

**Maestría en Economía**  
**Pontificia Universidad Javeriana**  
**Facultad de Ciencias Económicas y Administrativas**

**Impacto cuantílico de los determinantes de la duración del  
desempleo en Colombia**

**Trabajo elaborado por:**  
**Ricardo José Calderón Cómbita**  
**Alfonso Javier Peñuela Peña**

**Dirigido por:**  
**Martha Misas Arango**

**Julio de 2014**

## TABLA DE CONTENIDO

<b>1. INTRODUCCIÓN.....</b>	<b>3</b>
<b>2. REVISIÓN DE LA LITERATURA.....</b>	<b>6</b>
2.1 MODELOS PARAMÉTRICOS O SEMIPARAMÉTRICOS .....	6
2.2 MODELOS NO PARAMÉTRICOS .....	7
2.3 MODELOS DE DURACIÓN CON REGRESIÓN CUANTÍLICA .....	8
<b>3. MODELO BÁSICO DE BÚSQUEDA DE EMPLEO.....</b>	<b>9</b>
<b>4. REVISIÓN DE LOS MODELOS DE INTERÉS.....</b>	<b>12</b>
4.1 MODELO DE TIEMPO ACELERADO DE FALLA.....	12
4.2 REGRESIÓN CUANTÍLICA .....	13
4.3 REGRESIÓN CUANTÍLICA CENSURADA.....	15
<b>5. VARIABLES.....</b>	<b>18</b>
<b>6. RESULTADOS.....</b>	<b>21</b>
<b>7. ANÁLISIS DE LOS RESULTADOS.....</b>	<b>22</b>
7.1 COMPARATIVO PARA LA POBLACIÓN TOTAL .....	22
<b>8. CONCLUSIONES.....</b>	<b>26</b>
<b>REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>30</b>
<b>ANEXOS.....</b>	<b>33</b>

## 1. INTRODUCCIÓN

El fenómeno del desempleo suscita el interés de diferentes actores, sectores y grupos de la economía, así como diversas posturas desde una perspectiva académica. Desde las políticas o medidas de gobierno orientadas a su mitigación (que incluyen sistemas de incentivos para los empleadores, hasta la ampliación del Estado como principal empleador), pasando por los criterios y condiciones de los actores privados para la incorporación de personal a sus empresas, las organizaciones sindicales y actores sociales y políticos que reclaman medidas gubernamentales para su reducción o supresión, hasta las decisiones de las personas en la elección o rechazo de opciones laborales, todos manifiestan interés y preocupación por lo que atañe al empleo y, su contraparte, el desempleo. A su vez, desde una perspectiva académica, la valoración sobre el empleo, las condiciones que lo hacen posible, los límites económicos a su expansión, la conveniencia o posibilidad del pleno empleo, son aspectos que varían ampliamente.

Dada la importancia que reviste el fenómeno del desempleo, los intentos por ofrecer una explicación más precisa de sus características requieren especial atención. Tal es el propósito del presente trabajo. No pretende agotar todas las facetas del fenómeno sino, más bien, concentrarse en un aspecto muy específico del mismo, dentro del marco de la econometría. Por medio de una aproximación técnica, pretende ofrecer elementos explicativos en una de sus variables más sobresalientes: la *duración*.

La duración, que puede definirse en términos muy generales como *el tiempo que un individuo permanece en un estado determinado antes de pasar a otro*, es una variable cuyo estudio detallado es imprescindible en diversas áreas del conocimiento, sean estas naturales o sociales, y la economía no es la excepción. Existe abundante literatura que intente modelarla y darle adecuada explicación [1]. En nuestro contexto específico entenderemos por duración como el tiempo que permanece un individuo en desempleo.

Debemos ahora introducirnos en una pequeña consideración un poco más técnica. Dado que nuestro propósito es aportar a la explicación de la duración del desempleo, debemos prestar cuidadosa atención a las relaciones que se establecen entre el comportamiento de la variable *duración en desempleo* y el comportamiento de las *condiciones y características* de los individuos considerados. Cotidianamente, sabemos que el hecho de que una persona o un grupo social permanezca poco o mucho tiempo en el desempleo es algo que depende de, o se ve fuertemente afectado por, ciertas condiciones o características de dicha persona o grupo. También sabemos que si estas condiciones cambian, la duración en el estado de desempleo cambia a su vez. Sin embargo, conviene preguntarnos, ¿es posible estudiar estas relaciones con mayor precisión? ¿Podemos estimar un patrón o regularidad entre la duración de desempleo y sus condiciones? Para lograrlo, es necesario el uso de modelos, estructuras matemáticas, estadísticas o econométricas, que intentan aproximarse a las relaciones que existen entre, la duración del desempleo, y las condiciones y características de los individuos.

Ahora bien, las técnicas difieren en su conveniencia para explicar un fenómeno: unos pueden ser más apropiados para ciertos efectos y otros para otros. En nuestro caso, pretendemos examinar las virtudes que presentan dos modelos explicativos de la duración del desempleo. Uno de ellos ha recibido mayor atención (estimación de la media) y se encuentra mejor comprendido, incluso en sus limitantes (manejo de datos censurados); el otro, dada su complejidad, requiere un mayor estudio (estimación por regresión cuantílica censurada).

Este trabajo muestra el impacto diferenciado que tienen las variables fundamentales que explican la duración del desempleo en los diferentes cuantiles de su distribución, y lo compara con los impactos en la media. En términos econométricos, se realiza la estimación de un modelo de duración usando regresión cuantílica censurada, lo cual permite tener un panorama diferente, a los implementados por los investigadores en la actualidad, de la duración del

desempleo en Colombia. Hasta el momento, los modelos se han estimado exclusivamente para la media de la distribución, y este hecho, para efectos de políticas orientadas a mejorar la calidad del empleo y de la vida por medio de la reducción de la duración del desempleo, pueden resultar no equitativos para ciertos grupos.

En este artículo se contrastan los dos tipos de estimación, por medio de un modelo acelerado de falla, clásico, y uno estimado por regresión cuantílica censurada. Aunque en ambos casos se encuentran datos censurados de duración, éstos han recibido un tratamiento a profundidad en el caso de la visión clásica (es decir, en el modelo centrado en la estimación de la media) [3], y no tanto en el modelo de regresión cuantílica censurada. En particular, estos modelos presentan el problema de enfrentar una función objetivo no convexa. Esta característica implica usar un algoritmo específico en la estimación de los parámetros en cada cuartil. El algoritmo es desarrollado por Peng-Huang (2008), éste se basa en la teoría de Martingalas y la función de distribución Hazard acumulada de Nelson-Allen. En este trabajo, en particular, la censura se da en aquellos individuos que continúan en desempleo al momento de aplicar la encuesta<sup>1</sup>.

Los datos han sido tomados de la gran encuesta integrada de hogares para el tercer trimestre del año 2012, aplicada a las trece áreas metropolitanas y cabeceras. A partir de éstos podemos configurar nuestra variable de interés, esto es, el tiempo que han durado en desempleo los individuos y las principales características de los mismos que, según la literatura, resultan más importantes en la explicación de la duración del desempleo. Con base en los datos se estiman los parámetros, por una parte, de un modelo estándar con interés en la media de la distribución y, por otra, un modelo usando regresión cuantílica censurada, a fin de realizar un contraste de resultados que nos muestren las posibles diferencias.

El trabajo se divide del siguiente modo. La segunda sección contiene la revisión de la literatura, enfocada en las técnicas econométricas. La tercera

---

<sup>1</sup> La censura se define siguiendo la referencia bibliográfica [22].

sección expone la línea base del modelo de búsqueda de empleo. La cuarta presenta los modelos de duración y regresión cuantílica, así como la explicación del problema del modelo de regresión cuantílica censurada. La quinta describe las variables explicativas, seleccionadas para la estimación de los impactos en la duración del desempleo en los diferentes cuantiles de la distribución. La sexta sección presenta los gráficos de los resultados de estimar el modelo. La séptima sección presenta el análisis de los resultados del modelo por cuantiles y la comparación con el modelo en la media. En la octava sección se presentan las conclusiones. En la novena sección se presentan las referencias bibliográficas.

## **2. REVISIÓN DE LA LITERATURA**

La revisión bibliográfica se enfoca en las técnicas econométricas, paramétricas o no paramétricas, usadas en la estimación de la duración del desempleo tanto a nivel internacional como nacional.

### **2.1 MODELOS PARAMÉTRICOS O SEMIPARAMÉTRICOS**

Estos modelos tienen sus primeros referentes en Ehrenberg y Oaxaca (1976), donde se hace la estimación por medio de mínimos cuadrados ordinarios, al log-linealizar el modelo especificado para los niveles óptimos de seguros de desempleo. Un enfoque de tasas de riesgos que se presenta en Lancaster (1979) donde se hace el supuesto de proporcionalidad tipo Cox y se evalúa la variación de las posibilidades que tiene un desempleado de recuperar la condición de trabajador. En Rea y Ham (1987) usan éste enfoque para estimar el impacto de los cambios en las políticas de adjudicación de seguros de desempleo, en Canadá, teniendo en cuenta las variaciones en las probabilidades de salir del desempleo. El trabajo de Meyer (1990) busca la medición del efecto de los beneficios del seguro de desempleo sobre la duración del desempleo. En tanto que, en el trabajo de Pollman-Schult y Büchel (2005), consideran este tipo de modelos para estimar los efectos de los beneficios recibidos por estar desempleado en la calidad del próximo trabajo.

