CUADERNOS ECONÓMICOS ICE
INFORMACIÓN COMERCIAL ESPAÑOLA

Evaluación de políticas públicas con métodos cuasi-experimentales: una aplicación a las políticas activas de empleo en la Comunidad Valenciana

Evaluation of Public Policies Applying Quasi-Experimental Methods: An Application to Active Labour Market Policies in the Valencian Community

José Vila
Universidad de Valencia, ERI-CES, IDAL y DevStat

Yolanda Gómez
José Luis Cervera-Ferri
DevStat

Juan Martínez
Servicio Valenciano de Empleo y Formación

Resumen

Este artículo presenta un estudio comparativo de dos metodologías de medición de impacto de políticas públicas: los métodos baseline y cuasi-experimental. Ambos métodos son aplicados a la evaluación de las políticas activas de empleo implementadas en la Comunidad Valenciana en el periodo 2012-2015. La aplicación de ambos métodos a un mismo conjunto de políticas permite cuantificar un fenómeno (habitualmente citado en la literatura, pero escasamente validado empíricamente) como es el grado de confusión entre los efectos de una política y los de la evolución de su entorno, asociados a la utilización del método baseline. Los resultados de este análisis muestran que, en algunos casos, más de dos terceras partes del impacto medido a través de indicadores baseline es en realidad atribuible a variaciones del entorno y no a la política evaluada. Como consecuencia, los resultados de este artículo muestran los peligros de afrontar la evaluación de políticas desde un enfoque baseline y la conveniencia de recurrir a métodos alternativos con base experimental.

Palabras clave: *evaluación de políticas, métodos experimentales, cuasi-experimentos, políticas activas de empleo.*

Clasificación JEL: *C21, C93, J48.*

DOI: https://doi.org/10.32796/cice.2020.99.7006
Recibido: Febrero 2020 · Aceptado: Abril 2020

Abstract

This article presents a comparative study of two methodologies for measuring the impact of public policies: the baseline and quasi-experimental methods. Both methods are applied to the evaluation of Active Labour Market Policies implemented in the Valencian Community during the period 2012-2015. The application of both methods to the same set of policies allows us to quantify a phenomenon (that is usually cited in the literature, but scarcely empirically validated) such as the degree of confusion between the effects of a policy and those of the evolution of its environment associated with the use of the baseline method. The results of this analysis show that, in some cases, more than two thirds of the impact measured through baseline indicators is actually attributable to variations in the environment and not to the evaluated policy. As a consequence, the results of this article show the dangers of approaching policy evaluation from a baseline perspective and the convenience of resorting to alternative methods on an experimental basis.

Keywords: *policy evaluation, experimental methods, quasi-experiments, active labour market policies.*

JEL Classification: *C21, C93, J48.*

1. Introducción

Este artículo presenta un estudio comparativo de dos metodologías de medición de impacto de políticas públicas: *baseline* y experimental. El primer método compara la situación de los beneficiarios de la política antes y después de su aplicación. El segundo método es una variación de la metodología de experimentos aleatorios controlados (EAC), donde la aleatorización se realiza a posteriori de la implementación de la política a evaluar. En este trabajo, ambos métodos son descritos en detalle y aplicados al caso de estudio de la medición del impacto de las políticas activas de empleo (PAE) aplicadas en la Comunidad Valenciana en el periodo 2012-2015. Las PAE vienen definidas por el Real Decreto-ley 3/2011, de 18 de febrero, de medidas urgentes para la mejora de la empleabilidad y la reforma de las políticas activas de empleo (BOE núm. 43, de 19 de febrero de 2011) como:

> El conjunto de acciones y medidas de orientación, empleo y formación dirigidas a mejorar las posibilidades de acceso al empleo, por cuenta ajena o propia, de las personas desempleadas, al mantenimiento del empleo y a la promoción profesional de las personas ocupadas y al fomento del espíritu empresarial y de la economía social (art. 10).

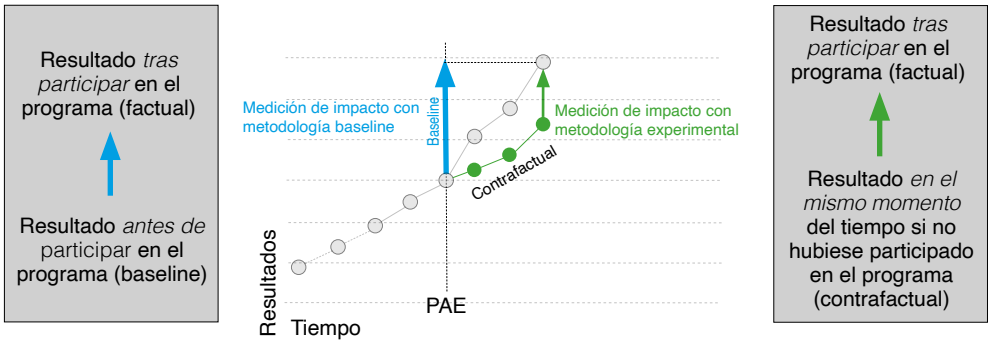
Las PAE juegan así un papel crucial por su contribución tanto a la creación de empleo como a la inserción de los desempleados. Aunque existen abundantes ejemplos de medición de impacto con métodos *baseline* o cuasi-experimentales (White y Sabarwal, 2014; Mueller y Gaus, 2017), la novedad de este artículo es la aplicación de ambos métodos a un mismo conjunto de políticas, lo que permite contrastar empíricamente su efectividad a la hora de evaluar políticas públicas.

