

Fedea Policy Papers - 2017/10

**Perfilado de parados: una propuesta de herramienta para
los servicios públicos de empleo***

Florentino Felgueroso
José Ignacio García-Pérez
Sergi Jiménez-Martín
Brindusa Anghel

fedea

* Este trabajo forma parte de un proyecto financiado parcialmente por la Fundación Ramón Areces que ha sido posible gracias a los datos aportados por el Ministerio de Empleo y Seguridad Social a través de un convenio con FEDEA. Agradecemos también la valiosa asistencia de Lucía Gorjón y Manu García.

Las opiniones recogidas en este documento son las de sus autores y no coinciden necesariamente con las de FEDEA.

Resumen

En este documento se presenta una propuesta para desarrollar una herramienta estadística que permita realizar un diagnóstico de la empleabilidad de los demandantes registrados en los servicios públicos de empleo (SPE) y facilite la gestión de las políticas activas y la asignación de itinerarios con criterios de eficacia y eficiencia.

Nuestra propuesta consta de tres etapas que por el momento están en distintos estadios de desarrollo. La primera, ya esencialmente operativa, sería un diagnóstico primario que utiliza datos administrativos procedentes de los SPE. A esto esperamos añadir en el futuro un diagnóstico detallado con datos complementarios que se obtendrían a partir de cuestionarios específicos y otras fuentes de información, y una herramienta auxiliar para la asignación de itinerarios que se basaría en los resultados de una evaluación estadística de la efectividad de los servicios ofrecidos por los SPE y otras políticas activas.

El núcleo de la propuesta es un modelo estadístico de perfilado que proporciona un indicador de empleabilidad para cada demandante desempleado y una primera indicación sobre los tipos de acciones que mejorarían dicha empleabilidad. En el momento actual ya se dispone de un modelo beta operativo de perfilado y de algunos resultados preliminares sobre la efectividad de algunas políticas activas de empleo.

El documento consta de cuatro partes. A modo de introducción, la primera parte trata del contexto en el que se desarrolla la propuesta y de sus principales objetivos. La segunda parte consiste en una guía en la que se explican de forma muy accesible los distintos elementos de la propuesta, desde cómo se realiza el perfilado estadístico hasta la evaluación de itinerarios. La tercera parte describe brevemente y con un lenguaje accesible al gran público los principales aspectos metodológicos. Finalmente, en la cuarta parte, se comentan brevemente los resultados a nivel de comunidad autónoma y se ofrecen algunos ejemplos de perfilado individual.

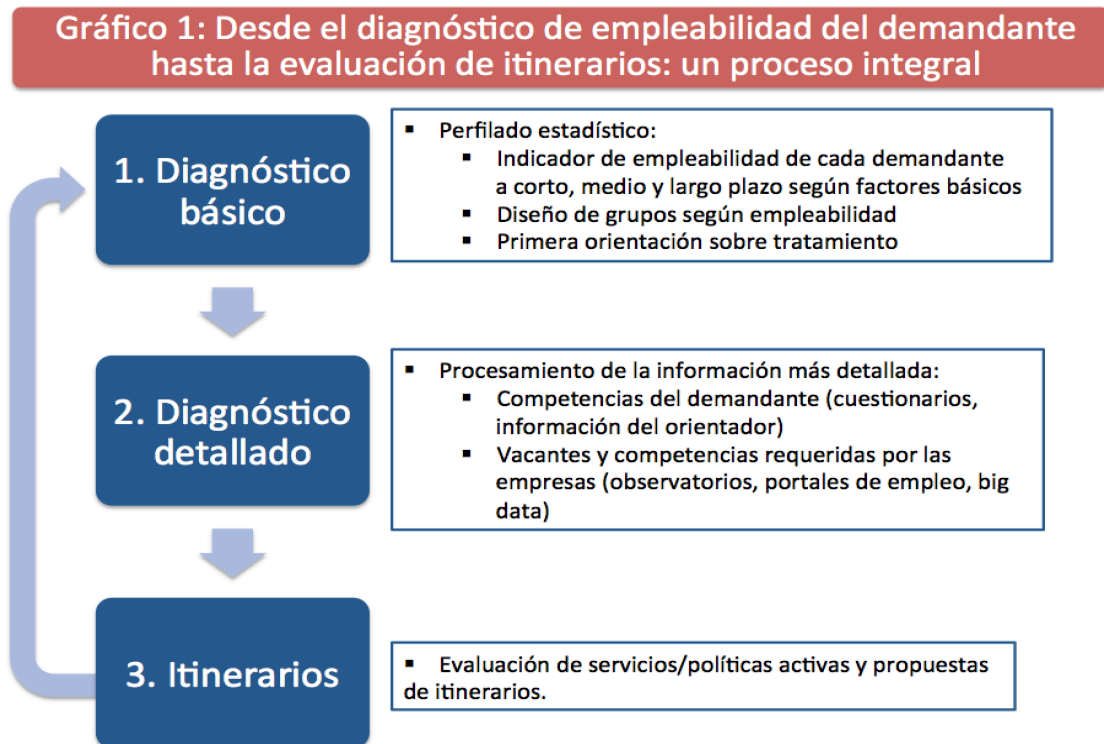
Códigos JEL: J64, J68

Palabras Clave: perfilado estadístico de desempleados, empleabilidad, competencias, evaluación de políticas activas.

1. Introducción

El objetivo principal del perfilado es mejorar la eficiencia de las políticas de empleo mediante la caracterización previa de los desempleados beneficiarios de las mismas. Esto es, se trata de valorar la empleabilidad de estos desempleados, entendida como la probabilidad de transitar a una situación de empleo en un período de tiempo determinado, de clasificar los demandantes de acuerdo con su grado de empleabilidad y de analizar los factores que determinan dicha empleabilidad y cómo podría mejorarse ésta. La propuesta que presentamos aquí consiste en el desarrollo y aplicación de un modelo integral cuyo núcleo es el diagnóstico básico de la empleabilidad de los usuarios de los Servicios Públicos de Empleo (SPE) y que esperamos enriquecer en el futuro a partir de una evaluación de los efectos de distintas políticas activas y de los servicios ofrecidos a los desempleados por los SPE.

Como se muestra en el Gráfico 1, nuestra propuesta consta de tres etapas: un diagnóstico primario que utiliza datos administrativos procedentes de los SPE, un diagnóstico detallado que se realizaría con datos complementarios obtenidos a partir de cuestionarios y otras fuentes de información, y finalmente una herramienta auxiliar para el diseño de itinerarios personalizados que se basaría en los resultados de la evaluación estadística de la eficacia de los servicios ofrecidos por los SPE y de otras políticas activas que buscan mejorar la empleabilidad de los demandantes.



El diagnóstico básico se realiza mediante un perfilado estadístico cuyo objetivo es obtener un indicador de empleabilidad para cada demandante de empleo. Por empleabilidad entendemos la probabilidad de que un demandante transite desde el desempleo al empleo antes de que transcurra un tiempo determinado, teniendo en cuenta sus características sociodemográficas, su historial laboral, sus competencias y la evolución del mercado de trabajo. Con este indicador se puede clasificar a los demandantes de empleo en distintos grupos de menor a mayor grado de empleabilidad. Esta clasificación por grupos tiene como objetivo proporcionar información sobre el grado o intensidad del apoyo que necesitará cada demandante para conseguir un empleo.

Una primera orientación sobre el tipo de servicios u otras políticas activas que mejorarían la empleabilidad de un demandante concreto se obtiene a partir de los efectos marginales sobre la probabilidad de empleo de distintas variables explicativas que se pueden asociar con políticas activas específicas. Por ejemplo, la formación complementaria puede ser una de las variables que permiten explicar el nivel de empleabilidad de un demandante. En efecto, una política activa que mejore la formación podría generar un efecto positivo sobre la empleabilidad de esta persona. Así, por tanto, cuando el modelo predice una mejora significativa en la empleabilidad por medio de un cambio en una o varias de las variables modificables incluidas en el modelo y relacionadas con este tipo de formación, se procede a recomendar una política o una combinación de políticas activas que mejore dicho nivel de formación complementaria.

Una vez realizado el diagnóstico primario se procederá un diagnóstico detallado para el que se incorporará información más específica sobre las competencias del desempleado, así como sobre el estado del mercado laboral en el que esta persona está buscando empleo. Esta información no está disponible en las bases de datos de tipo administrativo que hemos manejado hasta el momento. Las fuentes de datos necesarias para esta segunda etapa serían cuestionarios ad-hoc que profundicen en las habilidades cognitivas y no cognitivas del demandante, así como información procedente de otras fuentes tales como portales de empleo, observatorios ocupacionales y redes sociales.

El menú de itinerarios que eventualmente ayudaría a elaborar la herramienta propuesta en este informe con el fin de ayudar a los orientadores a prestar una atención individualizada, se apoyaría no solo en la información recabada en las fases anteriores del modelo, sino también en el análisis del impacto de las políticas activas que gestionan las Comunidades Autónomas (CCAA). La evaluación de las políticas activas es clave para identificar cuáles son aquellas que tienen un mayor

impacto sobre la mejora de la empleabilidad. El análisis distinguirá entre los efectos observados durante la aplicación de las políticas y aquéllos que se produzcan después de su finalización. Esta distinción es necesaria dado que algunas políticas activas pueden tener efectos negativos sobre la salida del paro al empleo durante su aplicación, por reducir la intensidad de la búsqueda de empleo.

La metodología científica para la evaluación de políticas activas es estándar y por lo tanto utilizada ampliamente en países donde están más avanzados en este tipo de herramientas. Concretamente, para evaluar el impacto de una política activa se compara la evolución de la empleabilidad de los beneficiarios de las políticas activas (el llamado *grupo de tratamiento*) con un *grupo de control* integrado por no beneficiarios con características similares a las de los tratados. La metodología exige también controlar por los posibles efectos de selección muestral, tanto en el grupo de beneficiarios como en el grupo de control. Por último, estas evaluaciones de política utilizan los mismos datos administrativos que se usan para estimar el modelo de perfilado tras cruzarlos con los ficheros de servicios de los SPE. Una evaluación más avanzada exigirá, además, poder cruzar dichos datos con otras fuentes de información tales como los historiales de afiliación de la Seguridad Social.

En este documento se describe el estado actual de desarrollo de los distintos componentes del modelo propuesto a marzo de 2017. Se trata aún en buena parte de un modelo en construcción, o modelo beta, en el que la etapa de diagnóstico primario con los datos de registros administrativos se encuentra bastante avanzada. Los avances que hemos realizado en la tercera etapa, de evaluación de políticas activas, se presentan en otro documento (ver García-Pérez, 2017). Ambos documentos se irán completando con los avances que se vayan realizando en los próximos meses.

