



Universidad de  
**La Sabana**

ESCUELA INTERNACIONAL DE CIENCIAS ECONÓMICAS Y ADMINISTRATIVAS  
ECONOMÍA Y FINANZAS INTERNACIONALES

TRABAJO DE GRADO PARA OPTAR POR EL TÍTULO DE  
ECONOMISTA CON ÉNFASIS EN FINANZAS INTERNACIONALES

**ANÁLISIS DE DURACIÓN DEL DESEMPLEO EN BOGOTÁ 2017-3**

**Directora:** Martha Misas Arango

**Presentado por:** David Steven Bernal Giraldo  
Mateo Cardoza González

Chía, 29 de diciembre de 2018

### *Resumen*

La tasa de desempleo de una economía es uno de los indicadores más importantes a la hora de observar lo que sucede en el mercado laboral. Sin embargo, este indicador no brinda la información suficiente sobre las dinámicas laborales en ciudades tan complejas como Bogotá. Una aproximación para dar una mejor información sobre el complejo mercado laboral de la ciudad viene dado por el modelo econométrico de duración del desempleo. Este modelo ha sido ampliamente estudiado en la literatura y es tema central de estudio en diversas investigaciones. A pesar de esto, es un tema vagamente estudiado y desarrollado en Colombia. El objetivo de este trabajo es llevar a cabo un modelo paramétrico de duración del desempleo, de manera que se logre aportar a la información existente de los procesos de búsqueda de empleo y del desempleo en la ciudad de Bogotá, así como aportar al desarrollo de políticas para el mercado laboral.

### *Abstract*

When evaluating the labor market of an economy, the unemployment rate is considered one of the most important indicators. Nevertheless, this index does not render enough information about the labor patterns in the wake of contemplating cities as complex as Bogotá. A more solid approach in order to provide more accurate information about this city's manifold labor market can be found in the unemployment span econometric model. Although this is not a new model and there seems to be a lot of literature about this topic, it has been vaguely studied and developed in Colombia. The objective of the research is to create an unemployment span econometric model that will provide further information regarding the unemployment in the city of Bogotá to the individual who is seeking job. Likewise, the model will contribute in developing effective policies in the labor market.

# TABLA DE CONTENIDO

I.	Introducción.....	4
II.	Revisión de literatura.....	5
III.	Metodología.....	8
	1. Modelo de búsqueda de empleo .....	8
	2. Análisis econométrico .....	10
IV.	Conformación de la base de datos. ....	12
V.	Conjunto de variables relevantes para la estimación econométrica. ....	12
VI.	Resultados .....	13
VII.	Pronóstico de la duración del desempleo. ....	17
VIII.	Bondad de ajuste. ....	19
IX.	Conclusiones .....	19
X.	Referencias .....	21
XI.	Anexo 1.....	22
	1. Base de datos individuos ocupados (no censurados).....	22
	2. Base de datos individuos desocupados (censurados). ....	22
	3. Conformación base de datos ocupados y desocupados. ....	23
	4. Conformación base de datos de características generales. ....	23
	5. Conformación de base de datos ocupados y desocupados con características generales. .	24
	6. Conformación de número de menores en el hogar y personas a cargo. ....	24
	7. Conformación base de datos ocupados y desocupados con características generales y número de menores en el hogar. ....	24
	8. Ingreso.....	24
	9. Conformación de la base de datos de ingreso individual. ....	27
	10. Conformación base de datos de ingreso familiar total.....	28
	11. Conformación de la data total por mes. ....	28
	12. Conformación DATAGLOBAL.....	28

# I. INTRODUCCIÓN

El desempleo, definido como el indicador “de la evolución en el tiempo (...) de la proporción de personas que, teniendo la intención de trabajar, se puedan emplear.” (Banco de la República, 2013), es un tema de gran profundidad e interés para la sociedad ya que por medio de este indicador se ve, en gran medida, el comportamiento de una economía y su mercado laboral.

El estudio se realiza con el fin de darle una mejor interpretación al indicador de desempleo y debido a la falta de estudios recientes para la ciudad de Bogotá. Este es muy importante para la economía, a pesar de que por construcción se puede llegar a un mismo resultado teniendo diferentes dinámicas de búsqueda de empleo y de rotación de este (Tenjo, Misas, Contreras, & Gaviria, 2012).

Tomando como base la tasa de desempleo (sin considerar cuánto tiempo llevan estos individuos en este estado), se puede llegar a tener una misma medición a pesar de que la duración entre individuos pueda variar. Por ejemplo, una tasa de desempleo en un mes dado del 7% debido a que en ese mes se finalizan los contratos a término fijo de ciertos trabajadores pero que, el siguiente mes, renovarían su contrato después de un breve periodo de paro. Para otro mes dado puede haber una tasa de desempleo equivalente al 7% ya que el número de desempleados en dicho mes es el mismo que en el primer caso, pero a diferencia de este, estos desempleados llevan 3 o más meses en esta situación. El caso mencionado es solo un ejemplo a las distintas variables y situaciones que hacen surgir la necesidad de profundizar en la interpretación del indicador de desempleo por medio de modelos de duración.

Los modelos econométricos de duración son modelos que miden el tiempo que se dura en un estado antes de pasar a otro (Cameron Colin, 2005). De manera que lo que se busca a través de estos modelos, para este caso, es en sí generar una interpretación más precisa de la tasa de desempleo, ya sea desempleo transitorio o estructural, y de esta forma poder sugerir políticas económicas laborales más apropiadas para fomentar el empleo. Al mismo tiempo se busca explicar en qué condiciones hay una menor duración del desempleo y a su vez qué condiciones lo aumentan.

El estudio se centra en la ciudad de Bogotá debido a su importancia económica, ya que históricamente ha contribuido en gran medida al PIB nacional, aportando en 2017 el 26,4% (DANE, 2018) y a su densidad poblacional (10,2 millones de habitantes), correspondiente al 20,83% de la población total colombiana (DANE, 2018). La importancia del empleo de Bogotá para Colombia es significativa, pues adicional a lo mencionado anteriormente, representó un alto porcentaje sobre el total de ocupados del país para el tercer trimestre de 2017. Teniendo en cuenta esto, es importante poder entender las dinámicas de empleo en Bogotá, y de esta manera generar un mejor panorama para la generación de políticas que ayuden a reducir el desempleo de una manera más efectiva.

## **Motivación del trabajo**

El desempleo es un tema que no solo les importa a las personas que se encuentran en esta condición, sino a los individuos que están próximos a ingresar al mercado laboral. Nosotros como futuros egresados estaremos en esta condición y debido a la falta de estudios recientes del tema, específicamente para la capital del país (Bogotá) no tenemos información plena acerca de esta problemática. Por medio de la presente investigación haremos un análisis del mercado laboral actual para así poder contribuir a la sociedad colombiana, dando explicación desde el punto económico y econométrico a las dinámicas de desempleo en la ciudad más importante frente a aporte al PIB y mayor densidad poblacional.

El objetivo de este trabajo es hacer la conexión de un modelo de economía laboral y un modelo econométrico realizando el análisis de los determinantes del desempleo a la luz de un modelo de duración del desempleo, tomando como base el salario de reserva. Estos dos modelos permiten brindar una mejor interpretación sobre el indicador de desempleo, ya que la duración del desempleo muestra si el desempleo es transitorio o si por el contrario es el resultado de dinámicas económicas que alargan la duración de este.

## **II. REVISIÓN DE LITERATURA**

Para efectos de la discusión de literatura, se agrupan los autores aquí estudiados por el modelo econométrico que se utiliza en sus respectivos trabajos. De esta forma, se mencionan a algunos autores que resultan relevantes para la investigación y que no utilizan modelos de duración y posteriormente se plantean tres grupos principales de metodologías (Paramétricas, no paramétricas y semi-paramétricas) así como combinaciones de los mismas, utilizadas en los diferentes modelos de duración aquí expuestos.

