

Instituto de Estudios Laborales y del Desarrollo (IELDE)
Facultad de Ciencias Económicas, Jurídicas y Sociales (FCEJyS)
Universidad Nacional de Salta (UNSa)

Seminario N° 11
14 de diciembre de 2010
Hs. 16:00 – Aula Moreno (FCEJyS-UNSa)

ACERCA DE LA ESTIMACIÓN DE LA DURACIÓN DEL DESEMPLEO EN LA ENCUESTA PERMANENTE DE HOGARES (*)

Juan Carlos Cid

Resumen

La duración promedio de un episodio completo de desempleo es uno de los indicadores más interesantes de la situación del mercado laboral. En la Argentina las diferentes estimaciones de ese período usaron los micro datos de la Encuesta Permanente de Hogares (EPH).

Este documento se propuso analizar la consistencia de la información sobre la antigüedad en el paro que proporcionan los entrevistados de la EPH.

Los desocupados estudiados corresponden al primer trimestre del corriente año en todos los aglomerados. Se chequeó la coherencia de sus respuestas a las 4 preguntas básicas del bloque de desocupación del cuestionario individual. Además de la lógica interna de esas respuestas puntualmente para el primer trimestre, se comparó con la información que las mismas personas suministraron en los trimestres anterior y posterior al mencionado (consistencia temporal).

La conclusión es que los datos de antigüedad en el paro de por lo menos la cuarta parte de los entrevistados serían inconsistentes, puntual o temporalmente.

A este problema se suma el diseño de la modalidad continua de la Encuesta aplicada desde 2003, que introduce una indeterminación en la fecha de inicio del episodio de desempleo.

(*) Comentarios, observaciones y sugerencias, dirigirlas a ielde@unsa.edu.ar.

ACERCA DE LA ESTIMACIÓN DE LA DURACIÓN DEL DESEMPLEO EN LA ENCUESTA PERMANENTE DE HOGARES

Juan Carlos Cid*

*Facultad de Ciencias Económicas, Jurídicas y Sociales. Universidad Nacional de Salta (UNSa-IELDE).

La Encuesta Permanente de Hogares (EPH) que se realiza actualmente en nuestro país permite estimar cada trimestre, entre otras cosas, la tasa de desocupación y la cantidad absoluta de desocupados en los principales aglomerados urbanos. Además, la EPH caracteriza mediante una serie de variables al contingente de personas desocupadas a la fecha de la entrevista. Estos datos (edad, sexo, nivel educativo, antigüedad en la búsqueda, etc.) resultan valiosos para los investigadores y para los responsables de diseñar y monitorear las políticas de empleo.

Sin embargo, al analizar el mercado laboral se debe distinguir entre el flujo y el stock, es decir, entre los que **quedan** desempleados por período y los que **están** desempleados a cierta fecha. Sucede que la Encuesta nos brinda una fotografía de un proceso esencialmente dinámico, el stock que describe está siendo modificado por sendos flujos de personas que ingresan al desempleo y de otras que lo abandonan.

Por ejemplo, en el último trimestre de 2009 se contabilizaron 958,6 miles de parados, entre los cuales había 50% de varones y otro tanto de mujeres. Pero en los 267,6 miles de desocupados que llevaban menos de un mes en el paro, cifra que reflejaría el flujo hacia el desempleo, la relación era de 55 varones por cada 45 mujeres. También dentro del flujo de recientes desempleados, 32% se ubicaban en un tramo central de edades, entre 30 y 44 años cumplidos. Sin embargo, la proporción de ese grupo etario en el stock de desocupados descendía a 25%. La explicación de estas diferencias radica en que el stock observado adolece de una sobre representación de los grupos que comparativamente sufren extensos períodos de desempleo –en el primer ejemplo, las mujeres– mientras que los que experimentan episodios breves tienden a quedar sub representados –en el segundo ejemplo, los activos de 30 a 44 años de edad–. En conclusión, las cifras de los desocupados en un punto del tiempo originarán estimaciones erróneas de la probabilidad de quedar desempleado de distintos grupos de población, en la medida en que la duración del desempleo difiera entre ellos.