El resto del presente documento se ha redactado para facilitar una lectura fluida, evitando los aspectos más técnicos que se presentarán en un documento aparte. En la segunda parte se describe el procedimiento completo de perfilado y evaluación con un lenguaje no técnico que esperamos resulte accesible a un lector no especializado que carezca de conocimientos avanzados de estadística. En la tercera parte, se presenta brevemente los aspectos metodológicos de la propuesta. En la cuarta parte, se comentan brevemente los resultados del modelo de perfilado a nivel nacional y se presentan algunos ejemplos de evaluación individual a nivel de comunidad autónoma. El documento también incluye una serie de anexos con definiciones y análisis descriptivos más completos.

2. Perfilado y evaluación: una guía para no especialistas

En esta sección del trabajo se resumen de forma no técnica los principales componentes de la herramienta propuesta, comenzando con el perfilado y terminando con el diseño de itinerarios personalizados para los parados.

2.1. Diagnóstico básico

2.1.1 ¿En qué consiste el perfilado estadístico?

El perfilado (*profiling* en inglés) hace referencia a la clasificación de los desempleados en distintos grupos con el fin de ofrecerles los servicios que mejor se adaptan a sus necesidades. Entre los modelos que se utilizan para realizar esta clasificación se encuentran los modelos de perfilado estadístico, en los que se utilizan técnicas estadísticas para estimar algún indicador de empleabilidad o de riesgo de permanecer en el paro que a su vez sirve para clasificar a los demandantes.

La principal ventaja de este tipo de perfilado es que los criterios de clasificación no se limitan a un pequeño número de variables elegidas de forma ad-hoc, como pueden ser la edad y/o la duración del paro, que son los indicadores que se han utilizado habitualmente en nuestro país. Por el contrario, este tipo de modelos permite incorporar un gran número de variables relevantes y las pondera en base a su impacto sobre la empleabilidad de los demandantes de empleo tal como se observa en los datos.

El perfilado estadístico se ha implando en muchos países, desde EEUU a Australia, pasando por los países escandinavos, Irlanda, Alemania y muchos otros países centro-europeos y, más recientemente, de Europa Oriental. En muchos de ellos, el perfilado se concibe como una herramienta complementaria de apoyo a los orientadores y no como un sustituto para ellos. Éste es también el espíritu de nuestra propuesta.

2.1.2. Medición de la empleabilidad individual

Tradicionalmente, los modelos de perfilado estadístico parten de la estimación de la probabilidad de convertirse en parados de larga duración, que es el riesgo que se

pretende evitar. En España, sin embargo, la implantación del modelo de perfilado se produciría en un momento en el que más de la mitad de los desempleados ya son parados de larga duración. Resulta por lo tanto necesario considerar una medida más genérica que pueda calcularse o aplicarse a todos los demandantes y en cualquier momento, y no sólo en el de su inscripción como parados.

El objetivo de esta propuesta es estimar la probabilidad de salir del paro hacia el empleo en un período determinado para todos los demandantes parados, independientemente del tiempo transcurrido desde su inscripción. Esta estimación se realiza utilizando técnicas estadísticas desarrolladas para modelos de duración en tiempo discreto y se repite con horizontes de 3, 6 y 12 meses, que identificaremos, respectivamente, con el corto, medio y largo plazo.

Esta probabilidad estimada de salir del desempleo será nuestro indicador de empleabilidad para cada demandante. La estimación se realiza utilizando un modelo estadístico sencillo que considera que la probabilidad de salida del desempleo al empleo es una función de las características personales del demandante, de sus competencias y experiencia laboral, del tiempo que lleva desempleado y de la evolución prevista del mercado de trabajo al que se dirige.

Los parámetros del modelo, que cuantifican la influencia de las distintas variables explicativas sobre la empleabilidad (o probabilidad de salir del desempleo) del demandante, se estiman utilizando de forma exhaustiva la información disponible en una amplia muestra de historiales laborales sobre episodios de salidas del paro al empleo. En esta versión beta, los parámetros del modelo de perfilado se estiman a varios niveles para cada comunidad autónoma.

Los datos

Los datos que se utilizan para estimar el modelo son los datos administrativos individuales registrados en el Sistema de Información de los Servicios Públicos de Empleo (SISPE) que sigue un procedimiento de registro común para todas las CCAA. Para el diagnóstico primario de la empleabilidad de los trabajadores, se utilizan los ficheros de Demanda, de Prestaciones y de Contratos registrados en los SPE. Para la evaluación de los servicios ofrecidos por los SPE también se utilizarán los microdatos procedentes de los Historiales de Servicios recibidos por los demandantes.

Los ficheros de Demandas contienen el universo de los registros administrativos mensuales de demandas de intermediación y otros servicios. Incluyen las características de todos los demandantes así como su situación laboral y administrativa a finales de mes. A partir de los microdatos de estos ficheros se obtiene, por tanto, la situación laboral a fin de mes de los demandantes 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. A partir de estos ficheros se construye el fichero principal o *fichero base* que se utiliza para estimar el modelo. El formato impuesto por estos ficheros, que reflejan valores a finales de cada mes, es el formato al que se adaptarán los *ficheros auxiliares* que se unirán al principal.

Los ficheros de Contratos contienen el universo de todos los contratos registrados cada mes. Contienen información sobre las características de los contratos registrados, incluyendo sus fechas de inicio y finalización, el tipo de contrato y ocupación, el sector de actividad de la empresa contratante, etc. La información contenida en estos ficheros se cruza con el fichero base para tener en cuenta la situación relativa de la demanda de trabajo en el ámbito de búsqueda de empleo del demandante (nacional, autonómico, provincial y municipal) para las distintas ocupaciones solicitadas por los demandantes.

Los ficheros de Prestaciones contienen información sobre los Beneficiarios y/o Perceptores de Prestaciones en cada mes y los movimientos de alta, renovación y baja. Esta información se une al fichero base para disponer de información más detallada sobre las prestaciones que ha percibido el demandante.

Finalmente, a efectos de la evaluación de políticas activas, el fichero base se fusionará también con el fichero de Historiales de *Servicios*, para así tener en cuenta las actuaciones de los SPE sobre cada demandante de empleo que pueden haber tenido un efecto sobre su empleabilidad. Estos ficheros incluyen información sobre los servicios recibidos (por ejemplo, orientación e información profesional, tutoría individual, búsqueda activa de empleo, etc.) a lo largo de toda la vida laboral del individuo, las fechas de inicio y de fin de dichos servicios, el número de horas y las vías de financiación de los mismos (fondos del Estado transferidos a las CCAA, fondos propios de las CCAA, fondos estatales no transferidos o vía de igualación).

Todos los ficheros incluyen el universo de individuos registrados en los SPE, excepto los de Historiales de *Servicios*, que proceden de una extracción realizada a una amplia muestra aleatoria de demandantes (1,5 millones de personas que han

sido demandantes en alta en algún momento del período de enero de 2011 a junio de 2015).

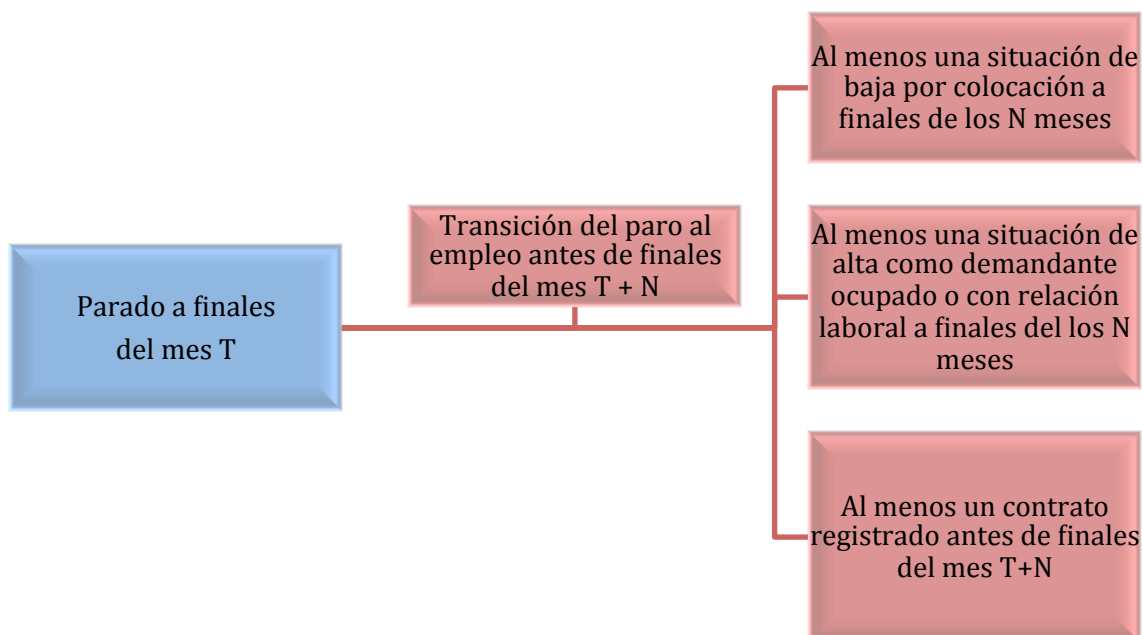
En la versión beta del modelo cuyos resultados se presentan en este documento, los parámetros del modelo son estimados con los datos procedentes de estos ficheros para el período comprendido entre septiembre de 2014 y octubre de 2016. La muestra utilizada está constituida por todos los demandantes de 16 a 64 años que han estado en situación de paro en alguno de los fines de mes de este período.

La salida del desempleo al empleo

La variable de referencia que se utiliza para medir la empleabilidad es si la persona desempleada ha salido de esta situación para incorporarse a una situación de empleo en un período determinado.

A efectos del perfilado se considerarán como parados todos aquellos demandantes que no tienen ninguna relación laboral ni están autoempleados a finales de cada mes (véase el Anexo 1 para un listado detallado de los grupos de demandantes que incluye y excluye esta definición).