En el caso colombiano los estudios sobre estos temas tienen su inicio en Tenjo y Ribero (1998), donde se estima el efecto de los determinantes de la participación laboral y el desempleo con un modelo de duración tipo Weibull. Núñez y Bernal (1998), ajustan un modelo de tasa de riesgo y uno de riesgos proporcionales de Cox, con el objetivo de estimar el impacto de los determinantes en la duración del desempleo y en la probabilidad de conseguir empleo en Colombia. Castellar y Uribe (2003) seleccionan un modelo de duración tipo Weibull para encontrar los determinantes más representativos en la duración del desempleo en el área metropolitana de Cali – Colombia. Otro trabajo, que hace su aporte a nivel nacional, es el de Martínez (2003), donde se presenta un análisis de los determinantes de la duración del desempleo y empleo para identificar grupos de individuos por género, con mayores duraciones usando modelos de duración tipo Gompertz y Weibull. De la misma forma, Viáfara y Uribe (2008) utilizan modelos de duración de riesgos proporcionales tipo Weibull para analizar el efecto del canal de búsqueda de empleo en la duración del desempleo. El trabajo de Tenjo *et al.* (2012) estima un modelo de tiempo acelerado de falla, o AFT, para evaluar el impacto de los determinantes en la duración del desempleo, y un modelo Probit para medir la probabilidad de que una persona se encuentre desempleada dados algunos determinantes del desempleo.

## **2.2 MODELOS NO PARAMÉTRICOS**

Pollman-Schult y Büchel (2005) consideran el análisis Kaplan – Meier para estimar y cotejar las distribuciones de probabilidad de salir del desempleo por las principales características de los beneficiarios o no beneficiarios de los auxilios por desempleo. Un trabajo que sigue esta línea es el de van den Berg y Dolton (2006), los autores comparan la distribución de la duración del desempleo para cada uno de los determinantes especificados bajo el mismo análisis.

Para el caso colombiano, el estimador Kaplan – Meier es utilizado por Martínez (2003), con el propósito de comparar las probabilidades de salir del desempleo de acuerdo con los determinantes de la duración del desempleo y el empleo, Viáfara y Uribe (2008), para analizar el efecto del canal de búsqueda de

empleo y de los determinantes en la distribución de la duración del desempleo, y por Tenjo *et al.* (2012) con el fin de identificar y comparar el impacto de los determinantes en la distribución de la duración del desempleo.

### **2.3 MODELOS DE DURACIÓN CON REGRESIÓN CUANTÍLICA**

Koenker y Biliás (2001), pioneros de la regresión cuantílica, argumentan la importancia de los métodos de regresión cuantílica en el análisis de duración y describen un procedimiento tradicional para estimar el efecto, por cuartiles, de la entrega de los bonos del nuevo empleo. Tomando como base el anterior trabajo, Fitzenberger y Wilke (2005) implementan un nuevo procedimiento o algoritmo de estimación de la regresión cuantílica en datos de duración del desempleo. Varios autores exponen el problema de la regresión y plantean algoritmos para su solución. Tal es el caso de Stengos y Dianqin (2007), quienes presentan el problema de la regresión cuantílica censurada mediante un ejemplo, donde se permite entender de manera gráfica y analítica el problema de la no convexidad de la función objetivo, lo que conlleva a un proceso de minimización más complejo que el usado en regresión cuantílica clásica. Además, proponen una solución, por medio de un nuevo algoritmo iterativo, que busca mínimos locales iniciando en un punto dado y generando recurrentes direcciones con el objetivo de minimizar la función de interés. Wichert y Wilke (2008) presentan una extensión del estimador convencional de Kaplan - Meier, no paramétrico, para la tasa de riesgo y la función cuantílica condicional. Dicha extensión se lleva a cabo considerando censura tanto en la variable dependiente como en la independiente. Bajo simulaciones se observa que el estimador tiene propiedades deseables en muestras finitas, y su implementación es sencilla.

Por último, el trabajo de Peng y Huang (2008) propone una metodología para regresión cuantílica censurada basada en el estimador Nelson-Aalen, no paramétrico, de la función de la tasa de riesgo, o *hazard rate*, acumulada. La metodología incorpora ecuaciones Martingala, que conllevan a un algoritmo de minimización de funciones convexas. Establecen una consistencia uniforme y convergencia de los estimadores obtenidos, y evalúan el rendimiento en muestras

finitas a través de extensos estudios de simulación. Este trabajo es de vital importancia en el estudio, debido a que es justamente el algoritmo que usamos en la estimación de los parámetros.

Es importante tener en cuenta que los artículos que consideran el caso colombiano se basan en técnicas clásicas, paramétricas o no paramétricas, de duración. Lo cual conlleva a que los resultados de dichos modelos se refieran a comportamientos en la media de la duración del desempleo. Este trabajo se diferencia de los ya existentes en Colombia en la forma como se estiman los parámetros del modelo de duración incorporando la regresión cuantílica.

### **3. MODELO BÁSICO DE BÚSQUEDA DE EMPLEO <sup>2</sup>**

Dentro del estudio de la economía de los mercados laborales aparece el modelo básico de búsqueda de empleo, el cual, bajo ciertos supuestos, permite dar un sustento teórico a la decisión de un desempleado de rechazar o aceptar una oferta laboral. En este modelo el desempleado toma dicha decisión luego de un proceso de maximización de sus beneficios esperados bajo la condición restrictiva de aceptar la oferta sólo cuando el salario ofrecido por el empleador sea mayor o igual a su salario de reserva.

Este modelo involucra los costos asociados a la búsqueda de empleo y los relaciona con la duración del desempleo para restarlos posteriormente de los ingresos percibidos por el trabajador una vez inicie su trabajo. Así, el salario óptimo de reserva se alcanza maximizando, en valor presente, el beneficio neto esperado.

Teniendo en cuenta el modelo de economía laboral expuesto anteriormente y los fundamentales del desempleo, estudiados a nivel nacional e internacional, se estiman y comparan los impactos en la media y en los diferentes cuantiles de la distribución de la duración del desempleo en Colombia.

---

<sup>2</sup> El enfoque del modelo básico de búsqueda de empleo presentado sigue las referencias bibliográficas [6] y [22].

A continuación, de manera simple, se presenta la estrategia establecida por el modelo básico de búsqueda de empleo: Se define la función de densidad  $g(w)$  con una media,  $\mu$ , y una desviación,  $\sigma$ , la cual describe la percepción del individuo de la distribución de ofertas salariales en un momento determinado.

Para la búsqueda de empleo, el individuo selecciona un salario de reserva  $w_r$  tal que: Si  $w \geq w_r$ , el individuo acepta la oferta. En caso contrario la rechaza.

La probabilidad de que la oferta sea rechazada es:

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

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

$$\Pr(w \geq w_r) = 1 - G(w_r) \quad (2)$$

Si las ofertas son un proceso aleatorio, el valor esperado del número de ofertas que se reciben antes que una sea aceptada es:

$$D = E(L) = \frac{1}{1-G(w_r)} \quad (3)$$

Suponiendo que el individuo recibe una oferta por período,  $D$  se interpreta como la duración esperada de la búsqueda. Si este supuesto no se cumple, entonces la duración de la búsqueda es una función de  $D$ .

El salario esperado,  $w_e$ , que obtiene el individuo una vez reciba una oferta aceptable es:

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

Una forma de plantear el problema de la determinación del salario de reserva,  $w_r$ , del individuo es suponer que él fija su salario de reserva de tal manera que maximiza los beneficios netos futuros de la búsqueda, diferencia entre beneficios y costos, medidos en términos de valor presente.

Los procesos de búsqueda tienen costos directos y de oportunidad. Los costos directos son los asociados con la búsqueda misma, como la adquisición de información (consulta de avisos, bases de datos, etc.), la generación de ofertas (envío hojas de vida, entre otras cosas). Los costos de oportunidad están asociados con los ingresos que deja de percibir el individuo por dedicar su tiempo a la búsqueda en vez de trabajar, es decir, el salario que deja de percibir por período. En general, los costos directos se pueden considerar como fijos, por período, pero el costo de oportunidad crece con la duración de la búsqueda. Si suponemos que se recibe una oferta por período de tiempo, podemos expresar los costos,  $C$ , como una función de la duración de la búsqueda:

$$C = FD + h(D) = C(w_r) \quad \text{y} \quad \frac{\partial C}{\partial w_r} > 0 \quad (5)$$

donde  $F$  representa el costo fijo directo por período,  $h$  es el costo de oportunidad, que es función de la duración.