Aunque no exentos de crítica, los EAC han sido definidos como el «patrón oro» de la medición de impacto (Bédécarrats *et al*., 2018; Grossman y Mackenzie, 2005), ya que son capaces de cuantificar el impacto realmente atribuible a una política.

Sin embargo, el método *baseline* puede confundir el impacto de la política y el de la evolución del entorno en el que esta se aplica (Athey y Imbes, 2017). A pesar de su creciente utilización en la fase de diseño de políticas públicas, la aplicación de EAC presenta dificultades logísticas y éticas para su evaluación. Esto se debe a que los EAC requieren una aleatorización previa entre los potenciales beneficiarios para determinar sobre quiénes de ellos se van a aplicar. Aunque esta aleatorización puede hacerse fácilmente a pequeña escala durante la fase de diseño, resulta complicada de realizar a mayor escala para evaluar la política una vez implementada. En este marco, surge la necesidad de plantear variaciones de los EAC que, manteniendo al máximo la capacidad de medir el impacto atribuible a una política, sean más sencillos de implementar en la práctica (White y Sabarwal, 2014). Ese es el caso del método cuasi-experimental analizado en este trabajo.

Para poder atribuir un impacto a una PAE, es necesario establecer una relación causa-efecto entre las acciones asociadas a esta política y sus indicadores clave de desempeño (ICD) (Lowndes *et al*., 2017). Sin embargo, la comparación de la situación de los beneficiarios antes de su participación en acciones de una PAE (medición *baseline*) con su situación después de dicha participación (medición factual) no es capaz de aislar el impacto de la PAE del de otras variables y establecer causalidad. Por ejemplo, en un entorno recesivo, un fuerte crecimiento de la proporción de desempleados que siguen sin encontrar trabajo tras su participación en una PAE no indica que esta tenga un efecto negativo: el crecimiento del desempleo podría haber sido aún mayor si no se hubiese aplicado la política para paliar la destrucción de empleo generada por la recesión. Como se muestra en la Figura 1, para establecer atribución es necesario comparar el nivel de empleo de los beneficiarios de la PAE después de participar en una acción (medición factual) con el que hubiesen tenido si no hubiesen participado en la acción asociada a la PAE (medición contrafactual). Sin embargo, es imposible efectuar ambas mediciones ya que un desempleado participa o no participa en la acción, pero no puede hacer ambas cosas a la vez.

FIGURA 1
ENFOQUES BASELINE Y EXPERIMENTAL

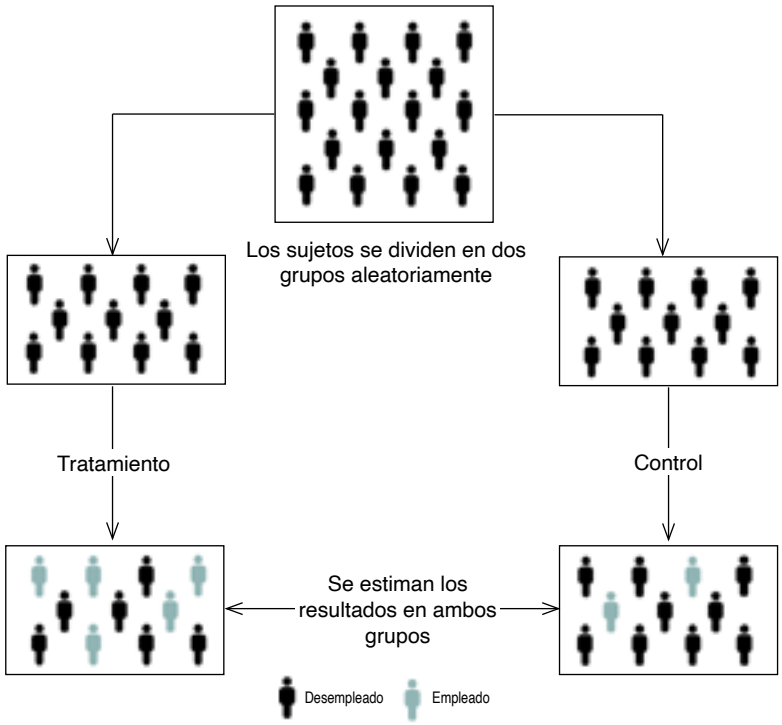


FUENTE: Elaboración propia.