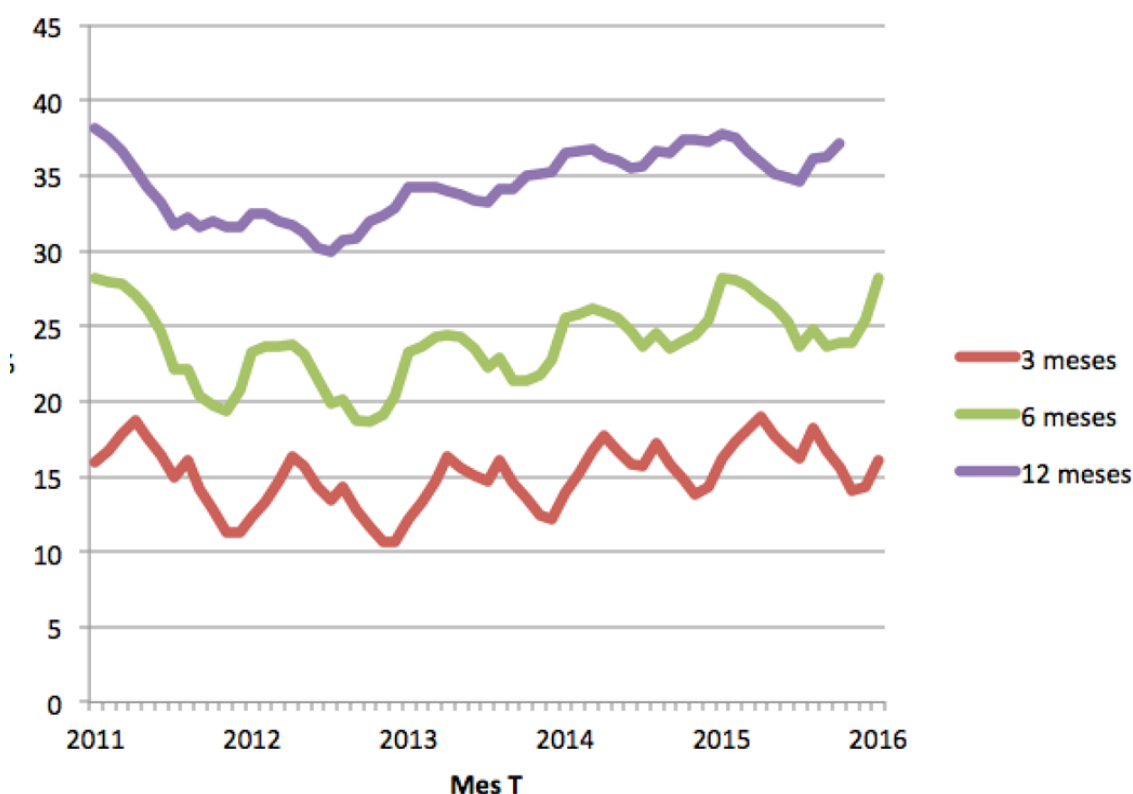
Gráfico 2: Transición del paro al empleo antes de N meses



Para detectar aquellos demandantes parados que consiguen emplearse antes de que trascurren N meses seguimos el procedimiento que se muestra en el Gráfico 2.

Definimos aquí, por tanto, las transiciones del paro al empleo como aquéllas que realizan los demandantes que, estando en situación de paro a finales de un determinado mes, hayan sido dados de baja por colocación, o hayan pasado a considerarse demandantes ocupados o con una relación laboral a finales de cualquiera de los meses de un período determinado, y/o hayan firmado algún contrato laboral durante este período. De esta forma, definimos la tasa de salida del paro al empleo antes de N meses como la proporción de personas paradas a finales del mes T que hayan conseguido un empleo antes de finales del mes $T+N$.

Gráfico 3: Tasas de salida del paro al empleo antes de 3,6 y 12 meses (Parados a finales del mes T que consiguen un empleo en los 3, 6, o 12 meses siguientes, 2011.1-2016.7)



En el Gráfico 3 se muestra la evolución de las tasas de salida del paro al empleo, en el período 2011-2016, para valores de N iguales a 3, 6 y 12 meses para toda la población de demandantes parados en los meses que aparecen en el eje de abscisas. Por ejemplo, el punto de la línea roja correspondiente a enero del 2011 indica que un 16,5% de los demandantes parados a finales de enero del 2011 transitaron hacia el empleo en algún momento de febrero del 2011, mientras que el punto correspondiente a enero del 2011 en la línea morada indica que cerca del

38% de los desempleados a finales de este mes realizaron una transición hacia el empleo antes de que se cumplieran 12 meses.

Las tres tasas de salida recogen el final de la fase recesiva y el inicio de la fase alcista. Las tasas de salida para valores mayores de N , es decir, para horizontes más largos, recogen el inicio de la fase expansiva con anterioridad, y, además, tienen una menor variabilidad debido a que están menos afectadas por la estacionalidad.

Las variables de control

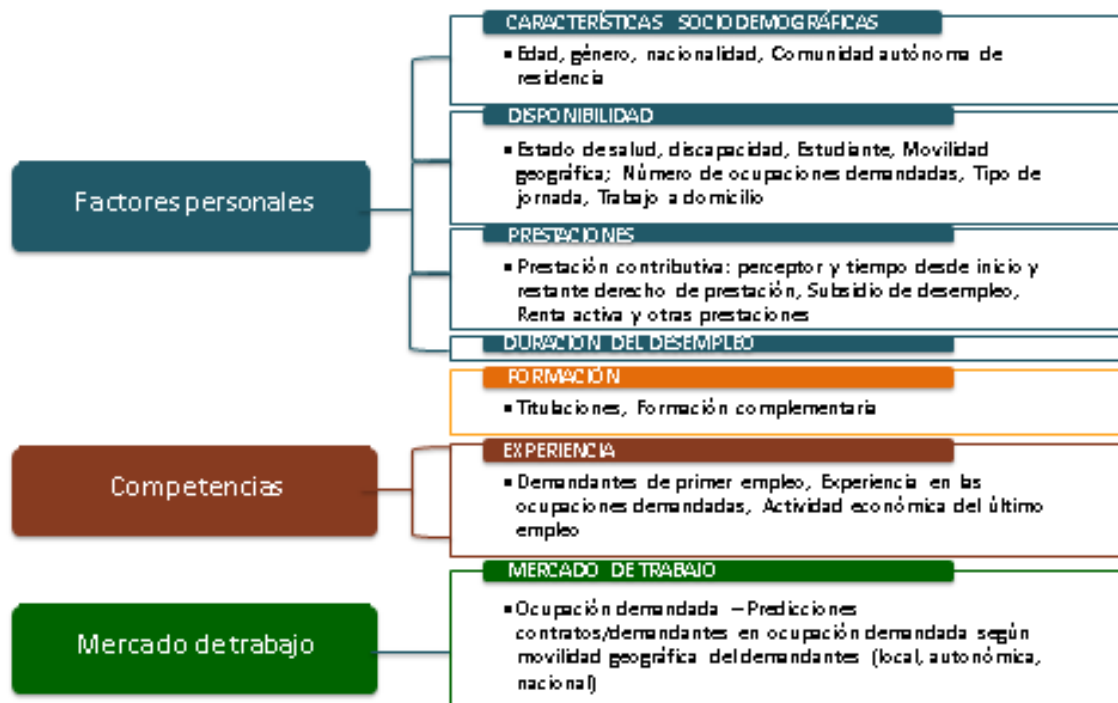
El objetivo de esta sección es describir las variables de control que se utilizan en el modelo de perfilado estadístico. Estas variables son aquellas que permiten explicar las diferencias de las tasas de salida del paro al empleo entre distintos parados y también su evolución a lo largo del tiempo. Los valores de estas variables permiten además estimar las salidas del paro al empleo y, con ello, clasificar a los parados en función de su mayor o menor dificultad para encontrar un empleo en el corto, medio y largo plazo.

Las variables de control o factores explicativos utilizados en los modelos de perfilado estadístico se suelen clasificar en distintos grupos que responden al enfoque que se elija en cada caso. Por ejemplo, en OCDE (2016) se sigue un enfoque esencialmente económico y se distingue entre los factores “capacitantes” (educación, experiencia laboral, problemas de salud, responsabilidades familiares o dependencia en el cuidados de niños y ancianos), los factores de incentivos (la tasa de reemplazamiento, la fiscalidad laboral, las prestaciones y las rentas no laborales) y los factores que determinan las oportunidades de empleo (la evolución cíclica del mercado de trabajo, las limitaciones a la contratación en el mercado de trabajo relevante o el nivel educativo). Otros modelos utilizan otros enfoques más acordes con la literatura sociológica.

En el presente estudio se ha optado por una clasificación más neutral, orientada a la predicción más que a la búsqueda de causalidad y acorde con los datos disponibles en los ficheros de los SPE. Hemos de tener en cuenta además dos limitaciones importantes. Primero, las personas registradas en los SPE sólo representan una parte de las personas que desean trabajar, y existen además potenciales e importantes sesgos de selección. No todos los parados solicitan la intermediación de los SPE, y la decisión de inscribirse en estos depende también de qué otros servicios puedan ofrecer (como la participación en políticas activas) y

de los requisitos para la percepción de prestaciones y otro tipo de ayudas. De ahí que tanto los estadísticos descriptivos de las variables de control como los resultados obtenidos en la estimación del modelo de perfilado sólo son válidos para este colectivo de personas registradas y han de ser interpretados teniendo en cuenta los posibles sesgos de selección. En segundo lugar, algunas variables importantes no están disponibles en los ficheros de demanda y contratos. Entre éstas cabe citar, por ejemplo la situación familiar y el cuidado de niños o personas de edad avanzada. Un objetivo de la segunda fase del desarrollo de esta herramienta este tipo de información.

Gráfico 4: Variables de control para el modelo de perfilado



Las variables de control del modelo de perfilado se pueden clasificar en siete grupos: las características sociodemográficas, la formación, la experiencia, la duración del paro, las prestaciones percibidas, la disponibilidad y la situación del mercado de trabajo al que se dirige el demandante parado. Las variables incluidas en cada grupo vienen detalladas en el Gráfico 4. Un análisis descriptivo detallado de estas variables se presenta en la sección 4.1.3.

2.1.3. Clasificación por niveles de empleabilidad

Una vez estimados los parámetros del modelo, estos se usan conjuntamente con los valores observados de las distintas variables explicativas para calcular la probabilidad de salida del paro al empleo de cada demandante antes de N meses. En esta versión del modelo, la estimación de los parámetros se realiza con datos del período de septiembre del 2014 a octubre del 2016, mientras que las probabilidades de salida hacia el empleo se calculan para los demandantes parados a 31 de octubre de 2016.

Nótese que el modelo se puede estimar a nivel nacional, a nivel regional e incluso, para un grupo de población determinado, todo ello dependiendo de los objetivos del análisis y del grado de especificidad que se quiera para las estimaciones de los parámetros. Por ejemplo, si el objetivo se centra en una sola comunidad o provincia el modelo adecuado sería el estimado con la muestra a nivel regional o provincial.

Una vez estimada la probabilidad de salida del desempleo al empleo a corto, medio y largo plazo para cada uno de los parados registrados en los SPE, este indicador se utiliza para clasificar a los parados en diversos grupos. Con ello se pretende facilitar la identificación de las personas con niveles de empleabilidad similares de cara a ofrecerles los servicios que mejor se adaptan a sus necesidades.