En (Arango, Luis Eduardo, Posada, 2001), se dice que una de las principales causas del desempleo colombiano es la rigidez del salario real, en cuanto a su falta de capacidad para ajustarse, a fin de eliminar, o al menos, reducir los movimientos del desempleo. En este artículo, por medio de un modelo de mercado laboral con características neoclásicas, que busca la optimización individual, se puede decir que la tasa de desempleo observada es de carácter transicional, y por eso es mayor a la tasa natural de desempleo. Esto, se ve reflejado en un efecto desánimo que ésta presente en la economía, porque los individuos llegaron a un punto de cansancio de búsqueda de empleo sin obtener resultados favorables.

En (Arango & Posada, 2003), los mismos autores buscan explicar cómo se ve afectada la participación laboral en siete ciudades de Colombia dependiendo del estado civil de hombres y mujeres (teniendo como significativo su género en la decisión de participación) y buscando aportar a la tasa de participación. Se utilizó un modelo en el cual la variable dependiente es de orden binario, siendo 1 si participa del mercado laboral y 0 de lo contrario. Se utilizó un modelo probit, con base en una función  $F$  que toma la forma de una distribución normal, para el cual se utiliza la máxima verosimilitud a la hora de realizar la estimación. Se llega a los resultados que indican que estar comprometido o no y el estatus de comprometido o no es

irrelevante para hombres en presencia de menores de seis años. El género es irrelevante para las personas comprometidas y el estatus de comprometido o no es irrelevante ante la riqueza. Para otros desempleados en el hogar es irrelevante el género tanto en personas comprometidas como no comprometidas y para la edad se obtuvo como resultado que es irrelevante para personas comprometidas. Como conclusión se dice que a mayor tasa de desempleo mayor es la participación laboral en los próximos trimestres. (Arango & Posada, 2003)

En su artículo (Shorrocks, 2009), Anthony habla sobre posibles alternativas a la medición del desempleo por medio de un análisis de duración de éste. Se plantea que, si bien en la literatura es reconocida la duración del desempleo como una medida complementaria al indicador de desempleo, nunca se ha considerado incorporar dicha duración este. Se propone entonces que la medida del desempleo debe tener cuatro propiedades básicas: *simetría*, *replicación invariable*, *monotonicidad* y *preferencia por la igualdad en la duración*. Posteriormente al hallar una función U que satisface las primeras propiedades, se enumeran dos propiedades más que se deben cumplir: la *homogeneidad frente a la tasa de desempleo* y *homogeneidad en duración*, generando de esta forma varios grupos de índices de desempleo que contemplan la duración del mismo, y por otra parte, cumplen en su mayoría las propiedades básicas que se observaron en los índices de desempleo estudiados (Shorrocks, 2009).

Tony Lancaster en (Lancaster, 1979), busca describir las formas de realizar análisis del desempleo en modelos de duración, resaltando errores y posibles soluciones a estos. Se usan como variables el *tiempo que lleva desempleado* al momento de responder la encuesta, *la edad*, una variable “*replacement*” que es una razón entre la cantidad de subsidios recibidos y cuanto ganaba después de deducciones en el trabajo que tenía. Se habla de los modelos paramétricos, semi-paramétricos y no paramétricos, y se explica que la diferencia entre el primer tipo y los últimos dos es que en los últimos se toma como censurada una observación que presenta una duración concluida si en ese mismo momento no existen otras observaciones con las que se pueda comparar, mientras que en el modelo paramétrico, al seguir una distribución, no se necesitan de otras observaciones para realizar la comparación, ya que esta se hace de acuerdo al comportamiento que toma la distribución en ese momento del tiempo.

Se proponen dos formas de especificación del Hazard para modelos paramétricos, dentro de las cuales están Parametric Hazard (un mismo Hazard para todas las variables y que no se modifica con el tiempo) y Accelerated Failure Time (en el cual el Hazard varía en el tiempo debido a un término de aceleración). El problema con el primer tipo de especificación es que, debido a regresores omitidos, se puede dar el caso de que algunos individuos tengan una tasa Hazard mayor a otros habiendo una mala especificación del modelo. En el segundo tipo de especificación el modelo presenta una nueva modificación en la que se incluye un término de error para capturar el efecto de los regresores omitidos.

Por último, se plantea que hay tres fuentes de variación para el modelo: un componente de suerte, ya que diversos individuos con unas mismas características encuentran trabajo en diferentes momentos. Todo esto bajo el supuesto de que los individuos están tomando decisiones con incertidumbre. La segunda y tercera fuente de variación se basan en la política que tomen o en el entorno que se encuentren.

Comenzando con los modelos de duración, se encontraron cuatro autores principales que manejan modelos de duración paramétricos (Belzil, 2000; Castellar & Uribe, 2006; Tenjo et al., 2012; Tenjo & Ribero, 1998). Estos autores justifican su escogencia de modelo ya que en los modelos semi-paramétricos se deben llevar a cabo una gran cantidad de estimaciones y que estos son los modelos más utilizados por la literatura econométrica correspondiente a modelos de duración. Los estudios, en términos generales, tiene como objetivo explicar la duración del desempleo y la relación de esta frente a los beneficios de desempleo.

Como resultados para los modelos paramétricos, se encuentra que variables como el género y el estado civil, junto con problemas de información del mercado laboral, son las principales determinantes de una mayor duración. Además, las características de las personas no influyen en la duración, pero si lo hacen los niveles de actividad económica y por lo tanto la implementación de políticas puede ayudar a manipular algunas de estas variables, haciendo que la duración de los individuos disminuya.

Frente a modelos no paramétricos, se encuentran estudios como los de (Narendranathan, 1993; Røed & Zhang, 2003). Los autores aseguran que, por medio de un modelo no restringido, se le pueden dar diferentes tasas Hazard a los individuos y a su vez usar dummies con mayor facilidad para llevar a cabo la estimación. El objetivo de los trabajos, al igual que en los anteriores es poder ver la influencia de los beneficios sobre la escogencia de empleos y transición del desempleo al empleo.

Estos autores encontraron que, ante una mayor duración, se observa una disminución en los beneficios, a la vez que la tasa de escape del desempleo (tasa Hazard) presenta un aumento para ambos géneros con un mayor efecto en los hombres, lo que implica una disminución en la duración.

En (Narendranathan, 1993), se va un paso más allá y se encuentra que el punto en el que los beneficios del desempleo dejan de tener relevancia es al llegar a los 3 meses. Adicionalmente, encuentra que el desempleo genera desutilidad y se plantea que el transcurso del tiempo no desincentiva la búsqueda de empleo.

Arulamplam y Stewart realizaron un análisis sobre la probabilidad condicionada de dejar el estado de desempleo en dos periodos de tiempo (1978 y 1987). Se busca estudiar este fenómeno partiendo del supuesto de que un individuo entre en desempleo y que un individuo salga del mismo, ambas siendo afectadas por las características de los individuos y tomando tres grupos de variables fundamentales que afectan las probabilidades de que un individuo encuentre una vacante y le ofrezcan el trabajo. Se utiliza un modelo semi-paramétrico partiendo del ingreso percibido durante el desempleo o empleo y se encuentra que, tras los primeros tres meses de duración del desempleo, se pierde el efecto de los beneficios del desempleo. Por otra parte, estos ingresos no solo tienen un impacto significativo en los tres primeros meses, sino que su mayor impacto está en la población de menos de 20 años. Finalmente se encuentra que un hombre tiene mayor posibilidad de escapar del desempleo si su esposa trabaja que si es soltero o su esposa no trabaja (Stewart & Arulamplam, 1995).

