



Departamento de Economía

Índice de empleabilidad para Colombia

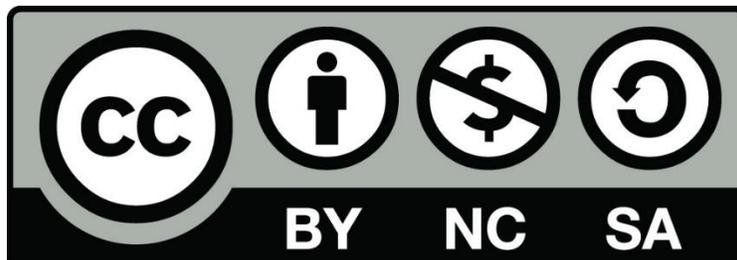
Tesis

Maestría en Econometría

Autor: Sergio Sánchez Delgado

Tutor: Martín González Rozada

Junio de 2017



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).



Reconocimiento: En cualquier explotación de la obra autorizada por la licencia se debe reconocer la autoría.



No Comercial: La explotación de la obra queda limitada a usos no comerciales.



Compartir igual: La explotación autorizada incluye la creación de obras derivadas siempre que mantengan esta misma licencia al ser divulgadas.

Resumen

Este trabajo tiene como objetivo construir una herramienta que facilite el diagnóstico de las personas que se encuentran buscando un trabajo en el mercado laboral colombiano, haciendo énfasis en las barreras que presentan. Adicionalmente, este instrumento permitirá identificar los programas y acciones específicos que buscador necesita para mitigar sus barreras y mejorar sus posibilidades de acceder a un puesto de trabajo, al igual que permitirá identificar barreras comunes a grupos específicos y diseñar nuevas políticas que sean aplicadas de manera transversal a estos grupos.

Para la construcción de esta herramienta, lo primero que se hace es identificar, a partir de un grupo de fuentes de información, la más idónea para establecer las diferentes características que inciden en la posibilidad de un buscador para conseguir un empleo. De esta identificación se establece que la Encuesta de Calidad de Vida 2016 es la fuente de información más completa y actualizada porque contiene información socioeconómica para una muestra representativa de la población colombiana. A partir de esta información, se identifican 23 preguntas que permiten construir las variables y se organizan las opciones de respuesta de cada una a partir de dos criterios: la incidencia de la tasa de desempleo y su relación y significancia en la probabilidad de estar ocupado.

Una vez establecidas las variables se debe reducir la dimensión del conjunto de información y para esto se estima un modelo de componentes principales, a partir del modelo se encuentra que de las 23 variables, 18 presentan las condiciones suficientes para ser utilizadas. Luego, se determina que hay 7 componentes que son significativos para la estimación del índice y se estima la proporción de varianza que aporta cada uno con respecto a la variabilidad total de la información y se normaliza para obtener los ponderadores de cada componente.

Una vez se obtienen el número de componentes y los ponderadores, se recurre a la matriz de cargas factoriales para determinar las variables que tienen mayor incidencia en cada componente. Se establece que toda variable que tenga una carga factorial dentro de un componente mayor o igual a 0,25 tiene una fuerte incidencia en la determinación de ese componente y por tanto hará parte de la estimación del mismo. Una vez identificadas las variables que inciden en cada componente se procede a calcular el valor del índice de empleabilidad.

Finalmente, se realiza el cálculo del índice para las más de 33 mil personas que hacen parte de la muestra de la encuesta y se comparan los resultados entre las personas que se encontraban ocupadas y las desempleadas, obteniendo que las personas ocupadas en promedio presentan un



Índice de empleabilidad para Colombia

índice superior en 10 puntos a los desocupados. Adicionalmente, a partir del cálculo del índice para toda la fuerza laboral, se divide la muestra en cuartiles y se obtienen cuatro grupos que permiten definir niveles de empleabilidad: bajo, medio bajo, medio alto y alto. Para cada uno de los grupos se calcula el índice promedio, se establecen los componentes en que presentan mayores brechas y se determinan algunas posibles acciones que permitirán mitigar las barreras y mejorar sus condiciones de empleabilidad.

Palabras clave: integer, blandit, pharetra, urna, id.



Contenido

Agradecimientos	¡Error! Marcador no definido.
Resumen	3
Abstract	¡Error! Marcador no definido.
Lista de tablas	6
Lista de gráficos	7
1. Introducción	9
2. Marco teórico	12
Análisis de componentes principales	16
Determinación de componentes	17
Selección de la cantidad de componentes	20
3. Metodología	22
4. Datos y Estimación	25
Conjunto de información	25
Resultados	61
Percentiles	65
5. Conclusiones y Recomendaciones	68
6. Bibliografía	73

Lista de tablas

Tabla 1. Variables construidas a partir de la ECV 2016.....	29
Tabla 2. Estadísticos descriptivos de las variables seleccionadas.	32
Tabla 3. Tasas de desempleo para cada categoría de las variables.	33
Tabla 4. Resultados modelos probit.....	36
Tabla 5. Variables seleccionadas con opciones ajustadas por criterios.....	37
Tabla 6. Matriz de correlaciones de las variables seleccionadas.....	54
Tabla 7. Resultados del test KMO.....	55
Tabla 8. Componentes estimados, autovalores asociados y proporción de varianza explicada ...	57
Tabla 9. Matriz de cargas factoriales.....	58
Tabla 10 composición de los componentes seleccionados.....	59
Tabla 11. Porcentaje de varianza explicada y varianza explicada normalizada.....	60
Tabla 12. Estadísticos descriptivos del índice de empleabilidad para el total de la muestra.....	61
Tabla 13. Estadísticos descriptivos del índice de empleabilidad para el total de personas ocupadas.....	62
Tabla 14. Estadísticos descriptivos del índice de empleabilidad para el total de personas desempleadas	63
Tabla 15. Test de medias para el índice entre el grupo de ocupados y desempleados.....	64

Lista de gráficos

Gráfico 1. Ubicación de la fuerza laboral en Colombia.....	39
Gráfico 2. Distribución de la fuerza laboral según tipo de departamento.	40
Gráfico 3. Distribución de la fuerza laboral según sexo.....	41
Gráfico 4. Distribución de la fuerza laboral según rangos de edad.	41
Gráfico 5. Distribución de la fuerza laboral según si tienen pareja o no	42
Gráfico 6. Distribución de la fuerza laboral según jefe de hogar	43
Gráfico 7. Distribución de la fuerza laboral según estrato socioeconómico.....	44
Gráfico 8. Distribución de la fuerza laboral según número de personas en el hogar	44
Gráfico 9. Distribución de la fuerza laboral según ingresos per cápita por rangos	45
Gráfico 10. Distribución de la fuerza laboral según pobreza.....	45
Gráfico 11. Distribución de la fuerza laboral según subsidio.....	46
Gráfico 12. Distribución de la fuerza laboral según nivel de satisfacción en el trabajo.....	47
Gráfico 13. Distribución de la fuerza laboral según ocupaciones	47
Gráfico 14. Distribución de la fuerza laboral según tiempo de búsqueda	49
Gráfico 15. Distribución de la fuerza laboral según trabajo por primera vez.....	49
Gráfico 16. Distribución de la fuerza laboral según frecuencia de uso del computador	50
Gráfico 17. Distribución de la fuerza laboral según frecuencia de uso del internet.	51
Gráfico 18. Distribución de la fuerza laboral según el nivel educativo.....	51
Gráfico 19. Gráfico de sedimentación con los autovalores asociados a los componentes estimados.....	56
Gráfico 20. Gráfico de distribución del índice de empleabilidad para el total de la muestra.	61
Gráfico 21. Gráfico de distribución del índice de empleabilidad para el total de personas ocupadas.....	62



Índice de empleabilidad para Colombia

Gráfico 22. Gráfico de distribución del índice de empleabilidad para el total de personas desempleadas	63
Gráfico 23. Puntuación promedio para el percentil 25 de la distribución del índice.....	65
Gráfico 24. Puntuación promedio para el percentil 50 de la distribución del índice.....	65
Gráfico 25. Puntuación promedio para el percentil 75 de la distribución del índice.....	66
Gráfico 26. Puntuación promedio para el percentil 100 de la distribución del índice.....	67



1. Introducción

Uno de los principales problemas que enfrentan las economías modernas a nivel mundial consiste en reducir el impacto económico y social que producen el desempleo, subempleo y la informalidad en la sociedad. En efecto, estos fenómenos se han convertido en grandes problemáticas, tanto económicas como sociales, dado que generan subutilización del capital humano, se desaprovecha el talento y las habilidades de la población, baja los niveles de empleabilidad, destruye el capital social, deteriora los niveles de ingreso, reduce la demanda interna y disminuye el potencial de crecimiento económico de un país.

Durante los últimos años diferentes países e instituciones multilaterales han promovido la creación e implementación de diferentes políticas activas y pasivas de empleo, con el propósito de contribuir a reducir esta problemática, facilitar el acceso a una vacante por parte de los buscadores de empleo y conseguir el recurso humano idóneo a los empresarios.

Sin embargo, la economía colombiana se caracteriza por la persistencia de distorsiones que generan dificultades en el proceso de emparejamiento entre oferta y demanda laboral. Entre las principales distorsiones se destacan la heterogeneidad del recurso humano, las barreras que presentan los buscadores de empleo y la poca información que limita realizar diagnósticos acordes a las realidades de los territorios y a su vez dificulta la implementación de acciones a la medida de sus necesidades.

Para dar solución a este problema, en diferentes países se han diseñado e implementado herramientas que facilitan el diagnóstico de las barreras que inciden en la población y a su vez permitan priorizar a la población más vulnerable en el mercado laboral de acuerdo a sus necesidades y realidades territoriales, de tal forma que se puedan implementar políticas que mitiguen esas barreras y les facilite el acceso a un empleo en condiciones de formalidad.

Dentro del conjunto de herramientas desarrolladas, diferentes autores han propuesto la creación de índices sintéticos como respuesta a la necesidad de generar este tipo de información y conocimiento. Este tipo de índices facilitan los diagnósticos de los buscadores de empleo y ayudan a diseñar estrategias más ajustadas a sus realidades, así mismo permiten identificar necesidades

comunes y focalizar acciones que permitan mitigar las barreras y cerrar brechas en el mercado laboral.

Estas herramientas se caracterizan por su diversidad de enfoques, tanto conceptual como en la operación. En la literatura se encuentran desde herramientas para diferentes sectores económicos (industria, comercio, agricultura, etc), poblaciones vulnerables específicas (como jóvenes, mujeres, informales, etc) y hasta ejercicios más aplicados a las necesidades de la operación, como los implementados en algunas oficinas de empleo de diferentes países. Sin embargo, a pesar de las diferencias existentes entre los diferentes ejercicios, en la mayoría se encuentran puntos comunes como la necesidad de definir conceptualmente que es una ventaja o limitación relativa que enfrenta el buscador de empleo, lo cual determina el alcance de la herramienta, hasta las diferentes metodologías econométricas que siempre concluyen en tomar un gran conjunto de información y reducirlo, de tal forma que se obtenga un único número que determine la posibilidad que tienen las personas de acceder a un empleo y las dimensiones críticas en que se deben implementar acciones para mitigar las barreras que enfrentan.

Por tanto, en el presente documento se plantea en primer lugar una definición del concepto de empleabilidad que permitirá determinar cuáles son los factores que inciden en la posibilidad de encontrar un empleo y posteriormente facilitará el diseño de un índice de empleabilidad para Colombia.

Este índice tendrá como objetivos establecer la ventaja o desventaja relativa de una persona en el mercado de trabajo. Así mismo, determinará cuales son las barreras que inciden en la posibilidad de engancharse laboralmente a una persona y de esta forma incidir a través de programas a la medida para contribuir en la mitigación de las mismas. Finalmente, el índice permitirá priorizar grupos vulnerables que puedan ser objeto de políticas que mitiguen sus barreras y faciliten su acceso a un empleo y facilitará la posterior la definición de políticas y acciones de mayor impacto en la mitigación de barreras y el cierre de brechas a poblaciones que presentan características similares.

En la construcción de este indicador lo primero que se realiza es una identificación de las fuentes de información, este proceso permitirá obtener los datos sobre las características del buscador, posteriormente, a partir de todo el conjunto de información, se establecerán las variables que inciden tanto en el desempleo, como en la posibilidad de engancharse. A continuación, se utiliza

la metodología de componentes principales con un doble objetivo: primero, reducir la dimensión de la información y segundo obtener los pesos relativos de cada uno de los componentes dentro de la posibilidad de conseguir un empleo. Finalmente, se realiza el cálculo del índice de empleabilidad para toda la fuerza laboral, se establecerán diferentes grupos de interés y se realizará un breve diagnóstico de las barreras que son comunes entre ellos.

El trabajo está dividido en cinco secciones. La primera es esta introducción, en la segunda se presenta el marco teórico que soporta todo el desarrollo conceptual y econométrico del índice. En la tercera sección se realiza todo el desarrollo metodológico de la herramienta. En la cuarta parte se presenta la identificación de las fuentes de información, el análisis del conjunto de información establecido, el procedimiento para la estimación de los componentes y los respectivos ponderadores que permitirán la construcción del indicador, los resultados para toda la muestra de la fuerza laboral y la identificación y diagnóstico de barreras comunes para los diferentes grupos priorizados a partir del indicador. Finalmente, en la última sección se presentan las conclusiones y recomendaciones derivadas de este estudio.



2. Marco teórico

El mercado laboral colombiano se caracteriza por diversas distorsiones derivadas de fallas de mercado, regulaciones, rigideces e imaginarios culturales que dificultan el encuentro entre la oferta y la demanda de trabajo. Esto genera como consecuencia una alta tasa de desempleo relativa a otras economías de la región, altas tasas de subempleo, informalidad y bajos salarios que impactan en mayor medida a diferentes grupos de personas que presentan altos niveles de vulnerabilidad.

Adicionalmente a estas distorsiones típicas de la economía colombiana, se suma el problema de la heterogeneidad en los trabajadores, lo cual dificulta el acceso a puestos de trabajo por parte de los buscadores, incrementa los tiempos de desempleo y aumenta los costos de búsqueda por parte de los empleadores. Esta heterogeneidad surge a partir de las diferencias en las características que presenta cada buscador y que pueden tanto potenciar como limitar su acceso a un puesto de trabajo. Poder identificar esas características que facilitan o limitan el acceso a un trabajo formal permitirá a los hacedores de política pública, a los empresarios y a los propios buscadores de empleo mitigar las barreras que inciden en mayor medida en la vinculación a un puesto de trabajo.

Sin embargo, al día de hoy no se cuenta en el país con instrumentos o metodologías que faciliten realizar este tipo de mediciones para incidir en las decisiones los diferentes actores y por tanto surge como respuesta la construcción de un índice que permita medir cual es la ventaja o limitación relativa que tiene una persona que se encuentra en el proceso de búsqueda de empleo.

Para poder construir una herramienta que permita realizar este tipo de mediciones, es necesario en primer lugar, tener claridad sobre el concepto de ventaja o desventaja relativa que enfrenta la fuerza laboral al momento de buscar un empleo. Este concepto ha sido ampliamente discutido en la literatura y si bien, no hay un consenso general al respecto, si es claro que diferentes autores recurren a conceptos relacionados con la empleabilidad o las vulnerabilidades que tienen las personas que se encuentran en un proceso de búsqueda laboral y de esta forma entender las ventajas o limitaciones de una persona en el mercado laboral.

Con respecto a la empleabilidad, diferentes autores definen este concepto a partir del alcance que pretenda tener su investigación o del contexto histórico o del mercado laboral en donde se utilice el término. En este sentido, es de bastante utilidad para el presente trabajo, la revisión que realizan

Grip, Loo y Sanders, en donde hacen un recuento histórico sobre las diferentes definiciones que han sido adoptadas desde principios del siglo XX y concluyen que la empleabilidad se puede entender como “la capacidad y la voluntad de los trabajadores de seguir siendo apreciados en el mercado de trabajo (factores de la oferta), anticipándose y reaccionando a los cambios de las tareas profesionales y del entorno laboral (factores de demanda), con la ayuda de los instrumentos de desarrollo de recursos humanos que se dispongan (instituciones)”¹

Si bien esta definición recoge elementos importantes del mercado laboral como la oferta y la demanda de trabajo, así como factores institucionales que facilitan el acceso a un empleo, dejan de lado elementos que van más allá de la capacidad y voluntad de un individuo y que tienen que ver con la posibilidad de obtener un empleo, dado que pueden haber personas que cuenten con estas características pero que aun así, dado el contexto del mercado de trabajo, pueden tener dificultades para engancharse. Así mismo, se excluye de esta definición otros factores del contexto del individuo e imaginarios culturales que pueden hacerlo más vulnerable.