Los EAC proporcionan una metodología para estimar la diferencia entre lo que ha pasado y lo que hubiese pasado sin aplicar una política. El punto clave de un EAC (McDermott, 2002) es la asignación aleatoria de los potenciales beneficiarios del programa a uno de dos grupos, control y tratamiento, con antelación a la implementación de las acciones de la PAE. Así, los miembros del grupo control no participan en acciones de la PAE, mientras que los del grupo tratamiento sí lo hacen.

El supuesto fundamental en la aplicación de EAC es que esta aleatorización *a priori* garantiza la comparabilidad entre los sujetos de los diferentes grupos, ya que estos tendrán características similares y habrá sufrido de igual manera los efectos de la evolución del entorno. En un EAC, el impacto atribuible a la PAE se puede estimar como la diferencia entre la medición de empleabilidad factual (proporcionada por el grupo tratamiento) y contrafactual (proporcionada por el grupo control) (Figura 2).

FIGURA 2
ESQUEMA CONCEPTUAL DE LOS MÉTODOS EXPERIMENTALES



FUENTE: Elaboración propia.

La implementación de EAC presenta dificultades, pues supone negar el acceso al programa a ciertos beneficiarios de forma aleatoria (grupo control). Como alternativa, el método cuasi-experimental plantea otra forma de definir los grupos control y tratamiento basada en aleatorizaciones que pueden realizarse después de la implementación de la PAE. En este caso, el grupo tratamiento se establece a partir una muestra aleatoria de beneficiarios que hayan participado en el programa y, para cada uno de ellos, se selecciona aleatoriamente un «espejo» con un perfil semejante, pero que no haya participado en la PAE.

Las PAE son políticas de mercado de trabajo dirigidas a mejorar los niveles de competitividad del sistema productivo (Vooren *et al*., 2019) y se caracterizan por ser anticipadoras y preventivas a la situación de desempleo. Así, las PAE están dirigidas a contrarrestar los efectos negativos que sobre el empleo provocan los cambios en las condiciones económicas y del entorno. Por esta razón, la discriminación entre el efecto de una PAE y el de la evolución del entorno económico es crítica para analizar su efectividad (Caliendo y Schmidl, 2016), convirtiendo a las PAE en candidatas ideales para la comparación de los métodos de medición de impacto *baseline* y cuasi-experimental efectuada en este trabajo.

Políticas activas de empleo

Las PAE están constituidas por tres grupos principales de políticas: formación, fomento de empleo e inserción laboral, cada una de las cuales se compone, a su vez, de una serie de acciones concretas. La definición de cada tipo de PAE se muestra en el Cuadro 1.

CUADRO 1
DEFINICIÓN DE LOS DIFERENTES TIPOS DE PAE

Políticas activas de empleo (PAE)
Formación
Políticas de formación, destinadas a fomentar la adquisición, mejora o puesta al día de competencias y cualificaciones profesionales.
1. Acciones de formación para desempleados (o formación profesional para el Empleo, prioritariamente para trabajadores desempleados): están dirigidas prioritariamente a personas en situación de desempleo, a fin de proporcionarles una cualificación profesional para desempeñar una ocupación adecuada a las necesidades del mercado de trabajo
2. Acciones de formación continua: son acciones de formación sectoriales y/o intersectoriales dirigidos tanto a la adquisición de competencias transversales como de competencias específicas dirigidas a la formación de personas trabajadoras.
3. Acciones de formación dual: es el conjunto de acciones e iniciativas formativas, mixtas de empleo y formación, que tienen por objeto la cualificación profesional de los trabajadores en un régimen de alternancia de actividad laboral en una empresa con la actividad formativa recibida en el marco del sistema de formación profesional para el empleo o del sistema educativo.

CUADRO 1 *(Continuación)*
DEFINICIÓN DE LOS DIFERENTES TIPOS DE PAE

Políticas activas de empleo (PAE)

Fomento de empleo

Políticas de promoción y creación de empleo.

4. Acciones de empleo público y desarrollo local: a través de este tipo de acciones se promueve la generación de empleo para la realización de obras o servicios de interés general y social, apoyando a las corporaciones locales y entidades que desarrollan este tipo de proyectos en el ámbito de la Comunitat Valenciana.
5. Acciones de empleo estable: conjunto de acciones de formación destinadas a aumentar el empleo y su estabilidad mediante medidas de fomento de empleo dirigidas tanto a la contratación inicial, como a la transformación de contratos temporales en indefinidos, favoreciendo la inserción laboral de trabajadores pertenecientes a los colectivos de desempleados con mayores dificultades de acceso al mercado de trabajo.
6. Acciones para discapacitados: las acciones de fomento del empleo para personas con discapacidad en Centros Especiales de Empleo tienen como finalidad mejorar el nivel de inserción laboral de este colectivo, rompiendo las barreras que dificultan su incorporación al mercado de trabajo, y potenciando la creación de empleo estable adecuado.

