

Juan Carlos Segura Ortiz*

República de Colombia, Unidad de Gestión
de Pensiones y Parafiscales
Ministerio de Hacienda y Crédito Público,
Bogotá, D. C., Colombia

Recibido: 15 de diciembre de 2015

Concepto de evaluación: 11 de abril de 2016

Aprobado: 27 de mayo de 2016

Artículo de investigación

© 2016 Universidad Católica de Colombia.

Facultad de Ciencias

Económicas y Administrativas.

Todos los derechos reservados

Aportes para la medición del impacto de la política de formación para el trabajo: propuesta para la evaluación del contrato de aprendizaje en Colombia¹

RESUMEN

El contrato de aprendizaje (CA) es una estrategia de formación para el trabajo que combina adiestramiento en aspectos vocacionales con el desarrollo de una fase práctica en empresas. Para este modelo dual de formación se reclaman ventajas representadas en mejores perspectivas de ocupabilidad y de ingresos para el trabajador, y menores riesgos de selección en las firmas. Dado el propósito formal que para este instrumento de formación se fija en Colombia, interesa evaluar sus resultados, con el objetivo de mejorar y corregir las condiciones actuales de desempeño del instrumento y de optimizar su impacto efectivo. Así, se ofrece una propuesta de evaluación del impacto del CA a partir de la medición de diferencias en niveles de ingresos de individuos que, habiendo optado por formación técnica vocacional, eligieron este contrato como alternativa en la fase práctica del proceso. El texto también ofrece una revisión de literatura relevante sobre CA, una descripción del tipo de modelo estadístico que se ha de utilizar en la evaluación y una descripción de los datos necesarios para adelantar esta empresa.

Palabras clave: medición de impacto, formación para el trabajo, contrato de aprendizaje, diferencia de ingresos, desempeño laboral.

JEL: H75, I26, I28

* Economista. Join Masters Programme Environmental and Natural Resource Economics. Actualmente es Asesor de Estrategia y Evaluación de la Dirección General de la Unidad de Gestión de Pensiones y Parafiscales (UGPP), Ministerio de Hacienda y Crédito Público. Correo electrónico: jsegura@ugpp.gov.co / j.c.segura@gmail.com.
Dirección de correspondencia:
Avenida Calle 26 N° 69B-45, piso 2,
Bogotá D. C., Colombia.

1 Investigación financiada a través del contrato CPS 134 de 2013, del Ministerio de Trabajo. Las opiniones que se expresan aquí son responsabilidad del autor y no involucran al Ministerio del Trabajo ni a ninguna de sus dependencias.

Contributions for the measurement of the impact of the training for work policy: A proposal for the assessment of the apprenticeship contract in Colombia

ABSTRACT

The apprenticeship contract (AC) is a training for work strategy that combines training in vocational aspects with a practical phase in a company. This dual training model presents advantages in terms of the improvement of workers' employability perspectives and income, together with a reduced employee selection risk. Given the formal purpose of this training instrument in Colombia, there is an interest in evaluating its results in order to improve and correct the instrument's current performance and to optimize its effective impact. Here, the authors offer a proposal for the assessment of the impact of ACs, based on the measurement of differences in levels of income of individuals who, having chosen technical vocational training, chose this contract as an alternative in the process' practical phase. The text also includes a review of the relevant literature on AC, a description of the type of statistical model that has to be used in the evaluation, and a description of the data necessary to help the company advance.

Keywords: impact assessment, training for work, apprenticeship contract, income gap, job performance

Contribuições para a medição do impacto da política de formação para o trabalho: proposta para a avaliação do contrato de aprendizagem na Colômbia

RESUMO

O contrato de aprendizagem (CA) é uma estratégia de formação para o trabalho que combina o ensino em aspectos vocacionais com o desenvolvimento de uma fase prática em empresas. Para esse modelo dual de formação, reivindicam-se vantagens representadas em melhores percepções de ocupabilidade e de ingressos para o trabalhador, e menores riscos de seleção nas firmas. Tendo em vista o propósito formal que para esse instrumento de formação se estabelece na Colômbia, interessa avaliar seus resultados com o objetivo de melhorar as condições atuais de desempenho do instrumento e corrigi-las bem como de otimizar seu impacto efetivo. Assim, oferece-se uma proposta de avaliação do impacto do CA a partir da medição de diferenças em níveis de ingresso de indivíduos que, tendo optado pela formação técnica vocacional, escolheram esse contrato como alternativa na fase prática do processo. Este texto também oferece uma revisão da literatura relevante sobre o CA, uma descrição do tipo de modelo estatístico que deverá ser utilizado na avaliação e uma descrição dos dados necessários para realizar essa empresa.

Palavras-chave: medição de impacto, formação para o trabalho, contrato de aprendizagem, diferenças de ingressos, desempenho profissional.

INTRODUCCIÓN

El contrato de aprendizaje (CA) en Colombia es una estrategia de formación para el trabajo que combina adiestramiento en aspectos vocacionales y teóricos, impartidos normalmente en aulas y espacios de aprendizaje dispuestos para tal efecto, con el desarrollo de una fase práctica que ha de tener lugar en empresas que, al acoger a los aprendices, adoptan un rol formador en el proceso. Se atribuye para este modelo dual de formación una serie de ventajas que, en el caso del aprendiz, está representada en mejores perspectivas de ocupabilidad y de ingresos, en tanto para las firmas las ventajas incluyen procesos de selección de personal y de *screening* que pueden reducir el riesgo de selección adversa, mejorar la productividad general, entre otros aspectos. En el agregado, según suele reclamarse en la literatura relacionada, el CA contribuye a reducir el desempleo que se suscita cuando las demandas por competencias del empleador no coinciden con las características de la oferta en el mercado laboral (McIntosh, 2007).