Los beneficios,  $Y$ , de la búsqueda están asociados con los ingresos que recibe el individuo y se pueden medir como una función de  $w_e$ .

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

El individuo determina su salario de reserva, óptimo, maximizando el valor presente del beneficio neto,  $V(B)$ , de la búsqueda, dado por:

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

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

El valor óptimo del salario de reserva,  $w_r^*$ , se obtiene resolviendo la ecuación:

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

Esto es equivalente a igualar el beneficio marginal con el costo marginal de la búsqueda. La ecuación (8) indica que el salario de reserva óptimo,  $w_r^*$ , es una función de la distribución de ofertas,  $g$ , la tasa de descuento,  $r$ , el horizonte para la toma de decisiones,  $N$ , el tiempo relevante para efectos de la planeación de la búsqueda y duración del empleo resultante, y un vector de otras variables,  $X$ .

Una vez determinado  $w_r^*$ , se puede obtener el valor de  $D$  reemplazando en la ecuación (3).

De la misma manera, la duración esperada de la búsqueda es:

$$D = d(g, r, N, X) \quad (9)$$

Una extensión del modelo es estimar la media,  $\mu$ , como una función de variables de capital humano y otras que determinan los niveles de ingresos, por ejemplo: género, años de estudio, jefe de hogar, etc. Esta es una forma de incluir las principales características de los individuos en los mercados laborales. En el trabajo se especifican y estiman los parámetros de los modelos de duración con el objetivo de realizar la extensión anteriormente expuesta.

## 4. REVISIÓN DE LOS MODELOS DE INTERÉS

### 4.1 MODELO DE TIEMPO ACELERADO DE FALLA <sup>3</sup>

Para analizar el tiempo final de ocurrencia,  $T$ , de un evento, en donde el evento es salir del desempleo y el tiempo hace referencia a la duración del desempleo, en semanas, hasta que ocurre el evento anteriormente descrito, se usa un modelo *Accelerated Failure Time* o modelo AFT. Adicionalmente, el evento puede ser recurrente o no, y, además, el tiempo de duración puede ser causado por diversos determinantes, y éstos pueden ser cuantitativos o cualitativos que dependan o no del tiempo en el que se observan.

---

<sup>3</sup> El enfoque del modelo se realiza siguiendo las referencias bibliográficas [1] y [4].

El modelo *Accelerated Failure Time* o modelo AFT, se especifica como:

$$\ln(t) = X'\beta + \ln(\tau) \quad (10)$$

Una característica de este modelo, se presenta en el término de perturbación  $\ln(\tau)$ , o error, debido a que diferentes distribuciones de probabilidad de  $\tau$ , generan diferentes modelos AFT.

La duración puede expresarse como:

$$t = \exp(X'\beta)\tau \quad (11)$$

La palabra “acelerado” se utiliza en la descripción de estos modelos porque, en lugar de asumir que el tiempo de falla  $t$  es exponencial, *weibull*, o alguna otra forma, se asume que la distribución es:

$$\tau = \exp(-X'\beta)t \quad (12)$$

y la expresión  $\exp(-X'\beta)$  es llamada parámetro de aceleración.

A continuación se describen las características del parámetro de aceleración para el individuo  $j$ :

- Si  $\exp(-X'_j\beta) = 1$ , entonces  $\tau_j = t_j$ , indicando que el tiempo pasa a su ritmo “normal”.
- Si  $\exp(-X'_j\beta) > 1$ , entonces el tiempo pasa más rápido para el individuo  $j$ , el tiempo se acelera, y se espera que la falla ocurra muy pronto.
- Si  $\exp(-X'_j\beta) < 1$ , entonces el tiempo pasa más despacio para el individuo  $j$ , el tiempo se desacelera, y se espera que la falla ocurra más tarde.

## 4.2 REGRESIÓN CUANTILICA

La regresión cuantílica, desarrollada por Konker y Baset (1978), se puede definir, en términos generales, como una técnica estadística que estima las funciones cuantílicas de una variable dependiente condicional a la información

observada por las variables explicativas en el contexto de un modelo lineal. Es decir, se cuenta con  $(Y_i, x_i)$  donde  $Y_i$  es la variable dependiente y  $x_i$  es un vector de covariables. A diferencia de la regresión lineal clásica, que estima la función de media condicional  $E(Y_i|x_i)$  a través de sus covariables de manera lineal, ahora se busca estimar las funciones cuantílicas condicionales  $Q_\theta(Y_i|x_i)$  por medio de parámetros  $\beta(\theta)$  para cada  $\theta \in (0,1)$ . Este problema matemáticamente puede resumirse en:

$$\min_{(\beta)} \sum_{i=1}^n \rho_\tau(y_i - x_i' \beta_\tau) \quad (13)$$

donde,  $\rho_\tau(u) = u(\tau - I(u < 0))$ , e  $I$  : Función indicadora clásica.

La expresión (13) es una suma ponderada de  $\tau$  con  $y_i - x_i' \beta_\tau$  cuando  $y_i > x_i' \beta_\tau$ , y  $(1 - \tau)$  con  $(x_i' \beta_\tau - y_i)$ , cuando  $y_i \leq x_i' \beta_\tau$ .

La solución del problema aporta los parámetros que permiten concluir el impacto lineal de las covariables sobre los diferentes cuantiles, para diferentes elecciones de  $\tau$ , de la distribución de la variable endógena. Esta solución se encuentra por medio de programación lineal, y no por los métodos clásicos, pues no superan el punto de la no diferenciabilidad.

El objetivo principal de la regresión cuantílica es responder la siguiente pregunta: *¿qué tan importante es cada una de las variables, cuando nos encontramos ubicados en diferentes cuantiles de la distribución de la variable endógena, Y?*

Para dar claridad acerca del objetivo de la regresión cuantílica, supongamos que la estimación de un modelo lineal, usando mínimos cuadrados, es  $\hat{Y} = 0.7425 + 1.0527X_1 + 2.0875X_2 + 1.3211X_3 - 1.8823X_4$ . Este modelo nos dice cómo las diferentes variables afectan nuestra variable dependiente  $Y$  en promedio. Ahora, supongamos que en el percentil 90 la estimación para el coeficiente de  $X_1$  fue de 2.0482. Esto nos dice que para el percentil 90 de la distribución de  $Y$  la regresión lineal subestima el parámetro llevando así a

conclusiones equivocadas para los individuos que se encuentran en dicho punto de la distribución.

Con el ejemplo anterior, podemos concluir que la regresión cuantílica permite hacer un análisis diferenciado, debido a que, si nos centramos en el promedio, probablemente este resultado no logre decir mucho sobre los individuos con medidas de Y que se encuentren distantes del promedio.

En nuestro contexto, la variable de interés es la duración del desempleo y queremos identificar las posibles diferencias en los impactos de las variables explicativas a través de los diferentes cuantiles de la distribución de la duración. El problema que se estudia en este trabajo presenta otra dificultad que la regresión cuantílica no logra superar porque no se cuenta con todas las observaciones completas de la variable dependiente debido a la censura, que es generada por aquellos individuos cuya situación de desempleo no ha concluido en el momento de la encuesta. Esto nos lleva a un marco diferente de estimación, un poco más complejo, que estudiaremos a continuación.

### **4.3 REGRESIÓN CUANTILICA CENSURADA**

Cuando se tienen datos censurados, la estimación por medio de regresión cuantílica es compleja, debido a la forma de la función a optimizar, porque además de no ser diferenciable tampoco es convexa. En nuestro caso particular, existe censura en la duración del desempleo porque los individuos que se encuentran en desempleo, en el momento de la encuesta, solo aportan el dato del tiempo que llevan en ese estado, pero no se sabe a ciencia cierta cuánto más pueden durar en ese estado, si saldrán pronto o no; en pocas palabras, no se conoce la duración del desempleo de estos individuos.

El problema de regresión cuantílica censurada fue estudiado en su inicio por Powell (1986), quien propuso el siguiente modelo lineal, con variable latente:

$$T_i = x_i' \beta + u_i \quad (14)$$

donde los errores,  $u_i$ , tienen una función de distribución  $F$ , y donde lo observado es de la forma  $Y_i = \max\{T_i, yc_i\}$ , donde  $yc_i$  son los puntos de censura. Como el objetivo de la regresión cuantílica es la estimación de las funciones cuantílicas condicionales de la variable dependiente  $Y_i$ .

Powell (2003) encuentra que la solución a este problema es minimizar la función  $\varphi$ , que describiremos a continuación. El problema de optimización puede ser expresado de la siguiente manera:

$$\min_{\beta} \varphi(\beta) \quad (15)$$

donde:

$$\varphi(\beta) = \sum_{i=1}^n (\theta * I(d_i > 0) + (1 - \theta) * I(d_i < 0)) |d_i| \quad (16)$$

$$d_i := y_i - \max(x_i' \beta, yc_i)$$

$$\theta \in (0,1) := \text{cuantil.}$$

$I$ : Función indicadora clásica.