Inserción laboral

Políticas de orientación, intermediación y colocación que ofrecen información y orientación laboral, al tiempo que se gestionan las colocaciones y posibles contrataciones a partir de los requerimientos de los empleados que precisan las empresas y de sus capacidades para optar a ellos.

7. Acciones de inserción laboral: formada por acciones cuyo objetivo la integración socio laboral de trabajadores en situación o riesgo de exclusión social en empresas de inserción y en empresas de trabajo.

FUENTE: Elaboración propia a partir de información de SERVEF.

2. Metodología

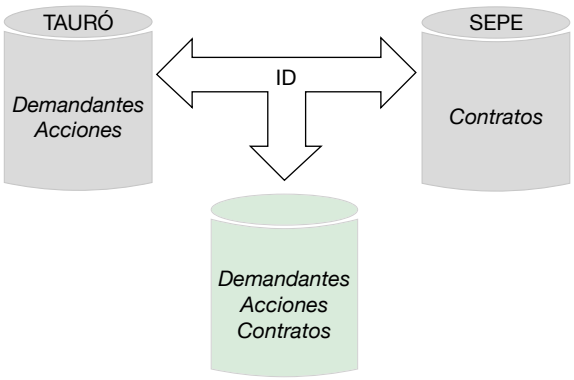
La comparación de métodos de medición de impacto presentada en este artículo cubre las PAE de formación, fomento de empleo e inserción laboral aplicadas en la Comunidad Valenciana en el periodo 2012-2015.

2.1. Fuentes de información utilizadas

La información utilizada para la medición de los indicadores se ha obtenido a partir de la integración de bases de datos TAURÓ del Servicio Valenciano de Empleo y Formación (SERVEF) y por el Servicio Público de Empleo Estatal (SEPE). La base de datos del SEPE contiene información sobre los contratos de trabajo firmados durante el periodo de análisis (2012-2015). En concreto, la base del SEPE contiene casi seis millones de registros que recogen la siguiente información: identificación del trabajador; información sobre la residencia; fecha de inicio y final del contrato; características del contrato (ocupación, tipo de contrato, tipo de jornada y horas de jornada). Por otro lado, la base de datos de SERVEF contiene información sobre los demandantes de empleo de la Comunidad Valenciana, incluyendo datos de

demandantes (más de dos millones de registros con información socio-demográfica de los demandantes de empleo) y de acciones implementadas en el marco de las PAE (tres millones de registros, el código de la acción, fecha de realización, participantes, etc.). Ambas bases de datos contienen información esencial para el cálculo de los indicadores: mientras la de SERVEF nos proporciona información sobre las acciones de formación que han implementado los demandantes de empleo y sus perfiles socio-demográficos, la de SEPE contiene información sobre los contratos de trabajo y su tipología. Para proceder al análisis, ambas bases se integraron por medio de un identificador único anonimizado por usuario. Este indicador permitió cruzar información de las dos bases de datos respetando la confidencialidad de los demandantes y la legislación en protección de datos (Figura 3).

FIGURA 3
RELACIÓN ENTRE LAS BASES DE DATOS



FUENTE: Elaboración propia.

2.2. Muestra

Los ICD has sido calculados a partir de una muestra de 44.236 participantes en las acciones de las diferentes PAE. Se ha utilizado un diseño estratificado no proporcional con 3 estratos definidos por los 3 tipos de acciones evaluadas (Cuadro 2).

