


Evaluación contrafactual de las políticas de empleo para jóvenes

Guía metodológica

MÁRTON CSILLAG

JUDIT KREKÓ

ÁGOTA SCHARLE



Con la colaboración de: MARCO CENTRA,
MASSIMILIANO DEIDDA, LUCÍA GORJÓN, VALENTINA
GUALTIERI, IGA MAGDA, MARTA PALCZYŃSKA,
FRANCESCO TRENTINI, CLAUDIA VILLOSIO

Evaluación contrafactual de las políticas de empleo para jóvenes

Guía metodológica

RESUMEN

La presente guía ofrece una introducción paso a paso a la evaluación contrafactual de las políticas del mercado laboral para los jóvenes, centrándose en la utilización de datos administrativos. El objetivo de las evaluaciones contrafactuales de las políticas del mercado laboral es estimar el efecto causal de un programa en diferentes resultados, por ejemplo, la probabilidad y la duración de ser empleado. Esto implica una estimación de los efectos de la política en comparación con el estado hipotético en que no se introdujo la política, manteniendo las demás condiciones inalteradas. El proyecto *Youth Employment PartnerSHIP* evalúa las políticas del mercado de trabajo actuando principalmente por el lado de la demanda, proporcionando incentivos monetarios a los empleadores para la contratación de jóvenes. Este documento se centra en las evaluaciones de cuatro países: *Hungría, Italia, Polonia y España*.

Márton Csillag

Judit Krekó

Ágota Scharle

Con la colaboración de: Marco Centra, Massimiliano Deidda, Lucía Gorjón, Valentina Gualteri, Iga Magda, Marta Palczynska, Francesco Trentini y Claudia Villosio.

IMPLEMENTED BY:

The „Youth employment partnerSHIP” project is funded by Iceland, Liechtenstein and Norway through the EEA and Norway Grants Fund for Youth Employment.

Índice de contenidos

1. Introducción	4
2. Datos administrativos	6
3. El proceso de evaluación	12
3.1. Elección de un programa específico.....	12
3.2. Elección de la variable dependiente.....	13
3.3. Estrategia de identificación: elección del método de evaluación del impacto contrafactual (EIC).....	15
3.4. Efectos heterogéneos.....	22
4. Cuestiones relevantes para los encargados de las políticas.....	24
4.1. Interpretación de los resultados	24
4.2. Vinculación de resultados al diseño de los programas y su implementación	24
4.3. Validez externa, comparación con otros estudios y resultados estadísticos	25
4.4. Efectos de desplazamiento y efectos de “Peso Muerto”	25
4.5. Análisis coste-beneficio y coste-efectividad.....	26
Bibliografía	27
Anexo 1: Breve descripción de las evaluaciones en los cuatro países de interés.	29
Anexo 2: Glosario y abreviaciones	34
Índice de recuadros.....	37
Índice de figuras.....	37

1. Introducción¹

La presente guía ofrece una introducción paso a paso a la evaluación contrafactual de las políticas del mercado laboral para los jóvenes, centrándose en la utilización de datos administrativos. Las cuestiones principales se ilustran con los problemas prácticos que se han encontrado al evaluar los subsidios a la contratación de jóvenes en los cuatro países participantes en el proyecto *Youth Employment PartnerSHIP* (*España, Hungría, Italia y Polonia*). Sin embargo, esta guía también puede aplicarse en la evaluación de otros programas.

La finalidad de esta guía es proporcionar un conjunto de instrumentos prácticos para investigadores y directores de instituciones que realizan evaluaciones (instituciones dedicadas a la investigación, ministerios, servicios públicos de empleo) que tienen previsto iniciar y coordinar evaluaciones de programas de empleo para jóvenes. En la realización de esta guía se asume una comprensión básica de los conceptos estadísticos y de evaluación por parte de los lectores, así mismo, esta contiene un glosario de términos técnicos claves para su comprensión. Dado que el contenido y el objetivo de los programas difieren de un país a otro, y que las fuentes de datos disponibles que pueden utilizarse para las evaluaciones también son diversas, esta guía no presenta un marco de evaluación unificado². En su lugar, resume e ilustra cuestiones y problemas prácticos comunes que son propios de la evaluación contrafactual de las políticas diseñadas para el empleo juvenil.

Si bien los programas del Sistema de Garantía Juvenil (SGJ) se introdujeron en todos los Estados miembros de la Unión Europea, los elementos de estos no se han evaluado rigurosamente en todos ellos. La evaluación de políticas públicas es fundamental para determinar su impacto y las evaluaciones cuantitativas basadas en datos son esenciales si se desea conocer si los objetivos de estas han sido alcanzados, además de para mejorar la eficiencia de los programas. Los datos administrativos podrían permitir evaluar a posteriori, incluso en los países en que los programas se introdujeron sin un marco amplio de supervisión y recopilación de datos.

El objetivo de las evaluaciones contrafactuales de las políticas del mercado laboral es estimar el efecto causal de un programa en diferentes resultados, por ejemplo, la probabilidad y la duración de ser empleado. Esto implica una estimación de los efectos de la política en comparación con el estado hipotético en que no se introdujo la política,

¹ Agradecemos a Namita Datta, Ilf Bencheikh, Tibor Keresztély y András Svraka sus valiosos comentarios.

² La presente guía se centra en las evaluaciones contrafactuales realizadas sobre conjuntos de datos administrativos. Para consultar guías de seguimiento y evaluación más completas, a partir del desarrollo de una teoría de cambio, véase, por ejemplo, OCDE (2009) y PNUD (2009), Gertler et al. (2016), Morris et al. (2013).

manteniendo las demás condiciones inalteradas. Al no poder observar en la práctica dichas situaciones hipotéticas, es clave encontrar un grupo de control verosímil que pueda considerarse similar en todos los aspectos a los participantes en el programa, salvo que sus miembros no participan en el mismo. En todas las etapas de la evaluación, las consideraciones teóricas y las restricciones resultantes de las limitaciones de los propios datos también influyen en la elección de los investigadores.

El proyecto *Youth Employment PartnerSHIP* evalúa las políticas del mercado de trabajo actuando principalmente por el lado de la demanda, proporcionando incentivos monetarios a los empleadores para la contratación de jóvenes. En dos países (*Hungría* y *Polonia*), el proyecto se centra en las subvenciones a la contratación ofrecidas en la iniciativa europea Garantía Juvenil, que es un compromiso de todos los Estados miembros de la UE para proporcionar a los jóvenes ninis una oferta de empleo, educación, formación o prácticas de calidad dentro de un breve período de tiempo. En otros dos países (*Italia* y *España*), las medidas evaluadas tienen por objeto aumentar la estabilidad del empleo de los jóvenes, incentivando la contratación indefinida. Los principales elementos de las evaluaciones se esbozan en los recuadros 1 y 2 y se resumen con más detalle en el apéndice 1.

Las cuatro evaluaciones del proyecto *Youth Employment PartnerSHIP* se llevaron a cabo utilizando datos administrativos a nivel individual (registros de desempleo, expedientes de la seguridad social, registros de notificaciones obligatorias de los flujos de entrada y salida del empleo con contrato formal, etc.) que están disponibles en la mayoría de los países europeos. La presente guía se centra especialmente en los retos y cuestiones que plantea la utilización de datos administrativos en las evaluaciones de políticas.

El contenido de la guía se estructura de la siguiente manera. En la primera parte se resumen las cuestiones más importantes relativas al acceso y uso de los datos administrativo, dado que las limitaciones de los datos repercuten en el proceso de evaluación en todas sus etapas. La segunda parte recorre el proceso de evaluación cuantitativa, centrándose en las decisiones que deben adoptarse para garantizar la validez de los resultados obtenidos por los investigadores. En la última parte se resaltan las cuestiones y los problemas de las etapas posteriores del proceso de evaluación, relacionados con las cuestiones que surgen al presentar los resultados a los responsables de la formulación de políticas. Las secciones se encuentran basadas e ilustradas en las experiencias y ejemplos concretos de los cuatro países del proyecto *Youth Employment PartnerSHIP*. Los principales elementos de las evaluaciones se esbozan en los recuadros 1 y 2, y son mostrados con más detalle en el Apéndice 1. Además, en el Glosario puede encontrarse una definición de los principales términos técnicos.

Recuadro 1 Detalles de las evaluaciones de Italia y Polonia

	Polonia	Italia
Política	<p>Evaluado conjuntamente:</p> <ol style="list-style-type: none"> 1) Reducción de los costes de despido para los contratos de duración indeterminada (contrato con indemnización al despido creciente). 2) Reembolso de la seguridad social a los nuevos contratos de duración indeterminada y a las conversiones contratos temporales a indefinidos. 	<ol style="list-style-type: none"> 1) Subvención al coste salarial. 2) Formación en el trabajo, formación en el aula, realización de obras públicas, subvención al coste salarial (trabajo de intervención), vales de formación en el trabajo y de formación en el aula.
Requisitos de participación	Personas que no hayan tenido un contrato indefinido en los 6 meses anteriores a su contratación.	Jóvenes desempleados menores de 30 años para (1) y (2).
Duración, tipo	<ol style="list-style-type: none"> 1) Cambios permanentes en el reglamento de despido. 2) 36 meses de subsidio/incentivo a la contratación. 	<ol style="list-style-type: none"> 1) 12 meses de subsidio al empleo + 12 meses de empleo con las mismas condiciones laborales pero sin subsidio. 2) Desde 1 a 12 meses dependiendo de la medida.
Subsidio	<ol style="list-style-type: none"> 1) Reestructuración de la reglamentación de los despidos. 2) Reembolso del 100% de los costes laborales no salariales. 	<ol style="list-style-type: none"> 1) Subvención del coste salarial a la empresa o empleador cuya cuantía alcanza el salario mínimo + seguridad social (~500 + 100 € en 2018) 2) subsidios al empleo público y trabajos de intervención y vales de formación en el empleo.
Datos	Sisco (Sistema estadístico de comunicación obligatoria en línea). Registro administrativo público con información elemental sobre contrataciones, conversiones y despidos.	Registros de empleo de los SPE
Muestra	Contratos de empleo registrados entre 2014 y 2015.	2015-2017
Metodología	DD, con corrección paramétrica para el problema de selección de muestra e interacciones entre el tratamiento y las diferentes categorías de edad, con el fin de captar los efectos en las cohortes jóvenes.	<ol style="list-style-type: none"> 1) RDR combinado con DD. 2) PSM.
Resultados	Proporción de nuevas contrataciones con indefinidas sobre el total de contratos de trabajo registrados en 2015.	<ol style="list-style-type: none"> 1) Estar fuera del registro de desempleo de 12 a 36 meses después de la inscripción en el mismo. 2) Estar fuera del registro y no haber participado en ninguna PAE de 12 a 36 meses después de la inscripción en el mismo

2. Datos administrativos

Los datos administrativos suelen caracterizarse por una cobertura alta o completa, grandes tamaños muestrales y, en muchos casos, existe la opción de enlazar diferentes fuentes de datos³. Aunque, la fiabilidad de los datos administrativos suele ser mayor que la de los procedentes de encuestas, en algunos casos estos pueden ser también incompletos o imprecisos. En las evaluaciones de las políticas del mercado de trabajo, los conjuntos de datos administrativos más comunes son los procedentes de la seguridad social, el registro de desempleo, los datos de los servicios públicos de empleo (SPE) -incluyen los detalles de los programas del mercado de trabajo y sus participantes- y las bases de datos de los empleadores. Dado que los conjuntos de datos administrativos contienen datos a nivel individual o de empresa, el acceso a los conjuntos de datos administrativos suele estar estrictamente regulado y restringido por consideraciones de protección de datos. Tanto la forma de reunir y almacenar los datos administrativos como el entorno jurídico de la obtención y el manejo de estos.

