

**LOS DETERMINANTES DE LA DURACIÓN DEL DESEMPLEO
EN COLOMBIA
2008**

SANDRA MILENA FLOREZ CALERO

**UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONÓMICAS
DEPARTAMENTO DE ECONOMÍA
SANTIAGO DE CALI
2011**

**LOS DETERMINANTES DE LA DURACIÓN DEL DESEMPLEO
EN COLOMBIA
2008**

SANDRA MILENA FLOREZ CALERO

**Trabajo presentado como requisito para optar al título de:
ECONOMISTA**

**Tutor:
DAVID ORLANDO RUIZ CASTRO**

**UNIVERSIDAD DEL VALLE
FACULTAD DE CIENCIAS SOCIALES Y ECONÓMICAS
DEPARTAMENTO DE ECONOMÍA
SANTIAGO DE CALI
2011**

NOTA DE ACEPTACIÓN

Prof. David Orlando Ruiz Castro

Prof. José Ignacio Uribe G

Santiago de Cali, octubre 7 de 2011

Ciudad y fecha (día, mes, año)

DEDICATORIA

*“A Dios, a mis padres y mi sobrino,
Los motores de mi vida,
Quienes me impulsan
Y me dan la fuerza
Para alcanzar mis metas”*

Sandra Milena

AGRADECIMIENTOS

A Dios por esta oportunidad que me brindó. Como uno de los tantos regalos que me ha dado en mi vida como son mis seres queridos a quienes amo tanto

A mis padres:

Jesús María Florez y Delfina Calero

Por haberme dado el ser, por su constante apoyo y por guiarme siempre en el camino recto, compartiendo mis momentos difíciles, comprendiendo la necesidad de dedicación de este proyecto, para que hoy pueda ver cristalizada esta significativa meta.

A mis familiares, amigos y compañeros por la colaboración, comprensión y apoyo que siempre me brindaron

Al profesor David Orlando Ruiz Castro, director de trabajo de grado, quien asumió el reto de incentivar me, orientarme y apoyarme en este proyecto.

A todos los que confiaron en mis ideas y me ayudaron a realizarlas

Gracias

TABLA DE CONTENIDO

	Pág.
INTRODUCCIÓN.....	1
1. MARCO TEORICO.....	3
1.1 MODELO DE BÚSQUEDA.....	3
2. ESTADO DEL ARTE.....	8
3. METODOLOGIA.....	15
3.1 MODELO.....	15
3.2 LA FUNCIÓN DE SUPERVIVENCIA KAPLAN - MEIER.....	20
4. RASGOS DESCRIPTIVOS DE LA DURACIÓN DEL DESEMPLEO EN COLOMBIA.....	22
5. LA DURACIÓN DEL DESEMPLEO: ANÁLISIS NO PARAMÉTRICO...	27
6. DETERMINANTES DE LA DURACIÓN DEL DESEMPLEO.....	35
7. LA DURACIÓN DEL DESEMPLEO: ANÁLISIS PARAMÉTRICO.....	38
8. CONCLUSIONES.....	46
REFERENCIAS BIBLIOGRAFICAS.....	48
ANEXOS.....	52

INTRODUCCIÓN

La duración del desempleo es uno de los elementos más importantes para comprender el fenómeno del desempleo, debido a que la tasa de desempleo es el resultado de la interacción entre la tasa de entrada al mismo y su duración media. Esta duración está determinada por diferentes elementos como: la dotación de capital humano del desempleado, los mecanismos usados para la búsqueda de empleo, el ingreso familiar a través del salario de reserva, el valor del ocio, los costos directos e indirectos de esa búsqueda, la tasa de descuento intertemporal, la tasa de llegada de ofertas salariales y la distribución salarial en la economía, entre otras.

Dicha duración no solo tiene efectos sobre los ingresos familiares del desempleado, si no que también hace que el desempleado experimente un sentimiento de rechazo social, debido a que el tiempo en que permanece en ese estado siente que su capital humano no es útil a la sociedad, generándole una importante pérdida de bienestar.

De esta manera, conocer los elementos que causan que una persona no consiga empleo y que hagan que su permanencia en este estado se prolongue es fundamental para no solo poder avanzar en el estudio de la duración del desempleo, sino también para poder realizar políticas focalizadas hacia esos determinantes. Teniendo en cuenta todo lo anterior, podemos decir que el análisis de la duración del desempleo es importante por el hecho de que el bienestar de los individuos desempleados está más relacionado al tiempo que ellos permanecen en esa condición y a las posibilidades de reincorporarse al mercado laboral que al mismo hecho de estar desempleados.

Por tal motivo y dada la actual situación del país en la que el desempleo sigue aumentando y la tasa de desempleo está en 11.2% para el mes de abril (DANE, 2011) y la duración promedio es de más de 22 semanas, vale la pena preguntarse: ¿Qué características de los individuos y del mercado laboral explican la permanencia de una persona en estado de

desempleo en Colombia?, ¿Afectan estas características de la misma forma a todos los individuos?, ¿Qué población es la que más dura en el desempleo?

El presente estudio tiene como finalidad identificar los principales determinantes de la duración del desempleo en Colombia para el año 2008. Para ello, se encuentran las características que afectan de la misma forma a todos los individuos y aquellas que afectan de forma diferente a los distintos grupos. Con este fin, se tomó la información de los desocupados de la Encuesta de Calidad de Vida (ECV) del segundo semestre de 2008 realizada por el Departamento Administrativo Nacional de Estadística (DANE) para el total nacional y no solo de 13 áreas metropolitanas como lo permite la Encuesta Continua de Hogares (ECH). Además, se utilizó la distribución Gamma generalizada que es la distribución que mejor se ajusta a los datos y a la forma de la función de riesgo de la duración del desempleo.

Este artículo está organizado de la siguiente manera: en una primera parte la presente introducción. En la segunda sección se planteará la teoría de la búsqueda como pilar fundamental del análisis microeconómico de los determinantes de la duración del desempleo. En un tercer lugar se esboza el estado del arte con relación a los determinantes de la duración del desempleo en Colombia y en otros países de América Latina y Europa. En cuarto lugar se presenta la metodología en donde se exponen los modelos empíricos implementados para alcanzar los objetivos trazados. Los resultados se han dividido en cuatro secciones diferentes donde se presenta un análisis descriptivo, un análisis no paramétrico, una descripción de las variables y un análisis paramétrico. Finalmente, las conclusiones y las referencias bibliográficas complementan la investigación.

1. MARCO TEÓRICO

1.1 MODELO DE BÚSQUEDA

La teoría de la búsqueda se basa en la noción de un comportamiento racional de los desempleados, en el cual ellos continuarán buscando hasta que encuentran una alternativa de trabajo que le ofrece un salario al menos igual a su salario de reserva, de esta manera la teoría de la búsqueda ha centrado su atención en la descripción de la información que debe adquirir y utilizar el agente para adoptar decisiones racionales en un entorno cambiante e incierto (Mortensen, 1986).

El adquirir información sobre el empleo implica unos costos monetarios y no monetarios como anuncios en el periódico, acceso a bases de datos de empleo, los gastos en transporte para acudir a las entrevistas, y lo más importante un elevado costo de oportunidad, el cual está asociado con el tiempo gastado en la utilización de los diferentes canales de búsqueda al no aceptar un empleo con menor salario que el de reserva (McConnell, et al, 2003).

Los primeros modelos que se desarrollaron sobre la teoría de la búsqueda se interesaron en el número de ofertas estudiadas por el buscador, mientras que los modelos actuales se interesan en la duración del desempleo, en el tiempo empleado en la búsqueda.

El modelo pionero fue el de Stigler (1962). Dicho modelo plantea que el buscador observa una oferta salarial y compara los beneficios esperados de tomarla, con los beneficios esperados de rechazarla; sólo acepta la oferta si los beneficios primeros exceden a los segundos, es decir,

$$W_0 \geq W_r \text{ Aceptación del empleo}$$

$$W_0 < W_r \text{ Seguir buscando}$$

Donde: W_0 representa el salario esperado del mercado y W_r representa el salario de reserva

En dicho modelo se considera la duración de la búsqueda t y la oferta salarial W como realizaciones de distribuciones de probabilidad que dependen de las características individuales del agente económico y del entorno en el cual este se mueve.

El modelo supone que el horizonte de vida es infinito, que el individuo es neutral al riesgo, que no tiene restricciones en el mercado de capitales, que todos los agentes inician en el estado desempleo y eventualmente hacen su tránsito hacia el estado empleo. También supone que los parámetros de las distribuciones de probabilidad son constantes y conocidos por el agente. Igualmente supone que no hay aprendizaje y que los procesos son estacionarios.

Antes de especificar el modelo que será utilizado en el presente trabajo, primero se mencionarán algunos trabajos que han utilizado la teoría de la búsqueda para estudiar el desempleo. Layard et al (1991), aborda el tema del problema del empleo mostrando que se debe analizar tanto los factores que presionan el alza de los salarios como la eficacia con que los desempleados buscan empleo. Por otro lado Blanco (1995) estudia el fenómeno del desempleo teniendo en cuenta la tasa de entrada y su duración. Plantea que una teoría de la búsqueda, en la cual, el comportamiento de un individuo racional y maximizador lleva a cabo un proceso por el cual va obteniendo información sobre las vacantes, de tal manera que el individuo primero encuentra una oferta y después decidirá si aceptarla o no.

El planteamiento y el desarrollo matemático del modelo que será empleado en la presente investigación, será el modelo desarrollado por Mortensen (1986).

Sea: $U_e(t)$ = flujo instantáneo de utilidad de estar empleado = $w(t)$

$U_u(t)$ = flujo instantáneo de utilidad de permanecer desempleado = $-c$

Donde c es el coste instantáneo de búsqueda.

Donde el individuo buscará obtener un nivel de utilidad mayor en el caso de estar empleado que en el caso de no estarlo: $U_e(t) > U_u(t)$.

El índice de utilidad se puede definir como V_u para la opción de estar desempleado y V_e para la opción de estar empleado. Con esto, el individuo debe elegir si estar empleado le genera mayor utilidad que no estarlo: $V_e > V_u$

De acuerdo con los autores, la oferta salarial está representada en un salario w , cuya función de distribución es $F(w)$ y la tasa de descuento es intertemporal. De acuerdo con lo anterior, el índice de utilidad del empleo será: $V_e(w, c) = \frac{w}{\rho}$

Es pertinente aclarar que el proceso de búsqueda se contextualiza en un periodo de tiempo continuo y pequeño h , durante el cual el agente incurre en un costo que al final se convierte en un flujo de utilidad $-ch$. El individuo recibe una oferta salarial con probabilidad λh ó no la recibe con una probabilidad $1 - \lambda h$. La probabilidad de recibir la oferta es $\lambda h + O(h)$ para capturar la eventual posibilidad de que una nueva oferta llegue en h . Donde $O(h)$ está definido de tal forma que $\lim_{h \rightarrow 0} \frac{O(h)}{h} = 0$. Si se recibe la oferta, entonces el agente económico debe elegir entre: recibir un índice de utilidad de $V_e(w, -c)$ ó continuar con $V_u(w, -c)$. En un mundo dinámico debe elegir:

$$E_w(\text{Max}\{V_e(w), V_u(w)\}).$$

Con el objetivo de encontrar el índice de utilidad de la búsqueda o de permanecer desempleado V_u , el autor utiliza el principio de optimalidad de Bellman para programación dinámica. Teniendo en cuenta que los valores futuros deben tener un factor de descuento, los autores utilizan $1 + \rho h$, llegando a que la función de valores es:

$$V_u = \frac{-ch}{1 + \rho h} + \frac{(1 - \lambda h)V_u}{1 + \rho h} + \frac{\lambda h}{1 + \rho h} E_w(\text{Max}\{V_e, V_u\}) + O(h) \quad (1)$$

El primer elemento es el valor descontado del costo de búsqueda. El segundo y el tercero conforman un típico valor esperado de una binomial, en este caso del índice de utilidad si se recibe ó no una oferta salarial.

A continuación se presenta el desarrollo matemático realizado por Castellar y Uribe (2003) para encontrar V_u :

Multiplicando la expresión (1) por $(1 + \rho h)$:

$$V_u (1 + \rho h) = -c h + V_u - \lambda h V_u + \lambda h E_W (\text{Max} \{V_e, V_u\}) + (1 + \rho h) O(h)$$

Agrupando a la izquierda para V_u :

$$h V_u (\rho + \lambda) = -c h + \lambda h E_W (\text{Max} \{V_e, V_u\}) + (1 + \rho h) O(h)$$

Dividiendo por h , tomando el límite cuando h tiende a cero y teniendo en cuenta que

$$\lim_{h \rightarrow 0} (1 + \rho h) \frac{O(h)}{h} = 0. \text{ Se llega a:}$$

$$V_u (\rho + \lambda) = -c + \lambda E_W (\text{Max} \{V_e, V_u\}) \quad (2)$$

Teniendo en cuenta que:

$$\text{Max} \{V_e, V_u\} = V_u + \{V_e - V_u \mid V_e > V_u\}$$

$$E_W (\text{Max} \{V_e, V_u\}) = V_u + E_W \{V_e - V_u \mid V_e > V_u\} \text{Prob} (V_e > V_u)$$

Recordando que $V_e = \frac{w}{\rho}$ se llega a:

$$E_W (\text{Max} \{V_e, V_u\}) = V_u + E_W \left\{ \frac{w}{\rho} - V_u \mid \frac{w}{\rho} > V_u \text{ Prob} \left(\frac{w}{\rho} > V_u \right) \right\} \quad (3)$$

Factorizando ρ^{-1} en E_W :

$$E_W (\text{Max} \{V_e, V_u\}) = V_u + \rho^{-1} E_W \{w - \rho V_u \mid w > \rho V_u\} \text{Prob} (w > \rho V_u) \quad (4)$$

Ahora de la definición de esperanza condicional:

$$E_W \{w - \rho V_u \mid w > \rho V_u\} = \frac{\int_{\rho V_u}^{\infty} (w - \rho V_u) f(w) dw}{\text{Prob}(w - \rho V_u)}$$

De donde:

$$E_W \{w - \rho V_u \mid w > \rho V_u\} \text{Prob} (w > \rho V_u) = \int_{\rho V_u}^{\infty} (w - \rho V_u) f(w) dw \quad (5)$$

Introduciendo (5) en (4) y dado que $dF(w) = f(w) dw$ se llega a:

$$E_W (\text{Max} \{V_e, V_u\}) = V_u + \rho^{-1} \int_{\rho V_u}^{\infty} (w - \rho V_u) f(w) dw \quad (6)$$

Llevando (6) a (2):

$$V_u (\lambda + \rho) = -c + \lambda V_u + \lambda \rho^{-1} \int_{\rho V_u}^{\infty} (w - \rho V_u) f(w) d(w)$$

Cancelando a ambos lados $V_u \lambda$ se tiene:

$$V_u \rho = -c + \frac{\lambda}{\rho} + \int_{\rho V_u}^{\infty} (w - \rho V_u) f(w) d(w) \quad (7)$$

En la ecuación anterior, tal y como lo mencionan los autores se advierte que ρV_u es creciente, en tanto que el lado derecho es decreciente en ρV_u y de esta manera existirá una solución única. Debido a que ρV_u es el valor que iguala a w , se le conoce como salario de reserva y su relación con el salario ofrecido por el mercado determina la duración de la búsqueda y la probabilidad de estar o no desempleado.

Esta ecuación muestra que la permanencia de los individuos en desempleo depende tanto de los costos de la búsqueda (c) como del valor presente de la diferencia entre las ofertas salariales recibidas (w) y el salario de reserva ρV_u . La duración del desempleo será mayor entre menores sean los costos de la búsqueda y entre mayor sea el salario de reserva de cada individuo.

2. ESTADO DEL ARTE

Los estudios sobre los determinantes de la duración del desempleo son abundantes en la literatura socioeconómica desde la década de los años ochenta. A nivel colombiano, los primeros análisis al problema de la duración del desempleo surgen de la Misión Chenery realizada entre 1985 y 1986, que en su informe final señala el preocupante aumento del desempleo en el transcurso del primer quinquenio de los ochenta. (Maddock, 1987).

