en un examen o los ingresos del hogar. Si, por el contrario, el acceso a una PAE dependiese de una entrevista, y no se registrase ninguna valoración del desempeño en la misma, estaríamos ante un rasgo no observable por el evaluador y, por tanto, ante la dificultad de determinar si ha existido un sesgo de selección a la PAE, que dificultará o llegará a impedir la atribución de un impacto.

Los tres tipos de métodos que se han presentado para la evaluación cuasiexperimental comparten objetivos, aunque rara vez podrán ser utilizados de manera conjunta. Es probable que evaluaciones de impacto para programas autonómicos o locales utilicen alguno de estos métodos, por lo que el informe de evaluación de impacto deberá incluir información sobre la técnica concreta y sobre las decisiones tomadas en su aplicación. A la hora de elegir el método más adecuado, entre los factores a tener en cuenta se encuentran los objetivos específicos que justifiquen la evaluación, el diseño concreto del programa evaluado y la disponibilidad de datos de calidad.

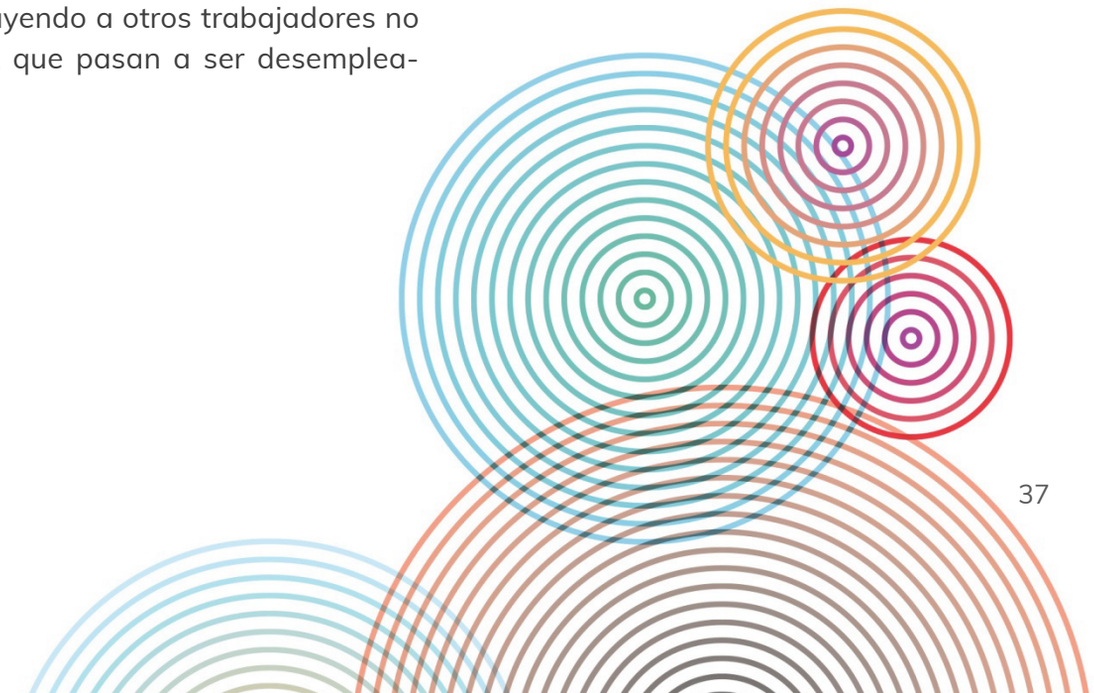
b. Evaluaciones macroeconómicas

Frente a las evaluaciones experimentales o cuasiexperimentales, que buscan obtener el impacto de las PAE sobre el colectivo objetivo, existe otro tipo de evaluación de impacto que trata de estimar el impacto sobre variables referidas al conjunto del mercado de trabajo. La diferencia es que las primeras proporcionarán estimaciones microeconómicas sobre, por ejemplo, la reducción del desempleo de los individuos participantes, o la creación de empleo adicional por parte de los autónomos que han recibido una subvención. En cambio, las segundas calcularán el impacto de las PAE sobre la tasa de desempleo agregada, la tasa de vacantes, o el aumento del empleo, todas ellas variables macroeconómicas.

