

El perfilado estadístico como instrumento para la evaluación del impacto del programa Incorpora*

Statistical Profiling: an instrument for evaluating the impact of Incorpora

Jorge Casanova
Compass Lexecon y FEDEA

Florentino Felgueroso
Universidad de Oviedo y FEDEA

José Ignacio García Pérez
Universidad Pablo de Olavide y FEDEA

Sergi Jiménez-Martín
Universitat Pompeu Fabra y FEDEA

Resumen

En este artículo presentamos un ejercicio de evaluación causal en el contexto del perfilado estadístico aplicado a un programa de empleo creado por la Fundación “La Caixa” y denominado Incorpora. Este programa consiste en un mecanismo de apoyo y recolocación en el mercado de trabajo de personas en riesgo de exclusión. Nuestros resultados apuntan a que la formación impartida dentro del programa contribuye a una mejora general de la empleabilidad media de los participantes en el programa, pero se observa cómo este efecto decae en el tiempo que el beneficiario está desempleado. Por otra parte, encontramos que el programa Incorpora mejora la probabilidad de encontrar y retener el empleo, así como el tiempo trabajado e incluso el salario una vez reemplado, para todos los colectivos analizados sin importar el género, nacionalidad, sector previo de empleo o experiencia laboral previa. Concretamente, la empleabilidad aumenta entre 3 y 4 puntos porcentuales gracias al programa y el efecto sobre el número de meses acumulados en el empleo se sitúa, dependiendo de la cohorte, entre 3 y 4 meses tres años después de salir del mismo.

Palabras clave: perfilado estadístico, análisis causal, contrafactual, tratados y grupo de control.

Clasificación JEL: C35, H43, J15, J64.

Abstract

In this article, we present a causal evaluation exercise in the context of statistical profiling applied to an employment program created by La Caixa Foundation and called Incorpora. This program consists of a mechanism of support and repositioning in the labor market of people at risk of exclusion. Our results suggest that the training given within the program contributes to a general improvement in the average employability of the participants in the program, but it is observed that this effect declines with unemployment duration. On the other hand, we find that the Incorpora program improves the probability of finding and retaining employment as well as the time worked and even the salary once re-employed for all the groups analyzed regardless of gender, nationality, previous sector of employment or previous work experience. Specifically, employability increases between 3 and 4 percentage points thanks to the program and the effect on the number of months accumulated in employment is situated, depending on the cohort, between 3 and 4 months three years after leaving it.

Keywords: statistical profiling, causal evaluation, counterfactual, treated and control groups.

* Agradecemos a la Fundación “La Caixa” la financiación recibida para elaborar este estudio y el permiso para publicar algunos resultados anticipándonos a la publicación del informe final. Las opiniones contenidas en este documento son responsabilidad exclusiva de los autores y no pueden atribuirse a Compass Lexecon ni a cualquier otra parte.

1. Introducción

La situación de alto desempleo y la persistencia en dicho estado que sufren muchos de los trabajadores que son despedidos o que buscan empleo por primera vez hacen cada vez más necesario la inversión en políticas de formación para hacer frente al desajuste entre las capacidades y competencias de una mano de obra en situación de desempleo que no corresponden con las necesidades y demandas de la economía productiva. Asimismo, una estancia prolongada en el desempleo puede deteriorar no solo las competencias de los desempleados, sino la motivación y actitud de estos hacia la búsqueda de empleo. Por ello, también se hace necesario reforzar la motivación para la búsqueda activa, el desarrollo de estrategias de activación y el fomento de iniciativas de emprendimiento empresarial o de incentivos a la contratación. Todas estas políticas o servicios se desarrollan normalmente por las oficinas públicas de empleo en base a lo que se conoce como políticas activas de empleo. También hay algunas instituciones privadas, normalmente fundaciones u organizaciones sin ánimo de lucro, que dedican parte de sus esfuerzos a desarrollar programas de apoyo a desempleados que pueden ser enmarcados dentro de este tipo de políticas activas de empleo. Así, en el conjunto de España, según datos de la OCDE, se dedica a estas políticas en torno al 0,7% del PIB (datos de 2018), cifra superior en dos décimas a la media del conjunto de países de dicha organización, pero lejana de lo que dedican países como Suecia (1,1%) o Dinamarca (1,81%). Los altos niveles de desempleo en España y, sobre todo, la persistente duración de este para ciertos colectivos, hacen muy necesario dedicar un mayor esfuerzo a identificar cuál es el efecto esperado de cada una de estas políticas en los distintos colectivos que componen el conjunto de demandantes de empleo. No es solamente por razones de eficiencia económica que se hace necesario analizar en qué partidas es más eficiente focalizar dicha inversión, sino que también se hace necesario por razones de oportunidad. Una vez clasificados los demandantes en función de su nivel esperado de empleabilidad, sería muy conveniente analizar qué políticas o servicios de entre los disponibles en una oficina de empleo son más eficaces de cara a reinsertar al demandante de nuevo en el *pool* de empleados. Por ello, necesitamos conocer qué programas o servicios son más eficaces en base a los efectos que cada uno de ellos tiene sobre el desempeño de sus usuarios en el mercado de trabajo durante y después de recibir dicho servicio.

Este análisis de desempeño se realiza en la literatura económica mediante técnicas de evaluación causal. Estas técnicas se han desarrollado dentro de una literatura de Evaluación de Políticas Públicas, en creciente expansión en los últimos años. Esta área de investigación persigue la obtención de información útil para valorar la racionalidad, coherencia, eficacia, idoneidad e impacto de las intervenciones públicas. Como herramienta al servicio del gestor público, la utilización de técnicas de evaluación mejora la intervención pública y aumenta su utilidad social; no solo por la información que suministra a todos aquellos interesados en conocer los resultados de las políticas evaluadas, sino por la posibilidad de aplicar sus conclusiones para retroalimentar y mejorar el diseño y elaboración de nuevas intervenciones, y así mejorar los resultados finales y sus impactos.

Esta metodología se complementa muy bien con las técnicas de perfilado estadístico que cada vez se usan más para clasificar a los demandantes de empleo en función de su empleabilidad (Desiere *et al.*, 2019). Concretamente, tras clasificar a estos demandantes en distintos segmentos en función de su empleabilidad, se distingue dentro de cada segmento entre colectivos elegibles y no elegibles de cada uno de los servicios a analizar y se analiza, mediante técnicas de emparejamiento o de análisis de discontinuidad, si se observan efectos significativos en cada segmento de las políticas o servicios a evaluar. El objetivo es comparar individuos elegibles con los no elegibles, pero de un mismo nivel de empleabilidad. Esto quiere decir que se estima el impacto comparando aquellos que reciben el servicio con aquellos que no, pero tenían las mismas probabilidades de encontrar un empleo en ese momento.

En este documento se presenta un ejercicio de evaluación causal en el contexto del perfilado estadístico aplicado a un programa de empleo creado por la Fundación «La Caixa» denominado Incorpora. Este programa consiste en un mecanismo de apoyo y recolocación en el mercado laboral de personas en riesgo de exclusión. A través de un asesoramiento individualizado y con recursos de formación y búsqueda de empleo, la intención del programa es lograr la inserción laboral y el mantenimiento del empleo de desempleados de difícil inserción, así como de colectivos en riesgo de exclusión social.

Concretamente, en nuestro análisis del programa Incorpora construimos una metodología de perfilado estadístico que podrá ayudar a incrementar la efectividad de la intermediación hecha por los propios técnicos del programa. A partir de los datos suministrados por la plataforma Incorpora, obtenemos un indicador cuantitativo que aproxima el nivel de empleabilidad del beneficiario, esto es, la verosimilitud de una eventual reincorporación al mercado de trabajo en un plazo dado. Una vez creado dicho indicador, evaluamos, entre otros aspectos, el efecto de la formación sobre dicho nivel de empleabilidad. En segundo lugar, y como objetivo central, abordamos el análisis del impacto causal a medio plazo del programa Incorpora sobre diversos indicadores de inclusión laboral de los beneficiarios, en comparación con un amplio conjunto de individuos que no formaron parte de este, pero han tenido vidas laborales previas muy similares y que, por tanto, pueden ser considerados como grupo de control.

Los resultados de la primera parte apuntan a que la formación impartida dentro del programa contribuye a una mejora general de la empleabilidad media de los participantes en el programa, pero se observa cómo este efecto decae en el tiempo que el beneficiario está desempleado. No obstante, observamos que la asignación de los cursos no prioriza sistemáticamente a los beneficiarios más vulnerables dentro de cada colectivo.

Por otra parte, encontramos que el programa Incorpora mejora los parámetros de integración laboral (probabilidad de encontrar y retener el empleo, tiempo trabajado e incluso salarios) a medio plazo de todos los colectivos analizados sin importar el género, nacionalidad, sector previo de empleo o experiencia laboral previa. Asimismo, la permanencia en el empleo es mayor para los beneficiarios del programa que

para los individuos del grupo de comparación o control, generando también mayores ganancias tanto en términos de salario como de meses acumulados en el empleo.

El uso de mecanismos de apoyo y formación para el empleo no es una novedad. En un metaanálisis, Card, Kluve y Weber (2018) evalúan 199 estudios y muestran como las políticas de formación y del sector privado obtienen mejores resultados en varias dimensiones (incluido salarios y tiempo empleado) tanto a medio como largo plazo.

En el caso de España, la evidencia es escasa debido a la falta de datos y la disposición a evaluar políticas públicas. Si bien es cierto que existen programas de inserción laboral, no tenemos conocimiento de algún estudio que los evalúe. Por el contrario, la literatura se ha centrado mayormente en estudiar políticas activas de empleo enfocadas en incentivar el empleo privado y en programas de formación de parados. Sobre este último, podemos mencionar el estudio de Arellano (2010), que evaluó los programas de formación dirigidos a desempleados por el Plan Nacional para la Formación y la Recolocación. Usando metodologías de estimación causal, encuentra que los cursos de formación fueron eficaces para reducir la duración de desempleo posterior entre 24 y 39 días.

La estructura del resto del artículo es la siguiente. Primero, describiremos el programa y su evolución. Segundo, dividiremos el análisis en dos partes (perfilado y análisis causal) y, sobre estas, haremos una descripción de los datos y la metodología utilizada. Terminamos con unas breves conclusiones.