Gracias a las inquietudes que surgieron de esta misión se realizan algunas investigaciones sobre el tema del desempleo en Colombia, como la de Hugo López (1988). Esta investigación estudia las tasas de desempleo y de incidencia, así como su duración media para las cuatro principales ciudades colombianas, utilizando las Encuestas de Hogares de los años de 1976 a 1988. López deriva un modelo de duración a partir de las relaciones entre población activa, desocupados y tiempo de búsqueda, obteniendo una relación en la que la duración incompleta y completa del empleo no tiene relevancia alguna.

De la investigación de López se pueden destacar los siguientes resultados: Las mujeres tienen mayor participación entre el grupo de población con duración mayor a un año. Los grupos de edad no presentan diferencias importantes entre los de duración mayor o menor a un año. Los más jóvenes predominan en ambos grupos, principalmente el grupo de 21-30 años que presenta una participación de 47,3% y 54,8% en los grupos de duración de menos y más de un año respectivamente.

Ya en la década de los noventa, se destaca el trabajo de Tenjo y Ribero (1998), quienes en su investigación incorporan la dimensión microeconómica al análisis de los determinantes de la duración del desempleo. Utilizando la información de los desempleados de la ENH de los años 1988, 1994 y 1996 estiman un modelo Weibull para hombres y mujeres, casados(as) y solteros(as). Además, incluyen de manera novedosa el tipo de contrato, con el propósito de observar las características del puesto de trabajo en la duración del desempleo. Los resultados más importantes de esta investigación sugieren que la duración se

incrementa con la edad, los años aprobados de educación, con la falta de experiencia, con el ser mujer y ser soltero. Adicionalmente tratan de captar el efecto de los cambios en el ritmo de actividad económica, encontrando que el desaceleramiento de la economía entre junio de 1994 y junio de 1996 incrementó la duración del desempleo tanto para hombres como para mujeres.

Núñez y Bernal (1998), estiman un modelo de duración bajo una función de distribución exponencial. La estimación se realizó con los datos para ocupados y desocupados de la Encuesta de Hogares para junio de 1988, 1992 y 1996, tomando las duraciones en el desempleo del modulo de informalidad de las diez principales ciudades del país. Específicamente, calculan cómo se incrementan las probabilidades de salir del desempleo en función de 35 variables, entre las que se encuentran edad, estado civil, educación, ciudad, sector económico, posición ocupacional, número de dependientes, ingresos no laborales y tasa de desempleo. Los autores encontraron que en promedio los desocupados tardaron 8 semanas buscando empleo; observando una mayor duración en los individuos de más de 45 años, los no calificados, los que vivían en las ciudades de Barranquilla y Pasto, los trabajadores por cuenta propia y las personas que buscaban trabajo en el sector de los servicios. Adicionalmente encontraron que la duración del desempleo es contra cíclica cuando se utiliza la tasa de desempleo regional y al utilizar el ingreso no laboral como Proxy del salario de reserva se encuentra una relación negativa.

Por otra parte Castellar y Uribe (2003) investigan qué determina que una persona permanezca en el desempleo en la ciudad de Cali en el periodo 1988-1998. Para su investigación los autores utilizaron la Encuesta Nacional de Hogares de los años: 1988, 1992, 1994, 1996, 1998 y la encuesta del Centro de información para el empleo (CIE) del SENA.

Castellar y Uribe estiman un modelo de duración que supone que la duración del desempleo sigue una distribución Weibull. Los resultados obtenidos de la ENH les permitieron deducir que el ser jefe de hogar disminuye la probabilidad de permanecer en el desempleo, de la misma manera que el ser hombre. Así mismo los ingresos no laborales y la dispersión

salarial del sector de búsqueda tienen un efecto positivo sobre la duración de esta búsqueda. También encontraron que la educación aumenta el tiempo de desempleo hasta los 8 años, y que luego disminuye; la experiencia también aumenta este tiempo en un 2%. Por otro lado, los resultados obtenidos con la CIE mantienen la dirección de los encontrados con la ENH pero no en magnitud: la experiencia tiene un efecto positivo y decreciente.

Martínez (2003) analiza la duración del desempleo y del empleo en Colombia y las probabilidades de salida del desempleo y del empleo con diferentes grupos. Su objetivo es encontrar los determinantes de la duración del desempleo y el empleo e identificar los grupos de personas con mayores duraciones. Para ello utiliza la Encuesta Nacional de Hogares para 1986, 1992, 1996, y la Encuesta Continua de Hogares para el año 2002. En este trabajo hace un primer análisis de la duración del desempleo netamente descriptivo desde el punto de vista de los flujos de entrada y salida de la situación del desempleo. Igualmente, estima cuatro modelos diferentes para encontrar los determinantes de la duración del desempleo y la estabilidad del empleo.

Según el autor, las personas con más de 45 años, así como las mujeres con mayor número de hijos dependientes, los bachilleres, las personas con universidad incompleta y quienes habían trabajado en el sector formal o buscan empleo en él, tienen bajas probabilidades de salir del desempleo. También encuentra que los grupos de mayores probabilidades de salir del desempleo son las personas menores de 45 años, el grupo de mujeres con mayor educación y los individuos que habían trabajado en el sector informal o buscan empleo en él. El autor concluye que la duración promedio del desempleo en Colombia es cada vez mayor.

Uribe y Viáfara (2009) estudian cómo la duración del desempleo se ve afectada por el canal de búsqueda que se utiliza para encontrar empleo. Para ello estiman un modelo paramétrico (distribución Weibull) y un modelo no paramétrico (Kaplan Meier), utilizando la Encuesta Continua de Hogares del año 2006 para 13 áreas Metropolitanas. Ellos encontraron que los canales de búsqueda son fundamentales, para entender el riesgo de salida del desempleo, al

igual que lo son ser jefe de hogar, hombre con mayor experiencia, y tener un nivel educativo alto.

Muchos de los trabajos existentes en el medio académico nacional, como los mencionados previamente, se han centrado en la explicación de los determinantes y la duración del desempleo. Antes de presentar el estado del arte a nivel internacional, se hará un pequeño resumen de los principales hechos estilizados de este tema. En general estas investigaciones muestran que una mayor dispersión salarial en el sector buscado causan una mayor duración del desempleo, la diferencia en salarios de reserva afecta de forma positiva la diferencia entre las tasas de desempleo de hombres y mujeres, tanto hombres como mujeres casados presentan menores tasas que los solteros. Además que existe una mayor tasa de desempleo entre los jóvenes. Los jefes tienen una tasa de desempleo menor al igual que ser hombre. Para los hombres se encuentra que es menor el desempleo de los que tienen niños, mientras que las mujeres con niños presentan una mayor tasa de desempleo frente a las otras. El análisis del desempleo por años de educación sugiere que las personas con nivel de educación medio son las más afectadas. La educación aumenta el tiempo de desempleo hasta los 8 años de educación, luego lo disminuye. La experiencia también aumenta este tiempo. De todos estos estudios, se puede concluir que las mujeres y los jóvenes son los grupos más afectados por el desempleo.

A nivel internacional también se han realizado investigaciones en torno a este tema en países como España, Perú, Paraguay y Argentina, por mencionar algunos. Entre estos trabajos tenemos el de Ahn y Ugidos- Olazabal (1985) cuyo propósito es analizar la duración del desempleo juvenil en España e identificar los principales determinantes de dicha duración. En este estudio se utiliza la encuesta de condiciones de vida y trabajo (ECVT) de 1985, para estimar la función de supervivencia de Kaplan y Meier y para estimar unos modelos de riesgo. Los resultados más sobresalientes de esta investigación muestran que es importante distinguir entre hombres y mujeres y entre dos posibles transiciones: desempleo-empleo y desempleo-no participación. De esta manera los autores encuentran que el seguro de desempleo aumenta la duración principalmente por retrasar la salida del mercado de trabajo para mujeres y por retrasar el tener un empleo para hombres. De la misma manera la situación familiar es un importante factor de la duración del

desempleo de un individuo. Así, los hombres cabeza de familia son más rápidos en encontrar empleo que los que no lo son, mientras que las mujeres solteras encuentran trabajo más rápido que las casadas. Tener una educación universitaria aumenta de forma sustancial la probabilidad de encontrar empleo de las mujeres jóvenes pero no tiene un efecto significativo en la probabilidad de los hombres jóvenes.

Ureta (1996) analiza las dos dimensiones básicas del fenómeno de desempleo: incidencia y duración en las áreas urbanas del Paraguay, utilizando la encuesta anual de hogares a partir de 1996 y un método no paramétrico conocido como la regresión de Cox. Ella encuentra que el desempleo en las áreas urbanas está concentrado en el grupo de trabajadores más jóvenes. De acuerdo con esta investigación la edad del trabajador parece ser el principal determinante de la probabilidad de estar desempleado. Y los trabajadores más jóvenes sufren periodos de desempleo significativamente más cortos que los trabajadores mayores de 50 años.

Arranz, et al (2000a) en su trabajo analizan los factores determinantes de la duración del desempleo masculino en Argentina. Para lo cual estiman un modelo discreto de riesgos proporcionales mixtos, modelos discretos de la tasa de salida del paro y modelos semiparamétricos, tomando como base de datos la encuesta permanente de hogares de Argentina (EPH) 1997-1998. Esta investigación muestra como las condiciones locales del mercado de trabajo desempeñan un papel importante en la explicación de la magnitud de las tasas de salida del paro en dicho país. El análisis de los determinantes de la duración del desempleo en ausencia de un sistema de protección por desempleo confirma que la renta familiar se convierte en el único soporte de un individuo desocupado.

Arranz, et al (2000b) en otra investigación estudian los principales determinantes de la duración del desempleo en Argentina en presencia de altas tasas de paro. Como base de datos utiliza la información sobre varones desocupados, de edades comprendidas entre 18 y 65 años, procedente del enlace de las cuatro ondas de la encuesta permanente de hogares (EPH) del instituto de estadísticas y censos (INDEC) correspondiente a los años 1997 y 1998. La metodología que utilizan en esa investigación se basa en el enfoque discreto de los

modelos de forma reducida de la tasa de salida del desempleo de riesgos proporcionales mixtos.

Los autores encontraron que la distribución de las tasas de salida del desempleo por tramos de edad presentaba forma de U invertida con un máximo para el grupo de edad entre 36 y 40 años; así mismo, las responsabilidades familiares asociadas con la convivencia en pareja incrementan relativamente las tasas de salida de la desocupación. El nivel de estudios de los individuos condiciona de manera irregular las tasas de salida del paro, como la elevada tasa de los individuos con estudios primarios y la reducida de los universitarios con estudios finalizados. La tasa de salida se eleva para los individuos cuyo lugar de nacimiento es otro país distinto de la Argentina y la experiencia laboral reduce notablemente las tasas de salida con relación a la de los individuos sin experiencia laboral.

El trabajo de Cerimedo (2004) analiza la relación entre la tasa de salida del desempleo y el ciclo económico para Argentina. Para ello utiliza un modelo discreto para la probabilidad de que un individuo salga del desempleo en el transcurso del mes siguiente al mes de desempleo considerado, teniendo en cuenta que no ha encontrado trabajo en determinada cantidad de meses; tomando como base de datos la Encuesta Permanente de Hogares para el periodo mayo 1995 - mayo 2003. Esta investigación muestra que cuando la economía crece 4% la probabilidad de experimentar un tránsito del desempleo al empleo es 16.2% mayor que cuando se contrae a igual tasa. Este efecto no es homogéneo entre sectores. Se halla que la tasa de salida del desempleo es 31% mayor para las personas con último empleo en el sector industrial respecto a aquellas que trabajaron por última vez en el sector servicios. También encuentra que cuando la economía se expande, la probabilidad de experimentar un tránsito hacia el empleo se incrementa y viceversa.

A manera de conclusión es posible decir que la revisión a nivel nacional e internacional de los estudios sobre los determinantes de la duración del desempleo, resaltan valiosas consideraciones sobre el tema. La mayoría las investigaciones mencionadas encuentran que variables como la edad, estado civil, educación, ciudad, sector económico, posición ocupacional, número de dependientes, ingresos no laborales, inciden en la duración del desempleo, más específicamente en la probabilidad de salir de él. También encuentran

diferencias entre las tasas de desempleo y algunas evidencias que muestran cuáles hechos afectan más a las mujeres que a los hombres. Sin embargo aun existen aspectos que justifican profundizar más en el tema. Uno de estos aspectos corresponde al estudio detallado de estos determinantes a nivel nacional y no solo para las principales ciudades que es la información que otorga la Encuesta Continua de Hogares. Por lo tanto, en este nuevo estudio se identificarán los determinantes de la duración del desempleo para toda Colombia utilizando la Encuesta de Calidad de Vida que provee información a nivel nacional.

3. METODOLOGIA

En busca de cumplir los objetivos propuestos anteriormente, a continuación se plantea la metodología propuesta para explicar cómo se modelan los determinantes de la duración del desempleo en Colombia para el 2008.

Para el tratamiento de los datos y su procesamiento se cuenta con la información suministrada por la última Encuesta de Calidad de Vida que corresponde al año 2008. A continuación se presentan los planteamientos formales de los modelos de duración tanto paramétricos como no paramétricos, siguiendo como referencia el desarrollado por Greene (2008)

3.1 MODELO DE DURACIÓN

Sea T una variable aleatoria continua, correspondiente a la duración de la búsqueda, siendo $T = 0$ el instante del inicio de la búsqueda. La función de distribución acumulada de T , $F(t)$, se define así:

$$F(t) = \text{Prob}(T < t) \quad t \geq 0 \quad (8)$$

La función de densidad, $f(t)$ es:

$$f(t) = \frac{dF(t)}{dt} \quad (9)$$

La función de distribución acumulada $F(t)$ es la probabilidad de que la duración de la búsqueda no llegue a T días mientras que la función de densidad corresponde a la probabilidad incondicional de que la duración sea exactamente de T días. De (8) y (9) se tiene que:

$$F(t) = \int_0^t f(x) dx \quad (10)$$

Una función importante en el análisis de la duración, es la función de supervivencia $S(t)$, la cual corresponde al complemento de la función de distribución acumulada $F(t)$, es decir, la probabilidad de que la búsqueda sea al menos de t períodos.

$$S(t) = 1 - F(t) = \text{Prob}(T \geq t) \quad (11)$$

Esta función también se puede entender como la fracción de buscadores que duran buscando al menos t períodos.

En este punto surge una pregunta relevante: ¿dado que la duración de la búsqueda ha llegado hasta t días, cuál es la probabilidad de que ella finalice en el próximo pequeño intervalo de tiempo dt ? Dicha probabilidad se le conoce como la tasa de escape ó función de riesgo $\theta(t)$ (hazard rate en el lenguaje anglosajón). Formalmente se trata de:

$$\text{Prob}(t \leq T < t + dt | T \geq t) \quad (12)$$

Es la probabilidad de que una persona que ha estado desempleada hasta el momento t , encuentre ocupación en el pequeño intervalo dt , ó después de él. Dividiendo por dt y tomando el límite cuando dt tiende a cero la expresión de la tasa de escape es:

$$\theta(t) = \lim_{dt \rightarrow 0} \frac{\text{Prob}(t \leq T < t + dt | T \geq t)}{dt} \quad (13)$$

Puede verse $\theta(t)$ como la tasa de salida instantánea por unidad de tiempo en el instante t . Esta tasa responde la pregunta relevante anteriormente señalada. Existe una íntima conexión entre las funciones antes descritas. La probabilidad condicional de la expresión (12) corresponde al aplicar la correspondiente definición a:

$$\text{Prob}(t \leq T < t + dt | T \geq t) = \frac{\text{Prob}(t \leq T < t + dt, T \geq t)}{\text{Prob}(T \geq t)} = \frac{\text{Prob}(t \leq T < t + dt)}{\text{Prob}(T \geq t)} = \frac{F(t+dt) - F(t)}{1 - F(t)} \quad (14)$$

Para llegar a lo instantáneo se divide la expresión (14) entre dt y en consecuencia:

$$\text{Prob}(t \leq T < t + dt | T \geq t) = \frac{F(t+dt) - F(t)}{dt} \times \frac{1}{1 - F(t)}$$

Tomando de nuevo el límite cuando dt tiende a cero para volver a llegar a $\theta(t)$ se tiene que al ser $\lim_{dt \rightarrow 0} \frac{F(t+dt) - F(t)}{dt} = F'(t) = f(t)$, pues es la definición de derivada, se llega a:

$$\theta(t) = \frac{f(t)}{1-F(t)} = \frac{f(t)}{S(t)} \quad (15)$$

Es decir que la tasa de escape es el cociente entre la función de densidad de la duración y la función de supervivencia. Esto provee la intuición de que la tasa de escape son los que salen del desempleo en el instante t sobre los que aún continúan desempleados.