Otra aproximación al mismo problema estudiado se basa en la utilización tanto de un modelo paramétrico como no paramétrico con el fin de poder comparar los resultados de ambas

metodologías y de esta forma poder llegar a mejores conclusiones. Estos casos se observaron en los estudios de (Arango & Ríos, 2015; Marcillo Yépez, 2011; Martínez, 2003; Oviedo Gil, 2007; Viáfara & Uribe, 2008).

En estos estudios se encontró, que el ser jefe de hogar implica una menor duración en ambas estimaciones. De igual forma, se encuentra que los hombres presentan una menor duración que las mujeres y que las mujeres que únicamente tienen un conyugue presentan una mayor duración del desempleo frente a mujeres con menores a cargo o jefes de hogar. Por otra parte, en el estudio de (Viáfara & Uribe, 2008) se encuentra que el efecto de las oportunidades es un factor importante en los canales de búsqueda de Colombia, al igual que variables como género, edad, jefe de hogar, años de educación y experiencia. Finalmente, en (Martínez, 2003) se plantean políticas de empleo, orientadas tanto a oferta y demanda del mismo, con el fin de generar una disminución en la duración del desempleo, aumentar la duración del empleo y mejorar la calidad de vida.

### III. METODOLOGÍA

#### 1. Modelo de búsqueda de empleo

El problema de la búsqueda de empleo ha sido extensamente estudiado por la literatura económica. La mayoría de los estudios aseguran que las personas que buscan empleo lo hacen bajo incertidumbre teniendo en cuenta que poseen información incompleta. Una estrategia de búsqueda se construye en torno a la fijación de un salario de reserva, tal como en (Lancaster, 1979) y (Tenjo et al., 2012), el cual sirve como umbral para la toma de decisiones sobre aceptar o no las ofertas de trabajo. El salario de reserva busca maximizar los beneficios netos futuros del proceso de búsqueda, para que solo sean aceptadas las ofertas de trabajo que sean iguales o superiores al salario de reserva. De manera más sencilla:

Siendo  $f(w)$  una función de densidad la cual describe la percepción del individuo de la distribución de las ofertas salariales ( $w$ ) en un momento dado.

El individuo fija un salario de reserva  $w^*$  en su búsqueda de empleo tal que

Si  $w \geq w^*$  Acepta la oferta

Si  $w < w^*$  Rechaza la oferta

De acuerdo con lo anterior, la probabilidad de que una oferta sea rechazada estaría dada por:

$$Prob(w < w^*) = G(w^*) = \int_0^{w^*} g(w)dw \quad (1)$$

La probabilidad de que una oferta sea aceptada estaría dada por:

$$Prob(w \geq w^*) = 1 - G(w^*) \quad (2)$$



Teniendo en cuenta que las ofertas son un proceso aleatorio, el valor esperado de ofertas que se recibirán antes de que una sea aceptable se presenta a través de:

$$D = E(\lambda) = \frac{1}{1-G(w^*)} \quad (3)$$

Haciendo el supuesto de que el trabajador recibe una oferta por período, D puede ser interpretado como la duración esperada de la búsqueda. De no cumplirse este supuesto, la duración esperada de la búsqueda sería una función de D.

El salario esperado que recibiría un trabajador una vez reciba una oferta aceptable, se puede expresar como:

$$w_e = E(w|w \geq w^*) = \frac{\int_{w^*}^{\infty} wg(w)dw}{1 - G(w^*)} \quad (4)$$

Determinación del salario de reserva ( $w^*$ ):

Una forma sencilla de plantear el salario de reserva es suponer que el trabajador fija el mismo con base al valor presente de la maximización de beneficios netos futuros de la búsqueda (diferencia entre beneficios y costos).

El proceso de búsqueda implica unos costos directos y unos costos de oportunidad, de manera que los costos directos son asociados al proceso de búsqueda en si, como la adquisición de información (consulta de avisos, bases de datos, etc.), la generación de ofertas (envío de hojas de vida, acudir a entrevistas y otras cosas), etc. Los costos de oportunidad se toman como el salario que el individuo deja de percibir por periodo teniendo en cuenta que está destinando dicho tiempo a la búsqueda. De esta forma se tiene que los costos directos se pueden considerar como costos fijos por periodo, mientras que el costo de oportunidad aumenta con la duración de la búsqueda.

Se puede expresar la función de costos en términos de la duración misma de la búsqueda:

$$C = FD + h(D) = C(w^*) \quad y \quad \frac{\partial c}{\partial w^*} > 0 \quad (5)$$

En donde F es el costo fijo por periodo, h es el costo de oportunidad que está en función de la duración de la búsqueda.

Los beneficios (Y) de la búsqueda están asociados a los ingresos que percibe el trabajador y están representados por la función:

$$Y = Y(w^*); \quad con \quad \frac{\partial y}{\partial w^*} > 0 \quad y \quad \frac{\partial^2 y}{\partial w^{*2}} < 0 \quad (6)$$

Como consecuencia de lo expuesto, el trabajador determina su salario de reserva maximizando el valor presente del beneficio neto de la búsqueda ( $V(B)$ ).

$$V(B) = \int_D^N Y(W^*)e^{-rt} dt - \int_0^D C(W^*)e^{-rt} dt$$

$$V(B) = \frac{1}{r} [Y(W^*)(e^{-rD} - e^{-rN}) - C(W^*)(1 - e^{-rD})] \quad (7)$$

El valor óptimo del salario de reserva ( $w^*$ ) se obtiene resolviendo la siguiente ecuación para  $w^*$ :

$$\frac{\partial V(B)}{\partial w^*} = 0 \quad \rightarrow \quad w^{**} = k(g, r, N, X) \quad (8)$$

Lo cual es equivalente a igualar el beneficio marginal con el costo marginal de la búsqueda. Donde  $W^*$  (8), es una función de distribución de ofertas, tasa de descuento, tiempo planeado en la búsqueda y un vector de otras variables.

De esta forma para obtener la duración esperada se utiliza la intersección del beneficio marginal de la búsqueda y del costo marginal de la misma, para así determinar el salario de reserva óptimo ( $W^*$ ) y de esta forma reemplazar el mismo en la función de duración.

## 2. Análisis econométrico

El análisis econométrico del desempleo en Bogotá se desarrolla a la luz de un modelo de duración del desempleo que toma como base el salario de reserva. Con tal propósito, se estima el valor esperado de la duración del desempleo partiendo de un modelo de regresión paramétrica de duración, en el cual la tasa Hazard es explicada por medio de un conjunto de variables, en las que se incluye el salario de reserva y que, según la literatura, resultan relevantes. A continuación, se presenta el procedimiento econométrico utilizado.

### A. Conceptos básicos:

Siendo una oferta de trabajo una variable aleatoria que es la realización de la función  $\lambda(t)$  (que es continua en  $t > 0$ ), tomando como referencia que existe una función de distribución de los salarios de reserva  $F(w)$ , y habiendo el trabajador escogido un salario de reserva  $w^*$ , la probabilidad de que el trabajador acepte un empleo en ese periodo de tiempo viene dada por:

$$\theta(t) = \{1 - F(w)\}\lambda(t) \quad (9)$$

Donde  $\theta(t)$  se define como la probabilidad instantánea de que el trabajador obtenga empleo observando que hasta el momento se encontraba desempleado.

Teniendo en cuenta lo anterior, y considerando que no se pueden observar las ofertas de salario que se hacen en  $t$ , se utiliza la duración para poder llevar a cabo la construcción de la tasa Hazard y posteriormente la estimación del modelo. La duración, a través de la censura indica si la oferta de salario fue mayor o no al salario de reserva ( $w^*$ ). La censura es una variable que toma el valor de 0 cuando se tiene la duración completa para un individuo (consiguió empleo y por tanto se sabe cuánto duro desempleado) o, por el contrario, toma el

valor de 1 cuando el individuo sigue en estado de desempleo y por lo tanto no se sabe cuándo terminara la duración del individuo.