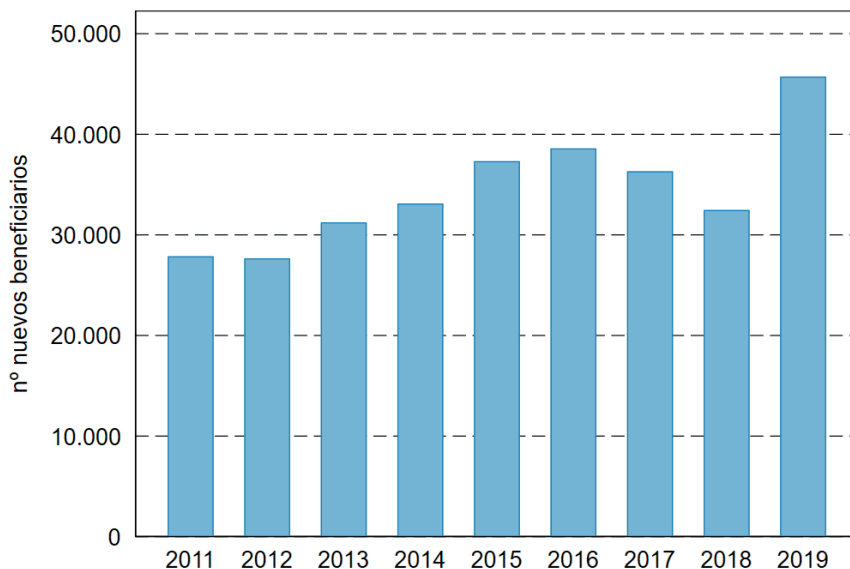
2. Descripción del programa

Desde su fundación en 2006, el programa Incorpora ha apoyado a 330.814 personas en riesgo de exclusión social a través de una red de 617 entidades y 3.307 técnicos especializados¹. Tal y como muestra la Figura 1, la media de incorporaciones al programa está en torno a las 40.000 anuales, lo que, a su vez, se ha visto complementado con una oferta de más de 450 cursos anuales en promedio que ofrecen herramientas necesarias en la búsqueda del empleo y en todo tipo de ocupaciones. Por otro lado, y gracias a la labor de las entidades asociadas al programa, se realiza una intensa tarea de búsqueda de nuevas oportunidades de acuerdos con empresas de todos los sectores, fruto del cual el número de ofertas de empleo acumuladas ha crecido hasta los 206.636 en 2019, habiéndose firmado hasta esta fecha más de 7.100 convenios con empresas colaboradoras. Esto ha hecho además que el número de inserciones laborales creciera tanto extensivamente como intensivamente, con 14.006 nuevos contratos laborales solo en 2019 y con una línea complementaria de apoyo al autoempleo que ha logrado la creación de microempresas en todo el territorio.

Desde el 2016, Incorpora complementa además el apoyo a los beneficiarios con una nueva línea de autoempleo que asesora a aquellos interesados en emprender.

¹ Datos actualizados en diciembre del 2019.

FIGURA 1
EVOLUCIÓN DEL NÚMERO DE ALTAS DE BENEFICIARIOS



FUENTE: Elaboración propia en base a datos programa Incorpora.

Esto ha resultado en la creación de nuevas empresas en todo el territorio y que, además, se construyen sobre la base de una estrategia planificada para su éxito.

La participación en Incorpora funciona de la siguiente manera: cualquier interesado puede acceder acudiendo a una de las entidades sociales donde, en primer lugar, se tomarán datos para hacer una evaluación integral de su situación. A partir de entonces, Incorpora mantiene el acompañamiento a la persona durante todo el proceso de integración sociolaboral. Permite a cada persona en situación o en riesgo de exclusión social definir un itinerario propio, con el asesoramiento del personal técnico. Incorpora impulsa y da apoyo a la persona, cualquiera que sea su punto de partida, para avanzar hacia la inserción laboral, por cuenta ajena o propia, ya que el autoempleo también está contemplado dentro de las líneas de actuación del programa.

El objetivo del presente trabajo se concreta en evaluar la efectividad de la metodología Incorpora. Para alcanzar nuestros objetivos, dividimos el análisis en dos partes.

En la primera parte, partimos de un modelo de perfilado estadístico basado con los datos de carácter mensual suministrados por la plataforma del programa Incorpora. Los resultados de la estimación son predicciones estadísticas derivadas de un modelo econométrico estimado exclusivamente a partir de la información de los beneficiarios. Todos estos datos son usados para obtener un indicador cuantitativo que aproxime el nivel de empleabilidad del beneficiario, esto es, la distancia con su probable reincorporación al mercado de trabajo.

Una de las dificultades para evaluar el efecto de la formación es la posible endogeneidad en la asignación de esta por parte de los técnicos del programa. Al haber cierta tendencia a formar parte de determinados colectivos, es necesario ajustar las regresiones añadiendo instrumentos que tengan relación con dicha formación, como son, por ejemplo, las experiencias de empleo y de formaciones previas.

En la segunda parte, hacemos un análisis causal con datos trimestrales comparando beneficiarios del programa Incorpora con personas activas de la muestra continua de vidas laborales de características similares y que se pueden considerar grupos de control. Para los colectivos de inmigrantes, discapacitados, jóvenes y personas de 30 a 44 años y mayores de 45, creamos grupos de tratamiento y control a partir de técnicas de análisis causal y, posteriormente, estimamos los efectos del programa en los tres-cinco años posteriores a la entrada en el mismo.

3. Perfilado estadístico

3.1. Datos

Incorpora posee una plataforma *online* a través de la cual todas las entidades registran los seguimientos de beneficiarios, las ofertas de empleo y los cursos de formación. Los datos provienen de dos ficheros distintos: uno referido a las características de los beneficiarios y otro a sus acciones. En el primer fichero, la información es única y se refiere a características demográficas, sociales (colectivos) y de situación previa a la incorporación.

En el segundo fichero, la fecha de cada acción indica el momento en que el técnico realiza el registro en la plataforma y, por tanto, se asume que entre una acción y otra no ha habido más cambios que los que se indican. A su vez, a partir de estos datos se obtiene la situación laboral de cada beneficiario, que será utilizada para generar la variable dependiente en la muestra de estimación y, también, los valores de la mayoría de las variables explicativas del modelo de empleabilidad.

Una de las ventajas de la plataforma es que permite detallar el tipo de acción que tiene el individuo. No obstante, y para simplificar el análisis, en una primera aproximación nos centraremos en agregar las acciones en seis categorías (alta, parado, formación, empleo, baja o censura).

- Alta: momento en que el individuo se registra en el programa (en el fichero se define como fecha de incorporación).
- Parado: momento en que el individuo no se encuentra inscrito en algún curso de formación o está trabajando. Esta categoría incluye acciones de asesoramiento puntual o entrevistas.
- Formación: los beneficiarios que se inscriban en un curso de formación entran dentro de esta categoría hasta que finaliza o cambian de acción en los siguientes registros.

- Empleo: los beneficiarios que estén trabajando (tanto por cuenta ajena como por cuenta propia) o realicen prácticas de empresa, entran en esta categoría.
- Baja o censura: en algunos casos, los técnicos registran la última acción del beneficiario para indicar que este se dio de baja del programa. En caso contrario, se desconocen las acciones a partir de entonces y se categorizan como censura.

Existen consideraciones metodológicas adicionales que surgieron durante el análisis y que exponemos a continuación.

En primer lugar, beneficiarios repetidos donde el identificador es el mismo pero las características son distintas. En este caso, nos quedamos con un solo registro (no hay un patrón que se repita y, por tanto, tampoco lo es la selección). En segundo lugar, individuos menores de 15 años y mayores de 65. Ninguno de estos casos se considera dentro del análisis porque el año de nacimiento es, en su mayoría, un valor atípico. En tercer lugar, múltiples acciones dentro de un mismo mes. Al ser un fichero de información mensual, nos quedamos con el último registro de cada mes. En caso de tener la misma fecha, se priman acciones referidas al empleo o la formación frente al paro.

Finalmente, individuos con un colectivo no asignado, pero características asignables. Hay beneficiarios cuyo colectivo no estaba asignado o, estando dentro de la categoría de «Otros», poseía características similares a otros. Para homogeneizar los grupos, solo se considera dentro de «Otros» a personas entre 31 y 44 años (edad *prime*) y con nacionalidad española (o sin asignar).

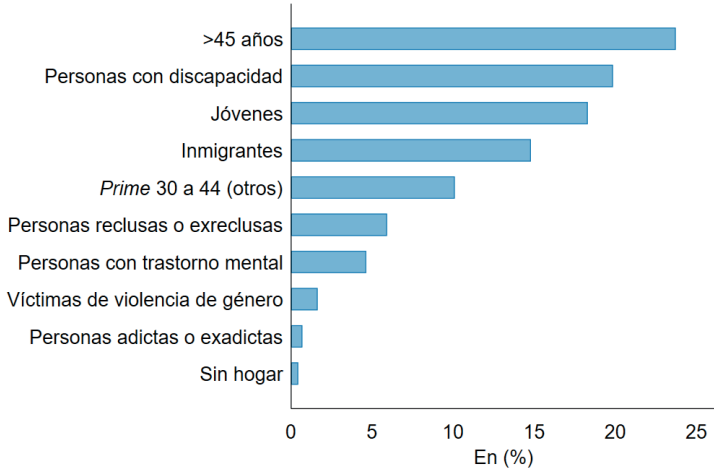
Una vez aplicadas estas restricciones, la muestra de análisis considera como beneficiarios aquellos incorporados al programa en el periodo 2011-2019. Este colectivo incluye un total de 302.632 personas, siendo entre estos los mayores de 45 años y los discapacitados aquellos con mayor peso. Son en su mayoría españoles (véase Figura 2), seguido de latinoamericanos, y mayoritariamente se han dado de alta en los grupos territoriales de Barcelona, Madrid o Andalucía. Por grupos de edad, la mitad de los beneficiarios tienen entre 31 y 50 años y, si lo vemos por colectivo, la distribución está más concentrada entre los 31 y 40 años para colectivos cuya clasificación no depende de la edad.

En cuanto al género, la distribución por colectivos es bastante desigual en casos como las personas adictas o exadictas o los reclusos o exreclusos (véase Figura 2). Esta heterogeneidad es la que buscamos captar al momento de estimar el efecto de la formación ya que nos permite, además, saber a quién beneficia esta, más o menos, y cómo podría enfocarse en el futuro.

Al ser la empleabilidad nuestro objeto de estudio, analizamos también gráficamente cómo se distribuye el porcentaje de beneficiarios que encuentran empleo por las características ya consideradas. En general, como se recoge en la Figura 3, no parece haber diferencias muy significativas entre cualquiera de las categorías, es decir, aproximadamente el 50 % de los beneficiarios consiguieron algún empleo durante su periodo de pertenencia al programa.