Entre los diferentes modelos de perfilado estadístico internacionales los criterios utilizados para la delimitación de los grupos suelen fijarse de forma ad-hoc, especificando valores determinados para el riesgo de permanencia en el desempleo. La clasificación que aquí proponemos se basa en los cuartiles de las distribuciones de probabilidades de salida del paro antes de 3, 6 o 12 meses.

En concreto, los demandantes parados se clasifican en 4 grupos, de mayor a menor empleabilidad, con las denominaciones que se muestran en el Gráfico 5.

Gráfico 5: Clasificación por niveles de empleabilidad



La asignación de cada individuo a uno de los grupos se realiza de forma excluyente utilizando las probabilidades estimadas de salida del paro a corto, medio y largo plazo. La utilización de sólo una de estas distribuciones supondría desaprovechar parte de la información disponible. Por ejemplo, si nos centramos exclusivamente en las probabilidades de salida antes de 3 meses, no distinguiríamos entre aquéllos que aun teniendo una baja probabilidad de salir en el corto plazo, pueden tener una probabilidad alta o media de encontrar empleo en el medio o largo plazo. De la misma forma, si sólo nos fijamos en el largo plazo no distinguiremos correctamente entre aquellos que pueden encontrar trabajo con mayor o menor rapidez. Por tanto, para aprovechar al máximo la información disponible la clasificación que se propone sería la siguiente:

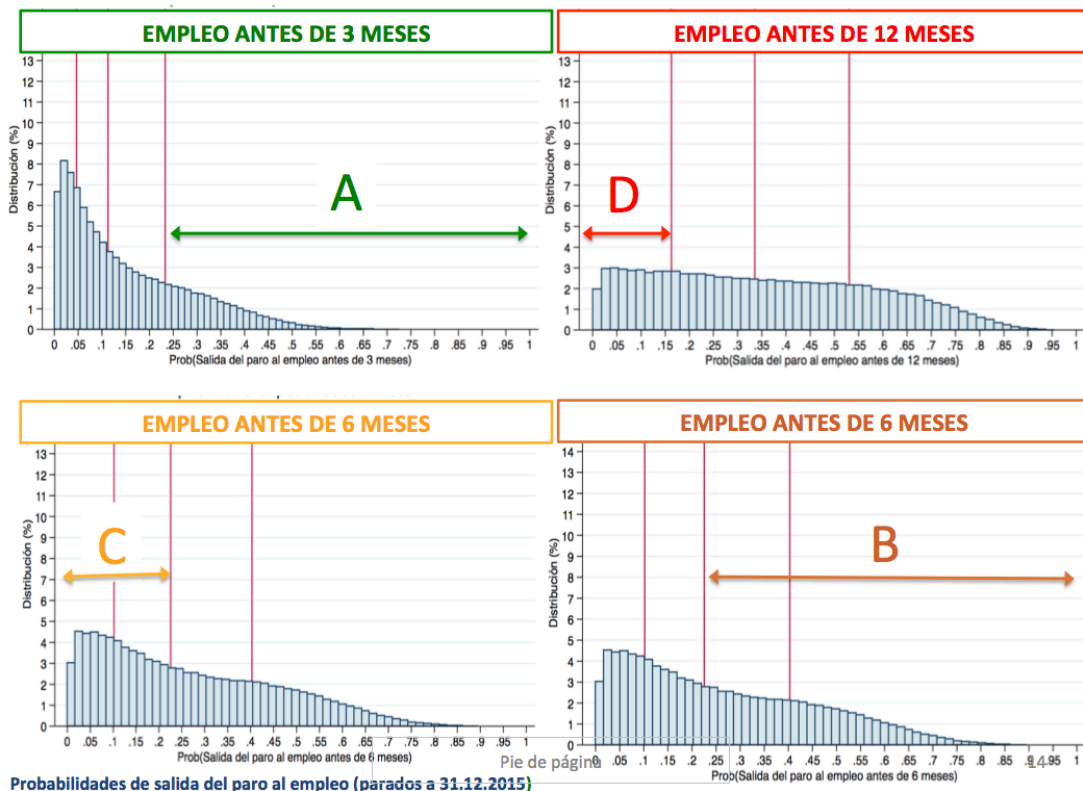
- En el **grupo A** estarían las personas con mayor nivel de empleabilidad en el corto plazo, definidos como **aquéllos cuya probabilidad de salida antes de tres meses se encuentra en el cuartil superior de la distribución de probabilidades de salida en menos de 3 meses.**
- Los parados asignados al **grupo D** son aquellos con las probabilidades más bajas de salida antes de 12 meses, esto es **aquéllos que se encuentran en el primer cuartil de la distribución de estas probabilidades.** En la

práctica no observamos personas que se puedan encontrar en ambos grupos de forma simultánea.

- Una vez clasificado a las personas pertenecientes a los grupos A (con mayor probabilidad de salida rápida o de corto plazo) y D (los que tienen una menor probabilidad de salida en el largo plazo), el resto de personas se clasifican en los grupos B y C. Los parados del **grupo B** son aquellos **individuos no integrados en los grupos A y C cuya probabilidad de salida en el medio plazo se sitúa por debajo de la mediana de la distribución de probabilidad de salida a seis meses**, mientras que los del **grupo C** serían aquellos que tienen una probabilidad superior a la mediana de la mencionada distribución.

El Gráfico 6 muestra cómo se distribuyen los desempleados en base a los resultados obtenidos para los parados a finales de octubre del año 2016.

Gráfico 6: Clasificación de los parados a 31 de octubre de 2016 por grupos de empleabilidad.



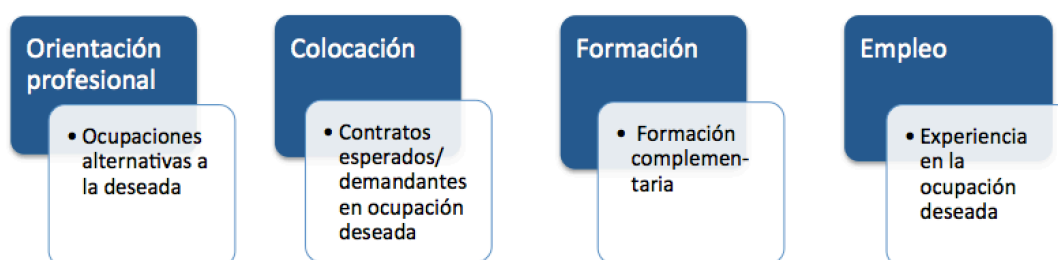
2.1.4 Orientación para para el tratamiento de desempleados

En esta fase, se completa el diagnóstico básico de empleabilidad mediante la identificación del conjunto de características modificables que el demandante no posee y que más podrían mejorar su nivel de empleabilidad. Se trata, pues, de obtener una primera orientación acerca de cuáles de los grandes grupos de servicios ofrecidos por los servicios públicos de empleo podrían ayudar más a mejorar la empleabilidad de un demandante determinado.

Esta primera orientación se obtiene utilizando el modelo estimado para identificar, en base a sus efectos marginales sobre la empleabilidad de los individuos, a aquellas variables que, en alguna medida, pueden modificarse a través de distintos tipos de políticas activas y que más podrían contribuir a mejorar su situación. Por ejemplo, el modelo permite valorar para un individuo sin experiencia, cuál sería el cambio en la empleabilidad que produciría adquirir un determinado nivel de formación.

Más en detalle (véase el Gráfico 7), consideramos cuatro grandes grupos de políticas activas que potencialmente puede afectar a la empleabilidad: i) los servicios de orientación profesional; ii) los servicios de colocación; iii) la formación; y iv) el fomento directo del empleo mediante medidas tales como subvenciones o la oferta de un empleo público. A cada uno de estos grupos, a modo de ejemplo, le asociamos una variable del modelo como sigue:

Gráfico 7: Primera orientación sobre los grupos de servicios que mejorarían la empleabilidad del demandante



En la sección 4, presentaremos los detalles específicos de cómo se obtiene esta primera orientación utilizando algunos ejemplos concretos.

3. Metodología para el perfilado estadístico y datos

El perfilado estadístico que se propone en el presente documento parte de un modelo estadístico que permite estimar la probabilidad de transitar del desempleo al empleo en un período determinado. El objetivo de este modelo es responder a la pregunta de cuál es la probabilidad de que un demandante inscrito en los registros de los SPE encuentre empleo a corto, medio y largo plazo, concretamente antes de 3, 6 y 12 meses. Ello permite clasificar a los individuos en función de su grado de empleabilidad.

Adicionalmente, el modelo permite estimar la contribución (o efecto marginal) de cada factor (o variable de control) a la probabilidad de salida del paro al empleo, para el conjunto de parados o para un grupo poblacional concreto. De esta forma, el modelo permite analizar el efecto sobre la empleabilidad que tendría un cambio en las variables de control de interés, como puede ser una mejora de las competencias del parado o de su disponibilidad.

El modelo pertenece a la familia de los modelos estadísticos de duración. Estos modelos son la herramienta natural para analizar la problemática del desempleado que busca empleo y que tiene una probabilidad de encontrarlo que no es constante a lo largo del tiempo, sino que depende de manera crucial del tiempo que lleve en el desempleo.²

La alternativa tradicional en los países en los que se implantaron modelos de perfilado estadístico ha sido la estimación de un modelo de elección discreta para la probabilidad de alcanzar un número concreto de meses o de semanas en situación de paro. Con este tipo de modelo, se obtiene un indicador del riesgo de que un demandante recientemente inscrito se convierta en parado de larga duración o de agotar las prestaciones por desempleo.³

² Ver Mortensen (1986) o van den Berg (2001), por ejemplo, para la caracterización de estos modelos y para la justificación de su optimalidad de cara al estudio del suceso objeto de estudio en este informe.

³ Para una comparación internacional de modelos de perfilado estadístico, véase Rebollo (2017). Ejemplos de países que estiman el riesgo de permanecer desempleado pasadas N semanas o meses son Estados-Unidos, Australia, Irlanda o Finlandia. En este sentido, nuestra propuesta tiene una mayor afinidad con los que se aplican o se aplicaron en países como Dinamarca o la República Checa.

El modelo básico con el que se desarrolla el perfilado es un modelo econométrico de elección discreta, concretamente, un modelo PROBIT. La variable dependiente es una variable binaria que adopta un valor de 1 si el demandante parado transita del desempleo al empleo antes del periodo T y de 0 en caso contrario. (T pueden ser 3, 6 o 12 meses). La expresión analítica del modelo a estimar en este caso es:

$$\Pr[\text{Salida al Empleo en el periodo } T | t, x] = F(\alpha't + x'\beta + \varepsilon) \text{ con } \varepsilon \sim N(0,1),$$

donde F denota una función de probabilidad y el término de error, ε , tiene una distribución normal estándar de varianza normalizada a 1. t hace referencia a la duración en el desempleo y x recoge el resto de variables de control, obtenidas de los registros del SPE.