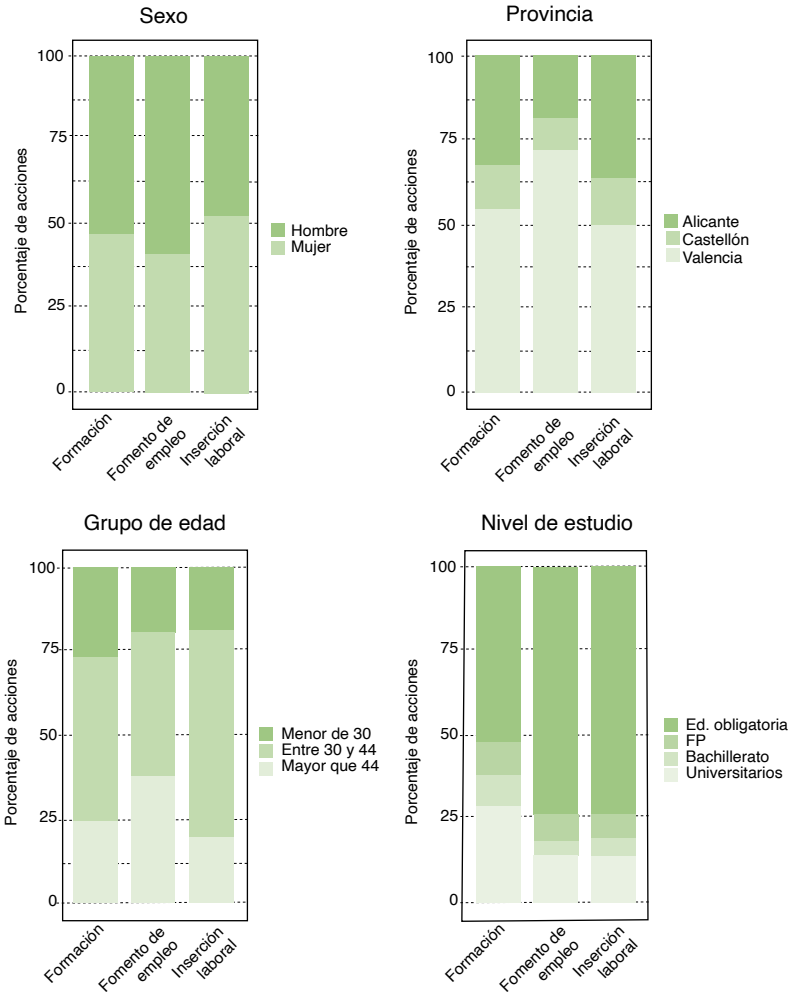
Atendiendo a las características socio-demográficas de los participantes (Figura 4), se observa que la participación de hombres y mujeres en las PAE está equilibrada. Los menores de 30 años participan más en las acciones de formación, los participantes entre 30 y 44 años participan en mayor medida en las acciones de inserción laboral y los mayores de 44 años participan más en las acciones de fomento del empleo. Los participantes que tienen educación obligatoria presentan un mayor porcentaje en las acciones de formación dual. Por el contrario, este grupo participa en menor medida en las acciones de formación para desempleados.

CUADRO 2
DISEÑO MUESTRAL

	Población	Muestra
Formación	145.372	20.715
Fomento de empleo	48.144	15.467
Inserción laboral	49.885	8.054
Total	**243.401**	**44.236**

FUENTE: Elaboración propia a partir de las bases de datos de SERVEF y de SEPE.

FIGURA 4
CARACTERÍSTICAS SOCIODEMOGRÁFICAS DE LOS PARTICIPANTES POR TIPO DE ACCIÓN



FUENTE: Elaboración propia a partir de las bases de datos de SERVEF y de SEPE.

2.3. Cálculo de los ICD: métodos baseline y cuasi-experimental

La definición de ICD es el paso inicial en la evaluación. La elección de indicadores debe responder a los criterios conocidos como SMART:

- Específicos (S de *Specific*): los indicadores deben poder ser traducidos en términos operativos. Debe quedar claro qué se está midiendo.
- Medibles (M de *Measurable*): los indicadores deben poder ser medidos en términos estadísticos, mediante la observación directa o la estimación siguiendo procedimientos aceptados científicamente. En este caso, deben estar basados en las fuentes de información disponibles que se describen posteriormente.
- Alcanzables (A de *Attainable* y de *Attributable*): los indicadores deben poder ser comparados con un objetivo a alcanzar. Al mismo tiempo, los cambios en los indicadores deben poder atribuirse a las intervenciones realizadas y no a factores ajenos. Este criterio hace necesario un enfoque experimental, que se describe más abajo.
- Realista (R de *Relevant*): los indicadores deben ser una medida válida del impacto de las PAE, es decir, que existe una relación causal entre las PAE y los valores alcanzados por los indicadores.
- En Tiempo (T de *Time*): los indicadores deben poder reflejar cambios en el tiempo, y ser obtenidos con un desfase temporal que los mantenga válidos.

Frente a la elección de un sistema óptimo de indicadores, hay que considerar aspectos pragmáticos de factibilidad y coste. Así, es necesario analizar las fuentes de información existentes antes de decidir una operación de recogida específica de información que sería muy costosa.

Para el presente estudio, se ha definido un sistema de ICD que recogen dos dimensiones diferentes de impacto de una PAE: su capacidad de incorporar al demandante al mercado laboral (*empleabilidad*) y de incrementar el número de días trabajados por este (*duración*). Estos grupos de indicadores son:

- Indicadores de *empleabilidad* que cuantifican el impacto de una acción a través del incremento (o decremento) de la probabilidad de contratación de los participantes en dicha acción (probabilidad que se estimará mediante una proporción).
- Indicadores de *duración* que cuantifican el impacto de una acción a través del incremento (o decremento) del número de días trabajados por los participantes en dicha acción calculado en forma de porcentaje para los diferentes plazos temporales.