Adicionalmente al tener en cuenta la derivada respecto a t del logaritmo de la función de supervivencia $\frac{d \log S(t)}{dt} = \frac{1}{S(t)} \times \frac{dS(t)}{dt} = -\frac{f(t)}{S(t)}$

Se llega a que la tasa de escape es el negativo de la anterior derivada, al observar la expresión (15)

$$\theta(t) = -\frac{d \log S(t)}{dt} \quad (16)$$

La anterior expresión permite encontrar una conexión biunívoca entre supervivencia y tasa de escape, pues dicha expresión contiene la ecuación diferencial $-\theta(t)dt = d \log S(t)$. Cuya solución al integrar ambos lados lleva a:

$$\log S(t) = -\int_0^t \theta(x) dx + C \quad (17)$$

Puesto que al inicio todos sobreviven, $S(0)$ es igual a uno y $\log S(0) = 0$, de donde $C = 0$ ya que $\int_0^0 \theta(x) dx$ también es cero. Por lo tanto

$$\log S(t) = -\int_0^t \theta(x) dx \quad (18)$$

La integral del lado derecho se conoce como la “*hazard*” *integrada*. Al tomar antilogaritmos en la expresión (18) se llega a la relación buscada entre función de supervivencia y tasa de escape

$$S(t) = e^{-\int_0^t \theta(x) dx} \quad (19)$$

Dada una función de riesgo se tiene una de supervivencia, dice la expresión (19) y de acuerdo a la (16), dada una de supervivencia se tiene la función de riesgo.

Es el momento de relacionar el instrumental analítico hasta ahora presentado, con el modelo teórico de búsqueda desarrollado en el marco teórico. En esencia, dicho modelo postula que el individuo saldrá del estado del desempleo en el intervalo de tiempo $t, t + dt$, sí y sólo sí se dan dos eventos: recibir una oferta (con probabilidad λdt) y aceptarla dado que excede el salario de reserva (probabilidad $S(\rho V_u)$).

La anterior regla de decisión llega a que la tasa de escape que la teoría de búsqueda es:

$$\theta(t) = \lambda S(\rho V_u) \quad (20)$$

Al observar el lado derecho de la expresión anterior se evidencia que no interviene el tiempo, esto es, se ha llegado a una tasa de escape constante. ¿Cuáles son las implicaciones sobre las funciones relevantes de una tasa de escape constante e igual a θ ?

De la ecuación (18) la tasa de escape integrada sería: $\int_0^t \theta(x) dx = \int_0^t \theta dx = \theta t$ (21)

Al aplicar la ecuación (19) se obtiene la función de supervivencia $S(t) = e^{-\theta t}$ (22)

Al derivar (22) y cambiar signo se llega a la función de densidad $f(t) = \theta e^{-\theta t}$ (23)

La teoría de la búsqueda presentada anteriormente lleva a que la duración del desempleo siga una ley de probabilidad exponencial y a que la tasa de escape del mismo sea constante. Como consecuencia, la distribución exponencial no tiene memoria, lo cual contradice la observación de las duraciones ya que es conocido que la tasa de escape disminuye con el tiempo. Por esta razón, los estudios de la búsqueda han propuesto utilizar otras funciones de densidad para la duración. Adicionalmente la teoría económica ha centrado su interés en la tasa de escape como función de la duración y sus características y pocas veces se ocupa únicamente de la duración. También se ha demostrado que la tasa de arribo de ofertas puede declinar con el paso del tiempo con lo cual una tasa de escape constante se hace demasiado poco realista. Una propiedad interesante de la distribución exponencial es la expresión del valor esperado del logaritmo de la duración:

$$E(\log T) = -\log \theta + \psi(1) \quad (24)$$

$$\text{Var}(\log T) = \psi'(1) \quad (25)$$

Donde $\psi(1)$ y $\psi'(1)$ son constantes conocidas y corresponden a la función digamma (la derivada de la función gamma). De (24) y (25) se puede plantear que:

$$\log T = -\log \theta + U \quad (26)$$

Siendo U un error de media y varianza conocidas

$$\log(T\theta) = U \quad (27)$$

Un hecho bastante reconocido en la literatura del tema son las posibilidades de censura en las observaciones de los tiempos de búsqueda, los cuales ocurren cuando al medir la duración de algún individuo este todavía permanece en el estado inicial, sin haber transitado hacia otro estado. Este hecho es otro argumento en contra de la estimación por mínimos cuadrados ordinarios (MCO). Además, la teoría económica ha centrado su interés en la tasa de escape como función de la duración y sus características, pero esta tasa pocas veces se ocupa exclusivamente de dicho suceso. También los teóricos han resaltado que suponer una tasa de escape constante se hace demasiado poco realista debido a que el salario de reserva va cayendo normalmente a medida que pasa el tiempo de búsqueda y con él la probabilidad de aceptar una oferta, además, las empresas miran el tiempo como señal de las habilidades de los trabajadores o medida de la depreciación de su capital humano.

Por estas razones se ha modelado la tasa de escape y se ha propuesto para la estimación métodos de máxima verosimilitud.

3.2 LA FUNCION DE SUPERVIVENCIA DE KAPLAN-MEIER

Los modelos paramétricos, como el presentado anteriormente, son atractivos porque permiten ver cómo afectan las diferentes variables la duración del desempleo, pero imponen una estructura rígida sobre los datos, lo cual puede distorsionar el estimador de razón de fallo.

Por esta razón se pueden utilizar técnicas no paramétricas como el del estimador límite del producto de Kaplan-Meier, que permite calcular las funciones de razón de fallo y de supervivencia desde un punto de vista estrictamente empírico.

La función de supervivencia de Kaplan-Meier de duración t periodos es definida como la probabilidad de un individuo de permanecer desempleado al menos t periodos. Esta función permite comprender la ocurrencia de un evento (abandonar el desempleo) durante un intervalo de tiempo dado.

En primer lugar, se asume que el evento ocurre precisamente en el tiempo t_1, t_2, \dots, t_j . Además se puede considerar una variable aleatoria positiva T que representa el tiempo desde el momento en que el individuo comienza a buscar empleo. Esta variable aleatoria podría caracterizarse por su función de distribución $F(t)$ y una función de densidad de probabilidad $f(t)$, como las definidas en las ecuaciones (8) y (9).

La función de supervivencia, la cual puede considerarse como la distribución de la duración del desempleo de T , puede definirse exactamente igual que la ecuación (11), donde $S(t)$ representa la probabilidad de que un individuo sobreviva (no abandone el desempleo), al menos hasta el momento t .

La distribución de T puede representarse a través de la tasa de riesgo o tasa de salida, definida por la ecuación (13), donde el numerador representa la probabilidad condicional de que se experimente en el evento el intervalo temporal $(t, t+dt)$, dado que el evento llegue

hasta la duración t , y el denominador es el tamaño de ese intervalo. Del mismo modo, la tasa de riesgo se puede escribir como en la ecuación (15).

Con lo anterior, la función de supervivencia puede escribirse:

$$S(t) = \prod_{k=1}^t [1 - \theta(t)] \quad (28)$$

Esta ecuación muestra que se puede usar uno menos $S(t)$ para calcular la probabilidad de abandonar el desempleo cuando se tiene una duración específica.

La estimación no paramétrica de la función de supervivencia es:

$$\tilde{S}(t_j) = \prod_{k=1}^t [1 - \tilde{\theta}(t_j)] \quad (29)$$

Donde $\tilde{S}(t)$ y $\tilde{\theta}(t_j)$ son respectivamente la función de supervivencia y la tasa de riesgo estimada. Además $\tilde{\theta}$ puede obtenerse como:

$$\tilde{\theta}(t_j) = \frac{h_j}{n_j} \quad (30)$$

Donde h_j representa el número de individuos que experimentaron el evento de salir del desempleo y n_j es el número de individuos en riesgo en el tiempo t_j . $\tilde{\theta}(t_j)$ es el estimador no paramétrico de Kaplan-Meier (Kaplan y Meier, 1958).

4. RASGOS DESCRIPTIVOS DE LA DURACIÓN DEL DESEMPLEO EN COLOMBIA

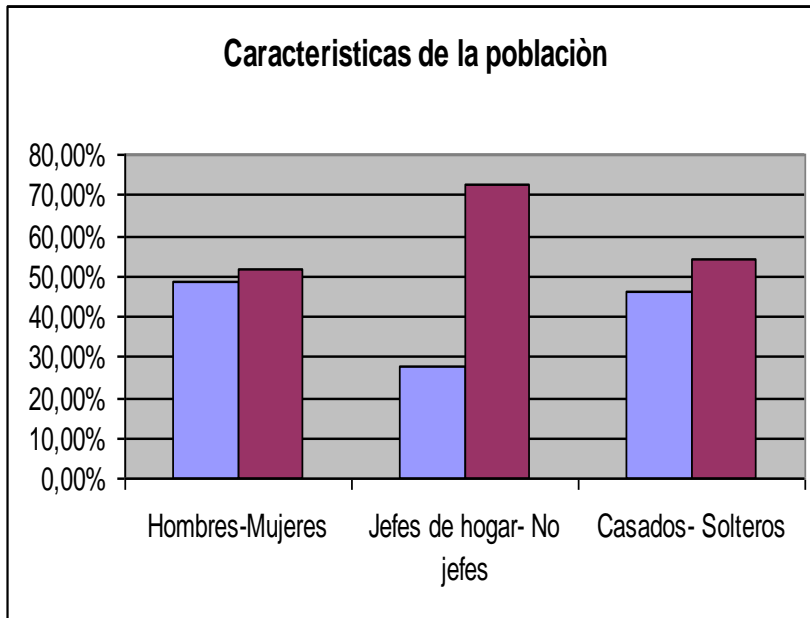
En este apartado se presenta una descripción de las características más relevantes asociadas con la duración del desempleo en Colombia. La muestra incluye 50.500 individuos que equivalen a una población de aproximadamente 44 millones de personas, de las cuales el 39%, unas 17 millones de personas para el 2008 pertenecían a la población económicamente activa.

La tasa de desempleo para esta matriz calculada como la proporción de la población desempleada sobre la población económicamente activa (PEA) era del 10,29%, es decir, para el 2008 teníamos aproximadamente 1.766.493 personas que se encontraban buscando empleo.

A continuación, se presenta una breve descripción de la muestra teniendo en cuenta los factores de expansión. Las Gráficas 1, 2, 3 y 4 proporcionan una imagen de un conjunto de características observadas en nuestra base de datos, estas son las variables personales y familiares siguientes: género, posición en el hogar, estado civil, región de residencia, nivel educativo y canal de búsqueda de empleo.

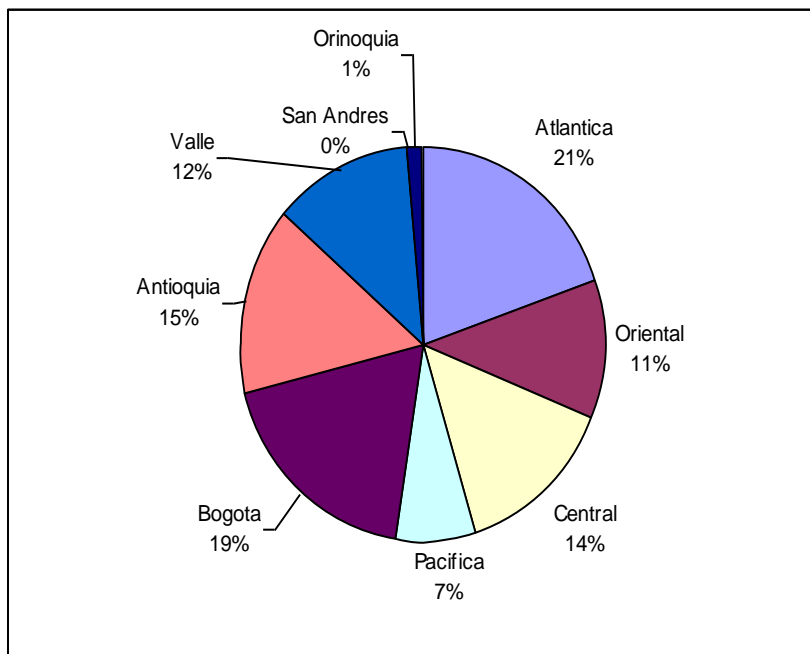
De las personas que se encontraban buscando empleo, el 48,55% eran hombres. El 27,39% eran jefes de hogar, mientras que el 46,02% eran casados. Esto nos muestra que la mayoría de los desempleados no constituye el proveedor principal en sus familias, y probablemente tienen ellos apoyo financiero de otros. El 21% de los desempleados vivían en la región Atlántica y el 19% en Bogotá, mientras que menos del 2% vivía en regiones como San Andrés y Orinoquía. Por otra parte, un 17,89% de esas personas buscaban empleo por primera vez.

Gráfica 1: Características de la población



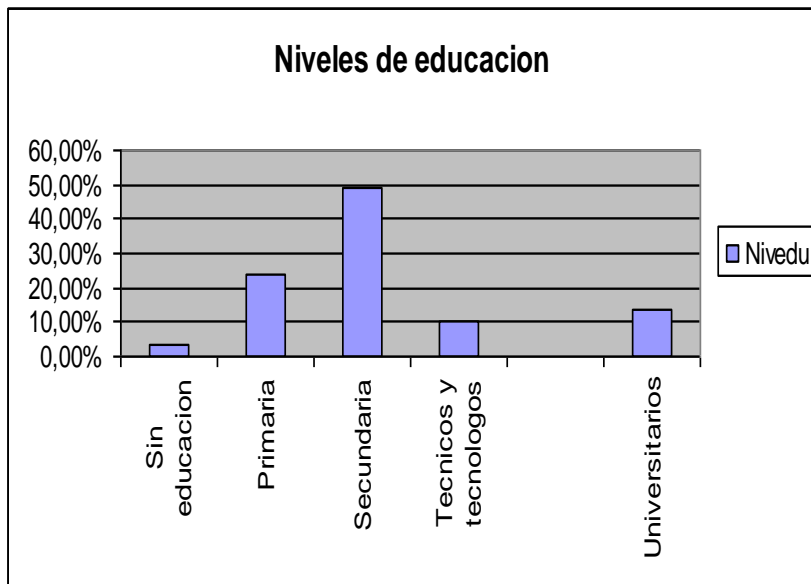
Fuente: ECV, segundo semestre 2008, cálculos propios

Gráfica 2: Región de residencia



Fuente: ECV, segundo semestre 2008, cálculos propios

Gráfica 3: Niveles de educación

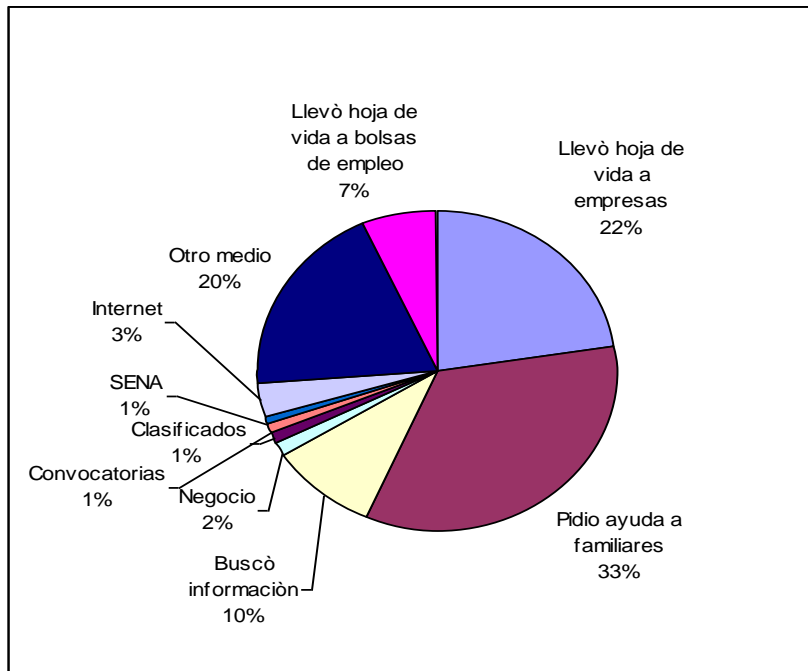


Fuente: ECV, segundo semestre 2008, cálculos propio

Al analizar los desempleados por nivel educativo, se encuentra que, el 3,38% no tenían ningún nivel de educación, el 24,09% solamente tenían primaria, el 49,05% tenían como nivel de educación secundaria, un 10,01% eran técnicos, mientras que un 13,47% eran universitarios.