$yc_i$ : Punto de censura;  $y_i$  está censurado por debajo en  $cy_i$

La función objetivo,  $\varphi(\beta)$ , presenta los problemas<sup>4</sup> de no diferenciabilidad y no convexidad que es causada por la censura.

Diversos autores han desarrollado algoritmos que permiten la optimización de funciones con éstas particularidades. Una lista de los algoritmos existentes, se puede consultar en la referencia bibliográfica [21], donde se muestran sus respectivas virtudes y falencias en términos de convergencia y tiempos de la estimación bajo regresión cuantílica censurada.

---

<sup>4</sup> En el anexo 1, por medio de un ejemplo se ilustran los inconvenientes de la función objetivo  $\varphi(\beta)$ .

El algoritmo que se utiliza en la estimación de los parámetros en la regresión cuantílica censurada, se basa en el estimador de Nelson-Aalen para la función de distribución acumulada Hazard. El estimador anteriormente mencionado es desarrollado por Peng y Hung (2008), quienes fusionan las teorías de la distribución acumulada Hazard, Martingalas y la optimización lineal.

El algoritmo que se utiliza en la estimación de los parámetros en la regresión cuantílica censurada, se basa en el estimador de Nelson-Aalen para la función de distribución acumulada Hazard. El estimador anteriormente mencionado es desarrollado por Peng y Hung (2008), quienes fusionan las teorías de la distribución acumulada Hazard, Martingalas y la optimización lineal.

#### 4.4 ALGORITMO DE ESTIMACIÓN

A continuación, se muestra de forma general el algoritmo<sup>5</sup> de estimación:

- Sea  $F_i(t|x) = P(T_i \leq t|x_i)$ , que mide la probabilidad de que el  $i$ -ésimo individuo tenga un tiempo de sobrevivencia  $T$  menor o igual a un valor específico  $t$ .
- Sea  $\Lambda_i(t|x_i) = -\log(1 - F_i(t|x_i))$ , que mide la tasa hazard acumulada para el  $i$ -ésimo individuo.
- Sea  $\Delta_i = I\{T_i < C_i\}$ , el indicador de censura.
- Sea  $N_i(t) = I\{T_i \leq t\} y \{\Delta_i = 1\}$ , un proceso de conteo.
- Fleming y Harrington (1991) demuestran que  $M_i(t) = N_i(t) - \Lambda_i(t \wedge Y_i|x)$  es un proceso Martingala asociado con el proceso de conteo  $N_i(t)$ . Donde  $\wedge$  significa operador mínimo.

---

<sup>5</sup> La presentación del algoritmo se realiza siguiendo las referencias bibliográficas [17] y [20]. Adicionalmente, en la referencia bibliográfica [20] se encuentra el método de estimación de los intervalos de confianza.

- Peng y Huang (2008) toman el proceso Martingala,  $M_i(t)$ , y proponen la siguiente ecuación de estimación:

$$n^{-1/2} \sum_{i=1}^n x_i \left[ N_i \left( \exp(x_i' \beta(\tau)) \right) - \int_0^\tau I \left( Y_i \geq \exp(x_i' \beta(\tau)) \right) dH(u) \right] = 0 \quad (20)$$

con  $H(u) = -\log(1 - u)$  y  $u \in [0,1)$ .

En la ecuación (20), se define  $t = \exp(x_i' \beta(\tau))$  por las propiedades del modelo AFT.

- La integral  $\int_0^\tau I \left( Y_i \geq \exp(x_i' \beta(\tau)) \right) dH(u)$ , se aproxima usando la siguiente malla:  $0 = \tau_0 < \tau_1 < \dots < \tau_j < 1$ .; y luego se encuentra el valor numérico recurrentemente de la siguiente expresión:

$$\alpha_i(\tau_k) = \sum_{j=0}^{k-1} I \left( Y_i \geq \exp \left( x_i' \hat{\beta}(\tau_j) \right) \right) \left( H(u_{j+1}) - H(u_j) \right) \quad (21)$$

- Los  $\beta(\tau_k)$ , con  $k = 1, \dots, M$ , pueden ser estimados resolviendo secuencialmente el siguiente problema de programación lineal:

$$n^{-1/2} \sum_{i=1}^n x_i \left[ N_i \left( \exp(x_i' \beta(\tau)) \right) - \alpha_i(\tau) \right] = 0 \quad (22)$$

## 5. VARIABLES

Teniendo en cuenta que el objetivo del trabajo es medir la duración del desempleo en Colombia, a través de sus fundamentales, con una perspectiva cuantílica, nuestra variable independiente será la duración del desempleo. Es importante recordar que en algunos casos se conoce la duración total del desempleo, pero, en otros, cuando los individuos aún siguen en desempleo, sólo se conoce la duración parcial. Este último evento se conoce como censura.

Las variables que se usan como explicativas en este trabajo coinciden con las propuestas en la referencia bibliográfica [22], a propósito de la duración del desempleo, basadas en el modelo básico de búsqueda de empleo. A continuación se da una breve explicación del aporte de cada una de ellas.

**5.1 Edad:** A partir de la edad de un individuo es posible realizar una aproximación tanto del tiempo de su experiencia laboral, como alcanzar una representación sobre el nivel de su desempeño; también está relacionada con las posibles razones del rechazo o acogimiento por parte del mercado laboral, ya que las empresas incluyen criterios de selección que, en algunos casos, están asociados a la edad del aspirante: atraen o excluyen por edad avanzada, escasa experiencia o, incluso, edades puntuales. Esta variable se incluye en el modelo de forma lineal y cuadrática. Esta última por cuanto en edades avanzadas hay una inflexión en la duración de la búsqueda de empleo.

**5.2 Género:** La mayoría de los trabajos realizados incluyen esta variable, por cuanto existen diferencias, en términos laborales, según el género. Por esta razón, el presente trabajo realiza una estimación, con el propósito de identificar posibles diferencias en los impactos de los cuantiles de la distribución de la duración.

**5.3 Estado Civil:** El grado de tolerancia de un individuo hacia su condición de desempleo puede estar relacionado con su estado civil: así, por ejemplo, si el individuo cuenta con el apoyo económico de su pareja, puede hallarse en mejores condiciones para sobrellevar su situación de desempleo y, en consecuencia, tolerar un mayor tiempo en dicha condición; caso contrario ocurre cuando, por no disponer de dicho apoyo, intenta reducir en lo posible el tiempo en condición de desempleo. La relación entre estado civil y duración del desempleo puede también observarse en casos en los que el individuo en desempleo tiene responsabilidades familiares que operan significativamente para reducir en lo posible la duración del desempleo; en contraste, un individuo sin responsabilidades económicas familiares podría tolerar una mayor duración en desempleo. Así, entre otros aspectos, resulta interesante analizar la posible relevancia que adquiere la variable estado civil en duraciones cortas de desempleo, en comparación con duraciones largas.

**5.4 Jefe de Hogar:** Esta variable amerita una consideración especial: claramente un jefe de hogar es la persona que tiene la mayor responsabilidad económica de una familia y su comportamiento ante periodos de desempleo es diferente al de una persona que no la tenga.

**5.5 Servicio Doméstico:** Dado que disponer de servicio doméstico puede estar asociado a la percepción de ingresos más altos y, en consecuencia, a un mayor nivel de confort, esta condición puede incidir significativamente en las medidas que el individuo tome para reducir en lo posible, o no, la duración del desempleo.