Las evaluaciones macroeconómicas resultan útiles porque, al abordar el impacto desde esta perspectiva más general, permitirían obtener estimaciones muy interesantes sobre aspectos como los efectos sustitución o desplazamiento. El efecto sustitución viene dado porque la mejora del empleo de cierto grupo de participantes en PAE se consiga a costa del empeoramiento de otros colectivos. Esto ocurre, por ejemplo, si las empresas contratan personas con bonificación sin crear empleo neto, sustituyendo a otros trabajadores no bonificados, que pasan a ser desempleadas. El efecto desplazamiento recoge la posibilidad de que empresas con menores costes laborales gracias a las PAE puedan expulsar del mercado a otras que no gozan de esas ventajas.

Sin embargo, las evaluaciones macroeconómicas también tienen algunas desventajas. La más importante es que resulta muy difícil aislar el impacto de las PAE de otros factores que actúan sobre el mercado de trabajo de manera simultánea, como la evolución del ciclo económico, la política monetaria o la política fiscal. Otro problema es que el gasto e intensidad de las PAE se ven influidos también por la evolución del mercado de trabajo, con lo que pueden dar lugar a un problema de bidireccionalidad (las PAE tratan de influir sobre la inserción laboral, pero se ven influidas, a su vez, por la crisis económica, que reduce el presupuesto disponible para estos programas).

Por último, un rasgo distintivo de las evaluaciones macroeconómicas, frente a los otros tipos de evaluación de impacto presentados en este Cuaderno, es que por su naturaleza se llevan a cabo con un nivel de análisis estatal.



c. Metaevaluaciones

La metaevaluación consiste en un método radicalmente distinto a todo lo que se ha presentado con anterioridad. De nuevo, la idea es muy sencilla: se trata de hacer acopio de un número de evaluaciones del impacto de políticas similares que sea lo suficientemente grande como para permitir usar cada una de ellas como un dato, calculando el impacto medio.

La elaboración de una metaevaluación no requiere datos primarios, ni diseño previo de una PAE, ni siquiera coordinación con los gestores de los programas o medidas. La materia prima de estas evaluaciones son evaluaciones de tipo microeconómico realizadas previamente. La metaevaluación se basa en trabajo de despacho y lo que precisa es definir la PAE de interés, buscar evaluaciones empíricas del impacto de dicha PAE, y seleccionar cuidadosamente los indicadores que pueden servir de base para la metaevaluación.

Existen metaevaluaciones en todos los campos científicos. En los ámbitos más próximos a la política de empleo, durante los últimos 15 años se han generado multitud de estudios sobre políticas sanitarias, de cooperación al desarrollo y de políticas educativas. En el ámbito de las PAE, se dispone de diversos estudios de este tipo realizados a partir de evaluaciones hechas en Europa, pero también existen metaanálisis para otras partes del mundo. Normalmente las metaevaluaciones abordan el papel de medidas por tipos generales: programas de formación, subvenciones al empleo, medidas de intermediación y orientación, y planes de empleo público. El Recuadro 7 presenta sintetiza una metaevaluación publicada en 2021 sobre PAE para trabajadores maduros.

En definitiva, las metaevaluaciones son trabajos de lectura muy recomendable para responsables de la toma de decisiones en políticas de empleo. Este tipo de estudios trasciende al impacto de una medida concreta para ofrecer un panorama que debería tenerse en cuenta a la hora de determinar las prioridades de las políticas de empleo tanto a nivel estatal como autonómico o local.

Recuadro 7. Metaevaluación de PAE dirigidas a personas mayores de 50 años

Un ejemplo notable de metaevaluación de PAE es el estudio realizado recientemente por Miguel Ángel Malo y Guillermo Orfao sobre el papel que juegan estas políticas en la inserción laboral de las personas mayores de 50 años. Estos autores seleccionaron evaluaciones de impacto realizadas entre 2008 y 2016 en cinco países europeos (Alemania, Bélgica, Francia, Países Bajos y Suiza), logrando hacer acopio de 82 estimaciones de impacto. En todos los casos las políticas evaluadas se dirigían al colectivo de personas mayores de 50 años y los impactos se calculaban sobre la probabilidad de encontrar empleo en este grupo de población. Esta homogeneidad permitió elaborar una base de datos que sirvió de base para la metaevaluación.