Uno entonces es el stock de los que están en el paro a cierta fecha –la semana de referencia de la EPH por ejemplo– y otro el flujo hacia el paro durante determinado período, tanto de los que acaban de perder su ocupación como de los recientes ingresantes al mercado laboral que aún no consiguieron emplearse.

El aumento en la tasa de desocupación puede deberse al incremento en el flujo de nuevos desocupados o al crecimiento de la duración promedio de los episodios de desempleo. De acuerdo a cuáles hayan sido las causas, diferirán las acciones que convendría aplicar.

Con respecto a la duración del desempleo, la EPH solo proporciona, al menos en forma directa, información sobre la extensión del episodio que está transcurriendo, sin que sepamos

hasta cuándo se extenderá. Esta limitación de los datos se conoce como truncamiento hacia la derecha. Por ejemplo, si sabemos que se entrevista durante mayo a un desempleado que declara que lleva 1 mes en el paro, se puede calcular que el episodio se inició en abril, pero no podemos precisar en qué mes concluirá, si bien descontamos que será en alguna fecha posterior a mayo.

En la literatura sobre el tema pueden hallarse diversos métodos en los que, a partir de esta información incompleta, se llega a estimar la duración promedio de un período completo de desempleo.

Por supuesto, para que esa estimación se ajuste a la realidad es fundamental que la medición del período corriente sea precisa y que las respuestas de los desocupados sean veraces. Un ejemplo sencillo bastará para ilustrar el problema. Supongamos que una persona queda cesanteada en su trabajo a fines del mes de diciembre. A comienzos de enero del año siguiente inicia la búsqueda de empleo presentándose a entrevistas de selección, etc. Luego de seis meses de infructuosas gestiones, decide retirarse temporalmente del mercado laboral para atender los problemas de salud de un familiar, lo que le demanda el mes de julio completo. Ya liberado de esa obligación en agosto, retoma inmediatamente la búsqueda de trabajo y a comienzos de octubre, cuando lo entrevistan, lleva dos meses en esa situación ¿Cuál será su respuesta si se le pregunta cuánto tiempo lleva buscando en forma continua un empleo?

Hemos mencionado más arriba la importancia de la precisión en la medición de la antigüedad del desocupado. Para nuestro propósito, convendrá entonces analizar el formulario individual actualmente en uso en la Encuesta. El bloque de preguntas del cuestionario que se inicia con la 10a se aplica a todos los desocupados. Entre la 10a y la 10e se plantea una serie de interrogantes que indagan sobre el tiempo de búsqueda y el período transcurrido desde el último trabajo.

Pregunta 10a. *¿Cuánto hace que está buscando trabajo?*

Pregunta 10c. *¿Durante ese tiempo hizo algún trabajo / changa?*

Pregunta 10d. *¿Ha trabajado alguna vez?*

Pregunta 10e. *¿Cuánto tiempo hace que terminó su último trabajo / changa?*

La pregunta 10a tiene cinco alternativas precodificadas de respuesta: menos de un mes de búsqueda; de 1 a 3 meses; más de 3 a 6 meses; más de 6 a 12 meses y más de 1 año. La pregunta 10e tiene esos mismos períodos excepto que el más extenso se divide en dos: más de 1 a 3 años (quinta alternativa de respuesta) y más de 3 años (sexta alternativa). Las preguntas 10c y 10d admiten únicamente la contestación por sí (valor 1) o por no (valor 2). La 10d se formula solamente a los parados que respondieron negativamente a la 10c. Al interrogante 10e se llega tras una respuesta afirmativa a 10c o a 10d. Si en ambas se contestó negativamente, se trata de un nuevo trabajador sin experiencia y no se hace la pregunta 10e. Si por el flujo establecido para la secuencia no corresponde formular alguna pregunta, la misma aparece en la base con valor cero.

Puede ocurrir que una persona ocupada empiece a buscar trabajo antes de quedarse sin puesto. En sentido contrario, es factible también que alguien que pierde su trabajo permanezca un cierto lapso inactivo y luego decida reingresar al mercado laboral comenzando la búsqueda. Teniendo entonces en cuenta esas dos situaciones, el tiempo que lleva el episodio de desempleo en transcurso al momento de la Encuesta resulta el menor período entre los declarados en 10a y, en caso de corresponder su formulación, en 10e.