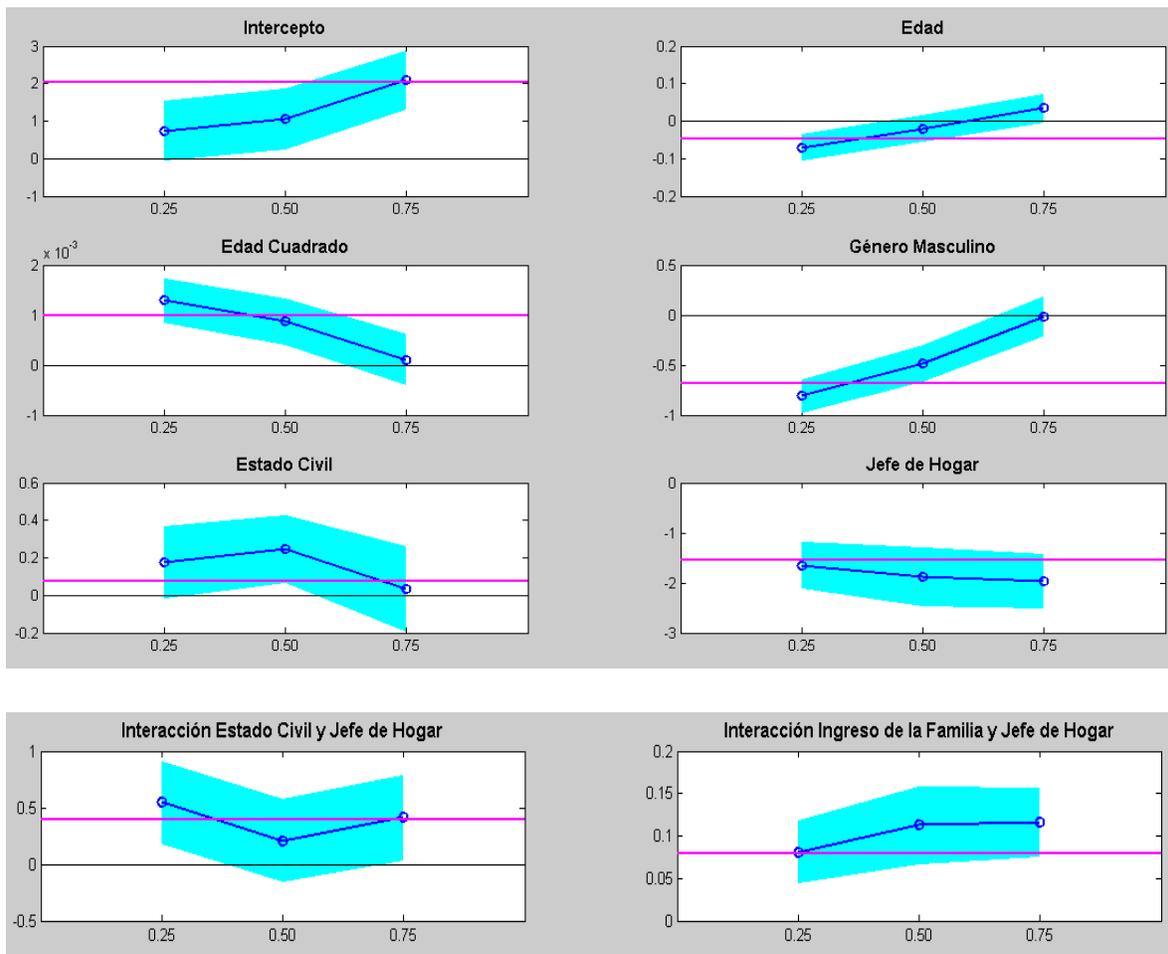
**5.6 Años de Educación:** Es una variable clásica y fundamental en el estudio teórico y empírico de la duración del desempleo. El tipo de relación existente entre la duración del desempleo y los años de educación, permite determinar la presencia de incentivos o desincentivos a la capacitación; así, por ejemplo, si la duración del desempleo tiende a disminuir cuanto mayores son los años de formación, se genera un incentivo a la capacitación; pero opera como un desincentivo, si la duración del desempleo está asociada a un mayor tiempo de formación del desempleado. El conocimiento de esta relación permite un mejor diseño de políticas públicas orientadas a la reducción de los periodos de desempleo, aquellas interesadas en aumentar los niveles de preparación académica de la población. Otro punto de interés, a observar en la presente investigación, es la existencia de una posible inflexión de los años de educación para duraciones largas.

**5.7 Ingreso Familiar Real Neto:** Esta variable se define como el soporte económico que un individuo tiene para poder sobrellevar el tiempo de desempleo y continuar con la búsqueda de trabajo. Su importancia radica en que la duración del desempleo puede ser sensible (tender a reducirse o a prolongarse) de acuerdo a las condiciones de vida en que se encuentre el desempleado, condiciones asociadas al ingreso.

## 6. RESULTADOS <sup>6</sup>

La estimación de los parámetros se presenta en el gráfico 1, en donde la estructura de los ejes es: el eje de las *abscisas*, o eje *x*, muestra los cuartiles y el eje de las *ordenadas*, o eje *y*, la estimación de los parámetros. En el gráfico se presenta la estimación de la media<sup>7</sup>, los cuartiles<sup>8</sup> y los intervalos de confianza del 95%, para la muestra completa.

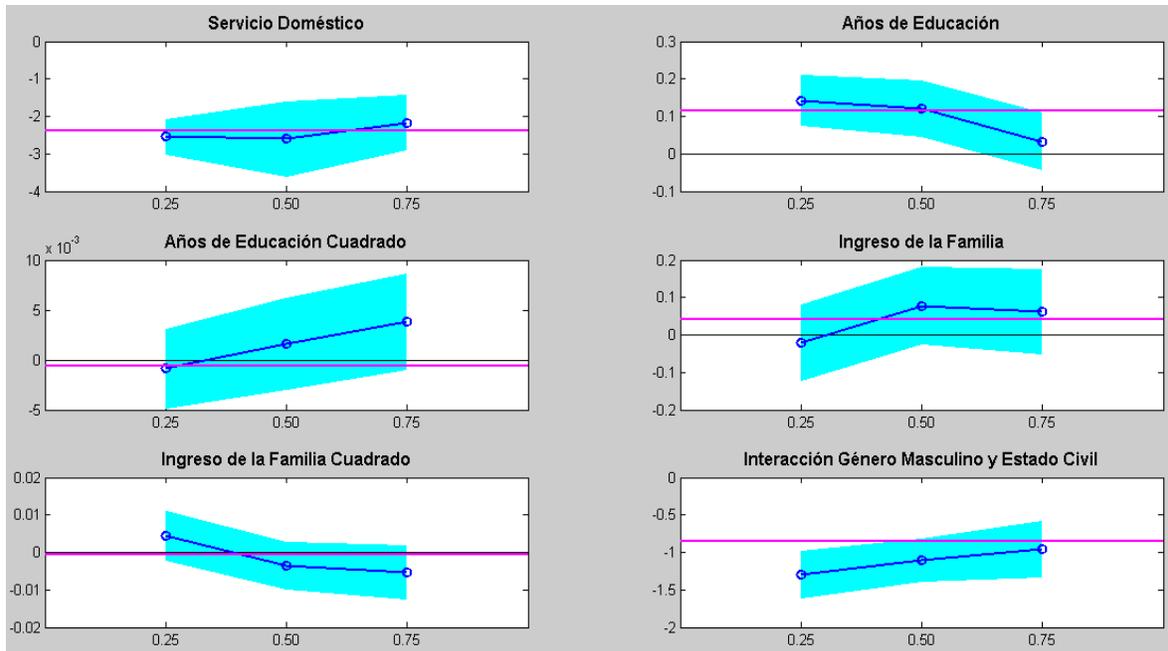
**Gráfico 1: Muestra completa**



<sup>6</sup> En el anexo 2 se encuentran las tablas con las estimaciones de los parámetros, en la media y los cuartiles, y de los intervalos de confianza del 95%.

<sup>7</sup> La estimación de la media se representa en los gráficos con una línea sólida de color magenta.

<sup>8</sup> Las estimaciones de los cuartiles se representa en los gráficos con una línea sólida de color azul y la región coloreada corresponden a los intervalos de confianza.



## 7. ANÁLISIS DE LOS RESULTADOS<sup>9</sup>

Entendiendo que el objetivo es la comparación entre la estimación clásica y la cuantílica, se presenta el análisis para la muestra total.

### 7.1 COMPARATIVO PARA LA MUESTRA TOTAL<sup>10</sup>

#### 7.1.1 EDAD

En el primer cuartil de la distribución de la duración del desempleo esta variable es significativa, tanto en el componente lineal como en el cuadrático; con el impacto esperado: negativo en el componente lineal y positivo en el cuadrático, es decir, la duración esperada es menor por cada año de edad, pero sólo hasta cierto punto, donde existe una inflexión que muestra que a mayor edad la duración esperada tiende a aumentar. En los

<sup>9</sup> La comparación de las estimaciones de los parámetros obtenidas para los cuantiles y la media, se realiza teniendo en cuenta la intersección de los intervalos de confianza correspondientes.

<sup>10</sup> A continuación se realiza la interpretación de los resultados para la media y los cuantiles, siguiendo la referencia bibliográfica [1]. Es importante resaltar que una de las propiedades del modelo AFT es la comparación de las funciones de supervivencia de dos individuos con diferentes características.

restantes cuantiles estudiados esta variable no es significativa de forma conjunta en sus dos componentes.

El impacto en la media es 1.71%; es decir, la duración esperada del desempleo para un individuo con un año más de edad, aumenta en dicho porcentaje. En el primer cuartil de la distribución de la duración del desempleo el impacto es 1.37%; es decir, la duración esperada del desempleo para un individuo con un año más de edad es mayor en dicho porcentaje. Al comparar el impacto en la media con el impacto en el primer cuartil se evidencia una sobrestimación.

### **7.1.2. GÉNERO MASCULINO**

En el primer y segundo cuartil de la distribución de la duración del desempleo, esta variable es significativa con impacto negativo. En el tercer cuartil de la distribución de la duración del desempleo, la variable no es significativa.

El impacto en la media es -64.55%; es decir, la duración esperada del desempleo para un hombre<sup>11</sup> es menor en 64.55% comparado con una mujer. En el primer y segundo cuartil de la distribución de la duración del desempleo los impactos son -74.29% y -61.37% respectivamente, es decir, la duración esperada del desempleo para un hombre es menor en dichos porcentajes comparado con una mujer.