Los resultados de este trabajo indican que las PAE recogidas muestran un impacto promedio ligeramente negativo sobre la probabilidad de encontrar empleo durante los dos años siguientes a la participación en las mismas. El impacto desaparece al cabo de ese plazo. Este resultado, que puede parecer sorprendente, tiene su explicación en el papel que juegan las diferentes PAE utilizadas por este grupo de población. Así, los programas de empleo público presentan un impacto promedio claramente negativo, reduciendo la probabilidad de empleo en cerca de un 4 por ciento durante más de un año. Este resultado coincide con el obtenido en otras metaevaluaciones anteriores sobre colectivos más amplios, que sitúan a la creación directa de empleo –los programas locales de empleo, se les denomina a veces- como la PAE menos efectiva.

En cambio, las políticas de formación muestran un impacto positivo sobre la población mayor de 50 años. Este resultado también puede sorprender, pues se considera que la formación es una inversión para el futuro y, a menos de 15 años de la edad de jubilación, optar por este tipo de programas no parecería la mejor idea. La metaevaluación concluye que la formación impacta positivamente sobre el acceso al empleo y que el mayor impacto se produce en torno a un año después de formarse, lo que puede suponer mucho tiempo para personas desempleadas en edad madura. Desafortunadamente, los datos no permitieron a estos autores establecer el impacto promedio de otra PAE muy atractiva para este grupo de población, como son las subvenciones a la contratación en el mercado de trabajo ordinario

7. Qué puede hacer la administración para facilitar las evaluaciones de impacto de políticas de empleo

En esta sección se presentan algunas pautas de interés para las relaciones entre los agentes encargados de hacer la evaluación de impacto y los agentes, normalmente personal de los Servicios Públicos de Empleo, que llevan a cabo la gestión y/o la supervisión de las políticas que se desean evaluar.

Dado que el objetivo último de la evaluación de impacto es proveer evidencia que resulte útil a los agentes que toman decisiones sobre PAE, parece razonable esperar que la actitud general de decisores, diseñadores, gestores y evaluadores de políticas confluya hacia la obtención de evidencia de la mayor calidad posible. Pero la actitud podría no ser suficiente, y por ello este Cuaderno trata de sintetizar aspectos clave para la evaluación de impacto.

a. La importancia de los datos

La calidad de los datos resulta crucial para la evaluación de políticas públicas en general y para la evaluación de impacto en particular. Este es uno de los aspectos en que la participación y colaboración del personal de la Administración resulta fundamental.

La **primera orientación** que cabe mencionar respecto a los datos en relación con las PAE es la necesidad de prever con suficiente antelación que las tareas evaluadoras no choquen con problemas derivados de la necesaria protección de datos personales. Dado que la participación en PAE antecede a la evaluación, y dado que las tareas de la evaluación no pretenden vulnerar ningún derecho de las personas ni de las empresas, es aconsejable que se requiera el permiso formal de individuos o empresas para utilizar sus datos de cara a la evaluación desde el primer momento, cuando aparezca el primer contacto entre la Administración y los interesados en los servicios de empleo. Esta indicación apela especialmente a los datos de registro, como las demandas de empleo.

Una **segunda recomendación** bastante evidente para todos los agentes implicados en las PAE, desde la perspectiva de contribuir a proveer evidencia y, con ello, a mejorar las políticas a medio y largo plazo, es el esmero en la recogida de datos. La importancia de esto no es fácil de transmitir y nunca resultará exagerada. En este caso podemos apoyarnos en que ya se han presentado los métodos cuasiexperimentales, y el lector comprenderá que sin buenos datos no va a funcionar bien ni el emparejamiento, ni la regresión en discontinuidad ni método estadístico alguno. Para lograr datos de calidad es aconsejable:

- Diseñar un sistema estable de recogida de datos y evaluarlo periódicamente.
- Favorecer la iniciativa de los diversos agentes a la hora de recoger datos adicionales a los ya existentes, de eliminar duplicidades o de revisar el sistema.
- Cuando se disponga de datos de registro, no por ello descartar, sino más bien valorar, por su complementariedad, el uso de datos adicionales procedentes de encuestas o cuestionarios.
- Dedicar recursos, tanto materiales como humanos, a la recogida y procesamiento de datos de calidad.
- Tratar de producir datos longitudinales cuando sea posible, prestando atención a la recogida de información sobre cambios de criterio estadístico a lo largo de las series resultantes.

En tercer lugar, para la evaluación de impacto resulta fundamental el poder disponer de datos lo más completos posible sobre el conjunto del colectivo al que se dirige cada programa. Esto es así con independencia de qué unidades de dicho colectivo resulten afectadas por el programa, cuales no se vean involucradas y cual sea el procedimiento de acceso al mismo.