Con respecto al uso de los canales de búsqueda, en la ECV se pregunta a los desempleados qué hizo en las últimas 4 semanas para conseguir empleo. En la gráfica 4 se puede apreciar la participación porcentual de los canales en la búsqueda ejercida por los desempleados del 2008. En ella es evidente que a la hora de buscar, los métodos predominantes son pedir información y ayuda a familiares, colegas o amigos (canales informales), pues de acuerdo con los datos, el 44,02% de los desempleados acuden a ellos a la hora de buscar empleo, mientras que menos del 34% emplean los canales formales. Llama la atención la gran participación del método llevar hojas de vida directamente a las empresas, el 22,41% de las personas acuden a él cuando buscan empleo.

Gráfico 4: Canales de búsqueda



Fuente: ECV, segundo semestre 2008, cálculos propios

En cuanto a la duración del desempleo, que es el tema de interés, en el cuadro 1 se presenta la función de supervivencia de la duración del desempleo, o sea, el valor estimado de la probabilidad que un periodo de desempleo dure más de X semanas. Por ejemplo la probabilidad de que el periodo de desempleo dure más de 2 semanas es de 92% y la probabilidad que dure más de 8 semanas es 67%. Aunque el valor esperado de la duración del desempleo es 22 semanas solo el 48% de los periodos de desempleo excederán dicho tiempo.

La variable duración, cuya tabla resumen se presenta en el anexo1, nos muestra que 67,53% de las personas que se encontraban buscando empleo salen del desempleo alrededor de las 13 semanas. El 79,19% de esas personas han salido del desempleo para las 26 semanas (6 meses). El 85,19% sale alrededor de las 39 semanas. Ya para el año, para las 52 semanas, ha salido el 92, 47% y para los 2 años, 104 semanas, han salido el 96,39%. Esta variable nos muestra como la tasa de salida del desempleo es mayor para los primeros 12 meses y de ahí en adelante se hace más lenta llevando a una tasa de desempleo de larga duración.

Cuadro 1

Función de supervivencia estimada

(Probabilidad de que un evento de desempleo dure X o más semanas)

Semanas	Función de Supervivencia	Semanas	Función de Supervivencia
2	92,77%	65	26,04%
4	80,53%	80	23,93%
6	74,84%	90	23,92%
8	67,40%	96	23,36%
13	59,41%	104	21,65%
22	52,17%	208	11,65%
26	48,34%	306	8,90%
39	41,08%	400	5,76%
52	29,12%	520	1,24%

Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts list

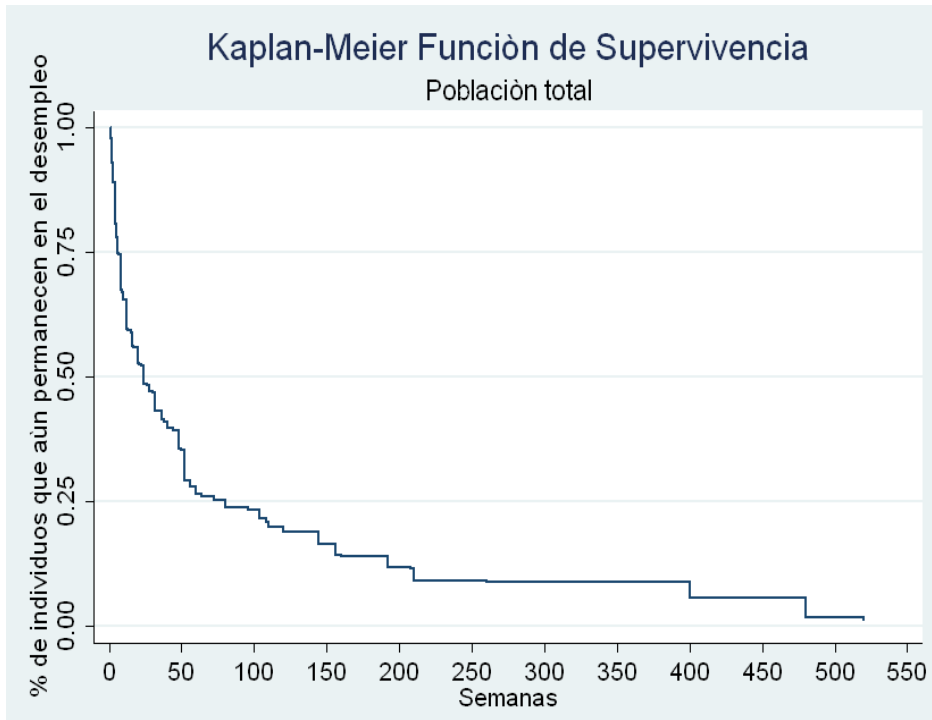
5. LA DURACIÓN DEL DESEMPLEO: ANÁLISIS NO PARAMÉTRICO

A continuación se realizará un análisis descriptivo de la duración del desempleo. Para ello se estiman funciones de supervivencia a través de la fórmula propuesta por Kaplan y Meier (1958), en ellas la probabilidad de supervivencia en el desempleo, o el número de personas que se encuentran aún en el desempleo después de determinado tiempo aparece en el eje de las ordenadas y el tiempo medido en semanas en el eje de las abscisas. Los resultados presentados en las gráficas se deben interpretar de la siguiente manera: entre más pegada la línea de cada grupo de discriminación al eje horizontal, mayor duración en el desempleo. Los indicadores resumen de la tabla de vida se presentan en el Anexo 2. En el Anexo 3 se calcularon los tests de Log-Rank y Wilcoxon para examinar la igualdad de las funciones de supervivencia y determinar si hay una diferencia significativa ($P < 0.05$) entre las curvas de supervivencia (ver, Hosmer y Lemeshow, 1999). Ambos, el Log-Rank test y Wilcoxon test, rechazan la hipótesis de equidad para las funciones de supervivencia.

La Gráfica 5 muestra la función de supervivencia estimada para los que se encontraban desocupados en el momento de la encuesta. En primer lugar se observa que el 25% de los desempleados, consiguieron empleo en un período inferior a un trimestre (desempleo de corta duración). En segundo lugar, en un período de 6 meses (unas 26 semanas), el 50% de los desempleados han conseguido empleo (desempleo de corta y mediana duración). Un 40% sale del desempleo en los siguientes 6 meses, es decir, al primer año de búsqueda. El otro 10% no consigue empleo en un período inferior a un año y por lo tanto se han convertido en desempleados de larga duración.

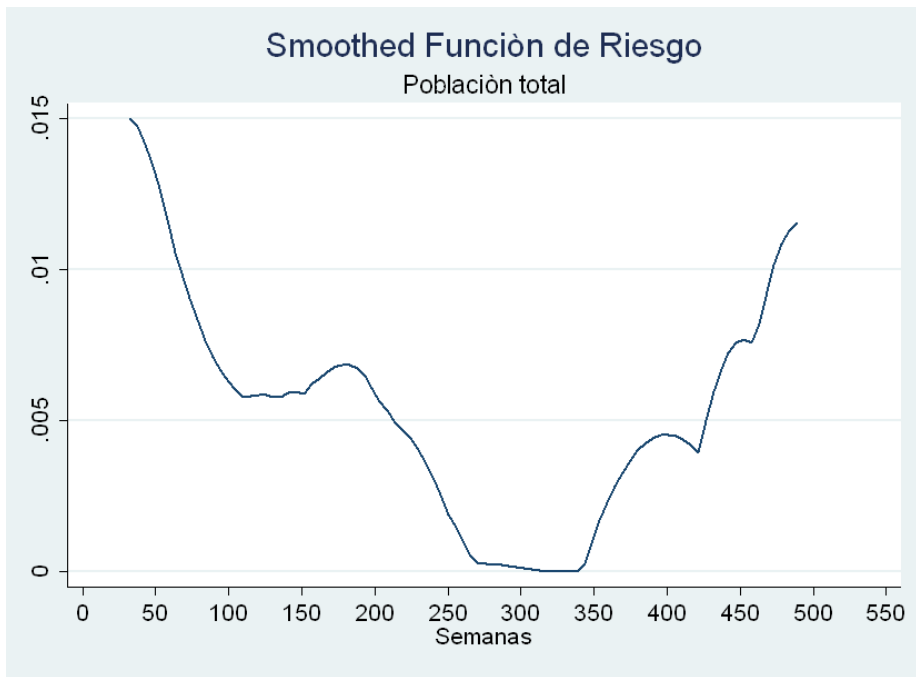
El gráfico 6 muestra la función de riesgo que surge de estos datos, es decir, nos está indicando el potencial instantáneo de salir del desempleo a medida que transcurre el tiempo. Esta función no tiene ninguna forma creciente o decreciente. Lo que nos lleva a buscar una distribución que cumpla con esta característica.

Gráfica 5: población total



Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

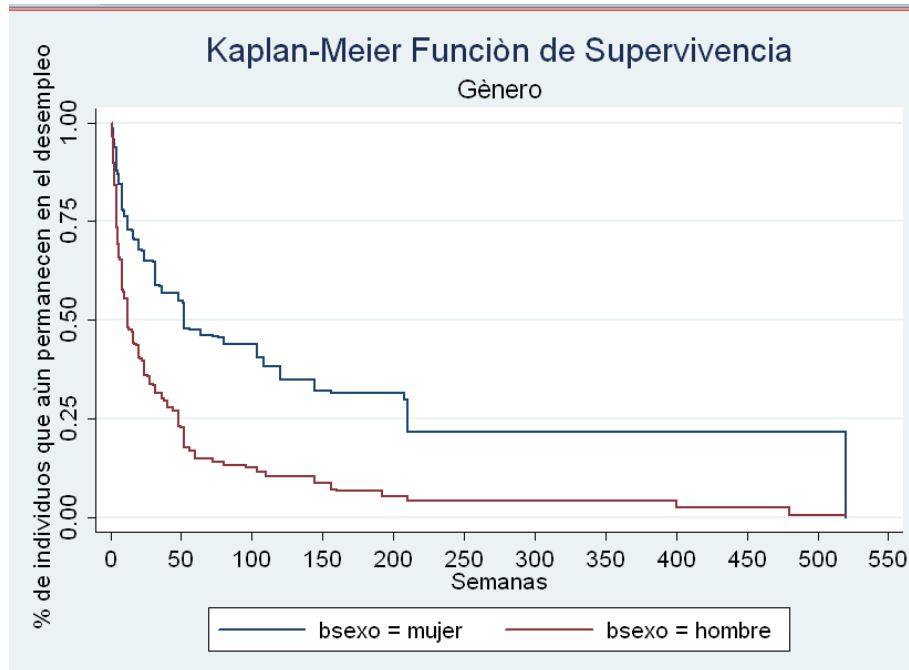
Gráfica 6: función de riesgo



Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

Realizando el análisis por género, las funciones de supervivencia indican que hay grandes diferencias entre hombres y mujeres en las probabilidades de abandonar el desempleo. Mientras que aproximadamente el 50% de los hombres consigue empleo a las 12 semanas, las mujeres necesitan 52 semanas para alcanzar tal proporción. La existencia de diferencias entre ambos sexos puede deberse, como lo explica Cerimedo (2004), a dos cuestiones, por un lado, los varones suelen tener un salario de reserva menor que las mujeres debido a que, en general, son el sostén del hogar, por lo que se ven obligados a aceptar las primeras propuestas laborales que reciben. Por otra parte, es posible que exista alguna forma de discriminación en contra de las mujeres lo que determine que los varones tengan una mayor cantidad de alternativas laborales. Hecho que corrobora una vez más las desventajas que enfrenta la mujer en el mercado laboral colombiano. Este resultado es soportado por los test tipo COX (los tests de Log-Rank y Wilcoxon) que se encuentran en el anexo 3.

Gráfica 7: género

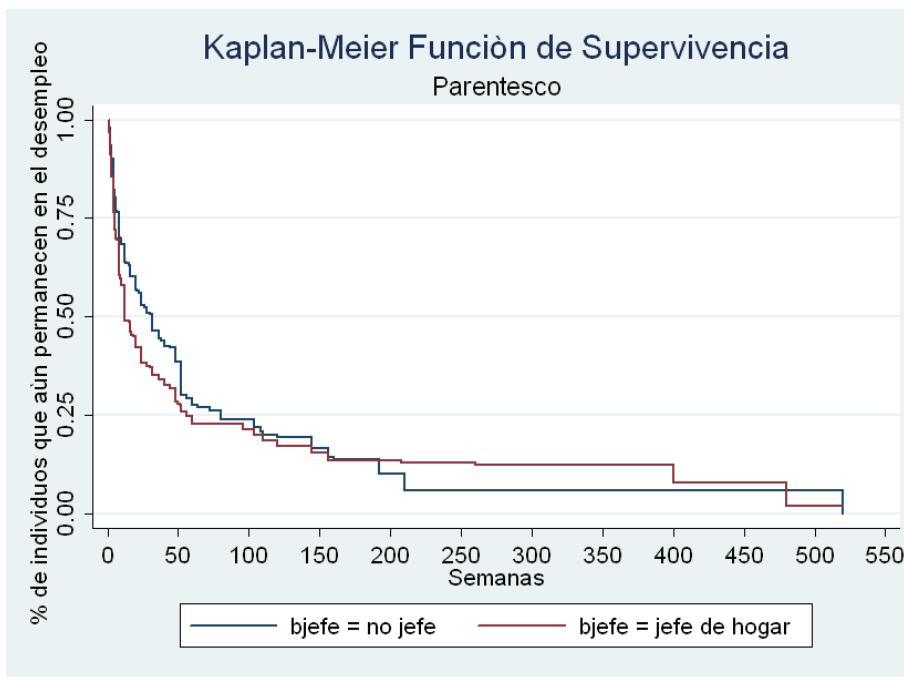


Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

Otra variable importante es la relación de parentesco de la persona en el hogar, esta permite valorar la situación familiar de los individuos y su efecto en la actividad laboral. Mientras

que aproximadamente el 50% de los jefes de hogar abandona el desempleo alrededor de las 12 semanas, los no jefes tardan 32 semanas en promedio. Este resultado está asociado con que los jefes de hogar por sus obligaciones familiares tienen que hacer una búsqueda más activa y también aceptar empleos de inferior calidad. Para los que no son cabeza de familia la situación laboral del cabeza de familia puede afectar su probabilidad de salir del desempleo. En un modelo de oferta de trabajo familiar, en el que las decisiones se toman conjuntamente, otros miembros de la familia tienen una menor probabilidad de trabajar si el cabeza de familia está empleado (Ahn y Ugidos- Olazabal, 1985).

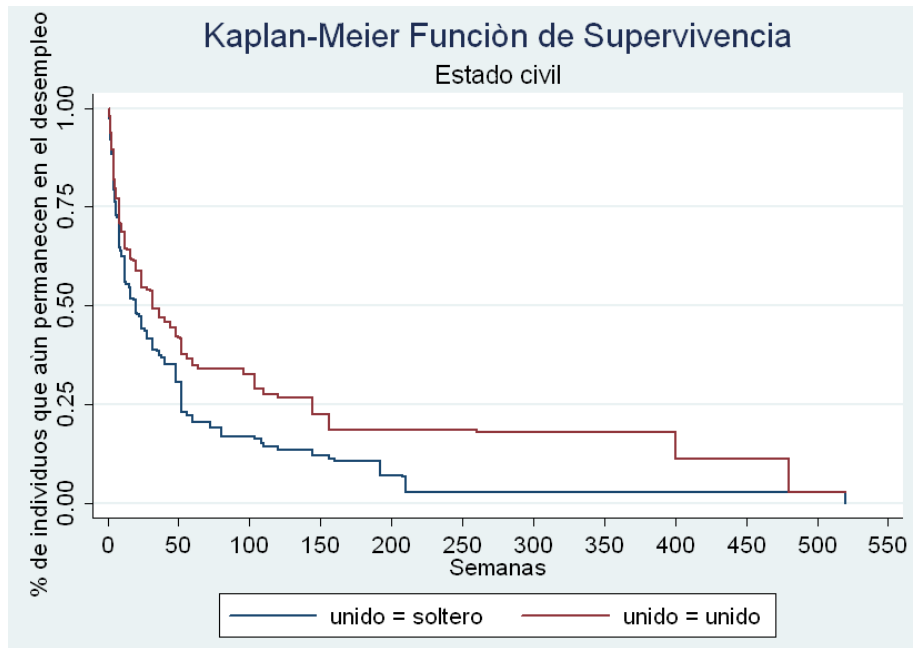
Gráfica 8: parentesco



Fuente: ECV, segundo semestre 2008, calculos propios, Stata 11, comando: sts graph

Con respecto al estado civil de los buscadores de empleo, las funciones de supervivencia indican que existen algunas diferencias entre casados y solteros en las probabilidades de abandonar el desempleo como se observa en la gráfica 9. Mientras que aproximadamente el 50% de los solteros sale del desempleo a las 20 semanas, los casados necesitan alrededor de 32 semanas para alcanzar tal proporción.