Al utilizar la duración como en este caso, se debe truncar la función de distribución hasta un momento “ $t_j$ ”, el cual es el momento del estudio, de manera que la duración de los individuos estudiados sea estrictamente mayor al tiempo de estudio.

Debido al truncamiento de la función de distribución de la duración, esta queda expresada como:

$$f(t) = \int_{t_j}^{\infty} \frac{g(t)}{1 - G(t)} dt \quad (10)$$

A partir de la cual se encuentra la función Hazard o de probabilidad instantánea de encontrar empleo:

$$\theta(t) = \frac{g(t)}{1 - G(t)} \quad (11)$$

De allí se saca a su vez la probabilidad de seguir desempleado después de un momento  $t$ :

$$1 - G(t) = e^{-\int_{t_j}^{\infty} \theta(t) dt} \quad (12)$$

En el análisis paramétrico de la duración del desempleo, que se explicará a continuación, se toma como punto central de interés la tasa Hazard que está condicionada a un conjunto de determinantes del desempleo (X).

#### B. Estimación a través del modelo AFT:

La estimación de la duración del desempleo se lleva a cabo mediante un modelo de regresión paramétrica. En este caso se utiliza el modelo Accelerated Failure Time (AFT). Este modelo tiene como característica que captura un Hazard base en función del error de especificación del modelo y un término de aceleración o desaceleración que muestra la variación de este en el tiempo.

El modelo de regresión AFT se especifica a partir de la ecuación (12)

$$\ln(t) = X'\beta + u \quad (12)$$

De manera que se llega a diferentes modelos AFT utilizando diferentes distribuciones del término de perturbación.

Teniendo en cuenta lo anterior, se puede reescribir la duración en términos de:

$$t = e^{(X'\beta)} v \quad (14)$$

Donde  $v = e^u$

Donde  $t$  tiene una tasa Hazard igual a:

$$\lambda(t) = \lambda_0(t e^{(-X'\beta)}) e^{(X'\beta)} \quad (15)$$

Donde  $v = t e^{(X'\beta)}$

Como ya se tiene la función de Hazard en términos del tiempo, se pasa a estimar la duración por medio de (16)

$$\hat{t} = e^{(X'\beta + E[ui])} \quad (16)$$

Donde  $E(ui) = \frac{\Gamma(1)}{\rho}$ , siendo  $\Gamma(1)$  la distribución Gamma evaluada en 1 y  $\rho$  el parámetro de forma de la distribución Weibull.

#### **IV. CONFORMACIÓN DE LA BASE DE DATOS**

Para la estimación de la duración en Bogotá en el tercer trimestre de 2017 se utiliza la información contenida en la Gran Encuesta Integrada Nacional de Hogares (GEINH) para los meses correspondientes a dicho trimestre. Se lleva a cabo la conformación de esta base de datos con dicho trimestre debido a que este no presenta variaciones estacionales y a la amplia utilización de este en la literatura. Para la conformación de la base de datos se utilizan datos mensuales, los cuales posteriormente se unen para conformar la base de datos final del trimestre. Ver anexo 1.

#### **V. CONJUNTO DE VARIABLES RELEVANTES PARA LA ESTIMACIÓN ECONÓMICA**

El modelo de duración del desempleo se estima a partir de un modelo AFT, en el cual la variable dependiente es la duración del desempleo y se define como:

1. Duración del desempleo: hace referencia a una dupla de variables que indican si la duración del desempleo para cada individuo está completa o no, y cuál fue la misma.
  - Censura: esta variable indica si la duración del desempleo para dicho individuo ya culminó o si por el contrario se encuentra censurada.
  - Duración: esta variable indica el tiempo que lleva el individuo buscando empleo, en caso de no haber conseguido aún, o cuanto duró buscando empleo dado que ya salió del estado de desempleo.

En particular, para este trabajo, los individuos que culminaron con su situación de desempleo se conforman con base al trimestre analizado.

El conjunto de variables explicativas que conforman la matriz de diseño X para cada una de las muestras se define a continuación:

2. Edad: La duración del desempleo, según la literatura, puede ser explicada mediante la edad, como un indicador de experiencia y conocimiento que el individuo tenga sobre el mercado laboral, así como la percepción de productividad que tiene el mercado laboral del individuo. Esta es una variable continua en el tiempo.

3. Educación (aeduc): Esta es una variable continua que mide los años aprobados de educación que tiene el individuo como medida de especialización o productividad.
4. Variables de Reemplazo: Las variables de reemplazo son aquellas variables que dan explicación al salario de reserva de un individuo. Se denominan variables de reemplazo ya que buscan determinar la urgencia que tiene un individuo para conseguir empleo, y por lo tanto explicar de cierta manera la duración del desempleo.
5.
  - Estado civil (estcivil): Una persona casada puede tener mayor urgencia para conseguir trabajo debido a que posee mayores responsabilidades y por tanto un salario de reserva más bajo. Esta variable toma el valor de 1 si el individuo se encuentra casado o vive en unión libre.
  - Jefe de Hogar (Jhogar): Al igual que en el caso anterior, esta variable puede tener una relación negativa con el salario de reserva debido a las obligaciones con las que debe cumplir el individuo.
  - Número de menores a cargo (Nmayor): Esta variable indica las personas dependientes del ingreso del individuo. Esta variable puede que aumente o disminuya el salario de reserva, dado que entre más personas a cargo se tengan, mayor va a ser el salario de reserva debido al costo de oportunidad implícito. Por el contrario, entre más individuos a cargo se tengan, puede haber un salario de reserva menor debido a la urgencia por conseguir trabajo.
  - Género (sm): Esta variable indica el género del individuo, tomando el valor de 1 para hombres y 0 para mujeres.

## **VI. RESULTADOS**

La estimación del modelo se lleva a cabo a través del modelo de falla acelerada o AFT que sigue una distribución Weibull, explicado previamente. La estimación de los modelos se realiza sobre la muestra total y sobre las sub muestras por género. Para las sub muestras, la variable sm quedara excluida.

Los cuadros 1, 2 y 3 muestran los resultados de la estimación del modelo considerando la muestra total y las sub muestras creadas por género respectivamente.

**Cuadro 1. Muestra total**

<b>_t</b>	<b>Coef.</b>	<b>P&gt; z </b>	<b>[90% Conf. Interval]</b>	
sm	-0,4748901	0,000	-0,6437822	-0,305998
Edad	-0,0199413	0,034	-0,0353949	-0,0044876
Edad2	0,0004086	0,002	0,0001885	0,0006287
estcivil	0,1853664	0,055	0,0262233	0,3445096
jhogar	-0,3735112	0,000	-0,5473409	-0,1996814
aeduc	-0,1161905	0,004	-0,1825153	-0,0498658
aeduc2	0,0049114	0,003	0,0021583	0,0076646
lnSw	0,0160019	0,823	-0,101886	0,1338903
lnSw2	-0,0030014	0,466	-0,0097794	0,0037767
Nmayor	-0,1788634	0,000	-0,2614309	-0,096296
_cons	4,327086	0,000	3,60387	5,050302
Parámetro de escala de la Weibull	1,008517		0,957701	1,062029
Parámetro de forma de la Weibull	0,9915553		0,9415942	1,044167