Un problema frecuente de las evaluaciones de impacto es que se dispone de datos sobre las unidades participantes, pero no sobre el conjunto del colectivo al que se dirigía la medida. Esto dificulta enormemente la posibilidad de realizar comparaciones entre participantes y no participantes. No obstante, incluso cuando esto ocurra, merece la pena invertir en la obtención de datos sobre las unidades participantes, aunque sea como un mal menor. Por ejemplo, si se trata de personas, la disponibilidad de datos sociodemográficos previos a la participación en el programa a evaluar podría permitir una comparación posterior con un grupo de no participantes procedente de una base de datos de registro. En este sentido, además de las variables sociodemográficas más típicas (sexo, edad, estudios, etc.) cabe plantearse el obtener información detallada del historial laboral y de participaciones anteriores en PAE. Si se tratase de empresas, el principio básico es válido: mejorar la cantidad y calidad de la información, incluso si se carece a priori de un grupo de control.



b. La selección y la participación en políticas de empleo

En los procesos de atribución de causalidad resulta fundamental conocer los mecanismos o árboles de decisión que han configurado la selección de los participantes en las políticas de empleo evaluadas.

Así, partiendo en todo momento del colectivo al que se dirige la actuación, sea éste un grupo de población o un grupo de empresas, la claridad sobre el tipo de decisiones de selección y el orden de las mismas es clave.

Por ejemplo, la Figura 2 recoge un caso frecuente, en el que las unidades (personas, empresas) que forman el colectivo al que se dirige la PAE toman la decisión, en un primer momento, de solicitar su participación. En un segundo momento, los gestores de la PAE seleccionan al grupo participante, entre los candidatos, de acuerdo con los criterios de cada caso (cumplimiento de requisitos, perfilado de destinatarios, disponibilidad de plazas, presupuesto, cuotas de participación por subgrupos del colectivo, etc.)

Figura 2. Un ejemplo de patrón de selección de participantes en una PAE



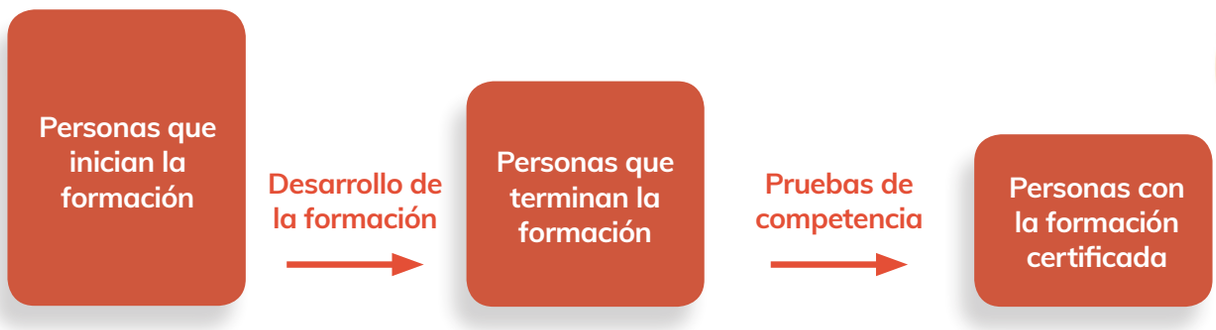
Fuente: elaboración propia.

Sin embargo, el patrón de selección puede ser diferente. Por ejemplo, en medidas de activación, el papel de los gestores de la PAE de seleccionar a los participantes puede preceder al de la autoselección de éstos. Por ejemplo, si desde los servicios públicos de empleo se selecciona a determinadas personas desempleadas a quienes se invita o motiva a acceder a una PAE, la participación efectiva dependerá de que finalmente acepten acceder a la misma.

De forma similar, la participación efectiva en una PAE puede sufrir variaciones a lo largo de su implementación. A lo largo del desarrollo de las medidas pueden producirse pérdidas de efectivos que, por razones diversas, lleven a que los participantes finales sean un subconjunto de los que iniciaron el proceso. Así, la Figura 3 recoge el ejemplo de un programa formativo donde a la pérdida de efectivos que podría acompañar al desarrollo de la formación se le añade la realización de pruebas de competencia. Así, las personas con la formación certificada podrán diferir en número de las que iniciaron los cursos.