Gráfica 9: estado civil



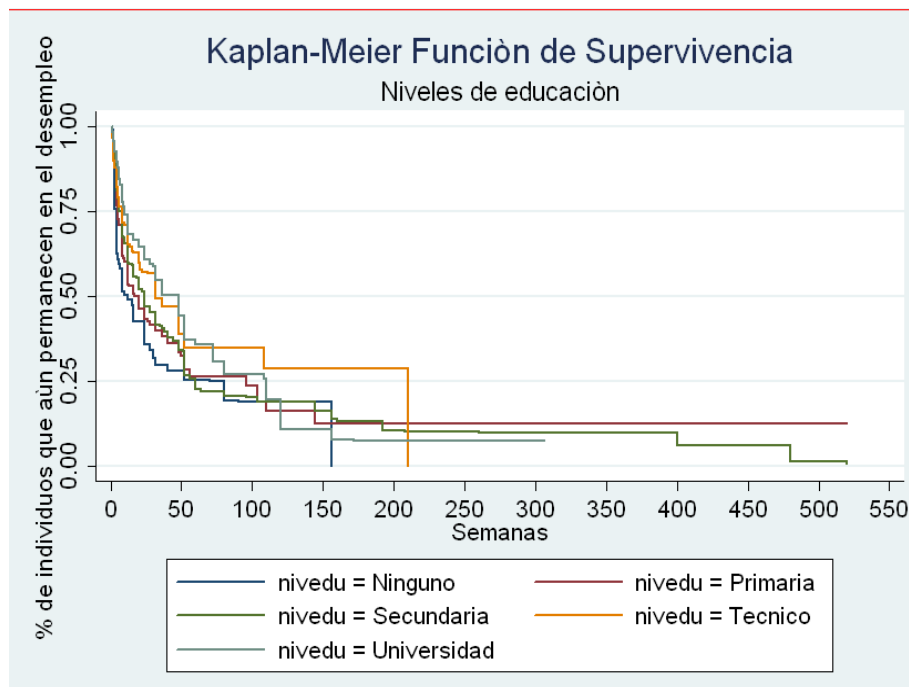
Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

Muchos de los estudios realizados tanto a nivel nacional como internacional encuentran una fuerte relación negativa entre educación y duración del desempleo. Nuestros resultados, así como los encontrados por Martínez (2003) y Nuñez y Bernal (1998), muestran un patrón diferente debido a que al igual que estos autores, sólo contamos con la información de los desocupados de la ECV. Según el nivel de educación, representados en la gráfica 10, se observa una diferencia apreciable entre las personas sin ningún año de educación y los demás grupos. El grupo que ostenta la mayor duración del desempleo está asociado a los años de educación más altos (universidad). Estas personas se demoran más tiempo en abandonar el desempleo, aunque su función de supervivencia se trunca alrededor de las 300 semanas. Esto sugiere que las personas con mayor nivel educativo podrían ser más selectivas a la hora de aceptar un empleo que aquellos menos educados.

La mayor probabilidad de salir del desempleo que tienen los individuos de menor nivel educativo seguramente está relacionada con el hecho de que estas personas están dispuestas a trabajar en cualquier empleo por lo que encuentran trabajo más fácilmente, disminuyendo

así su tasa de desempleo. Mientras que las personas con alto nivel de educación tienen mayores oportunidades de encontrar un buen empleo debido a la mayor selectividad en la selección de empleos que se deriva del salario de reserva más elevado que suelen tener.

Gráfica 10: años de educación



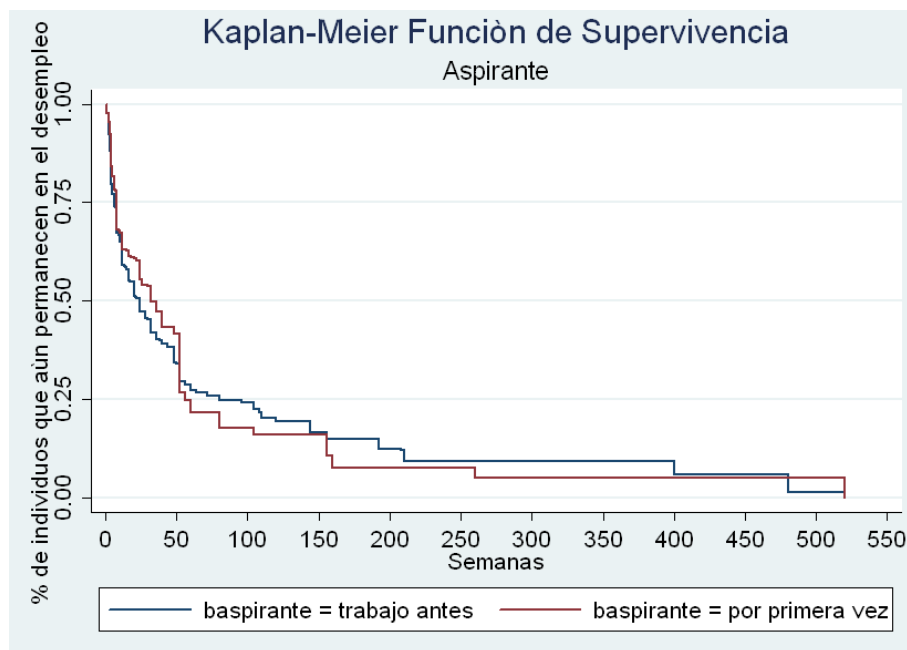
Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

Otra variable que resulta interesante mirar en este análisis es aquella que nos muestra si los individuos están en búsqueda de su primer empleo o si ya han trabajado antes. Dicha variable esta representada en la gráfica 11. La diferencia entre ambos grupos no es tan notoria, mientras que el 50% de aquellas personas que han trabajado antes abandonan el desempleo alrededor de las 24 semanas, los que buscan empleo por primera vez tardan 32 semanas en promedio. Este resultado puede estar asociado con el hecho de que la mayoría de los puestos de trabajo exigen una experiencia.

El comportamiento inicial de la función de supervivencia varía significativamente por canal de búsqueda de empleo como se observa en la Gráfica 12. Se encuentra que quienes escogieron canales de búsqueda formales son los individuos más propensos a sobrevivir en

el desempleo. Mientras que aquellos que se acogieron a canales informales e informal moderado resultan tener una probabilidad sistemáticamente más pequeña de permanecer en el desempleo, destacando que la probabilidad de emplearse resulta mayor para los usuarios de los métodos informales. La duración mediana del desempleo (probabilidad de supervivencia igual a 0,5) para estos últimos se encuentra en 17 semanas. Es necesario tener en cuenta que, aunque las redes sociales permiten conseguir empleo, los empleos que se encuentran a través de estos canales son, normalmente, informales (Uribe et al., 2007).

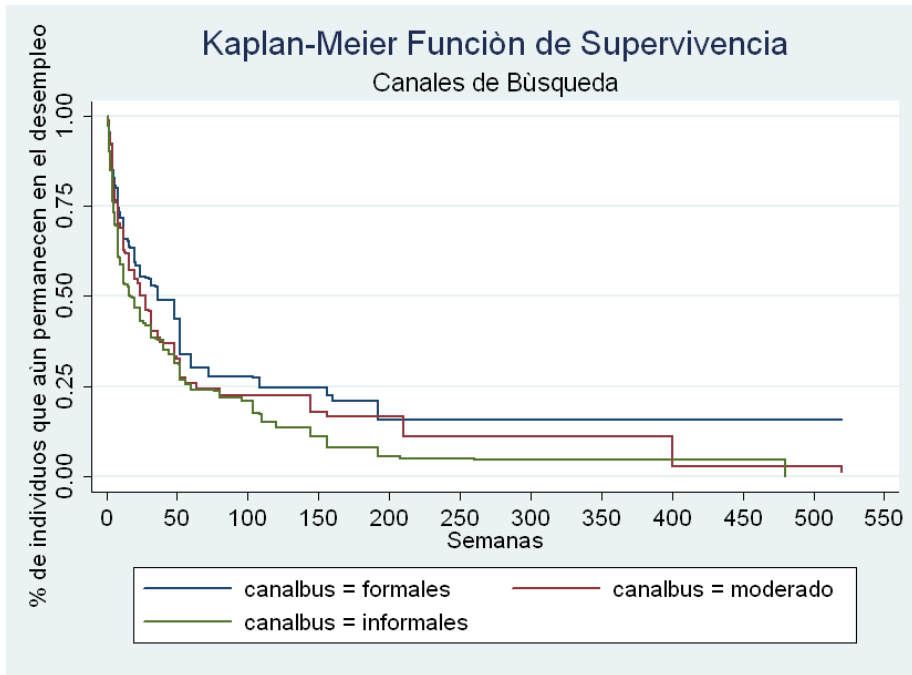
Gráfica 11: aspirante



Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

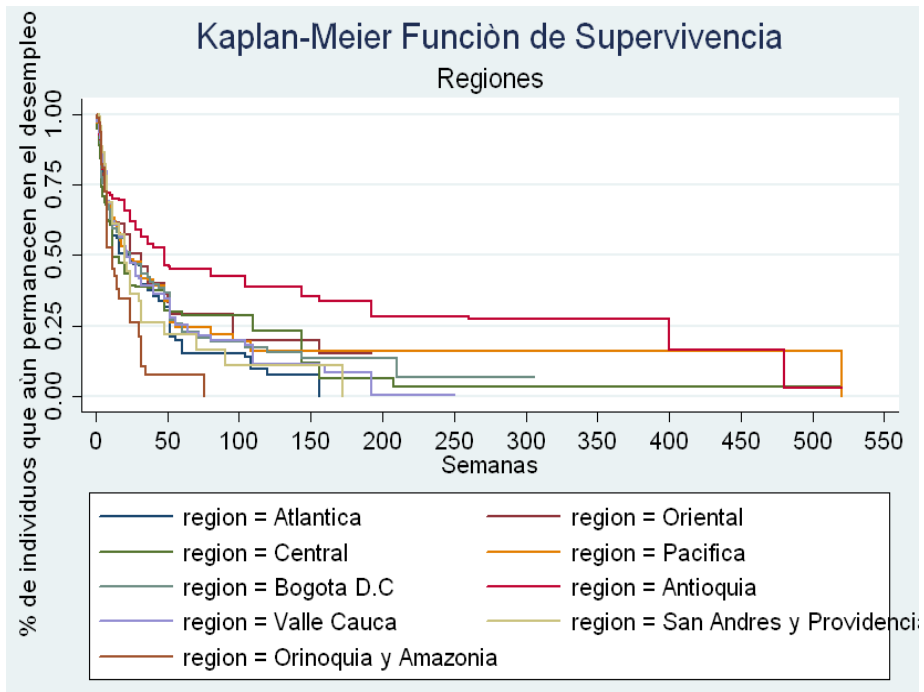
Resulta interesante observar también como es el comportamiento de la función de supervivencia de acuerdo a la región de residencia. La gráfica 13 nos muestra que quienes viven en la Orinoquía y Amazonía o en la región Central salen del desempleo alrededor de las 12 semanas, seguidos por quienes habitan en San Andrés y el Valle del Cauca que lo hacen alrededor de las 20 semanas y las 22 semanas respectivamente. Mientras que quienes viven en regiones como la Pacífica, Bogotá, la región Atlántica, la Oriental y Antioquia tienen que esperar 24, 32 y hasta 48 semanas para alcanzar tal proporción.

Gráfica 12: canal de búsqueda



Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

Gráfica 13: región de residencia



Fuente: ECV, segundo semestre 2008, cálculos propios, Stata 11, comando: sts graph

6. DETERMINANTES DE LA DURACIÓN DEL DESEMPLEO

Las variables utilizadas para estimar los determinantes de la duración del desempleo son variables de tipo sociodemográfico que se encuentran en la ECV. Como variables personales y familiares están: sexo, nivel educativo, responsabilidades familiares derivadas de la posición dentro de la familia, entre otras. Como variables económicas: la experiencia laboral previa y el ingreso mensual del hogar.

A continuación se expondrán cada uno de los determinantes que han sido propuestos para ser sometidos a la evidencia empírica. El ser jefe de hogar se supone disminuye el salario de reserva y por ende su efecto en el tiempo de búsqueda será negativo. La variable BJEFE es una variable ficticia que toma el valor de 1 para los jefes de hogar y 0 en caso contrario. Igual acontece con el género masculino. Como se mencionó anteriormente y de acuerdo con Cerimedo (2004) esto puede deberse a dos cuestiones, por un lado los hombres tienen un menor salario de reserva y por otro puede existir algún tipo de discriminación en contra de las mujeres. Si se define un binario como BSEXO, igual a 1 para los hombres y 0 para las mujeres, se espera un impacto negativo en la duración.

Otra variable que se incluye dentro del modelo es la que nos muestra si el individuo busca empleo por primera vez o si ha trabajado antes, la cual ha sido definida como BASPIRANTE, que toma el valor de 1 para los aspirantes y 0 para los demás. La cuarta variable que se incluye vía salario de reserva son los ingresos del hogar, LOGING, los cuales aumentan dicho salario pues permiten financiar la búsqueda y de ahí se espera su efecto positivo. También se ha incluido una variable que refleja cual es el estado civil de los individuos buscadores, la variable UNIDO toma el valor de 1 para los casados o en unión libre y 0 si es viudo, soltero, separado-divorciado. Se espera de acuerdo con Arranz, et al (2000a) que esta variable tenga una relación negativa con la variable dependiente debido al hecho de que las responsabilidades familiares incrementan las tasas de salida vía el aumento de los esfuerzos de búsqueda.

Por el lado del capital humano se han tomado la educación y la experiencia como regresores. De los años aprobados de educación formal, ANOSEDU, se ha postulado que los primeros años disminuyen el salario ofrecido y a partir de un punto lo aumenta, es decir, existen rendimientos marginales crecientes en la búsqueda. Esto debido a que en los primeros años de educación el individuo valora más cada año de educación que lo que lo hace el mercado y a partir de un punto el mercado valora más que el individuo (Uribe y Viáfara, 2009). Con relación a la experiencia potencial, EXPER, aproximada como la edad (EDAD) menos ANOSEDU menos 6, se ha supuesto un efecto positivo en el salario ofrecido que disminuirá el tiempo de búsqueda. En la interpretación de la inclusión de la experiencia es importante anotar que se trata de la potencial pues se supone que no hay periodos de desempleo y que es homogénea.

Basándonos en Tenjo (1998), en este trabajo supondremos que la duración del desempleo esta asociada con dos elementos. El primer grupo de variables, presentadas anteriormente, hacen parte del primer elemento al que llamaremos *efecto aspiraciones*, el cual capta la influencia de características personales y del nivel de capital humano de los individuos, lo cual a manera de hipótesis definiría las aspiraciones que los trabajadores pueden tener en el mercado laboral. Este efecto tiene en cuenta, como se puede observar, tanto elementos del lado de la oferta laboral determinantes del salario de reserva como elementos del lado de la demanda reflejados en el salario potencial de mercado.

El segundo elemento es el *efecto oportunidades*, efecto que está asociado con el acceso a la información por parte de los individuos pero también del capital humano, el cual depende de las redes relacionales en la sociedad. De acuerdo con sus características, los individuos definirán el canal de búsqueda a utilizar. Este efecto ha sido incluido dentro del modelo a través del canal de búsqueda utilizado por el individuo para encontrar empleo. Los canales formales, bolsas de empleo, empresas de trabajo temporal, cooperativas de trabajo asociado, convocatorias, clasificados y medios reconocidos se codifican como 1. Los canales informales, referencias personales (familiares, colegas, amigos), búsqueda directa y emprender su propio trabajo como 2 y el canal informal moderado, llevar hoja de vida

directamente a las empresas como 3. Con estas variables se construyeron dos dummies CANAL2 y CANAL3 en las cuales la categoría de referencia es el canal formal. La hipótesis es que los canales de búsqueda más eficientes, es decir, los formales y el informal moderado, producen duraciones del desempleo inferiores.