Al comparar el impacto en la media con los impactos cuantílicos para un hombre en contraste con una mujer, se evidencia una sobrestimación en el primer cuartil y una subestimación en el tercer cuartil. Adicionalmente, en el segundo cuartil el impacto es aproximadamente igual al de la media.

### **7.1.2 ESTADO CIVIL**

En el segundo cuartil de la distribución de la duración del desempleo, esta variable es significativa con impacto positivo, indicando así que los

---

<sup>11</sup> El impacto se cuantifica usando el estado civil.

individuos casados en comparación con los no casados tienen un aumento en la duración esperada del desempleo. En el primer y tercer cuartil de la distribución de la duración del desempleo, esta variable no es significativa, indicando así que el estado civil no es un determinante relevante en la duración del desempleo cuando el individuo se encuentra en puntos extremos de la distribución.

En el segundo cuartil de la distribución de la duración del desempleo el impacto es -23.21%; es decir, la duración esperada del desempleo para un individuo casado es menor en 23.21% comparado con uno que no lo es. Adicionalmente, en el segundo cuartil el impacto es aproximadamente igual al de la media.

### **7.1.3 JEFE DE HOGAR**

En todos los cuartiles de la distribución de la duración del desempleo, esta variable es significativa, indicando así que los individuos que son jefes de hogar en comparación con los que no lo son, tienen una reducción en la duración esperada del desempleo.

El impacto en la media es -29.98%; es decir, la duración esperada del desempleo para un individuo jefe de hogar<sup>12</sup> es menor en 29.98% comparado con uno que no lo es. En los tres cuartiles de la distribución de la duración del desempleo el impacto es -32.21%, -30.66% y -27% respectivamente es decir, la duración esperada del desempleo para un individuo jefe de hogar es menor en tales porcentajes comparado con uno que no lo es.

En resumen, para un individuo jefe de hogar se observa una reducción de la duración esperada del desempleo en todos los cuartiles de la duración del desempleo.

---

<sup>12</sup> El impacto se cuantifica usando el estado civil y el ingreso de la familia promedio.

Al comparar el impacto en la media con los impactos cuantílicos para un individuo jefe de hogar versus uno que no lo es, se evidencia un impacto similar al estudiado usando cuantiles.

#### **7.1.5 SERVICIO DOMÉSTICO**

En todos los cuartiles de la distribución de la duración del desempleo, esta variable es significativa, indicando así que los individuos que tienen servicio doméstico en comparación con los que no lo tienen, presentan una reducción en la duración del desempleo. Adicionalmente, se evidencia un comportamiento similar a través de los cuartiles de la distribución de la duración del desempleo.

El impacto en la media es -90.8%; es decir, la duración esperada del desempleo para un individuo que tiene servicio doméstico es menor en 90.8% comparado con uno que no lo tiene. En los tres cuartiles de la distribución de la duración del desempleo el impacto es -92.07%, -92.61% y -88.51%, es decir, la duración esperada del desempleo para un individuo que tiene servicio doméstico es menor en dichos porcentajes para cada cuartil comparado con uno que no lo tiene. En resumen, para un individuo con servicio doméstico se observa una reducción de la duración esperada del desempleo en todos los cuartiles de la duración del desempleo comparada con aquellos que no poseen dicho servicio.

Al comparar el impacto en la media con los impactos cuantílicos para un individuo con servicio doméstico versus uno que no lo tiene, se evidencia un impacto similar.

#### **7.1.6 AÑOS DE EDUCACIÓN**

En el primer y segundo cuartil de la distribución de la duración del desempleo, esta variable es significativa en el componente lineal, pero no en el componente cuadrático. En el tercer cuartil de la distribución de la duración del desempleo, la variable no es significativa indicando así que

para los individuos con duraciones muy altas los años de educación no resultan relevantes en la duración del desempleo.

El impacto en la media es 10.95%; es decir, la duración esperada del desempleo para un individuo con un año más de educación<sup>13</sup> es mayor en 10.95%. En el primer y segundo cuartil de la distribución de la duración del desempleo el impacto es 13.31% y 16.5%; es decir, la duración del desempleo para un individuo con un año más de educación es mayor en dichos porcentajes. En resumen, para un individuo con un año más de educación la duración del desempleo aumenta a través de los dos primeros cuartiles de ésta.

Al comparar el impacto en la media con los impactos cuantílicos, se evidencia una subestimación para el segundo cuartil y un comportamiento similar en el primer y tercer cuartil de la duración del desempleo.

#### **7.1.7 INGRESO FAMILIAR**

Esta variable no es significativa en ninguno de los cuartiles. Sin embargo, el impacto en la media es 4.81%; es decir, la duración esperada del desempleo para un individuo con un incremento en el ingreso familiar<sup>14</sup> es mayor en 4.81%.

### **8. CONCLUSIONES**

Este trabajo se basa en la aplicación de una técnica econométrica que incorpora la regresión cuantílica en un modelo de duración, para identificar los impactos diferenciados de los fundamentales en los diferentes cuantiles de la duración del desempleo. Adicionalmente se hace una estimación del impacto en la media y se compara con los impactos a través de los cuantiles para establecer diferencias o similitudes estadísticas entre los resultados obtenidos en los dos

---

<sup>13</sup> El impacto se cuantifica usando los años de educación promedio.

<sup>14</sup> El impacto se cuantifica usando el ingreso familiar promedio y jefe de hogar.

enfoques, permitiendo un espectro más amplio en la medición y caracterización de la duración del desempleo en Colombia.

Desde la perspectiva de las políticas económicas tendientes a mejorar la calidad de vida de los individuos, por medio de la reducción de la duración del desempleo, es importante la visión cuantílica, debido a que se puede evidenciar el comportamiento particular de ciertas variables en los diferentes cuantiles de la duración del desempleo, y así poder establecer las principales características de éstos individuos.

Al estudiar la forma de la distribución de la duración del desempleo, podemos observar duraciones extremas; esto es, duraciones muy por debajo o por encima del promedio; hecho que puede influir la estimación del impacto promedio de las variables en la duración del desempleo y puede generar imprecisiones e inequidad en el uso de los resultados, como por ejemplo en la adjudicación de los seguros de desempleo o posibles ayudas al desempleado. Lo anteriormente expuesto se puede corregir, o mitigar, al implementar la metodología econométrica que involucra, en la estimación de los impactos, la regresión cuantílica.

Este estudio muestra hallazgos de diferencias y similitudes entre las dos formas de estimación. En cuanto a las diferencias de los enfoques, un ejemplo claro lo constituye el caso de los individuos *jefe de hogar*, quienes, de acuerdo al estudio de regresión cuantílica, y para todos los cuantiles considerados, presentan una menor duración de tiempo en el desempleo que aquellos que no son jefes de hogar. Al comparar las magnitudes absolutas de los coeficientes de la regresión cuantílica se puede ver una superioridad frente a la estimación clásica en los dos primeros cuantiles, con lo cual se infiere que la estimación clásica calcula una duración esperada mayor que aquella que pronostica la regresión cuantílica. En el tercer cuartil, en cambio, la estimación clásica pronostica una salida más rápida del desempleo ante duraciones altas; es decir, manteniendo lo demás constante,

un individuo que lleva un tiempo largo en desempleo tiende a salir más rápido de este estado comparado con lo que estima la regresión por cuantiles.

Lo anteriormente expuesto nos lleva a reflexionar sobre la relación existente entre la reducción del tiempo de búsqueda de empleo y la calidad del empleo conseguido. No necesariamente, en la mayoría de los casos se encuentra un trabajo mejor que el anterior, pero la urgencia o necesidad sí presiona para tomarlo. En situaciones como estas es importante que el Estado disponga de alertas, de modo que, una vez el individuo llegue a una situación de desempleo, pueda ofrecerle alguna ayuda que le permita buscar con tranquilidad y objetividad un empleo que se ajuste más a sus expectativas y necesidades.

Por otro lado, es importante notar que una política pública que tome como base una estimación clásica, afectaría a aquellos individuos ubicados en el tercer cuartil, en la medida en que propende a concebir que estos salen más rápidamente del desempleo y, en consecuencia, la ayuda generada por el Estado no se orientaría a satisfacer de manera adecuada las necesidades de este grupo de individuos. Aunque estos coeficientes no están demasiado alejados muestran la diferencia potencial que los dos enfoques tienen y que redundaría en dejar de tener en cuenta algunas poblaciones para las cuales el promedio no es una buena aproximación. Este es justamente el problema del que se ocupa la regresión cuantílica.

Así como se ha puesto de manifiesto que los individuos que son jefes de hogar denotan claras diferencias con aquellos que no los son, en términos de duraciones a través de los diferentes cuantiles de la duración de desempleo, así también se hace evidente una clara diferencia en el caso del género. Sin embargo, en este segundo caso, la dificultad es aún mayor por parte de las mujeres en la consecución del empleo, en comparación con los hombres, lo cual evidencia la pervivencia, en los empleadores, de un sesgo hacia la contratación de hombres más que mujeres. Así también, pone de presente la falta de apoyo de las

instituciones del Estado para lograr corregir estos hechos poco favorables para la sociedad femenina. Esta relación desfavorable se mantiene, con ciertos cambios de magnitud, a través de todos los cuantiles de la distribución. En el caso del tercer cuartil se denota una diferencia importante en comparación con la estimación en la media.