Analicemos ahora el detalle de las distintas combinaciones de respuesta que se pueden presentar en esta secuencia de preguntas.

- i) No trabajó durante el período de búsqueda ni trabajó alguna vez. Como ya vimos, esta alternativa corresponde a los ingresantes al mercado laboral sin experiencia previa. La secuencia de respuestas es, si por ejemplo llevan menos de un mes de búsqueda, 1 – 2 – 2 – 0.
- ii) Hizo algún trabajo o changa durante el lapso que buscó empleo. Si se produce esta segunda combinación de respuestas, el tiempo que transcurrió desde la finalización del último trabajo tiene que ser menor que el que la persona lleva buscando empleo. Una posible secuencia es 2 – 1 – 0 – 1. Debido a los intervalos de tiempo preestablecidos, se admite que el período en 10e sea menor pero también igual que el informado en 10a. Por consiguiente, si lleva de uno a tres meses de búsqueda la finalización de ese trabajo o changa debió producirse en el último mes o, a lo sumo, en el tramo de 1 a 3 meses. Llamaremos inconsistencia del tipo I a la secuencia en que no se cumple la regla. Por ejemplo, es obvio que resulta incoherente una secuencia 2 – 1 – 0 – 5.
- iii) No hizo ningún trabajo en el período de búsqueda, pero sí trabajó alguna vez previamente. Si se produce esta tercera combinación de respuestas, el tiempo transcurrido desde el último desempeño laboral tiene que ser mayor o igual al que se lleva buscando trabajo. También se admite aquí que el período indicado en 10e sea igual al de 10a. A la violación de esta condición, por ejemplo una secuencia 3 – 2 – 1 – 1, la denominaremos inconsistencia del tipo II.

Vayamos ahora a los datos del desempleo del primer trimestre de 2010. Los desocupados en los aglomerados urbanos de la EPH fueron 1.933 casos, que ponderados representaron a 943.681 personas. Los más frecuentes fueron los que llevaban de 1 a 3 meses buscando, no hicieron ningún trabajo durante ese período pero habían tenido empleo previamente, también dentro del lapso de 1 a 3 meses (secuencia de respuestas 2 – 2 – 1 – 2). Fueron 189 casos que expandidos significaron 90.805 parados. La segunda combinación de respuestas más frecuente, en casos y en datos ponderados, fue la de los que llevaban también de 1 a 3 meses pero diferían en que su desempeño laboral se solapó brevemente con la búsqueda. La secuencia 2 – 1 – 0 – 2 se halló en 178 casos equivalentes a 66.541 personas. No precisamente entre las combinaciones más frecuentes, se detectan sin embargo incoherencias evidentes en las respuestas. La primera de ellas, en datos expandidos, es la serie 2 – 2 – 1 – 1, con 7.082 parados. Luego sigue la secuencia 5 – 2 – 1 – 3, otra inconsistencia de tipo II, que correspondió a 6.663 desempleados.

En ese trimestre en total existen 62 combinaciones diferentes de respuestas en el bloque de desocupación, dos de ellas con valores total o parcialmente ignorados. De las 60 restantes, hubo 19 series inconsistentes, once eran del tipo I y ocho del tipo II. Sin embargo, podría

afirmarse que las inconsistencias fueron relativamente poco frecuentes, ya que significaron 5,0% de los casos de la muestra y 5,5% de los datos expandidos.

Estas inconsistencias de tipo I y II se detectan analizando exclusivamente los desocupados de un trimestre. Podríamos denominarlas puntuales, porque surgen de las declaraciones de los desocupados hechas a determinada fecha, en un punto en el tiempo.

Otra clase de inconsistencias, que cabría llamar temporales, se descubren al comparar las respuestas de las personas en dos entrevistas consecutivas. Como es sabido, la EPH no fue diseñada como encuesta de panel pero aproximadamente 50% de los hogares de la muestra de un trimestre son visitados nuevamente en el trimestre siguiente. Por consiguiente, con sus integrantes pueden armarse pseudopaneles. Como el objetivo es analizar la calidad de la variable que mide el tiempo de desocupación, se hizo el *matching* con los entrevistados que seguían parados en el trimestre abril-junio. Si las personas están ocupadas en la segunda entrevista, se cuenta con su declaración sobre el tiempo que llevan en el puesto. Pero el dato no es pertinente en nuestro estudio, en todo caso serviría para conocer la calidad de la variable antigüedad en el empleo. Tampoco es aprovechable la información proporcionada por los inicialmente desocupados que pasaron a la inactividad.