Una vez realizada la estimación de dicho modelo, tal y como ya comentamos en la sección 2, éste puede utilizarse para clasificar a los demandantes en base a su probabilidad estimada de salida. La clasificación se basa en la predicción obtenida dado un determinado horizonte temporal. Esto es, a corto plazo (3 meses), medio (6 meses) y largo plazo (12 meses).

Sin embargo, el mayor interés del análisis de empleabilidad es que permite, tal y como ya apuntamos en la sección 2, por un lado, analizar la posición relativa de un demandante respecto a la distribución de su ámbito de referencia, por ejemplo, su comunidad de residencia; y por otro, evaluar, como cambia el grado de empleabilidad de dicho demandante cuando se modifican, a través de acciones o itinerarios formativos, algunas características asociadas al demandante.

Este proceso de primer diagnóstico individual de (potencialmente) cada demandante específico, conlleva las siguientes etapas:

1. Estimación del modelo para la región en la que vive el demandante para calcular grupos de empleabilidad de referencia.
2. División de la muestra total en celdas representativas definidas por combinaciones de características de la población, léase, género, nivel de educación y grupo de edad. Así por ejemplo, si consideramos cuatro niveles de educación y tres grupos de edad, este proceso comportaría 24 celdas en una región dada. De hecho para muestras muy grandes podríamos añadir características adicionales a la definición de celdas.

3. Estimación del modelo en cada una de las celdas predefinidas para calcular parámetros específicos para cada subgrupo de población.
4. Cálculo del grado de empleabilidad del demandante con los parámetros de la celda que le corresponde.
5. Evaluación del cambio del grado de empleabilidad como consecuencia del cambio de características específicas, representativas de las grandes tipos de orientación profesional que consideramos.

En la siguiente sección veremos dos ejemplos detallados de este procedimiento.

3.1. La muestra de estimación

Los resultados que presentamos aquí utilizan los datos del SISPE para todas las CCAA en el período comprendido entre septiembre de 2014 y octubre de 2016. En la Tabla 1 se muestra el número de personas registradas como demandantes de empleo que cumplen con la definición de paro extendido (véase Felgueroso et al., 2017), en algún momento de este período.

**Tabla 1. Descripción de la muestras de demandantes por CCAA.
Sept 2014 - Oct 2016.**

	Demandantes parados	Episodios mensuales	Tamaño muestras para estimación		
			3 meses	6 meses	12 meses
Andalucía	2,090,398	26,951,363	8,030,535	3,529,578	1,250,119
Aragón	250,356	2,643,418	793,316	352,186	126,217
Asturias	210,842	2,679,704	800,213	352,194	124,659
Baleares	205,109	1,819,038	551,023	249,529	86,953
Canarias	546,947	6,820,594	2,025,645	885,499	312,727
Cantabria	113,901	1,302,294	390,173	172,584	59,574
Castilla - La Mancha	484,127	6,105,258	1,829,468	809,218	286,367
Castilla y León	468,006	5,620,374	1,684,216	744,551	264,101
Cataluña	1,371,695	14,551,268	4,364,508	1,936,356	693,540
C. Valenciana	1,077,101	13,261,480	3,956,657	1,743,493	623,437
Extremadura	276,679	3,642,782	1,085,962	478,494	168,509
Galicia	541,052	6,503,177	1,952,992	862,126	304,781
C. de Madrid	1,201,469	13,081,232	3,904,778	1,721,189	612,704
Murcia	325,972	3,585,850	1,069,806	472,724	168,539
Navarra	115,638	1,225,874	366,513	161,663	57,120
País Vasco	377,310	4,501,078	1,342,120	591,269	210,506
Rioja	58,348	626,126	188,317	83,369	29,591
Ceuta	22,866	372,917	110,324	48,085	16,889
Melilla	24,676	376,749	112,294	49,414	17,605
Total con movilidad	9,762,492	115,670,576	34,558,860	15,243,521	5,413,938
Total sin movilidad	9,588,395	115,670,576	34,558,860	15,243,521	5,413,938

- Nota: el número total de demandantes parados no coincide con la suma por CCAA debido a la movilidad existente.

Con este conjunto de datos estimamos los modelos descritos en la sección 3 tanto con la muestra nacional como con la muestra de cada CCAA. Para que la estimación nacional sea tratable, hemos decidido utilizar una sub-muestra aleatoria de 1,5 millones de personas, es decir, algo más de un 15% de la muestra total. En el caso de las CCAA hemos utilizado un 80% de las muestras, reservando el 20% restante a efectos de comprobación de la bondad de ajuste del modelo.

4. Resultados y una primera aplicación del perfilado

En esta sección se presentan los principales resultados obtenidos para la primera fase del proyecto, es decir, los resultados necesarios para la elaboración de un diagnóstico básico de los demandantes desempleados. La sección se estructura en cuatro partes. Iniciamos la sección con un breve análisis de los resultados obtenidos con la muestra nacional, seguido por un análisis descriptivo de la distribución del grado de empleabilidad resultante. A continuación detallamos los principales ingredientes de nuestra propuesta de orientación. Finalmente presentamos algunos ejemplos de diagnóstico individual.

4.1. Los resultados para la muestra nacional

En esta sección presentamos brevemente los principales resultados de la estimación con un 15 por ciento de la muestra de demandantes de empleo entre septiembre del 2014 y julio del 2016. La Tabla 2 presenta la distribución de la muestra según las principales características individuales (género, edad, nivel de educación, formación, duración en el desempleo, etcétera). Así, vemos como el 47 por ciento de las observaciones corresponden a varones y el 53 por ciento restante a mujeres. El grueso de los desempleados (36.5%) son de 35-44 años de edad, de nacionalidad española (88.7%), con educación primaria (45.3%) cuentan con algún tipo de formación complementaria (83.7%) y tienen una duración de desempleo inferior a tres meses (53.0%). La mayor parte de los demandantes incluidos en la muestra de estimación no disponen de prestaciones (60.6%) y tienen experiencia previa en el sector servicios (63.0%).

Tabla 2. Estadísticos descriptivos de las principales variables de la muestra de demandantes a sept-14 y tasas de salida a 3, 6 y 12 meses.

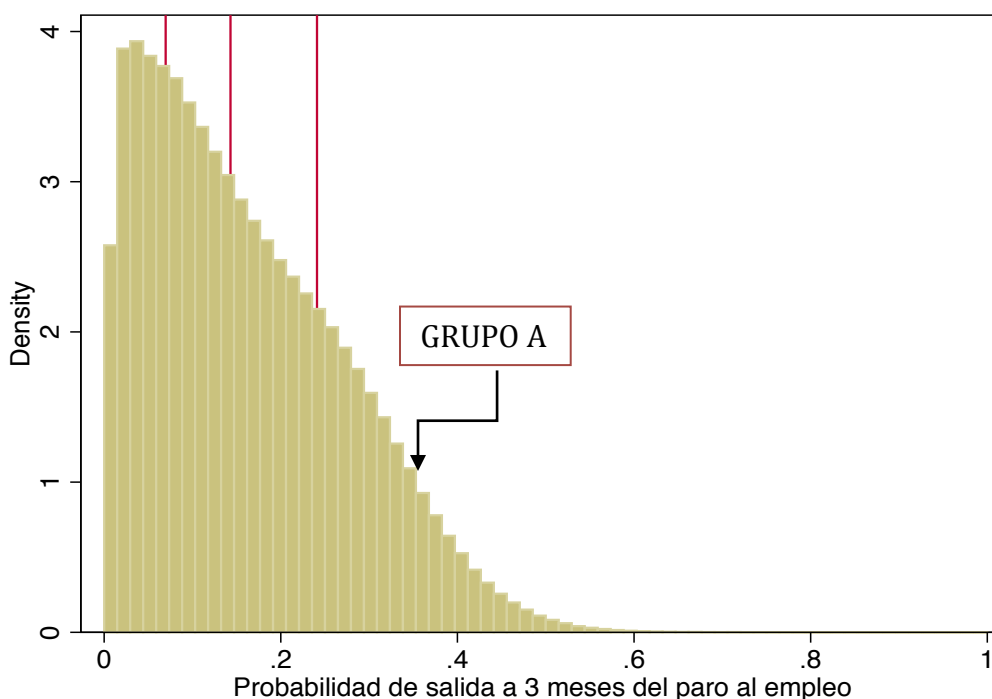
	Distribución sept-14 (%)	Tasas de salida del paro al empleo antes de		
		3 meses	6 meses	12 meses
TODA LA MUESTRA	100	14,54	22,6	31,87
Género				
<i>Varón</i>	47,0%	16,5%	25,6%	35,7%
<i>Mujer</i>	53,0%	12,6%	19,8%	28,2%
Edad				
<i>16-24 años</i>	9,9%	16,6%	25,8%	36,9%
<i>25-34 años</i>	25,0%	18,5%	28,6%	39,9%
<i>35-44 años</i>	36,5%	16,5%	25,7%	36,1%
<i>45-54 años</i>	21,9%	8,5%	13,5%	19,4%
<i>55-64 años</i>	6,6%	3,1%	4,9%	6,9%
Nacionalidad				
<i>Española</i>	88,7%	14,5%	22,6%	31,9%
<i>Extranjero</i>	11,3%	15,0%	22,9%	31,5%
Nivel educativo				
<i>Educación primaria o inferior</i>	45,3%	12,3%	19,2%	27,2%
ESO	23,8%	15,4%	24,0%	34,0%
Bachiller	7,4%	13,5%	21,2%	30,3%
FP Grado Medio	7,0%	18,4%	28,4%	39,4%
FP Grado Superior	6,0%	18,3%	28,3%	39,6%
Título Universitario	10,5%	19,3%	29,3%	40,5%
Formación complementaria				
Si	83,7%	13,9%	21,6%	30,5%
No	16,3%	18,6%	28,9%	40,3%
Duración de la situación de desempleo				
< 3 meses	52,0%	20,8%	31,3%	42,0%
3-6 meses	12,9%	13,9%	22,1%	32,7%
6-12 meses	8,1%	9,5%	16,6%	26,7%
12-24 meses	10,7%	6,7%	11,6%	18,6%
> 24 meses	16,2%	3,6%	6,4%	10,2%
Prestación				
Sin prestación	60,6%	12,9%	19,5%	27,3%
Contributiva	13,5%	21,2%	33,0%	45,7%
Otras	25,9%	12,8%	20,8%	30,2%
Experiencia				
Sin experiencia	7,5%	6,5%	10,7%	17,1%
Agricultura	8,7%	22,2%	32,9%	43,3%
Industria	9,5%	12,9%	20,4%	29,3%
Construcción	11,3%	15,4%	24,3%	35,0%
Servicios	63,0%	15,3%	23,6%	33,0%

En Felgueroso et al. (2017) se presentan los resultados de la estimación del modelo de perfilado para la muestra utilizada en este artículo aplicada a los tres horizontes temporales considerados: a corto (3 meses), medio (6 meses) y largo plazo (12 meses).