FIGURA 2
DESCRIPTIVOS SOBRE LOS BENEFICIARIOS

a) Porcentaje de beneficiarios por colectivo



b) Distribución del género por colectivo

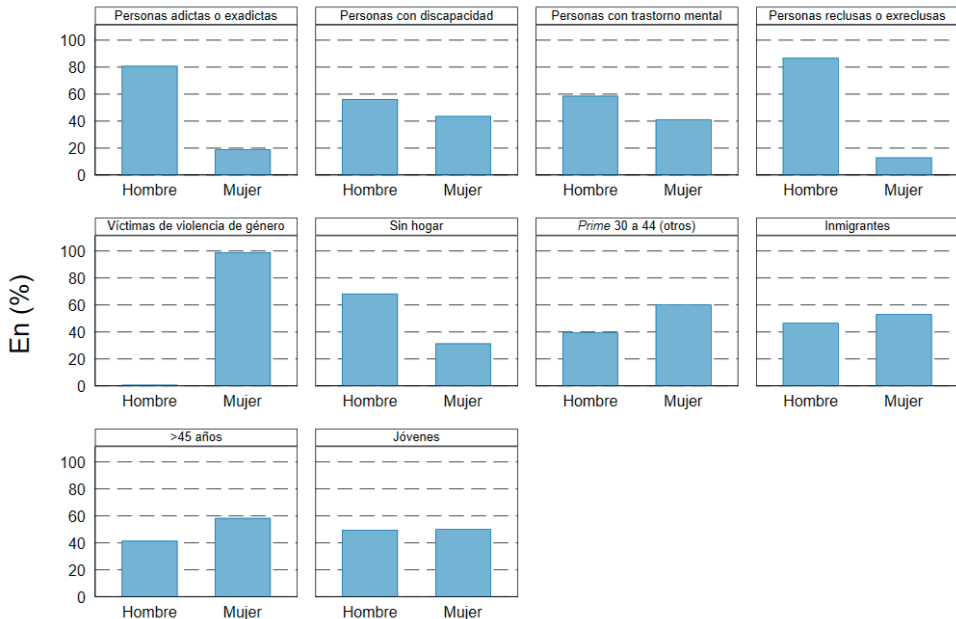
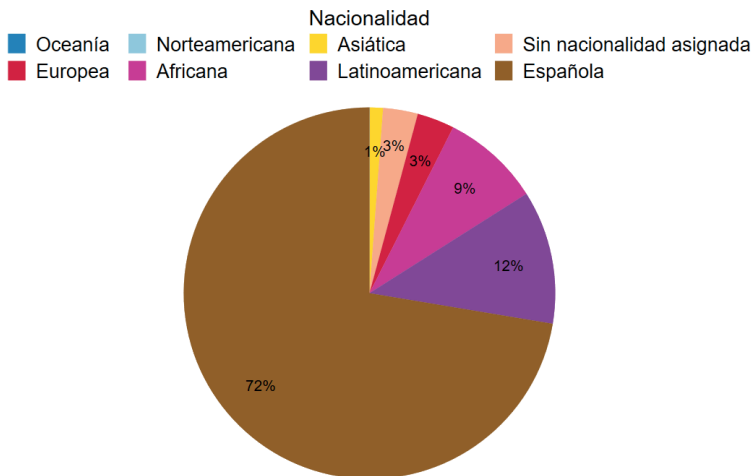
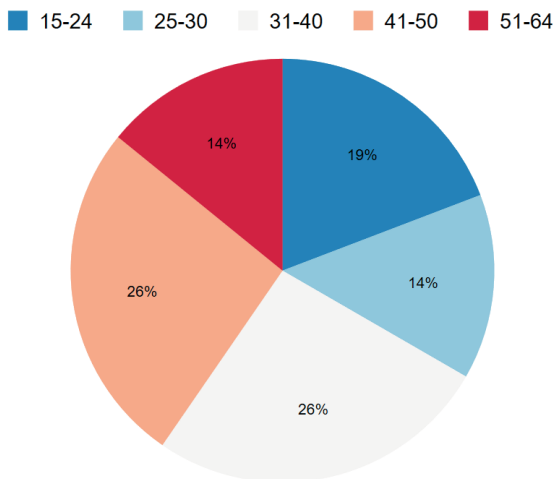


FIGURA 2 (Continuación)
DESCRIPTIVOS SOBRE LOS BENEFICIARIOS

c) Porcentaje de beneficiarios por nacionalidad

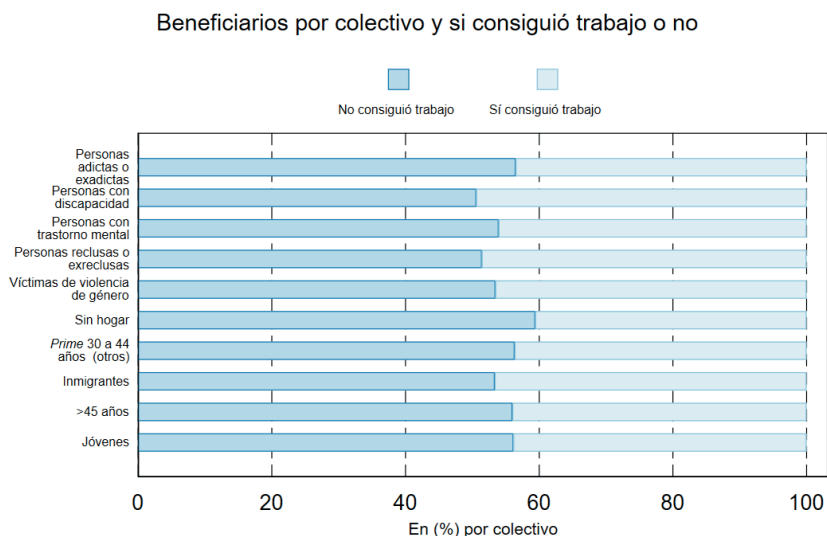


d) Porcentaje de beneficiarios por edad



FUENTE: Elaboración propia en base a datos programa Incorpora.

FIGURA 3
DESCRIPTIVOS SOBRE LOS BENEFICIARIOS Y EL PORCENTAJE DE EMPLEADOS



NOTA: La información sobre el estado laboral de todos los participantes proviene de la plataforma Incorpora bajo el estado «Trabajo por cuenta ajena» o «Trabajo por cuenta propia». Condicionamos a tener trabajo si se encuentra más de un mes empleado tras pasar por el programa.

FUENTE: Elaboración propia en base a datos programa Incorpora.

3.2. Metodología

Una vez construida la base de datos estimamos un modelo econométrico de empleabilidad para los beneficiarios del programa, teniendo en cuenta la posible endogeneidad en la asignación de formación al participante en el programa.

Una de las características de Incorpora es la posibilidad de ofrecer cursos de formación para mejorar los conocimientos y habilidades de los beneficiarios. Esta asignación no es aleatoria y depende, no solo de sus características, sino también del técnico que los evalúa y la entidad en la que se registra. Por tanto, una dificultad del análisis es determinar aquella información que permita corregir la posible endogeneidad en la asignación de la formación.

Otra de las dificultades para evaluar la formación es que, durante ese tiempo, el beneficiario puede o no estar trabajando y el efecto que observemos no sea homogéneo, sino que deba desagregarse.

Resulta necesario, en ambos casos, incluir la información específica sobre el técnico que acompaña al beneficiario y el tipo de formación que realiza para poder cuantificar el efecto de esta.

El modelo con el que se desarrolla el perfilado es un modelo econométrico de elección discreta, concretamente un modelo Probit. Siguiendo la metodología propuesta por Felgueroso *et al.* (2018), la variable dependiente es una variable binaria que adopta un valor de 1, si el beneficiario transita del desempleo al empleo antes del periodo T ; y de 0, en caso contrario.

La expresión analítica del modelo a estimar en este caso es:

$$\Pr[\text{Salida al empleo} | x] = \Phi(X'\beta + \varepsilon) \quad \text{con } \varepsilon \sim N(0,1)$$

donde Φ denota una función de probabilidad y el término de error tiene una distribución normal estándar de varianza normalizada a 1.

Las variables de control (X) que se consideran son las siguientes:

- *Características de los beneficiarios*: edad, nacionalidad, colectivo, número de hijos, nivel de formación, sexo, grado de discapacidad, ocupación solicitada, disponibilidad para viajar vehículo propio, idioma, situación laboral, proximidad del empleo deseado, tipo de jornada deseada y efecto fijo de técnico de atención en el programa.
- *Variable de formación*: cursos cortos (menos de 3 meses de duración) y cursos largos (más de 3 meses). Cada una representa una variable discreta con tres etapas: sin formación, durante la formación y después de la formación.
- *Características de las entidades*: tipo de entidad (de red o coordinadora), servicios que ofrece (punto formativo, punto reincorpora, punto retorno educativo, Servicio comunidad, punto auto-ocupación, Fondo Social Europeo y/o punto Incorpora joven), tipo de colectivo principal y secundario.
- *Características del mercado Incorpora*: número de ofertas de empleo (por entidad compartida y mes), desagregadas entre duración del contrato (1 semana, 1 mes, 3 a 12 meses y 1 a 3 años), número de horas (al menos 40 horas o menos), colectivo de preferencia (discapacitado, en riesgo de exclusión social o indiferente), sector de actividad (industria, comercio, agricultura o servicios).
- *Características de las empresas inscritas*: Número de empresas inscritas en Incorpora por mes, sector y si tienen o no convenio firmado.
- *Características temporales*: año y mes.
- *Características de las acciones*: duración del paro, duración hasta la formación, si tiene o no formación en el episodio previo y si tiene o no un empleo previo.

Siguiendo, en ese caso, la siguiente especificación:

$$X = X'_{\text{benef}}\beta_{\text{benef}} + X'_{\text{formación}}\beta_{\text{formación}} + X'_{\text{entidad}}\beta_{\text{entidad}} + X'_{\text{mercado}}\beta_{\text{mercado}} + X'_{\text{empresas}}\beta_{\text{empresas}} + Z'\alpha + X'_{\text{acciones}}\beta_{\text{acciones}}$$

Y, en el caso de formación, tenemos que:

$$X'_{\text{formación}} \beta_{\text{formación}} = \sum_{j=1}^2 X_{\text{durante } j} \beta_{\text{durante } j} + X_{\text{después } j} \beta_{\text{después } j} + \sum_{i=1}^9 \text{mes } i \times X_{\text{después } j \text{ mes } i} \beta_{\text{después } j \text{ mes } i}$$

donde $j = 1$ si el curso es corto y $j = 2$ si es largo.

Una vez estimado el modelo, con una determinada periodicidad, la predicción para cada uno de los individuos se define como el grado de empleabilidad de dicho individuo (*empleabilidad* = $\Phi(X' \hat{\beta})$)

La *probabilidad estimada de salida del paro al empleo en un mes* para cada beneficiario se puede considerar como un indicador de empleabilidad individual que se utiliza a continuación con los siguientes propósitos:

En primer lugar, clasificar a los beneficiarios en grupos de empleabilidad y así poder analizar las características de los beneficiarios incluidos en cada grupo. La clasificación propuesta aquí se realiza en función de la pertenencia a uno de los cuartiles de las distribuciones de la probabilidad de salir del paro al empleo. Así, establecemos cuatro grupos: empleabilidad alta, media alta, media baja y baja. Esto nos permite ver cómo están repartidas las características de los beneficiarios entre cada umbral e identificar aquellas personas que presentan mayores dificultades para encontrar un empleo.

En segundo lugar, estimar, entre otros posibles factores, el efecto neto de la formación y cómo este varía en el tiempo. Recurrimos al efecto marginal promedio para cuantificar el efecto durante y después de la formación a partir de la siguiente fórmula:

$$\text{Efecto marginal promedio} = \frac{1}{n} \sum_{i=1}^n [\Phi(x'_i \beta | x_i^k = 1) - \Phi(x'_i \beta | x_i^k = 0)]$$

Consideramos tres posibles tipos de estado respecto a la formación:

- Antes: si el beneficiario no ha realizado ningún curso.
- Durante: si en el periodo t se encuentra realizando la formación.
- Después: una vez finalizado el curso y hasta su censura o baja del programa.