Figura 3. Un ejemplo de selección a lo largo de un programa formativo



Fuente: elaboración propia.

La razón de que los gestores de las PAE estén llamados a colaborar clarificando los procesos citados en esta sección es clara: la evaluación de impacto tratará de determinar el momento o momentos más adecuados para encontrar grupos de comparación con los que contrastar la evolución posterior del grupo participante en la PAE.

Por ejemplo, es posible que las personas que inician, pero no terminan, la formación, no constituyan un grupo de interés para la comparación, quizá por su reducida entidad numérica. Pero también cabe la posibilidad de que este grupo, unido al de personas que no obtienen la certificación, sí sume los efectivos necesarios para configurar un grupo de comparación. En ese caso se podría evaluar el impacto de culminar con éxito la formación. Para ello se compararía a las personas que inician la formación, pero no la certifican al final, con quienes obtienen el certificado. También se puede considerar distintos tipos de tratamiento: personas formadas o personas formadas con certificación final, como se reflejaba en la evaluación recogida anteriormente en el Recuadro 4.

En definitiva, la recogida de datos de calidad (apartado a de esta sección), junto con la claridad en la representación de los árboles de decisión o acceso y de culminación de las PAE (apartado b) constituyen la mejor forma en que la Administración puede facilitar la tarea de los agentes evaluadores del impacto de las políticas

c. Procedimientos de recogida de datos durante y después de la PAE

La obtención de información durante la realización de una PAE y después de la misma puede interesar no solamente para la evaluación del impacto, sino para otro tipo de evaluaciones, como la de productos o la de resultados.

La elaboración de procedimientos es conveniente porque los agentes involucrados en la implementación de una PAE pueden ser numerosos y diversos. Por ejemplo, a las empresas de inserción que participen en un programa de inserción de colectivos desaventajados, o a las entidades formativas que participen en un programa de formación, o a las empresas de trabajo temporal que participen en un programa de orientación, debe trasladárselas un procedimiento específico de recogida de datos que facilite obtener información con criterios homogéneos.

El tipo de datos a recabar puede ser muy diverso, dependiendo del tipo de PAE. A título de ejemplo, algunos datos relativos a la realización, que pueden interesar posteriormente para la evaluación de impacto, son los que aparecen a continuación (no se explicita el tipo de PAE al que correspondería cada uno).

- Durante la realización de la medida:
  - o aceptación de participación en la PAE
  - o grado de implicación, asistencia o participación
- Al finalizar la medida:
  - o datos de finalización de la medida (compleción, finalización con éxito)
  - o grado de satisfacción
  - o datos de costes (costes totales, costes unitarios)
  - o datos de inversión realizada
  - o evaluación de resultados
- Con posterioridad a la medida:
  - o intensidad de búsqueda de empleo
  - o acceso al empleo
  - o mantenimiento del empleo
  - o creación de empleo
  - o otros datos de empleo (duración, tipo de contrato, jornada, etc.)
  - o grado de activación o participación en otras medidas de empleo

Los datos incluidos en el listado anterior pueden resultar necesarios para evaluar el impacto. En algunos casos su trascendencia es evidente, pues se trata del dato típico para valorar el impacto de la PAE, como el acceso al empleo. En otros casos la información se precisa porque el impacto va a tomar en consideración variables más específicas, adaptadas al programa en cuestión. Un ejemplo es el grado de activación, utilizado como referencia para valorar el impacto del proyecto piloto ERSISI en Navarra. Finalmente, diversos datos del listado anterior se necesitan porque se trata de variables que hay que controlar al tratar de establecer causalidad. Por ejemplo, la creación de empleo adicional como impacto de un programa de apoyo al autoempleo debería tener en cuenta la inversión realizada.



8. La integración de la evaluación de impacto en el proceso de la política

Esta última sección aborda algunas observaciones sobre la coordinación y la temporalización de las evaluaciones de impacto. De lo expuesto en las secciones anteriores se desprende que una evaluación de impacto no puede acometerse sin una planificación previa y sin considerar, por una parte, el encaje de las diversas tareas que la componen en el proceso de puesta en marcha de la política evaluada. Por otra parte, también será preciso establecer el horizonte temporal que se considera adecuados para que se pueda poner de manifiesto el impacto de las medidas a evaluar.