En Colombia, el alcance y la naturaleza del CA corresponden a lo prescrito en la Ley 789 de 2002. Dado el propósito formal que se fija para este instrumento de formación, interesa evaluar sus resultados con el objetivo típico de corregir y mejorar las condiciones actuales de desempeño del instrumento y optimizar su impacto efectivo. Este artículo contiene una revisión de la literatura reciente y relevante sobre relaciones de aprendizaje o *apprenticeships* —que es la categoría general a la que el CA corresponde— y sobre la experiencia en relación con la evaluación y medición de impacto de este tipo de programas. Dicha revisión da lugar a una propuesta metodológica para la valoración del impacto de la estrategia y facilitar su seguimiento, y a una indagación de los datos necesarios y disponibles para tal propósito.

En la siguiente sección se introduce el concepto de *contrato de aprendizaje* según es proposición de la Organización Internacional del Trabajo (OIT), se describen los atributos que *idealmente* estarían presentes en esta suerte de relaciones de

aprendizaje, así como las oportunidades que se espera se deriven de su aplicación. Esta información es contrastada con las disposiciones que, para el caso colombiano, operan de acuerdo con la Ley 789 de 2002 y con sus desarrollos normativos vigentes.

Como se quiere medir el impacto del CA, parece conveniente identificar a) las variables instrumentales del operador de la política o de sus agentes (i. e., los tratamientos) y b) las variables objetivo o de resultado del tratamiento (Bernal y Peña, 2011; Gertler *et al.*, 2011; Khander, Koolwal y Samad, 2010; D. Lee, 2005). Las variables objetivo y los tratamientos han de aparecer relacionadas dentro del sistema de toma de decisiones relevante (Munk, 2005), tanto en lo que se refiere al postulado normativo de la OIT como en el caso específico colombiano; por tanto, vale contrastar empíricamente la hipótesis de causalidad entre unas y otras variables.

En la medición del impacto de los CA, la mayor parte de los esfuerzos combina postulados de motivación de la política con contenidos teóricos bien definidos dentro de la economía laboral (Cahuc y Zylberberg, 2004; Laing, 2011; Borjas, 2012). De cualquier modo, estos contenidos suelen referirse a los aportes clásicos habituales (Becker, 1964; Mincer, 1975) sobre acumulación de capital humano, principalmente. En general, la literatura sobre evaluación del impacto de los CA suele concentrarse en la estimación econométrica de las ecuaciones de Mincer, pues trata de indagar los efectos del entrenamiento sobre los ingresos del aprendiz y sobre los resultados de largo plazo en términos de ocupabilidad y movilidad laboral. La evaluación de los resultados de las relaciones de aprendizaje sobre las economías privadas de las firmas son de menor producción, en parte porque los datos sobre costo-beneficio necesarios son difícilmente disponibles, visto que las empresas son renuentes a difundirlas.

En la sección “Datos” se consolida una propuesta de evaluación elaborada con base en la revisión de la sección previa. El alcance del proceso de modelamiento supone múltiples estrategias que pueden resultar deseable utilizar y que han

sido utilizadas de facto en algunas investigaciones como, por ejemplo, el trabajo precursor de Jiménez y Kugler (1987), que se concentra en el la medición del efecto de la duración del proceso de formación en los ingresos de los aprendices; o aquel más reciente de Estacio *et al.* (2010), que evalúa el impacto de dos programas de formación del Servicio Nacional de Aprendizaje (SENA) a través de estrategias computacionales como la de diferencias en diferencias (DD) o la de apareamiento.

Todo esfuerzo de valoración cuantitativa está condicionado por la información disponible y por la calidad de los datos. En relación con los CA, la principal fuente de información es el SENNA, que registra datos sobre individuos en formación para el trabajo a su cargo en el Sistema Optimizado para la Formación Integral del Aprendizaje (SOFÍA), que viene operando desde 2010. Las características de dicho sistema garantizan el registro de un número amplio de individuos que se han vinculado a procesos de formación para el trabajo, entre los cuales se encuentran precisamente aquellos que han optado por el CA como una alternativa para el desarrollo de la etapa productiva de los programas de formación a los que se encuentran vinculados¹.

La oportunidad de disponer de información sobre registro de aprendices y de pagos a la seguridad social tiene alcance limitado porque, aparte de la información aproximada sobre el ingreso y quizás la continuidad de la ocupabilidad —en cualquier caso difusa y con error—, la Planilla Integrada de Liquidación de Aportes (PILA) y el Registro Único de Afiliados (RUAF) tampoco incluyen datos sobre características socioeconómicas que puedan ser utilizadas como controles para disminuir las consecuencias adversas del sesgo de (auto)selección. En efecto, información relativa a la experiencia y otras características individuales,

1 La etapa productiva de los procesos de formación de los aprendices del SENNA —que es el ámbito en el que los individuos transfieren los conocimientos teóricos a contextos productivos reales— es susceptible de ser desarrollada según distintas modalidades: a) contrato de aprendizaje, b) vinculación laboral o contractual, c) participación en proyectos productivos (en SENNA-Empresa; SENNA proveedor SENNA; en producción de centros), d) asesoría a mipymes y e) monitorias (SENA, Resolución 02367 de 2010, artículo 1).

que comúnmente forman parte de los esquemas de estimación de ingresos à là Mincer, no aparece registrada explícitamente allí. Si bien la indagación de fuentes alternativas con dichas características es obligada, puede esperarse poco éxito en este propósito. La sección “Datos” se ocupa de examinar los requerimientos de datos y las oportunidades que sobre este particular se presentan para el caso de esta investigación.

EL CONTRATO DE APRENDIZAJE

El CA en Colombia se identifica con la noción que sobre relaciones de aprendizaje reguladas adopta la OIT² (2012):

Con el término “contrato de aprendizaje” se denota a un conjunto de programas de formación y entrenamiento que combinan educación vocacional con aprendizajes basados en el lugar de trabajo, en alguna competencia ocupacional intermedia (i. e., más allá de los trabajos rutinizados), y que está sujeto a estándares de formación impuestos externamente, en particular al componente que tiene lugar en las empresas [traducción propia] (p. 3).