También se tendrá en cuenta dentro del análisis la región de residencia, porque es necesario conocer que ocurre con la duración del desempleo cuando existen mercados segmentados. Veremos en cuales regiones el salir del desempleo toma más tiempo. Para ello se ha dividido el país en nueve regiones de la siguiente manera: Atlántica codificada como 1, Oriental como 2, Central como 3, Pacífica como 4, Bogotá como 5, Antioquia como 6, Valle como 7, San Andrés como 8 y finalmente Orinoquía como 9. Con estas variables se construyeron 8 dummies REGGION1, REGGION2, REGGION3, REGGION4, REGGION6, REGGION7, REGGION8 Y REGGION9. Se tomo como categoría de referencia a Bogotá.

De estos elementos en conjunto dependerá tanto el tiempo de búsqueda cómo el tipo de empleos al que se pueden aspirar.

7. LA DURACIÓN DEL DESEMPLEO: ANÁLISIS PARAMÉTRICO

En esta sección se presenta primero una prueba para determinar cuál es la mejor distribución que explica los datos, para después presentar los resultados de las estimaciones. Las seis distribuciones utilizadas en este artículo son: Exponencial, Log-logística, Log-normal, Weibull, Gamma Generalizada y Gompertz. Para encontrar cuál es la mejor distribución que debe ser aplicada al modelo de duración se realizó la prueba de Akaike (AIC), la cual está basada en el nivel de ajuste del modelo. AIC es un estimador no sesgado del parámetro de contenido de información de Kullback – Leibler el cual es una medida de la información perdida al usar un modelo para aproximar la realidad, por tanto, a menor valor de AIC mejor ajuste del modelo a los datos. Al penalizar por cada parámetro adicional tiene en cuenta tanto la bondad de ajuste como la varianza asociada con la estimación de los parámetros (Vinuesa, 2008). Akaike (1974) propuso ese criterio que se puede definir como:

$$AIC = -2(\log \text{verosimilitud}) + 2(c+p+1) \quad (37)$$

donde \log verosimilitud es el logaritmo de la función de verosimilitud del modelo estimado, c es el número de variables de que consta el modelo y p es el número de parámetros auxiliares. Como puede observarse este criterio consta de dos componentes, uno que se refiere a la minimización de los errores y otro que es un término de penalización por la incorporación de parámetros adicionales. Se debería elegir entonces aquel modelo que presenta el menor valor del criterio. En el cuadro 2 se muestran las pruebas realizadas a las seis distribuciones.

El criterio de AIC se basa en el valor de máximo verosimilitud, el cual nos da una medida del ajuste del modelo ponderado por el número de variables; entre más pequeño mejor. Hay ligeras diferencias en el valor de la función de verosimilitud entre los modelos Weibull y Gamma Generalizado, en favor de éste último. Siguiendo el AIC, el mejor modelo es el

Gamma Generalizado. Esto último se puede corroborar al observar la gráfica 6 de la página 28, donde claramente la función de riesgo (suavizada) no es monótona a diferencia de la que pertenezca a una distribución Weibull.

Cuadro 2.
Prueba de AIC

Distribución	Log Likelihood	AIC
Exponencial	-786,78728	1611,575
Weibull	-769,91885	1579,838
Gamma	-716,22679	1474,454 *
Log normal	-726,8563	1493,713
Log Logística	-730,63446	1501,269
Gompertz	-752,03638	1544,073

Fuente: ECV, agosto-octubre 2008, cálculos propios, Stata 11, comando: estat ic

La función de supervivencia y la función de densidad de la distribución Gamma se presentan a continuación:

$$S(t) = \begin{cases} 1 - I(\gamma, u), & \text{si } \kappa > 0 \\ 1 - \Phi(z), & \text{si } \kappa = 0 \\ I(\gamma, u), & \text{si } \kappa < 0 \end{cases} \quad (38)$$

$$f(t) = \begin{cases} \frac{\gamma^\gamma}{\sigma t \sqrt{\gamma} \Gamma(\gamma)} \exp(z\sqrt{\gamma} - u), & \text{si } \kappa \neq 0 \\ \frac{1}{\sigma t \sqrt{2\pi}} \exp(-z^2/2), & \text{si } \kappa = 0 \end{cases} \quad (39)$$

Donde $\gamma = |\kappa|^{-2}$, $z = \text{sign}(\kappa)\{\log(t) - \mu\}/\sigma$, $u = \gamma \exp(|\kappa|z)$, $\Phi(z)$ es la función de distribución acumulada.

Afortunadamente, los modelos Gamma Generalizado y Weibull están anidados, esto debido a que la función de riesgo gamma es muy flexible y puede acomodar muchas formas incluyendo como casos especiales el modelo Weibull cuando $k=1$, el exponencial cuando $k=1$ y $\sigma=1$, y el modelo log-normal cuando $k=0$. Así, la estimación de uno de los

parámetros auxiliares k para los modelos con distribución Gamma Generalizada permitirá rechazar la hipótesis de $k=0$ (no es apropiada la distribución log-normal) y de $k=1$ (se rechaza un modelo Weibull). En este caso como se puede observar en el anexo 4 las pruebas permiten rechazar la distribución log-normal y aceptar la distribución Weibull para modelar los datos. La distribución Gamma es usada principalmente para evaluar y ayudar a seleccionar el modelo paramétrico más adecuado, que para este caso es él mismo porque como vimos anteriormente es el que tiene el menor valor de AIC.

Con lo anterior, queremos decir que en este trabajo se utilizara el modelo Gamma debido a que este modelo, de acuerdo con el criterio AIC, es el modelo que mejor se aproxima a la realidad, es decir, es el modelo que realiza el mejor ajuste de los datos.

El modelo a contrastar es:

$$T = \text{Gamma}(bjefe, unido, bsex, baspirante, loging, anosedu, exper, exper2, canal2, canal3, regionn2, regionn3, regionn4, regionn5, regionn6, regionn7, regionn8, regionn9) \quad (40)$$

Donde T representa la duración del desempleo.

La interpretación de los coeficientes resulta en cierto modo compleja, dada la naturaleza discreta de casi todas las variables, pero si se pueden interpretar los signos de los coeficientes estimados. Si el signo de la estimación es positivo, esa variable aumenta la duración de los individuos en la situación de desempleo. En caso contrario, la duración se reduce.

El modelo estimado se presenta en el cuadro 3. La regresión se realiza teniendo en cuenta la función de riesgo acelerada, las censuras de la variable y errores estándar robustos. En el anexo 5 se presentan cada una de las salidas de Stata de las diferentes distribuciones.

Cuadro 3
Estimación del modelo de supervivencia paramétrica para
la duración del desempleo en Colombia 2008

Variables		Gamma	Exponencial	Weibull	Log normal	Log logistica	Gompertz
CONSTANTE	Coef	4,19	-5,87	-5,046	5,20	5,19	-5,65
	P>z	0,000	0,000	0,000	0,000	0,000	0,000
	E.E robusto	0,6536703	0,8717902	0,755552	0,6549005	0,6828937	0,641601
BJEFE	Coef	0,05	-0,05	0,01	0,004	-0,008	-0,01
	P>z	0,772	0,803	0,996	0,980	0,961	0,921
	E.E robusto	0,1286585	0,1986252	0,155817	0,1512261	0,1589804	0,1504551
UNIDO	Coef	0,04	-0,14	-0,16	0,13	0,14	-0,13
	P>z	0,734	0,426	0,277	0,332	0,345	0,332
	E.E robusto	0,1275447	0,1788151	0,14395	0,1382735	0,1435258	0,1373707
BSEXO	Coef	-0,55	1,05	0,95	-0,88	-0,97	0,99
	P>z	0,005	0,000	0,000	0,000	0,000	0,000
	E.E robusto	0,1943738	0,188048	0,152102	0,1362151	0,1456411	0,1428744
BASPIRANTE	Coef	0,85	-0,62	-0,54	0,79	0,76	-0,52
	P>z	0,000	0,024	0,022	0,001	0,002	0,021
	E.E robusto	0,1985699	0,2742314	0,236684	0,2367665	0,2452069	0,2265544
LOGING	Coef	-0,25	0,28	0,24	-0,29	-0,30	0,27
	P>z	0,000	0,000	0,000	0,000	0,000	0,000
	E.E robusto	0,0453557	0,0563149	0,049283	0,0504147	0,0538288	0,0449454
ANOSEDU	Coef	0,08	-0,16	-0,13	0,12	0,13	-0,12
	P>z	0,001	0,000	0,000	0,000	0,000	0,000
	E.E robusto	0,0239812	0,022176	0,01963	0,0186738	0,019859	0,0187067
EXPER	Coef	0,03	-0,05	-0,04	0,04	0,04	-0,03
	P>z	0,018	0,005	0,014	0,012	0,019	0,025
	E.E robusto	0,0125764	0,0174678	0,014775	0,0145234	0,0155773	0,0145103
EXPER2	Coef	-0,0002	0,0003	0,0002	-0,0002	-0,0002	0,0002
	P>z	0,208	0,234	0,385	0,349	0,435	0,444
	E.E robusto	0,0001846	0,0002726	0,000228	0,000227	0,0002479	0,0002248
CANAL2	Coef	-0,65	0,69	0,65	-0,72	-0,72	0,67
	P>z	0,000	0,001	0,000	0,000	0,000	0,000
	E.E robusto	0,1502112	0,1983958	0,171161	0,1662597	0,1743956	0,1675042
CANAL3	Coef	-0,39	0,19	0,26	-0,39	-0,42	0,45
	P>z	0,03	0,569	0,321	0,077	0,070	0,050
	E.E robusto	0,1782668	0,3269212	0,263796	0,2188941	0,2326054	0,2275714
REGIONN1	Coef	0,56	0,37	0,18	0,38	0,50	0,05
	P>z	0,034	0,333	0,586	0,236	0,173	0,873
	E.E robusto	0,2641372	0,3851342	0,332688	0,324291	0,3646234	0,3269665

REGIONN2	Coef	0,23	0,63	0,39	0,02	0,10	0,32
	P>z	0,39	0,17	0,322	0,955	0,796	0,399
	E.E robusto	0,2632897	0,4565881	0,389604	0,3530648	0,3840281	0,3774392
REGIONN3	Coef	0,10	0,55	0,40	-0,08	-0,01	0,39
	P>z	0,66	0,193	0,259	0,791	0,974	0,299
	E.E robusto	0,238484	0,4204494	0,353142	0,3063758	0,3405452	0,3211062
REGIONN4	Coef	0,17	0,49	0,30	0,13	0,31	0,23
	P>z	0,521	0,219	0,372	0,692	0,385	0,487
	E.E robusto	0,2654754	0,3997848	0,338099	0,3198148	0,3526603	0,3257819
REGIONN6	Coef	0,42	0,10	0,02	0,34	0,43	-0,01
	P>z	0,1	0,81	0,955	0,318	0,25	0,972
	E.E robusto	0,2578432	0,4210811	0,354243	0,3320651	0,3782367	0,339972
REGIONN7	Coef	0,38	0,24	0,11	0,26	0,37	-0,05
	P>z	0,132	0,55	0,755	0,433	0,332	0,888
	E.E robusto	0,2514281	0,4032049	0,343772	0,3301751	0,3782367	0,3368769
REGIONN8	Coef	0,76	0,38	0,25	0,29	0,27	0,21
	P>z	0,04	0,533	0,629	0,543	0,608	0,646
	E.E robusto	0,3683935	0,6168335	0,513506	0,4761055	0,5220182	0,464667
REGIONN9	Coef	0,63	0,20	-0,0008	0,43	0,55	-0,08
	P>z	0,048	0,702	0,999	0,313	0,252	0,863
	E.E robusto	0,3197806	0,5261977	0,458092	0,4248163	0,4808308	0,4521595

Fuente: ECV, agosto-octubre 2008, cálculos propios, Stata 11, comando: streg

El primer conjunto de variables incluidas en el modelo, como se menciona en el apartado anterior, corresponde a aquel que afecta el salario de reserva establecido por el buscador. En este grupo se encuentran el ser jefe de hogar, el estado civil, el género del buscador, si es aspirante o no y los ingresos de su hogar.

La primera variable de este grupo en el modelo es ser jefe de hogar, que no tiene ningún efecto sobre la duración, pero que extrañamente y contrario a la hipótesis planteada en la descripción de las variables, si fuese significativa tendría una relación positiva con la variable dependiente tal y como lo encontró Oviedo (2007). Esto puede deberse a que la mayoría de los desempleados no son jefes de hogar. Estos individuos con pocos compromisos se adaptan más fácilmente a las condiciones del trabajo y pueden ser más flexibles a los horarios y remuneraciones que estipule el empleador. Además, se debe tener en cuenta que actualmente han ganado peso los trabajos temporales y poco estables dentro de las ofertas laborales que no son muy apetecidas por los jefes de hogar debido a la condición de estabilidad que prima entre ellos. Sin embargo, como se mencionó

anteriormente la carencia de significancia económica y estadística del estimador de la variable impide hacer estas inferencias.

Al igual que la variable anterior, la variable estado civil no parece tener relación con la duración del desempleo, esto puede deberse a que la mayoría de los desempleados son solteros.

La tercera variable incluida en el modelo es el género del buscador, esta variable cuyo coeficiente tiene signo negativo, nos muestra que el hecho de ser mujer aumenta la probabilidad de permanecer en el desempleo. En otras palabras, con relación al género, se observa que los hombres tienen una duración del desempleo inferior que las mujeres. Una explicación a este fenómeno según Polacheck y Stanley (1994) esta relacionado con el efecto de los factores culturales que confiere al hombre una mayor responsabilidad en el sostenimiento económico de los hogares, lo que los obliga a salir del desempleo rápidamente.

En cuarto lugar, con relación a la variable aspirante, se observa que aquellos individuos que buscan empleo por primera vez tienen una mayor duración del desempleo que aquellos individuos que han trabajado antes. Es decir, el ser aspirante aumenta la duración del desempleo. Este es un resultado lógico porque posiblemente estas personas tienen un menor conocimiento del mercado, tienen problemas de información, falta de señal, costos de entrenamiento y por lo tanto requieren mayor tiempo para adquirir la información necesaria.

El ingreso del hogar, en teoría debería incrementar la duración del desempleo porque estos ingresos le permitirían al individuo hacer una búsqueda más larga que les permita encontrar mejores ofertas como lo han encontrado Tenjo y Ribero (1998), Martínez (2003) o Castellar y Uribe (2003). Sin embargo, y contrario a lo que se esperaba, el estimador que se obtiene al incluir esta variable es negativo, lo que puede estar asociada a redes sociales que no son captadas por los canales de búsqueda.

Por otro lado, el segundo grupo de variables incluidas en el modelo corresponde a aquel que afecta el salario ofrecido. Ese grupo está compuesto por la educación, la experiencia y

la experiencia al cuadrado. Se encuentra que un año adicional en educación incrementa la duración del desempleo, es decir, los más educados desean tener un mayor tiempo de búsqueda o lo que se conoce en la literatura un efecto aspiraciones de esta forma mayor educación producirían mayores ingresos esperados, haciendo que los más educados tengan mayores salarios de reserva.

Con respecto a la experiencia, se presenta un resultado extraño, que deberá ser revisado, ya que las cifras actuales dicen que aquélla reduce notablemente las tasas de salida con relación a las de los individuos sin experiencia laboral. Por su parte, la variable experiencia al cuadrado muestra un efecto reductor y no significativo en la duración del desempleo. La prueba conjunta, que se encuentra en el anexo 6, nos muestra que la experiencia y la experiencia al cuadrado son variables que se deben tener en cuenta debido a que si afectan la duración. Aquí de nuevo podría estar dominando el efecto aspiraciones.

El penúltimo grupo de variables a considerar es el compuesto por el uso de los diferentes canales de búsqueda. Este grupo distingue por influir en la probabilidad de recibir una oferta salarial y por ello en la duración del desempleo. Los que utilizan canales de búsqueda informales e informal moderado frente a los que utilizan los canales formales, controlando las demás variables, permanecen un menor tiempo en el desempleo. Este resultado puede deberse al hecho de que las personas primero agotan los canales que inicialmente le son más accesibles como son la ayuda de amigos y familiares. Ya después de haber utilizados estos recursos sin ningún resultado, pasan a utilizar los canales formales. Es decir, lo que aquí podemos tener es un problema de dependencia entre el canal utilizado y el tiempo que se lleva buscando empleo, por lo que no se debe pensar que utilizar medios formales para conseguir empleo implicara un mayor tiempo de búsqueda.