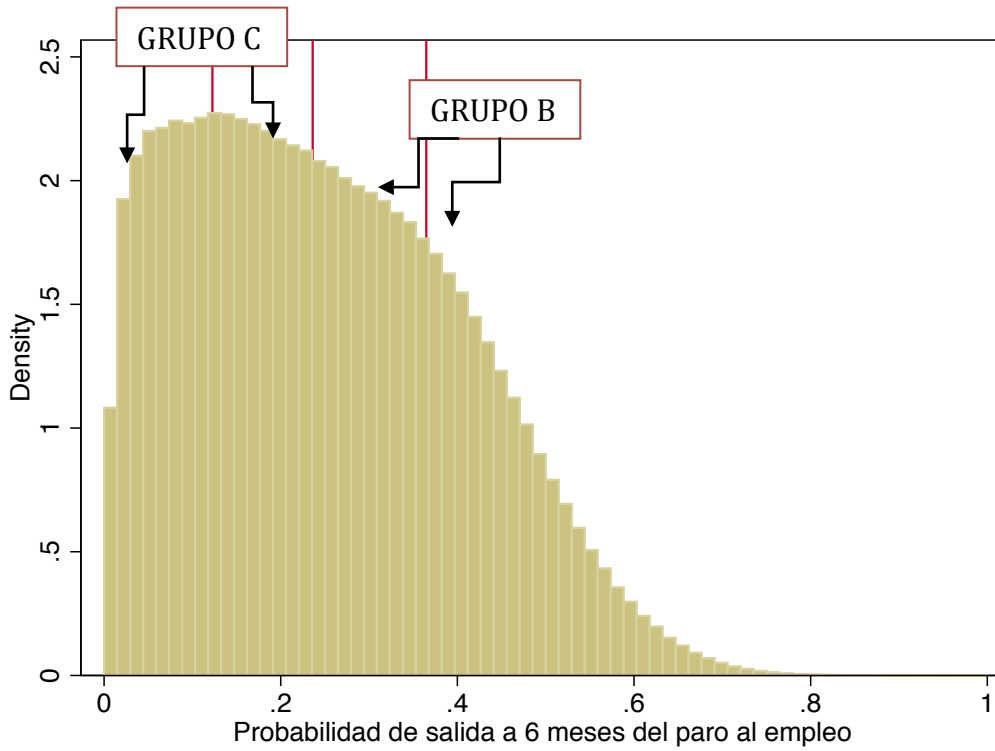
4.2. Clasificación de los demandantes en función de su empleabilidad para las CCAA

En base a los resultados del modelo estimado, se puede calcular el grado de empleabilidad individual siguiendo el procedimiento descrito en la Sección 2.1.3. Las figuras 8 a 10 presentan la distribución de este grado de empleabilidad implícito en dichas estimaciones. En cada una de estas figuras, las líneas verticales de color rojo representan los límites entre cuartiles de la distribución de probabilidades de empleo predichas. Así, por ejemplo, los individuos situados en el cuartil superior (de la derecha) de la Figura 8 son los de mayor empleabilidad o grupo A. Los situados en el cuartil inferior de la distribución a 12 meses (Figura 10) son los de más difícil inserción o grupo D. Finalmente, los que están en los dos cuartiles de la derecha de la distribución (y que no están en alguna de las anteriores) a medio plazo (Figura 9) son los de grupo B y el resto se clasifican como grupo C.

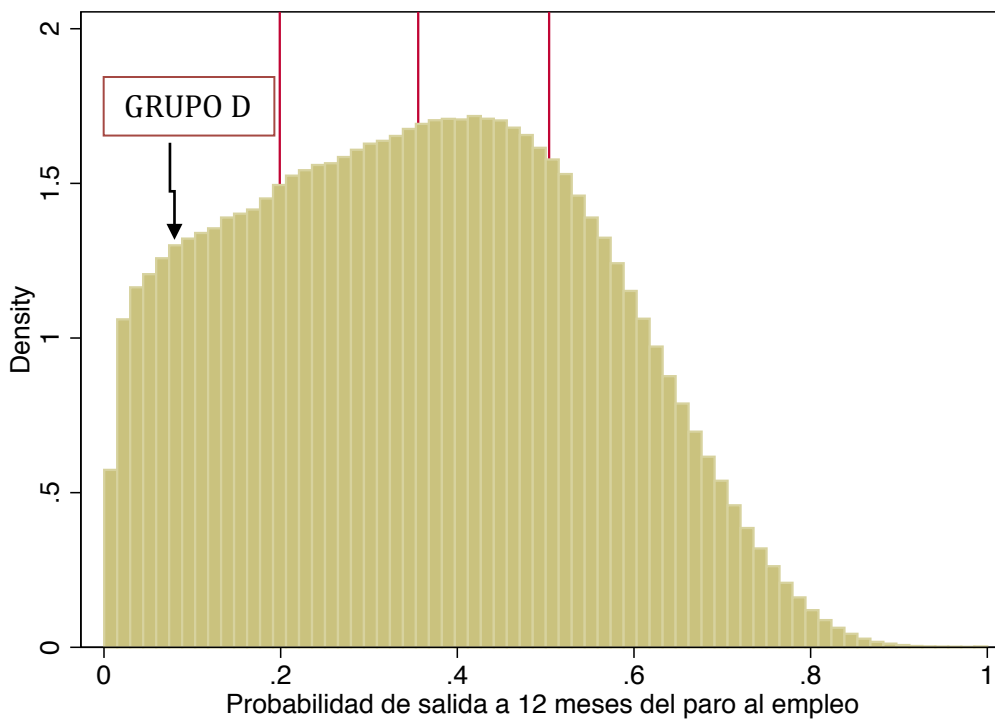
Figura 8. Distribución de empleabilidad en la muestra nacional a corto plazo Sept 2014-Julio 2016, (Basada en las estimaciones del modelo a 3 meses, Tabla B1)



**Figura 9. Distribución de empleabilidad en la muestra nacional a medio plazo
Sept 2014-Julio 2016, (Basada en las estimaciones del modelo a 6 meses, Tabla B1)**



**Figura 10. Distribución de empleabilidad en la muestra nacional a largo plazo
Sept 2014-Julio 2016, (Basada en las estimaciones del modelo a 12 meses, Tabla B1)**



La Tabla 3 resume los resultados obtenidos en términos de empleabilidad para cada región en relación a la distribución de empleabilidad a nivel nacional. En particular se muestra la distribución en grupos de empleabilidad (de A, mejor, a D, peor empleabilidad) de los parados en cada CCAA si se utilizan los umbrales determinados por las estimaciones a nivel nacional para la definición de grupos. Ello nos permite valorar si la distribución de los desempleados de una comunidad es relativamente mejor (tienen mayor peso los grupos A y B que el resto) o peor (tienen, por contra, mayor peso los grupos C y D) que el total nacional.

Es fácil detectar importantes diferencias entre comunidades en la localización de los desempleados respecto a la distribución estandarizada a nivel nacional. Así por ejemplo, Ceuta tiene relativamente pocos desempleados en el mejor grupo a nivel nacional (3.1%) y muchos en el tercer grupo (46.3%). Alternativamente Baleares tiene relativamente más en los grupos A (31.6%) y B (33.3%) que en los grupos C y D (19.3% y 15.8%, respectivamente). Finalmente, la distribución de Madrid prácticamente replica la distribución a nivel nacional.

Tabla 3: Distribución de los demandantes por grupos de empleabilidad para cada CCAA usando umbrales a nacionales (octubre 2016)

CCAA\Grupos	A	B	C	D
Andalucía	26.3	26.0	26.0	21.7
Aragón	31.2	19.9	23.4	25.5
Asturias	19.0	19.3	24.7	37.1
Baleares	31.6	33.3	19.3	15.8
Canarias	21.3	25.1	25.2	28.4
Cantabria	19.6	28.3	26.0	26.0
Castilla - La Mancha	27.0	29.8	27.7	15.5
Castilla y León	23.4	24.4	25.5	26.7
Cataluña	25.2	24.9	21.0	28.9
C. Valenciana	23.5	23.2	26.7	26.6
Extremadura	33.7	28.1	25.0	13.1
Galicia	21.9	22.3	26.2	29.6
Madrid	24.2	26.6	25.1	24.1
Murcia	24.7	25.4	24.7	25.3
Navarra	28.7	22.4	22.0	26.9
País Vaco	25.4	19.6	21.9	33.1
Rioja	31.0	18.7	24.3	25.9
Ceuta	3.1	17.2	46.3	33.5
Melilla	12.9	16.8	28.2	42.2
Total	25.0	25.0	25.0	25.0

4.3 Una primera orientación sobre el tratamiento en base al perfilado

Tal y como comentamos al final de la sección 2 el diagnóstico básico de empleabilidad se completa mediante la identificación del conjunto de características que el demandante no posee y que podrían modificarse a través de políticas activas o servicios de empleo, mejorando su nivel de empleabilidad. Consideramos cuatro grandes grupos de políticas activas que potencialmente puede afectar a la empleabilidad: i) los servicios de orientación profesional; ii) los servicios de colocación; iii) la formación; y iv) el fomento directo del empleo mediante medidas tales como subvenciones o la oferta de un empleo público. Para identificar a los que se beneficiarían más de cada servicio, utilizamos los valores estimados de los coeficientes del modelo. Tal y como se recogía en el Gráfico 7 cada una de estas políticas se relaciona con los coeficientes estimados de algunas de las variables del modelo. Concretamente, la necesidad de recibir servicios de orientación profesional se infiere de la estimación de un efecto marginal positivo y significativo para las variables que representan ocupaciones distintas de las deseadas por el individuo porque ese resultado nos alerta de que la probabilidad de encontrar trabajo del individuo sería mayor en otras ocupaciones. De la misma forma, la conveniencia de ofrecer al demandante servicios de colocación se infiere de la ausencia de efecto significativo en la variable que recoge el ratio entre el número de contratos esperados en la ocupación deseada sobre demandantes en dicha ocupación en el mercado en el que busca dicho demandante. Así, si la probabilidad de empleo no aumenta significativamente cuando hay más contratos disponibles, cabe suponer que el individuo no está buscando trabajo correctamente. En tercer lugar, las políticas de formación se asocian con los efectos estimados del número de horas de formación complementaria; y, finalmente, las políticas de fomento de empleo se identifican con las variables que miden el efecto de la experiencia en la ocupación deseada por el individuo desempleado. Veamos, a continuación, y con un cierto detalle, la interpretación de cada una de las variables representativas de los grupos escogidos.