Para complementar esta definición, se recurre en primera medida a una definición de empleabilidad que presentan Grip, Loo y Sanders en su investigación, donde explican que “en los años cincuenta y sesenta la empleabilidad se entendía como las posibilidades de ser contratado que tenía un individuo. Se describía básicamente como la distancia que separa a las personas de un empleo estable”²

Esta segunda definición acerca más hacia el concepto de empleabilidad entendido como una medición que permite identificar las brechas que presenta una persona para conseguir un empleo y esas brechas pueden ser diferentes en grado e intensidad dependiendo de las barreras que enfrente el individuo.

Adicionalmente a estas dos anteriores definiciones, es importante recurrir al concepto de vulnerabilidad laboral desarrollado por Weller (2009). Este concepto básicamente explica que “la vulnerabilidad laboral consistiría en que (ciertos grupos) a causa de las carencias de sus activos (conocimientos, habilidades, experiencia, acceso a capital, etc.), no tienen acceso a actividades y

¹Grip, A., Loo, J., Sanders, J. (2004) The Industry Employability Index: Taking account of supply and demand characteristics. Ginebra, Suiza: International Labour Review, Vol.143 (2004), N°3.

² Ob. Cit. p. 245



oportunidades laborales que les facilitarían los ingresos para mejorar significativamente sus condiciones de vida.”³

A partir de estos tres elementos, se puede construir un concepto de empleabilidad que permita dar el soporte conceptual que necesita el índice de empleabilidad para Colombia. Básicamente, se entenderá como empleabilidad a la posibilidad que tiene una persona para ocupar un trabajo a partir de sus características (capacidad, voluntad, conocimientos, habilidades, experiencia, etc) y demás variables de contexto (personales, territoriales, económicas, etc) y que son susceptibles de mejorar, a partir de acciones de política pública que permitan mitigar sus barreras y cerrar las brechas.

Una vez se tiene claridad sobre los elementos que componen la definición de empleabilidad y que serán susceptibles de ser medidos para el cálculo del índice, se debe proceder a identificar cómo se debe realizar esa medición y cómo se deben articular todos esos elementos en un solo indicador compuesto. En este sentido el siguiente paso de la investigación, consiste en identificar en la literatura la forma en que se realiza la construcción de este tipo de mediciones y cómo se debería adecuar para lograr medir los elementos anteriormente mencionados.

En este sentido, en la literatura se encuentran varios autores que han implementado este tipo de herramientas para determinar el grado de empleabilidad de los buscadores. Entre ellos se destaca, por ejemplo, el índice desarrollado por Grip, Loo y Sanders, en donde construyen un índice de empleabilidad sectorial a partir de un conjunto de indicadores que miden la empleabilidad de la fuerza de trabajo, la necesidad de empleabilidad sectorial y las condiciones de consecución. A partir de esto construyen el índice para trece diferentes sectores de la economía y obtienen un escalafón de los sectores con mayor y menor empleabilidad.

Por otra parte, se destaca el ejercicio realizado por Rodríguez, en donde implementa un índice de empleabilidad para los jóvenes. Este instrumento permite medir diversas habilidades que los buscadores adquieren durante su proceso de formación y capacitación, así como la incidencia de otros factores de índole personal, económica o social que pueden condicionar su probabilidad de engancharse en un empleo. Como resultado, el autor encuentra que los jóvenes que tienden a ser

³Weller, J. (2009). El fomento de la inserción laboral de grupos vulnerables. Consideraciones a partir de cinco estudios de caso nacionales. Santiago de Chile, Chile: CEPAL.

más empleables tienen en común un alto desarrollo de sus habilidades básicas, así como acceso a las tics y están concluyendo sus procesos de formación.

Otro ejercicio destacado de implementación de este tipo de herramientas es el instrumento de clasificación de buscadores de empleo (JSCI por sus siglas en inglés) desarrollado por el Servicio Público de Australia. Este instrumento permite calcular la probabilidad de que una persona permanezca en situación de desempleo en caso de no recibir el acompañamiento necesario en su proceso de búsqueda laboral y además, provee una medida de la desventaja relativa en el mercado de trabajo con base en las características del individuo.

Este ejercicio, tiene un importante aporte, dado que va más allá del planteamiento conceptual y aplica la herramienta a los buscadores que van al centro de empleo. En efecto, la aplicación del índice se realiza a través de un cuestionario que es complementado con preguntas sobre temas que influyen en la búsqueda de empleo y el nivel relativo de desventaja se encuentra expresado en un puntaje que representa medida relativa y no absoluta de las circunstancias de cada buscador de trabajo.

El instrumento se divide en factores (que incluyen subfactores), en donde se agrupan las preguntas relacionadas, las cuales arrojan un puntaje: a más alto el valor del indicador mayor es la probabilidad de desempleo por un año o más. Los factores utilizados para la construcción de la herramienta contienen información sobre las características de la persona, educación, vulnerabilidades, ubicación geográfica y cualificación técnica entre otros.

Para estos índices y otros que se destacan en la literatura, se encuentra que si bien, las fuentes de información con las que se construyen pueden variar, la metodología siempre consiste en buscar algún procedimiento o algoritmo que permita reducir la dimensionalidad de las variables y así crear un único índice de fácil interpretación, de forma tal que sea de uso práctico para la implementación de políticas y programas de cierre de brechas.

A partir de los elementos mencionados, es claro que el índice de empleabilidad debe ser un instrumento que permita calcular un indicador relativo sobre las características y demás variables de contexto que facilitan o dificultan el acceso a un trabajo por parte de la fuerza laboral, así mismo permita identificar el grado y la intensidad de las barreras que enfrentan los buscadores para facilitar su mitigación y el cierre de brechas.



Una vez se tiene claro el alcance del concepto de empleabilidad y cómo se mide a través de un indicador, se puede establecer que el índice de empleabilidad para Colombia permitirá alcanzar tres grandes objetivos:

1. Establecer cuál es la ventaja o desventaja relativa de un buscador en el mercado laboral.

En efecto, el primer objetivo derivado de la construcción de un índice de empleabilidad es poder aplicarlo a cada una de las personas que se encuentran buscando un empleo y de esta forma obtener un puntaje que indique la posibilidad que tiene de ser ocupado en el mercado laboral.

2. Determinar cuáles son las barreras que más inciden en la posibilidad de enganche de una persona y de esta forma incidir a través de programas a la medida para contribuir en la mitigación de las mismas.

El segundo gran objetivo del índice consiste en identificar cuáles son las dimensiones en las que se presentan las mayores brechas y de esta forma, a través de programas y acciones, incidir en la superación de sus barreras y facilitar el ingreso a un trabajo formal.

3. Identificar grupos vulnerables con barreras transversales que sean objeto de políticas activas y pasivas de empleo.

Finalmente, el tercer objetivo del índice consiste en poder identificar grupos poblacionales que presenten barreras similares en forma e intensidad de manera tal, que se puedan implementar programas masivos de gestión del recurso humano que les facilite el acceso a un trabajo.

Ahora bien, una vez se tiene claro qué se espera determinar con el índice, el problema surge en identificar cuál debe ser la metodología que permita hacer esa medición teniendo en cuenta que debe tener la capacidad de tomar un gran conjunto de información y sintetizarlo de tal forma que se asegure la menor pérdida de información posible. A continuación se presenta la metodología seleccionada para poder llevar a cabo este procedimiento y realizar el cálculo del índice.

Análisis de componentes principales

La construcción del índice de empleabilidad requiere de una metodología que permita reducir el número de variables, dado que el gran conjunto de información relacionado con la determinación de la empleabilidad de un buscador dificulta cualquier tipo de análisis.

En efecto, para realizar la construcción del indicador se recurre a la metodología de análisis de componentes principales (ACP). Esta metodología desarrollada por Pearson en 1901, busca

explicar la estructura de la varianza de un conjunto de variables a través de un conjunto mucho menor de componentes o constructos que recogen la mayor proporción de la variabilidad. Es decir que se parte de un conjunto de p variables aleatorias correlacionadas y se transforma, a partir de combinaciones lineales de las variables originales, en otro conjunto más pequeño de variables hipotéticas o latentes no correlacionadas, que presentan la menor pérdida de información posible y facilitan la interpretación de la información contenida en los datos iniciales.

Luego, la metodología de ACP pretende encontrar un subespacio de dimensión $r < p$ tal que al proyectar los puntos sobre ese espacio conserven su estructura con la menor distorsión posible. Adicionalmente, la metodología de componentes principales permite descubrir e interpretar interdependencias entre las variables y examinar relaciones que puedan existir entre los individuos. Para implementar el análisis de componentes principales se parte de un conjunto de información para todos los individuos, sea $X = \{x_1, \dots, x_p\}$, la matriz de datos de p variables para la muestra de individuos y se busca reducir la dimensión a partir de una combinación lineal de las variables X con la menor pérdida de información posible, para ello se busca un espacio de dimensión $r < p$ tal que al proyectar los puntos sobre ese espacio conserven su estructura con la menor pérdida posible, es decir, las distancias entre los puntos originales y sus proyecciones sobre el subespacio de dimensión r deben ser lo más pequeñas posibles.

Formalmente, las componentes principales se determinan como:

$$z_i = a_{11}x_{i1} + \dots + a_{1p}x_{ip} = a_1'x_i$$

Donde

$a_1 = (a_{11}, a_{12}, \dots, a_{1p})$ es el vector de norma uno que define la dirección de la proyección sobre x_i en el espacio r .

z_i es el vector de nuevas variables o componentes que representan la proyección de x_i en el espacio r .

Y por tanto, $z_i = a_1'x_i$ se denomina la transformación por componentes principales.

Determinación de componentes

El primer paso de la metodología de ACP consiste en determinar los componentes que surgirán a partir de las variables originales. En este sentido, el primer componente principal se



define como la combinación lineal de las variables originales que tienen máxima varianza. Luego, el primer componente para los n individuos se define como

$$z_1 = Xa_1$$

Como suponemos que las variables originales estaban centradas, el vector z_1 también tendrá media cero y su varianza es:

$$\frac{1}{n} z_1' z_1 = \frac{1}{n} a_1' X' X a_1 = a_1' S a_1$$

Sin embargo, se encuentra que la varianza se puede maximizar de manera arbitraria aumentando el módulo del vector a_1 y por tanto, para garantizar que la maximización de la varianza tenga solución, se impone, sin pérdida de generalidad, que la norma sea unitaria $a_1' a_1 = 1$.

Luego, como resultado se tiene un problema de maximización con una restricción, el cual se soluciona a partir del multiplicador de Lagrange

$$\max_{a_1} a_1' S a_1 \quad \text{s. t. } a_1' a_1 = 1$$

Obteniendo la siguiente condición de primer orden,

$$\frac{\partial \mathcal{L}}{\partial a_1} = 2S a_1 - 2\lambda a_1 = 0$$

$$S a_1 = \lambda a_1$$

Esta solución implica que a_1 es un autovector de la matriz S de varianzas y covarianzas de las observaciones y λ su correspondiente autovalor.

Ahora bien, para determinar el autovalor de S que corresponden a la solución, se pre-multiplica la solución por a_1' de forma tal que:

$$a_1' S a_1 = \lambda a_1' a_1$$

Y por tanto se obtiene que λ es la varianza de z_1 . Dado que esta es la cantidad que se desea maximizar, λ será el autovalor más grande de la matriz S . Así mismo, el autovector a_1 contiene los coeficientes de cada variable en la combinación lineal del primer componente principal.

A continuación se procede generalizar el procedimiento de componentes principales. En general, la matriz X y por tanto S , tienen rango p , entonces existen tantas componentes principales como variables. Luego, para encontrar el mejor plano de proyección de las variables, se debe establecer una función en donde se maximice la suma de las varianzas de los $Z = (z_1, z_2, \dots, z_p) = (Xa_1, Xa_2, \dots, Xa_p)$ donde a_i son los vectores que definen el plano:

$$Q = \sum_{i=1}^p a_i' S a_i - \sum_{i=1}^p \lambda_i (a_i' a_i - 1)$$

Obteniendo las condiciones de primer orden:

$$\begin{aligned} \frac{\partial Q}{\partial a_1} &= 2S a_1 - 2\lambda_1 a_1 = 0 \\ \frac{\partial Q}{\partial a_2} &= 2S a_2 - 2\lambda_2 a_2 = 0 \\ &\vdots \\ \frac{\partial Q}{\partial a_p} &= 2S a_p - 2\lambda_p a_p = 0 \end{aligned}$$

Solucionando el sistema se obtiene:

$$\begin{aligned} S a_1 &= \lambda_1 a_1 \\ S a_2 &= \lambda_2 a_2 \\ &\vdots \\ S a_p &= \lambda_p a_p \end{aligned}$$

Esto indica que los a_i deben ser los autovalores de S . Ahora, sustituyendo los autovectores de norma unitaria en Q se obtiene que en el máximo, la función es $Q = \sum_{i=1}^p \lambda_i$ y por tanto queda claro que los λ_i deben ser los i autovalores más grandes de la matriz S y los a_i sus correspondientes autovectores.

Como resultado de este procedimiento, se encuentra que los componentes principales no están correlacionados:

Para todo los $i \neq j$ se tiene que

$$Cov(z_i, z_j) = Cov(X a_i, X a_j) = \frac{1}{n} a_i' X' X a_j = a_i' S a_j = \lambda_j a_i' a_j = 0$$

A partir de la estimación de las componentes, se derivan las siguientes propiedades:

1. Las componentes principales conservan la variabilidad inicial, es decir que la suma de las varianzas de los componentes es igual a la suma de las varianzas de las variables originales y la varianza generalizada de los componentes es igual a la original:

$$\sum_{i=1}^p Var(X_i) = \sum_{i=1}^p \lambda_i = \sum_{i=1}^p Var(z_i)$$



2. La proporción de la variabilidad explicada por un componente es la relación entre su varianza y la suma de autovalores de la matriz S

$$\text{Varianza explicada } (z_i) = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i}$$

3. La correlación entre una componente principal y una variable X_k es proporcional al coeficiente de esa variable en la definición de la componente (a_{ik}), donde el coeficiente de proporcionalidad es la relación entre la desviación estándar del componente y la desviación de la variable.

$$\text{Corr}(z_i, X_k) = a_{ik} \frac{\sqrt{\lambda_i}}{\sqrt{\text{Var}(X_k)}}$$

Selección de la cantidad de componentes

Una vez se establece la forma en que se estiman los componentes y sus propiedades, se procede a determinar cuántos de esos componentes son significativos para explicar la mayor variabilidad posible de todo el conjunto de información. En la literatura se encuentran varios criterios que pueden facilitar la determinación del número de componentes a utilizar:

1. Criterio de porcentaje

El número de componentes principales se determina de forma tal que las primeras k componentes expliquen un determinado porcentaje de la varianza total.

2. Criterio de Kaiser

A partir de la matriz de correlaciones R se puede obtener las componentes principales suponiendo que las variables tienen varianza 1, por tanto toda componente principal que no tenga una varianza superior o igual a 1 explicará menos variabilidad que cualquiera de las variables observadas, luego el criterio de Kaiser determina que mantenemos las primeras k componentes tales que $\lambda_k \geq 1$.

3. Test de esfericidad

El test de esfericidad consiste en plantear una hipótesis nula en la cual hay k componentes principales con máxima variabilidad y a partir del componente $k + 1$ ya no hay direcciones adicionales que aporten significativamente a la variabilidad, es decir que la distribución de los

datos es esférica. Para realizar el test se parte del supuesto que la matriz de datos se distribuye normal multivariada $N_p(\mu, \Sigma)$. El test se basa en el estadístico chi-cuadrado y se aplica de manera secuencial: el primer test considera como hipótesis nula $k = 0$, si se rechaza se procede a aplicar el test con la hipótesis nula de $k = 1$ y así sucesivamente se repite el proceso hasta que no haya evidencia estadística para rechazar H_0 y se pueda determinar el k^* .

Finalmente, se ha determinado el número de componentes se procede a estimar la cantidad de varianza que aporta cada uno y en total y de esta forma ya se tienen los elementos necesarios (componentes y ponderadores) para calcular el índice de empleabilidad.

Una vez se tienen definidos todos los elementos que inciden en la construcción de un índice de empleabilidad, los cuales abarcan desde la definición del concepto de empleabilidad, la forma en que se debe definir el indicador, sus objetivos y la herramienta econométrica para su medición, a continuación se presenta la metodología para establecer la construcción del indicador.



3. Metodología

A continuación se presenta la metodología para la construcción del índice de empleabilidad, en la cual se especifica cada uno de los pasos a seguir para obtener este indicador. Lo primero que se presenta es la identificación de la fuente de información y las variables que componen el conjunto de información que caracteriza a las personas que hacen parte del mercado laboral. Esta información debe contener las características y variables de contexto que inciden realmente en la posibilidad, que tiene un buscador, de engancharse en un puesto de trabajo.

Para garantizar que esta información es la que realmente incide en la probabilidad de enganche, se procede a determinar las variables que efectivamente servirán para el cálculo del índice. Con este propósito, se implementan dos criterios: la relación con la tasa de desempleo y la significancia en la probabilidad de estar ocupado, a partir de diferentes herramientas econométricas y de análisis multivariado como los modelos probit.