³ Véase, por ejemplo, Caliendo et al. (2017) sobre los beneficios añadidos que se derivan del acceso a los datos basados en encuestas para la evaluación de las PAE.

Recuadro 2 Detalles de las evaluaciones de España y Hungría

	España	Hungría
Política	Contrato en prácticas	Periodo de prueba de 90 días
Requisitos de participación	Grado o Ciclo Formativo, menores de 30 años, desempleados.	Jóvenes menores de 25 años registrados como desempleados.
Duración, tipo	Mínimo 6 meses y máximo dos años. A jornada completa o parcial.	90 días máximo.
Subsidio	50-75% reducción en las cuotas de seguros sociales. Bonificaciones a la conversión indefinida de los contratos: 500€ por año en el caso de los hombres y 700€ por año en el caso de las mujeres, durante los primeros 3 años.	100% de los costes salariales (salario + contribución del empleador).
Datos	(1) Registros de empleo de los SPE, registros completos de contratos firmados, vinculados a	1) Registros de empleo de los SPE vinculados a
Muestra	(2) Muestra de los registros de la seguridad social: Muestra Continua de Vidas Laborales (MCVL).	2) Registros de la seguridad social e historial educativo.
Metodología	2002-2008	2015-2017 (variables relacionadas con el historial de empleo son construidas desde el año 2003)
Resultados	Regresión logística multinomial y modelo logit para comparar individuos participantes en el programa (CP) con los*** aptos que no participaron (contratos temporales).	1) <i>Nearest neighbour</i> y <i>PS matching</i>
	1) Probabilidad de que el individuo permanezca en la empresa después del CP.	2) DD. Analizando que el programa comenzó 9 meses después en la región de Hungría Central.
	2) Probabilidad de que el individuo firme un contrato indefinido en la misma empresa.	1) Estar trabajando 6 y 12 meses después de la aplicación del subsidio.
	3) Probabilidad de que el individuo firme un contrato indefinido en otra empresa	2) Salarios acumulados en 6 y 12 meses después de la aplicación del subsidio.

Dicho entorno jurídico difiere de un país a otro, por lo que los encargados de la evaluación podrían toparse con dificultades y un largo proceso administrativo para acceder a los datos requeridos. Se plantean, por tanto, muchas cuestiones al utilizar los datos administrativos⁴:

- **¿Existen bases de datos administrativos a nivel nacional?**

En el ámbito de la evaluación influye el hecho de que los datos administrativos en cuestión se registren y procesen en una base de datos a nivel nacional o a nivel de divisiones administrativas territoriales. En *Hungría*, *España* y *Polonia*, los datos de los Servicios Públicos de Empleo (SPE) están integrados en una base de datos a nivel nacional, mientras que *Italia* tiene un registro administrativo de las relaciones laborales a nivel nacional denominado SISCO.

- **¿Existe la posibilidad de vincular una base de datos administrativa con otra con el fin de enriquecer el conjunto de variables de control y de resultados?**

⁴ Para una guía paso a paso sobre el uso de datos administrativos en las evaluaciones contrafactuales del impacto de los Fondos Sociales Europeos, véase también EC (2020).

La información de una determinada persona suele registrarse en diversas bases de datos de diferentes organismos, por ejemplo, instituciones de la seguridad social, oficinas de empleo, servicios de formación o autoridades tributarias. La vinculación de diferentes fuentes de datos tiene un gran potencial para aumentar el alcance de la información disponible. Por ejemplo, en *Hungría*, los datos de los Servicios Públicos de Empleo pueden vincularse con los datos administrativos de la autoridad encargada de las pensiones que contienen datos sobre el empleo y los salarios de todos los solicitantes de empleo que aparecen registrados en los datos del SPE. Esto permite utilizar más variables de resultados (como la calidad del empleo), así como un conjunto más rico de características observables (incluido un historial de períodos de empleo). En *España*, FEDEA tiene acceso a dos bases de datos administrativas que pueden ser enlazadas entre sí, como es el caso de la Muestra Continua de Vidas Laborales (MCVL) y los registros de desempleo y de contratos firmados mensualmente en *España* de los SPE. Sin embargo, en muchos casos, la vinculación de diferentes fuentes de datos está estrictamente regulada y limitada o prohibida por consideraciones de protección de datos y de privacidad. Por ejemplo, en *Polonia*, los datos sobre empleo y salarios de los SPE no pueden vincularse a los datos de la Institución de la Seguridad Social, salvo en los casos estrictamente definidos por ley como, por ejemplo, la presentación de informes a la Comisión Europea.

- **¿Cuál es el procedimiento de acceso a los datos?**

No existe una reglamentación a nivel de la UE sobre el acceso a los datos ya que los requisitos administrativos difieren de un país a otro. Por ejemplo, en *España*, la base de datos de la MCVL puede obtenerse solicitando el acceso a la Seguridad Social. En general, los investigadores no pueden acceder fácilmente a los datos de los SPE a los que, sin embargo, FEDEA tiene acceso, tras la firma de un acuerdo de cooperación entre ambas partes. No obstante, no existen procedimientos formales para acceder a los datos de la Seguridad Social y de los servicios públicos de empleo de *Polonia*, donde el proceso es discrecional⁵.

- **Antecedentes institucionales y uso de datos administrativos**

Se aconseja encarecidamente a los investigadores que se pongan en contacto con los propietarios de los datos y aclaren los antecedentes institucionales, la finalidad original de los mismos, así como, del proceso y las pautas seguidas en su recopilación. La utilización y, por consiguiente, la fiabilidad de la calidad de dichos datos puede diferir incluso dentro de una misma base de datos, dependiendo de si una determinada información se utiliza en algún procedimiento contable, de evaluación o administrativo. Conocer en profundidad los mecanismos de recopilación, almacenamiento y

⁵ *Polonia* está construyendo actualmente una plataforma analítica integrada, constituyendo un sistema central que facilitará el análisis de los datos administrativos de diversas fuentes.

procesamiento de datos es esencial para establecer la estrategia de estimación. En algunos casos, especialmente si la recopilación y evaluación de los datos es competencia de la misma organización u organizaciones cooperantes, la institución evaluadora puede influir en la forma y contenido de la recopilación de los datos, mejorando así el propósito específico de la misma. La retroalimentación entre el proceso de evaluación y la recolección de datos es especialmente útil en el caso de evaluaciones a largo plazo o periódicas.

- **Calidad e integridad de los datos**

Como los conjuntos de datos administrativos suelen abarcar un conjunto completo de observaciones -o una muestra aleatoria de todas las observaciones-, la representatividad de los datos suele estar garantizada. Sin embargo, los conjuntos de datos administrativos también pueden contener fallos importantes y/o datos faltantes, especialmente en el caso de las variables cuya cumplimentación es opcional. El investigador debe evaluar siempre la magnitud y el carácter aleatorio de los datos que faltan y aplicar correcciones si fuera necesario. La integridad y calidad general del conjunto de datos puede comprobarse, por ejemplo, comparándolo con datos agregados de fuentes alternativas.

- **¿Se puede disponer de los mismos datos para los grupos de tratamiento y de control?**

Un problema frecuente de las evaluaciones de políticas que utilizan datos procedentes de los servicios públicos de empleo es que la base de datos (o la base de datos vinculada) contiene datos diferentes (distinta estructura, otras variables de diferentes fuentes) para los participantes en el programa y para el posible grupo de control, lo que podría dificultar definir unas variables de resultados y de control comunes en la base de datos. En muchos casos, los registros de desempleo únicamente contienen información limitada sobre la situación real de empleo de aquellas personas registradas como solicitantes de empleo después de salir del registro o de completar un programa de mercado de trabajo.

En *Hungría*, los datos administrativos de los servicios públicos de empleo sólo contienen los datos sobre la situación laboral de los participantes en un programa de empleo hasta 180 días después de la finalización de este. Sin embargo, no se disponen datos sobre el empleo de los solicitantes de empleo registrados que no hayan participado en una PAE. Una posible solución es utilizar la condición de desempleado inscrito como variable de resultado que también está disponible para los solicitantes de empleo no participantes. Se puede observar si una persona sigue siendo un solicitante de empleo registrado, y se puede suponer que la persona que ya ha dejado el registro se encuentra, en cambio, empleada.

Sin embargo, el investigador debe tener presente que el hecho de estar fuera del registro de desempleo puede ocurrir por muchas razones además de estar de empleado de nuevo, ya que el conjunto de personas que ya han salido del registro contiene muchas personas inactivas. En la evaluación de *Hungría*, esta cuestión se resolvió utilizando los datos de empleo, de la base de datos vinculada de la seguridad social, como variable de resultado. La situación es similar a la de los datos de los servicios público de empleo en *Polonia*, donde la situación laboral sólo puede determinarse si el individuo se encuentra inscrito en el registro empleo y no ha participado en ninguna PAE o utilizando la información disponible sobre el motivo por el que se ha dado de baja en la oficina de empleo.

- **¿Se puede identificar en los datos a los participantes en el programa o sólo a aquellos que cumplen los requisitos para hacerlo?**

En muchos casos, la base de datos no contiene información que permita identificar a los participantes en el programa, sino sólo a los que reúnen los requisitos para participar en él. En este caso, el investigador puede estimar el efecto del programa sólo en la subpoblación que reúne las características para participar, pero, sin embargo, no puede identificar el efecto del programa en los participantes. Utilizando la terminología de evaluación de programas, el investigador no puede identificar ni el efecto medio del tratamiento (conocido como ATE por sus siglas en inglés) ni el efecto medio del tratamiento en los tratados (conocido como ATT por sus siglas en inglés), sólo la intención de tratar (conocido como ITT por sus siglas en inglés).