Según esta definición, los contratos de aprendizaje involucran a) educación y entrenamiento fuera del lugar del trabajo y b) regulación externa de los estándares de entrenamiento tanto en la parte externa como interna del lugar de trabajo. En la noción aportada por la OIT (2012), los CA *deben observar*, en adición, una serie de atributos concretos:

- a) Suponen un apoyo económico o una retribución salarial.
- b) Tienen un marco legislativo para su gestión y regulación.
- c) Se desarrollan de acuerdo con un *programa o currículo*.

2 La definición adoptada por OIT corresponde a aquella de Ryan *et al.* (2010).

- d) Requieren entrenamiento *en* el lugar de trabajo.
- e) Requieren formación *fuera* del lugar de trabajo.
- f) Deben ser sometidos a procesos formales de evaluación.
- g) Dan lugar a una certificación socialmente reconocida.
- h) Tienen duración fija³.

En Colombia, los CA toman forma con las previsiones de la Ley 789 de 2002 y de sus desarrollos normativos. Según ese conjunto de disposiciones, estas relaciones de aprendizaje pueden expresarse así:

Una persona natural desarrolla formación teórico-práctica en una entidad autorizada, a cambio de que una empresa patrocinadora proporcione los medios para adquirir formación profesional metódica y completa requerida en el oficio, actividad u ocupación y esto le implique desempeñarse dentro del manejo administrativo, operativo comercial o financiero propios del giro ordinario de las actividades de la empresa por cualquier tiempo determinado no superior a dos (2) años y por esto reciba un apoyo de sostenimiento mensual, el cual en ningún caso constituye salario (Ley 789 de 2002, artículo 30).

Desde la norma colombiana vigente, los CA son una relación entre un aprendiz y una empresa patrocinadora en cuya planta tiene lugar una fase práctica que complementa la fase lectiva del proceso, desarrollada previamente en aulas e instalaciones provistas por centros de formación reconocidos como el SENA. En la Ley 789, el objeto del contrato es facilitar la formación del aprendiz en una serie de ocupaciones de interés, en un ambiente en el que se observa una relación de subordinación

que se referirá únicamente a actividades del aprendizaje.

Hay un contexto socioeconómico específico para el CA⁴: a) los individuos más jóvenes llevan la mayor parte de la carga del desempleo en Colombia; b) la situación de los jóvenes tiene explicación, así sea parcial, en la falta de experiencia y en la carencia de competencias y habilidades relativas al trabajo; c) el desempleo es tanto más alto cuanto menor es el grado de educación alcanzado; d) la desocupación es mayor en el caso de quienes no asisten a algún nivel educativo, que es, justamente, el grupo de los más jóvenes entre la población.

Dada la promesa subyacente a los CA y las aproximaciones metodológicas para develar sus aportes socioeconómicos, es posible concebir que tanto aprendices como firmas puedan considerar el *input* trabajo como una especie de mercancía compuesta *à la* Rosen, cuyo precio viene definido por los atributos que integran el *paquete* (Rosen, 1974). Respecto a los beneficios que precisamente se atribuyen a los CA vale citar, entre otros, los siguientes:

- a) Los CA superan el enfoque de “aprender para las ocupaciones” y los sustituyen por el de “aprender a través de las ocupaciones” (Ryan, 2011).
- b) Las habilidades y competencias adquiridas son reales, porque surgen de un contacto concreto con los procesos fabriles en las empresas. Son habilidades y competencias productivas.
- c) Como las competencias son reales y aportan a la producción, la transición del estudio al trabajo deberá ser suave y sin vicisitudes.
- d) Estudios recientes relacionados con la formación técnica y tecnológica del SENA indican que las perspectivas en cuanto a ingresos, estabilidad y seguridad social son mejores para aquellos que optan por estos tipos de formación.

³ En el caso de los programas de formación titulada del SENA, la duración máxima estimada (DME) es de 12 meses para programas del nivel técnico, y de 24 meses para programas del nivel tecnológico.

⁴ Con base en observaciones de la Gran Encuesta Integrada de Hogares (GEIH), desarrollada por el DANE (2003).

- e) Algunas investigaciones son enfáticas al afirmar que, en la medida en que los aprendices se involucran en los procesos productivos de las firmas donde desarrollan sus prácticas, el valor de las contribuciones productivas puede superar la inversión de las empresas en entrenamiento (Wolter, 2012).
- f) Las empresas pueden usar los CA como un método para detectar a los aprendices más dinámicos, talentosos y motivados. Luego de un proceso de aprendizaje exitoso en la firma, es probable que esta contrate a los aprendices por un término mayor que el del CA (Wolter, 2012).

Las firmas tienen la necesidad y el interés de entrenar aprendices en competencias específicas a sus procesos fabriles, por lo cual transfieren valor a los aprendices; un valor que se transforma en mejores condiciones laborales, ingreso y estabilidad incluidos.

REVISIÓN DE LITERATURA

El impacto de los CA tiene por lo menos tres esferas de manifestación: a) la economía individual de los aprendices, b) los resultados operativos de las firmas y c) la economía como un todo. En general, los estudios de impacto son contrastaciones empíricas de hipótesis que parten de postulados sobre el comportamiento microeconómico de agentes individuales; pueden representarse analíticamente como unidades de decisión que optimizan funciones objetivo —de utilidad en el caso de los productores, de beneficio en el caso de los productores (Ginsburgh y Keyzer, 1997)—, mediante la elección racional de planes de consumo o de producción en sus conjuntos de elección.

En las economías modernas, los niveles agregados de capital humano están asociados a los niveles de desarrollo; la acumulación acelerada de capital humano es condición indispensable para el crecimiento moderno, y este, bajo

ciertas circunstancias, es un *trigger* del desarrollo⁵. El conjunto de conocimientos que sirven a un individuo para interactuar con otros, modificar en su favor los estados de la naturaleza y transformar unas mercancías en otras —esto es, el conjunto de competencias de que está dotado— constituyen para él un activo del que puede disponer para garantizar y mejorar sus condiciones de vida. En el mundo actual, los vehículos para la adquisición de dichas competencias son la formación (general y específica) y la práctica.