En una fase ulterior, podemos analizar dicho efecto para cada colectivo, y dentro de estos, por género o grupo de edad, para tomar en cuenta la heterogeneidad de beneficiarios que hay.

3.3. Resultados

A continuación, se presentan los principales resultados del modelo previamente descrito. Todas las regresiones incluyen los controles previamente señalados (efectos fijos sobre características demográficas, variación temporal (año y mes), características y herramientas del programa Incorpora (cursos, número de ofertas, tipos de entidad, entre otros) y características previas a la entrada en Incorpora (vida laboral previa, situación laboral). En primer lugar, en el Cuadro 1 presentamos los efectos marginales referidos a las variables de formación ($X'_{\text{formación}}$).

CUADRO 1
EFFECTOS MARGINALES DE LOS CURSOS DE FORMACIÓN SOBRE
LA REGRESIÓN DE PERFILADO

Variables	Efectos marginales	Variables	Efectos marginales
Cursos cortos		Cursos largos	
Durante	0,0637*** (0,00058)	Durante	0,0250*** (0,00116)
Después	0,00517*** (0,000532)	Después	0,00620*** (0,00142)
Después × mes 1	0,0221*** (0,00111)	Después × mes 1	0,0171*** (0,00338)
Después × mes 2	0,00562*** (0,00114)	Después × mes 2	-0,0159*** (0,00347)
Después × mes 3	-0,0179*** (0,00115)	Después × mes 3	-0,0123*** (0,00352)
Después × mes 4	-0,0197*** (0,00116)	Después × mes 4	-0,0169*** (0,0036)
Después × mes 5	-0,0155*** (0,00118)	Después × mes 5	-0,0177*** (0,00367)
Después × mes 6	-0,0122*** (0,00121)	Después × mes 6	-0,00882** (0,00376)
Después × mes 7	-0,0146*** (0,00123)	Después × mes 7	-0,0117*** (0,00387)
Después × mes 8	-0,0184*** (0,00126)	Después × mes 8	-0,0182*** (0,00395)
Después × mes 9	-0,0150*** (0,00128)	Después × mes 9	-0,0141*** (0,00405)
Observaciones	4.285.429	Observaciones	4.285.429
R-cuadrado	0,05	R-cuadrado	0,05

NOTA: Errores estándares en paréntesis. * $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$. Los regresores omitidos en la tabla incluyen características demográficas (X'_{benef}), de las entidades (X'_{entidad}), de las ofertas de empleo publicadas en Incorpora (mercado Incorpora), de las empresas inscritas en el programa (X'_{empresas}), *dummies* de año y mes (Z) y características de las acciones de los beneficiarios (X'_{acciones}). El listado más detallado se encuentra en la sección 3.2.

FUENTE: Elaboración propia.

El efecto de la formación sobre la probabilidad de salir del paro resulta significativo tanto durante como después de haberla percibido, para todos los beneficiarios del programa considerados en el periodo 2011-2019. Este resultado es importante ya que indica que la formación no se considera un periodo de inactividad, sino que mejora las oportunidades de empleo. Esto tiene que ver con la orientación de los cursos, ya que en muchos casos ofrecen herramientas relacionadas con la búsqueda de trabajo.

En la Figura 4 presentamos un resumen de los resultados respecto a la formación: efecto marginal durante el periodo formativo y después del mismo. Asimismo, representamos la evolución del efecto marginal en el tiempo, tanto para cursos cortos como para cursos largos. En este último caso, también presentamos los resultados según el colectivo estudiado.

A partir de los datos suministrados por la plataforma Incorpora, se estimó que los cursos de formación aumentan las probabilidades de encontrar un empleo entre un 3 y un 6% adicional, siendo los de menor duración aquellos que mayor efecto tienen. Este, además, se observa desde el momento en que el beneficiario inicia la formación, ya que el hecho de asistir parece tener un impacto inmediato sobre su empleabilidad. No obstante, una vez finalizada la formación las ganancias del curso disminuyen conforme transcurren los meses.

FIGURA 4
EFFECTO DE LA FORMACIÓN SOBRE LA SALIDA DEL PARO

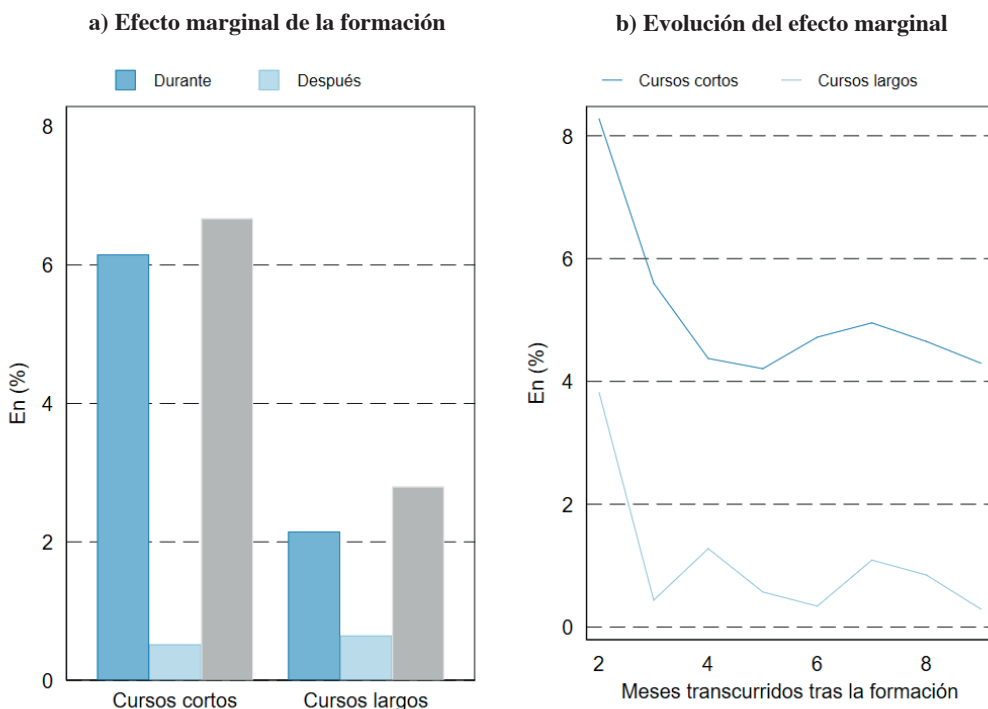


FIGURA 4 (Continuación)
EFFECTO DE LA FORMACIÓN SOBRE LA SALIDA DEL PARO

c) Efectos de la formación

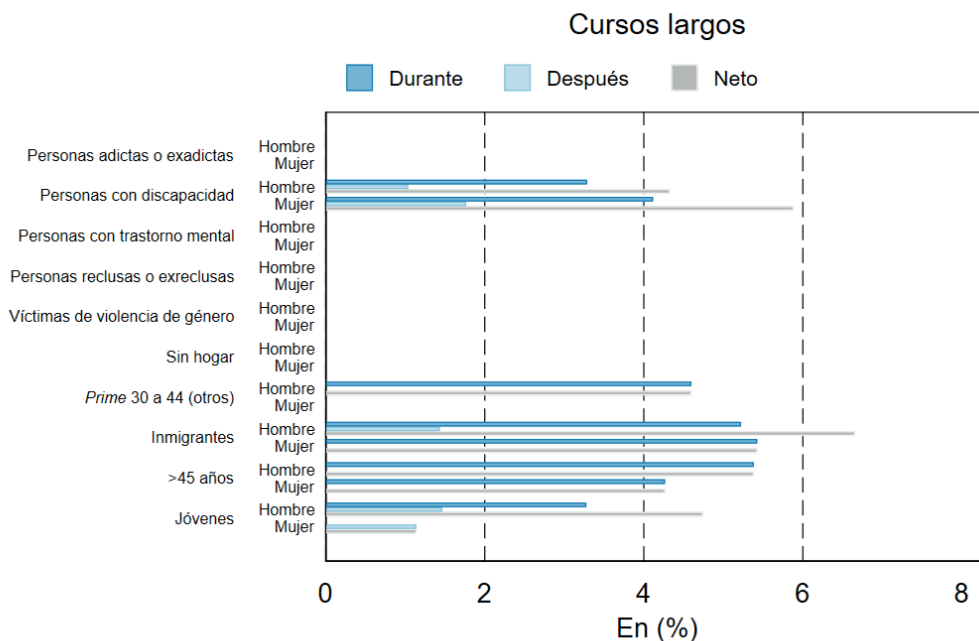
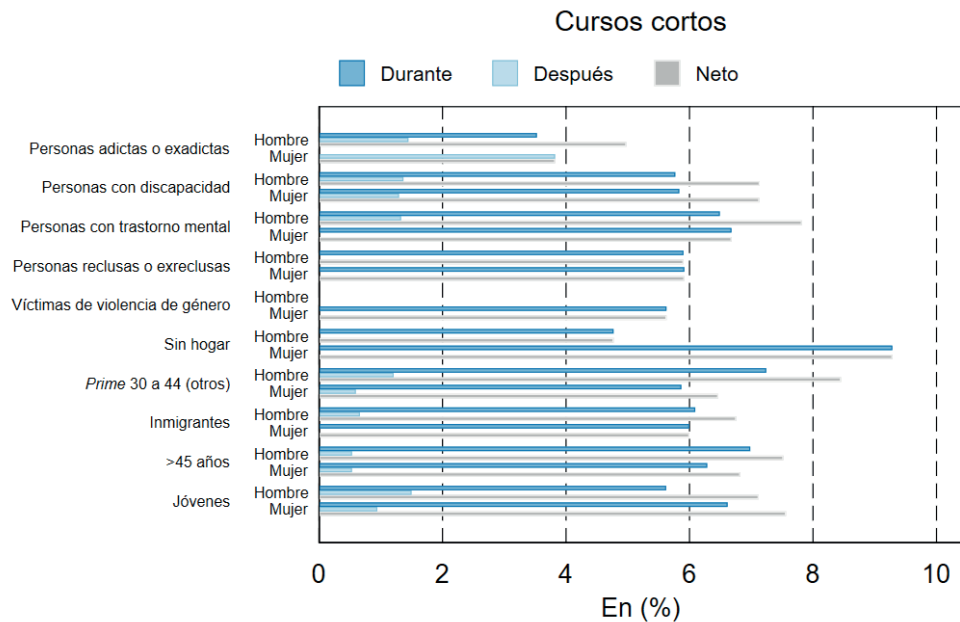
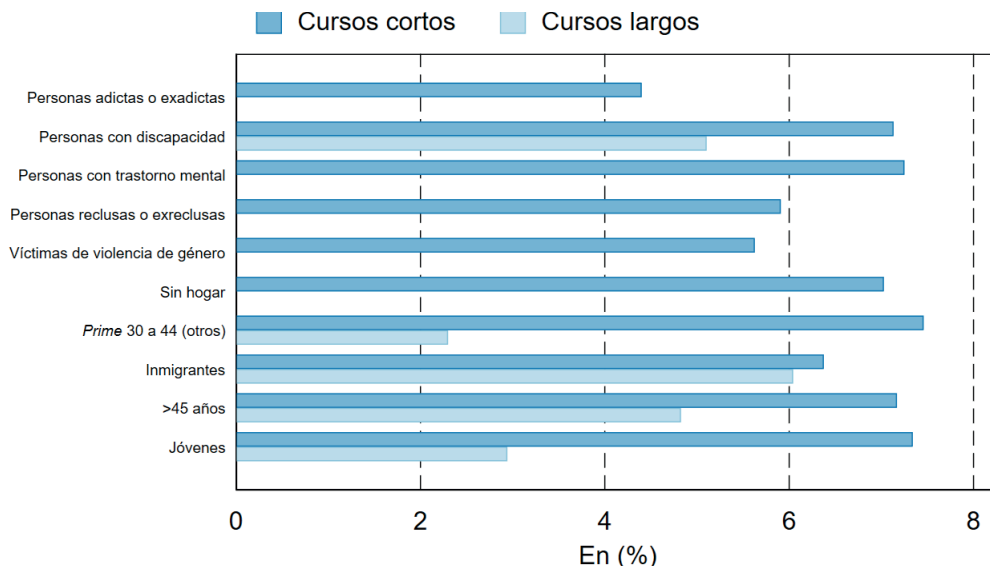