Esta cuestión es menos preocupante si el programa es prácticamente universal (por ejemplo, un vale o cuasivale universal al que tienen derecho todos los jóvenes desempleados o uno que fuera obligatorio para algunos grupos). Sin embargo, la distinción podría cobrar importancia si la participación es parcial y la decisión sobre la participación en el programa no está claramente definida. Por ejemplo, si el efecto estimado del programa en el conjunto de la población que cumple los requisitos de participación es bajo, el investigador podría no estar seguro de si el resultado del programa es débil debido a la baja tasa de participación en el programa o a su escaso efecto. Esta distinción puede ser importante, ya que las consecuencias de la aplicación de una política son diferentes en el caso de una baja tasa de participación y en el caso de un efecto bajo sobre los participantes en el programa. La primera podría requerir la intensificación de las actividades de divulgación, mientras que la segunda podría derivarse de un mal diseño del programa.

Los conjuntos de datos procedentes de los servicios públicos de empleo suelen contener detalles de las PAE y, además, permiten identificar a los participantes en el programa. En *Hungría*, la base de datos de los SPE contiene los detalles de los programas de políticas del mercado de trabajo, incluyendo la especificación del tipo de programa y su fecha de inicio y finalización. En el caso de los programas de subvenciones al coste

salarial, ni la cuantía del subsidio ni los salarios se incluyen en la base de datos, aunque estos datos se pueden obtener al vincular el conjunto de datos con los de la seguridad social.

En *España*, ambas bases de datos permiten la perfecta identificación del Contrato en Prácticas (CP), contrato en el que se centra este análisis. En las dos bases de datos utilizadas en la evaluación española (la MCVL y los registros de desempleo de los SPE), todos los individuos que firman un CI están perfectamente identificados, al igual que las fechas de inicio y finalización de sus respectivos contratos. Además de información sobre los CI, disponemos de información sobre el resto de los diferentes contratos, por lo que podemos utilizarlos como controles. Sin embargo, en *Italia*, el registro administrativo (SISCO) no contiene información sobre los participantes en el programa, por lo que el modelo de evaluación italiano sólo estima los efectos sobre la población que cumple los requisitos de participación. En *Polonia*, la base de datos de los SPE contiene los detalles de los programas de políticas del mercado de trabajo, incluyendo la especificación del tipo de programa y su fecha de inicio y finalización.

- **¿Qué grado de detalle tiene el conjunto de características observables?**

El conjunto de características observables disponibles que pueden utilizarse como variables, influyen tanto en la estrategia de identificación como en el alcance para analizar los efectos heterogéneos en diferentes subpoblaciones. Por consiguiente, antes de elaborar la estrategia de evaluación, es fundamental determinar cuáles son las variables disponibles o identificar las que pueden obtenerse. En los conjuntos de datos administrativos individuales el género, la edad y las variables regionales suelen estar incluidas, sin embargo, más allá de esta información básica, las bases de datos administrativas difieren en cuanto a las variables disponibles.

El nivel de educación suele ser una variable fundamental que afecta en gran medida en los indicadores de empleo, por lo que suele ser una de las variables más importantes en la evaluación de las políticas del mercado laboral. El nivel de educación está disponible en los datos de los SPE de *Hungría* y *Polonia* y en el SISCO de *Italia*. Sin embargo, en *España*, el nivel educativo proviene del censo y sólo se actualiza cada diez años.

Otras variables útiles son la información sobre los hogares (por ejemplo, el número de cohabitantes en *España*), la información sobre diferentes prestaciones (por ejemplo, las prestaciones de maternidad en *Hungría*) y la presencia de un/a menor en el hogar (*Polonia*).

En el caso de los registros de empleo, el SISCO de *Italia* contiene información sobre la industria (NACE) y la ocupación, aunque esta última no figura en la base de datos de la seguridad social española.

Además de las características observables de los individuos, los contratos y las empresas, las variables macroeconómicas externas también podrían ser útiles para

controlar el entorno del mercado laboral. Por ejemplo, en *Italia*, el valor añadido por sector económico después de la contratación con el fin de descontar la expectativa de crecimiento de los empresarios en el momento en que deciden contratar.

Además de la variedad de características observables, también podría ser relevante la duración de los episodios de empleo antes del programa presentes en la base de datos. En algunos casos, podría ser útil para la evaluación disponer de datos anteriores a la introducción o a la reestructuración de un programa, de modo que la situación anterior al cambio de política pueda compararse con la que prevalece después en un marco de Diferencias-en-Diferencias (DD). En otros casos, pueden utilizarse como controles las variables de series extensas de datos, como sucede en *Hungría*, donde el historial completo de empleo y los registros disponibles en la base de datos de la seguridad social permiten elaborar un historial laboral de una persona y esto permitir captar la experiencia y la motivación de una persona.

3. El proceso de evaluación

3.1. Elección de un programa específico

Muchas políticas del mercado laboral están compuestas por numerosos elementos procedentes de distintos programas. El organismo evaluador (en colaboración con los ejecutores) deben decidir cuáles son los elementos objeto de evaluación. El proceso de evaluación puede abarcar todo el programa, que a su vez puede contener varios subprogramas, o puede centrarse en un único programa homogéneo bien definido. Esta elección está limitada por las normas del programa y por su implementación práctica. Por ejemplo, si los elementos del programa suelen combinarse para satisfacer las necesidades específicas de cada demandante de empleo, tal vez no sea posible evaluar un elemento separado (a menos que se disponga de datos completos y detallados sobre las necesidades de los demandantes de empleo). Por ejemplo, en el caso de la evaluación realizada en *Hungría*, y dado que el plan del programa permite esta combinación en casos concretos, el periodo de prueba de 90 días suele ir seguido de una subvención al coste salarial parcial de una duración de hasta 8 meses. No tener en cuenta dicha combinación podría ocasionar un grave sesgo al alza en la tasa de empleo 6 meses después de la finalización del periodo de prueba del empleo, ya que para los que tienen acceso a este subsidio posterior, la tasa de empleo será cercana al 100%. Sin embargo, en la mayoría de los casos, los programas de mercado laboral ofrecidos por los servicios públicos de empleo en *Polonia* se componen de un solo elemento, bien sea formación, prácticas o subsidios empresariales. La combinación de los diferentes elementos es muy poco común (aunque la ayuda para la búsqueda de empleo está disponible para todos).

La elección del programa influye en la elección del grupo de tratamiento y de los posibles grupos de control. Si los participantes no están identificados en los datos y los investigadores observan sólo el grupo que cumple los requisitos de participación, la evaluación de un elemento específico del programa continúa sin ser una opción posible.

Es fundamental analizar y precisar todo el diseño y los detalles del programa con el fin de valorar y determinar el alcance de una posible evaluación:

¿Cuál es el problema que intenta solventar el programa y cómo los diseñadores son conscientes del mismo? ¿El programa era obligatorio o voluntario, cuáles son los criterios de selección y cuál es el alcance geográfico y temporal del programa? ¿Qué sabemos sobre el objetivo, el desarrollo temporal del programa y la escala del programa? ¿Qué otros detalles son importantes aclarar para observar todos los efectos y deficiencias del programa? ¿Cuáles son las combinaciones de programas más frecuentes y cuáles son sus consecuencias en la evaluación/interpretación de estos?

En *Italia*, por ejemplo, las políticas que se están evaluando son aquellas relacionadas con los incentivos a la contratación -en forma de reembolsos a la seguridad social -y la sustitución del contrato indefinido por un contrato con indemnización al despido creciente- lo que da lugar a una reducción de los costes de despido. La primera se aplicó sólo a aquellos solicitantes voluntarios durante un periodo de tiempo limitado, mientras que la segunda modificó la disposición de la ley y se aplicó a todos los nuevos contratos indefinidos después de la introducción de la ley. Ambas políticas entraron en vigor con poco tiempo de diferencia, por lo que la evaluación estima el efecto simultáneo de las dos políticas.

3.2. Elección de la variable dependiente

El punto de partida para la elección de la variable sobre la que se mide el efecto y del método es establecer el mecanismo de impacto y el objetivo del programa. ¿Cuáles son sus efectos esperados? ¿Cuál es la cronología de sus efectos? ¿Causa su aplicación algún efecto secundario? Aclarados y definidos los efectos del programa, surgen las siguientes preguntas:

- ¿Cuál es el resultado de interés: situación laboral, empleo en la misma empresa, empleo en el mercado laboral general, salarios (acumulados), etc.
- ¿Qué horizonte temporal se analiza, corto o largo plazo (más de un año después del programa)? Lo ideal sería que la investigación analizara los resultados en diferentes horizontes y evaluara cómo varía el impacto en los distintos horizontes analizados. En la evaluación de los programas de subvenciones al coste salarial, los investigadores deben tener en cuenta que las subvenciones a los costes salariales pueden estar vinculadas a una obligación de empleo

posterior (no subvencionado), por lo que sólo deben considerar el resultado una vez transcurrido dicho periodo de compromiso.

Las consideraciones conceptuales pueden verse muy limitadas por la disponibilidad de datos. En muchos casos, no se dispone de resultados a largo plazo, además de la ausencia de datos salariales en las bases de datos de los SPE.

Las subvenciones al coste salarial suelen tener por objetivo aumentar la probabilidad de encontrar un (buen) empleo, por lo que los resultados más relevantes son la probabilidad de estar empleado durante un determinado periodo de tiempo después de completar el programa, así como también podrían ser importantes los ingresos salariales totales. Por ejemplo, el objetivo del periodo de prueba de 90 días, programa en el que se centra la evaluación de *Hungría*, es introducir en el mercado laboral a los jóvenes y disminuir el riesgo de su contratación asociado a su falta de experiencia. En teoría, la participación en el programa mejora las perspectivas de empleo. En concreto, se espera que después de la finalización del programa, la probabilidad de estar empleado aumente e, incluso, que el programa aumente los salarios al mejorar las perspectivas de encontrar un trabajo mejor remunerado por proporcionar experiencia y aumentar el valor del currículum vitae de la persona joven en cuestión. Por lo tanto, las principales variables sobre las que se mide el efecto son, por una parte, la situación laboral transcurridos 0,5, 1, 2, etc. años después del programa y, por otra parte, el diferencial salarial acumulado en el mismo horizonte temporal. Sin embargo, la disponibilidad de datos limita el alcance temporal de las variables dependientes, en este caso concreto no se dispone de resultados a largo plazo, ya que la política se introdujo en 2015 y la muestra disponible incluye sólo datos de tres años, abarcando el período 2015-2017.