El aprendizaje en el lugar de trabajo es una de las principales fuentes de acumulación de capital humano (Becker, 1964; Mincer, 1975; Ashenfelter, 1978; Acemoglu y Pishke, 1998). Más aún, las variaciones de los salarios individuales y el incremento de los perfiles salariales pueden ser explicados por diferencias en el capital humano y por las habilidades inducidas en la experiencia y la formación vocacional continuada (Kuckulenz y Zwick, 2003). En Heckman (2000, citado en Kuckulenz y Zwick, 2003, p. 1), la capacitación posterior a la entrada al mercado laboral puede ser considerada como una de las mayores inversiones en capital humano.

Según la visión neoclásica, la financiación de estas inversiones deberá estar a cargo del trabajador, cuando supone para este una mayor productividad individual en un amplio rango de trabajos y ocupaciones en el mercado laboral que, por ende, se tornen en mayores salarios; en contraste, la formación deberá estar financiada por la firma, si esta se refiere únicamente a procesos específicos de la empresa que no le permiten al trabajador obtener rentas de dicha formación en ningún otro mercado (Acemoglu y Pishke, 1998, p. 1).

Las competencias y habilidades que un trabajador ofrece en los mercados actuales pueden ser tenidas como un bien compuesto de competencias

5 Para una discusión de economía política sobre cuando el crecimiento no se torna en desarrollo, cuando este es agenciado por instituciones extractivas, véase el trabajo de Acemoglu y Robinson (2013). En general, el crecimiento no tiene por qué tornarse en desarrollo, según así se postula en las caducas teorías de *trickledown*; no obstante, en ausencia de crecimiento o en presencia de crecimiento económico débil, la búsqueda de desarrollo se complica.

generales y competencias específicas, y el salario estará constituido por los precios implícitos de los atributos que caracterizan al trabajador. El precio (hedónico) del trabajo se fija en el mercado relevante y supone la compatibilidad de las disposiciones a pagar del empresario (DAP) y a recibir del trabajador (DAR). Contrario a las prescripciones competitivas, la renta de la formación es capturada por el trabajador o por el productor, o puede ser repartida entre los dos, dependiendo del poder de negociación de las partes, según se propone en Kuckulenz y Zwick (2003). Se deriva de aquí que si las rentas de la formación son capitalizadas por el trabajador, el empresario no tendrá incentivo para invertir, independientemente de si las competencias son generales o específicas a la firma. En contraste, si las rentas han de ser apropiadas por la firma, el trabajador no debería invertir en formación.

Acemoğlu y Pishke (1998, p. 80) plantean una historia creíble que explica por qué las firmas deberían pagar por el entrenamiento en competencias generales y ofrecer contratos de aprendizaje de sus trabajadores, con base en un mecanismo de selección adversa: en forma muy general, las habilidades y competencias de los nuevos trabajadores es desconocida, de modo que no es posible saber cuáles individuos son más productivos que otros; no obstante, el proceso de entrenamiento, desde su inicio, habilita para la firma un mecanismo de recolección de información sobre las habilidades del que se está entrenando; una información que en su mayor parte es recabada por el empresario y permanece oculta a otras firmas, a las que no se transmite.

Los incentivos para financiar la formación del joven trabajador están entonces vinculados a un poder informacional de monopsonio, al cual debería hacerse quien paga por la formación y tiene el poder para extraer rentas de ella: si la firma puede apropiarse del producto marginal del trabajador, tendrá un incentivo para invertir en él, y naturalmente, si los trabajadores se percatan de esto, no tendrán interés en financiar su propia formación:

Como resultado, en marcado contraste con el modelo de formación beckeriano estándar, una firma que quiera atraer trabajadores a sus contratos de aprendizaje, podría tener que pagar no únicamente el entrenamiento, sino también ofrecer a los aprendices un salario de entrenamiento por encima de su productividad inicial [traducción propia] (Acemoğlu y Pishke, 1998, p. 80).

En el caso colombiano no resulta difícil interpretar la historia de Acemoğlu y Pishke (1998) en términos de los incentivos a los que responden los agentes involucrados en la relación de aprendizaje. En Colombia, el apoyo privado al CA se limita normalmente a la fase práctica del proceso, e incluso hay firmas que, en circunstancias de no restricción de oferta de aprendices, tienden a monetizar⁶ la obligación de incluir aprendices en sus procesos fabriles.

Si fuera posible sostener que con los trabajadores entrenados el empresario puede extraer rentas que justifiquen el pago del proceso de formación completo (i. e., si las expectativas sobre retornos de su inversión son ciertos a la luz del riesgo de elegir en forma inadecuada a su personal), la condición estándar de acuerdo con la cual un empresario enganchará trabajo, en la medida en que el costo marginal de este trabajador no sea mayor que el beneficio marginal de contratarlo, debería observarse en el caso en el cual el empresario decide financiar todo el CA. Sin embargo, si el aprendiz no contribuye a la recuperación de los costos de su formación, una nueva razón para explicar el retiro del individuo bajo el tratamiento es el desbalance costo-beneficio, que obliga al empresario a prescindir del aprendiz, o bien, a incorporarlo bajo condiciones que impliquen mínimo coste.

El impacto de los programas de formación para el trabajo ha sido examinado recientemente en el caso colombiano, con resultados a veces contradictorios. Por ejemplo, una investigación adelantada por Fedesarrollo (Estacio *et al.*, 2010) muestra que los salarios de las personas que se

⁶ Véase el artículo 34 de la Ley 789 de 2002.

han sometido a un proceso de formación para el trabajo son mayores que los de los individuos que no lo han hecho.

Por otra parte, un estudio a cargo de la red de investigadores del Banco Interamericano de Desarrollo (BID) proporciona resultados que no suponen ventaja de los programas del SENA respecto de los que ofrecen los privados; incluso, bajo argumentos de costo-beneficio, encuentra que los programas del SENA son mucho menos rentables que los de otras entidades privadas de formación (Medina y Núñez, 2005). Dichos estudios se sustentan cuantitativamente en estimaciones de funciones de ingresos de Mincer, en las que el indicador de tratamiento aparece acompañado de controles de tipo dotacional (i. e., variables asociadas al capital humano como las competencias —o alguna *proxy* de estas—, el polinomio habitual de la experiencia, el grado de educación formal, entre otros).