Se trata, por tanto, de indicadores que miden diferencias, no niveles. Estos indicadores se han calculado para tres plazos temporales: corto plazo (6 meses), medio plazo (12 meses), y largo plazo (18 meses).

Los ICD se han calculado de dos formas distintas. En primer lugar, se ha aplicado un enfoque *baseline*, comparando la situación laboral del participante (empleado o desempleado) antes y después de su participación en la PAE. Los ICD de empleabilidad *baseline* se han calculado pues como la diferencia entre el porcentaje de contratados en la muestra antes y después de realizar una acción. Análogamente, los ICD de duración *baseline* se han obtenido como la diferencia entre el porcentaje de los días trabajados antes y después de realizar la acción.

La aplicación del método cuasi-experimental requiere identificar para cada uno de 44.236 participantes de la muestra a otra persona que no haya participado en dicha acción y que tenga un perfil comparable, a la que denominaremos «espejo». El espejo de cada demandante en la muestra ha sido seleccionado aleatoriamente entre todos los sujetos de la base de datos integrada que no han realizado ninguna acción y tienen un perfil similar al del demandante original en función de todas las variables disponibles. Así, cada espejo tiene el mismo sexo, provincia de residencia, grupo de edad, nivel de estudios y una diferencia menor de 30 días de experiencia laboral que el demandante de la muestra al que refleja.

Una vez identificados los espejos, los ICD de empleabilidad cuasi-experimentales se han calculado como la diferencia entre el porcentaje de participantes de la muestra contratados tras ser beneficiario de la PAE y el porcentaje de espejos contratados en ese mismo momento. Del mismo modo, los ICD de duración cuasi-experimentales se calculan como la diferencia en los porcentajes de días trabajados por los participantes de la muestra y sus espejos.

3. Resultados

La Figura 5 y el Cuadro 3 comparan los resultados de la estimación de los ICD con ambos métodos para los tres tipos de PAE y los tres horizontes temporales considerados. Hay que destacar que, dado los elevados tamaños muestrales en los tres grupos, la potencia estadística de los t-test efectuados para comparar la hipótesis nula de igualad de indicadores obtenidos con ambos métodos de cálculo es muy elevada. Como resultado, los t-test efectuados proporcionan evidencia empírica suficiente para rechazar que no existen diferencias entre los valores de los ICD *baseline* y cuasi-experimentales (el *p*-valor es inferior a 0.00 para los indicadores de empleabilidad y duración de los tres tipos de PAE y en los tres horizontes temporales).

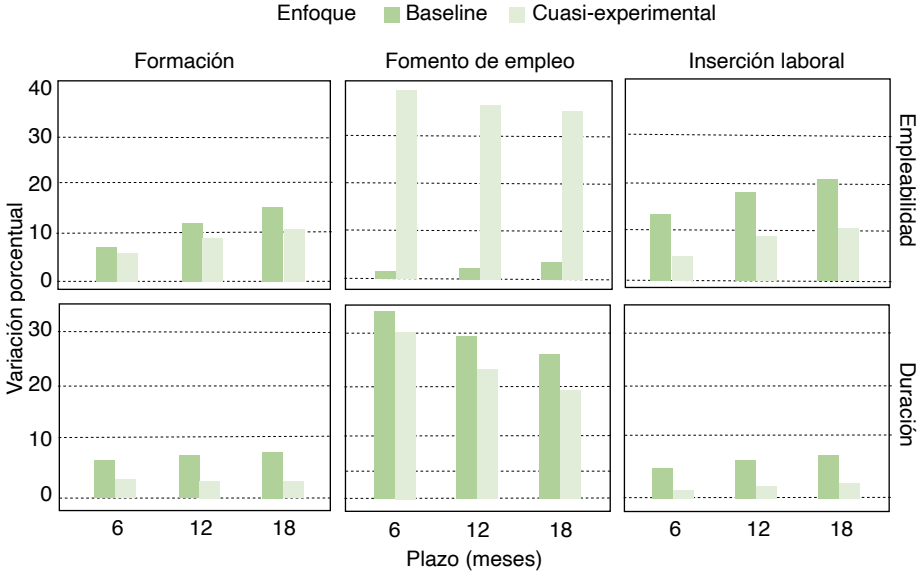
Con la única excepción de los ICD de empleabilidad para las PAE de fomento de empleo, la comparación de los ICD *baseline* y cuasi-experimentales exhiben un mismo patrón. Si bien ambos métodos de medición establecen que las PAE han tenido un impacto positivo tanto sobre empleabilidad y como duración del empleo, la estimación *baseline* es siempre superior a la cuasi-experimental. Esta diferencia sistemática puede explicarse por el hecho de que las estimaciones *baseline* confunden el efecto de las PAE con el de la evolución del ciclo económico. Dado que las políticas se han implementado en una fase expansiva del ciclo (2015-2017), la