La utilización de los datos de los servicios públicos de empleo (SPE) para evaluar las políticas activas empleo (PAE) en *Polonia* ofrece una selección limitada de variables resultado. En dichos datos, se puede observar si una persona regresó al registro de desempleo después de un determinado número de meses. Sin embargo, estar fuera del registro de desempleo no sólo incluye a las personas que se encuentran empleadas sino también a las inactivas y a las participantes en el programa. El resultado, por tanto, puede reducirse a no estar en el registro de desempleo y a no participar en ninguna PAE.

El objetivo final del contrato en prácticas en *España* y de los incentivos a la contratación en *Italia* es mejorar la estabilidad del empleo e incrementar la proporción y duración de los contratos indefinidos. Por lo tanto, las variables relevantes sobre las que se mide el efecto muestran si la situación laboral de la persona joven en cuestión es más estable después de participar en el programa. Las principales variables resultado de la evaluación española son tres: en primer lugar, la probabilidad de permanecer en la misma empresa después de la finalización del contrato en prácticas; el segundo resultado es la probabilidad de permanecer en la empresa con un contrato indefinido y, en tercer resultado es la probabilidad de encontrar un contrato indefinido en una

empresa diferente después del contrato en prácticas, lo que también indica la estabilidad de la situación laboral en el mercado de trabajo como un efecto indirecto de las prácticas. El objetivo de las políticas evaluadas en *Italia*, como son los descuentos en la contribución a la seguridad social y la reducción de los costes de despido, es similar: aumentar la estabilidad del empleo y prolongar la duración de los contratos. Por consiguiente, la variable principal de resultado de la evaluación de los incentivos a la contratación indefinida es la proporción de nuevos contratos indefinidos en 2015 respecto del total de contratos de trabajo registrados durante dicho año.

3.3. Estrategia de identificación: elección del método de evaluación del impacto contrafactual (EIC)⁶

La principal dificultad para medir el impacto causal de cualquier política es que no se observa la situación hipotética o contrafactual en la que la misma persona no recibe el tratamiento. Los investigadores deben inferir las situaciones contrafactuales a partir de los datos observados.

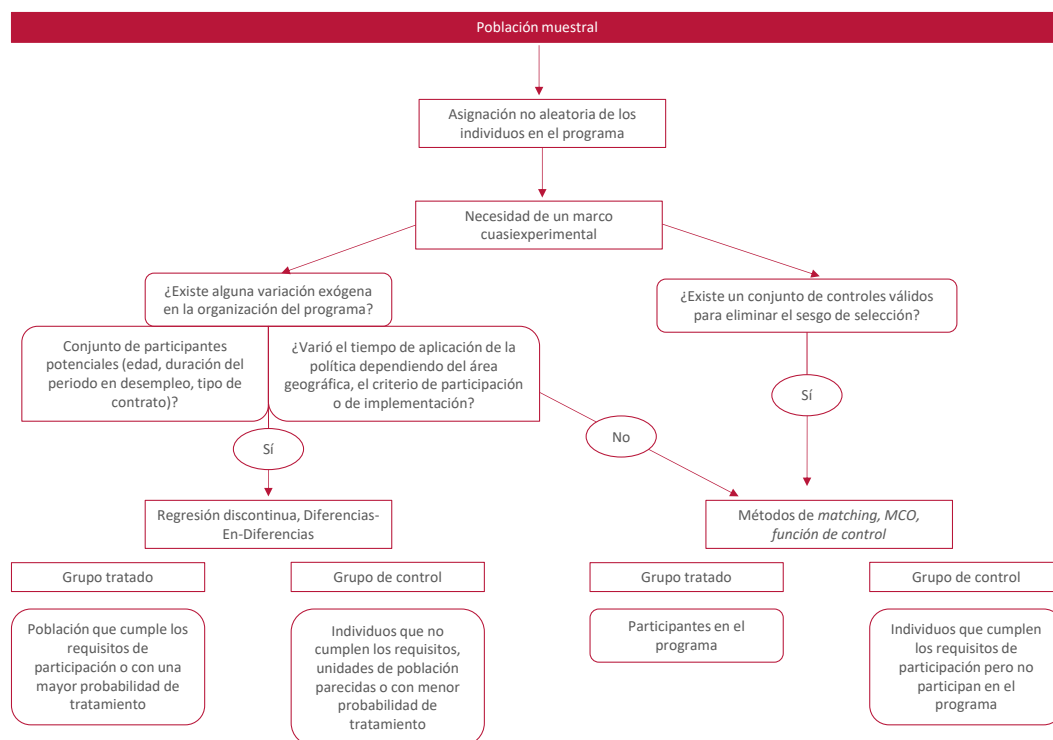
Una situación ideal sería que la evaluación de las políticas del mercado de trabajo se basara en la creación de un marco experimental mediante la selección aleatoria de participantes dentro de un grupo de personas que reúnan las condiciones necesarias para participar en el programa. Cuando no se garantiza dicho marco experimental, la selección de participantes en el programa no es aleatoria y, por tanto, los no participantes pueden ser diferentes en cuanto a características observables e inobservables. Esto podría sesgar la evaluación de dos maneras diferentes. En primer lugar, la variable dependiente o resultado (que puede ser, por ejemplo, tener un empleo, u oportunidades salariales) de los participantes en el programa, puede ser diferente de la de los no participantes incluso sin tener en cuenta el programa (esto sería la diferencia de base entre los dos grupos, participantes y no participantes). En segundo lugar, el programa podría afectar a los participantes de manera diferente de lo que afectaría a los no participantes si éstos participasen en el programa (efecto de tratamiento heterogéneo).

Además, un tercer sesgo surge cuando la aplicación del tratamiento afecta a los posibles resultados del grupo de control, así como del grupo tratado. Por ejemplo, cuando los efectos de los compañeros influyen en los resultados incluso cuando estas no son participantes (violación del supuesto de Valores Estables de la Unidad de Tratamiento, SUTVA). En ausencia de circunstancias experimentales, los investigadores deben aplicar

⁶ Los conceptos y términos necesarios más importantes y los métodos se encuentran descritos en el Glosario.

un marco cuasiexperimental que simule circunstancias experimentales ocupándose de la selección no aleatoria en el programa (véase Figura 1).

Figura 1 Esquema de la lógica de la evaluación del impacto contrafactual

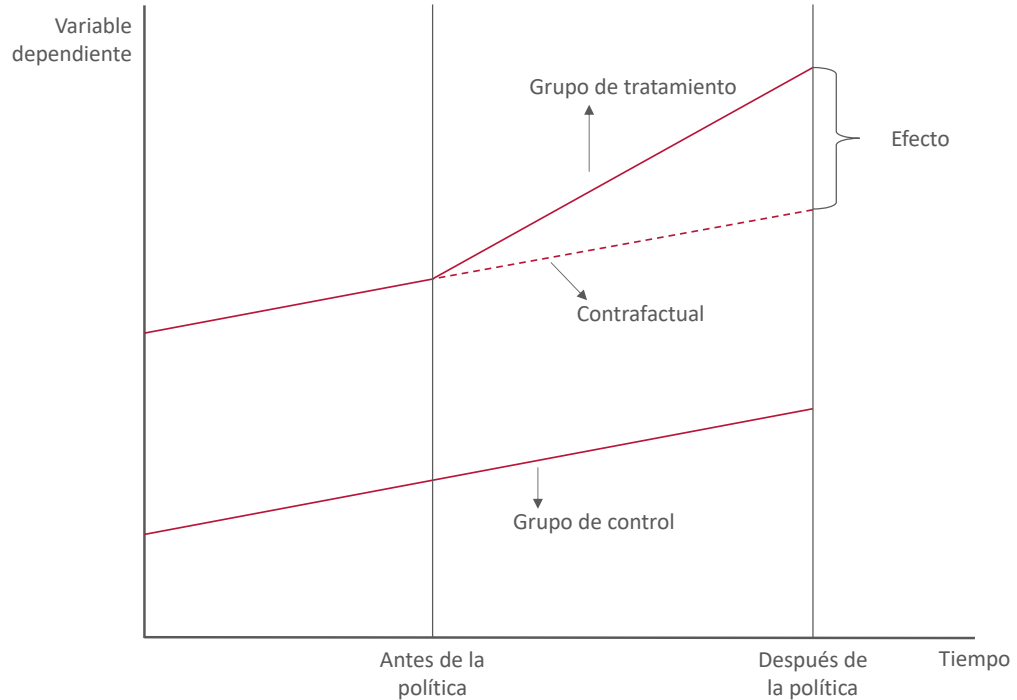


Existen dos vías principales para garantizar un marco cuasiexperimental. Una opción es encontrar alguna variación exógena en el marco del programa y aprovechar dicha exogeneidad. El objetivo de este proceso es distinguir a las personas que están excluidas exógenamente del programa porque no cumplen con los criterios de participación definidos en la normativa del programa, de las que son seleccionadas exógenamente en el programa o de las que cumplen los requisitos para participar en el mismo. Los ejemplos más comunes de restricciones exógenas que podrían identificar el efecto de la política son los siguientes:

- Grupo de solicitantes que cumplen los requisitos (por ejemplo, limitaciones relacionadas con la edad, con el tiempo en el desempleo o sin contrato a largo plazo)

La estrategia de identificación en *Italia* utiliza el hecho de que sólo tienen derecho al incentivo a la contratación aquellos que no han sido empleados con un contrato indefinido que haya vencido en los últimos 6 meses. Esta estrategia de identificación se basa en un modelo DD aplicado a los contratos inscritos en el SISCO, entre el 1 de enero y el 31 de diciembre de 2015, en comparación con los registrados en los años anteriores en los dos grupos, los que cumplen los requisitos para participar y el grupo de control.