Para la formación vocacional y, en particular, para los CA y la formación dual, según se reconoce ampliamente en la literatura (Gunderson y Harry, 2011; Mohrenweiser, 2010; Korpi y Antje, 2003; Wolter, 2012), se reclama que proporciona a los aprendices competencias que pueden coincidir con las demandas del mercado laboral en escenarios de selección adversa, al igual que aumenta sus oportunidades de transitar suavemente a la vida laboral con mayor permanencia y baja rotación interocupacional (Acemoğlu y Pishke, 1998; Korpi y Mertens, 2003). Dichas ventajas, sin embargo, no se explotan en forma extensa y general; un reporte reciente en *The Economist* denuncia que, independientemente de los beneficios que podrían involucrar, los CA no se han extendido en forma generalizada en Europa, en razón a la baja oferta de *apprenticeships* que se observa en la práctica (Bagehot, 2011).

Una explicación particular para la baja oferta de CA es el temor que los empresarios tienen respecto al ofrecimiento de plazas para aprendices, por cuanto podría estar asociado con costos netos sustanciales que no se repagan si el aprendiz se moviliza hacia otro empleador. Este es un argumento

comparable al de Acemoğlu y Pishke (1998): si el empresario no puede imponer algún límite a la movilidad del aprendiz, suponiendo que este hace aportes efectivos a la productividad, la firma resultará renuente a ofrecer *apprenticeships*. En el caso positivo, y con un lenguaje más común y accesible, Wolter (2010) resume una serie de formas en que una firma puede lucrarse de los CA:

- a) En la medida en que los aprendices trabajan para la firma patrocinadora, el valor de su contribución productiva puede superar costos del entrenamiento relacionados normalmente con costos materiales y de infraestructura, salarios de los instructores y apoyos de los aprendices.
- b) Las firmas pueden usar los CA como un medio de detección de los aprendices más motivados y talentosos. Aunque el CA termina con el final del proceso de entrenamiento, siempre es posible ofrecer un contrato normal de trabajo a los individuos más hábiles —dado un mejor conocimiento de las aptitudes del aprendiz, en general no disponible para otras firmas— y, por tanto, evitar los costos de la selección adversa.
- c) Las firmas pueden entrenar aprendices teniendo en cuenta sus necesidades específicas de competencias, lo cual constituye una importante ventaja en tiempos de escasez de trabajadores calificados en distintas industrias, situación que dificulta llenar las vacantes con trabajadores del mercado laboral externo. En efecto, un trabajador vinculado del mercado laboral externo siempre requerirá satisfacer la necesidad de adquirir habilidades específicas durante su periodo de adaptación, un hecho que restringe la productividad del nuevo trabajador. En muchos casos, los costos de la contratación externa resultan substancialmente altos en relación con

la inversión potencial neta asociada al entrenamiento de un aprendiz.

A las de Wolter (2010), que no son en realidad razones novedosas, suelen añadirse las ventajas que pueden reportar estas modalidades de formación para la economía global. Hace ya cerca de cuarenta años, Ashenfelter (1978) atribuía para los *apprenticeships* una serie de ventajas concretas: la formación para el trabajo ayuda a reducir la inflación mediante la provisión de trabajadores más especializados para paliar escasez de competencias; los CA contribuyen a reducir el desempleo de ciertos grupos; la formación para el trabajo reduce la pobreza, al contribuir a aumentar las capacidades de algunos segmentos poblacionales. Aun cuando el argumento relacionado con la inflación resulte excéntrico en los tiempos actuales, los otros dos, según se ha revisado, siguen siendo de importancia.

La importancia relativa de los aportes del CA y de la formación para el trabajo varía con frecuencia entre individuos, firmas, ramas productivas, ocupaciones y países. Normalmente, la importancia del instrumento también varía con el enfoque de la política laboral, que puede ir de consideraciones de largo plazo a la definición de metas en el corto. En muchos casos, las variaciones en los resultados están relacionadas con la estructura del mercado laboral de referencia. En sociedades con mercados de trabajo más flexibles, las firmas que patrocinan CA enfrentan un riesgo importante de perder su inversión ante la imposibilidad de imponer barreras a la movilidad de aprendices entre firmas, y tienden a ofrecer menos CA si la perspectiva de percibir beneficios netos de ellos no es halagüeña; el beneficio que surge del *screening* disminuye si otras firmas pueden contratar con facilidad trabajadores en el mercado externo de trabajo (*second hand market*) y despiden a aquellos que resultaron una mala elección.

En estas condiciones, las empresas patrocinadoras tienen que confiar en la posibilidad de cubrir sus gastos de formación con el valor de la productividad de los aprendices durante el periodo de formación. En Suiza, donde el mercado laboral es relativamente flexible, hay evidencia

de costo-beneficio de que una firma típica puede recuperar la inversión en entrenamiento cerca del final del proceso de formación de aprendices; según dicha evidencia, el aporte productivo de los aprendices excede el costo del entrenamiento, incluido el apoyo para la formación en más de 6000 €/aprendiz (Mohrenweiser, 2010).

¿De qué depende que una firma se beneficie de los CA en condiciones de flexibilidad laboral? Ryan (2011) afirma que la obtención de tales niveles de beneficio neto depende crucialmente de que los aprendices se involucren tempranamente en procesos de trabajo productivo y que los apoyos económicos de los aprendices sean menores comparados con las remuneraciones de los trabajadores más experimentados; en otro caso, los CA constituirían observancias de la ley, cuyo resultado y desarrollo, por estar desligado del curso real de los negocios de la firma, será artificial y espurio.