Una vez se han definido las variables, se realiza la reducción de la dimensión del conjunto de información a través de la metodología de componentes principales, por medio de la cual se estiman los constructos que contienen la mayor cantidad de información de las variables originales, así mismo, se estiman los ponderadores de cada uno de estos componentes y finalmente se procede a estimar el índice teniendo en cuenta la varianza que explica cada factor y las cargas factoriales derivadas de la respectiva matriz.

Para iniciar con el proceso de construcción del indicador, el primer paso consiste en identificar un conjunto de información que caracterice a cada uno de los individuos que hacen parte de la fuerza laboral. En Colombia existen diversas fuentes de información que podrían permitir realizar el cálculo de este indicador, por ese motivo es necesario realizar un primer levantamiento de las fuentes de información existentes. Entre las fuentes de información identificadas se tuvieron en cuenta tanto encuestas de hogares como la Gran encuesta Integrada de Hogares, la Encuesta de Calidad de vida y la Encuesta Longitudinal de Protección Social. Así mismo se analizaron registros administrativos como la base de oferentes del Servicio Público de Empleo. Sin embargo, dado que el objetivo de la fuente de información es contener datos representativos de toda la fuerza laboral del país y que sea robusta tanto en temas de mercado laboral como en otras variables de tipo social,

económico e institucional, se determinó que la fuente de información más fiable sería la Encuesta de Calidad de Vida 2016, dado que cumple con todos los requisitos enunciados.

En efecto, la Encuesta de Calidad de Vida 2016, es una encuesta de hogares que presenta datos socioeconómicos de la población colombiana y permite conocer la situación de bienestar de todos los colombianos a partir de diferentes dimensiones como acceso a bienes y servicios, salud, educación, empleo, entre otros. Todo este conjunto de información permitirá identificar y construir las variables que inciden en la probabilidad de que una persona acceda a una oportunidad laboral. Una vez establecida la fuente de información, se procede a realizar la identificación de las preguntas que podrían tener incidencia en la empleabilidad de una persona. Para realizar esta identificación se recurre a la literatura relacionada con los temas de mercado laboral y empleabilidad.

Posteriormente, Identificadas las preguntas a utilizar, se procede a construir las variables y se categorizan ordinalmente, de tal forma que el puntaje se incrementará a medida que esa característica incida en menor medida en la tasa de desempleo y en mayor medida en la probabilidad de colocarse en el mercado laboral. Luego, se procede a constatar que cada variable se relacione de manera positiva con la probabilidad de estar ocupado, para esto se implementan modelos probit entre la variable ocupado y cada una de las variables definidas, en el caso que el estimador no sea significativo o el signo no sea el esperado, se procede a retirar o modificar la variable analizada respectivamente.

Con este procedimiento se definen las variables que inciden en la posibilidad de estar ocupado y así se realiza la implementación de la metodología de análisis de componentes principales. Como se mencionó anteriormente, esta metodología permite reducir la dimensión del conjunto de información, definiendo un conjunto de nuevas variables o componentes que explican la mayor parte de la variabilidad del conjunto de información, con una menor cantidad de variables que las originales. Así mismo, este procedimiento facilita el cálculo de los ponderadores del índice a partir de la estimación de la varianza explicada por cada componente. En esta etapa se tienen en cuenta diferentes criterios como el de adecuación de la muestra (KMO) para establecer las variables que definitivamente harán parte del índice y cuales se deberán descartar.

Una vez definidas las variables, se determina el número de factores que explican la mayor cantidad de variabilidad de los datos, en este punto se establece que todos los componentes con autovalores mayores a la unidad serán tenidos en cuenta para el cálculo del índice. Una vez determinado el



número de componentes, se procede a definir las variables que hacen parte de cada componente y a estimar sus respectivos valores (dimensiones). Para estimar estos valores se recurre a la matriz de cargas factoriales y se define que todos los valores de las cargas factoriales que sean superiores a 0,25 indican una fuerte incidencia de esa variable en ese constructo y por tanto se utilizará para calcular el valor del componente.

Paso seguido, se procede a calcular el índice de empleabilidad, para esto se realiza el siguiente procedimiento: en primer lugar se toman las puntuaciones que ha obtenido una persona para cada una de las variables y se agrupan según la identificación de los componentes, se suman y luego se normaliza por la máxima puntuación que se puede obtener en cada componente.

$$C_{k,i}^* = \frac{\sum_{k=1}^K x_{k,i}}{\max \sum_{k=1}^K x_k}$$

Una vez quedan definidos los componentes para cada persona, se procede a ponderar por la proporción de la varianza que es explicada por el componente, para esto, lo primero que hay que realizar es una normalización de la varianza explicada por el número de componentes seleccionados, es decir se toma la proporción de varianza y se divide por la suma del total de la varianza explicada de todos los componentes escogidos.

$$\% \text{ Varianza explicada normalizada } C_k = \frac{\text{Varianza explicada } C_k}{\sum_{k=1}^K \text{Varianza explicada } C_k}$$

Finalmente, para calcular el valor del índice, se multiplica cada componente estimado por su proporción de varianza explicada normalizada y se suma para todos los componentes:

$$\text{Índice de empleabilidad}_i = \sum_{k=1}^K (C_{k,i} \times \% \text{ Varianza explicada normalizada } C_k)$$

Con este cálculo se obtiene el índice de empleabilidad para aplicarlo a cada una de las personas que hacen parte de la fuerza laboral de la encuesta y proceder a realizar los análisis derivados de los resultados del mismo.

4. Datos y Estimación

A continuación se presenta todo el proceso de estimación del índice de empleabilidad, en primer lugar se muestra cómo se determina el conjunto de información que permitirá construir las variables que harán parte del indicador. Posteriormente, se caracterizan todas las variables determinadas y se procede a desarrollar una metodología de priorización de las variables a partir de dos criterios: relación entre la tasa de desempleo y la relación de cada variable y su significancia en la probabilidad de estar ocupado. Una vez determinado el conjunto de información, se procede a desarrollar la metodología de análisis de componentes principales para determinar el número de componentes que harán parte del índice, cuál es su composición y el ponderador asociado a cada uno.

Finalmente, se presentan los resultados de la estimación del índice para todos los individuos que hacen parte de la fuerza laboral, se analiza el grado de empleabilidad promedio y de los extremos de la distribución y se realiza la estimación del mismo para diferentes grupos de interés: en primer lugar se comparan las distribuciones entre ocupados y desocupados y se verifica si existen en promedio diferencias significativas entre los dos grupos, en segundo lugar se divide la muestra en cuartiles y se estima para cada uno el índice promedio, estos grupos se definen por grados de empleabilidad bajo, medio bajo, medio alto y alto y se realiza un ejercicio de identificación de brechas para cada componente y las principales barreras que inciden en esa brecha, por último se analiza las posibles acciones de política pública que podrían ayudar a mitigar tales barreras.

Conjunto de información

El primer paso para la determinación de un índice de empleabilidad para Colombia consiste en identificar y definir la información que permitirá estimarlo. Para esto se identifican diferentes fuentes de información y se realiza un ejercicio de validación para determinar cuál de estas es la mejor opción para construir el índice. En este sentido, se analizaron diferentes encuestas de hogares como la Gran Encuesta Integrada de Hogares, la Encuesta de Calidad de Vida y la Encuesta Longitudinal de Protección Social, así mismo se valora la información de registros administrativos como el reporte de oferentes del Servicio Público de Empleo. Para cada una de estas fuentes se plantearon las ventajas y desventajas de ser utilizadas como insumo en la construcción del índice y se concluyó que por el contenido de información tanto social, económica, institucional y



relacionada con posibles vulnerabilidades asociadas al individuo, así como datos del mercado laboral que son consistentes con los estimados a través de otros instrumentos y su representatividad con relación al total de la fuerza laboral del país, se estableció que la Encuesta de Calidad de Vida 2016 es la más idónea para realizar esta investigación.

En efecto, a partir de estos criterios se escoge la Encuesta de Calidad de Vida 2016 y a partir de diferentes elementos establecidos en la literatura relacionada con mercado laboral y empleabilidad se seleccionan las preguntas que permitirán identificar las características que pueden incidir en la consecución de un empleo para todas las personas que hacen parte de la fuerza laboral.

A continuación se procede a realizar la construcción de cada una de las variables. Para hacer este procedimiento, se toman las posibles opciones de respuesta de las preguntas identificadas y se determina una escala ordinal en donde se asigna un puntaje que se incrementa a medida que la opción incide positivamente en la posibilidad de acceder a un empleo. Como se muestra en la tabla 1, inicialmente se tomaron de la ECV 2016, 23 preguntas para construir las variables que incidirán en la estimación del índice. A continuación se presentan las variables organizadas en 7 grupos que se establecieron de forma preliminar y se define cada una de sus posibles opciones:

1. Variables relacionadas con el territorio: las variables relacionadas con el territorio permiten capturar información acerca de las particularidades de los lugares de residencia de los buscadores, establecer diferencias entre los territorios y cómo se puede potenciar o limitar a una persona en su afán de buscar un empleo, a partir de su espacio geográfico. En esta categoría se definen dos variables:
 - Ubicación: esta variable contiene la información relacionada con el lugar de residencia de la persona, el cual puede ser en la cabecera (ciudades grandes o intermedias), un centro poblado (ciudades pequeñas, veredas o corregimientos) y áreas rurales dispersas.
 - Departamento: la variable departamento se construye a partir de los 23 principales departamentos del país, se toma la información y a partir del tamaño de su mercado laboral se clasifica en tres categorías.
2. Variables relacionadas con características del individuo: este grupo de variables explica aspectos propios del individuo que inciden en la probabilidad de encontrar un empleo y no

son susceptibles de ser modificados a través de alguna acción tanto del individuo como de un tercero. En esta categoría se clasifican cuatro variables:

- Sexo: esta variable se define a partir del sexo reportado por la persona en ECV 2016, masculino o femenino. Es importante aclarar que en este caso, la variable captura exclusivamente la información sobre el sexo del buscador y no de su género.
 - Edad: la variable edad se determina a partir de la agrupación de las personas en seis rangos, la primera categoría contiene a todas las personas menores de 20 años y las restantes contienen la población dividida en grupos de 10 años y finalmente la última categoría agrupa a todos los mayores de 60 años.
 - Pareja: esta variable contiene la información relacionada con el estado civil actual del encuestado y se agrupa en dos opciones, las personas que tienen pareja (casados y unión libre) y los que no tienen pareja (solteros, viudos y divorciados).
 - Jefe de hogar: Esta variable contiene información sobre si la persona es jefe de hogar o no lo es.
3. Variables relacionadas con características del hogar: Este conjunto de variables agrupa la información relacionada con características que son comunes a todos los miembros del hogar e inciden en el proceso de búsqueda laboral, dado que puede generar barreras a la entrada a un empleo. Este grupo se compone de tres preguntas:
- Estrato: esta variable contiene la información del estrato socioeconómico al que pertenece la persona y se divide en cuatro opciones: sin estrato, bajo (estratos 1 y 2), medio (estratos 3 y 4) y alto (estratos 5 y 6).
 - Personas en el hogar: En esta variable se agrupan en cinco categorías a las personas encuestadas, cada opción recoge dos grupos de personas e inicia con las personas que tienen hogares con 1 o 2 personas y aumenta de a pares hasta la última opción en donde se ubican a todas las personas con más de 8 miembros en el hogar.
 - Ingresos per cápita: esta variable se construye a partir de dividir a la población en cuartiles y por tanto cada uno contiene el 25% del total de la muestra.
4. Variables relacionadas con vulnerabilidades del individuo: este tipo de variables presentan la información sobre las vulnerabilidades que enfrenta un individuo en su vida cotidiana, que si bien no están relacionadas directamente con el mercado de trabajo, si tienen



incidencia en su posibilidad de colocarse en un puesto de trabajo. Adicionalmente, estas variables satisfacen la condición de ser susceptibles de mitigarse a través de políticas sociales que mejoren la calidad de vida de la persona. Este grupo de información está conformado por seis variables:

- Pobre: esta variable se determina a partir de una pregunta de auto reconocimiento por parte de la persona si se considera pobre o no.
 - Subsidios: esta variable se determina si la persona ha recibido algún subsidio en dinero o en especie durante los últimos doce meses o no.
 - Cambio de municipio: esta variable se especifica a partir de si la persona se ha cambiado de municipio durante los últimos tres años o no.
 - Servicio de salud: esta variable determina si la persona se encuentra afiliada a algún régimen de salud o no.
 - Estado de salud: la variable estado de salud contempla cuatro opciones entre malo, regular, bueno o muy bueno.
 - Discapacidad: define si la persona tiene alguna incapacidad o no.
5. Variables laborales: Este grupo de variables tienen una alta incidencia en las posibilidades que tiene un buscador de encontrar un empleo y todas se relacionan con aspectos laborales del individuo que pueden generar ventajas o limitaciones para acceder a un puesto de trabajo. Este grupo está conformado por cuatro variables:
- Satisfacción en el trabajo: Se construye a partir de la pregunta que tan satisfecho o insatisfecho se encuentra en su trabajo y contempla cuatro opciones de respuesta desde muy insatisfecho hasta muy satisfecho.
 - Ocupación: esta variable se construye a partir de la clasificación de ocupaciones (CIUO-08) y permite clasificar a las personas que hacen parte de la encuesta en 10 grandes grupos ocupacionales.
 - Semanas de búsqueda: Esta variable se construye a partir del número de semanas que la persona ha estado o estuvo buscando trabajo y se agrupan en ocho opciones desde las personas que llevan más de 24 semanas, hasta las personas que llevan cero semanas buscando (ocupados).

- Trabajo por primera vez: esta variable hace distinción entre las personas que son aspirantes y las que se encuentran cesantes u ocupadas.
6. Variables relacionadas con el uso de las tecnologías de la información y las comunicaciones: el grupo que contiene las variables relacionadas con el uso y tenencia de tecnologías de la información es relevante para la medición dado que se hace cada vez necesario tener dominio de estas tecnologías para desempeñar de mejor forma un empleo y por tanto es relevante para medir la facilidad o dificultad relativa que presenta una persona para acceder a un empleo formal. Este grupo lo componen tres variables:
- Uso del computador: esta variable mide la frecuencia de uso del computador y lo clasifica en tres categorías, alto, medio o bajo.
 - Uso del internet: esta variable mide la frecuencia de uso del internet y lo clasifica en tres categorías, alto, medio o bajo.
 - Tenencia de celular: se define a partir de si la persona tiene o no teléfono celular.
7. Variables relacionadas con la educación: El último bloque de información está relacionada con el nivel educativo de los buscadores de empleo. para este grupo solo se considera una pregunta relacionada con la información del nivel educativo más alto alcanzado:
- Nivel educativo: Esta variable recoge la información sobre el nivel educativo más alto alcanzado (terminado o sin terminar) por parte de la persona.

Una vez definidas las variables seleccionadas, se presenta un resumen de los puntajes inicialmente establecidos y sus estadísticos descriptivos.

Tabla 1. Variables construidas a partir de la ECV 2016.

Variable	Puntuaciones	Variable	Puntuaciones
Ubicación	1. Cabecera	Servicio de salud	1. No
	2. Centro Poblado		2. Si
	3. Rural		
Estrato	1. Sin Estrato	Estado de salud	1. Malo
	2. Bajo		2. Regular
	3. Medio		3. Bueno
	4. Alto		4. Muy bueno



Índice de empleabilidad para Colombia

Personas en el hogar	1. 1-2 personas	Discapacidad	1. Si
	2. 3-4 personas		2. No
	3. 7-6 personas		
	4. 7-8 personas		
	5. Más de 8 personas		
Ingresos por rangos	1. menos de 220.500	Educación	1. Ninguno - Prescolar
	2. Entre 220.500 y 391.500		2. Primaria
	3. Entre 391.500 y 700.000		3. Secundaria
	4. Más de 700.000		4. Media
			5. Técnico
			6. Tecnólogo
			7. Universitario
			8. Posgrado
Pobre	1. Si	Ocupación	1. Fuerzas militares
	2. No		2. Técnicos y profesionales de nivel medio
			3. Trabajadores de los servicios y vendedores de comercios y mercados
			4. Directores y gerentes
			5. Ocupaciones elementales
			6. Personal de apoyo administrativo
			7. Profesionales, científicos e intelectuales
			8. Operadores de instalaciones fijas y maquinas
			9. Oficiales, operarios, artesanos y oficios relacionados
			10. Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros

Subsidio	1. Si 2. No	Semanas de Búsqueda	1. Más de 24 semanas 2. Entre 20 y 24 semanas 3. Entre 16 y 20 semanas 4. Entre 12 y 16 semanas 5. Entre 8 y 12 semanas 6. Entre 4 y 8 semanas 7. Menos de 4 semanas 8. Ocupado
Satisfacción en el Trabajo	1. Muy Insatisfecho 2. Insatisfecho 3. Satisfecho 4. Muy Satisfecho	Trabajo por primera vez	1. Si 2. No
Sexo	1. Mujer 2. Hombre	Frecuencia de uso del computador	1. Alto 2. Medio 3. Bajo
Edad por rangos	1. Menos de 20 años 2. Entre 20 y 30 años 3. Entre 30 y 40 años 4. Entre 40 y 50 años 5. Entre 50 y 60 años 6. Más de 60 años	Frecuencia de uso del internet	1. Alto 2. Medio 3. Bajo
Jefe de hogar	1. No 2. Si	Tiene celular	1. No 2. Si
Pareja	1. Si 2. No	Departamento	1. Tipo 1 2. Tipo 2 3. Tipo 3
Cambio de municipio	1. Si 2. No		

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Adicionalmente, en el tabla 2 se presentan los estadísticos de cada una de las variables anteriormente definidas. En todos los casos, la muestra cuenta con 33.281 observaciones y se analiza la media, la desviación estándar y el valor mínimo y máximo de cada variable.