CUADRO 3
ICD DE EMPLEABILIDAD Y DURACIÓN ESTIMADOS CON LOS MÉTODOS BASELINE Y CUASI-EXPERIMENTAL

Indicador y grupo	Baseline			Cuasi-experimental		
	6 meses	12 meses	18 meses	6 meses	12 meses	18 meses
Empleabilidad para formación	6,9	12,3	14,8	5,6	9,0	10,5
Empleabilidad para fomento de empleo	1,1	2,3	3,2	38,6	35,6	34,7
Empleabilidad para inserción laboral	13,2	17,4	20,2	4,9	8,9	10,9
Duración para formación	6,3	7,7	8,2	3,1	3,2	3,0
Duración para fomento de empleo	33,6	28,6	25,9	29,5	22,8	19,0
Duración para inserción laboral	5,4	6,8	7,5	1,1	2,2	2,5

FUENTE: Elaboración propia.

FIGURA 5
COMPARACIÓN DE LOS ICD DE EMPLEABILIDAD Y DURACIÓN ESTIMADOS CON LOS MÉTODOS BASELINE Y CUASI-EXPERIMENTAL



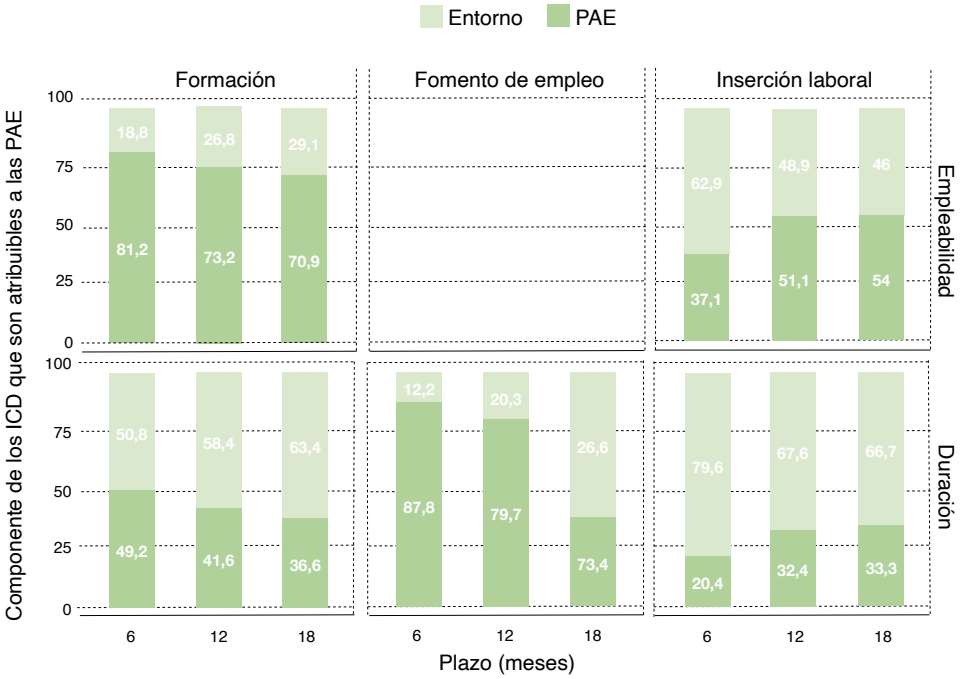
FUENTE: Elaboración propia

medición *baseline* cuantifica tanto el impacto positivo de la PAE en el incremento de empleabilidad y duración como el que se produce como consecuencia de la mejora de la situación económica (y que se hubiese producido de todas formas, aunque las PAE no hubiesen sido implementadas). Sin embargo, las estimaciones

cuasi-experimentales aíslan el impacto que puede ser atribuido estrictamente a la implementación de la política. La Figura 6 presenta cómo el impacto total de política y entorno (recogido por el valor de los indicadores *baseline*) se distribuye entre el efecto que puede ser atribuido a las PAE y a la evolución de la economía, respectivamente. Por ejemplo, para las PAE de formación, un 18,8 %, 26,8 % y un 29,1 % del valor estimado para los ICD *baseline* de empleabilidad a 6, 12 y 18 meses no se puede atribuir a la política sino al entorno. Los errores de atribución son incluso mayores en otros casos, suponiendo en algunos casos más de dos terceras partes del impacto estimado con los ICD *baseline* (indicadores de duración para inserción laboral).

Conviene señalar que los ICD de empleabilidad para PAE de fomento de empleo son los únicos que muestran un patrón diferente. En este caso, la estimación por el método cuasi-experimental proporciona valores mucho más elevados que los de la estimación *baseline* para los tres horizontes temporales considerados. Este resultado se debe a una particularidad de este tipo de PAE, que suelen incluir la firma de un contrato o la mejora de las condiciones de contrato ya existente en el momento concreto en que comienza la acción.