Para finalizar vamos a analizar el último conjunto de variables compuesto por las diferentes regiones de residencia. Con estas variables vamos a observar en cuales la duración del desempleo es mayor. Aquellos individuos que viven en regiones como la Atlántica, San Andrés o la Orinoquía presentan una mayor duración del desempleo frente a aquellos que

viven en Bogotá. Con respecto a las demás regiones no es mucho lo que podamos decir debido a que no son muy diferentes a lo que sucede en Bogotá. La prueba conjunta de estas variables, que se encuentra en el anexo 6, nos muestran que no es importante la ciudad donde de residencia ya que esta no afecta la duración. Podríamos decir entonces, que no existe una diferencia sistemática en la búsqueda de empleo en las diferentes regiones del país, lo cual podría dar indicios de una integración del mercado laboral colombiano.

8. CONCLUSIONES

El interés del estudio de los determinantes de la duración del desempleo tiene varias razones de ser. Por una parte, ofrecer una base empírica que permita edificar medidas de política económica conducentes a la mejora de las condiciones del mercado de trabajo. Por otra, facilitar la cuantificación de los efectos de un conjunto de variables socioeconómicas sobre la duración de los períodos de desempleo en un país en el que no existe un sistema universalizado de protección para este fenómeno.

Con este fin se planteó un modelo paramétrico para la duración de desempleo que se contrastó con la información que ofrece la ECV del 2008 para toda Colombia. De acuerdo con este modelo la cifra de duración promedio en el desempleo fue de 22 semanas aproximadamente. La tasa de desempleo calculada para esa matriz era del 10.29%, es decir, para ese año en Colombia había aproximadamente 1.766.493 personas desempleadas.

A diferencia de otras investigaciones en este trabajo se ha utilizado la distribución Gamma generalizada para identificar los principales determinantes de la duración del desempleo, esto debido a que es la distribución que mejor se ajusta a los datos de acuerdo con el criterio AIC y a la forma de la función de riesgo de la duración.

Se encontró que los jefes de hogar y los casados, *ceteris paribus* los otros determinantes, no tienen ningún efecto sobre la duración del desempleo. Mientras que el ser hombre a diferencia del hecho ser mujer hace que la duración del desempleo sea menor. El ser aspirante también aumenta la duración del desempleo. Por otro lado los ingresos del hogar tienen una incidencia negativa en la duración del desempleo, lo que se puede asociar con redes sociales que no son captadas por los canales de búsqueda. En cuanto a las variables de capital humano, la educación y la experiencia aumentan el tiempo de búsqueda de empleo, lo que se asocia con un efecto aspiraciones.

De igual manera, los canales de búsqueda utilizados por los individuos en el mercado laboral son una variable determinante en la duración del desempleo, de esta forma quienes utilizan canales de búsqueda informales e informal moderado frente a los que utilizan los canales formales, permanecen un menor tiempo en el desempleo.

El lugar de residencia también es una variable que se ha tenido en cuenta como determinante de la duración del desempleo. Se encontró que quienes habitan en regiones como la Atlántica, San Andrés o la Orinoquía tienen una mayor probabilidad de permanecer en el desempleo frente a los que viven en Bogotá, pero se también se halló que de manera conjunta la región de residencia no afecta la duración.

A partir de los resultados obtenidos, es recomendable que se establezca un seguro o subsidio de desempleo que ayude a las personas a financiar su período de búsqueda, permitiéndoles buscar un empleo que se ajuste más a sus características y que cumpla con sus expectativas de salario. También se pueden establecer políticas de entrenamiento y de reentrenamiento de la población más afectada por las duraciones, al igual que una política de primer empleo para facilitarles el acceso al mercado laboral a aquellos individuos que son aspirantes. Además, sería recomendable, dirigir ciertas políticas laborales para las mujeres. El Estado podría intervenir estableciendo leyes que faciliten la inclusión y permanencia de la mujer en el mercado laboral sin que esto genere un costo adicional para las empresas.

REFERENCIAS BIBLIOGRÁFICAS

Akaike, H. (1974): «A new look at the statistical model identification», *IEEE transaction and Automatic Control AC-19*, 716-723.

Arranz, JM. Cid, JC. Muro, J (2000a). “La duración del desempleo en Argentina”. En *Asociación Argentina de Economía Política (AAEP). Borradores*. Universidad de Alcalá. INDEC y Universidad Nacional de Salta.

Arranz, JM. Cid, JC. Muro, J (2000b). “La duración del desempleo en presencia de altas tasas de paro: el caso de la Argentina”. En *Asociación Argentina de Economía Política (AAEP). Documento de Trabajo No 1465*. Universidad de Alcalá y Universidad Nacional de Salta.

Ahn, Namkee. Ugidos- Olazabal, Arantza (1985). “Desempleo juvenil en España: qué determina su duración”. *Ekonomiaz, No 31-32*, Universidad del País Vasco. Págs. 245.265

Blanco J.M. (1995), “La duración del desempleo en España, en Dolado, J.J y Jimeno, J.F. (Comp.)(1995); *Estudios sobre el funcionamiento del mercado de trabajo español*, Op cit, Pág.123-151.

Castellar, Carlos E. Uribe G, José Ignacio (2003). “Determinantes de la duración del desempleo en el área metropolitana de Cali 1988-1998”. *Archivo de Economía*, No. 218. Departamento Nacional de Planeación.

Castillo Caicedo, Maribel. (2004). “Determinantes de la probabilidad de estar desempleado en el área metropolitana de Cali: evidencias micro y macroeconómicas en el período 1988-1998.” Tesis presentada para optar al título de economista, Universidad del Valle, Cali, Colombia.

Cerimedo, Federico (2004). “Duración del desempleo y ciclo económico en la Argentina”. Centro de estudios distributivos, laborales y sociales. *Documento de Trabajo Nro. 8*. Facultad de Ciencias Económicas, Universidad de la Plata

DANE (2011). “Principales indicadores del mercado laboral”. Boletín de Prensa, Abril 2011. Bogotá, 31 de mayo de 2011.

Greene, William. (2008), *Econometric Analysis*, Sixth Edition, Practice- Hall International.

Hosmer D. W. y Lemeshow, S., 1999, *Applied survival analysis: regression modeling of time to event data*, John Wiley and Sons, Inc, New York.

Kaplan, E. L. y Meier P (1958). “Nonparametric estimation from incomplete observations”, *Journal of the American Statistical Association*, 53, pp. 457–481.

Layard R., Nickell s. and Jackman r. (1991), “*Unemployment, Macroeconomic Performance and the Labor Market*”. Oxford University Press. Traducido como “*El Paro, los resultados Macroeconómicos y el Mercado de Trabajo*”. Ministerio de Trabajo y Seguridad Social. España 1994.

López, Hugo (1988), “La duración del desempleo y el desempleo de larga duración en Colombia”. Fedesarrollo, *Coyuntura Económica: Análisis y perspectivas de la economía colombiana*, Vol. XVIII No 4, Bogotá, diciembre 1988. Págs. 163-186

Maddock R. (1987). “A propósito de la Misión Chenery: La Importancia de las Medidas de Duración del Desempleo”, *Lecturas de Economía*, N° 22, Medellín. Págs. 207-221

Martínez, Hermes Fernando (2003). “¿Cuánto duran los colombianos en el desempleo y el empleo? Un análisis de supervivencia”. Departamento Nacional de Planeación, *Archivos de Macroeconomía*, Documento N° 236. Bogotá.

McConnell, Campbell R. Brue, Stanley L. Macpherson, David A (2003). “La búsqueda de empleo: externa e interna” en *Economía Laboral*, sexta edición adaptada. McGraw-Hill.

Mortensen, D.T. (1986), “Job Search and Labor Market Analysis” en Ashenfelter, O. y Layard, R (Eds.), *Handbook of Labor Economics*, North Holland, Amsterdam. Edición en castellano: “Búsqueda de empleo y análisis del mercado de trabajo” en *Manual de Economía del trabajo*, Vol. II, Ministerio de Trabajo y Seguridad Social.

Murmis, Miguel. Feldman, Silvio (1997). “De seguir así” en *Sin Trabajo: las características del desempleo y sus efectos en la sociedad argentina*. Buenos Aires: Unicef. Págs.189-219.

Núñez, Jairo. Bernal, Raquel (1998), “El desempleo en Colombia: Tasa Natural, Desempleo cíclico y estructural y la duración del desempleo, 1976 – 1998”. DNP, *Archivos de Macroeconomía*, Documento 97, Bogotá.

Oviedo, Yanira Marcela, (2007). "Canales de búsqueda de empleo y duración del desempleo en el mercado laboral Colombiano 2003", *Sociedad y Economía*, No. 13, Facultad de Ciencias Sociales y Económicas, Universidad del Valle, Cali, pp.153-173.

Polachek S. W. y W. Stanley S., (1994), “Gender in the labour market”, in David B Grusky (ed), *Social stratification: class, race, and gender in sociological perspective*, First Edition. Boulder, CO: Westview Press, pp. 583-589.

Stigler (1962), “Information in the labor market”, *Journal of Political Economy*, Vol. LXX, pp. 94-105

Tenjo J (1998). “La duración y la incidencia del desempleo en Colombia: una nueva aproximación”, *Indicadores de Mercado Laboral*, SENA, Bogota, pp. 9-26

Tenjo J. y Ribero R., (1998), “Participación, desempleo y mercados laborales en Colombia”, DNP, *Archivos de Macroeconomía, Documento No. 81*, Bogotá.

Ureta, Manuelita (1996). “Incidencia y duración del desempleo en áreas urbanas del Paraguay”. *Estudios de Economía*. Vol. 25 – N° 1, Junio 1998. Págs. 51-70.

Uribe, José Ignacio. Viáfara, Carlos Augusto. Oviedo, Yanira Marcela, (2007). “Efectividad de los canales de Búsqueda de empleo en Colombia en el año 2003”. *Lecturas de Economía, No 67*, Facultad de Ciencias Económicas, Universidad de Antioquia, Medellín, Págs. 43-70.

Uribe, José Ignacio. Viáfara, Carlos Augusto (2009). “Duración del desempleo y canales de Búsqueda, 2006”. *Economía Institucional*. Vol. II – N° 21, Págs. 130-160.

Vinuesa, Pablo (2008). “Tema 7: máxima verosimilitud, estima de parámetros, selección de modelos y contraste de hipótesis” www.ccg.unam.mx. Mayo 10 de 2011.

Anexo 1

Resumen de la variable duración en semanas

duracion	Freq.	Percent	Cum.
1	136,978.36	7.75	7.75
2	158,630.86	8.98	16.73
3	108,738.03	6.16	22.89
4	249,849.07	14.14	37.03
5	55,960.196	3.17	40.20
6	62,368.699	3.53	43.73
7	14,802.0379	0.84	44.57
8	228,898.27	12.96	57.53
9	5,645.2436	0.32	57.85
10	23,677.6	1.34	59.19
12	140,155.53	7.93	67.12
13	7,121.8418	0.40	67.53
14	463.721344	0.03	67.55
15	8,341.6778	0.47	68.02
16	62,773.879	3.55	71.58
17	3,552.8427	0.20	71.78
18	1,804.8747	0.10	71.88
19	796.928528	0.05	71.93
20	52,244.73	2.96	74.88
21	1,955.2709	0.11	74.99
22	3,019.4803	0.17	75.16
24	65,874.888	3.73	78.89
26	5,286.186	0.30	79.19
28	20,559.033	1.16	80.36
30	9,748.9983	0.55	80.91
32	38,827.223	2.20	83.11
33	2,876.8009	0.16	83.27
34	1,641.0461	0.09	83.36
35	2,307.8188	0.13	83.49
36	26,213.781	1.48	84.98
38	2,461.62207	0.14	85.12
39	1,349.8633	0.08	85.19
40	14,209.036	0.80	86.00
42	188.780518	0.01	86.01
44	5,284.5608	0.30	86.31
45	1,608.8854	0.09	86.40
48	43,436.589	2.46	88.86
49	556.541626	0.03	88.89
50	6,424.6926	0.36	89.25
51	596.119629	0.03	89.29
52	56,178.6623	3.18	92.47
54	552.424942	0.03	92.50
56	4,541.2889	0.26	92.75
60	11,605.66	0.66	93.41
64	1,667.3439	0.09	93.51
65	2,007.4957	0.11	93.62
70	2,256.1798	0.13	93.75
72	9,653.3678	0.55	94.29
76	825.138779	0.05	94.34
80	10,578.42	0.60	94.94
84	580.918335	0.03	94.97
90	36.6959305	0.00	94.97
96	7,752.8557	0.44	95.41
98	5,180.8954	0.29	95.71
99	33.616024	0.00	95.71
100	401.618134	0.02	95.73
102	1,623.0387	0.09	95.82
104	9,968.1669	0.56	96.39
108	3,578.2578	0.20	96.59
110	3,142.1542	0.18	96.77
120	2,344.9608	0.13	96.90
125	161.818954	0.01	96.91
140	2,831.766	0.16	97.07
144	9,958.7799	0.56	97.63
150	323.660126	0.02	97.65
156	12,274.71	0.69	98.35
160	668.546692	0.04	98.38
172	40.088089	0.00	98.39
180	1,856.0676	0.11	98.49
192	7,919.3353	0.45	98.94
208	734.592346	0.04	98.98
210	3,948.2859	0.22	99.21
240	1,982.4132	0.11	99.32
250	127.03643	0.01	99.32
260	256.9833374	0.01	99.34
306	2,014.7882	0.11	99.45
336	2,485.5623	0.14	99.59
400	2,330.689	0.14	99.74
480	3,163.5522	0.18	99.92
520	1,474.0273	0.08	100.00
Total	1,766,493	100.00	

Anexo 2

Resumen estadístico del conjunto de datos sobre la duración del desempleo

Resumen por género

bsexo	time at risk	incidence rate	no. of subjects	survival time		
				25%	50%	75%
mujer	18398579.31	.0141478	908857.5	12	52	210
hombre	20476114.19	.0293277	857635.9	4	12	48
total	38874693.5	.0221434	1766493	6	24	80

Resumen por parentesco

bjefe	time at risk	incidence rate	no. of subjects	survival time		
				25%	50%	75%
no jefe	26834643.41	.0217081	1282666	8	32	80
jefe de	12040050.09	.0231135	483827.9	5	12	56
total	38874693.5	.0221434	1766493	6	24	80

Resumen por estado civil

unido	time at risk	incidence rate	no. of subjects	survival time		
				25%	50%	75%
soltero	19311232.99	.0269611	953479.7	6	20	52
unido	19563460.51	.0173877	813013.8	8	32	144
total	38874693.5	.0221434	1766493	6	24	80

Resumen por nivel de educación

nivedu	time at risk	incidence rate	no. of subjects	Survival time		
				25%	50%	75%
Ninguno	1033004.495	.0356642	59749.48	4	12	70
Primaria	7795641.99	.0266361	425495.8	5	17	96
Secundar	20095965.63	.0217938	866495.7	6	24	60
Tecnico	4219377.031	.0182409	176729.2	8	32	210
Universi	5730704.35	.0176936	238023.3	10	48	110
total	38874693.5	.0221434	1766493	6	24	80

Resumen según sea aspirante o no

baspir~e	time at risk	incidence rate	no. of subjects	Survival time		
				25%	50%	75%
trabajo	32967771.32	.0219846	1450466	6	24	80
por prim	5906922.175	.0230293	316027.5	8	32	56
total	38874693.5	.0221434	1766493	6	24	80

Resumen por canal de búsqueda

canalbus	time at risk	incidence rate	no. of subjects	Survival time		
				25%	50%	75%
canal 1	12130889.45	.0191376	593021.3	8	36	108
canal 3	11038561.44	.0190768	395864	8	28	64
canal 2	15705242.62	.0266204	777608.1	5	17	60
total	38874693.5	.0221434	1766493	6	24	80