FIGURA 4 (Continuación)
EFFECTO DE LA FORMACIÓN SOBRE LA SALIDA DEL PARO
d) Efecto marginal de la formación sobre la empleabilidad (según el colectivo)



FUENTE: Elaboración propia en base a datos programa Incorpora.

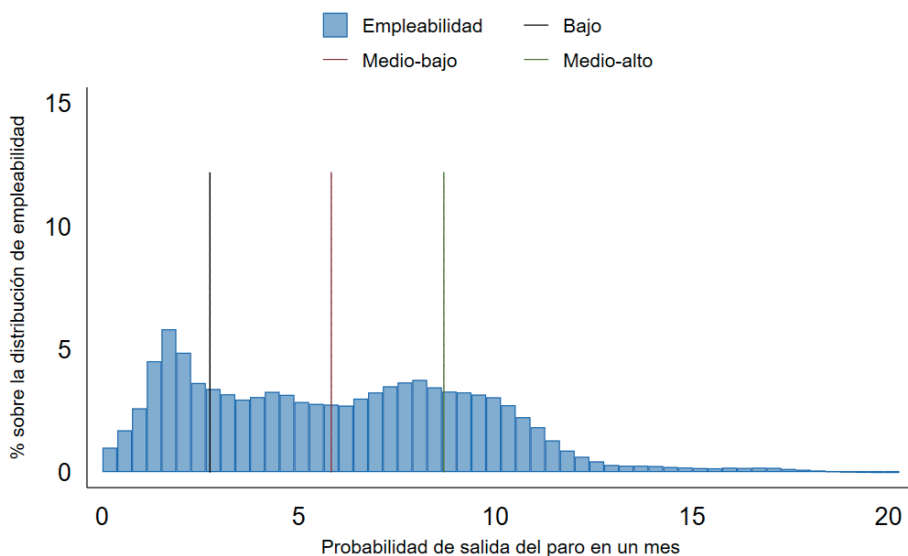
Por otra parte, es importante analizar si estas diferencias también se ven reflejadas en la distribución de la empleabilidad. En la Figura 5 presentamos la distribución de empleabilidad mensual (probabilidad de dejar el no empleo en un mes) para el conjunto de beneficiarios. En el conjunto de los beneficiarios la gran masa de probabilidad se encuentra entre el 0 y el 13%, estando la moda de la distribución muy a la derecha, entre el 1 y el 2%. El percentil 50 en algo menos del 6% y el percentil 75 entorno al 8%.

En segundo lugar, comparamos la empleabilidad de los beneficiarios al momento de inscribirse en el programa. El Cuadro 2 señala que las probabilidades de salida del paro en un mes más altas son las de los jóvenes (6,61 puntos) y las menores las de las personas adictas o exadictas (4,53 puntos). En términos de dispersión, detectamos que la diferencia entre el primer y el tercer cuartil oscila entre 5,7 (jóvenes) y 6,15 (inmigrantes) puntos de probabilidad. Solo para las personas sin hogar la dispersión es sensiblemente menor (solo 4,6 puntos).

La Figura 6 señala importantes diferencias en la distribución de aquellos con y sin formación. Mientras que la media se encuentra entorno al 6% para aquellos que no la percibieron, esta es superior (8%) para el caso contrario.

Finalmente, una de las preguntas más relevantes para los coordinadores del programa es si los recursos en términos materiales y de capital humano se están

FIGURA 5
DISTRIBUCIÓN DE LA EMPLEABILIDAD PROMEDIO DE TODOS
LOS BENEFICIARIOS



NOTA: Las rectas indican los cuartiles estimados.

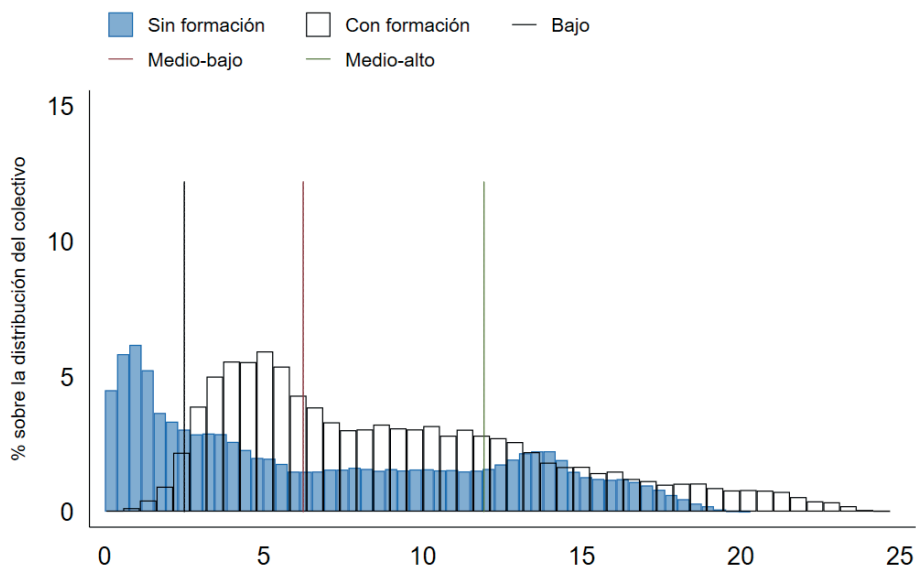
FUENTE: Construcción propia en base a datos programa Incorpora.

CUADRO 2
PERCENTILES DE EMPLEABILIDADES INICIALES POR COLECTIVO

Probabilidad de salida del paro en un mes	Percentil		
	25	50	75
Colectivo	25	50	75
Personas adictas o exadictas	2,03	4,53	7,83
Personas con discapacidad	2,07	4,56	7,89
Personas con trastorno mental	2,14	4,87	7,98
Personas reclusas o exreclusas	2,80	6,14	8,81
Víctimas de violencia de género	2,65	5,56	8,31
Sin hogar	2,19	4,56	6,86
Prime (otros) 30 a 44 años	2,94	6,01	8,77
Inmigrantes	2,87	6,03	9,02
Mayores de 45 años	3,14	6,24	8,92
Jóvenes	3,47	6,61	9,11

FUENTE: Elaboración propia.

FIGURA 6
DISTRIBUCIÓN DE LA EMPLEABILIDAD SEGÚN LA FORMACIÓN RECIBIDA



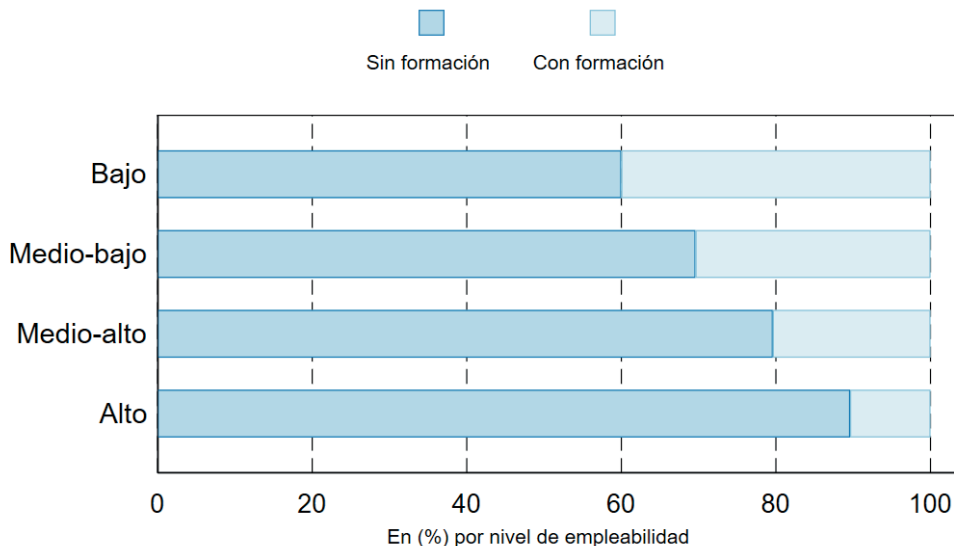
NOTA: Las rectas marcan los límites de la empleabilidad en general.

FUENTE: Elaboración propia en base a datos programa Incorpora.

aportando a aquellos que más lo necesitan. La Figura 7 muestra la distribución de la formación según el nivel de empleabilidad que tienen los beneficiarios antes de recibir dicha formación. Si bien los beneficiarios con una baja empleabilidad que reciben formación representan un porcentaje relativamente alto, también ocurre con aquellos cuyas probabilidades de salir del paro al momento de acceder a Incorpora son relativamente buenas. Esto podría mejorarse, si se enfocara el programa más en aquellos beneficiarios que más necesitan la formación como medio para aumentar su empleabilidad. Y más aun habiendo demostrado que efectivamente los cursos generan ganancias significativas sobre las probabilidades de encontrar un empleo.

En general encontramos, por tanto, efectos positivos de la formación, más importantes para los cursos cortos que para los largos. Los efectos de los primeros son importantes para prácticamente todos los colectivos estudiados. En cambio, los segundos solo encontramos que son significativos para los colectivos más importantes (discapacitados, inmigrantes, 30-45, 45+ y jóvenes).

FIGURA 7
EL EFECTO DE LA FORMACIÓN SEGÚN EL CUARTIL DE EMPLEABILIDAD
Beneficiarios por nivel de empleabilidad inicial y formación percibida en Incorpora



FUENTE: Elaboración propia en base a datos programa Incorpora.