También se puede observar que en los años de educación, las dos estimaciones resultan ser muy cercanas, de modo que la estimación por cuantiles resulta ser un elemento de medición equivalente a la medición del efecto en la media.

Uno de los aspectos más importantes que se ha pretendido resaltar a lo largo de este trabajo, y cuyos beneficios para la investigación económica y econométrica puede explotarse aún más, lo constituye el énfasis en la medición del impacto sobre los cuantiles de la duración del desempleo. Se trata de una aproximación que captura con más precisión las especificidades del objeto de estudio y provee una base más sólida para la toma de decisiones de los diferentes agentes económicos. En particular, las investigaciones con un enfoque en el impacto de las variables sobre cuantiles de duración del desempleo, tienen una especial utilidad para los agentes gubernamentales, por cuanto les provee de información indispensable para el adecuado diseño e implementación de políticas públicas de empleo, en la medida en que pueden realmente estar focalizadas en los sectores que la política busca beneficiar (eficacia), al tiempo que permite realizar una estimación y uso adecuado de los recursos públicos y el capital humano (eficiencia).

## REFERENCIAS BIBLIOGRÁFICAS

- [1] **Allison, P. (2012)** “Survival Ananlysis Using SAS, A Practical Guide” *SAS Publishing, 2<sup>th</sup> Edition*
- [2] **Cameron, A. and P. Trivedi (2005)** “Microeconometrics: Methods and Applications” *Cambridge University Press*
- [3] **Castellar, C y Uribe, J. I. (2003)** “Determinantes de la duración del desempleo en el Área Metropolitana de Cali, 1988-2000” *Archivo de Economía, No. 218 Departamento Nacional de Planeación*
- [4] **Cleves, M.A. (2008)** “An Introduction to Survival Analysis Using Stata” *A Stata Press Publication, 2<sup>th</sup> Edition*
- [5] **Ehrenberg, R. y Oaxaca, R. (1976)** “Unemployment Insurance, Duration of Unemployment, and Subsequent Wage Gain” *AER*
- [6] **Fallon, P. & Donald V. (1988)** “The economics of Labour Markets” *PhilipAllan Publishers Limited, Oxford U.K.*
- [7] **Fitzenberger, B. and Wilke, Ralf A. (2005)** “Using Quantile Regression for Duration Analysis” *ZEW Discussion Paper, 05-58, Mannheim*
- [8] **Heckman, J. and B. Singer (1984)**, “A method for minimizing the impact of distributional assumptions in econometric models for duration data” *Econometrica, 52*
- [9] **Heij, C., de Boer, P., Franses, P. H., t. Kloek and H. van Dijk (2004)** “Econometric Methods with Applications in Business and Economics” *Oxford university Press*
- [10] **Hosmer, D., Lemeshow, S. and May, S. (2008)** “Applied Survival Analysis” *Wiley Series in Probability and Statistics, 2<sup>th</sup> Edition*

- [11] **Koenker, R. (2001)** “Quantile Regression” *Encyclopedia of Environmetrics*
- [12] **Koenker, R. and Biliyas, Y. (2001)** “Quantile regression for duration data: A reappraisal of the Pennsylvania Reemployment Bonus Experiments” *Empirical Economics*, Vol.26, pp. 199-220
- [13] **Lancaster, T. (1979)** “Econometric Methods for the Duration of Unemployment” *Econometrica*, Vol.47, No.4, pp. 939-956
- [14] **Martínez, H. (2003)** “Cuánto duran los colombianos en el desempleo y en el empleo: un análisis de supervivencia”, *Archivos de Economía* 236
- [15] **Meyer, B. (1990)** “Unemployment Insurance and Unemployment Spells” *Econometrica*, Vol.58, No.4, pp. 757-782
- [16] **Núñez J. y R. Bernal (1998)** “El desempleo en Colombia: tasa natural, desempleo cíclico y estructural y la duración del desempleo (1976-1998)” *Archivos de Macroeconomía, Documento No. 97, Departamento Nacional de Planeación*
- [17] **Peng, L. and Huang, Y. (2008)** “Survival Analysis with Quantile Regression Models” *Journal of the American Statistical Association*, 103, 637–649
- [18] **Pesaran, M. H. and P. Schmidt (1999)** “Handbook of Applied Econometrics, Volume II: Microeconometrics” *Blackwell Handbooks in Economics*
- [19] **Pollmann-Schult, M. and Büchel, F. (2005)** “Unemployment Benefits, Unemployment Duration and Subsequent Job Quality: Evidence from West Germany” *Acta Sociologica* Vol. 48 No. 1, pp. 21-39
- [20] **SAS Institute Inc. (2012)** “SAS/STAT 12.1 User’s Guide – The Quantile Procedure” *Cary, NC: SAS Institute Inc.*

**[21] Stengos, T. y Dianqin W. (2007)** “An algorithm for censored quantile regressions” *Economics Bulletin*, Vol. 3, No. 1 pp. 1-9

**[22] Tenjo, J., Misas, M., Contreras, A. y Gaviria, A. (2012)** “Modelos de duración del desempleo en Colombia” *Vniversitas Económica*, Pontificia Universidad Javeriana, Vol. 12

**[23] Tenjo, J. y Ribero, R. (1998)** “Participación, desempleo y mercados laborales en Colombia” *Archivos de Macroeconomía*, Documento No. 81 Departamento Nacional de Planeación

**[24] van den Berg, G., Lindeboom, M and Dolton, P. (2006)** “Survey Non-Response and the Duration of Unemployment” *Royal Statistical Society*

**[25] Viáfara C. y Uribe, I. (2008)** "Duración del Desempleo y Canales de Búsqueda de Empleo en Colombia 2006" *Revista de Economía Institucional*. Vol 11, # 21, 2009. Universidad Externado de Colombia.

**[26] Wichert, L. and Wilke, R. (2008)** “Simple non-parametric estimators for unemployment duration analysis” *Royal Statistical Society* Vol. 57, Part1, pp117-126

## Anexo 1

A continuación se presenta un caso particular para ilustrar los inconvenientes de la función objetivo. Sea  $\varphi$  la función objetivo, y algunos valores para estudiar su comportamiento de manera analítica y gráfica. Siguiendo la referencia bibliográfica [21], Sean:

$$\theta = 0.5, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} cy_1 \\ cy_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad (20)$$

En este punto es de aclarar que el valor  $y_2=3$ , se toma sin pérdida de generalidad ya que lo importante es que sea una cantidad superior al punto de censura  $cy_2=2$ .

Dado que en (16) encontramos las cantidades  $d_i := y_i - \max(x'_i\beta, yc_i)$ , para nuestro caso particular escribiremos las cantidades correspondientes a  $d_i$  para  $i=1,2$  son:

$$d_1 := y_1 - \max(x'_1\beta, yc_1), \text{ y reemplazando } d_1 := 1 - \max(3\beta, 1)$$

$$d_2 := y_2 - \max(x'_2\beta, yc_2), \text{ y reemplazando } d_2 := 3 - \max(4\beta, 2)$$

Ahora, estudiando los casos obtenemos los siguientes valores para las cantidades  $d_i$  son:

$$d_1 = \begin{cases} 0 & \text{si } \beta \leq \frac{1}{3} \\ 1 - 3\beta & \text{si } \beta > \frac{1}{3} \end{cases} \text{ y } d_2 = \begin{cases} 1 & \text{si } \beta \leq \frac{1}{2} \\ 1 - 4\beta & \text{si } \beta > \frac{1}{2} \end{cases}; \text{ en un caso, la función indicadora}$$

se definen como:

$$I(d_1 > 0) = \begin{cases} 1 & \text{si } \beta < \frac{1}{3} \\ 0 & \text{si } \beta > \frac{1}{3} \end{cases},$$

Y análogamente, tenemos los restantes casos dados por:

$$I(d_1 < 0) = \begin{cases} 1 & \text{si } \beta > \frac{1}{3}, \\ 0 & \text{si } \beta < \frac{1}{3}, \end{cases}$$

$$I(d_2 > 0) = \begin{cases} 1 & \text{si } \beta < \frac{1}{2}, o, \beta > \frac{3}{4} \\ 0 & \text{si } \frac{1}{2} < \beta < \frac{3}{4} \end{cases}$$

$$I(d_2 < 0) = \begin{cases} 1 & \text{si } \frac{1}{2} < \beta < \frac{3}{4} \\ 0 & \text{si } \beta < \frac{1}{2}, o, \beta > \frac{3}{4} \end{cases}$$

Reemplazando en la ecuación (16), tenemos:

$$\varphi(\beta) = \sum_{i=1}^2 (0.5 * I(d_i > 0) + (1 - 0.5) * I(d_i < 0)) |d_i|$$

Realizando las operaciones y factorizando, se tiene la siguiente expresión:

$$\varphi(\beta) = 0.5[I(d_1 > 0) + I(d_1 < 0)]|d_1| + 0.5[I(d_2 > 0) + I(d_2 < 0)]|d_2| \quad (21)$$

Como se conoce la naturaleza de las cuatro funciones indicadoras y sus puntos singulares  $\beta = \frac{1}{3}, \beta = \frac{1}{2}$  y  $\beta = \frac{3}{4}$ . Se hace el análisis de los diferentes intervalos para lograr expresar (18) como una función a trozos sencilla y poder realizar su gráfica:

Si  $\beta < \frac{1}{3}$ , se tiene que  $\varphi(\beta) = 0.5[1 + 0]|0| + 0.5[1 + 0]|1| = 0.5$

Si  $\frac{1}{3} \leq \beta < \frac{1}{2}$ , se tiene  $\varphi(\beta) = 0.5[0 + 1]|1 - 3\beta| + 0.5[1 + 0]|1| = 0.5[(3\beta - 1) + 1] = 1.5\beta$ .