Para el caso alemán, donde se han adelantado análisis similares con base en información de empresas compilada a lo largo de varias décadas, se tiene evidencia según la cual la firma típica está dispuesta a cargar con el coste neto del entrenamiento durante la fase de formación porque las regulaciones del mercado laboral facilitan la recuperación de los beneficios de la inversión en periodos posteriores a la culminación del proceso de formación. No obstante, la crítica puntualiza que, en la medida en que el mercado laboral alemán se ha venido tornado más flexible, las firmas han adoptado cambios en sus prácticas de entrenamiento y han involucrado cada vez más a los aprendices en labores productivas que les obligan a percibir rentas que antes estaban parcialmente financiadas por el estado. Este movimiento hacia una política de formación orientada a la producción ha redundado en costos de entrenamiento netos mucho más bajos (Wolter, 2010).

La importancia de la regulación es mayúscula y no necesariamente de efecto positivo. En relación con los resultados en Alemania, críticos ingleses señalan que muchos de los oficios en ese país están sujetos a una "licencia de práctica" que solo puede ser obtenida por un trabajador si completa

en forma exitosa un proceso de aprendizaje con práctica en el interior de las firmas. En Alemania es ilegal ofrecer trabajos a jóvenes en ocupaciones reconocidas si no es bajo los términos de un CA.

Al contrastar esta experiencia con el caso británico, donde los empleadores son mucho más renuentes a enganchar aprendices, no deja de ser interesante notar que en realidad no están obligados por la ley a vincular aprendices. Con esta visión, la piedra angular del sistema de aprendizaje alemán es un arraigado corporatismo: las prácticas laborales limitadas por certificación suponen una restricción que opera en detrimento de la movilidad laboral del trabajador, pero que permite a las firmas alemanas recuperar el costo neto de su inversión en formación (Bagehot, 2011).

En Colombia, el CA aparece fuertemente regulado no únicamente por las prescripciones sobre el mercado laboral en general, sino bajo los mandatos asociados en forma concreta a este contrato (Ley 789 de 2002). De este modo, en algunas instancias se ha considerado el CA como una carga parafiscal que afecta negativamente la productividad de las firmas.

Ahora bien, del lado del aprendiz se argumentan tres puntos fundamentales como ventajas intrínsecas de la formación para el trabajo y del esquema dual de formación: a) mejores remuneraciones salariales respecto de quienes no se han sometido a un proceso de formación para el trabajo, b) suavidad en el tránsito hacia la vida laboral y c) baja movilidad ocupacional, es decir, menos retiros y pérdidas de empleo durante la carrera laboral. Estacio *et al.* (2010) proponen además, para el caso colombiano, d) mejores condiciones laborales —medidas con un índice de calidad del empleo— y e) mejores condiciones de vida. Sin embargo, los resultados de evaluación, en relación con estos dos puntos, siguen siendo provisionales.

El aludido estudio de Fedesarrollo asegura que los programas de formación técnica y tecnológica del SENA tienen impactos positivos sobre aspectos relacionados con las condiciones laborales de los aprendices. La empleabilidad de los sujetos tratados es superior a la de aquellos que

no se sometieron al tratamiento, aun cuando no se encuentran diferencias importantes en el nivel de ingreso. Los egresados del SENA tienen una probabilidad de entre el 7,4% y el 9,6% mayor de estar empleados en comparación con personas de características socioeconómicas similares que no se han formado allí (Estacio *et al.*, 2010, p. 18). No obstante, los resultados, según indican los autores, presentan una sensibilidad no despreciable a las estrategias de agregación de los datos adoptada.

Los hallazgos de Kuckulenz y Zwick (2003), para el caso alemán, señalan que en lo concerniente al ingreso, el impacto de la formación para el trabajo bajo el sistema dual es significativamente positivo. Sin embargo, los efectos pueden variar entre agentes heterogéneos. Trabajadores altamente especializados se benefician comparativamente más que los de baja cualificación. Luego de experimentar algún proceso de formación, los trabajadores entrantes perciben incrementos salariales mayores que trabajadores de mayor permanencia en el trabajo. Trabajadores con contratos de corto plazo se benefician menos que aquellos con contrato permanente de trabajo. Si los trabajadores que no disfrutaban de un efecto salarial positivo experimentan un incremento en su productividad, inducido por procesos de formación, los empleadores retienen todos los beneficios de la inversión en capital humano.

Gunderson y Krashinsky (2011) encuentran resultados similares para el caso canadiense. Usando datos del Censo General de 2006, encuentran que, en comparación con otras alternativas de formación (i.e., bachillerato, *community colleges*, *non-apprenticeship trades*), los aprendices varones ganan un 24% más que aquellos cuyo nivel de educación más alto es el bachillerato, 15% más que aquellos asociados a los *non-apprenticeship trades* y 2% más que los egresados de *community colleges*. También se observan diferencias de género importantes: para los hombres, el *apprenticeship* supone un *premium* sobre el ingreso medio, mientras que en el caso de las mujeres hay, por contraste, déficit. Para ellas, la formación bajo CA produce menores retornos que los que implicarían

la terminación del bachillerato, y muchísimo más bajos aún que los que le reportarían concluir el *community college*. El resultado está asociado probablemente a que los CA están relacionados con oficios y ocupaciones de baja remuneración en las ramas productivas de alimentos y servicios. Estas barreras institucionales no visibles (por ejemplo, *glass ceilings*) y sus consecuencias socioeconómicas deben ser investigadas en cualquier estudio sobre este particular en Colombia.

Para aislar y controlar por componentes políticas e institucionales, Korpi y Mertens (2003) adelantan una comparación entre países con similares modelos de formación. En los casos estudiados por esos autores (Alemania y Suecia), se encuentra que mayores tasas de movilidad están relacionadas con menores niveles de preparación y trabajadores con educación tradicional (bachillerato no vocacional). Estas diferencias no parecen estar relacionadas con la movilidad en el interior de la firma o entre firmas e industrias. Los CA no eliminan el proceso de búsqueda al comienzo de la vida laboral, pero reducen la movilidad ocupacional; si hay instancias en las que las diferencias relacionadas con formación vocacional permanecen, se dan en el caso de la movilidad ocupacional.