4. Análisis causal

Hasta ahora, el análisis solo permitía comparar los beneficios del programa entre colectivos, pero no estudiar su impacto. Para ello, incorporamos en esta sección una base de datos con la vida laboral de los beneficiarios del programa. Esto nos permite analizar causalmente el efecto de Incorpora a corto y medio plazo, construyendo un grupo de control y cuantificando el efecto promedio entre los tratados.

4.1. Datos

La información para realizar el análisis causal está basada en la Muestra Continua de Vidas Laborales (MCVL)², que es un extracto de datos individuales anonimizados, procedentes de las bases de datos de la Seguridad Social y a los que se añaden otros que se toman del Padrón Continuo Municipal (INE) (García Pérez, 2008). En nuestro caso, esta muestra se refiere a las vidas laborales de una muestra aleatoria de beneficiarios del programa Incorpora y que no forman parte de la MCVL de libre acceso. En el

² www.seg-social.es/wps/portal/wss/internet/EstadisticasPresupuestosEstudios/Estadisticas/EST211

Cuadro 3 se puede observar la distribución de los beneficiarios según el tipo de colectivo, siendo los jóvenes aquel con un mayor peso y las personas adictas con el menor³.

Una vez depurados los datos, nos quedamos con la siguiente información de tratados por el programa Incorpora.

CUADRO 3
FRECUENCIAS SOBRE EL NÚMERO DE BENEFICIARIOS CON DATOS
DE LA SEGURIDAD SOCIAL

Colectivo	Freq.	Percent
Personas adictas o exadictas	28	0,03
Personas con discapacidad	15.535	19,29
Personas con trastorno mental	2.951	3,66
Personas reclusas o exreclusas	1.997	2,48
Víctimas de violencia de género	1.111	1,38
Sin hogar	146	0,18
<i>Prime</i> 30 a 44 (otros)	15.883	19,72
Inmigrantes	12.490	15,51
>45 años	12.187	15,13
Jóvenes	18.203	22,6

FUENTE: Elaboración propia.

Esta información luego es cruzada con un *pool* de Muestras Continuas de Vidas Laborales (2007-2018) donde tenemos información equivalente a la que tenemos sobre tratados de los potenciales controles⁴.

En este punto es conveniente destacar que restringimos el análisis causal a los colectivos que podemos identificar suficientemente bien en la muestra de control. Concretamente, el análisis se realizará para los siguientes colectivos: personas con discapacidad, *prime* (30-44), jóvenes, inmigrantes y mayores de 45 años. Otros colectivos requerirían el cruce con bases de datos adicionales (de salud, penitenciarias, o sociales) para poder recolectar información de población equivalente que no haya participado en el programa. Por esta, y no otra razón, los mantenemos al margen del análisis causal presentado en el presente artículo.

³ Condicionar el análisis a la presencia de los trabajadores en los datos de la Seguridad Social podría generar un sesgo de selección, ya que obvia a aquellos que acceden al mundo laboral gracias al programa y no encuentran empleo (Heckman, 1979). No obstante, no excluimos a estos individuos de la muestra y tomamos en cuenta su vida laboral.

⁴ Es posible que existan beneficiarios del programa Incorpora cuya vida laboral se encuentre en la MCVL. Para excluir estos casos, emparejamos individuos en ambas bases de datos usando sus datos personales y de afiliación. El resultado es que 23 individuos se encuentran en ambas bases de datos y fueron excluidos de la base de datos de la MCVL para evitar duplicados.

4.2. Metodología

Para analizar el efecto causal del programa Incorpora, es necesario crear antes una muestra de control comparable a la muestra de tratados y que permita evaluar el efecto del tratamiento en los tratados (*average treatment effect over the treated*, ATT).

El problema surge en el momento de crear dicha muestra de control. Si todos los beneficiarios se dieran de alta al mismo tiempo, existiría un periodo de tratamiento consistente a partir del cual es posible crear una muestra única. No obstante, las altas en Incorpora se producen de manera sucesiva desde enero del 2011 hasta el día de hoy y, por tanto, resulta necesario crear muestras que tengan características similares no solo inherentes (demográficas), sino también en el momento de evaluarse.

Por ello, nuestro análisis procede en primer lugar a separar por colectivos objeto de estudio y años de alta en el programa a los tratados y a la muestra que en cada caso se usa como control⁵. En el caso de los potenciales controles, se consideran, además, aquellos que cumplan con unas características demográficas previas:

- Discapacitados: personas que tengan un grado de discapacidad.
- Jóvenes: personas que tengan entre 16 y 29 años entre 2011 y 2016, y no sean discapacitadas ni inmigrantes.
- Mayores de 45 años: personas que tengan más de 45 años entre 2011 y 2016, y no sean discapacitadas ni inmigrantes.
- *Prime* (Otros): personas que tengan entre 30 y 44 años entre 2011 y 2016, y no sean discapacitadas ni inmigrantes.

Para cada año, consideramos como punto de referencia el mes previo al alta en Incorpora (tratados) o diciembre del año previo para los controles (como ejemplo, intentaremos crear un grupo de control para los beneficiarios del 2011 con la información de los potenciales controles en diciembre del 2010).

4.2.1. Primera etapa: *propensity score matching*

El primer caso consiste en crear dicho grupo de control a partir de técnicas del *propensity score matching* (Rosenbaum & Rubin, 1983), donde se estima primero el siguiente modelo Probit:

$$\Pr(T1 = 1|Z) = \Phi(Z'\beta)$$

Siendo $T1$ la variable de tratamiento (o participación del programa Incorpora en este caso) y Z las variables de control. Los regresores Z van a condicionar el

⁵ En este trabajo analizamos el efecto promedio del programa sobre aquellos que cumplen dichas características y pueden ser emparejados con la muestra. Un análisis más detallado y separado por colectivos aparecerá publicado próximamente por la Fundación Obra Social La Caixa.

análisis ya que, según la información que usemos, el modelo resultará en un grupo de tratamiento y control distinto. En nuestro criterio base⁶, construimos las regresiones siguiendo lo que a partir de ahora denominaremos «criterio Incorpora», esto es, basaremos la selección en información que los técnicos registran en la base de datos de Incorpora pero que también está en los ficheros de la Seguridad Social. En concreto, las variables que incluimos son:

$$Z_i = edad_alta_i + género_i + educa_i + 1\{prestación_i\} + meses_paro_i$$

Donde $edad_alta_i$ corresponde a la edad de la primera alta laboral, $género_i$ a una dicotómica de hombre o mujer, $educa_i$ a siete niveles de nivel educativo (sin estudios, primaria, secundaria, FP, superior, estudios no reglados y no consta), indicador de si percibió o no prestación por desempleo en el último empleo, y número de meses de desempleo ($meses_paro_i$).

Al mismo tiempo, se añade información adicional dependiendo del tipo de colectivo. Así, para los inmigrantes, la nacionalidad está en el emparejamiento, mientras que para los discapacitados se incluye el grado de la discapacidad.

Para cada año (2011-2015) se estima un modelo con 3 años pre y 3 años postratamiento, considerando únicamente a beneficiarios del programa dados de alta en ese año (ejemplo: para los dados de alta en 2011, consideramos el periodo 2008-2014).

El procedimiento del *propensity score matching* es el siguiente:

En primer lugar, se estima el modelo Probit con las variables de control mencionadas y se hace una predicción de la probabilidad de ser un tratado. Luego, con estos datos se estima una densidad de kernel para representar la distribución de la predicción y se empareja a cada tratado con aquel control cuya probabilidad sea lo más parecida posible usando la siguiente expresión:

$$K_i = K\left(\frac{p(x_i) - p(x_k)}{h_n}\right)$$

Donde h es la banda (*bandwidth*) seleccionada y K es el kernel de Epanechnikov (1969). De esta forma, se obtiene un control con las mismas probabilidades de ser elegido. Un elemento importante de mencionar es que los controles no son excluyentes entre años, es decir, un control para los tratados del año t también puede serlo luego para los tratados del año $t + 1$.

Obtenido el grupo de control para cada colectivo y año de análisis, procedemos a analizar el efecto causal de Incorpora sobre resultados en el mercado de trabajo de los diferentes colectivos considerados en Incorpora.

⁶ Están disponibles, a petición del lector interesado, resultados con criterios alternativos, más estrictos, que validan los resultados obtenidos en el escenario base.

4.2.2. Segunda etapa: análisis de impacto

En esta sección evaluamos causalmente el impacto del tratamiento *Incorpora* en diversas variables de impacto en base al modelo de regresión. Todas las regresiones se estiman por Mínimos Cuadrados Ordinarios⁷ en base a:

$$Y_{it} = \beta_0 + \beta_1 T1\text{ Incorpora}_{it} * Post_{it} + \beta T X_{it} + \alpha_i + \alpha_t + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Donde Y es la variable objeto de análisis, $T1\text{ Incorpora}$ es una variable ficticia que vale 1 para los beneficiarios emparejados en la primera etapa, X son todas las variables de control, α_i son efectos fijos de individuo, α_t son efectos fijos de tiempo, ε el residuo y la unidad de tiempo t es trimestral⁸.

Se plantearon tres preguntas principales que corresponden a variables dependientes (Y) diferentes, o a valores del parámetro β distintos, esto es, a efectos del tratamiento definidos de manera diferente:

1. ¿Cuál es el efecto del programa *Incorpora* sobre la probabilidad de salir del paro?

$$Y_{it} = \begin{cases} 1, & \text{si en } t + 1 \text{ el individuo } i \text{ sale del paro} \\ 0, & \text{si en } t \text{ se encuentra parado} \end{cases}$$

2. ¿Cuál es el efecto del programa *Incorpora* sobre el salario posterior a *Incorpora*?

$$Y_{2i} = \ln(\text{base de cotización})_{it} \text{ del individuo } i \text{ en el trimestre } t$$

3. ¿Cuál es el efecto del programa *Incorpora* sobre el número de meses acumulados en el empleo?

$$Y_{3i} = \text{Meses acumulados en el empleo desde el alta en } \textit{Incorpora} \text{ a } n \text{ trimestres}$$

Para todos los casos, los regresores (X_{it}) se mantienen constantes y son los siguientes:

- Edad de la primera alta.
- Género.

⁷ Los modelos de probabilidad lineal (MPL) se definen como aquellos modelos de decisión discreta donde la estimación se hace a través de Mínimos Cuadrados Ordinarios. Estos modelos tienen ventajas y desventajas: por una parte, la estimación es menos exigente (no es el caso del Probit, donde se estiman los coeficientes por Máxima Verosimilitud) y los coeficientes son muy similares a los estimados en un modelo con distribución normal o logística; por otra, las magnitudes del valor promedio de Y_{it} para casos extremos son difíciles de interpretar ya que son Y_{it} mayores a 1 o menores a 0.

⁸ Este modelo sigue una especificación *Twoway Fixed Effects DD* con variación temporal en el tratamiento (Goodman-Bacon, 2018).