Tabla 2. Estadísticos descriptivos de las variables seleccionadas.

Variable	Observaciones	Media	Desviación Estándar	Mínimo	Máximo
Ubicación	33.281	1,64	0,854	1	3
Estrato	33.281	2,14	0,472	1	4
Personas en el hogar	33.281	2,19	0,949	1	5
Ingresos por rangos	33.281	2,50	1,116	1	4
Pobre	33.281	1,60	0,489	1	2
Subsidio	33.281	1,98	0,141	1	2
Satisfacción en el Trabajo	33.281	2,73	0,555	1	4
Sexo	33.281	1,60	0,489	1	2
Edad por rangos	33.281	3,43	1,448	1	6
Jefe de hogar	33.281	1,50	0,500	1	2
Pareja	33.281	1,57	0,495	1	2
Cambio de municipio	33.281	1,92	0,277	1	2
Servicio de salud	33.281	1,94	0,234	1	2
Estado de salud	33.281	2,88	0,564	1	4
Discapacidad	33.281	1,98	0,154	1	2
Educación	33.281	3,68	1,798	1	8
Ocupación	33.281	5,32	2,285	0	9
Semanas de Búsqueda	33.281	7,78	1,003	1	8
Trabajo por primera vez	33.281	1,99	0,108	1	2
Computador	33.281	2,40	0,819	1	3
Internet	33.281	2,14	0,902	1	3
Celular	33.281	1,88	0,326	1	2
Departamento	33.281	2,45	0,738	1	3

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Una vez se tienen todas las variables determinadas se procede a calcular la tasa de desempleo que hay para cada categoría de la variable, en el caso en que la relación no sea inversa al puntaje, es

decir que a medida que aumenta el puntaje aumenta la tasa de desempleo, se deberá realizar una recategorización de la variable.

Tabla 3. Tasas de desempleo para cada categoría de las variables.

Variable	Opciones de respuesta	Tasa de desempleo
Ubicación	1. Cabecera	8.78
	2. Centro Poblado	8.88
	3. Rural	3.47
Estrato	1. Sin Estrato	6.05
	2. Bajo	8.62
	3. Medio	6.89
	4. Alto	4.76
Personas en el hogar	1. 1-2 personas	5.57
	2. 3-4 personas	7.88
	3. 7-6 personas	9.13
	4. 7-8 personas	11.62
	5. Más de 8 personas	12.22
Ingresos por rangos	1. menos de 220.500	13.58
	2. Entre 220.500 y 391.500	9.66
	3. Entre 391.500 y 700.000	7.24
	4. Más de 700.000	4.04
Pobre	1. Si	9.87
	2. No	7.17
Subsidio	1. Si	14.33
	2. No	7.96
Satisfacción en el Trabajo	1. Muy Insatisfecho	22.6
	2. Insatisfecho	15.69
	3. Satisfecho	5.08
	4. Muy Satisfecho	4.03
Sexo	1. Mujer	10.56
	2. Hombre	6.36
Edad por rangos	1. Menos de 20 años	19.2



Índice de empleabilidad para Colombia

	2. Entre 20 y 30 años	11.19
	3. Entre 30 y 40 años	6.99
	4. Entre 40 y 50 años	5.36
	5. Entre 50 y 60 años	4.65
	6. Más de 60 años	4.01
Jefe de hogar	1. No	11.52
	2. Si	4.42
Pareja	1. Si	10.93
	2. No	5.99
Cambio de municipio	1. Si	12.2
	2. No	7.71
Servicio de salud	1. No	22.96
	2. Si	7.18
Estado de salud	1. Malo	11.44
	2. Regular	8.91
	3. Bueno	7.95
	4. Muy bueno	7.24
Discapacidad	1. Si	9.74
	2. No	8.01
Educación	1. Ninguno - Prescolar	4.82
	2. Primaria	4.6
	3. Secundaria	9.24
	4. Media	9.32
	5. Técnico	12.09
	6. Tecnólogo	9.83
	7. Universitario	8.42
	8. Posgrado	3.27
Ocupación	1. Fuerzas militares	37.72
	2. Técnicos y profesionales de nivel medio	7.81
	3. Trabajadores de los servicios y vendedores de comercios y mercados	6.71
	4. Directores y gerentes	12.08
	5. Ocupaciones elementales	6.85
	6. Personal de apoyo administrativo	10.76
	7. Profesionales, científicos e intelectuales	2.38
	8. Operadores de instalaciones fijas y maquinas	4.2
	9. Oficiales, operarios, artesanos y oficios relacionados	4.9

10. Agricultores y trabajadores calificados 7.2
 agropecuarios, forestales y pesqueros

Semanas de Búsqueda	1. Más de 24 semanas	-
	2. Entre 20 y 24 semanas	-
	3. Entre 16 y 20 semanas	-
	4. Entre 12 y 16 semanas	-
	5. Entre 8 y 12 semanas	-
	6. Entre 4 y 8 semanas	-
	7. Menos de 4 semanas	-
	8. Ocupado	-
Trabajo por primera vez	1. Si	-
	2. No	6.91
Frecuencia de uso del computador	1. Alto	8.06
	2. Medio	12.23
	3. Bajo	6.62
Frecuencia de uso del internet	1. Alto	8.25
	2. Medio	12.24
	3. Bajo	6.1
Tiene celular	1. No	11.29
	2. Si	7.71
Departamento	1. Tipo 1	9.59
	2. Tipo 2	7.24
	3. Tipo 3	8.07

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
 Cálculos propios del autor.

A partir de los resultados de la tabla 3, se encuentra que según el criterio de relación inversa entre tasa de desempleo y puntaje de la variable se debe realizar una recategorización de las variables: ubicación, estrato, discapacidad, educación, ocupación, uso del computador e internet y departamento.

Una vez se realiza el ajuste señalado anteriormente a las variables, se procede a correr un modelo probit entre la variable ocupado (que toma el valor 1 si la persona está ocupada o 0 en caso contrario) como dependiente, contra cada una de las variables de interés y así se determina si el signo del coeficiente asociado es coherente con los supuestos iniciales y en caso de no ser así se debe realizar otra recalibración de la variable. Adicionalmente, se procede a revisar los niveles de significancia de la variable y en caso de que no tenga la significancia suficiente se supone que no



hay evidencia de que la variable influya en la determinación de la probabilidad de estar ocupado y por tanto se retira del modelo. A continuación se presentan los resultados de los modelos probit.

Tabla 4. Resultados modelos probit.

Variable	Coficiente	Significancia
Ubicación	0,236	-
Estrato	0,083	-
Personas en el hogar	0,127	-
Ingresos por rangos	0,202	-
Pobre	0,153	-
Subsidio	0,361	-
Satisfacción en el Trabajo	0,515	-
Sexo	0,309	-
Edad por rangos	0,188	-
Jefe de hogar	0,510	-
Pareja	0,353	-
Cambio de municipio	0,232	-
Servicio de salud	0,627	-
Estado de salud*	0,020	0,373
Discapacidad*	0,003	0,961
Educación	0,100	-
Ocupación	0,125	-
Semanas de Búsqueda**	-	-
Trabajo por primera vez**	-	-
Computador	0,184	-
Internet	0,194	-
Celular	0,118	-
Departamento	0,080	-

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

A partir de los resultados derivados de los modelos probit especificados, se encuentra que las variables estado de salud y discapacidad no inciden en la probabilidad de estar ocupado y por tanto se retirarán del modelo de componentes principales. Una vez determinadas las variables que inciden en la probabilidad de estar ocupado, se procede a estimar el modelo de componentes

principales y determinar si las variables satisfacen con el criterio de adecuación del muestreo, se considera que cualquier variable que tenga un puntaje inferior a 0,5 se retirará del modelo de componentes principales. Finalmente el modelo seleccionado incluirá las siguientes 18 variables con una nueva categorización⁴:

Tabla 5. Variables seleccionadas con opciones ajustadas por criterios.

Variable	Puntuaciones	Variable	Puntuaciones
Ubicación	1. Centro Poblado	Jefe de hogar	1. No
	2. Cabecera		2. Si
	3. Rural		
Estrato	1. Sin Estrato	Pareja	1. Si
	2. Bajo		2. No
	3. Medio		
	4. Alto		
Personas en el hogar	1. 1-2 personas	Educación	1. Técnico
	2. 3-4 personas		2. Tecnólogo
	3. 7-6 personas		3. Media
	4. 7-8 personas		4. Secundaria
	5. Más de 8 personas		5. Universitario
			6. Ninguno - Prescolar
			7. Primaria
			8. Posgrado
Ingresos por rangos	1. menos de 220.500	Ocupación	1. Fuerzas militares
	2. Entre 220.500 y 391.500		2. Técnicos y profesionales de nivel medio
	3. Entre 391.500 y 700.000		3. Trabajadores de los servicios y vendedores de comercios y mercados
	4. Más de 700.000		4. Directores y gerentes
			5. Ocupaciones elementales
			6. Personal de apoyo administrativo
			7. Profesionales, científicos e intelectuales
			8. Operadores de instalaciones fijas y

⁴ Se descartaron las variables cambio de municipio, servicio de salud y celular por tener un kmo menor a 0,5, lo que implica que no presentan la suficiente capacidad para asociarse con las demás variables al momento de realizar el procedimiento de ACP.



Índice de empleabilidad para Colombia

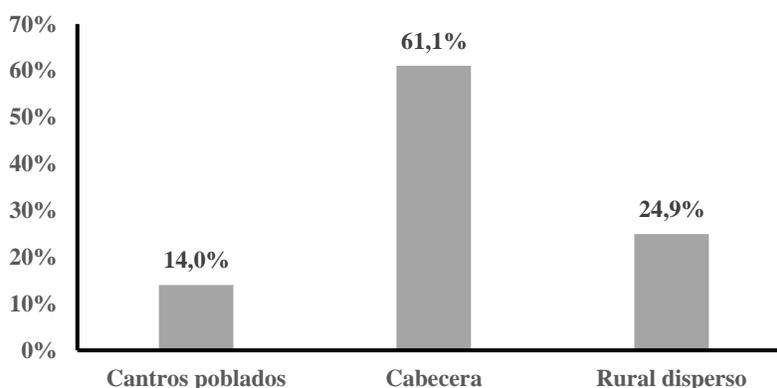
			maquinas
			9. Oficiales, operarios, artesanos y oficios relacionados
			10. Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros
Pobre	1. Si 2. No	Semanas de Búsqueda	1. Más de 24 semanas 2. Entre 20 y 24 semanas 3. Entre 16 y 20 semanas 4. Entre 12 y 16 semanas 5. Entre 8 y 12 semanas 6. Entre 4 y 8 semanas 7. Menos de 4 semanas 8. Ocupado
Subsidio	1. Si 2. No	Trabajo por primera vez	1. Si 2. No
Satisfacción en el Trabajo	1. Muy Insatisfecho 2. Insatisfecho 3. Satisfecho 4. Muy Satisfecho	Uso del computador	1. Medio 2. Alto 3. Bajo
Sexo	1. Mujer 2. Hombre	Uso del internet	1. Medio 2. Alto 3. Bajo

Edad por rangos	Departamento	Departamento
1. Menos de 20 años	1. Tipo 1	
2. Entre 20 y 30 años	2. Tipo 3	
3. Entre 30 y 40 años	3. Tipo 2	
4. Entre 40 y 50 años		
5. Entre 50 y 60 años		
6. Más de 60 años		

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Una vez determinadas las variables que harán parte del índice y las categorías según su incidencia tanto en la tasa de desempleo como en la probabilidad de estar ocupado, se procede a realizar un breve análisis descriptivo de las mismas. Para cada una de las variables se presenta un gráfico donde se distribuye la muestra para cada una de las opciones de la variable y se realiza un breve análisis sobre la distribución y la incidencia en las posibles barreras que enfrentan los buscadores en cada caso.

Gráfico 1. Ubicación de la fuerza laboral en Colombia.



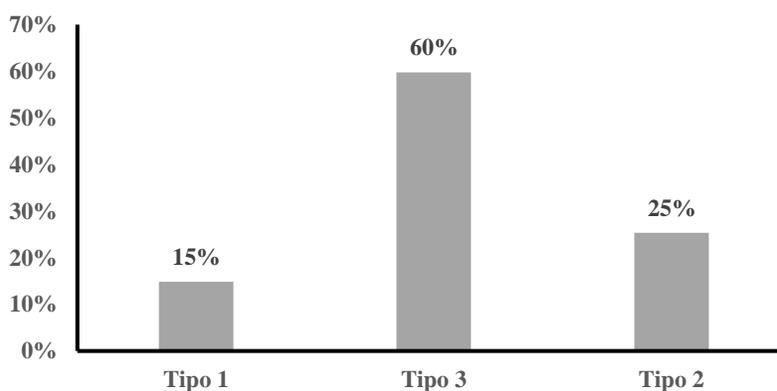
Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Colombia se caracteriza por ser un país altamente urbano, en donde 6 de cada 10 personas están trabajando o buscando un empleo en las ciudades, estas personas en su mayoría se dedican a actividades relacionadas con los servicios, industria y construcción. En segundo lugar se destaca la proporción de fuerza laboral ubicada en territorios rurales dispersos donde se desarrollan actividades relacionadas con el sector agropecuario y la minería, esta población se caracteriza por tener tasas de desempleo menores a los demás tipos de territorios, sin embargo enfrentan



vulnerabilidades altas derivadas de la informalidad y los bajos salarios. Finalmente, se encuentran las personas ubicadas en los centros poblados, los trabajadores de estas zonas realizan labores relacionadas con el comercio y los servicios personales de baja productividad y se caracterizan por tener las mayores tasas de desempleo relativas a los otros tipos de territorios con salarios bajos muy cercanos a los rurales dispersos.

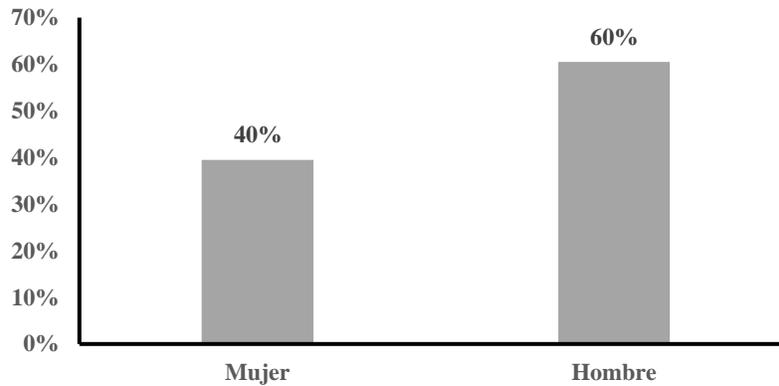
Gráfico 2. Distribución de la fuerza laboral según tipo de departamento.



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

A partir de la categorización realizada a los 23 departamentos del país, se encuentra que la mayoría de la fuerza laboral se encuentra en los departamentos tipo 3. Este hallazgo es coherente con la distribución del gráfico anterior dado que estos departamentos se caracterizan por tener los mercados laborales más grandes del país y las economías más dinámicas. En segundo lugar están los departamentos tipo dos, estos departamentos se caracterizan por tener mercados laborales medianos y menores tasas de desempleo relativas a los demás tipos. Finalmente, se encuentran los departamentos tipo 1, estos departamentos cuentan con los mercados laborales más pequeños del país y las mayores tasas de desempleo y por tanto es donde se encuentran las mayores barreras para que los buscadores puedan acceder a un empleo en condiciones de formalidad.