El apareamiento de observaciones correspondientes a desocupados en los dos primeros trimestres fue de 245 casos, que expandidos correspondían a 128.437 personas. Aquí debe analizarse la coherencia de las respuestas a las 4 preguntas en los dos trimestres sucesivos, es decir una secuencia de 8 valores. Por ejemplo, es consistente una serie 2 – 2 – 1 – 2 – 3 – 2 – 1 – 3, que indica que el parado lleva de 1 a 3 meses en esa situación en el primer trimestre y, cuando se lo vuelve a entrevistar tres meses más tarde, declara que el período ya se extiende a más de 3 a 6 meses. Según nuestro criterio, se admiten series como 1 – 2 – 1 – 1 – 2 – 2 – 1 – 2, que indica un corto intervalo en el empleo seguido por un nuevo episodio de paro, o 5 – 2 – 1 – 5 – 5 – 1 – 0 – 2, que también corresponde a un breve lapso con ocupación, pero en este caso sin interrupción de la búsqueda. Claramente no es consistente una combinación de respuestas 3 – 2 – 1 – 3 – 3 – 2 – 1 – 3, porque al transcurrir tres meses entre una entrevista y la otra, el intervalo inicial “más de 3 a 6 meses”, o bien se convierte en “más de 6 a 12 meses” (si siguió acumulando antigüedad en el paro) o bien disminuye a “menos de un mes” o a “de 1 a 3 meses” (si la desocupación se interrumpió por un breve período de empleo o de inactividad). Además, consideramos inconsistente una secuencia 5 – 2 – 2 – 0 – 5 – 2 – 1 – 6 porque, más allá de que el tiempo de permanencia en el paro sea sin ambigüedad de más de un año, en la primera ocasión se declaró ser un nuevo trabajador sin experiencia y en la segunda se reconoció un desempeño laboral en una fecha que antecedió a la de la primera entrevista.

En los 245 desocupados en los dos trimestres existen 12 que tenían inconsistencias puntuales de tipo I y II en el primer trimestre. Esos casos fueron retirados del panel porque de lo contrario se computarían dos veces. Tampoco consideramos 10 casos de respuestas inconsistentes de tipo I y II del segundo trimestre. Es decir que aplicamos un criterio conservador, suponiendo que la inconsistencia puntual del segundo período podía influir en el análisis de la temporal, llevando a sobre estimar esta última. Nos quedamos por consiguiente con 223 casos representando a 113.297 personas. De ellos, los entrevistados con coherencia temporal en las respuestas del bloque de desocupación fueron 100, que ponderados significaron 48.270 personas. El resto, 123 entrevistas del panel y 65.027 personas, proporcionaron respuestas inconsistentes del punto de vista temporal. Esa cifra equivale a 57,4% de los desocupados apareados, aunque no sabemos qué proporción del error se originó en cada trimestre. Suponiendo en forma conservadora que una de las dos veces se haya

contestado bien, un criterio razonable sería asumir que 50% de los errores de medición se originaron en los datos del primer trimestre. Además supondremos que no existen diferencias en la calidad de las respuestas entre la población de desocupados que rotó y la que permaneció en el panel, ni tampoco hay diferencias entre las personas del panel que siguieron desocupadas y las que pasaron a otro estado. Este 28,7% de inconsistencia temporal se sumaría entonces al 5,5% puntual detectado inicialmente.

La consistencia temporal en las declaraciones del tiempo de desocupación que acabamos de establecer era “hacia adelante”. Parados en un punto del tiempo, en el primer trimestre, revisamos si lo declarado por los que seguían desocupados un trimestre después era compatible con lo manifestado previamente. Pero también resulta factible la comparación “hacia atrás”, analizando la consistencia temporal de las respuestas de los desocupados del primer trimestre de 2010 con su situación en el cuarto trimestre de 2009. Obsérvese que, en virtud de que el objetivo es revisar la calidad de la variable que mide el tiempo de desocupación declarado en el primer trimestre, el *matching* se debe hacer con los entrevistados del trimestre previo, cualquiera haya sido su condición de actividad.