**Cuadro 2. Sub-muestra hombres**

<b>_t</b>	<b>Coef.</b>	<b>P&gt; z </b>	<b>[90% Conf. Interval]</b>	
Edad	-0,011294	0,498	-0,0387359	0,016148
Edad2	0,0002996	0,165	-0,0000549	0,0006541
estcivil	0,0571232	0,756	-0,2452559	0,3595022
jhogar	-0,3280584	0,080	-0,6359218	-0,020195
aeduc	-0,086711	0,160	-0,1881229	0,014701
aeduc2	0,0057803	0,046	0,0010166	0,0105441
lnSw	0,0244824	0,814	-0,1465126	0,1954773
lnSw2	-0,0064117	0,277	-0,0161079	0,0032844
Nmayor	-0,1733735	0,051	-0,3192969	-0,0274501
_cons	3,857547	0,000	2,726291	4,988803
Parámetro de escala de la Weibull	1,05397		0,9694693	1,145837
Parámetro de forma de la Weibull	0,9487933		0,8727247	1,031492

**Cuadro 3. Sub-muestra mujeres**

<b>_t</b>	<b>Coef.</b>	<b>P&gt; z </b>	<b>[90% Conf. Interval]</b>	
<b>Edad</b>	-0,0207506	0,072	-0,0397494	-0,0017517
<b>Edad2</b>	0,0004422	0,012	0,0001528	0,0007317
<b>estcivil</b>	0,2021437	0,169	-0,0395567	0,4438442
<b>jhogar</b>	-0,3581299	0,030	-0,6292967	-0,0870631
<b>aeduc</b>	-0,1805166	0,001	-0,2689418	-0,0920915
<b>aeduc2</b>	0,0062853	0,002	0,002902	0,0096686
<b>lnSw</b>	0,0899626	0,334	-0,0630583	0,2429834
<b>lnSw2</b>	-0,0042604	0,432	-0,0131792	0,0046584
<b>Nmayor</b>	-0,1893153	0,001	-0,2873856	-0,0912449
<b>_cons</b>	4,036503	0,000	3,114492	4,958513
<b>Parámetro de escala de la Weibull</b>	1,01931		0,9533927	1,089786
<b>Parámetro de forma de la Weibull</b>	0,9810554		0,9176116	1,048886

En el cuadro 1 se presentan los resultados de la estimación de la muestra total. Se encuentra que los hombres presentan una duración menor en cuanto al tiempo de búsqueda en comparación con las mujeres. Esto puede explicarse dado a un fenómeno cultural, ya que las mujeres presentan un salario de reserva mayor que el de los hombres porque son las principales encargadas de las labores domésticas y del cuidado de los niños.

La edad y la edad al cuadrado son variables significativas en este modelo. La estimación muestra que edad y edad al cuadrado tienen signos inversos, lo cual indica que la duración, en un principio, disminuye con la edad, y después de cierto punto comienza a aumentar. Acorde con la literatura presentada anteriormente (Tenjo et al., 2012), para las personas mayores a 30 años la duración aumenta con la edad, esto, debido a que la percepción de productividad de una persona con el tiempo tiende a ser menor que en una persona joven. De la misma forma, a una persona joven a medida que va ganando experiencia, presenta una disminución en la duración.

El ser jefe de hogar (jhogar) presenta signo negativo en todos los resultados. Esta variable tiene una fuerte influencia sobre la duración para todas las submuestras, y está ligada económicamente a que el jefe de hogar tiene otros individuos a su cargo y por lo tanto presenta mayores responsabilidades. Por esta razón, el individuo establece un salario de reserva menor y por ende la duración de la búsqueda a su vez será menor. Esta variable es la más influyente en cuanto a disminuir la duración después de la variable género.

Los años de educación (aeduc y aeduc2) resultan significativos para casi todas las muestras a excepción de los hombres, donde solo es significativa la variable en su forma cuadrática. Estas dos variables presentan coeficientes con signos opuestos, lo cual indica que los años de educación generan una disminución en la duración en un primer momento, pero a medida

que aumenta la educación, llega un punto en la que esta pasa a aumentar la duración. Esto se debe a que el individuo, después de cierta educación, pasa a estar sobre capacitado para muchos trabajos y por tanto se demora más tiempo en encontrar trabajo. Por otra parte, los años de educación conllevan a un mayor salario de reserva, y por ende a una mayor duración.

El logaritmo natural del ingreso familiar o salario de reserva ( $\ln Sw$ ), como se explicó previamente, se relaciona con la búsqueda de empleo ya que las personas utilizan su salario de reserva para financiar la búsqueda de empleo y, por lo tanto, entre mayor sea el salario de reserva, mayor será el tiempo de duración de la búsqueda. No obstante, en países como Colombia la búsqueda de empleo está estrechamente relacionada a las redes de contactos que posean los individuos, y por ende, un mayor salario de reserva puede implicar una mejor red de contactos y por lo tanto una menor duración, tal como se expone en (Tenjo et al., 2012). El modelo arroja que esta variable no resulta significativa bajo la transformación logarítmica para explicar la duración.

El estado civil ( $Estcivil$ ), variable que indica si el individuo se encuentra en unión marital o no, aumenta la duración del desempleo para la muestra total. Esto puede ser explicado debido a que un individuo en unión marital tiene mayores gastos y por lo tanto un mayor salario de reserva.

La variable de menores a cargo ( $N_{mayor}$ ), resulta significativa y tiene un efecto bastante considerable a la hora de explicar la duración, reduciendo la misma. Este efecto es explicable debido a que, ante la necesidad de recursos, el individuo tiende a disminuir su salario de reserva y a ser menos selectivo a la hora de buscar trabajo.

En los cuadros 2 y 3 se presentan los resultados de la estimación de las submuestras diferenciadas por género respectivamente. Con algunas excepciones, estos resultados confirman en general lo anteriormente mencionado.

El principal cambio cuando se comparan las submuestras por género es la edad y los años de educación. Estas resultan significativas para mujeres y no para hombres. Retomando lo anterior, se puede decir que una mayor edad, en un principio, reduce la duración para las mujeres, mientras que en la forma cuadrática se tiene que esta variable aumenta la misma tanto para ambos géneros. Esto indica que las mujeres jóvenes tienen una menor duración que las mujeres de mayor edad. Por otra parte, los años de educación exponen que, ante altos niveles de estos, hay una mayor duración tanto para ambos sexos, mientras que cambios en la educación básica solo reducen la duración en desempleo de las mujeres. Ambos fenómenos se pueden deber a que las mujeres en Latinoamérica tienden a formar hogares a una temprana edad (entre los 20 y 23 años) (García & Rojas, 2002) y, debido a que por lo general estas son las encargadas del cuidado del hogar, tienen un costo de oportunidad mayor al momento de acceder a un empleo (un mayor salario de reserva) y a una mayor educación, por lo que el mercado laboral tiene una mayor demanda por mujeres jóvenes con mayores niveles educativos.



## VII. PRONÓSTICO DE LA DURACIÓN DEL DESEMPLEO

A continuación, se presentan los pronósticos de la duración del desempleo por cada una de las sub muestras (hombres y mujeres). Cabe resaltar que, las variables estado civil, jefe de hogar y edad fueron fijados en rangos seleccionados y las variables restantes: salario de reserva y menores a cargo se consideraron en el promedio de la submuestra determinada por dichos rangos.

De manera general se puede observar que, la duración del desempleo es creciente con la educación y la edad. Por lo tanto, una persona con más años de educación o mayor edad tardará más en conseguir empleo en comparación con una persona que tenga menos años de educación o edad. Esto se debe a que con mayor educación los individuos establecen un salario de reserva mayor y a su vez las empresas tienen procesos de selección más complejos según las especificaciones del trabajo y las cualidades académicas del individuo. Por otra parte, el tiempo en desempleo de las mujeres es mayor al de los hombres.

En los hombres se observa que tienen una menor duración al ser jefes de hogar y no estar en una unión marital. Por otra parte, se evidencia una mayor duración para casi todas las edades en el caso de los hombres en unión marital, siendo mayor la duración de los hombres que no son jefes de hogar ni se encuentran casados a los 35 años. Finalmente, el ser jefe de hogar es un factor de gran importancia, ya que esto reduce la duración significativamente para las primeras tres edades (25, 35 y 45 años).