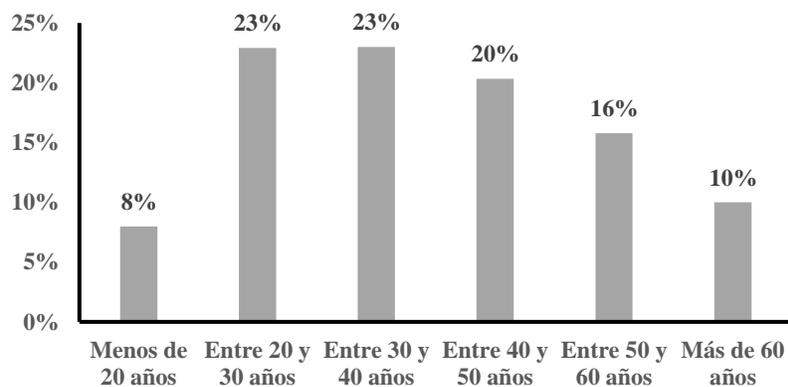
Gráfico 3. Distribución de la fuerza laboral según sexo.



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En el país aún persiste una brecha en el acceso al mercado laboral entre hombres y mujeres. En efecto, al analizar la información de la ECV 2016 se encuentra que 60 de cada 100 personas que se encuentran en el mercado laboral son hombres y tan solo 40 de cada 100 son mujeres. Adicionalmente, la tasa de desempleo y los salarios en la economía colombiana son desfavorables en mayor medida a las mujeres respecto a los hombres. Según diferentes estudios y análisis estas brechas surgen por barreras que enfrentan las mujeres relacionadas con el uso del tiempo, el cuidado de los niños y los ancianos, la masculinización de ciertas ocupaciones, los imaginarios culturales, entre otras razones, que generan mayores vulnerabilidades y dificultades tanto para acceder al mercado laboral como para colocarse en una vacante.

Gráfico 4. Distribución de la fuerza laboral según rangos de edad.

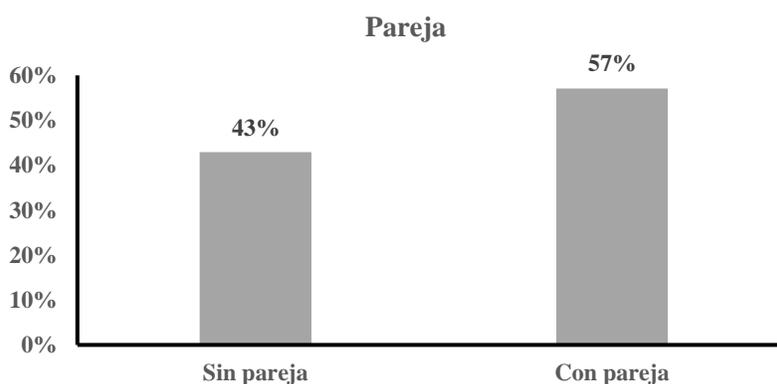


Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.



En cuanto a la de edad de las personas que se encuentran en el mercado laboral, es evidente que la mayoría de la fuerza laboral tiene edades entre los 20 y los 50 años. En efecto, 6 de cada 10 personas que están ocupadas o buscando un empleo en Colombia hacen parte de este grupo etario. En cuanto a las vulnerabilidades que enfrentan las personas según su edad se encuentran dos fenómenos muy marcados. En primer lugar se evidencia una relación inversa entre la edad y la tasa de desempleo, es decir que las personas más jóvenes tienden a ser más vulnerables al desempleo con períodos de duración cortos, este hecho se puede deber en gran medida a factores relacionados con la poca experiencia laboral o la falta de capacitación y formación para el trabajo que presentan los más jóvenes. Por el contrario, las personas de mayor edad tienden a tener menores tasas de desempleo con mayores períodos de duración, esto implica que las barreras que enfrentan con respecto a los jóvenes son totalmente diferentes y se enfocan en la falta de capacitación o reentrenamiento para ocupar una vacante, la destreza para usar tecnologías de la información, su capacidad de resiliencia o el desarrollo de algunas habilidades blandas.

Gráfico 5. Distribución de la fuerza laboral según si tienen pareja o no

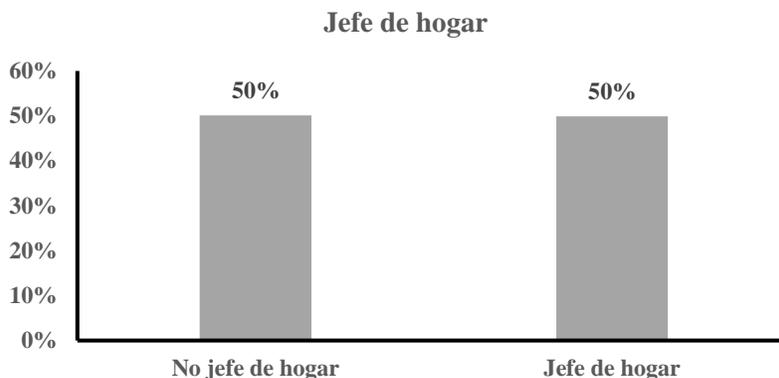


Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Con respecto a la distribución de las personas que tienen o no tienen pareja se encuentra que 6 de cada 10 personas que hacen parte de la muestra declararon tener pareja (unión libre, casado) y el 43% reportó no tener pareja (soltero, viudo o divorciado). En cuanto a la relación con la tasa de desempleo para cada uno de los grupos, se evidencia que las personas sin pareja tienen una tasa de

desempleo levemente superior a las que no tienen pareja, este resultado, en principio, se puede deber a que las personas solteras en su mayoría son jóvenes que presentan barreras para acceder a una vacante y adicionalmente, en muchos de los casos aún viven con sus familias les permite tener un salario de reserva más alto.

Gráfico 6. Distribución de la fuerza laboral según jefe de hogar

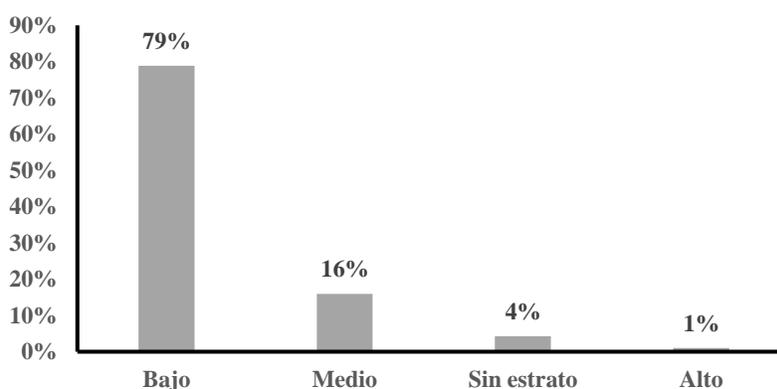


Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Por su parte, al analizar la distribución de los jefes de hogar, no se encuentra una diferencia significativa entre estos dos grupos. En efecto, la fuerza laboral se divide en partes iguales entre jefes y no jefes de hogar. Sin embargo, al analizar su relación con la tasa de desempleo se encuentra que el grupo de no jefes de hogar presenta mayores tasas de desempleo y esto se puede deber a que este grupo está compuesto en su mayoría por mujeres y jóvenes que presentan diversas barreras para acceder a un puesto de trabajo relacionadas con uso del tiempo, imaginarios culturales, masculinización de algunas ocupaciones y falta de experiencia, entre otras.

Índice de empleabilidad para Colombia

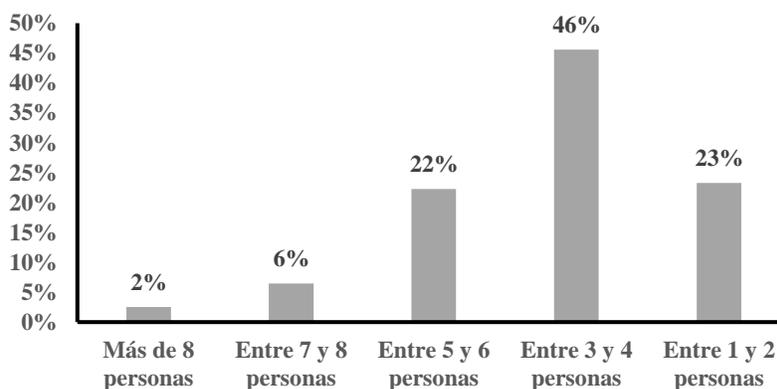
Gráfico 7. Distribución de la fuerza laboral según estrato socioeconómico



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Con respecto al estrato socioeconómico de la fuerza laboral colombiana, se encuentra que la mayoría hace parte de los estratos más vulnerables, cerca de 8 de cada 10 personas pertenecen a los estratos 1 y 2. Estas personas se caracterizan por presentar las mayores tasas de desempleo y una gran variedad de tipos de barreras que impiden su acceso a un empleo como el nivel educativo, el acceso a tics, los bajos niveles de ingreso per cápita, la ubicación alejada de los grandes centros de producción, entre otras. Por el contrario, el restante de la fuerza laboral (17%) se encuentra en los estratos medios y altos, estas personas presentan características totalmente diferentes al primer grupo con tasas de desempleo sustancialmente menores y barreras más relacionadas con la falta de acceso a información sobre vacantes o salarios de reserva más altos.

Gráfico 8. Distribución de la fuerza laboral según número de personas en el hogar

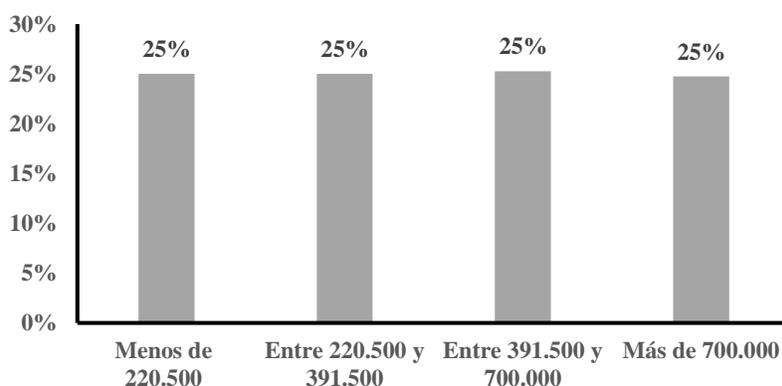


Fuente: Encuesta de Calidad de Vida. DANE. 2016.

Cálculos propios del autor.

En cuanto a la composición del hogar, se encuentra que la mayoría de la fuerza laboral tiene una composición del hogar inferior a 4 personas. En efecto, 7 de cada 10 personas hacen parte de este tipo de hogares, los cuales tienen menores tasas de desempleo y por tanto menores barreras para acceder a un empleo. Por el contrario los hogares con más de 5 personas se caracterizan por presentar muchas más barreras que se reflejan en tasas de desempleo más altas, estas barreras se relacionan, en la mayoría de casos, con el nivel educativo, pobreza, falta de acceso a bienes públicos, entre otros.

Gráfico 9. Distribución de la fuerza laboral según ingresos per cápita por rangos



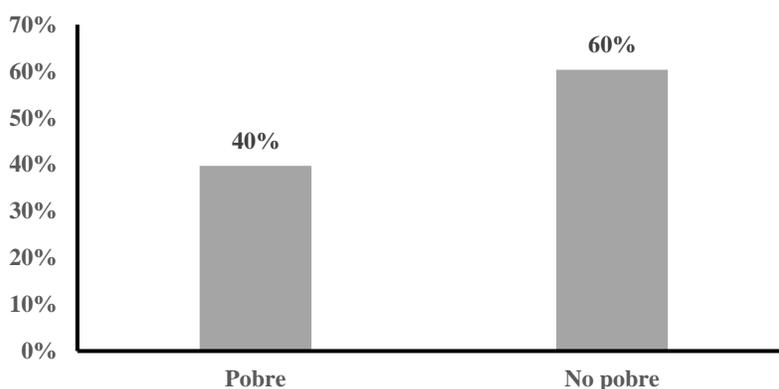
Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Como se mencionó anteriormente, la variable ingresos per cápita se construyó de tal forma que se crearon rangos de ingresos tal que se distribuyera la muestra de manera uniforme en los 4 cuartiles. Para este caso se encontró que existe una relación inversa marcada entre el nivel de ingresos per cápita y la tasa de desempleo. En efecto, las personas que presentan los menores ingresos per cápita tienden a presentar mayores niveles de vulnerabilidad en el mercado laboral, dificultando su colocación en una vacante. Estas barreras están asociadas a las descritas anteriormente en el caso del número de personas en el hogar, adicionalmente se identifican barreras relacionadas con la ubicación, tamaño del mercado laboral y poco acceso a canales formales de búsqueda de empleo.

Gráfico 10. Distribución de la fuerza laboral según pobreza



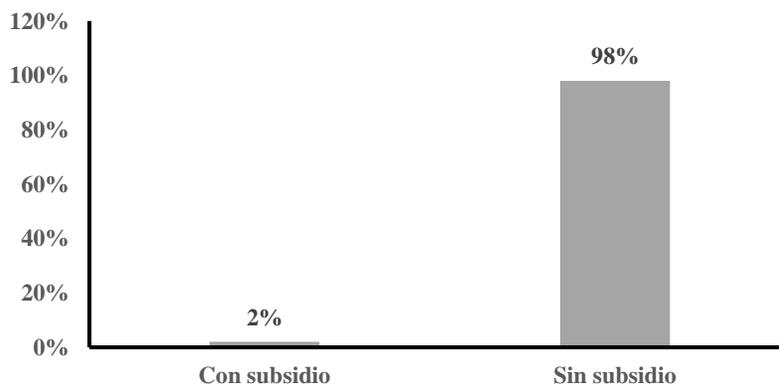
Índice de empleabilidad para Colombia



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

La variable de pobreza parte de un auto reconocimiento por parte de la persona que considera si se encuentra o no en situación de pobreza. El 40% de la fuerza laboral se auto reconoce en situación de pobreza mientras que 6 de cada 10 personas no se consideran pobres. Al analizar esta variable con relación a la tasa de desempleo, se encuentra que el grupo de personas pobres realmente presenta mayores dificultades para ocupar una vacante. En efecto, la tasa de desempleo de este grupo es superior a la que presenta el grupo no pobre. En este caso las dificultades en el acceso a un puesto de trabajo se enfocan en temas de educación y variables relacionadas con la vulnerabilidad del individuo y del hogar.

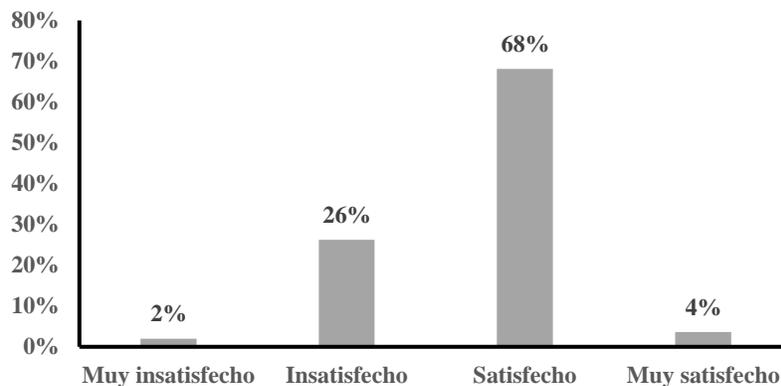
Gráfico 11. Distribución de la fuerza laboral según subsidio



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En cuanto al acceso a subsidios por parte de la fuerza laboral, tan solo un 2% de las personas que se encuentran en el mercado laboral accede a uno de los subsidios que se consultan en la ECV 2016. Al relacionar esta variable con la posibilidad de estar empleado, se encuentra que las personas con subsidios tienden a presentar tasas de desempleo más altas. Estas mayores tasas pueden ser explicadas por diferentes razones entre las que se destacan las mayores barreras que enfrentan estas personas o que su salario de reserva se incrementa por el subsidio recibido.

Gráfico 12. Distribución de la fuerza laboral según nivel de satisfacción en el trabajo



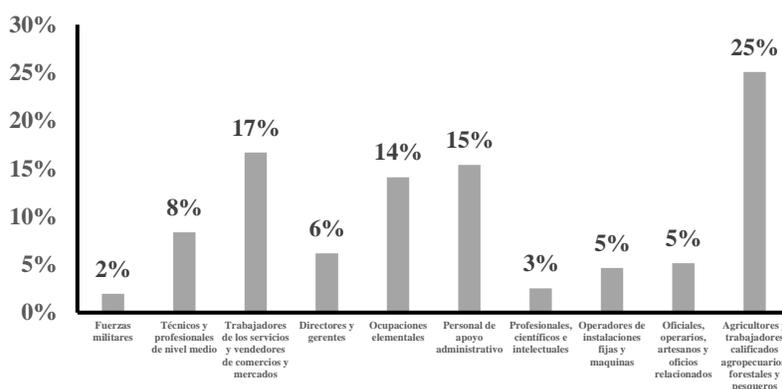
Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En cuanto al nivel de satisfacción en el trabajo, las personas que hacen parte de la fuerza laboral declaran en su mayoría que se encuentran satisfechas o muy satisfechas con las actividades que desarrollan (72%), por el contrario tan solo 3 de cada 10 personas declaran que están insatisfechas o muy insatisfechas en su trabajo. Al analizar la relación entre la satisfacción en el trabajo y la tasa de desempleo se evidencia que una relación inversa, es decir que entre más satisfecho está una persona en su trabajo, es menor la tasa de desempleo de ese grupo en particular.