Figura 2 Método de diferencias-en-diferencias (DD) ilustrado



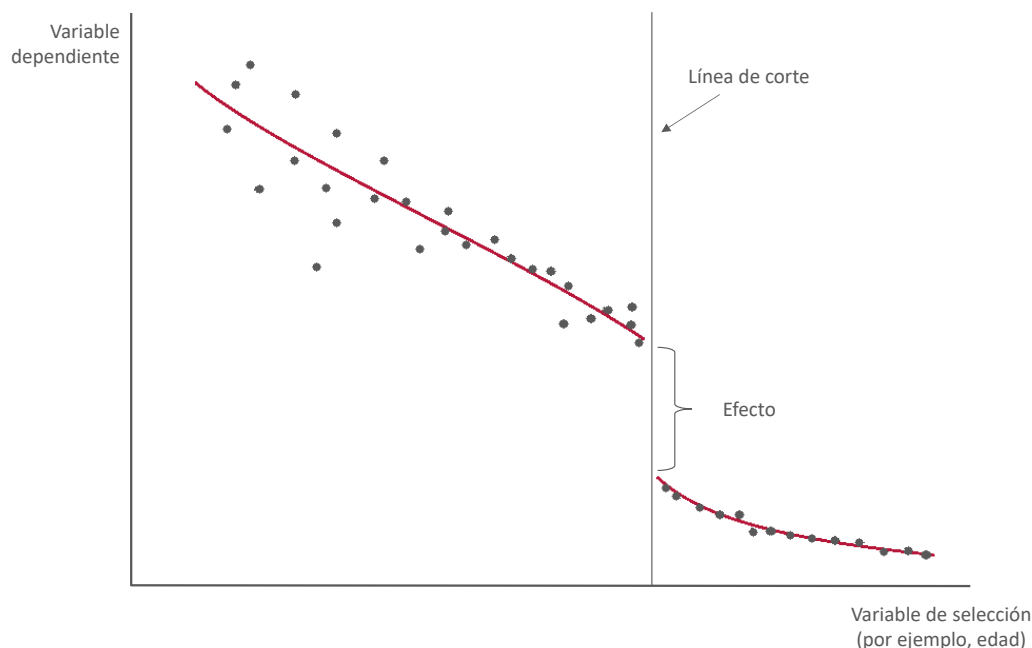
Se definieron dos grupos diferentes para estimar la proporción adicional de nuevos contratos de empleo de duración indefinida que no se habrían firmado de no aplicarse ambas políticas, respecto de los registrados durante el año 2015. Los dos grupos se definen mediante el criterio de selección. El grupo compuesto por aquellos individuos que cumplen los requisitos incluye los contratos de trabajo de las personas contratadas durante el periodo 2014-2015 y no empleadas anteriormente con un contrato de duración indeterminada que haya expirado en los seis meses anteriores. El grupo de control contiene a personas que habían sido contratadas en 2015, pero que habían sido despedidas hasta 6 meses antes de comenzar el nuevo contrato indefinido, estando ya contratadas con un contrato indefinido anterior, y que por tanto no optaban a la seguridad social (para una ilustración del modelo DD, véase Figura 2)

Como la base de datos no proporciona información directa sobre las prestaciones y permite identificar únicamente a las personas que reúnen los requisitos para participar, la evaluación estima el efecto de la política en dicho grupo de personas aptas (intención de tratar, efecto ITT).

Dado que los programas de Garantía Juvenil sólo están disponibles para jóvenes menores de cierta edad, lo que limita la participación en los mismos bajo dicho criterio, una solución sencilla sería explorar el límite de edad en un marco de regresión discontinua, como se realizó en una de las evaluaciones de *Polonia*. El programa de subvenciones al coste salarial durante 12 meses en *Polonia* estaba previsto sólo para los

desempleados menores de 30 años, lo que permite utilizar el límite de edad en el marco del diseño de regresión discontinua (en lo sucesivo, DRD) (véase Figura 3).

Figura 3 Ilustración del Diseño de Regresión Discotinuua (DRD)



Sin embargo, hay otras normas institucionales que cambian en el umbral de los 30 años. En primer lugar, los trabajadores desempleados menores de 30 años tienen a su disposición una mayor variedad de programas (vales de formación, de reasignación y de periodo de prueba) y algunas PAE siguen reglas diferentes (por ejemplo, periodos de prueba con una duración de hasta 12 meses en el caso de los trabajadores más jóvenes y hasta 6 meses en el caso de los trabajadores de más edad). En segundo lugar, las fuentes de financiación difieren en parte entre los desempleados más jóvenes y los de más edad. En el caso de los trabajadores más jóvenes, se distribuye a través de un programa operativo gestionado de manera centralizada, mientras que en el caso de los trabajadores de mayor edad se financia a través de programas regionales gestionados a nivel de voivodías⁷. Esto puede causar diferencias en la ejecución y la existencia de los programas entre los trabajadores desempleados menores y mayores de 30 años. En la estrategia de identificación se utiliza el hecho de que antes de la implantación del programa de subvenciones al coste salarial ya existían todas las diferencias institucionales entre el grupo de tratamiento y el grupo de control, con la excepción del programa de subvención de coste salarial que se está evaluando. Por lo tanto, se puede combinar el diseño de regresión discontinua (DRD) con el marco de diferencias-en-diferencias (DD). Cabe mencionar que los resultados obtenidos aplicando este enfoque

⁷ Un voivodado es la subdivisión administrativa de más alto nivel de *Polonia*. En general, corresponde al concepto de una provincia o estado en otros países.

se basan en la comparación de los grupos de tratamiento y control que se encuentren sólo cerca del umbral. Esto puede plantear algunas limitaciones en cuanto a la validación externa de los resultados, además de requerir precaución a la hora de extrapolar los resultados a otros grupos de edad.

- **Heterogeneidad regional en la intensidad, fecha de inicio, cifras objetivo del programa, entre otros.**

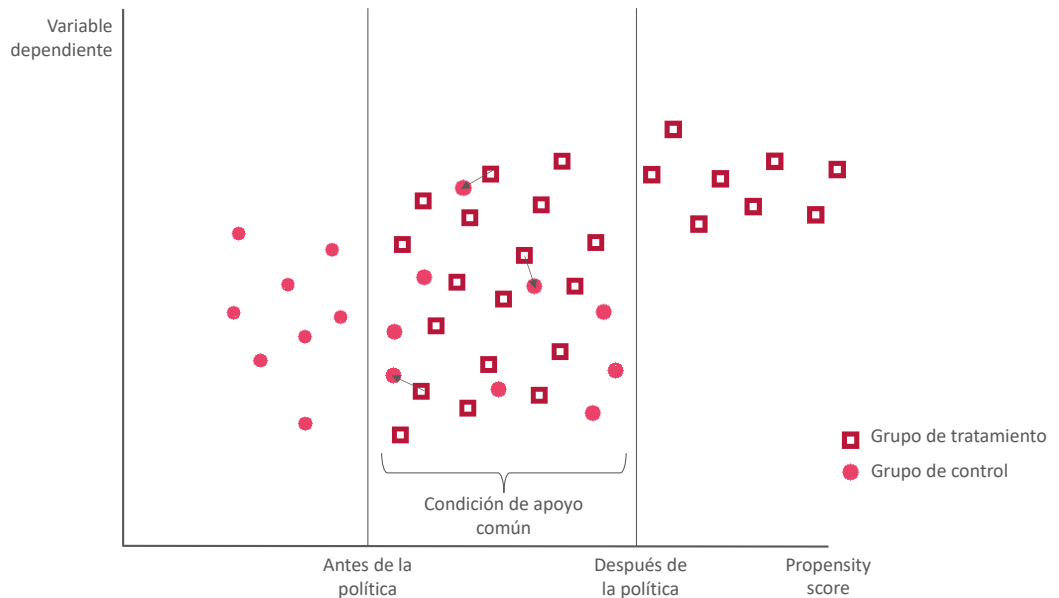
En *Hungría*, el Programa de Garantía Juvenil de la región de *Hungría* Central comenzó en octubre de 2015, nueve meses más tarde que en las demás regiones. La razón de este retraso es que la región de *Hungría* Central, a diferencia de las otras seis regiones húngaras, no es una región de convergencia según el objetivo de convergencia de los Fondos Estructurales de la UE, por lo que las fuentes de financiación del programa son distintas. Dicho retraso se debe a razones administrativas que pueden considerarse como exógenas. El programa iniciado en *Hungría* Central presenta elementos y requisitos de participación idénticos. La heterogeneidad exógena relacionada con la fecha de inicio permite aplicar un método DD como el que se detalla a continuación. El grupo de tratamiento está formado por las personas que cumplen los requisitos de participación y que buscan empleo en las regiones de convergencia mientras el grupo de control está compuesto por personas que buscan empleo situadas en la región de *Hungría* Central. Los grupos de tratamiento y de control se comparan en los nueve meses anteriores y posteriores al inicio del programa en las regiones de convergencia, es decir, en enero de 2015. Una limitación evidente de esta estrategia es que, al analizar la implementación por etapas, se comparan básicamente los grupos de tratamiento y de control durante un periodo corto, en este caso, de 9 meses.

Otros trucos que aparecen en la literatura son las "reglas de la casa" exógenas o el rigor de los investigadores. Hay ocasiones en las que las diferencias en los criterios de selección o en la probabilidad de ser tratado están basadas en diferencias en las "normas de la casa" informales, que difieren entre agencias de los Servicios Públicos de Empleo, municipios, personas, etc. (Maestas Mueller, 2013). Las diferencias en la probabilidad de ser tratado motivadas por las "normas de la casa" pueden ser analizadas desde un marco de variables instrumentales.

La otra vía principal en el proceso de identificación consiste en captar el sesgo de selección con un gran conjunto de variables observables como el historial de empleo y las características personales. El sesgo de selección puede reducirse utilizando un conjunto de variables que representen las características personales tanto de los participantes en el programa como de los no participantes que cumplieran los requisitos para hacerlo, constituyendo así un grupo de no participantes que se asemeje lo más posible a aquellos que sí participaron. En este caso, la estrategia de identificación se basa en el supuesto de que, en función de las variables observadas, el supuesto de independencia condicional está asegurado, es decir, ajustar por diferencias en las

variables observadas, elimina los sesgos que surgen en la comparación de unidades tratadas y unidades de control.

Figura 4 Ilustración del método de Propensity Score Matching



Sin embargo, la principal dificultad radica en que ciertas características no observables - en primer lugar, la motivación y la capacidad de los individuos- están correlacionadas con las variables que miden el efecto. Por ejemplo, una persona joven más motivada tiene más probabilidades de solicitar participar en una PAE, aunque probablemente sus perspectivas de empleo son mejores incluso sin participar. Un razonamiento similar podría aplicarse en otras aptitudes y capacidades que no son captadas por el nivel de educación. Los investigadores pueden tratar de encontrar variables que reflejen las capacidades y la motivación no observadas como, por ejemplo, la obtención de resultados más altos en las pruebas de competencia y en el historial de empleo o en los programas de formación podrían indicar una mejor motivación del individuo.

Sin embargo, como regla básica, esos factores no observables dan lugar a una sobreestimación del sesgo de selección del efecto del tratamiento, por lo que el efecto estimado debe considerarse como un límite superior.