- Nivel educativo en 7 niveles.
- CC. AA. de la última cuenta de cotización.
- Indicador de si percibió o no prestación por desempleo en el último desempleo.
- Indicador de si percibió o no prestación por discapacidad en el último desempleo.
- Sector de actividad económica (CNAE) del último empleo.
- Número de meses parados desde el último empleo.
- Último tipo de contrato.

Mientras que en el modelo que estima el impacto en el salario se añaden, además, los siguientes regresores referidos a la empresa donde se encuentra el empleo:

- Sector del empleo de alta.
- Número de trabajadores en la cuenta de cotización (6 tramos que van desde 1 hasta ≥ 1.500 trabajadores).

4.3. Resultados

En primer lugar, presentamos el número de observaciones resultantes de la primera etapa. El método de emparejamiento de Kernel nos permite asignar más de un control a un tratado como se puede observar en el primer Cuadro 4. Además, el número de controles es, en líneas generales, constante para cada periodo de análisis.

Más interesante aún es el test de medias que elaboramos para analizar posibles diferencias en términos de características observables (Rosenbaum & Rubin, 1985). No existen diferencias importantes entre ambos grupos y para ninguno de los años, lo que parece indicar que el emparejamiento es bueno y permite inferir causalidad ante la presencia de una formación como Incorpora. Cabe recordar que en ningún momento estimamos el efecto promedio del programa (ATE) sino el efecto de Incorpora entre los tratados (ATT), ya que la decisión de participar es voluntaria.

En segundo lugar, una vez realizada la segunda etapa encontramos que el efecto del programa sobre la probabilidad de salir del desempleo es positivo y creciente con el año estudiado (Cuadro 5). Empieza en algo más de cuatro puntos de probabilidad y termina en nueve. Asimismo, son significativos, importantes y crecientes en el tiempo los efectos encontrados sobre la probabilidad de estar empleado a uno, cuatro, ocho y once trimestres vista.

En el promedio también encontramos efectos significativos sobre la base reguladora asociada al empleo encontrado y el número de meses en el empleo acumulados en todos los horizontes analizados. Esto se debe a unos errores estándares relativamente menores respecto al resto de colectivos, lo que supone que la inferencia tenga mayores probabilidades de indicar efectos significativos.

En definitiva, en el corto y medio plazo el programa Incorpora ocasiona un impulso efectivo en la empleabilidad del conjunto de los colectivos estudiados.

CUADRO 4
NÚMERO DE OBSERVACIONES Y CUADRO DE MEDIAS PARA EL PROMEDIO
DE TODOS LOS BENEFICIARIOS ANALIZADOS

	<i>Matching de criterios Incorpora</i>					
	2011	2012	2013	2014	2015	2016
Controles	439.988	439.555	439.700	439.479	439.849	440.355
Tratados	3.173	4.924	7.875	9.117	10.903	14.425
	Medias					
	2011		2012		2013	
	Tratados	Controles	Tratados	Tratados	Controles	Controles
Edad de la primera alta	23,478	30,316	23,355	30,321	22,758	30,315
Últimos meses paro	10,996	11,309	13,864	15,061	16,777	18,827
Mujeres	0,469	0,485	0,499	0,485	0,510	0,485
Sin estudios	0,031	0,133	0,023	0,133	0,022	0,133
Primaria	0,256	0,157	0,239	0,157	0,233	0,156
Secundaria	0,292	0,235	0,289	0,235	0,289	0,235
FP	0,178	0,168	0,222	0,168	0,228	0,168
Superior	0,102	0,125	0,130	0,125	0,146	0,125
Estudios no reglados	0,141	0,180	0,097	0,181	0,083	0,182
	2014		2015		2016	
	Tratados	Controles	Tratados	Controles	Tratados	Controles
Edad de la primera alta	22,766	30,318	22,922	30,304	23,031	30,278
Últimos meses paro	21,425	22,648	27,627	26,464	32,750	30,084
Mujeres	0,502	0,485	0,497	0,485	0,505	0,486
Sin estudios	0,019	0,134	0,024	0,133	0,025	0,133
Primaria	0,232	0,157	0,249	0,157	0,262	0,158
Secundaria	0,308	0,235	0,327	0,236	0,305	0,235
FP	0,231	0,167	0,219	0,168	0,225	0,168
Superior	0,136	0,126	0,121	0,125	0,138	0,125
Estudios no reglados	0,073	0,182	0,060	0,181	0,045	0,181

FUENTE: Elaboración propia.

CUADRO 5
RESULTADOS PRINCIPALES PARA EL PROMEDIO DE TODOS
LOS BENEFICIARIOS ANALIZADOS

Efectos sobre la salida del paro						
	2011	2012	2013	2014	2015	2016
Tratados Incorpora	0,04148*** (0,00372)	0,04924*** (0,00301)	0,06021*** (0,00258)	0,06780*** (0,00253)	0,07426*** (0,00239)	0,09631*** (0,00199)
Observaciones	5.076.284	5.233.094	5.424.347	5.525.972	5.592.003	5.604.019
Efectos sobre los salarios						
	2011	2012	2013	2014	2015	2016
Tratados Incorpora	0,03541 (0,02443)	0,02795 (0,01838)	0,03786** (0,01544)	0,05955*** (0,01513)	0,04881*** (0,01592)	0,07134*** (0,01345)
Observaciones	306.579	292.748	303.907	315.361	328.479	343.453
Efectos sobre meses acumulados en el empleo						
Meses acumulados a 1 trimestre						
	2011	2012	2013	2014	2015	2016
Tratados Incorpora	1,78033*** (0,17712)	1,55174*** (0,14883)	1,67671*** (0,10127)	1,81286*** (0,11342)	1,78590*** (0,10096)	1,97372*** (0,09249)
Observaciones	442.582	443.634	446.425	447.446	449.490	453.562
Meses acumulados a 4 trimestres						
	2011	2012	2013	2014	2015	2016
Tratados Incorpora	3,26881*** (0,59846)	2,25714*** (0,44261)	3,11020*** (0,28008)	3,08167*** (0,33266)	3,13656*** (0,28039)	3,34951*** (0,25066)
Observaciones	442.431	443.317	445.892	446.864	448.708	452.557
Meses acumulados a 8 trimestres						
	2011	2012	2013	2014	2015	2016
Tratados Incorpora	3,87795*** (1,16215)	2,70467*** (0,8273)	4,39270*** (0,5066)	3,74612*** (0,60294)	4,79458*** (0,53755)	4,94722*** (0,46414)
Observaciones	441.546	441.901	443.482	443.981	445.366	448.085
Meses acumulados a 11 trimestres						
	2011	2012	2013	2014	2015	2016
Tratados Incorpora	2,75309* (1,52591)	1,60359 (1,072)	4,27771*** (0,70467)	3,39323*** (0,83186)	4,40315*** (0,76505)	4,86818*** (0,69351)
Observaciones	440.039	439.642	439.921	439.642	440.060	440.612

FUENTE: Elaboración propia.

Los resultados indican que, en el promedio de beneficiarios analizados, el programa Incorpora tiene un efecto positivo sobre la probabilidad de salida del paro una vez dentro del programa, positivo sobre el número de meses acumulados en el empleo y sobre los salarios (aunque este último desaparece cuando se analiza colectivo a colectivo). Asimismo, cuando analizamos los resultados para cada colectivo por separado encontramos efectos algo más importantes para el grupo de inmigrantes (entre un 4,8 % y un 12 %) y para los discapacitados (entre un 5,9 % y un 6,9 %), que para los otros tres grupos poblacionales donde distinguimos por tramos de edad. Para estos últimos grupos el efecto es significativo, pero en algunos casos sustancialmente menor, en torno al 3 %⁹.

Para el conjunto de todos los beneficiarios encontramos, en un contexto donde la empleabilidad media de los que integran el colectivo cae en el tiempo, un efecto sobre la probabilidad de encontrar empleo significativo y bastante estable en el tiempo. En concreto, la empleabilidad aumenta entre 3 y 4 puntos porcentuales, dependiendo del año (Figura 8). En términos relativos, el efecto es muy importante, entre el 20 y 30 % de la probabilidad promedio *ex ante* de los controles dependiendo del año (Figura 8). El efecto sobre la base de cotización mensual es poco significativo y, si acaso, pequeño al final del periodo considerado (Figura 9). Sí es importante, en cambio, el efecto sobre el número de meses acumulados en el empleo, que se sitúa, dependiendo de la cohorte, entre 3 y 4 meses al cabo de tres años (Figura 9).

Las diferencias entre beneficiarios, en términos de empleabilidad previa, pueden ser un factor importante, ya que es posible que Incorpora tenga un mayor efecto sobre personas ya de por sí más aventajadas en el mercado laboral.

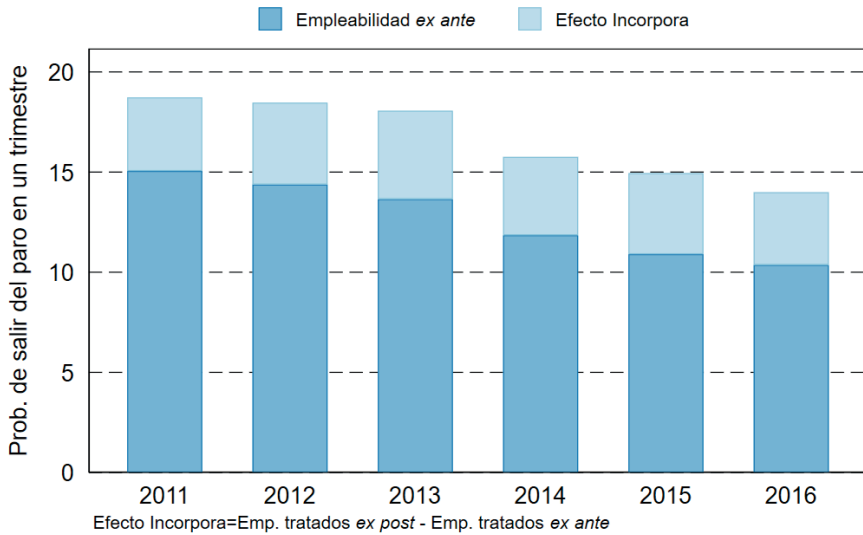
Para comprobar si este es el caso, volvemos a repetir todos estos ejercicios distinguiendo entre tres niveles de empleabilidad previa a la entrada del programa tanto para tratados como para controles. Los resultados de estos nuevos ejercicios (disponibles bajo petición) indican que no existen diferencias en el impacto del programa para aquellos que se dan de alta en Incorpora teniendo un nivel de empleabilidad alto y aquellos que tienen uno bajo. Este efecto es además consistente entre los cinco colectivos analizados.

⁹ Estos resultados, así como los que más adelante se citan en función del nivel previo de empleabilidad, están disponibles, a petición del lector interesado, en la versión completa del informe que da origen a este artículo.