Gráfico 13. Distribución de la fuerza laboral según ocupaciones



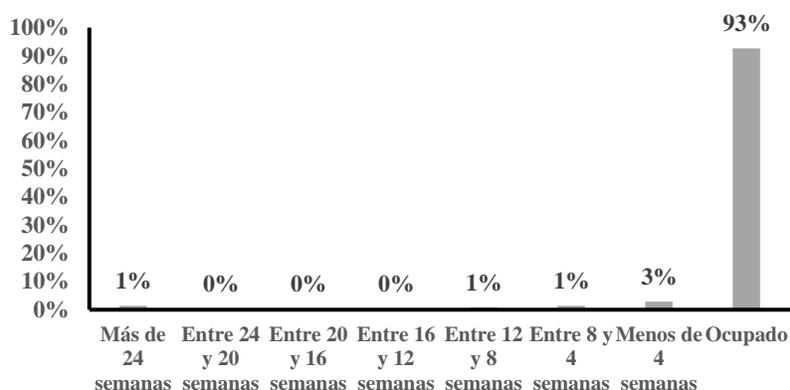
Índice de empleabilidad para Colombia



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En cuanto a las ocupaciones, se encuentra que 1 de cada 4 trabajadores en el país se desempeña en actividades relacionadas con la agricultura y trabajos calificados agropecuarios, forestales y pesqueros. Así mismo se destaca la alta proporción de personas que se dedican a las actividades relacionadas con los servicios como: los servicios y comercio (17%), personal de apoyo administrativo (15%) y ocupaciones elementales (14%). Con respecto a la relación entre las ocupaciones y la tasa de desempleo, se encuentra que las personas más vulnerables han estado vinculadas a ocupaciones relacionadas con las fuerzas militares, los técnicos profesionales del nivel medio y los trabajadores relacionados con los servicios y comercio. Estas personas enfrentan diferentes tipos de barreras relacionadas con la falta de capacitación para el trabajo, poco acceso a redes de información sobre ofertas laborales, entre otros. Por su parte, las ocupaciones relacionadas con el campo presentan una tasa de desempleo relativa menor a las demás ocupaciones, este hallazgo es consistente con el hecho que el campo presenta tasas de desempleo menores al promedio nacional, sin embargo, las tasas de formalización están muy por debajo a las demás ocupaciones y la brecha salarial es mayor con respecto a ocupaciones que se desarrollan en las ciudades.

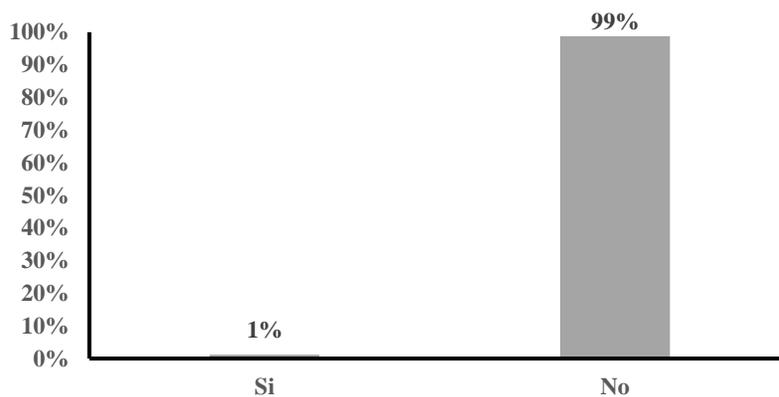
Gráfico 14. Distribución de la fuerza laboral según tiempo de búsqueda



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En cuanto al tiempo de búsqueda, al no poder establecer una relación entre la tasa de desempleo y el tiempo por la manera en que se construye esta variable, se determinó que a mayor tiempo de búsqueda de empleo, la persona se encontrará con mayores barreras derivadas de la desactualización de los conocimientos, la falta de redes para tener mayor información acerca de un empleo, y la desconexión con el mercado.

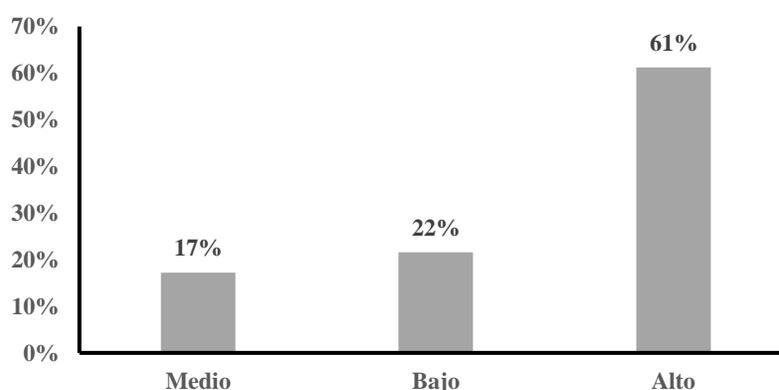
Gráfico 15. Distribución de la fuerza laboral según trabajo por primera vez



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Al igual que en el caso anterior, se considera que ser aspirante en el mercado laboral es una barrera que dificulta el acceso a un trabajo formal. Del total de la fuerza laboral, tan solo el 1% es aspirante y el restante 99% es cesante u ocupado. Las principales barreras que enfrentan las personas que intentan conseguir un trabajo por primera vez están relacionadas con la falta de experiencia laboral, no tener referencias laborales y la falta de habilidades blandas relacionadas con el trabajo en equipo, la autogestión, entre otros.

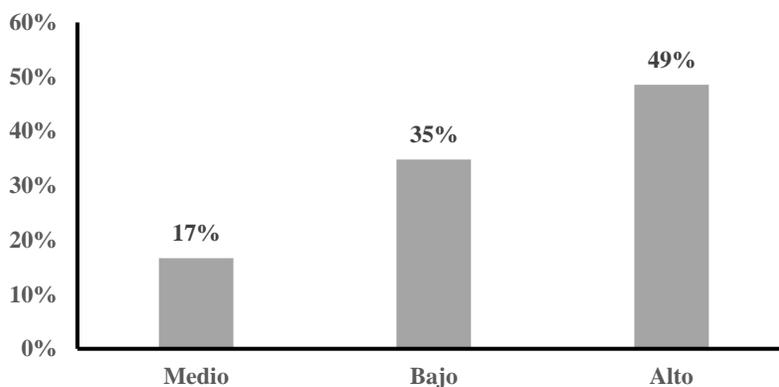
Gráfico 16. Distribución de la fuerza laboral según frecuencia de uso del computador



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Cada vez más, se hace necesario el uso de las tecnologías de la información para acceder a un puesto de trabajo. En efecto, en el mundo de hoy el uso de las tics es un requisito indispensable para desempeñar las funciones relacionadas con un trabajo. Al analizar la frecuencia de uso del computador, se encuentra que 6 de cada 10 personas que hacen parte del mercado laboral lo utilizan de manera frecuente, seguido de las personas con un uso bajo (22%) y el 17% lo usan de manera intermedia. En cuanto a la relación entre el uso del computador y la tasa de desempleo se destaca que las personas con una frecuencia de uso media o baja presentan mayores tasas de desempleo que la población que usa esta herramienta de manera frecuente, esta brecha se explica en gran medida porque el bajo uso indica barreras en el uso de las tics, lo cual implica dificultades al momento de colocarse en un empleo.

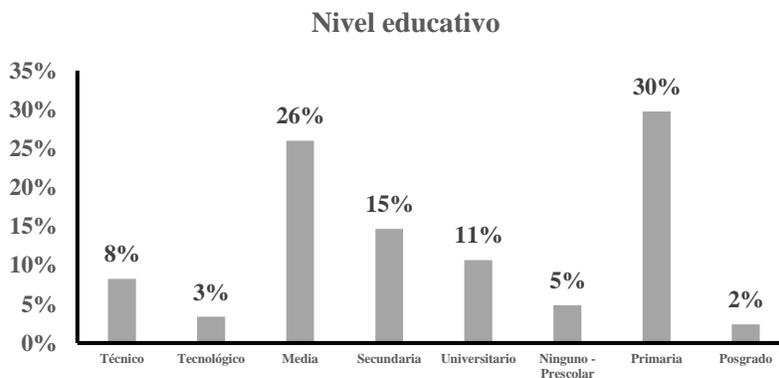
Gráfico 17. Distribución de la fuerza laboral según frecuencia de uso del internet.



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Al igual que en el caso del uso del computador, la frecuencia de uso del internet está relacionada con la destreza que tiene el buscador en el uso de las tics. En este caso, la mayoría de las personas, 5 de cada 10, hacen alto uso de internet, esta proporción se reduce con respecto al uso del computador en 10 puntos porcentuales. Por el contrario, el 50% de las personas encuestadas en la ECV 2016 declararon que usan el internet de manera media o baja, esta población, al igual que en el caso anterior, es la que presenta mayores vulnerabilidades al momento de buscar un empleo debido a las barreras que presentan en cuanto al uso de las tics.

Gráfico 18. Distribución de la fuerza laboral según el nivel educativo



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.



Finalmente, la última variable que se seleccionó para hacer parte del índice de empleabilidad es el nivel educativo. El nivel educativo es una de las características más importantes a la hora de buscar un trabajo, dado que la formación académica puede facilitar o dificultar el acceso a ciertas vacantes y claramente es una característica que el mercado de trabajo recompensa con mayores salarios.

A pesar de esto, en Colombia aún existe un alto porcentaje de personas que no han realizados estudios superiores. En efecto 8 de cada 10 personas no tiene más de educación media y de este grupo se destaca el alto porcentaje de personas que solo tienen educación primaria (30%) o educación media (26%). Por su parte, las personas que han accedido a educación superior tan solo alcanza el 21% del total de la fuerza laboral del país, de estos la mayoría son universitarios (11%), técnicos o tecnólogos (8%) y posgrado (2%).

En cuanto a la relación entre el nivel educativo y la tasa de desempleo se encuentran varios hallazgos importantes. En primer lugar, el desempleo no afecta directamente según el nivel educativo, este resultado está influenciado en parte por la alta informalidad de la economía colombiana y el ciclo económico que facilita el enganche de personas con bajos niveles educativos pero en puestos de trabajo que no presentan condiciones de empleo formal o autoempleo, sin afiliación al sistema de seguridad social y bajos niveles salariales, inclusive por debajo del salario mínimo legal. Este es el caso de las personas con niveles de educación primaria y preescolar o ninguno, que si bien tienen menores tasas de desempleo relativas a los demás niveles, presentan brechas educativas, formación en tics, formación para el trabajo y habilidades blandas que les dificulta acceder a un trabajo formal.

Por otro lado se encuentran las personas con niveles técnico y tecnológico de educación que presentan las mayores tasas de desempleo relativas a los demás niveles y que si bien tienen alguna formación superior, sus barreras se concentran en temas relacionados con la falta de experiencia y la especificación de las vacantes por parte de los empresarios que no tienen claridad entre las funciones que desempeña un técnico o tecnólogo y un profesional universitario.

Finalmente, se encuentran los niveles profesionales (universitario y posgrado) presentan diferentes dificultades para acceder a un trabajo. En el primer caso, los universitarios por lo general presentan barreras relacionadas con la insuficiente experiencia laboral o la deficiencia en competencias claves para desempeñarse en un trabajo. Por su parte las personas con niveles de posgrado

presentan la menor tasa de desempleo según el nivel educativo y las principales dificultades están relacionadas directamente con redes de contactos e información sobre vacantes disponibles.

Una vez determinadas las variables que se incluirán en el índice y descritas sus características y las posibles barreras relacionadas que enfrentan los buscadores, se procede a implementar la metodología de ACP. En este sentido, el primer paso consiste en analizar la matriz de correlaciones para revisar las posibles relaciones entre las variables seleccionadas. En esta matriz se puede observar las variables que tienen altos niveles de asociatividad y de esta forma empezar a tener una primera idea de las interrelaciones de las variables.



Tabla 6. Matriz de correlaciones de las variables seleccionadas.

	Ubicación	Estrato	Personas en el hogar	Ingresos por rangos	Pobre	Subsidio	Satisfacción en el Trabajo	Sexo	Edad por rangos	Jefe de hogar	Pareja	Educación	Ocupación	Semanas de Búsqueda	Trabajo por primera vez	Computador	Internet	Departamento	
Ubicación	1																		
Estrato	0.0252	1																	
Personas en el hogar	-0.0056	0.0711	1																
Ingresos por rangos	-0.1062	0.1807	0.2923	1															
Pobre	-0.0748	0.0939	0.0613	0.3013	1														
Subsidio	0.0146	0.0052	0.0332	0.0621	0.0554	1													
Satisfacción en el Trabajo	0.0286	0.0732	0.0863	0.2662	0.2506	0.079	1												
Sexo	0.0927	0.0005	-0.0058	-0.0476	-0.0477	0.0062	0.0088	1											
Edad por rangos	0.042	0.0356	0.2139	0.0578	-0.0407	0.0245	0.0256	0.0451	1										
Jefe de hogar	0.0299	0.0237	0.2908	-0.0135	-0.0073	0.0072	0.0284	0.2583	0.3869	1									
Pareja	0.0336	0.0021	-0.0266	-0.042	0.041	0.0023	0.0371	0.1131	0.2303	0.2146	1								
Educación	0.1744	0.0395	0.0479	-0.15	-0.1417	-0.0106	-0.0038	0.1227	0.3619	0.185	0.104	1							
Ocupación	0.3165	-0.0248	-0.0036	-0.2332	-0.149	0.0009	0.0053	0.31	0.1417	0.1318	0.0971	0.3182	1						
Semanas de Búsqueda	0.059	0.0115	0.0448	0.0835	0.0321	0.0286	0.1499	0.0695	0.0923	0.1014	0.0768	0.0923	0.1469	1					
Trabajo por primera vez	0.0262	0.0092	0.0465	0.0554	0.0107	0.0197	0.0491	0.0304	0.1315	0.0918	0.0857	0.0677	0.0819	0.3028	1				
Computador	0.1205	-0.0635	0.023	-0.2229	-0.1544	-0.005	-0.045	0.0893	0.2975	0.1564	0.1081	0.3539	0.2504	0.0765	0.0563	1			
Internet	0.1419	-0.0448	0.0302	-0.2155	-0.1697	-0.0028	-0.0406	0.1056	0.3544	0.1719	0.1112	0.4011	0.2722	0.0729	0.0516	0.6732	1		
Departamento	0.1679	-0.0344	-0.0126	-0.0283	-0.0292	-0.0011	0.0031	-0.0025	0.0204	-0.0107	0.0036	0.0469	0.068	0.0212	0.0127	0.0123	0.0186	1	

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

De la anterior matriz de correlaciones, se destacan las interrelaciones entre las variables computador e internet; educación e internet; edad y educación; edad e internet; educación y computador; ubicación y ocupación; sexo y ocupación; y pobreza e ingresos. Estos resultados ya van dando alguna idea de cómo se podrán ir conformando los componentes a partir de las relaciones evidenciadas en la matriz. El siguiente paso consiste en aplicar el test de adecuación muestral Kaiser-Meyer-Olkin (KMO) para las variables definitivas y de esta forma se podrá establecer el modelo de componentes principales en su conjunto es significativo o no.

Tabla 7. Resultados del test KMO.

Variable	kmo
Ubicación	0.7071
Estrato	0.6673
Personas en el hogar	0.5752
Ingresos por rangos	0.6542
Pobre	0.7586
Subsidio	0.7258
Satisfacción en el Trabajo	0.6651
Sexo	0.5946
Edad por rangos	0.7385
Jefe de hogar	0.6556
Pareja	0.7313
Educación	0.8442
Ocupación	0.7423
Semanas de Búsqueda	0.6411
Trabajo por primera vez	0.6345
Computador	0.714
Internet	0.7145
Departamento	0.5908
Overall	0.704

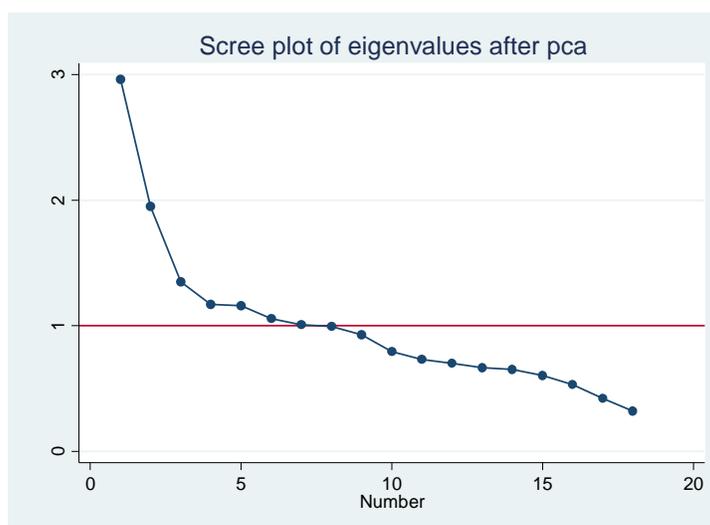
Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En efecto, para las variables que fueron seleccionadas se calcula el estadístico KMO y se encuentra que para todos los casos supera el 0,5; es decir que estas variables son adecuadas para el uso de la metodología de componentes principales, además el KMO total es superior al 0,7 lo que nos da un buen indicador sobre el uso conjunto de las mismas. Una vez se ha demostrado que todas las variables escogidas son pertinentes para realizar el ACP, se procede a estimar el modelo.