En *Hungría*, los resultados del periodo de prueba laboral de 90 días se evalúan en relación con los resultados de los participantes en otros programas: programas de formación y de empleo público. Si el grupo de control está formado por participantes en otros programas en vez de por no participantes, el sesgo resultante de las características no observables es menor, a la vez que la validez externa de los resultados es también menor. Los resultados, por tanto, muestran la eficacia relativa del programa en comparación con otros programas. Cabe mencionar que no se llega a ninguna conclusión sobre el efecto del programa en comparación con la situación, cuando las

personas que buscan empleo no han participado en ningún programa, ya que las oficinas de los SPE procuran inscribir a los solicitantes de empleo inscritos en algún programa. Esto supone que el grupo de jóvenes que no ha participado en ningún programa pero que sí se ha mantenido registrado, probablemente no sea un grupo de control adecuado, ya que son o los que tienen una menor motivación o los que tienen mayores opciones de reinserción laboral. Por ello, se preveía que incluso los participantes en el programa de empleo público difirieran significativamente de otros participantes en cuanto a las perspectivas y la motivación implícitas en el mercado de trabajo. El nivel de educación y el historial de empleo completo disponible para cada persona en la base de datos permite atenuar la heterogeneidad no observada. La segunda evaluación polaca utiliza la misma estrategia para evaluar la eficacia relativa de la formación en el empleo, del vale de formación en el empleo, la formación en el aula, el bono de formación en el aula, el subsidio salarial (trabajo de intervención) y el empleo público.

En la evaluación de los contratos en prácticas realizada en *España* se comparan también los resultados entre un grupo bien definido de participantes en el programa y un conjunto de no participantes que cumplían los requisitos mediante modelos de regresión. El grupo de tratamiento está formado por jóvenes menores de 30 años que han firmado un contrato en prácticas, mientras que el grupo de control está formado por jóvenes desempleados, y que cumplen los requisitos de participación, que han firmado contratos temporales. La muestra de registros de la seguridad social permite construir variables relacionadas con el historial laboral de los individuos tales como la permanencia en el mercado de trabajo, el número de periodos laborales anteriores, el número de empresas en las que se ha estado empleado anteriormente y el tiempo en situación de desempleo. Estas variables pueden capturar una parte considerable de las diferencias entre el grupo de control y el de tratamiento. Además, los grupos de tratamiento y de control se definen de manera que permiten obtener un análisis preciso, eliminando del mismo a todos aquellos individuos que presenten algunos casos particulares. Por ejemplo, un enfoque es centrarse en aquellos que acaban de incorporarse al mercado laboral, ya que el objetivo del contrato en prácticas (CP) es servir como herramienta de acceso al empleo, restando inestabilidad a la empresa. Además, los individuos que acceden al mercado laboral a través de un CP pueden compararse con los que firmaron un CP habiendo trabajado antes. Dicha comparación permite identificar si el impacto del CP depende de si existe una trayectoria previa en el mercado laboral o no.

En *Italia*, la evaluación se refiere a las diferencias en la probabilidad de ser contratado con un contrato indefinido entre los trabajadores que cumplen los requisitos y los que no lo hacen, controlando por las características de referencia de los dos grupos. El principal supuesto en el que se basa el proceso de identificación de dicho efecto mediante una metodología DD, es que el incentivo no afecta al resultado del grupo de control, es decir, la política destinada a mejorar la probabilidad de contratar con un

contrato indefinido a un trabajador que cumple los requisitos no afecta a la probabilidad de empleo de los trabajadores que no los cumplen. Las estadísticas descriptivas muestran que este supuesto no se cumple puesto que estas reflejan una discontinuidad (reducción) en el año 2015 en la serie temporal de la variable dependiente para el grupo de control. La razón de dicha interrupción puede ser que algunos empleadores hayan cambiado sus preferencias de contratación de los trabajadores que no cumplen los requisitos (los que tienen un mayor vínculo al mercado laboral) a aquellos que sí lo hacen (aquellos que presentan una mayor fragmentación en sus carreras profesionales) debido al nuevo coste de oportunidad entre los dos grupos causado por la implantación de la política⁸.

En otras palabras, no puede descartarse la posibilidad de que el grupo de control se viese afectado por el de tratamiento. Ese posible sesgo se ha corregido con un procedimiento de dos etapas. En la primera, se estiman los valores teóricos de la variable dependiente - los valores que habría tenido si no se hubiesen producido los cambios regulatorios- y se atribuyen al grupo de control. En la segunda etapa, se estima el efecto de los subsidios a la contratación en la muestra corregida utilizando un modelo de mínimos cuadrados ordinarios (MCO) con interacciones entre el tratamiento y las diferentes categorías de edad, con el fin de obtener los efectos en la cohorte más joven.

3.4. Efectos heterogéneos

Un aspecto importante de la evaluación de políticas es la heterogeneidad del efecto de estas. ¿Tiene un programa un impacto diferente en los distintos grupos en función, por ejemplo, del sexo, edad, región, año en el que se implanta, condiciones macroeconómicas o dependiendo de si se combina con otro/s programa/s? Comprender dicha heterogeneidad en el efecto del tratamiento ayuda a comprender y evaluar

- La validez externa de los resultados y la comparación con los resultados de otras evaluaciones.
- La eficiencia de coste-beneficio del programa.
- Las debilidades del programa y las posibilidades de mejorarlo adaptando el mismo.
- La perspectiva de género del programa.

El efecto del programa puede variar según el nivel de educación de los participantes en el mismo. Los encargados de las políticas probablemente reciban con satisfacción la conclusión de que una subvención al coste salarial tiene un impacto más fuerte en aquellos que tienen un nivel de educación más bajo, ya que posiblemente este grupo

⁸ Como se explica en la siguiente sección, esa cuestión se complica aún más por el hecho de que el cambio en el coste de oportunidad es desigual entre los grupos de edad.

tiene resultados menos favorables en el mercado de trabajo. Por consiguiente, si la evaluación concluye que las personas con educación secundaria o inferior se benefician de dicha política de subvención al coste salarial, pero, por el contrario, no tiene repercusiones en los graduados universitarios, las subvenciones al coste salarial deberán destinarse al primer grupo.

Además, como los costes de una subvención al coste salarial a un empleo que requiere mayor nivel educativo y que es mejor remunerado son presumiblemente más altos, la orientación hacia los demandantes de empleo con menor nivel educativo podría mejorar la eficiencia de los costes del programa. Este resultado supone, de nuevo, que las subvenciones al coste salarial deben dirigirse al primer grupo (personas con educación menor o igual a secundaria). En *Hungría*, el periodo de prueba de 90 días demostró ejercer un impacto mayor en los participantes con educación elemental que en las personas con educación secundaria o superior en comparación con el empleo público. Este resultado no difiere con nuestras expectativas al respecto, ya que, desde el punto de vista de las perspectivas de empleo, los jóvenes con un nivel educativo bajo son el grupo más vulnerable. Al mismo tiempo, los beneficiarios del periodo de prueba tienen, en media, una mejor educación que los participantes en el programa de empleo público, lo que indica la autoselección en el programa los jóvenes desempleados más aptos para el empleo. Esto sugiere que priorizar a los jóvenes desempleados con un menor nivel educativo podría aumentar el efecto medio de tratamiento y, al mismo tiempo, disminuir la pérdida de eficiencia del programa. Por otra parte, los evaluadores podrían encontrar que la duración del programa influye en sus efectos. Dicho hallazgo ayudaría a determinar la duración de un programa similar en el próximo periodo.

El análisis español es abordado desde una perspectiva de género, concluyendo que las mujeres tienen menos probabilidades que los hombres de quedarse desempleadas después del contrato en prácticas. Sin embargo, no se encuentran diferencias entre hombres y mujeres estadísticamente significativas en las otras cuestiones analizadas.

La evaluación italiana encontró diferencias significativas en el efecto de los programas entre los distintos grupos de edad. El modelo de diferencias-en-diferencias estima un impacto positivo del cambio regulatorio en la proporción de contratos indefinidos sobre el total de contratos registrados en el año 2015, siendo dicho efecto positivo decreciente según categoría de edad.

La evaluación realizada en *Polonia* determinó que la formación en el aula y los vales de formación en el aula funcionan relativamente bien para los hombres, pero no para las mujeres. El empleo público es menos eficaz para las personas con un nivel educativo más bajo y para aquellas que viven en regiones con un alto nivel de desempleo. Los vales formativos en el empleo son más eficaces que la formación habitual en el empleo, independientemente del género, nivel educativo o lugar de residencia de las personas desempleadas.

4. Cuestiones relevantes para los encargados de las políticas

Hasta ahora, hemos analizado los elementos necesarios para garantizar que los investigadores puedan estimar el efecto causal de las políticas (validez interna). Sin embargo, existe una variedad de cuestiones que los investigadores querrán considerar cuando examinen los resultados de las evaluaciones con los encargados de las políticas. Nuestro objetivo en esta guía metodológica es destacar dichas cuestiones, mencionando la importancia de los datos para solucionar dichas cuestiones, sin ofrecer una orientación detallada al respecto.

4.1. Interpretación de los resultados

Es muy importante que los investigadores presenten detalladamente el efecto que son capaces de estimar en su estudio de evaluación. Muy a menudo, el diseño de la evaluación conduce a una estimación internamente válida del efecto medio del tratamiento (ATE) en el grupo tratado o el efecto medio local del tratamiento - los que actualmente adoptan las medidas, o un (pequeño) subgrupo de ellos -, sin embargo, al presentar estos resultados a los políticos, es fundamental hacer hincapié en que los resultados pueden generalizarse a poblaciones más amplias con supuestos adicionales. Así pues, vale la pena considerar también la posibilidad de recopilar datos sobre las características de poblaciones más amplias, para contribuir a esta generalización. Por el contrario, las evaluaciones que tienen por objeto estimar la intención de tratar - por lo tanto, la oferta de la medida - o el efecto medio del tratamiento (ATE) - efecto del tratamiento sobre todas las personas que cumplen los requisitos de participación, no sólo sobre aquellas que se inscriben realmente en el programa - podrían producir resultados directamente más útiles para los encargados de las políticas, ya que incorporan la posible no aceptación (para el ITT) o incluyen la posibilidad de estimar el efecto causal sobre los que actualmente no participan en la medida (para el ATE).

4.2. Vinculación de resultados al diseño de los programas y su implementación

Más allá de la estimación cuantitativa de los efectos de un programa, la evaluación debe conducir a conclusiones sobre los elementos que se encuentran detrás de los resultados. Por ejemplo, si la evaluación llega a la conclusión de que el programa no muestra ningún efecto significativo en el empleo, es importante saber si dicho fracaso se debe a un mal diseño del programa (elección errónea de la población objetivo, criterios de selección o

participación, recursos financieros y de personal), o a problemas en la ejecución de este (falta de información, selección arbitraria de los participantes, etc.).