Conviene entonces dividir el análisis en las tres situaciones iniciales posibles: ocupado, desocupado e inactivo. Aquí los apareamientos de observaciones fueron 808 y representan a 385.904 individuos. También en esta oportunidad no corresponde considerar los casos que presentan inconsistencias puntuales de los tipos I y II en el primer trimestre de 2010, porque sería contarlos dos veces. El detalle antes y después de esa depuración y según la condición de actividad previa es el siguiente:

<i>Situación (4º trim. 2009)</i>	<i>Casos (1)</i>	<i>Personas (1)</i>	<i>Casos (2)</i>	<i>Personas (2)</i>
Ocupados	313	142.265	300	137.956
Desocupados	212	103.869	204	100.233
Inactivos	283	139.770	267	133.637
Total	808	385.904	771	371.826

En las columnas indicadas con (1) figuran los casos originales de apareamiento y las personas de acuerdo con la ponderación de la muestra y en las señaladas con (2) los mismos conceptos pero después de eliminar los que estaban ya contabilizados como inconsistentes puntuales.

Para los que estaban inicialmente fuera del mercado laboral la atención debe fijarse en los tiempos declarados en las variables 10a y 10e al hallárseles desocupados en el primer trimestre de 2010. Esos períodos tienen que ser inferiores o a lo sumo iguales a tres meses. Por ejemplo, la secuencia de respuestas 5 – 2 – 1 – 6 se halló en la mayor cantidad de personas en esta situación. Si bien es consistente cuando se considera puntualmente, el error surge cuando se la compara con la inactividad detectada en el trimestre precedente. Las incongruencias temporales afectaron a 55,0% de las declaraciones de los inicialmente inactivos, lo que constituyó el porcentaje más alto de las 3 situaciones. Seguramente esto se vincula con el ejemplo que habíamos planteado en la página 2: muchas personas en nuestro país alternan períodos de búsqueda de trabajo con otros de inactividad y tienden a computar como tiempo de búsqueda el transcurrido desde el inicio.

Los que estaban ocupados en el cuarto trimestre de 2009, simultáneamente con su desempeño laboral pudieron estar buscando otro empleo (pregunta 3j del bloque de ocupados del cuestionario). No existen límites entonces para el tiempo de búsqueda declarado en la segunda entrevista. Sin embargo, no puede exceder a tres meses el tiempo que se lleva desocupado, es decir el menor período de las respuestas a las preguntas 10a y 10e. En esta categoría se

presentó la proporción más reducida de inconsistencia temporal, fueron solamente 18,6% de los casos ponderados.

Por último, los desocupados en los 2 trimestres consecutivos permiten comparar las dos series de variables como ya se hizo también en la consistencia temporal “hacia adelante”. Después de la depuración de las inconsistencias puntuales, se detectaron 167 combinaciones de respuestas diferentes sobre los 204 casos de *matching*. Teniendo en cuenta que los 4 primeros valores corresponden a las respuestas de la primera entrevista y los 4 últimos a las del primer trimestre de 2010, se consideró válida por ejemplo la serie 2 – 2 – 1 – 6 – 3 – 1 – 0 – 1. En la entrevista del cuarto trimestre del año pasado, esta persona llevaba buscando de uno a tres meses y habían transcurrido más de tres años desde su último empleo. Cuando se lo encuestó nuevamente en el primer trimestre de 2010 el tiempo de búsqueda pasó al tramo superior, de 3 a 6 meses, pero hacía menos de un mes había cesado en una ocupación, seguramente transitoria. Otro ejemplo de una secuencia válida de respuestas es 5 – 1 – 0 – 5 – 2 – 2 – 1 – 5. Inicialmente el desocupado llevaba más de un año en la búsqueda y su salida del empleo se había producido dentro de ese lapso; luego respondió que acumulaba entre 1 y 3 meses de antigüedad, que en ese período no había tenido ninguna ocupación aunque había trabajado alguna vez y que el cese se remontaba a más de un año hacia atrás. Por el contrario, se consideró inconsistente temporalmente la serie 2 – 1 – 0 – 3 – 5 – 2 – 1 – 5. Esta respuesta es incoherente de tipo I en el cuarto trimestre, pero registra por añadidura una inconsistencia temporal al saltar de una antigüedad de uno a tres meses en el paro a más de un año un trimestre después. Se detectaron inconsistencias temporales en 47,0% de los desocupados en trimestres consecutivos.