Resumen por región

region	time at risk	incidence rate	no. of subjects	Survival time		
				25%	50%	75%
Atlantic	5407534.869	.0307567	349382.6	6	24	52
Oriental	2922998.628	.0264029	198480.5	6	32	96
Central	4342513.329	.0300953	250539.9	4	12	110
Pacifica	2495007.167	.0244279	127400.6	8	24	56
Bogota D	8618054.889	.0222973	329479.2	6	24	60
Antioqui	10668840.57	.0105292	269429.6	8	48	400
Valle Ca	4168477.593	.0260434	216760.6	8	22	64
San Andr	43795.64479	.0291388	2499.483	8	20	48
Orinoqui	207470.809	.0547271	22520.95	6	12	30
total	38874693.5	.0221434	1766493	6	24	80

Anexo 3

Log-rank y Wilcoxon (Breslow) test para la igualdad de las funciones de supervivencia a través de los grupos

Comparación de las funciones de supervivencia por género

Log-rank test for equality of survivor functions

bsexo	Events observed	Events expected
mujer	260299.82	426916.07
hombre	600516.94	433900.69
Total	860816.76	860816.76

chi2(1) = 139503.90
Pr>chi2 = 0.0000

Wilcoxon (Breslow) test for equality of survivor functions

bsexo	Events observed	Events expected	Sum of ranks
mujer	260299.82	426916.07	-1.588e+11
hombre	600516.94	433900.69	1.588e+11
Total	860816.76	860816.76	0

chi2(1) = 109220.71
Pr>chi2 = 0.0000

Comparación de las funciones de supervivencia por parentesco

Log-rank test for equality of survivor functions

bjefe	Events observed	Events expected
no jefe	582529.65	624697.74
jefe de hogar	278287.11	236119.03
Total	860816.76	860816.76

chi2(1) = 11370.96
Pr>chi2 = 0.0000

wilcoxon (Breslow) test for equality of survivor functions

bjefe	Events observed	Events expected	Sum of ranks
no jefe	582529.65	624697.74	-5.665e+10
jefe de hogar	278287.11	236119.03	5.665e+10
Total	860816.76	860816.76	0

chi2(1) = 17671.02
Pr>chi2 = 0.0000

Comparación de las funciones de supervivencia por estado civil

Log-rank test for equality of survivor functions

unido	Events observed	Events expected
soltero	520652.42	456588.86
unido	340164.35	404227.90
Total	860816.76	860816.76

chi2(1) = 20719.07
Pr>chi2 = 0.0000

wilcoxon (Breslow) test for equality of survivor functions

unido	Events observed	Events expected	Sum of ranks
soltero	520652.42	456588.86	4.612e+10
unido	340164.35	404227.90	-4.612e+10
Total	860816.76	860816.76	0

chi2(1) = 9271.19
Pr>chi2 = 0.0000

Comparación de las funciones de supervivencia por nivel de educación

Log-rank test for equality of survivor functions

nivedu	Events observed	Events expected
Ninguno	36841.319	25562.54
Primaria	207645.4	186226.85
Secundaria	437968.04	427264.67
Tecnico	76965.258	90574.72
Universidad	101396.74	131187.98
Total	860816.76	860816.76

chi2(4) = 17754.54
Pr>chi2 = 0.0000

Wilcoxon (Breslow) test for equality of survivor functions

nivedu	Events observed	Events expected	Sum of ranks
Ninguno	36841.319	25562.54	1.538e+10
Primaria	207645.4	186226.85	3.012e+10
Secundaria	437968.04	427264.67	1.082e+09
Tecnico	76965.258	90574.72	-7.975e+09
Universidad	101396.74	131187.98	-3.861e+10
Total	860816.76	860816.76	0

chi2(4) = 23968.92
Pr>chi2 = 0.0000

Comparación de las funciones de supervivencia según sea aspirante o no

Log-rank test for equality of survivor functions

baspirante	Events observed	Events expected
trabajo antes	724784.45	713014.67
por primera vez	136032.31	147802.09
Total	860816.76	860816.76

chi2(1) = 1220.25
Pr>chi2 = 0.0000

wilcoxon (Breslow) test for equality of survivor functions

baspirante	Events observed	Events expected	Sum of ranks
trabajo antes por primera vez	724784.45 136032.31	713014.67 147802.09	1.970e+10 -1.970e+10
Total	860816.76	860816.76	0

chi2(1) = 2860.49
Pr>chi2 = 0.0000

Comparación de las funciones de supervivencia por canal de búsqueda

Log-rank test for equality of survivor functions

canalbus	Events observed	Events expected
canal 1	232156.52	287489.18
canal 3	210580.77	218611.04
canal 2	418079.47	354716.54
Total	860816.76	860816.76

chi2(2) = 23999.72
Pr>chi2 = 0.0000

wilcoxon (Breslow) test for equality of survivor functions

canalbus	Events observed	Events expected	Sum of ranks
canal 1	232156.52	287489.18	-5.831e+10
canal 3	210580.77	218611.04	-1.601e+10
canal 2	418079.47	354716.54	7.433e+10
Total	860816.76	860816.76	0

chi2(2) = 26229.13
Pr>chi2 = 0.0000

Comparación de las funciones de supervivencia por región

Log-rank test for equality of survivor functions

region	Events observed	Events expected
Atlantica	166317.95	148572.50
Oriental	77175.739	78052.36
Central	130689.28	108679.85
Pacifica	60947.832	62460.97
Bogota D.C	192159.28	183237.75
Antioquia	112334.71	165338.17
Valle Cauca	108561.53	105835.72
San Andres y Providencia	1276.1521	1157.02
Orinoquia y Amazonia	11354.284	7482.43
Total	860816.76	860816.76

chi2(8) = 29377.09
Pr>chi2 = 0.0000

wilcoxon (Breslow) test for equality of survivor functions

region	Events observed	Events expected	Sum of ranks
Atlantica	166317.95	148572.50	3.775e+09
Oriental	77175.739	78052.36	1.311e+09
Central	130689.28	108679.85	3.215e+10
Pacifica	60947.832	62460.97	-4.869e+09
Bogota D.C	192159.28	183237.75	6.719e+09
Antioquia	112334.71	165338.17	-3.233e+10
Valle Cauca	108561.53	105835.72	-8.701e+09
San Andres y Providencia	1276.1521	1157.02	-1.647e+08
Orinoquia y Amazonia	11354.284	7482.43	2.109e+09
Total	860816.76	860816.76	0

chi2(8) = 17291.16
Pr>chi2 = 0.0000

Anexo 4

Salida de stata y

Pruebas de hipótesis de la distribución Gamma

Gamma regression -- accelerated failure-time form

```

No. of subjects      =          715          Number of obs      =          715
No. of failures     =           336
Time at risk        =          12179
Log pseudolikelihood = -716.22679
Wald chi2(18)       =          116.02
Prob > chi2         =           0.0000
    
```

_t	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
bjefe	.0457294	.1286585	0.36	0.722	-.2064365	.2978954
unido	.0433425	.1275447	0.34	0.734	-.2066405	.2933256
bsexo	-.5511712	.1943738	-2.84	0.005	-.9321368	-.1702055
baspirante	.8533637	.1985699	4.30	0.000	.4641739	1.242553
loging	-.2483434	.0453557	-5.48	0.000	-.3372389	-.159448
anosedu	.0782704	.0239812	3.26	0.001	.0312681	.1252727
exper	.0297262	.0125764	2.36	0.018	.0050769	.0543754
exper2	-.0002322	.0001846	-1.26	0.208	-.000594	.0001296
canal2	-.6548309	.1502112	-4.36	0.000	-.9492394	-.3604224
canal3	-.3871826	.1782668	-2.17	0.030	-.7365791	-.037786
regionn1	.5605919	.2641372	2.12	0.034	.0428925	1.078291
regionn2	.2262449	.2632897	0.86	0.390	-.2897934	.7422832
regionn3	.1048493	.238484	0.44	0.660	-.3625707	.5722693
regionn4	.1702315	.2654754	0.64	0.521	-.3500907	.6905538
regionn6	.4246702	.2578432	1.65	0.100	-.0806932	.9300336
regionn7	.3782353	.2514281	1.50	0.132	-.1145547	.8710252
regionn8	.7577046	.3683935	2.06	0.040	.0356665	1.479743
regionn9	.6330426	.3197806	1.98	0.048	.0062841	1.259801
_cons	4.194209	.6536703	6.42	0.000	2.913038	5.475379
/ln_sig	.265408	.0601398	4.41	0.000	.1475361	.3832799
/kappa	-1.033416	.3817892	-2.71	0.007	-1.781709	-.2851231
sigma	1.303963	.0784201			1.158975	1.467089

- La prueba de la hipótesis nula $k=0$ (que el modelo es log-normal) usando la prueba de wald, es la que se presenta en los resultados, $p=0.007$. Esto sugiere que el modelo log-normal no es el adecuado para estos datos

- La prueba de chi-cuadrado de la hipótesis nula $k=1$

```
. testnl [kappa]_cons=1
```

```
(1) [kappa]_cons = 1
```

```

      chi2(1) =          28.37
    Prob > chi2 =          0.0000
    
```

Acepta el modelo weibull

Anexo 5

Salidas de Stata de las diferentes distribuciones

• Weibull

weibull regression -- log relative-hazard form

```

No. of subjects      =          715          Number of obs   =          715
No. of failures     =           336
Time at risk        =          12179
Log pseudolikelihood = -769.91885
Wald chi2(18)       =          189.00
Prob > chi2         =           0.0000
  
```

_t	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
bjefe	.0007725	.1558166	0.00	0.996	-.3046224	.3061674
unido	-.1565752	.1439503	-1.09	0.277	-.4387127	.1255623
bsexo	.9545729	.1521022	6.28	0.000	.656458	1.252688
baspirante	-.5436775	.2366835	-2.30	0.022	-1.007569	-.0797864
loging	.2436101	.0492829	4.94	0.000	.1470174	.3402029
anosedu	-.1273926	.0196304	-6.49	0.000	-.1658674	-.0889178
exper	-.0362239	.014775	-2.45	0.014	-.0651825	-.0072654
exper2	.0001981	.000228	0.87	0.385	-.0002487	.000645
canal2	.6548881	.1711608	3.83	0.000	.319419	.9903571
canal3	.2618407	.2637959	0.99	0.321	-.2551898	.7788711
regionn1	.1814295	.3326882	0.55	0.586	-.4706273	.8334864
regionn2	.3859561	.3896037	0.99	0.322	-.3776531	1.149565
regionn3	.3986353	.353142	1.13	0.259	-.2935102	1.090781
regionn4	.3018322	.3380986	0.89	0.372	-.3608288	.9644932
regionn6	.0198162	.3542425	0.06	0.955	-.6744864	.7141187
regionn7	.1072402	.3437718	0.31	0.755	-.5665401	.7810205
regionn8	.2483503	.513506	0.48	0.629	-.7581029	1.254804
regionn9	-.0007634	.4580917	-0.00	0.999	-.8986067	.8970799
_cons	-5.046349	.755552	-6.68	0.000	-6.527203	-3.565494
/ln_p	-.2027629	.0425978	-4.76	0.000	-.2862531	-.1192727
p	.8164718	.0347799			.7510725	.8875657
1/p	1.224782	.052173			1.126677	1.331429

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	715	-880.9505	-769.9188	20	1579.838	1671.283

Note: N=Obs used in calculating BIC; see [\[R\] BIC note](#)

- Log normal

Lognormal regression -- accelerated failure-time form

No. of subjects	=	715	Number of obs	=	715
No. of failures	=	336			
Time at risk	=	12179			
Log pseudolikelihood	=	-726.8563	wald chi2(18)	=	232.95
			Prob > chi2	=	0.0000

_t	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
bjefe	.0038829	.1512261	0.03	0.980	-.2925147	.3002806
unido	.134035	.1382735	0.97	0.332	-.1369761	.405046
bsexo	-.8793741	.1362151	-6.46	0.000	-1.146351	-.6123975
baspirante	.7857676	.2367665	3.32	0.001	.3217138	1.249821
loging	-.2920015	.0504147	-5.79	0.000	-.3908125	-.1931905
anosedu	.1212623	.0186738	6.49	0.000	.0846624	.1578622
exper	.0365021	.0145234	2.51	0.012	.0080367	.0649674
exper2	-.0002128	.000227	-0.94	0.349	-.0006578	.0002322
cana12	-.7151149	.1662597	-4.30	0.000	-1.040978	-.3892518
cana13	-.3875786	.2188941	-1.77	0.077	-.8166033	.041446
regionn1	.3846182	.3242907	1.19	0.236	-.25098	1.020216
regionn2	.0195526	.3530648	0.06	0.956	-.6724417	.711547
regionn3	-.0813689	.3063758	-0.27	0.791	-.6818545	.5191167
regionn4	.1267996	.3198148	0.40	0.692	-.5000259	.753625
regionn6	.3372388	.3320651	1.02	0.310	-.3135968	.9880745
regionn7	.2586674	.3301751	0.78	0.433	-.3884639	.9057988
regionn8	.2894341	.4761055	0.61	0.543	-.6437156	1.222584
regionn9	.4289474	.4248163	1.01	0.313	-.4036772	1.261572
_cons	5.198786	.6549005	7.94	0.000	3.915205	6.482367
/ln_sig	.3037758	.0387951	7.83	0.000	.2277389	.3798128
sigma	1.354965	.052566			1.255757	1.462011

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	715	-825.521	-726.8563	20	1493.713	1585.158

Note: N=Obs used in calculating BIC; see [R] BIC note

- Gompertz

Gompertz regression -- log relative-hazard form

```

No. of subjects      =          715          Number of obs      =          715
No. of failures     =           336
Time at risk        =          12179
Log pseudo-likelihood = -752.03638
Wald chi2(18)       =          197.97
Prob > chi2         =           0.0000
  
```

_t	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
bjefe	-.0148453	.1504551	-0.10	0.921	-.3097319	.2800413
unido	-.1332474	.1373707	-0.97	0.332	-.402489	.1359942
bsexo	.9883023	.1428744	6.92	0.000	.7082737	1.268331
baspirante	-.5229089	.2265544	-2.31	0.021	-.9669473	-.0788705
loging	.2662563	.0449454	5.92	0.000	.178165	.3543477
anosedu	-.1247302	.0187067	-6.67	0.000	-.1613946	-.0880657
exper	-.0325363	.0145103	-2.24	0.025	-.0609759	-.0040967
exper2	.0001721	.0002248	0.77	0.444	-.0002684	.0006127
canal2	.6658243	.1675042	3.97	0.000	.3375221	.9941266
canal3	.4468568	.2275714	1.96	0.050	.000825	.8928885
regionn1	.0523321	.3269665	0.16	0.873	-.5885106	.6931747
regionn2	.3184453	.3774392	0.84	0.399	-.4213219	1.058213
regionn3	.3859305	.3211062	1.20	0.229	-.243426	1.015287
regionn4	.2262696	.3257819	0.69	0.487	-.4122511	.8647903
regionn6	-.0119381	.3399712	-0.04	0.972	-.6782694	.6543932
regionn7	-.047241	.3368769	-0.14	0.888	-.7075077	.6130256
regionn8	.2137501	.464667	0.46	0.646	-.6969806	1.124481
regionn9	-.0777669	.4521595	-0.17	0.863	-.9639832	.8084494
_cons	-5.647936	.641601	-8.80	0.000	-6.90545	-4.390421
/gamma	-.0132043	.0033119	-3.99	0.000	-.0196955	-.0067131

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	715	-867.5308	-752.0364	20	1544.073	1635.518

Note: N=Obs used in calculating BIC; see [R] BIC note

Anexo 6

Pruebas conjuntas

- **Experiencia**

```
. test (exper exper2)
( 1) [_t]exper = 0
( 2) [_t]exper2 = 0
      chi2( 2) =    9.78
      Prob > chi2 =  0.0075
```

- **Canales de búsqueda**

```
. test (cana12 cana13)
( 1) [_t]cana12 = 0
( 2) [_t]cana13 = 0
      chi2( 2) =   19.12
      Prob > chi2 =  0.0001
```

- **Región de residencia**

```
. test ( regionn1 regionn2 regionn3 regionn4 regionn6 regionn7 regionn8 regionn9)
( 1) [_t]regionn1 = 0
( 2) [_t]regionn2 = 0
( 3) [_t]regionn3 = 0
( 4) [_t]regionn4 = 0
( 5) [_t]regionn6 = 0
( 6) [_t]regionn7 = 0
( 7) [_t]regionn8 = 0
( 8) [_t]regionn9 = 0
      chi2( 8) =   14.29
      Prob > chi2 =  0.0746
```