4.3. Validez externa, comparación con otros estudios y resultados estadísticos

Aquellas personas interesadas en los estudios de evaluación y, a su vez, en la formulación de políticas, desearán comparar las medidas que constituyen los objetivos del estudio con otros resultados. Es probable que se trate de comparaciones tales como la eficacia de medidas similares para los jóvenes de otras regiones/países; de otras medidas (alternativas) para los jóvenes; y de la eficacia (potencial) de las medidas evaluadas para grupos destinatarios distintos. Para todos estos objetivos, los investigadores deben analizar una serie de cuestiones. En primer lugar, cómo presentar los resultados de un modo que permitan realizar comparaciones con otros estudios. En segundo lugar, necesitan expresar la forma en que sus resultados pueden generalizarse a otros escenarios, con el fin de evitar las extrapolaciones simplistas y destacar sus limitaciones. En tercer lugar, los investigadores tienen también que documentar con claridad los elementos de diseño más relevantes de las medidas evaluadas, así como las condiciones económicas existentes, de modo que otros interesados puedan comprender cómo éstas pueden haber afectado al éxito de las políticas. Por ejemplo, las políticas evaluadas en *Hungría*, *Italia* y *Polonia* se pusieron en marcha en un período de recuperación económica, mientras que la medida analizada en *España* se pudo observar tanto en un período de crecimiento económico como durante la Gran Recesión.

Al interpretar y presentar los resultados en términos de empleo e ingresos, los políticos deben tener presente la vinculación del indicador en cuestión con otras estadísticas conocidas y sus posibles diferencias. Por ejemplo, el análisis podría centrarse en los desempleados registrados cuyas cifras podrían diferir considerablemente de las cifras basadas en el concepto de desempleo de la encuesta de población activa. La interpretación debe abordar estas diferencias mediante una explicación cuidadosa.

4.4. Efectos de desplazamiento y efectos de “Peso Muerto”

A fin de que los responsables de la formulación de políticas tengan un panorama más completo de la eficacia de las medidas, es necesario enfatizar dos cuestiones. En primer lugar, se produce un efecto de desplazamiento cuando la medida contribuye al empleo de los jóvenes al tiempo que perjudica las perspectivas de las personas que son muy similares a ellos en cuanto a sus habilidades, pero que no, por el contrario, tienen acceso a la medida. En segundo lugar, un programa genera efectos de “peso muerto” si una medida aumenta el empleo de los jóvenes, pero una parte considerable del apoyo

prestado podría no haber sido necesario ya que los empleadores habrían contratado a esas personas concretas incluso en ausencia de las medidas. Considerar la importancia relativa de los objetivos alternativos conduce a analizar cuestiones relacionadas con las preferencias sociales (políticas). Sin embargo, aportar evidencias sobre estas cuestiones no es tarea sencilla y, como mínimo, requiere mayor cantidad de datos tales como información sobre los trabajadores de la misma empresa y de empresas competidoras, preferentemente donde los datos de los empleados se correspondan con las empresas que los contrataron.

4.5. Análisis coste-beneficio y coste-efectividad

La evaluación del impacto que hemos analizado hasta ahora expresa el efecto de una política en términos de resultados en el mercado laboral. Sin embargo, también es necesario tener en cuenta la evaluación de los costes totales del programa. Esto cobra especial relevancia ya que los encargados de las políticas querrán comparar unas políticas con otras en términos de rendimiento relativo (monetario).

Hay dos tipos de análisis que vinculan los costes y los beneficios. El análisis coste-efectividad compara los costes del programa con los resultados expresados en unidades materiales, por ejemplo, el coste-efectividad de un programa podría ser un contrato a tiempo completo por x euros. En el análisis coste-beneficio, los efectos del programa se expresan también en unidades monetarias, lo que permite evaluar si el beneficio total supera los costes totales. Resulta evidente que este último análisis es más complejo, ya que para contabilizar y expresar los beneficios se requiere más información.

Incluso el análisis coste-efectividad puede resultar complejo si la base de datos administrativa no contiene los costes del programa, en cuyo caso suele ser lo más probable. Si se diese dicha situación, los investigadores podrían utilizar datos agregados de otras fuentes para calcular los costes por unidad de programa. Por otra parte, en el caso de los programas de subvenciones al coste salarial o de formación, los datos sobre salarios procedentes de una base de datos vinculada con la de la seguridad social, junto con la normativa del programa, podrían ofrecer una estimación de los costes de este. En cualquier caso, los investigadores deberían planificar con antelación la obtención de este tipo de datos al comienzo de la evaluación.

Bibliografía

Abadie, Alberto and Cattaneo, Matias D., (2018): Econometric Methods for Program Evaluation (August 2018). Annual Review of Economics, Vol. 10, pp. 465-503, 2018, <http://dx.doi.org/10.1146/annurev-economics-080217-053402>

Angrist J.D., Pischke J.S. (2015), Mastering Metrics: The Path from Cause to Effect, Princeton University Press, Princeton.

Angrist, Joshua and Pischke, Jörn-Steffen. (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Caliendo M., Mahlstedt R., Mitnik O. (2017). Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labor market policies. Labour Economics, 46, 14-25.

Card D., Krueger A.B. (1994), Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania, American Economic Review, 84, n. 4, pp. 772–793. <http://davidcard.berkeley.edu/papers/njmin-aer.pdf>

European Commission (2013) : Evalsed Sourcebook, Methods and Techniques

https://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/evaluation_sourcebook.pdf

European Commission (2020): How to use administrative data for European Social Funds counterfactual impact evaluations, Directorate-General for Employment Social Affairs and Inclusion, Directorate G <https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8303>

Imbens, Guido, and Thomas Lemieux (2008): Regression Discontinuity Designs: A Guide to Practice, Journal of Econometrics, 142, 615-635.

Gertler, Paul J.; Martinez, Sebastian; Premand, Patrick; Rawlings, Laura B.; Vermeersch, Christel M. J. (2016). Impact evaluation in practice (English). Washington, D.C.: World Bank Group. <http://documents.worldbank.org/curated/en/698441474029568469/Impact-evaluation-in-practice>

Morris, Stephen & Tödtling-Schönhofer, Herta & Wiseman, Michael. (2013). Design and Commissioning of Counterfactual Impact Evaluations - A Practical Guidance for ESF Managing Authorities. 10.2767/94454. Publisher: European Commission, Directorate-General for Employment, Social Affairs and Inclusion

OECD (2009). Guidelines for Project and Programme Evaluations

UNDP (2009). Handbook on Planning Monitoring and Evaluating for Development Result

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>

Anexo 1: Breve descripción de las evaluaciones en los cuatro países de interés.

ITALIA

Programa:

- 1) Reducción de los costes de despido para los contratos indefinidos (contrato con indemnización al despido creciente).
- 2) Reembolso de la seguridad social a los nuevos contratos indefinidos y a las conversiones de un contrato temporal en contrato indefinido (evaluados conjuntamente).

Datos: contratos registrados en la Base de Datos del Ministerio de Trabajo y Políticas Sociales (SISCO) referidos a la comunicación enviada por los empleadores por ley (COB), entre el 1 de enero de 2005 y el 31 de diciembre de 2015, en comparación con los registrados en los años anteriores.

Resultado: Proporción de nuevas contrataciones con contrato indefinido sobre el total de contratos de trabajo registrados en 2015.

Estrategia de identificación: Modelo DD con corrección por sesgo de selección (violación del SVTUE)

Grupo que cumple los requisitos de participación:

- Personas contratadas con un contrato indefinido (excluida la formación) durante el periodo 2014-2015
- Todos los empleadores sin especificar industria o ubicación geográfica. Incluye las asociaciones y empresas públicas, pero excluye la administración pública y la agricultura.
- Empleadores sin atrasos de contribuciones pendientes o cualquier otra situación irregular en términos de convenios colectivos o acuerdos territoriales.

Grupo que no cumple los requisitos de participación:

- Personas contratadas durante el año 2015 que hayan sido despedidas dentro de los 6 meses anteriores al inicio del nuevo contrato indefinido.

- Personas contratadas durante 2015 que en los 3 meses anteriores a la reforma (desde el 1 de octubre de 2014 hasta el 31 de diciembre de 2014) no fueron despedidas por el mismo empleador que las había contratado o por otro empleador relacionado con este último, evitando así incentivos perversos.

Definición de la variable que mide el efecto: el resultado es una *dummy*, *Y*, que toma dos posibles valores, $Y=1$ si el contrato es indefinido; $Y=0$ si el contrato es temporal.

Definición de la variable de tratamiento: $T=1$, para un contrato indefinido seleccionado para tratamiento; $T=0$, contrato indefinido no seleccionado para tratamiento.

Se crean interacciones de la variable tratamiento con las categorías de edad, con el fin de captar los efectos en los grupos jóvenes.

Definición del periodo: $P=0$, 2014; $P=1$, 2015.

HUNGRÍA

Programa: La evaluación húngara centra el análisis en el programa que contempla un periodo de prueba de 90 días diseñado para jóvenes menores de 25 años. El periodo de prueba laboral proporciona una subvención al coste salarial del 100% durante 90 días sin obligatoriedad de continuar con el contrato. Este programa se inició en 2015 como parte del Sistema de Garantía Juvenil.

Datos: La evaluación se basa en dos conjuntos de datos administrativos. Por una parte, los datos de los SPE contienen todos los solicitantes de empleo registrados, además de los participantes en el programa. Por otra parte, dichos datos están vinculados a un conjunto de datos administrativos del organismo encargado de las pensiones que contiene la situación laboral y el historial de empleo de, aproximadamente, la mitad de la población total.

Resultados: La principal variable dependiente es la situación laboral a los 6 y 12 meses de la finalización del programa y los salarios acumulados en comparación con el salario mínimo.

Estrategia de identificación: La evaluación aplica dos estrategias de identificación diferentes. En la primera se comparan los resultados de los participantes en el programa que contempla un periodo de prueba de 90 días con aquellos de los participantes en otros programas (empleo público y formación), aplicando *propensity score matching* (PSM). En la segunda, la estrategia de identificación se basa en una especificación Diferencia-en-Diferencia que aprovecha el hecho de que en la región de Hungría Central el programa se inició nueve meses más tarde que en otras partes del país por razones meramente administrativas. El grupo de tratamiento contiene jóvenes que cumplen los requisitos de participación de

tres regiones de convergencia y el grupo de control está formado por jóvenes que cumplen los requisitos de participación de la región centrohúngara. Dicha estimación compara los resultados de los grupos de tratamiento y de control, es decir, los que se inscribieron en los nueve meses anteriores y posteriores al 1 de enero de 2015, día en que se implanta el Sistema de Garantía Juvenil en la región de *Hungría* Central. La especificación permite una identificación ITT que muestra el efecto de todo el SGJ en la subpoblación que cumple los requisitos de participación.