El mismo estudio hace énfasis en la baja movilidad interocupacional de los aprendices, resultado que sugiere que las competencias obtenidas son menos generales que las que se ofrecen en las escuelas de formación vocacional. En relación con este tipo de movilidad, el mercado laboral alemán aparece más estructurado alrededor de los sistemas de formación y entrenamiento para el trabajo que en Suecia, país en el que la educación vocacional tiene lugar principalmente en la escuela. No se detectan diferencias en la movilidad entre firmas que puedan relacionarse con el tipo de formación vocacional. La proporción de competencias realmente específicas a la firma en un CA en Alemania es más bien baja en relación con las competencias transferibles obtenidas, por lo cual parece ser claro que las diferencias institucionales, relativas al grado de regulación del mercado laboral, son de importancia.

Haciendo uso, como resulta habitual en este tipo de estudios, de estimaciones de ecuaciones de Mincer, McIntosh (2007) indaga acerca de los costos y beneficios de los contratos de aprendizaje financiados por el estado en el Reino Unido. Las estimaciones muestran que la diferencia salarial de las relaciones de aprendizaje que son objeto de estudio —los llamados *modern apprenticeships*— puede ser de hasta el 18% en los niveles de cualificación vocacional NVQ3 y de hasta el 16% en el NVQ2 respecto de individuos de los menores niveles de cualificación.

El resultado está asociado a la baja oferta de *apprenticeships*, por lo que los empleadores pueden utilizarlas como un *screening device* que les facilita detectar a los aprendices más hábiles: parte del resultado viene explicado por diferencias dotacionales, esto es, por el valor de las habilidades *per se*. Si se incluye un *dummy* de género, se detectan diferencias salariales de relieve para las mujeres: 14% en NVQ3 respecto de los otros niveles de formación. En el campo sectorial también se detectan diferencias significativas: los CA modernos [NVQ3] incrementan el salario promedio de un individuo del sector de la construcción en un 32% respecto de un individuo del mismo sector con NVQ2, mientras que en el comercio no se observan diferencias sustanciales.

En relación con la empleabilidad en el Reino Unido, todas las modalidades de *apprenticeships* mejoran la probabilidad de estar empleado. El resultado obtenido en McIntosh (2007) es comparable con el resultado empírico en Acemoğlu y Pishke (1998), según el cual el número mediano de trabajos diferentes por los que pasa un trabajador norteamericano con 10 años de experiencia es 6, mientras que en Alemania es 1 o, cuando más, 2.

El caso de Dinamarca resulta de gran interés, toda vez que, junto con el caso alemán, supone una de las experiencias en formación dual de mayor éxito. Westergård-Nielsen y Rasmussen (1997) indagan por el impacto sobre los ingresos del trabajador atribuibles al sistema dual de formación para el trabajo. A partir de una aproximación econométrica, los investigadores estiman ecuaciones de ingresos

con ocupabilidad sobre un tramo muestral de, al menos, 60.000 individuos registrados en el Danish Longitudinal Market Register: en general, el salario de un trabajador calificado es mayor que el de uno no calificado, lo que no supone ya un hallazgo novedoso. El *premium* de un *apprenticeship* oscila alrededor del 25% y el 30% respecto del ingreso de trabajadores sin formación bajo esa modalidad.

Al igual que en el caso canadiense, un sesgo altamente relevante proviene del hecho de que una alta proporción de trabajadores daneses no calificados son mujeres, en tanto que solo una proporción menor de mujeres forma parte del conjunto de los trabajadores calificados. Estos, por su parte, tienen una menor probabilidad de estar desempleados que los trabajadores no calificados (i. e., las diferencias entre trabajadores con competencias y sin competencias es el resultado de la decisión de haber optado por una relación de aprendizaje) (Westergaard-Nielsen y Rasmussen, 1997).

En las experiencias analizadas se identifican dos variables centrales por tener en cuenta en una propuesta de evaluación como la presente: a) el nivel de ingreso, que puede estar asociada con la variable de resultado, y b) que el individuo haya pasado por un proceso de formación dual, que constituiría el indicador de tratamiento. Un conjunto de variables relacionadas con aspectos dotacionales, institucionales y sectoriales contribuiría a moderar los efectos nocivos del sesgo de selección y del sesgo de variable omitida (sesgo de especificación) que se presenta habitualmente en estos ejercicios; no obstante, tienen una intención de mayor alcance: cualificar el impacto del tratamiento, a partir de las características de cada uno de los individuos que se registran en la base de datos. Dichas características no se limitan al *background* socioeconómico del individuo: incluyen características demográficas, de ambiente sectorial y de ambiente territorial, así como de las características mismas del tipo de formación al que el individuo fue sometido, como en el estudio de Kuckulenz y Zwick (2003), que indaga con detalle por estas características.

PROPUESTA METODOLÓGICA

La propuesta metodológica que se describe en esta sección es, sin ir más lejos, una evaluación económica de impacto con base en microdatos. Parece conveniente revisar de cerca —ya no en cuanto a resultado sino en cuanto a procedimientos— algunas de las propuestas reseñadas en la revisión de literatura para, a partir de este desarrollo, discutir aspectos empíricos generales relacionados con la práctica de la *evaluación de impacto*, que para el caso presente tiene una relación muy estrecha con la ya mencionada *ecuación de Mincer*. El tema relativo a los datos disponibles se pospone para la sección siguiente.

Antecedentes: síntesis de experiencias

En las fuentes revisadas se ha identificado la tendencia a concentrarse en el desarrollo de aproximaciones cuantitativas del resultado de los contratos de aprendizaje o *apprenticeships* sobre la condición económica y laboral del aprendiz. Esta práctica no supone que los aspectos relacionados con la economía del empresario no resulten de interés o no sean tema de concentración especial —como sucede en el caso de Mohrenweiser (2010), por ejemplo—, o que no sea de interés el conjunto de resultados sobre el mercado de trabajo, dadas variaciones en las condiciones macroeconómicas prevalentes, como es el caso del estudio de Brunello (2009), que reúne evidencia sobre el carácter procíclico del *ratio* aprendices a trabajadores, dado el efecto de cambios de régimen en el desempeño económico general.