FIGURA 8
SALIDA DEL PARO EN UN TRIMESTRE

Empleabilidad de los tratados Incorpora *ex ante* y *ex post*

Todos los beneficiarios



Todos los beneficiarios

Efecto Incorpora sobre la salida del paro

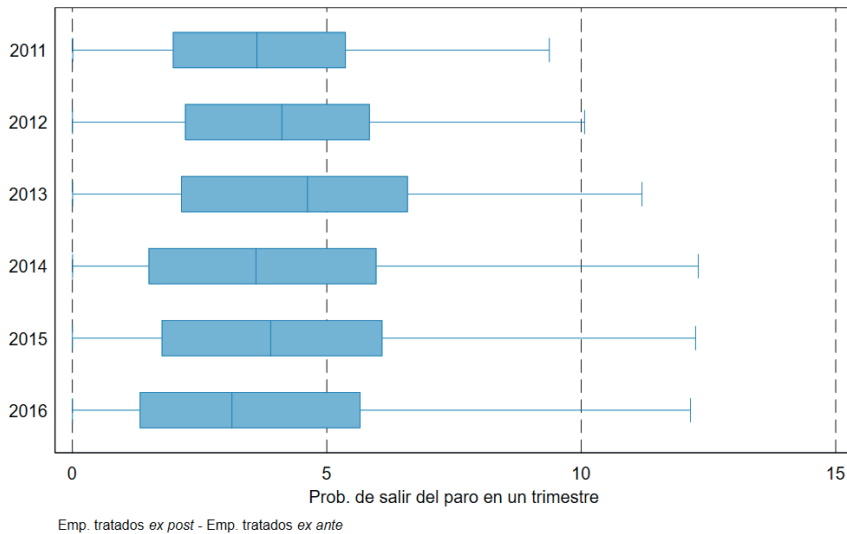
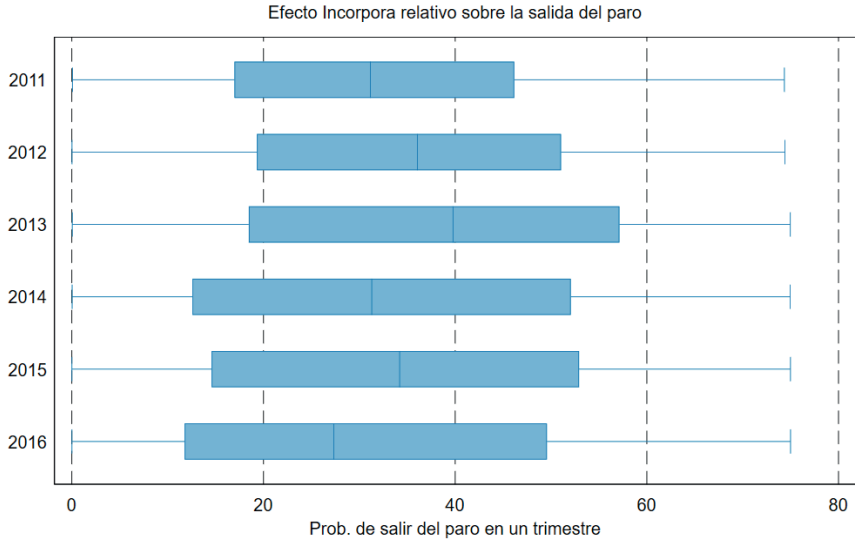


FIGURA 8 (Continuación)
SALIDA DEL PARO EN UN TRIMESTRE

Todos los beneficiarios



FUENTE: Elaboración propia.

FIGURA 9
SALARIOS Y NÚMERO DE MESES ACUMULADOS EN EL EMPLEO

a) Efecto sobre la base de cotización mensual

a1) *Efecto Incorpora sobre la base de cotización mensual*

Todos los beneficiarios

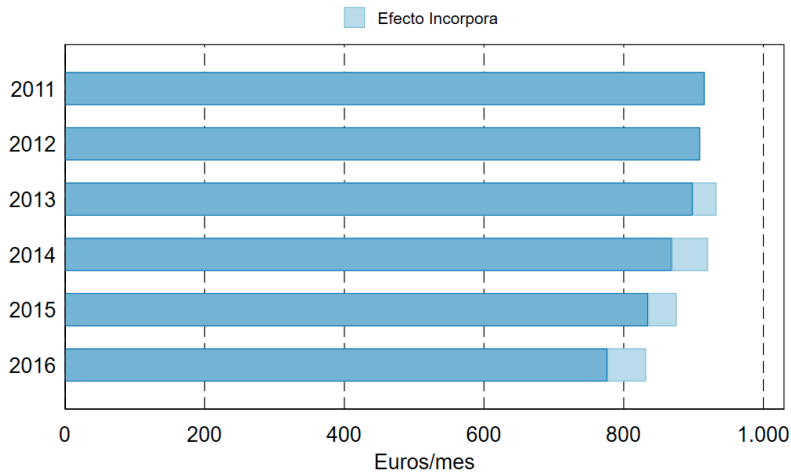
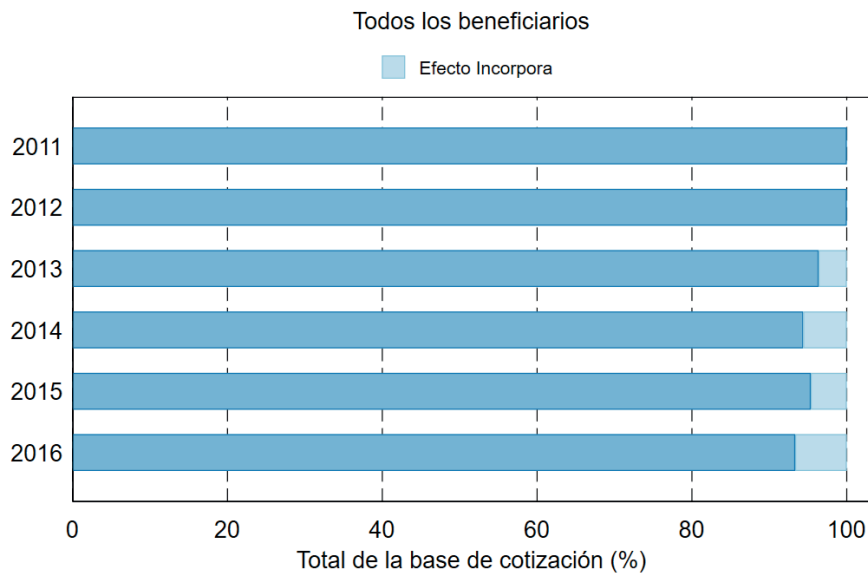
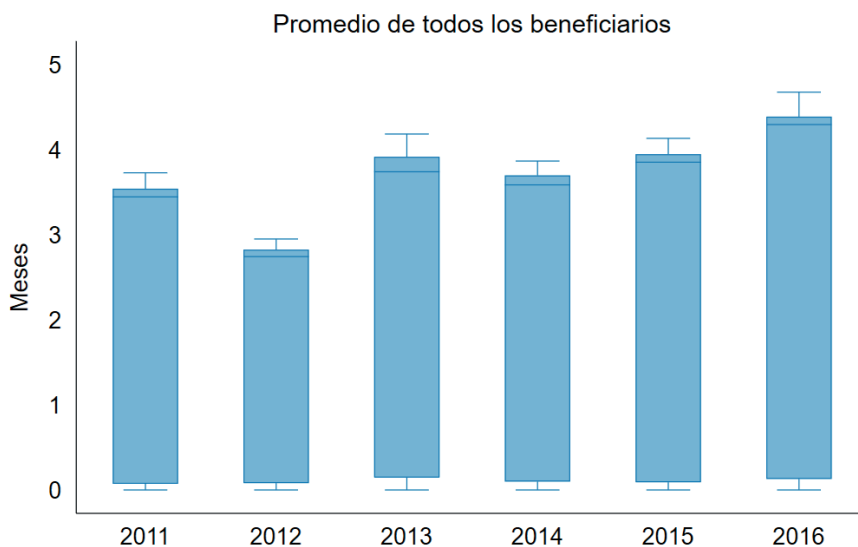


FIGURA 9 (Continuación)
SALARIOS Y NÚMERO DE MESES ACUMULADOS EN EL EMPLEO

a2) Peso relativo del efecto Incorpora sobre la base de cotización mensual



b) Meses acumulados en el empleo



NOTA: Los datos son mes de diferencia respecto a los controles.

FUENTE: Elaboración propia.

5. Conclusiones y recomendaciones

El presente trabajo tiene como objetivo arrojar luz sobre el impacto del programa Incorpora en términos, aunque no exclusivamente, de empleabilidad. Distinguiamos dos partes principales: una evaluación interna, de impacto a corto plazo sobre la empleabilidad, a partir de la información suministrada por los técnicos; y otra, complementaria, evaluando el impacto a medio plazo, en base a la información que proporciona el historial laboral inmediatamente posterior a la participación en el programa, de los participantes en el mismo.

Los resultados indican que Incorpora mejora la empleabilidad de aquellos que acceden al programa. Al mismo tiempo, encontramos que la formación contribuye a una mejora de la empleabilidad, pero que este efecto decae en el tiempo.

No obstante, observamos que la asignación de los cursos no parece destinarse a los beneficiarios más vulnerables dentro de cada uno de los colectivos estudiados. Si bien los beneficiarios con una baja empleabilidad que perciben formación representan un porcentaje relativamente alto, también ocurre con aquellos cuyas probabilidades de salir del paro al momento de acceder a Incorpora son relativamente buenas.

La herramienta de perfilado estadístico usada en la primera parte del artículo, adaptada convenientemente para este propósito, podría contribuir a mejorar esta asignación y enfocar el programa en aquellos beneficiarios que más la necesitan. Una forma de mejorar la asignación es generando una evaluación inicial del beneficiario en el momento del alta, donde se indiquen las competencias a mejorar y las herramientas disponibles que mejores resultados generarían.

Referencias bibliográficas

- Arellano, F. A. (2010). Do training programmes get the unemployed back to work? A look at the Spanish experience. *Revista de Economía Aplicada*, 18(53), 39-65.
- Card, D., Kluve, J., & Weber, A. (2018). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3), 894-931.
- Desiere, S., Langenbucher, K., & Struyven, L. (2019). *Statistical Profiling in Public Employment Services: An international comparison*. OECD Social, Employment and Migration (Working Paper No. 224).
- Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability and Its Applications*, 14(1), 153-158.
- Felgueroso, F., García-Pérez, J. I., & Jiménez-Martín, S. *Perfilado estadístico: un método para diseñar políticas activas de empleo*. FEDEA y Fundación Ramón Areces.
- García-Pérez, J. I. (2008). La muestra continua de vidas laborales: una guía de uso para el análisis de transiciones. *Revista de Economía Aplicada*, 16(1), 5-28.
- Goodman-Bacon, A. (2018). *Difference-in-difference with variation in treatment timing*. NBER, National Bureau of Economic Research (Working Paper Series No. 25018).

- Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1), 153-161. <https://doi.org/10.2307/1912352>
- 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.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician*, 39(1), 33-38.