Para las mujeres se puede observar que presentan un aumento en la duración al encontrarse casadas y no ser jefes de hogar. Esto puede deberse al mayor costo de oportunidad que, como hemos mencionado anteriormente, se presenta por aspectos culturales. Por otra parte, el ser jefe de hogar genera una disminución en la duración, tal como sucede en los hombres.

Finalmente, las diferencias mas importantes frente a la duración entre hombres y mujeres es que estas ultimas tienen una mayor duración si son casadas y no son jefes de hogar, especialmente a la edad de 55 años. Además, la duración de los hombres, no excede en ninguno de los casos expuestos las 20 semanas, mientras que para las mujeres este es un factor recurrente a los 55 años.

HOMBRES											
Jefe de Hogar=0, Estado civil=0											
Jefe de Hogar=0, Estado civil=1											
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	11,61005	12,60328	14,54896	23,00443	11,54040	12,52767	14,46168	22,86643	12,13499	13,17313	15,20678
					24,04456	15,14495	16,44059	18,97867	30,00857		
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	13,20770	14,33761	16,55103	26,17006	11,13642	12,08914	13,95544	22,06598	18,61710	20,20978	23,32974
					36,88837	15,86375	17,22089	19,87943	31,43283		
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	8,36146	9,07678	10,47804	16,56760	8,26553	8,97264	10,35782	16,37752	10,66561	11,57805	13,36545
					21,13310	12,14781	13,18705	15,22284	24,06996		
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	9,51594	10,33003	11,92476	18,85512	9,89899	10,74584	12,40477	19,61409	11,10500	12,05503	13,91606
					22,00372	14,06926	15,27289	17,63069	27,87719		
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	16,48140	17,89138	20,65343	32,65666	19,23441	20,87991	24,10332	38,11153	18,51820	20,10243	23,20581
					36,83498	25,61597	27,80741	32,10028	50,75612		
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	18,14615	19,69855	22,73958	35,95523	18,81459	20,42418	23,57723	37,27970	23,12352	25,10174	28,97690
					45,81752	43,58637	47,31517	54,61962	86,36311		
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	11,02710	11,97046	13,81845	21,84937	11,99273	13,01870	15,02851	23,76269	13,42266	14,57097	16,82041
					26,59599	13,42266	14,57097	16,82041	26,59599	13,42266	14,57097
Edad	25	35	45	55							
Años de educación	0	5	11	16	0	5	11	16	0	5	11
Duración del desempleo	14,68558	15,94193	18,40302	29,09837	14,69629	15,95355	18,41643	29,11958	13,15633	14,28185	16,48666
					26,06827	20,85582	22,64003	26,13516	41,32424		

## **VIII. BONDAD DE AJUSTE**

La bondad de ajuste se hizo por medio de la diferencia entre la duración observada y la duración estimada en valor absoluto. Esta última, se construyó a partir de los coeficientes estimados y la matriz de diseño para cada uno de los individuos. Para hallar la diferencia se crea una variable y se establece la regla de que si el individuo tiene una diferencia en la duración menor a 24 semanas (6 meses) toma el valor de 1 y en caso contrario de 0. Por último, se calcula la media de esta variable para hallar el porcentaje correspondiente al ajuste del modelo, dando como resultado que el 96,039% de los individuos entran en el rango establecido de las 24 semanas.

## **IX. CONCLUSIONES**

Nuestros resultados sugieren que la duración del desempleo entre individuos tiende a aumentar principalmente con el género debido a los aspectos culturales y al costo de oportunidad que representa para las mujeres factores como la conformación de una familia. De esta forma, se sugiere implementar políticas de igualdad de género que propicien una equitativa distribución de las labores domésticas que culturalmente han sido otorgadas a las mujeres, y generar incentivos en el mercado laboral de manera que no haya una penalización a las mujeres que se ausenten del mercado laboral debido a la conformación de una familia o el nacimiento de un hijo.

La educación y la edad tienen una alta significancia en todas las estimaciones y representan la diferencia más destacada que se encuentra entre géneros. Los resultados dejan en evidencia que para las mujeres disminuye la duración de desempleo en la juventud y si presentan niveles de educación bajos. Esto se puede explicar debido a que las mujeres jóvenes en América Latina tienden a conformar hogares a una menor edad, generando una disminución en la oferta de mujeres en edad de trabajar, lo que da explicación a la mayor facilidad de estas mujeres para la consecución de empleo. Por otra parte, los resultados muestran que hay un punto de inflexión en la educación, en la que esta alarga la duración del desempleo debido a la poca disponibilidad de empleos que requieren tal nivel de estudios, por lo que es necesario fomentar la generación de empleos de alto nivel ya bien sea en empresas nacionales o por medio de la atracción de inversores internacionales al país.

También vale la pena resaltar que indistintamente del género, el hecho de ser jefe de hogar reduce la duración de desempleo. Esto, debido a que las personas jefes de hogar tienen una mayor responsabilidad frente a las personas que tienen a cargo, y debido a esto, establecen un menor salario de reserva, lo cual contribuye a tener una menor duración del desempleo. Como se puede comprobar, este fenómeno tiene un efecto mayor al de todas las variables.

Por otra parte, estado civil afecta la duración alargando la misma. Esto dado a que posiblemente, los individuos en unión marital dedican menor tiempo a la búsqueda de empleo, lo cual, genera un aumento en su duración o posiblemente debido a que tienen responsabilidades económicas más altas con sus conyugues y por ende establecen un salario de reserva más alto, lo cual alarga su duración.

Se concluye que el salario de reserva no influye de manera directa en la duración del desempleo. Sin embargo, es visible que la mayoría de los fenómenos observados en el estudio son explicados por medio de este, resaltando la importancia de este.

Finalmente, se plantea que es posible realizar un análisis de mayor profundidad a partir del modelo desarrollado en este artículo y adicionándole un modelo no paramétrico para llevar a cabo una comparación de resultados entre metodologías. Por otra parte, es posible concebir un análisis del mismo periodo entre diferentes años, con el fin de realizar una aproximación a la evolución del mercado laboral, así como trabajar en formas de incluir la duración del desempleo al indicador de desempleo como lo sugieren algunos autores aquí estudiados.