Así, dependiendo del tipo de evaluación que se pretenda llevar a cabo, tanto el diseño de la misma como los procedimientos de recogida de datos deben preceder a la implementación de la PAE. Lo ideal es que las personas encargadas de la evaluación participen en la puesta en marcha de la política desde su inicio, no tanto porque vayan a influir sobre el diseño de las medidas, como porque puedan anticipar problemas de la evaluación y evitarlos o, al menos, reducir sus efectos.

Existen multitud de ejemplos en que la falta de previsión de la evaluación de impacto impidió su realización. Uno muy reciente es el de la evaluación de PAE realizada por la AIReF en 2019, que trataba de abarcar un conjunto amplio de medidas en todas las comunidades autónomas y acabó limitándose a evaluar el impacto de dos programas: las medidas de diagnóstico individualizado en Aragón, y el PREPARA a nivel estatal.

Concretamente, haciendo mención a algunos de los contenidos presentados a lo largo del Cuaderno, cabe señalar la conveniencia de que las personas encargadas de la evaluación de impacto sean partícipes de las siguientes actividades:

- La selección de datos y los procesos de recogida de los mismos, como se ha expuesto en la sección 7, exigen una cuidada planificación en la que deben participar los evaluadores del impacto.

- Los tres mantras presentados en la sección 4 deben estar presentes de alguna forma en el proceso de diseño e implementación de la PAE, y no solo en evaluaciones experimentales. Especialmente, el conocimiento y control de los posibles sesgos de selección resulta una pieza fundamental de la evaluación de impacto y requiere atención en todo momento. Aunque la evaluación a realizar no sea experimental, los encargados de la misma tendrán como objetivo controlar los sesgos y su participación en el proceso contribuirá a que el método de evaluación elegido sea el más adecuado para cada caso.

- En el caso de que la selección de personas, empresas o entidades a una PAE se realice mediante entrevistas u otros procedimientos cualitativos, los evaluadores requerirán la oportunidad de documentar dichos procedimientos y, en la medida de lo posible, cuantificar en alguna medida el proceso de selección. El indicador cuantitativo puede ser muy útil, precisamente, para controlar sesgos de selección.

- El conocimiento de la literatura procedente de las metaevaluaciones de impacto de PAE, expuestas en la sección 6, es aconsejable en cualquier fase de una política, pues puede contribuir a mejorar el perfilado de los destinatarios de la misma.

- La posibilidad de que una política produzca efectos de desborde (spillovers) también debe ser tenida en cuenta, pues puede afectar y/o sesgar los resultados de la evaluación de impacto. Estos efectos remiten a que la participación en una medida de empleo desborde al colectivo al que va destinada. Un ejemplo es que una medida que no ha completado las plazas disponibles se abra a participantes de municipios inicialmente no previstos, en los que, por consiguiente, tampoco estaba previsto que se produjesen impactos.

Por último, la coordinación entre agentes encargados del diseño, de la implementación y de la evaluación atenderá a los plazos de la evaluación y, concretamente, a los horizontes temporales en los que cabe evaluar el impacto. Más concretamente, es aconsejable atender a dos cuestiones sobre plazos y tiempos:

- En primer lugar, si una política va a ser objeto de evaluaciones diferentes y complementarias, parece razonable que se establezcan mecanismos de coordinación entre las mismas. Por ejemplo, aunque puede parecer obvio, cabe recordar que la evaluación de impacto deberá atender a plazos más largos que la evaluación de resultados. Esto es así no solamente porque el impacto requiere un período de maduración mayor que el resultado, como se ha expuesto en la sección 2, sino también porque la evaluación de resultados puede ser un input muy útil para la evaluación del impacto.

- En segundo lugar, un tema clave es el horizonte temporal en el que cabe esperar que se pongan de manifiesto los impactos de PAE. La literatura de la evaluación de impacto recoge muchos casos en los que el impacto de las medidas deja de ser significativo a partir de un plazo medio-largo. Para pensar sobre esto, la intuición puede estar en comparar un programa de formación a desempleados con un programa de enseñanza reglada. Parece razonable esperar que el impacto de estudios técnicos o profesionales sea más duradero que el de cursos de menor duración o entidad.

Sin embargo, también se encuentra con frecuencia que el impacto de las PAE a corto plazo no resulta significativo o resulta negativo. Un ejemplo de ello es el conocido como efecto de bloqueo de la búsqueda de empleo, pues mientras una persona está participando en una PAE es posible que reduzca su intensidad de búsqueda o su disponibilidad para aceptar un empleo.