POLONIA (EVALUACIÓN 1)

Programa: Se evalúa un programa de subvenciones al coste salarial para los desempleados menores de 30 años en *Polonia*. Estuvo operativo en los años 2016 hasta 2018. El subsidio para la empresa o empleador consistía en una cuantía equivalente al salario mínimo más la seguridad social que se abonaba durante 12 meses. Además, los empleadores estaban en la obligación de prolongar el empleo durante otros 12 meses después de la fecha de finalización del subsidio.

Requisitos de participación: Pueden optar al subsidio los desempleados registrados menores de 30 años con el perfil I o II (el perfil III no cumple los requisitos de participación). Asimismo, los empleadores que redujeron el empleo por iniciativa propia en los últimos 6 meses no fueron considerados.

Datos: Se utilizan datos del registro de los servicios públicos de empleo (SPE) con un historial completo de inscripciones de desempleados y participación en el programa. Dichos datos no están vinculados a los datos de la seguridad social.

Muestra: Personas que se incorporaron al desempleo entre los años 2015 y 2016, excluyendo a aquellas que ya se encontraban desempleadas o que habían participado durante los últimos 6 meses en un PAE. Todo ello, con el fin de garantizar que analizamos un nuevo registro y no una parte de un período de desempleo más largo. Por otro lado, para poder observar el resultado 36 meses después de la inscripción, limitamos nuestra muestra a las inscripciones fechadas hasta abril de 2016.

Resultados: Analizamos dos tipos de resultados. El primero, estar fuera del registro de 12 a 36 meses después de la inscripción y, el segundo, el número acumulado de días fuera del registro de 12 a 36 meses después de la inscripción.

Estrategia de identificación: DRD usando el criterio de edad combinado con DD (2016 vs. 2015: otras reglas institucionales cambian en el límite de los 30 años y estaban vigentes en el año 2015).

POLONIA (EVALUACIÓN 2)

Programa: Un conjunto de PAE dirigidas los jóvenes, tales como formación en el trabajo, formación en el aula, subsidio salarial (trabajos de intervención), empleo público, vales de formación en el trabajo y de formación en el aula.

Grupo que cumple los requisitos de participación: Desempleados registrados menores de 30 años.

Datos: Se utilizan datos del registro de los servicios públicos de empleo (SPE) con un historial completo de registros de desempleo y participación en el programa, no vinculados a los datos de la seguridad social.

Muestra: Restringimos nuestra muestra a las PAE finalizadas que comenzaron entre el año 2015 y 2016, y que duraron al menos un día. Incluimos únicamente los participantes que tenían entre 18 y 29 años cuando se iniciaron las PAE y excluimos a aquellas personas que hayan estado alguna vez en la cárcel. Analizamos 6 tipos de intervenciones. Desde la formación en el trabajo, formación en el aula, subvención salarial (obras de intervención), a obras públicas, vales de formación en el trabajo y de formación en el aula. La muestra total cuenta con 319 610 observaciones.

Resultados: estar fuera del registro y no estar en las PAE 12 a 36 meses después del registro.

Estrategia de identificación: Comparación por parejas de participantes en el programa que coincidan en el PSM.

ESPAÑA

Programa: Contrato en prácticas. Se implantó en *España* en 1998 y, de acuerdo con la legislación española, tiene como objetivo principal aumentar la estabilidad laboral desarrollando prácticas profesionales en relación con el nivel y el campo de estudios e incentivando a la empresa a contratar jóvenes trabajadores una vez finalizado el periodo de prácticas.

Requisitos de participación: Las restricciones de los criterios de selección se refieren a la educación y la edad. Los individuos deben haber terminado la licenciatura o el ciclo formativo de formación profesional y ser menores de 30 años o, haber terminado el grado hace un máximo de 5 años.

Datos: Muestra Continua de Vidas Laborales (MCVL). Se trata de una base de datos en la que se incluyen los registros de la seguridad social del 5% de la población española. Esta base de datos incluye el historial completo de empleo de muestra representativa de los trabajadores españoles.

Muestra: Individuos menores de 30 años que accedieron al empleo a partir de 2002, con un alto nivel de educación, es decir, potenciales usuarios del CP.

Resultados: Se analizan dos tipos de resultados. La probabilidad de permanecer en la empresa después del CP y la probabilidad de estar empleado mediante un contrato indefinido después del CP, en la misma empresa (para los que permanecen) o en otra empresa (para los que cambian).

Estrategia de identificación: Para el primer análisis se estima mediante un multinomial *logit* en la que la variable dependiente toma tres valores diferentes, en concreto, cambiar a otra empresa, permanecer en la misma y quedarse desempleado. En el segundo paso, se estima aplicando un modelo *logit* la probabilidad por separado de firmar un contrato indefinido (comparado con uno temporal) dependiendo de si permanecen en la empresa o cambian a otra empresa.

Anexo 2: Glosario y abreviaciones

ATE, Average Treatment Effect, ATT, Average Treatment Effect on the Treated, Efecto Medio del Tratamiento y el Efecto Medio del Tratamiento en los Tratados: un efecto de tratamiento es el efecto causal de un tratamiento (una variable binaria, 0-1) sobre una variable que mide el efecto. Representa la diferencia entre el resultado potencial de un individuo seleccionado aleatoriamente de la población con y sin el tratamiento (exposición a la política, participación en un programa específico, etc.). Existen dos conceptos principales del efecto medio del tratamiento. El ATE muestra la ganancia esperada por recibir tratamiento en el par de resultados potenciales promediados sobre toda la población de interés. Esta medida es relevante si toda la población está expuesta a la política en cuestión.

$$ATE = E(Y_i(1) - Y_i(0))$$

Donde $Y_i(1)$ es el resultado de la unidad i cuando esta recibe tratamiento y $Y_i(0)$ es el resultado cuando la unidad i no recibe tratamiento.

El ATT, Efecto Medio del Tratamiento en los Tratados muestra el efecto medio de tratamiento en la subpoblación tratada:

$$ATT = E(Y_i(1) - Y_i(0) | D_i = 1)$$

CO, COB, Comunicazioni Obbligatorie on-line, Registro de comunicaciones obligatorias gestionado por el Ministerio de Trabajo y Políticas Sociales (Ley Nº 296, de 27 de diciembre de 2006 (Ley Financiera de 2007).

Diferencias-en-Diferencias, DD, Método de estimación cuasiexperimental en el que el efecto del programa se calcula comparando el grupo de tratamiento y un grupo de control antes y después del tratamiento. El efecto del tratamiento es el cambio medio a lo largo del tiempo en la variable resultado del grupo de tratamiento menos el cambio medio del grupo de control. El método DD se basa en el supuesto de que, en ausencia del tratamiento, el resultado del grupo de tratamiento habría seguido la tendencia de las unidades no tratadas, es decir, que los dos grupos habrían tenido tendencias paralelas (para más detalles, véase, por ejemplo, Card y Krueger, 1994, Angrist y Pischke, 2008).

DRD, Diseño de Regresión Discontinua. Método cuasiexperimental que permite estimar el efecto del tratamiento del programa, cuando los individuos cumplen los requisitos de participación o la probabilidad de participar en el programa depende de una determinada característica observable como, por ejemplo, la edad (variable forzada, o corriente). El efecto del programa se mide comparando los resultados por debajo y por encima del umbral en las cercanías de este. El DRD calcula el efecto medio local del tratamiento ajustando un polinomio (local) a ambos lados del umbral y se basa en el supuesto de que la selección entre

el grupo de tratamiento y el de control es aleatoria. La regresión discontinua puede ser *fuzzy* y *sharp* (para más detalles, véase, por ejemplo, Imbens y Lemieux, 2008; Angrist y Pischke, 2008).

EIC, Evaluación de Impacto Contrafactual

INAPP, Istituto Nazionale per l'Analisi delle Politiche Pubbliche.

ISCED, Clasificación Internacional Normalizada de la Educación.
https://ec.europa.eu/education/international-standard-classification-of-education-isced_it

ITT, Intención de tratamiento: muestra el efecto de la política en la población que cumple con los requisitos de participación. En otras palabras, nos muestra el efecto causal de la propuesta de tratamiento. Si no todos los miembros de la población que cumple con los requisitos de participación reciben el tratamiento si no que muchos de ellos rechazan la propuesta de recibirlo, la ITT diferirá del efecto medio del tratamiento.

MCO, Mínimos Cuadrados Ordinarios.

MCVL, Muestra Continua de Vidas Laborales. Es una muestra del 5% de los registros de la Seguridad Social Española.

NACE, Clasificación Estadística de Actividades Económicas de la Comunidad Europea, Rev. 2 (2008), Rev. 2 (NACE Rev. 2)
https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2&StrLanguageCode=IT&IntPcKey=&StrLayoutCode=HIERARCHIC

PAE, Políticas Activas de Empleo

PSM, *Propensity Score Matching*, método de evaluación contrafactual que compara la variable dependiente de los individuos tratados con la variable de resultado de los individuos del grupo de control que han sido emparejados por ser similares en características a las unidades de tratamiento. En este método, las unidades de tratamiento se emparejan con unidades de control con una puntuación similar o *propensity score*, es decir, la probabilidad de ser tratado dada una serie de variables observables. Este método se basa en el supuesto de que, condicionada a una serie de características observables, la selección en el grupo de tratamiento es aleatoria (para más detalles, véanse, por ejemplo, Rosenbaum y Rubin, 1983; Angrist y Pischke, 2008).

SISCO, Sistema Statistico delle Comunicazioni Obbligatorie on-line, mandatory communication collected by the Ministry of Labour and Social Policies (MLSP).

SPE, Servicios Públicos de Empleo.

SUTVA, Stable Unit Treatment Value Assumption, El Supuesto de Valores Estables de la Unidad de Tratamiento es aplicado en la inferencia causal: usually made in causal inference: a treatment applied on one unit does not affect the outcome for another unit (no interference between the units).

Supuesto de Independencia Incondicional, se refiere al caso cuando controlar por las diferencias en el conjunto fijo de variables pretratamiento elimina los sesgos en las comparaciones entre las unidades tratadas y las unidades de control, de modo que, condicionado a dichas variables de control, la asignación de tratamiento puede considerarse aleatoria.

Índice de recuadros

Recuadro 1 Detalles de las evaluaciones de Italia y Polonia	6
Recuadro 2 Detalles de las evaluaciones de España y Hungría	7

Índice de figuras

Figura 1 Esquema de la lógica de la evaluación del impacto contrafactual	16
Figura 2 Método de diferencias-en-diferencias (DD) ilustrado	17
Figura 3 Ilustración del Diseño de Regresión Discotinuа (DRD)	18
Figura 4 Ilustración del método de Propensity Score Matching	20

FEDEA

Calle de Jorge Juan, 46

28001

Madrid

+34 91 435 90 20

www.fedea.net