FIGURA 6
COMPONENTE DE LOS ICD QUE SON ATRIBUIBLES A LAS PAE
(En %)



FUENTE: Elaboración propia.

4. Conclusiones

Este artículo presenta dos evaluaciones alternativas de las políticas activas de empleo (PAE) implementadas en la Comunidad Valenciana durante el periodo 2015-2017, aplicando un enfoque *baseline* y cuasi-experimental respectivamente. Dicho análisis ofrece evidencia empírica de cómo la estimación de ICD con el método *baseline* puede llegar a proporcionar información imprecisa, la cual puede dificultar la evaluación y la optimización de políticas públicas. Este método de cálculo, basado en la comparación de la situación de los beneficiarios de la política antes y después de participar en ella, confunde el resultado atribuible a dicha política con los efectos de la variación del entorno. Así, los ICD de las PAE evaluadas en un periodo expansivo (como sucede en el caso de estudio presentado) aparecen sobrestimados si se aplica la metodología *baseline*. Esta sobrestimación es en algunos casos crítica (Figura 6), siendo, en algunos, atribuible al entorno y no a la implementación de la política más de un 60 % del valor de los ICD *baseline*.

Sin embargo, la contribución de este artículo no es señalar una debilidad del método *baseline*, ampliamente discutida en la literatura, ni de los peligros que su utilización puede conllevar en la evaluación y diseño de políticas públicas. Más allá de esto, el artículo propone un método de evaluación de fácil implementación que no requiere una privación aleatoria de los efectos de la política a una parte de la población (como es el método cuasi-experimental descrito en la sección 2) y permite obtener mediciones de los ICD más fiables y con capacidad de discriminar el nivel de impacto que puede ser efectivamente atribuible a la política misma.

En términos generales, las PAE analizadas han tenido un impacto positivo, tanto en empleabilidad, como en duración. Descontado el efecto del contexto económico (enfoque pseudo-experimental) el impacto sobre ambos tipos de indicadores sigue siendo positivo. Por tanto, podemos concluir que las PAE aplicadas están siendo efectivas. Como es de esperar dado su diferencia de diseño, los diferentes grupos de acciones tienen impacto en dimensiones diferentes del mercado laboral. Así las acciones de formación e inserción laboral incrementan en mayor medida la empleabilidad, mientras que las acciones de fomento de empleo son más efectivas en cuanto a duración.

Como conclusión, los métodos de evaluación de políticas basados en variaciones de la metodología de EAC puede proporcionar una alternativa a los habituales métodos *baseline*. Esta alternativa no presenta las dificultades logísticas ni éticas que conllevan los EAC y son capaces de discriminar cuál es el impacto que se puede atribuir a una política. Sería deseable que los métodos con base cuasi-experimental vayan sustituyendo pronto a los habituales métodos *baseline* para optimizar la efectividad de las políticas y el uso de los recursos públicos.

Referencias bibliográficas

Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, *31*(2), 3-32.

Bédécarrats, F., Guérin, I., & Roubaud, F. (2018). *The gold standard for randomised evaluations: from discussion of method to political economy* (Working Papers CEB 15-009). Universite Libre de Bruxelles.

Real Decreto-ley 3/2011, de medidas urgentes para la mejora de la empleabilidad y la reforma de las políticas activas de empleo. *Boletin Oficial del Estado,* núm. 43, de 19 de diciembre de 2011, pp. 19240 a 19260.

Caliendo, M., & Schmidl, R. (2016). Youth unemployment and active labor market policies in Europe. *IZA Journal of Labor Policy*, *5*(1), 1.

Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: gold standard, or merely standard? *Perspectives in Biology and Medicine*, *48*(4), 516-534.

Lowndes, V., Marsh, D., & Stoker, G. (Eds.). (2017). *Theory and methods in political science*. Macmillan International Higher Education.

McDermott, R. (2002). Experimental methodology in political science. *Political Analysis*, *10*(4), 325-342.

Mueller, C. E., & Gaus, H. (2017). Quasi-experimental comparison group designs for social policy evaluation. In *Handbook of Social Policy Evaluation* (pp 38-56). Edward Elgar Publishing.

Vooren, M., Haelermans, C., Groot, W., & van den Brink, H. M. (2019). *The Effectiveness of Active Labour Market Policies: A Systematic Meta-Analysis*. https://doi.org/10.1111/joes.12269

White, H., & Sabarwal, S. (2014). *Quasi-Experimental Design and Methods: Methodological Briefs-Impact Evaluation No. 8*. https://www.unicef-irc.org/publications/753-quasi-experimental-design-and-methods-methodological-briefs-impact-evaluation-no.html