Figura 11. Variables específicas en la primera orientación sobre los grupos de servicios que mejorarían la empleabilidad del demandante



i) Orientación profesional

El procedimiento que utilizamos para detectar la necesidad de servicios de orientación profesional se basa en la detección de posibles ocupaciones en las que el demandante tendría mayores probabilidades de encontrar empleo que en las que ha solicitado. El modelo de perfilado se ha estimado incluyendo como variables explicativas una serie de variables ficticias que corresponden a distintos tipos de ocupaciones. Cuando la ocupación deseada por un individuo difiere de aquella que maximizaría su probabilidad de empleo en ocupaciones alternativas, parece razonable concluir que el demandante podría beneficiarse de un servicio de orientación profesional que le ayude a dirigir de la forma más eficiente posible su búsqueda de trabajo (tal y como indica la Figura 11).

ii) Colocación

La variable que consideramos indicativa de la necesidad de servicios de colocación es el ratio entre el número de contratos y el de demandantes para la ocupación

deseada por cada demandante.⁴ El número de contratos se refiere a los registrados en los siguientes N meses (3, 6 o 12, según se trata de la estimación del modelo a corto, medio o largo plazo), mientras que el número de demandantes se refiere a los registrados a final de mes.

Dados los valores de las demás variables, a mayor número de contratos por demandante para una ocupación, mayor debería ser la probabilidad de salida del paro al empleo del demandante que desee trabajar en esta ocupación. Si el modelo no ofrece este resultado (es decir, si el parámetro correspondiente a esta variable no resulta significativamente positivo), quiere decir que esta persona requiere de una ayuda específica en la intermediación para conseguir este puesto de trabajo, por ejemplo, la intervención de una agencia privada de colaboración.

Es importante remarcar que esta variable puede estar censurada por no disponer aún de información sobre la evolución real de los contratos en los próximos N meses.⁵ Para resolver esta censura, en esta versión beta del modelo, se supone que las tasas de crecimiento interanual de los contratos por ocupaciones son similares a las registradas en año anterior.

iii) Formación

En el modelo se incluyen una serie de variables ficticias que nos indican el número de horas de formación complementaria que ha recibido cada demandante. Los efectos marginales calculados en base a los parámetros estimados para estas variables nos dan una primera indicación sobre si el demandante podría beneficiarse o no de una mejora de su formación. Es decir, si a través de un curso de formación complementaria podría, eventualmente, mejorar su empleabilidad.

iv) Fomento del empleo

El cuarto grupo de servicios que consideramos incluye todas aquellas políticas que pueden mejorar la experiencia profesional en una ocupación deseada. Aquí se incluyen medidas tales como las subvenciones y bonificaciones al empleo de determinados colectivos, los planes públicos de empleo y las escuelas taller. La indicación sobre la necesidad de este tipo de política se basa en los efectos

⁴ La clasificación de ocupaciones utilizada para esta variable es la CNO a cuatro dígitos.

⁵ En nuestra versión beta, la censura se produce a partir de agosto de 2016 en la estimación de las salidas al empleo antes de tres meses, a partir de mayo de 2016 en el modelo a 6 meses, y a partir de noviembre de 2015 en el modelo a 12 meses.

marginales calculados en base a los coeficientes estimados de las variables de experiencia profesional, esto es, una variable indicador de si ha trabajado en la ocupación deseada y una serie de variables ficticias para el número de años que ha trabajado en dicha ocupación. Es decir, esta variable, valora si mejorar la experiencia del individuo mediante un programa de empleo en la ocupación deseada mejora o no su empleabilidad futura.

Finalmente, las conclusiones que se obtienen a partir del análisis de estas variables se pueden resumir en una ficha individual que podría tener un formato parecido al que se muestra en Figura 12. En la primera columna se mostraría el grupo de empleabilidad al que pertenece el demandante y en las columnas 3-6, se mostraría una recomendación para cada uno de los grandes grupos de servicios de los SPE. Las filas permiten diferenciar las recomendaciones en función del horizonte considerado. Con esta tabla, culminaría el diagnóstico básico que se puede realizar con nuestro modelo de perfilado. En la siguiente sub-sección veremos ejemplos prácticos de cómo funciona esta herramienta básica.

Figura 12. Cuadro de orientación

GRUPO	Salida antes de	ORIENTACIÓN	COLOCACIÓN	FORMACIÓN	EMPLEO
A,B, C o D	3 meses				
	6 meses				
	12 meses				

4.4. Ejemplos de diagnóstico individual

En esta sección presentamos algunos ejemplos de diagnóstico individualizado siguiendo los criterios establecidos en la metodología y en base a las variables comentadas en la sección anterior. Es decir, planteamos el análisis de la situación de un demandante concreto respecto a la distribución de empleabilidad de su ámbito de referencia (en general la comunidad donde reside) y también el análisis del efecto del cambio en algunas variables del modelo sobre el grado de empleabilidad de dicho demandante que, recordemos, está caracterizado por una

combinación de características concreto basadas en su región de residencia, su género, su nivel educativo y su grupo de edad, entre otras características.

Para cada demandante evaluamos el efecto sobre la empleabilidad del cambio de variables representativas de los cuatro grandes aspectos de la orientación profesional que detallamos en la sección anterior.

1. **Orientación Profesional:** Cambios en la ocupación.
2. **Colocación:** Cambio en la demanda de la ocupación deseada a nivel local, autonómico o nacional, dependiendo del ámbito de búsqueda del individuo, lo que permite evaluar los cambios en el grado de empleabilidad del demandante derivados de cambios en las condiciones de mercado de trabajo. En nuestro ejercicio aproximamos las condiciones del mercado de trabajo con el número de contratos firmados en la ocupación deseada respecto al número de demandantes en la ocupación deseada. (Colocación).
3. **Formación:** Cambio en el número de horas de formación, caracterizado en 6 intervalos: 1-50, 51-100, 101-250, 251-500, 501-1000, 1000 o más.
4. **Empleo:** Cambios en el nivel de experiencia de la ocupación deseada.

A modo de ejemplo consideraremos un colectivo concreto consistente en mujeres de entre 30 y 34 años, residentes en la provincia de Toledo, que llevan desempleadas entre 6 y 11 meses, cuentan con un nivel de estudios de Bachillerato y tienen una experiencia previa en el sector de Administrativo.

El análisis de diagnóstico llevado a cabo para realizar las recomendaciones a este colectivo conlleva los siguientes tres pasos principales:

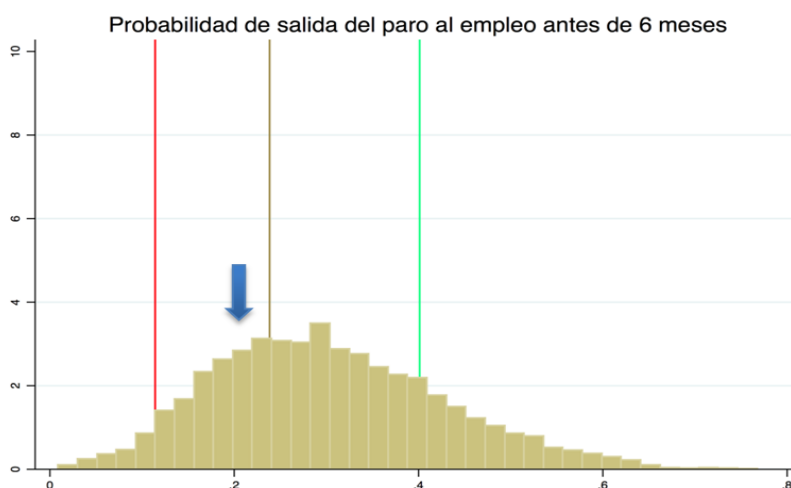
- a) **la estimación del modelo para la muestra autonómica para obtener los grupos de empleabilidad de referencia** (Castilla la Mancha para el colectivo 1).
- b) **la estimación de los parámetros en la muestra de la celda que engloba al colectivo de referencia dentro de la región escogida** (mujeres de Toledo con nivel educativo Bachiller o FP de grado medio y edad entre 30 y 44 años) para calcular los parámetros específicos que determinan el grado de empleabilidad del individuo.
- c) **El análisis del efecto sobre el grado de empleabilidad del cambio en las características escogidas para un demandante específico dentro del colectivo.** Para este demandante, todas las variables del modelo toman un valor específico. Por ejemplo en nuestro caso tiene experiencia

administrativa, demanda trabajo a tiempo completo y demanda empleada de oficina sin atención al público.

Figura 13. Probabilidad de salida del demandante tipo para el colectivo 1 a corto, medio y largo plazo.

Salida del paro al empleo antes de	Distribución de la probabilidades de salida en la CCAA de residencia del demandante				Demandante	
	Cuartil 1	Cuartil 2	Cuartil 3	Cuartil 4	Prob. salida	Segmento
3 meses	< 4,4	4,4%-10,5%	10,5%-21,1%	>= 21,1%	13,1%	C
6 meses	< 11,4%	11,4%-23,9%	23,9%-40,1%	>= 40,1%	21,5%	
12 meses	< 22,2%	22,2%-40,1%	40,1%-58,6%	>= 58,6%	35,5%	

Distribución de probabilidades de salida en la muestra utilizada para la simulación