El tercer paso consiste en estimar el modelo de componentes principales y se procede a identificar el número de componentes a estimar. Para determinar el número de componentes se tomará como punto crítico los autovalores asociados a cada componente cuyo valor es mayor a uno, en el gráfico de sedimentación (ver gráfico 19) se identifica que el número de autovalores que cumplen con este criterio es de 7.

Gráfico 19. Gráfico de sedimentación con los autovalores asociados a los componentes estimados.



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Adicionalmente, se presenta la tabla de resumen de los autovalores de cada componente, con la respectiva proporción de la varianza total que explica cada uno y la varianza acumulada. A partir del criterio descrito anteriormente se determinó que el número de componentes es de 7, dado que son los autovalores mayores que uno. Estos componentes explican el 59% de la varianza total del conjunto de información.



Tabla 8. Componentes estimados, autovalores asociados y proporción de varianza explicada

Componente	Autovalor	Diferencia	Proporción	Acumulado
Comp1	2.9608	1.01071	0.1645	0.1645
Comp2	1.95009	0.599918	0.1083	0.2728
Comp3	1.35017	0.181246	0.075	0.3478
Comp4	1.16892	0.00901368	0.0649	0.4128
Comp5	1.15991	0.104822	0.0644	0.4772
Comp6	1.05509	0.0469567	0.0586	0.5358
Comp7	1.00813	0.0144522	0.056	0.5918
Comp8	0.993681	0.0664575	0.0552	0.647
Comp9	0.927223	0.133884	0.0515	0.6986
Comp10	0.79334	0.0605284	0.0441	0.7426
Comp11	0.732811	0.0315131	0.0407	0.7833
Comp12	0.701298	0.0349403	0.039	0.8223
Comp13	0.666358	0.0138425	0.037	0.8593
Comp14	0.652515	0.0486549	0.0363	0.8956
Comp15	0.60386	0.0714002	0.0335	0.9291
Comp16	0.53246	0.110298	0.0296	0.9587
Comp17	0.422162	0.100987	0.0235	0.9822
Comp18	0.321174	.	0.0178	1

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Una vez identificado el número de componentes y la proporción de la varianza que explican, se procede a determinar las variables que harán parte de cada componente y el ponderador en el cálculo final del índice de empleabilidad. Para esto se recurre a la matriz de cargas factoriales y se impone como criterio de selección que una variable tiene incidencia en un componente si su coeficiente de carga es superior a 0,25⁵. A continuación en la tabla 9, se presenta la matriz con las cargas factoriales que satisfacen esta condición:

⁵ Por lo general en la literatura se encuentra que el coeficiente debería ser superior a 0,3, sin embargo, con el objetivo de que todas las variables tengan incidencia en los componentes se optó por reducir el criterio a 0,25 lo cual no afecta la consistencia de los mismos.



Tabla 9. Matriz de cargas factoriales.

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Unexplained
Ubicación			0.4316		0.3987			0.4287
Estrato					0.3118			0.4759
Personas en el hogar		0.3859						0.3609
Ingresos por rangos		0.4806						0.3949
Pobre		0.3473				0.3697		0.4987
Subsidio						0.2947	0.3601	0.7095
Satisfacción en el Trabajo		0.3639				0.3434		0.4607
Sexo			0.2525					0.3485
Edad por rangos	0.3329	0.2633						0.3907
Jefe de hogar	0.264	0.2867						0.3313
Pareja						0.305	0.3467	0.4885
Educación	0.3804							0.5089
Ocupación	0.3381		0.3736					0.3886
Semanas de Búsqueda			0.3711					0.3607
Trabajo por primera vez			0.2511					0.3338
Computador	0.4075			0.3102				0.3013
Internet	0.429			0.2942				0.2639
Departamento			0.2819		0.2982		0.5427	0.3014

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Por tanto, a partir de la matriz de cargas factoriales los componentes quedan determinados de la siguiente manera:



Tabla 10 composición de los componentes seleccionados.

Componentes	Variables	Componentes	Variables
C1 (General)	1. Edad por rangos	C5 (Territorio)	1. Ubicación
	2. Jefe de hogar		2. Estrato
	3. Educación		3. Departamento
	4. Ocupación		
	5. Computador		
	6. Internet		
C2 (Características del hogar)	1. Personas en el hogar	C6 (Vulnerabilidad de la persona)	1. Pobre
	2. Ingresos por rangos		2. Subsidio
	3. Pobre		3. Satisfacción en el Trabajo
	4. Satisfacción en el Trabajo		4. Pareja
	5. Edad por rangos		
	6. Jefe de hogar		
C3 (Características laborales)	1. Ubicación	C7 (General 2)	1. Subsidio
	2. Sexo		2. Pareja
	3. Ocupación		3. Departamento
	4. Semanas de Búsqueda		
	5. Trabajo por primera vez		
	6. Departamento		
C4 (Tics)	1. Computador		
	2. Internet		

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Entre los siete componentes determinados se pueden identificar cinco que presentan asociaciones muy claras sobre las variables que recogen y están relacionados con temas del hogar C2, laboral C3, Tics C4 y vulnerabilidad de la persona C6. Así mismo, se estimaron dos componentes que no presentan una clara relación sobre barreras particulares que inciden en el buscador de empleo y son el C1 y el C7, y por tanto definen conceptualmente en términos de componentes que abarcan temas generales acerca del buscador.

Una vez establecidas las variables que harán parte de cada uno de los componentes se procede a determinar la proporción de la varianza que cada componente aporta, para esto se realiza una normalización de la varianza aportada por los primeros 7 componentes de forma tal que la suma sea igual a uno. Los resultados de la proporción de varianza explicada por cada componente se presentan a continuación.



Tabla 11. Porcentaje de varianza explicada y varianza explicada normalizada.

Componente	Varianza explicada	Varianza explicada Normalizada
C1	0.16	0.28
C2	0.11	0.18
C3	0.08	0.13
C4	0.06	0.11
C5	0.06	0.11
C6	0.06	0.10
C7	0.06	0.09

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Una vez se tienen tanto los siete componentes definidos como la proporción de la varianza que aportan (ponderadores) se procede a realizar la estimación del índice. Para la estimación del indicador se divide el puntaje que obtiene cada buscador en cada componente sobre el máximo puntaje posible, posteriormente se multiplica por su respectivo ponderador (proporción explicada de la varianza) y se suma para todos los componentes y de esta forma se obtiene el cálculo del índice. En la siguiente sección se presentan los principales resultados de la estimación del índice de empleabilidad para las 33.281 personas que componen la muestra.



Resultados

A partir del cálculo del índice para las más de 33 mil personas que componen la muestra se obtienen los siguientes descriptivos con su correspondiente distribución.

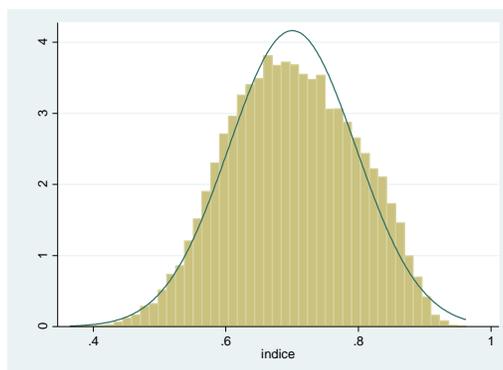
Tabla 12. Estadísticos descriptivos del índice de empleabilidad para el total de la muestra.

		Índice			
Percentiles		Smallest			
1%	0.48805	0.36495			
5%	0.54477	0.37204			
10%	0.57598	0.38604	Obs	33281	
25%	0.62987	0.39425	Sum of Wgt.	33281	
50%	0.6991		Mean	0.7003444	
		Largest	Std. Dev.	0.0958251	
75%	0.77245	0.9408			
90%	0.83038	0.94592	Variance	0.0091824	
95%	0.8569	0.95461	Skewness	-0.0453777	
99%	0.89501	0.96257	Kurtosis	2.422621	

Fuente: Encuesta de Calidad de Vida. DANE. 2016.

Cálculos propios del autor.

Gráfico 20. Gráfico de distribución del índice de empleabilidad para el total de la muestra.



Fuente: Encuesta de Calidad de Vida. DANE. 2016.

Cálculos propios del autor.

Como se presenta en la tabla 12 y gráfico 20, el promedio de la distribución del índice se encuentra en 70,03 puntos, con una desviación estándar de 9,5 puntos. Adicionalmente, se calcula el índice tanto para la población que se encontraba ocupada y desocupada al momento de la encuesta, al realizar esta distinción se encuentran diferencias marcadas tanto en los descriptivos analizados



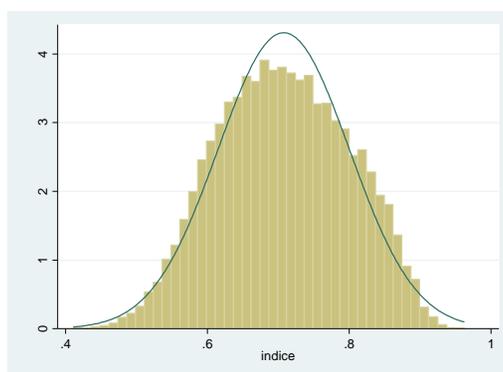
como en la función de distribución que conducen a pensar que efectivamente el índice permite identificar la posibilidad de emplearse relativa a las condiciones del buscador, a continuación se presentan los descriptivos y los histogramas para ambos grupos.

Tabla 13. Estadísticos descriptivos del índice de empleabilidad para el total de personas ocupadas

		Índice		
	Percentiles	Smallest		
1%	0.50793	0.41083		
5%	0.55811	0.42184		
10%	0.58681	0.42233	Obs	30851
25%	0.63852	0.42686	Sum of Wgt.	30851
50%	0.70648		Mean	0.7076789
		Largest	Std. Dev.	0.0925817
75%	0.77762	0.9408		
90%	0.83349	0.94592	Variance	0.0085714
95%	0.85903	0.95461	Skewness	-0.0191036
99%	0.89595	0.96257	Kurtosis	2.361791

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Gráfico 21. Gráfico de distribución del índice de empleabilidad para el total de personas ocupadas



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

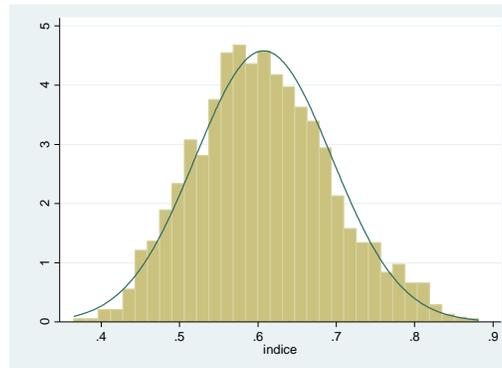


Tabla 14. Estadísticos descriptivos del índice de empleabilidad para el total de personas desempleadas

		Índice		
	Percentiles	Smallest		
1%	0.43319	0.36495		
5%	0.47132	0.37204		
10%	0.4969	0.38604	Obs	2430
25%	0.54694	0.39425	Sum of Wgt.	2430
50%	0.60276		Mean	0.6072257
		Largest	Std. Dev.	0.0871174
75%	0.66492	0.86132		
90%	0.72487	0.86204	Variance	0.0075894
95%	0.76467	0.86806	Skewness	0.2617507
99%	0.81661	0.8818	Kurtosis	2.774852

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Gráfico 22. Gráfico de distribución del índice de empleabilidad para el total de personas desempleadas



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En efecto, al analizar las distribuciones del índice para los subgrupos de ocupados y desempleados se encuentra que, para el caso de los ocupados, la distribución está corrida hacia la derecha, con un puntaje promedio de 0,70 puntos y una desviación estándar de 0,09 puntos. Por su parte, el grupo de los desocupados presenta una distribución corrida hacia la izquierda, con un puntaje promedio de 0,60 puntos y una desviación de 0,08 puntos.



Al comparar los dos grupos se encuentra que en promedio el índice de los ocupados supera al de los desocupados en 10 puntos y adicionalmente presenta una mayor dispersión de los datos, es decir el grupo de los ocupados es un grupo mucho más heterogéneo que los desocupados. Esta heterogeneidad en el grupo de ocupados puede deberse principalmente a que no se hace diferencia entre empleos formales e informales, lo cual puede generar amplias diferencias entre los trabajos que se realizan y las características de las personas que los ocupan.

Tabla 15. Test de medias para el índice entre el grupo de ocupados y desempleados.

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	2430	0.6072257	0.0017673	0.0871174	.6037602	.6106912
1	30851	0.7076789	0.0005271	0.0925817	.7066458	.708712
combined	33281	0.7003444	0.0005253	0.0958251	.6993148	.7013739
diff		-0.1004532	0.0019425		-0.2009064	

diff = mean(0) - mean(1) t = -51.7132

Ho: diff = 0 degrees of freedom = 33279

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0

Pr(T < t) = 0.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 1.0000

Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

Para establecer si existen diferencias significativas, se realizó un test de comparación de medias entre los grupos de ocupados y desocupados. El test da como resultado que si existe diferencia estadística significativa entre los dos grupos y en particular, las personas que se encuentran ocupadas, en promedio, presentan un mayor índice a las personas desocupadas y por tanto este último grupo presenta mayores barreras para conseguir un trabajo y por tanto se considerará más vulnerable.

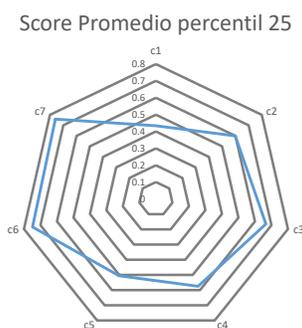
Finalmente, se presenta el análisis para cada uno de los niveles de empleabilidad identificados por los percentiles 25, 50, 75 y 100, lo cual permitirá clasificar a las personas en niveles de empleabilidad bajo, medio bajo, medio alto y alto,. Esta identificación es importante para la priorización de grupos vulnerables e implementación de políticas que permitan mitigar barreras comunes a los buscadores que hacen parte de estos grupos y que son priorizados por el índice.



Percentiles

Al analizar los percentiles, la población que se encuentra en el cuartil más bajo de la distribución obtiene puntuaciones entre los 36,4 y 62,9 puntos, es decir que todo buscador de empleo que obtenga un puntaje inferior a 63 puntos se considerará parte de la población con un grado de empleabilidad bajo y deberá acceder servicios que faciliten el cierre las brechas en temas como: primer empleo, educación, formación para el trabajo y tics con el fin de mejorar sus condiciones de empleabilidad y adicionalmente se le deberán facilitar servicios de gestión y colocación de empleo en el territorio que le permitan el acceso a un trabajo formal. (Ver gráfico 23).

Gráfico 23. Puntuación promedio para el percentil 25 de la distribución del índice.



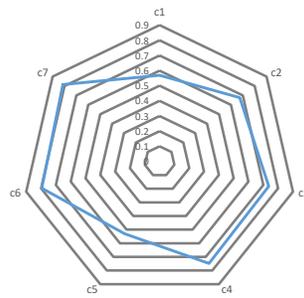
Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

En el segundo percentil se ubican las personas con puntajes entre 62,9 y 69,9 puntos, estas personas presentan un grado de empleabilidad medio bajo. Esta población, al igual que la anterior, se caracteriza por tener deficiencias en capacitación tanto en educación formal como formación en el uso de las tecnologías de la información. Adicionalmente, estas personas presentan una brecha significativa en el componente 2, el cual contiene variables relacionadas con vulnerabilidades que afectan a todo el hogar. Con el fin de reducir la brecha en el C2, se recomienda iniciar acciones que mitiguen las barreras relacionadas con el nivel de pobreza, la edad, los ingresos, el cuidado de las personas que componen el hogar y otras variables del entorno familiar. Estas acciones deberán ser implementadas de forma conjunta entre los gobiernos nacional y local de forma tal que facilite la articulación de la oferta de programas sociales.

Gráfico 24. Puntuación promedio para el percentil 50 de la distribución del índice.