## X. REFERENCIAS

- Arango, Luis Eduardo, Posada, C. E. (2001). El Desempleo En Colombia. *Borradores de economica banco de la republica*, 176.
- Arango, L. E., & Posada, C. E. (2003). La participación laboral en Colombia. *Kavilando*, 1, 63-83.
- Arango, L. E., & Ríos, A. M. (2015). Duración del desempleo en Colombia : género , intensidad de búsqueda y anuncios de vacantes. *BID*.
- Arulampalam, W., & Stewart, B. M. (1995). The Determinants of Individual Unemployment Durations in an Era of High Unemployment. *The Economic Journal*, 105(429), 321-332.
- Banco de la República. (2013). Tasas de empleo y desempleo | Banco de la República (banco central de Colombia). Recuperado 15 de noviembre de 2018, de <http://www.banrep.gov.co/es/tasas-empleo-y-desempleo>
- Belzil, C. (2000). Unemployment Insurance and Subsequent Job Duration: Job Matching vs Unobserved Heterogeneity. *IZA*, (116).
- Cameron Colin, P. T. (2005). *Microeconometrics*. New York: CAMBRIDGE UNIVERSITY PRESS. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Castellar, C., & Uribe, J. I. (2006). Determinantes de la duración del desempleo en el área metropolitana de Cali 1988-1998. *Revista Sociedad y Economía*, (11), 8-38.
- DANE. (2018). Boletín técnico Cuentas Departamentales 2017 preliminar, 1-13.
- García, B., & Rojas, O. L. (2002). Cambios en la formación y disolución de las uniones en América Latina. *Papeles de población*, 8(32), 11-30. Recuperado de [http://www.scielo.org.mx/scielo.php?pid=S1405-74252002000200002&script=sci\\_arttext](http://www.scielo.org.mx/scielo.php?pid=S1405-74252002000200002&script=sci_arttext)
- Lancaster, B. Y. T. (1979). Econometric Methods for the Duration of Unemployment. *Econometrica*, 47(4), 939-956.
- Marcillo Yépez, E. (2011). Determinantes de la duración del desempleo en Colombia en sus trece principales áreas metropolitanas. *Gestión & Desarrollo*, 8(2), 61-79.
- Martínez, H. F. (2003). ¿Cuánto Duran Los Colombianos En El Desempleo Y El Empleo? : Un Análisis De Supervivencia. *IDEAS Working Paper Series from RePEc*, 7191, 1-42. [https://doi.org/10.1016/S0300-2896\(10\)70034-X](https://doi.org/10.1016/S0300-2896(10)70034-X)
- Narendranathan, W. (1993). Job Search in a Dynamic Environment-An Empirical Analysis. *Oxford Economic Papers*, 45(1), 1-22.
- Oviedo Gil, M. Y. (2007). EN EL MERCADO LABORAL COLOMBIANO 2003. *Sociedad y economía*, 22.
- Røed, K., & Zhang, T. (2003). Does unemployment compensation affect unemployment

duration? *Economic Journal*, 113(484), 190-206. <https://doi.org/10.1111/1468-0297.00086>

Shorrocks, A. (2009). On the measurement of Unemployment. *Journal of Economic Inequality*.

Tenjo, J., Misas, M., Contreras, A., & Gaviria, A. (2012). Modelos de duración del desempleo en Colombia. *Vniversitas Económica*, 12, 1-55.

Tenjo, J., & Ribero, R. (1998). Participación, desempleo y mercados laborales en Colombia. *Archivos de macroeconomía*, 1985-1995.

Viáfara, C. A., & Uribe, J. I. (2008). Duración del desempleo y canales de búsqueda de empleo en Colombia, 2006. *DNP, Documento*, 1-29.

## **XI. ANEXO 1.**

Conformación base de datos.

### **1. Base de datos individuos ocupados (no censurados).**

La base de datos de individuos no desempleados, no censurados, se construye a partir de la información obtenida en el módulo 60 de la encuesta, utilizando como llave de información: Directorio, secuencia\_p, orden, hogar y mes. Adicionalmente se toman como variables relevantes las preguntas: Factor de expansión, **p6426** “Tiempo que lleva desocupado en la actualidad” y **p760** “Tiempo que estuvo sin empleo entre el último trabajo y el actual”.

- Para el mes de Julio se toma que si **p760=.** (valor de omisión), se elimine la observación y si **p6426>1** a su vez se descarte la observación, de manera que se conforme la **Data1a**.
- Para el mes de agosto se toma que si **p760=.** (valor de omisión), se elimine la observación y si **p6426>2** a su vez se descarte la observación, de manera que se conforme la **Data1b**.
- Para el mes de septiembre se toma que si **p760=.** (valor de omisión), se elimine la observación y si **p6426>3** a su vez se descarte la observación, de manera que se conforme la **Data1c**.

Adicionalmente se crean las siguientes variables: **Duración= p760** (duración del desempleo en meses), **subsidio=0** (subsidios relacionados con desempleo) y **censura=0**. En el código de programación **Data1**.

### **2. Base de datos individuos desocupados (censurados).**

La base de datos de individuos censurados, desempleados, se construye a partir de la información obtenida en el módulo 70 de la encuesta, utilizando como llave de información: Directorio, secuencia\_p, orden, hogar y mes. Se utilizan las siguientes preguntas para generar

las variables relevantes: Factor de expansión, **p7250** “Semanas en que busco o ha estado buscando trabajo” y **p9460** “¿Recibe en la actualidad subsidio?”.

En caso de que **p7250=.** (valor de omisión), se elimina la observación, adicionalmente, si **p9460=.** (valor de omisión), se cambia por 0. Se procede a generar las siguientes variables: **Duración=(p7250)/4**, de manera que la duración quede con la misma periodicidad de la base de datos de ocupados. Además, se genera la variable **Censura=1**, y **Subsidio=1** en caso de que la respuesta a **p9460=1**. En el código de programación **Data2**.

### 3. Conformación base de datos ocupados y desocupados.

En esta base de datos se reúnen los datos anteriormente generados por medio de un merge o pegue, utilizando la llave de información para este fin. En el código de programación **Data3**.

### 4. Conformación base de datos de características generales.

El objetivo de esta base de datos es otorgarles a los individuos que conforman las anteriores bases de datos, algunas características generales que resulten relevantes a la hora de explicar la duración del desempleo. Estas características se obtienen del módulo 10 de la encuesta.

La base de datos se construye a partir de la información obtenida del módulo 10 de la encuesta, utilizando como llave de información: Directorio, secuencia\_p, orden, hogar y mes. Como preguntas relevantes se tienen: Factor de expansión, **p6040** “Años cumplidos”, **p6020** “Sexo”, **p6070** “¿Con quién vive?”, **p6210** “nivel de educación”, **p6220** “Títulos educativos”, **p6050** “Parentesco con el jefe de hogar” y **p6210s1** “Grados: años aprobados”. Con estas preguntas se procede a generar las variables: **edad=p6040**, **estcivil=1** si **p6070=1** o **2** o **3** (vive en unión marital), y 0 en caso de ser diferente, **jhogar=1** en caso de que **p6050=1** y 0 de lo contrario, **sm=1** si **p6020=1** y 0 de lo contrario (indicando que el individuo es mujer), se construyen a su vez **tit\_bach**, **tit\_tec**, **tit\_univ** y **tit\_post** utilizando la **p6210** de título alcanzado y finalmente, la variable **aeduc**, la cual indica los años de educación cursados y aprobados, se construye de la siguiente forma:

- Si **p6210=1** o **p6210=2**, **aeduc=0**
- Si **P6210=3** y **p6210s1=0**, **aeduc=0**
- Si **P6210=3** y **p6210s1>0**, **aeduc=p6210s1**
- Si **P6210=4** y **p6210s1=0**, **aeduc=5**
- Si **P6210=4** y **p6210s1>0**, **aeduc=p6210s1**
- Si **P6210=5** y **p6210s1=0**, **aeduc=9**
- Si **P6210=5** y **p6210s1>0**, **aeduc=p6210s1**
- Si **P6210=6** y **p6210s1=0**, **aeduc=11**
- Si **P6210=6** y **p6210s1>0**, **aeduc=11+p6210s1**

En el código de programación **Data4**.

## **5. Conformación de base de datos ocupados y desocupados con características generales.**

En esta base de datos se encuentra la información correspondiente a ocupados y desocupados, y se le agregan las características generales por medio de un merge o pegue de **Data3** y **Data4**. En el código de programación **Data5**.

## **6. Conformación de número de menores en el hogar y personas a cargo.**

La construcción de la base de datos del número de menores en el hogar y personas a cargo se lleva a cabo con base en el módulo 10 de la encuesta.

Se define **Nmayo=1** si la pregunta **p6040<12**, es decir, si el individuo es menor a 12 años. Luego se realiza la suma de personas menores en el hogar, **Nmayor** por medio del directorio, de manera que los individuos que tengan el mismo directorio tendrán el mismo número de menores en el hogar. En el código de programación **Data6**.