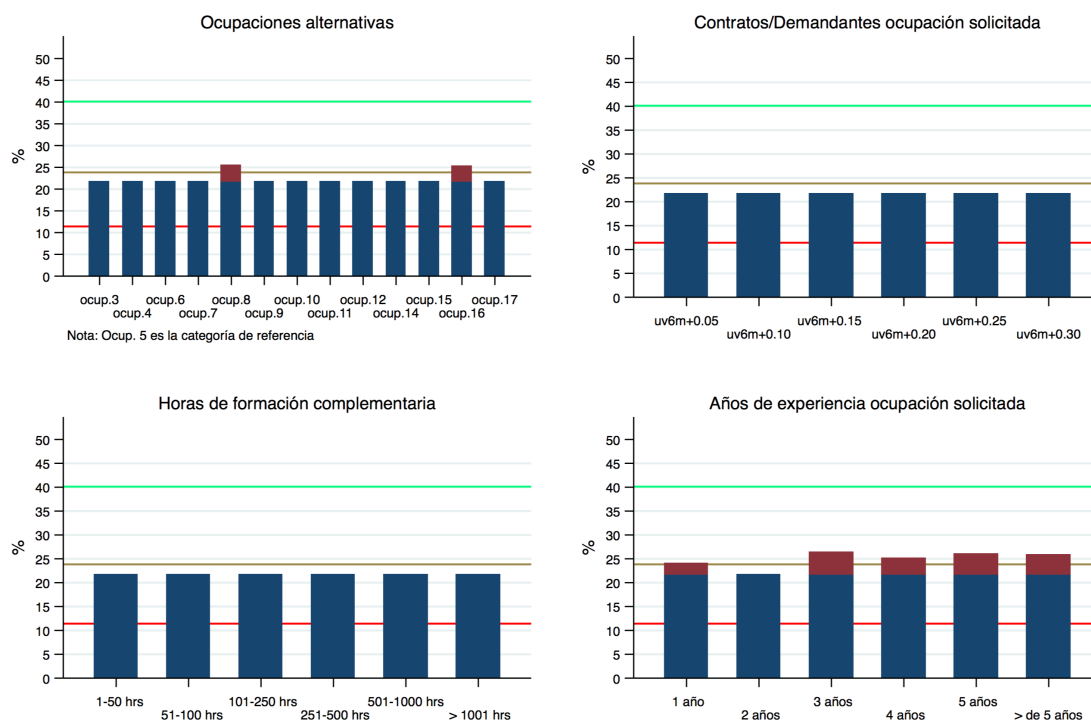
Dadas las estimaciones detalladas antes, la Figura 13 muestra la posición o ranking relativo de la demandante escogida en la distribución de desempleados de Castilla La Mancha, a corto, medio y largo plazo. En cada uno de los plazos presentamos las probabilidades que delimitan la pertinencia a uno u otro cuartil. Por ejemplo, a corto plazo (en tres meses) la demandante tiene una probabilidad de encontrar empleo del 13,1%, lo que la sitúa en el tercer cuartil. A seis meses vista, tiene una probabilidad de salir del 21.5%, lo que la sitúa en el segundo cuartil. Finalmente, a un año vista tiene una probabilidad del 35.5%, que la sitúa también en el segundo cuartil. Dada la situación de la demandante en las tres distribuciones consideradas, la clasificación de su empleabilidad viene determinada por su posición a medio plazo (6 meses), que al ser cuartil 2 la lleva al segmento o grupo C, esto es, a un nivel de empleabilidad medio-baja.

A continuación analizamos si las diversas acciones que hemos considerado en este caso son capaces de mejorar la empleabilidad de esta demandante y llevarle hasta un segmento superior, grupo B en su caso. Como pertenece al segmento C, debemos fijarnos sobre todo en los efectos a medio y por eso representamos en la Figura 15 los efectos de estas acciones sobre la probabilidad de salir del desempleo antes de 6 meses.

En esta figura, la barra de color azul representa la probabilidad de encontrar empleo en el plazo considerado para la demandante de referencia, mujer de 30-34, que reside en Toledo, es bachiller, lleva 6-11 meses en el desempleo y sus último empleo fue como administrativa. Dicha probabilidad (véase la Figura 15) se sitúa en torno al 21,5 por ciento a medio plazo. Por encima de la barra azul, la franja en rojo, que solo se representa en el caso de ser estadísticamente significativa, representa la ganancia que se obtiene en términos de probabilidad al adquirir una determinada característica. Como ya hemos comentado con anterioridad, las líneas horizontales marcan los límites entre los cuartiles de empleabilidad de la comunidad o provincia en la que el individuo vive.

En lo que respecta a los resultados encontramos que solo dos ocupaciones alternativas (empleo en servicios de salud y el cuidado y empleo no cualificados de servicios) parecen tener efectos significativos sobre el nivel de empleabilidad a medio plazo (Figura 15), entre cuatro y cinco puntos porcentuales, que además la llevan a cambiar de segmento (desde la parte superior del C a la inferior del B). Por otra parte, ni el ratio de contratos/demandantes en la ocupación solicitada, ni las horas de formación parecen tener efectos significativos sobre su empleabilidad. Finalmente, la experiencia sí que parece ser un factor en la empleabilidad, con un efecto máximo de casi 5 pp a partir de 3 años de experiencia como se puede ver en la Figura 15.

Figura 15. Análisis del efecto sobre la probabilidad de salida antes de 6 meses de cambios en características. Mujer de Toledo, bachiller, de 30-34, desempleada 6-11 m. Diciembre 2015.



Finalmente, la Figura 16 resume los efectos encontrados en relación a las variables escogidas de orientación profesional (ocupación), colocación (demanda), formación y empleo (experiencia). Respecto a la orientación, encontramos que nuevas opciones de empleo pueden mejorar su empleabilidad y le llevan a cambiar de grupo. Encontramos que la demanda no parece tener un efecto significativo sobre su empleabilidad, por lo que necesita asesoramiento cara a la colocación. Finalmente, la formación no tiene efecto significativo sobre este tipo de demandante y la experiencia aumenta significativamente su empleabilidad y, como en el caso de la orientación, conduce a la demandante a un segmento superior. En resumen un orientador que usara esta herramienta, recomendaría a los demandantes en el colectivo 1 intentar buscar en ocupaciones alternativas (en concreto en servicios de salud y el cuidado y empleos no cualificados de servicios) y aumentar la experiencia en la ocupación solicitada, sugiriendo la participación del demandante en los programas de empleo disponibles.

Figura 16. Resumen/recomendaciones para la demandante de Toledo, bachiller, de 30-34, desempleada 6-11 m. Diciembre 2015.

SEGMENTO	Salida antes de	ORIENTACIÓN	COLOCACIÓN	FORMACIÓN	EMPLEO
C	3 meses	↑	↑↑		↑↑
	6 meses	↑	↑↑		↑↑
	12 meses	↑	↑↑		

5. Conclusiones

En el presente documento detallamos nuestra propuesta de perfilado de demandantes de empleo en España en base a los datos sobre demandantes que dispone el SEPE. El diagnóstico básico se realiza mediante un perfilado estadístico cuyo objetivo es conseguir un indicador de empleabilidad de cada demandante de empleo. Por empleabilidad entendemos la probabilidad de que un demandante transite desde el desempleo al empleo antes de que transcurra un tiempo determinado, teniendo en cuenta sus características sociodemográficas, su historial laboral, sus competencias y la evolución del mercado de trabajo.

La propuesta caracteriza a los individuos según su grado de empleabilidad en cuatro grandes grupos con diferente orden de prioridad lo que facilita la asignación de recursos públicos. En una segunda etapa, nuestra propuesta permite evaluar la respuesta del grado de empleabilidad a diferentes itinerarios de acción por parte del mediador de empleo. Este tipo de evaluación se puede llevar a cabo con prácticamente cualquier combinación de residencia, género, nivel de educación y grupo de edad. Tal y como hemos comprobado los resultados de evaluación varían ampliamente entre colectivos, indicando que debemos ser cuidadosos en la aplicación de políticas o itinerarios formativos ya que el efecto no es homogéneo entre demandantes y/o colectivos aparentemente similares pero de diferentes regiones.

Nuestro proyecto no se acaba aquí ya que comporta dos etapas complementarias de suma importancia y que se desarrollan en documento aparte: el diagnóstico detallado y la evaluación de políticas públicas.

En primer lugar, una vez completado el diagnóstico primario, se pretende plantear un diagnóstico detallado para el que se podría incorporar información más específica sobre las competencias del desempleado, así como sobre el estado del mercado laboral en el que esta persona está buscando empleo. Esta información no está disponible en las bases de datos de tipo administrativo a nuestra disposición. Las fuentes de datos necesarias para esta segunda etapa serían cuestionarios ad-hoc que profundizasen en las habilidades cognitivas y no cognitivas del demandante, así como información procedente de otras fuentes tales como los portales de empleo, observatorios ocupacionales y las redes sociales.

En segundo lugar, resulta imprescindible complementar el diagnóstico primario y detallado con una buena evaluación de las políticas de empleo implementadas en una comunidad específica. La evaluación de las políticas activas es clave para identificar las que tienen un mayor impacto en la mejora de la empleabilidad y por tanto en el diseño del menú de itinerarios que eventualmente ayudaría a elaborar la herramienta propuesta en este informe con el fin de ayudar a los orientadores a prestar una atención individualizada.

En definitiva, nuestra propuesta de perfilado estadístico plantea un menú completo que combina el análisis de empleabilidad de los demandantes de empleo, con la recomendación de itinerarios y la evaluación de políticas públicas.

Referencias bibliográficas

- Carrasco, R. (2002): “Modelos de elección discreta para datos de panel y modelos de duración: una revisión de la literatura”, Cuadernos Económicos del ICE, nº 66, pp. 21-49.
- Felgueroso, F., J. I. García Pérez y S. Jiménez Martín (2017): “Herramienta de Perfilado de parados: modelización y resultados preliminares”, MIMEO
- García Pérez J.I. (2009) (Editor): Metodología y Diseño de Estudios para la Evaluación de Políticas Públicas, Ed. Antoni Bosch.
- Heckman, J (1979): “Sample selection bias as a specification error”, *Econometrica*, 47, pp. 153-61.
- Heckman, J. J., y Singer, B. (1984): «A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data», *Econometrica*, 52(2), 271-320
- Lancaster, T. (1979): “Econometric Methods for the Duration of Unemployment”, *Econometrica*, 47(4), 939-56
- Mortensen, D. (1986): “Job search and labor market analysis” en O. Ashenfelter & R. Layard (ed.), *Handbook of Labor Economics*, Elsevier, volumen 2, numero 2, páginas 849–919.
- O'Connell, Philip J. & McGuinness, Seamus & Kelly, Elish, 2010. "A Statistical Profiling Model of Long-Term Unemployment Risk in Ireland," Papers WP345, Economic and Social Research Institute (ESRI).
- OECD (2016): “Faces of joblessness. Understanding employment barriers to inform policy”.
- Rebollo (2017): El Modelo de Perfilado Estadístico: una herramienta eficiente para caracterizar a los demandantes de empleo, FEDEA, DT2017-1
- Ridder, G. (1987): “The Sensitivity of Duration Models to Misspecified Unobserved Heterogeneity and Duration Dependence”, Mimeo, University of Amsterdam, Amsterdam.
- Rosholm, Michael & Svarer, Michael & Hammer, Bo, 2004. "A Danish Profiling System," IZA Discussion Papers 1418, Institute for the Study of Labor (IZA)
- Van den Berg, G. (2001): “Duration models: specification, identification and multiple durations” en J.J. Heckman & E.E. Leamer (ed.) *Handbook of Econometrics*, Elsevier, volumen 5, número 5, páginas 3381-3460.