Score Promedio percentil 50

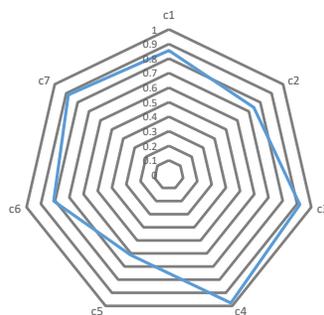


Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

El tercer grupo está compuesto por personas que obtienen puntuaciones entre 69,9 y 77,2 puntos. Estas personas se caracterizan por presentar en promedio barreras relacionadas con aspectos del territorio y por tanto es necesario realizar acciones que faciliten la conectividad a mercados laborales que sean accesibles desde su lugar de residencia y promuevan las conmutaciones en los territorios. Para esto es importante implementar acciones desde mejorar la movilidad entre los lugares donde se ubica el recurso humano y los sitios donde se concentra el tejido empresarial, hasta ferias de empleo y la promoción y gestión con las empresas para facilitar la colocación de estas personas.

Gráfico 25. Puntuación promedio para el percentil 75 de la distribución del índice.

Score Promedio percentil 75



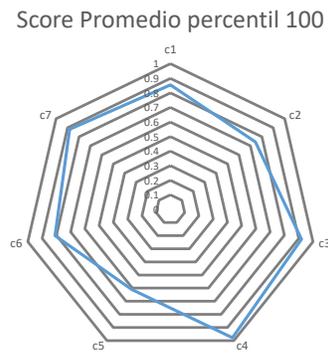
Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.

El último grupo identificado está compuesto en promedio por personas que tienen menores barreras para el acceso a un puesto de trabajo: menor tiempo de desempleo, mayor destreza y conocimientos en el uso de las tecnologías de la información, mayores niveles de educación formal



y competencias claves y transversales más desarrolladas. Las dificultades para estas personas surgen principalmente por la falta de información sobre vacantes disponibles y pocas redes de contactos que faciliten la búsqueda de un empleo, por tanto las acciones para facilitar su colocación deben ir dirigidas a proveer una oferta de servicios más enfocada en la gestión para la colocación junto con algún acompañamiento para mitigar vulnerabilidades relacionadas con el territorio.

Gráfico 26. Puntuación promedio para el percentil 100 de la distribución del índice.



Fuente: Encuesta de Calidad de Vida. DANE. 2016.
Cálculos propios del autor.



5. Conclusiones y Recomendaciones

Una de las principales dificultades en los mercados laborales modernos es la falta de herramientas, que permitan diagnosticar oportunamente las barreras que enfrentan los buscadores de empleo al momento de acceder a una oportunidad laboral. Esta dificultad incide en las bajas tasas de colocación que presentan las personas en condición de vulnerabilidad, dado que no se cuenta con la suficiente información que permita generar políticas y estrategias tanto por parte del sector público, los empresarios y los propios buscadores para mitigar las barreras de acceso.

Por esta razón, la presente investigación pretende entregar una herramienta que sea lo suficientemente robusta para identificar esas barreras que enfrentan las personas al momento de realizar un proceso de búsqueda laboral, de tal forma, que permita establecer el grado de empleabilidad de una persona en el mercado laboral. El concepto de empleabilidad se entiende como la posibilidad que tiene una persona para ocupar un trabajo a partir de su características como la capacidad, voluntad, conocimientos, habilidades, experiencia, etc y otras variables del contexto del buscador como el mercado laboral, el territorio y el tejido productivo que son susceptibles de modificarse a partir de acciones de política pública que permitan mitigar barreras y cerrar brechas.

Así mismo, el índice de empleabilidad permite determinar las barreras que más inciden en el proceso de enganche laboral de forma tal que se pueda diagnosticar de manera precisa las necesidades que presenta y los programas que le permitirán mitigar esas barreras y finalmente establecer grupos vulnerables que presenten brechas transversales que sean objeto de definición de nuevas políticas activas y pasivas de empleo.

Para la construcción del índice se identifica un conjunto de información lo suficientemente completo y representativo de la fuerza laboral colombiana y a partir de esto se construyen las variables que permitirán medir la incidencia de ciertas características en la empleabilidad de una persona. Estas variables se construyen a partir de la definición de opciones de respuesta con puntajes que son ordenados de acuerdo a dos criterios: la relación con la tasa de desempleo y la su incidencia y significancia en la probabilidad de estar ocupado, para este último se utilizan modelos probit que permitan medir la dirección del efecto y su nivel de significancia.

Una vez determinadas las variables que satisfacen los dos criterios establecidos, se identificaron 18 variables para la construcción del índice. Debido a la gran cantidad de información se recurre



al uso de la metodología de componentes principales con el fin de reducir la dimensionalidad de del conjunto de información y es de esta forma que se pasa de tener 18 variables a 7 componentes que recogen cerca del 60% de la variabilidad total de los datos. Estos siete componentes se definen como dos componentes de generalidades que contienen información relacionada en un caso con la edad, jefatura de hogar, nivel educativo, ocupación y uso de tics y en el otro caso el componente contiene información relacionada con acceso a subsidios, si tiene pareja y el departamento de residencia.

Los restantes 5 componentes contienen información sobre características del hogar (número de personas, ingreso per cápita, pobreza, jefe de hogar, edad y satisfacción en el trabajo); características laborales (ubicación, sexo, ocupación, semanas de búsqueda, trabajo por primera vez y departamento); tecnologías de la información (uso de computador e internet); territorio (ubicación, estrato y departamento) y vulnerabilidad de la persona (pobreza, subsidios, satisfacción en el trabajo y pareja).

Luego de establecer la composición de los constructos, se estiman los ponderadores que definirán la incidencia de cada componente dentro del cálculo del índice. En efecto, dado que la intensidad de cada una de los componentes en la medición del grado de empleabilidad es distinto, se debe tener un ponderador que asigne el peso correspondiente. Para definir los ponderadores se utiliza la varianza normalizada y se obtiene que los componentes, del primero al séptimo, tendrán ponderaciones de 0,28, 0,18, 0,13, 0,11, 0,11, 0,1 y 0,09 respectivamente.

Estos dos importantes resultados, componentes y ponderadores, son la base fundamental para calcular el índice de empleabilidad, dado que ya se conoce como se estructuran y su peso relativo dentro del cálculo. Ahora bien, el siguiente paso consiste en testarlo y para eso se replicó el cálculo a las más de 33 mil personas que componen la muestra de la fuerza laboral en Colombia y se obtuvo que en promedio la fuerza laboral presenta un índice de 0,70 puntos, con una desviación de 0,09 puntos. La persona más vulnerable de toda la fuerza laboral obtuvo un indicador de 0,36 puntos, por el contrario la persona que obtuvo el mejor puntaje fue de 0,96 puntos.

Luego de calcular el índice para toda la muestra, con el propósito de validar su robustez, se estimó el índice para diferentes grupos poblacionales y de esta forma determinar si se comporta de acuerdo a lo esperado para cada grupo. En este sentido se realizaron dos ejercicios, el primero consistió en comparar el indicador para la población ocupada y los desempleados y en el segundo se calculó el índice para cada uno de los cuartiles de la muestra total.



En el primer caso se encontraron diferencias significativas para los índices promedio de los ocupados y desempleados. En efecto, mientras la población ocupada en promedio alcanza un índice de 0,70 puntos, en el caso de los desempleados el indicador se reduce en 10 puntos y tan solo alcanza los 0,60 puntos, este resultado implica que en promedio la población que se encontraba desocupada al momento de responder la encuesta presenta mayores barreras que los ocupados. Así mismo, al comparar los extremos de la distribución se encuentra que el valor mínimo que alcanza el índice para los ocupados es de 0,41 puntos, mientras que en los desempleados la persona más vulnerable obtiene un valor de 0,36 puntos, es decir que entre las personas más vulnerables de ambos grupos existe una diferencia de más de 15 puntos. Estos resultados sugieren que el índice presenta consistencia al momento de ser replicado en diferentes poblaciones y en efecto permite medir el grado de empleabilidad que tiene cada buscador en el mercado laboral.

Ahora bien, en el segundo ejercicio se dividió la muestra de la fuerza laboral en cuartiles, clasificándolos en los siguientes niveles de empleabilidad: bajo, medio bajo, medio alto y alto y se replicó el ejercicio para un individuo promedio de cada grupo. Adicionalmente, para cada grupo se identificaron los componentes críticos en donde se presentan las mayores brechas y se diagnosticaron las barreras que se deben mitigar en cada caso, a partir del diagnóstico se propusieron algunas acciones para mejorar la empleabilidad en cada grupo.

El primer grupo identificado es el de personas con empleabilidad baja, este grupo es el que presenta menores niveles de empleabilidad y mayores brechas en cada componente. Las principales barreras se encuentran en temas relacionados con la falta de experiencia laboral, altos tiempos de búsqueda, bajos niveles de educación formal y no formal, analfabetismo tecnológico, pobreza, bajos ingresos, y ubicación en territorios alejados de los centros económicos y productivos del país (zonas rurales) y por tanto surge como recomendación para esta población iniciar acciones articuladas entre los gobiernos nacional y local, las entidades del sector trabajo y los sectores productivos y educativo con el fin de ampliar el acceso a educación formal y a programas de formación para el trabajo, ofrecer programas de primer empleo y facilitar alternativas para reducir las vulnerabilidades derivadas de los bajos niveles de ingreso y la alta pobreza que presentan estas personas. Así mismo se deben implementar estrategias que permitan mejorar la productividad del campo y promuevan los emprendimientos campesinos.



El segundo grupo identificado está compuesto por las personas que presentan niveles de empleabilidad medio bajo. Con respecto al anterior grupo, las personas con nivel medio bajo de empleabilidad presenta mejores condiciones en varios de los componentes analizados, sin embargo, presentan brechas significativas en los componentes relacionados con las características del hogar y características de localización o del territorio. Estas brechas se explican en gran parte por la existencia de barreras relacionadas con el número de personas en el hogar, ingresos y pobreza, edad, estrato, ubicación y departamento. En efecto, estas personas se caracterizan por vivir en territorios alejados de los grandes centros de producción (rural disperso y ciudades pequeñas), en donde la pobreza es alta y las familias están conformadas por una gran cantidad de personas. Para este grupo de personas se recomienda implementar acciones que reduzcan los niveles de pobreza y faciliten el acceso a trabajos en el territorio, para esto los gobiernos locales deben iniciar acciones que mejoren la infraestructura de los territorios con proyectos que contraten mano de obra local, promuevan la competitividad del territorio y mejoren la atracción de inversiones y empresas que reduzcan los niveles de desempleo.

En el tercer grupo se clasificaron a las personas con empleabilidad media alta, este grupo se caracteriza por estar por encima del puntaje promedio de toda la muestra y presenta una notable mejoría en términos de empleabilidad con respecto a los dos grupos anteriores. En efecto, este grupo presenta menores niveles de pobreza y mayores ingresos, mayores niveles de educación y más uso de tecnologías de la información. No obstante, su principal brecha está relacionada con la ubicación, dado que se encuentran en su mayoría en territorios con bajos niveles de productividad y altas tasas de desempleo y por tanto las principales recomendaciones para este grupo de personas están dirigidas a facilitar la conectividad con territorios que ofrezcan mayores oportunidades laborales promoviendo las conmutaciones, en los caso que es posible, entre los lugares de residencia y los territorios en donde se ubican las empresas.

El último grupo identificado está conformado por las personas que presentan niveles de empleabilidad altos. Estos buscadores presentan las menores brechas en comparación a los restantes grupos y se caracterizan por tener mayores facilidades para acceder a un empleo, son personas con altos niveles de formación, pocas vulnerabilidades del individuo, buenos niveles de ingreso y experiencia laboral. La principal barrera de este grupo se presenta en la poca información sobre oportunidades laborales. Y por tanto se recomienda fortalecer los servicios de gestión y colocación de empleo para que las personas con niveles de empleabilidad altos tengan un mayor



conocimiento de las oportunidades laborales que existen y de esta forma acceder de manera rápida y efectiva a un empleo.

Este segundo ejercicio permitió cumplir con los dos objetivos restantes de la investigación, dado que se comprobó que la herramienta permite diagnosticar barreras que inciden en la posibilidad de emplearse y de esta forma proponer acciones para mitigarlas. Adicionalmente, el índice facilitó la identificación de grupos vulnerables con barreras comunes que necesitan de políticas de empleo, sociales y económicas transversales para facilitar el cierre de brechas.

A partir de estos resultados, se verifica que el índice de empleabilidad para Colombia es una potente herramienta que permite caracterizar y diagnosticar a la población que se encuentra en medio de un proceso de búsqueda laboral, facilitando la información y conocimiento sobre el estado actual de las personas, cuáles son sus barreras, determinando su grado e intensidad para que a través de acciones puedan ser mitigadas. Así mismo, permite priorizar grupos vulnerables para los cuales es necesario implementar políticas que reduzcan sus brechas en el mercado laboral y facilitar su colocación.

Finalmente, es importante mencionar que el trabajo con el índice no concluye acá, esta investigación tan solo es el primer esfuerzo para mejorar la información en el mercado laboral colombiano. Sin embargo es necesario continuar con el proceso de implementación de la herramienta, aplicándola por medio de entrevistas a los buscadores de empleo y constatar los resultados obtenidos en este documento y en caso de ser necesario realizar los ajustes pertinentes a las necesidades identificadas. Así mismo, se espera que a partir de la información que se obtenga de la aplicación del instrumento se ajusten las políticas actuales e implementen nuevas estrategias que permitan mitigar las barreras que se identifiquen y faciliten la colocación de las personas en empleos formales.



6. Bibliografía

- Arango, L. (2011). Mercado de trabajo de Colombia: suma de partes heterogéneas. Bogotá, Colombia: Borradores De Economía, Banco de la República.
- Barón, J., (2010). Primeras experiencias laborales de los profesionales colombianos: probabilidad de empleo formal y salarios. Cartagena, Colombia: Documentos de trabajo sobre economía regional, Banco de la República.
- Clemente, A., (segundo trimestre de 2012). Análisis multivariado. Presentación realizada en el curso de análisis multivariado. Universidad Torcuato di Tella. Buenos Aires, Argentina.
- Departamento Administrativo Nacional de Estadística. (2016). Encuesta Nacional de Calidad de Vida 2016. Bogotá, Colombia: DANE.
- Department of Education, Employment and Workplace Relations. (2012). Job seeker classification instrument—factors and points. Canberra, Australia: Department of Education, Employment and Workplace Relations.
- Department of Employment. (1998). Job seeker classification instrument. Canberra, Australia: Commonwealth Department of Employment, Workplace Relations and Small Business.
- Forrier, A., Sels, L. (2003). The concept employability: a complex mosaic. Lovaina, Bélgica: Int. J. Human Resources Development and Management, Vol. 3, No. 2.
- Grip, A., Loo, J., Sanders, J. (2004) The Industry Employability Index: Taking account of supply and demand characteristics. Ginebra, Suiza: International Labour Review, Vol.143 (2004), N°3.
- Hair, J. (2000). Análisis Multivariante. Prentice Hall.
- Cuadras, C. (2014). Nuevos métodos de análisis multivariante. Barcelona, España: CMC Editions.
- Hofstetter, M. (2012). Otra mirada a datos laborales urbanos colombianos: Muchas preguntas y algunas respuestas tentativas. Bogotá, Colombia: Borradores De Economía, Banco de la República.
- Lebart, L., Morineau, A., Piron, M. (1995). Statistique exploratoire multidimensionnelle. París. Francia: Dunod Editions.
- Leibovich, J., Nigrinis, M., Ramos, M. (2006). Caracterización del mercado laboral rural en Colombia. Bogotá, Colombia: Borradores De Economía, Banco de la República.



- Ortiz, J. (2004). Descripción multivariada de datos. Componentes principales. Bogotá, Colombia: Universidad Nacional de Colombia.
- Rencher, A. (2007). Methods of multivariate statistical analysis, 2 ed. Wiley-Interscience.
- Rodríguez, J. (2009). Índice de empleabilidad de los jóvenes. Lima, Perú: Serie Avances de Investigación n° 32 Centro de Investigaciones Sociológicas, Económicas, Políticas y Antropológicas (CISEPA).
- Timm, N. (2002): Applied multivariate analysis. Pittsburgh, USA: Springer.
- Weller, J. (2007). La inserción laboral de los jóvenes: características, tensiones y desafíos. Santiago de Chile, Chile: CEPAL.
- Weller, J. (2009). El fomento de la inserción laboral de grupos vulnerables. Consideraciones a partir de cinco estudios de caso nacionales. Santiago de Chile, Chile: CEPAL.
- Weller, J. (2012). Vulnerabilidad, exclusión y calidad del empleo: Una perspectiva latinoamericana. Santiago de Chile, Chile: CEPAL.
- Wooldridge, J., (2002). Econometric Analysis of Cross Section and Panel Data. Londres, Inglaterra: The MIT Press.