En definitiva, parece razonable detenerse y considerar cuidadosamente el horizonte temporal que se considera adecuado para que la medida evaluada produzca su impacto. Idealmente se dispondrá de datos de evolución a lo largo de un período amplio, que permita conocer impactos a corto, medio y largo plazo. No obstante, cuando esto no sea posible, es aconsejable considerar un horizonte cercano a los 12 meses tras la finalización de la medida a evaluar.

Por último, una opción que resulta muy económica para obtener estimaciones de impacto a medio y largo plazo es la de utilizar indicadores de síntesis en aquellos casos en que el impacto se evalúe sobre aspectos de cantidad. Ejemplos de indicadores de síntesis, en este sentido, son el número de días trabajados por una persona durante un período, o el número de días que la persona ha estado registrada como desempleada, o el número de días que una iniciativa de autoempleo ha estado funcionando desde su inicio.

Bibliografía

- AIReF (2019). Evaluación Del Gasto Público 2018. Estudio Programa Políticas Activas de Empleo. Disponible en: https://www.airef.es/es/politicas-activas/.

- AIReF (2020). Evaluación del gasto Público 2019. Estudio Incentivos a la Contratación. Disponible en: https://www.airef.es/wp-content/uploads/2020/10/INCENTIVOS/PDF-Web-Incentivos-a-la-contratacio%CC%81n-3.pdf.

- Arranz, J. M., García Serrano, C. y V. Hernanz (2013). Evaluación macroeconómica de las políticas activas del mercado de trabajo en España: ¿mucho ruido y pocas nueces? Revista Internacional del Trabajo, vol. 132 (2013), núm. 2, pp. 367-390. Disponible en: https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1564-9148.2013.00185.x.

- Card, D. y A. Krueger (2000). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. American Economic Review. Disponible en: https://pubs.aeaweb.org/doi/pdf/10.1257/aer.90.5.1397.

- Cebrián López, I., Cueto Iglesias, B. y F. J. Mato Díaz (2021). La evaluación de impacto de las políticas activas de empleo en España: el deseo y la realidad. Cuadernos Económicos de ICE nº 102, pp. 133-155. DOI: https://doi.org/10.32796/cice.2021.102.7314.

- Cueto, B., Toharia, L., García Serrano, C. y J. A. Alujas Los efectos de la formación ocupacional: ¿Importa la duración de las acciones? Hacienda Pública Española 195¬ (4/2010), pp. 9-¬36. Disponible en: https://ww.ief.es/docs/destacados/publicaciones/revistas/hpe/195_Art1.pdf.

- Laparra, M. y Martínez Sordoni, L. (2021). La integración de servicios sociales y de empleo en el debate entre protección y activación. Papers, en prensa. Disponible en: https://doi.org/10.5565/rev/papers.2839.

- Mato Díaz, F. J., Alonso Domínguez, A., González Arias, R. e I. Escudero Castillo (2021). ¿Funcionan las políticas de prevención del abandono escolar temprano en España? Análisis de impacto del Programa de Mejora del Aprendizaje y del Rendimiento (PMAR). Disponible en: https://elobservatoriosocial.fundacionlacaixa.org/es/-/funcionan-las-politicas-de-prevencion-del-abandono-escolar-temprano-en-espana.

- Orfao, G. y Malo, M. (2021). Are active labour market policies effective for the older unemployed? A meta-evaluation. Ageing and Society, 1-21. https://doi.org/10.1017/S0144686X21001288.

- Rebollo-Sanz, Y. y J. I. García Pérez (2021). Evaluación de Impacto de Políticas Activas de Empleo para Colectivos de Difícil Inserción Laboral. Cuadernos Económicos de ICE nº 102, pp. 157-187. DOI: https://doi.org/10.32796/cice.2021.102.7315.

- Van der Klaauw, B y L. Ziegler (2019). A Field Experiment on Labor Market Speeddates for Unemployed Workers. Disponible en: https://www.oecd.org/employment/emp/S2.2.%20vd%20 Klaauw_PAPER.pdf.


servicio navarro
de empleo -
nafar lansare
Gobierno de Navarra
Nafarroako Gobernua
Departamento de Derechos Sociales
Eskubide Sozialetako Departamentua
sárHe
CONSULTORES