Los porcentajes de error que se vinieron mencionando corresponden a los tres subconjuntos de desocupados del primer trimestre de 2010 que se logró aparear. La medida resumen es el promedio ponderado por la frecuencia de cada una de las tres situaciones. Se llega así a una cifra de 39,3% de inconsistencia temporal “hacia atrás”. Con un criterio similar al aplicado anteriormente la mitad, aproximadamente 19,6%, sería atribuible a los errores de declaración del primer trimestre de 2010.

Resumamos ahora el resultado de nuestras estimaciones. Cuando a la medición puntual se agrega el cotejo de la coherencia “hacia adelante”, se concluye que 34,2% de las respuestas que nos interesaban del bloque de desocupación son erróneas. Si la coherencia temporal se mide “hacia atrás”, la proporción de inconsistencias sería menor, un 25,1%. Es razonable pensar que el verdadero porcentaje se ubica en algún valor intermedio, quizá más cerca del límite inferior que del superior. Esto se debe a que en el análisis “hacia adelante” supusimos que la calidad de las respuestas era igual para los que permanecían en el paro que para lo que tenían tránsitos hacia el empleo o la inactividad. Sin embargo, una de las conclusiones del análisis “hacia atrás” es que los desocupados que en el trimestre inmediato anterior transitaron por el empleo tienden a responder con más exactitud.

De todas maneras, existen razones para sospechar que nuestro cálculo subestima la verdadera frecuencia de inconsistencias en el bloque de desocupación. A algunas mencionadas en los párrafos previos agreguemos que, de acuerdo con el sistema de rotación que viene aplicando la Encuesta de Hogares, la vivienda vuelve a ser visitada después de un año. De modo que también se podrían revisar las respuestas de los desempleados con ese intervalo, verificando la consistencia temporal hacia atrás anual, lo que no se ha hecho en este estudio.

Por supuesto, todo esto no sería de mayor relevancia si no fuera porque la antigüedad en la desocupación es una variable que constituye uno de los principales insumos en los distintos procedimientos de cálculo de la duración promedio de un período completo de desempleo.

Otras consideraciones pueden hacerse, en el tema de la estimación de la duración del desempleo, sobre la metodología actual de la Encuesta y el diseño de su formulario individual. Para ello será útil mencionar el anterior procedimiento usado por este programa, la Encuesta Permanente de Hogares en su modalidad puntual. El sistema se aplicó hasta mayo de 2003, generalmente en los meses de mayo y octubre de cada año. En el cuestionario individual el bloque de desocupación se iniciaba con la pregunta 32, cuya formulación era *¿Cuánto tiempo hace que está buscando trabajo?* La pregunta era abierta, el entrevistado tenía la posibilidad de precisar días, meses y años que acumulaba en la búsqueda de empleo. Con los archivos para usuarios de la EPH puntual conocíamos que una persona desocupada en el mes de octubre había declarado que tenía 7 meses de búsqueda infructuosa. La conclusión entonces era que el episodio de desempleo que se estaba desarrollando había arrancado en marzo de ese año.

Con la instauración de la modalidad continua en la Encuesta Permanente de Hogares, a partir del segundo semestre de 2003, los usuarios que intentan estimar la duración del desempleo se encuentran con ambigüedades en la fecha de la observación y en la extensión del período corriente: Una persona que es entrevistada en algún mes del primer trimestre de 2010 y se halla desocupada declara que lleva más de 6 meses pero hasta un año de búsqueda. De acuerdo con esta información, el episodio en transcurso puede haber comenzado en cualquier mes del lapso comprendido entre enero (si la encuesta se hizo en enero de 2010 y la respuesta había sido un año) y septiembre de 2009 (si la fecha de la entrevista fue marzo de 2010 y la respuesta fue 6 meses y un día). Sin embargo, el factor estacional es un elemento fundamental en todos los modelos de duración del paro.