## **7. Conformación base de datos ocupados y desocupados con características generales y número de menores en el hogar.**

A la base de datos referenciada como **Data5** se le añade la **Data6** correspondiente a menores a cargo por medio de la llave de información. En el código de programación **Data\_m07**.

## **8. Ingreso.**

### **Ocupados: Data7**

Para la conformación de la base de datos de ocupados se utiliza el módulo 60 de la encuesta.

- Ingreso referente a ocupados asalariados: **Ingasal**

El ingreso correspondiente a ocupados asalariados se construye con base a las siguientes preguntas:

**P6500** “Antes de descuentos, ¿cuánto ganó en este empleo? (valor mensual)”

**P6510** “¿Recibió ingresos por concepto de horas extras? Si o no”

**P6510s1** “Valor”

**P6510s2** “¿Incluyó ese valor en p6500? Si o no”

**P6590** “¿El mes pasado recibió alimentos como parte de pago? Si o no”

**P6590s1** “¿En cuánto estima lo que recibió? Valor”

**P6600** “¿El mes pasado recibió vivienda como parte de pago de trabajo? Si o no”

**P6600s1** “¿En cuánto estima lo que recibió? Valor”

**P6610** “¿El mes pasado recibió vivienda como parte de pago de su trabajo? Si o no”

**P6610s1** “¿En cuánto estima lo que recibió? Valor”



**P6620** “¿El mes pasado recibió otros ingresos como parte de pago? Si o no”

**P6620s1** “¿En cuánto estima lo que recibió? Valor”

Otros subsidios:

**P6585s1** “¿Auxilio o subsidio de alimentación? Si o no”

**P6585s1A1** “Valor”

**P6585s1A2** “¿Incluyo ese valor en p6500? Si o no”

**P6585s2** “¿Auxilio o subsidio de transporte? Si o no”

**P6585s2A1** “Valor”

**P6585s2A2** “¿Incluyo el valor en p6500? Si o no”

**P6585s3** “¿Subsidio familiar? Si o no”

**P6585s3A1** “Valor”

**P6585s3A2** “¿Incluyo ese valor en p6500? Si o no”

**P6585s4** “¿Subsidio educativo? Si o no”

**P6585s4A1** “Valor”

**P6585s4A2** “¿Incluyo ese valor en p6500? Si o no”

¿El mes pasado recibió?

**P6545** “¿Prima (técnica, de antigüedad, clima, orden publica, otras)? Si o no”

**P6545s1** “Valor”

**P6545s2** “¿Incluyo ese valor en p6500? Si o no”

**P6580** “¿Bonificaciones mensuales? Si o no”

**P6580s1** “Valor”

**P6580s2** “¿Incluyo ese valor en p6500? Si o no”

¿En los últimos doce meses recibió?

**P6630s1** “¿Prima de servicios? Si o no”

**P6630s1A1** “Valor”

**P6630s2** “¿Prima de navidad? Si o no”

**P6630s2A1** “Valor”

**P6630s3** “¿Prima vocaciones? Si o no”

**P6630s3A1** “Valor”

**P6630s4** “¿Viáticos permanentes y/o bonificaciones? Si o no”

**P6630s4A1** “Valor”

**P6630s6** “¿Pagos por accidentes de trabajo? Si o no”

**P6630s6A1** “Valor”

- Ingreso referente a ocupados independientes: **Ingindep**

El ingreso referente a ocupados independientes se construye con base a las siguientes preguntas:

**P6750** “¿Cuál fue la ganancia neta o los honorarios de esta actividad? Valor”

**P6760** “¿A cuántos meses corresponde lo recibido? Meses”

**P550** “¿Cuál fue la ganancia neta del negocio o la cosecha durante los últimos 12 meses? Valor”

- Ingreso referente a trabajo secundario: **Ingsecun**

El ingreso referente a trabajo secundario se construye con base a las siguientes preguntas.

**P7040** “¿El mes pasado, además del trabajo principal, tenía otro trabajo? Si o no”

**P7070** “¿Cuánto recibió o ganó el mes pasado en ese segundo trabajo? Valor”

De esta forma la variable ingreso ocupados se construye de la siguiente forma:

**Data7=Ingasal+Ingindep+Ingsecun**

#### **Desocupados: Data8**

Para la conformación de esta base de datos se utiliza el módulo 70 de la encuesta. El ingreso de desocupados, **Data8**, se construye con base en las siguientes preguntas:

**P7422** “¿Recibió o gano el mes pasado ingresos por concepto de trabajo? Si o no”

**P7422s1** “Valor”

#### **Inactivos: Data9**

Para la conformación de esta base de datos se utiliza el módulo 80 y el módulo 95 de la encuesta. El ingreso de inactivos, **Data9**, se construye con base en las siguientes preguntas:

**P7472** “¿Recibió o ganó el mes pasado ingresos por concepto de trabajo? Si o no”

**P7472s1** “Valor”

#### **Otros ingresos: Data10**

Para la conformación de esta base de datos se utiliza el módulo 95 de la encuesta. Los otros ingresos, **Data10**, se construye con base en las siguientes preguntas:

- Ingreso referente a otros ingresos mes pasado: **Data10a**

**P7500** “¿El mes pasado recibió ingreso por arriendo de casa, apartamento, finca, etc.? Si o no”

**P7500s1** “Valor”

**P7500s2** “Pensión o jubilación por vejez, invalidez o sustitución pensional? Si o no”

**P7500s2A1** “Valor”

**P7500s3** “Pensión alimenticia por paternidad, divorcio o separación? Si o no”

**P7500s3A1** “Valor”

- Ingreso referente a otros ingresos últimos 12 meses: **Data10b**

El ingreso últimos 12 meses se divide en 12 para homogeneizar a mes.

**P7510s1** “¿Recibió dinero de otros hogares o personas residentes en el país? Si o no”

**P7510s1A1** “Valor”

**P7510s2** “¿Recibió dinero de otros hogares o personas residentes fuera del país? Si o no”

**P7510s2A1** “Valor”

**P7510s3** “¿Ayuda de dinero de instituciones del país o fuera del país? Si o no”

**P7510s3A1** “Valor”

**P7510s5** “¿Dinero por intereses de préstamos o CDT, depósitos de ahorro, utilidades? Si o no”

**P7510s5A1** “Valor”

**P7510s6** “¿Intereses cesantía? Si o no”

**P7510s6A1** “Valor”

**P7510s7** “¿Ingresos de otras fuentes? Si o no”

**P7510s7A1** “Valor”

Así, la variable de otros ingresos se construye de la siguiente forma:

**Data10=Data10a+Data10b**

## **9. Conformación de la base de datos de ingreso individual.**

Esta base de datos está conformada por el merge o pegue de la información anteriormente mencionada en el numeral 8, a través de la llave de información.

De esta forma se cuenta con la información necesaria para generar la siguiente variable:

**Ingresoindividuo=Data7+Data8+Data9+Data10**

#### **10. Conformación base de datos de ingreso familiar total**

Se genera la variable **Ingreso\_familiar\_total**, la cual es igual a la suma de los ingresos individuales correspondientes a cada observación del directorio que de acuerdo con la variable parentesco, este relacionado con el jefe de hogar con **P6050<6** y en caso contrario se le asigna como **Ingreso\_familiar\_total** el valor que tiene en **Ingresoindividuo**.

#### **11. Conformación de la data total por mes.**

Esta base de datos surge de adicionar la variable **Ingreso\_familiar\_total** a **Data-m07**. En el código de programación **DatatotX**, **X=1**, **X=2**, **X=3**.

#### **12. Conformación DATAGLOBAL**

Finalmente se reúne la información de las datas de cada mes para conformar la **DATAGLOBAL**, en donde queda contenida toda la información relevante para llevar a cabo el análisis y la estimación econométrica del modelo de duración del desempleo en Bogotá para el tercer trimestre de 2017.