Si  $\frac{1}{2} \leq \beta < \frac{3}{4}$ , se tiene que  $\varphi(\beta) = 0.5[0 + 1]|1 - 3\beta| + 0.5[0 + 1]|3 - 4\beta| = 0.5[(3\beta - 1) + 3 - 4\beta] = -0.5\beta + 1$ .

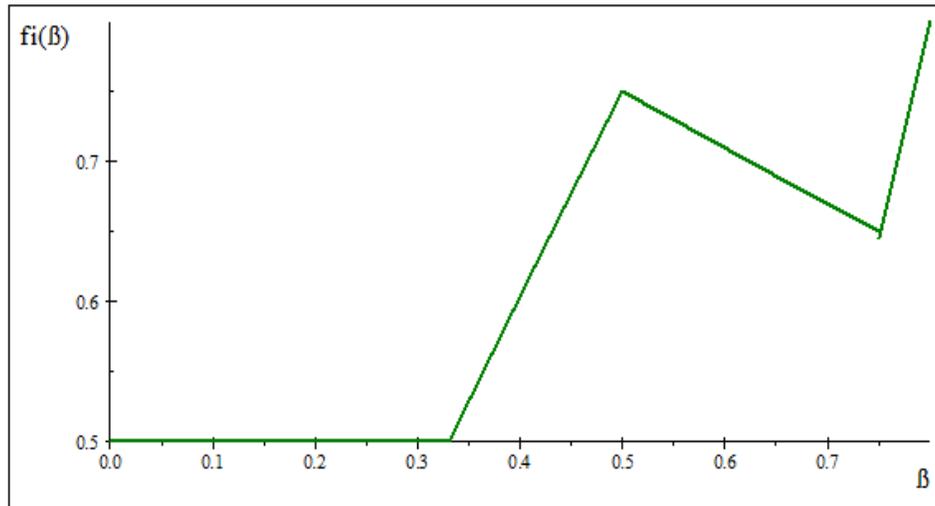
Si  $\beta \geq \frac{3}{4}$ , se tiene que  $\varphi(\beta) = 0.5[0 + 1]|1 - 3\beta| + 0.5[1 + 0]|3 - 4\beta|$   
 $= 0.5[(3\beta - 1) + (4\beta - 3)] = 3.5\beta - 2.$

Con base en los resultados obtenidos se puede expresar la función, de interés, como:

$$\varphi(\beta) = \begin{cases} 0.5 & \text{si } \beta < \frac{1}{3} \\ 1.5\beta & \text{si } \frac{1}{3} \leq \beta < \frac{1}{2} \\ -0.5\beta + 1 & \text{si } \frac{1}{2} \leq \beta < \frac{3}{4} \\ 3.5\beta - 2 & \text{si } \beta \geq \frac{3}{4} \end{cases} \quad (22)$$

Una vez estudiada la función obtenemos la siguiente representación gráfica donde se pueden apreciar las características mencionadas de linealidad a trozos, no suave en los puntos críticos y no convexa, que dificulta el proceso de minimización.

**Gráfico 2: Función  $\varphi(\beta)$**



## Anexo 2

### Población Total

Variables	Cuartil 1			
	Estimación	Lím. Inferior*	Lím Superior*	Media**
<b>Años de Educación</b>	0.14260	0.07540	0.20990	0.11590
<b>Años de Educación Cuadrado</b>	-0.00088	-0.00482	0.00306	-0.00060
<b>Edad</b>	-0.07000	-0.10300	-0.03700	-0.04790
<b>Edad Cuadrado</b>	0.00129	0.00086	0.00172	0.00100
<b>Estado Civil</b>	0.17750	-0.01100	0.36590	0.07760
<b>Género Masculino</b>	-0.81050	-0.97600	-0.64490	-0.67900
<b>Ingreso de la Familia</b>	-0.02030	-0.12030	0.07980	0.04250
<b>Ingreso de la Familia Cuadrado</b>	0.00448	-0.00197	0.01090	-0.00070
<b>Interacción Estado Civil y Jefe de Hogar</b>	0.54610	0.18750	0.90470	0.39610
<b>Interacción Género Masculino y Estado Civil</b>	-1.29930	-1.60930	-0.98920	-0.84950
<b>Interacción Ingreso de la Familia y Jefe de Hogar</b>	0.08100	0.04460	0.11740	0.07980
<b>Intercepto</b>	0.74250	-0.03720	1.52230	2.04820
<b>Jefe de Hogar</b>	-1.64600	-2.10860	-1.18340	-1.53520
<b>Servicio Doméstico</b>	-2.53460	-2.99480	-2.07440	-2.38630

\* Límites del intervalo de confianza del 95%.

\*\* Estimación del parámetro con el modelo clásico.

Variables	Cuartil 2			
	Estimación	Lím. Inferior*	Lím Superior*	Media**
<b>Años de Educación</b>	0.12010	0.04680	0.19340	0.11590
<b>Años de Educación Cuadrado</b>	0.00163	-0.00291	0.00616	-0.00060
<b>Edad</b>	-0.01990	-0.05420	0.01450	-0.04790
<b>Edad Cuadrado</b>	0.00087	0.00042	0.00132	0.00100
<b>Estado Civil</b>	0.24900	0.07190	0.42610	0.07760
<b>Género Masculino</b>	-0.48470	-0.66360	-0.30580	-0.67900
<b>Ingreso de la Familia</b>	0.07770	-0.02510	0.18040	0.04250
<b>Ingreso de la Familia Cuadrado</b>	-0.00358	-0.00974	0.00258	-0.00070
<b>Interacción Estado Civil y Jefe de Hogar</b>	0.20670	-0.15440	0.56790	0.39610
<b>Interacción Género Masculino y Estado Civil</b>	-1.10660	-1.38670	-0.82650	-0.84950
<b>Interacción Ingreso de la Familia y Jefe de Hogar</b>	0.11270	0.06730	0.15810	0.07980
<b>Intercepto</b>	1.05270	0.25000	1.85530	2.04820
<b>Jefe de Hogar</b>	-1.88230	-2.45700	-1.30760	-1.53520
<b>Servicio Doméstico</b>	-2.60550	-3.60540	-1.60560	-2.38630

\* Límites del intervalo de confianza del 95%.

\*\* Estimación del parámetro con el modelo clásico.

Variables	Cuartil 3			
	Estimación	Lím. Inferior*	Lím Superior*	Media**
<b>Años de Educación</b>	0.03310	-0.04350	0.10970	0.11590
<b>Años de Educación Cuadrado</b>	0.00383	-0.00092	0.00858	-0.00060
<b>Edad</b>	0.03420	-0.00214	0.07060	-0.04790
<b>Edad Cuadrado</b>	0.00011	-0.00038	0.00060	0.00100
<b>Estado Civil</b>	0.03330	-0.19070	0.25720	0.07760
<b>Género Masculino</b>	-0.01330	-0.20720	0.18050	-0.67900
<b>Ingreso de la Familia</b>	0.06130	-0.05120	0.17370	0.04250
<b>Ingreso de la Familia Cuadrado</b>	-0.00531	-0.01230	0.00171	-0.00070
<b>Interacción Estado Civil y Jefe de Hogar</b>	0.41200	0.03520	0.78880	0.39610
<b>Interacción Género Masculino y Estado Civil</b>	-0.95780	-1.33150	-0.58410	-0.84950
<b>Interacción Ingreso de la Familia y Jefe de Hogar</b>	0.11600	0.07560	0.15650	0.07980
<b>Intercepto</b>	2.08750	1.32110	2.85390	2.04820
<b>Jefe de Hogar</b>	-1.95930	-2.49600	-1.42250	-1.53520
<b>Servicio Doméstico</b>	-2.16330	-2.90140	-1.42530	-2.38630

\* Límites del intervalo de confianza del 95%.

\*\* Estimación del parámetro con el modelo clásico.