A grandes rasgos, los estudios reseñados se concentran en consideraciones de equilibrio parcial⁷, en oposición a aquellos de naturaleza estructural y microfundamentada con que se abordan los efectos de equilibrio general de una iniciativa de política activa de empleo sobre el bienestar de un segmento específico de población a lo largo del

7 No debe dejarse de notar, en cualquier caso, que el enfoque de los *salarios hedónicos* es, en realidad, de equilibrio general competitivo.

ciclo de vida. Un estudio que cae en esta categoría es, por ejemplo, el de D. Lee (2005), que utiliza un modelo computable de generaciones traslapadas para analizar decisiones sobre escolaridad, empleabilidad y elección de la ocupación, y calibrado para reproducir los primeros momentos de series sobre empleabilidad individual durante el ciclo de vida en los Estados Unidos.

La práctica que caracteriza las investigaciones revisadas es la estimación de relaciones entre el ingreso de un individuo típico que se ha formado en sistemas *duales* y varias características que registran este hecho, es decir, el que el individuo haya pasado por una relación de aprendizaje. En la parte empírica del estudio de Acemoğlu y Pishke (1998) se adopta como variable de resultado al (*log del*) salario promedio por hora, en tanto que en el lado derecho (RHS), además del indicador de *tratamiento*, se incluyen los controles típicos de la ecuación de Mincer: el máximo grado de educación general alcanzado y un polinomio de cuarto grado en la experiencia del individuo. En este estudio se introducen además algunas variables dicotómicas que identifican al individuo como un *stayer* —i. e., el individuo trabaja en la firma donde el contrato de aprendizaje tuvo lugar— o un *quitter* —el individuo se ha retirado por razones como la obligación de prestar servicio militar— (Acemoğlu y Pishke, 1998, p. 109).

Westergård-Nielsen y Rasmussen (1997) adoptan una aproximación metodológica similar para examinar el caso danés: estiman el aporte marginal sobre el (*log del*) salario por hora, producido por cambios en las variables usuales del modelo de Mincer; se incluyen interacciones entre la experiencia —mensurada como el tiempo durante el cual el individuo ha participado en el mercado laboral— y las competencias, que son representadas por un *dummy* que es igual a 1 en el caso de trabajadores que han pasado por un proceso de formación para el trabajo y 0 en caso contrario. Se incluyen controles que identifican diferencias de género y *dummies*

sectoriales para seis diferentes ramas productivas⁸ (Westergård-Nielsen y Rasmussen, 1997, p. 21).

Gunderson y Krashinsky (2011) se valen del mismo tipo de aproximación econométrica: el (*log del*) ingreso semanal se regresa en los años de educación y en una serie de variables categóricas dispuestas con el propósito de capturar el efecto credencial asociado a los grados de educación alcanzados, certificados mediante un diploma (bachillerado, pregrado, etc.). Entre otras, son incluidas las variables convencionales de control, un polinomio de la experiencia, el estatus marital, un *dummy* que identifica la pertenencia del individuo a alguna minoría socioeconómica, la región y la rama productiva.

En las evaluaciones para el caso colombiano se hace otro tanto. Jiménez y Kugler (1987) indagan el impacto sobre el ingreso laboral de la formación del SENA, para lo cual distinguen explícitamente entre cursos cortos (educación continuada) y cursos largos (formación titulada). La aproximación parte de comparaciones entre los ingresos de individuos tratados y no tratados, dadas ciertas características socioeconómicas. Las estimaciones no restringidas del modelo adolecen de sesgo de selección, un problema que se enfrenta al controlar argumentos que explican la elección de formación del individuo. El modelo resultante explora, por tanto, la selección de una variable instrumental que mapea sus ocurrencias en las condiciones socioeconómicas de los individuos; el procedimiento de estimación es un *probit* ordenado que corrige el sesgo antedicho, pero que disminuye el efecto del SENA en los ingresos. Parte de este resultado tendría que ver con que los datos no proporcionan medidas explícitas de las habilidades de los sujetos observados (Jiménez y Kugler, 1987).

Fedesarrollo, en su investigación de 2010, incluye variables relacionadas con las condiciones del hogar del individuo, como si el sujeto es jefe del hogar, el tamaño de este y si hay menores de 5 años (Estacio *et al.*, 2010, p. 20).

8 Estas variables, sin embargo, no resultaron estadísticamente distintas de cero.

Por su parte, en el estudio de Medina y Núñez (2005) se incluyen controles como la educación del padre y de la madre del individuo, la tasa de desempleo regional, la región en la que vive el individuo (se incluyen 7 regiones, incluyendo San Andrés y Providencia), así como el número de niños menores de 6 años en el hogar y el número de personas que allí residen. Un regresor que se incluye en Medina y Núñez (2005) y que no aparece en ninguno de los estudios citados es la *edad a la cual el individuo obtuvo un trabajo remunerado por primera vez*.

La *ocupabilidad* del individuo, que, en general, es tratada a través información relativa al estatus de este en relación con la búsqueda de trabajo, es estudiada en Westergård-Nielsen y Rasmussen (1997, p. 11): regresan un indicador de desempleo en los controles utilizados en la ecuación del ingreso inicialmente estimada. Respecto a la *calidad del empleo*, Estacio *et al.* (2010, p. 14) construyen un índice de calidad según la propuesta de OIT, que pretende sintetizar diferentes dimensiones como el ingreso, la modalidad de contratación, la afiliación a seguridad social y el horario de trabajo.

Con estos antecedentes parece atractivo mensurar el efecto de los CA mediante una aproximación metodológica similar que pueda dar cuenta de si los postulados que se formulan para el instrumento se verifican en la práctica: mejores ingresos, mayor empleabilidad y mayor estabilidad laboral.

Evaluación econométrica de impacto: generalidades⁹

La investigación sobre la valoración del impacto de programas o intervenciones de política sobre aspectos diversos de las economías actuales ha

venido produciendo una variedad de métodos aplicables en la medición de los resultados que se esperan de esas intervenciones. La literatura sobre la metodología econométrica de evaluación de resultados incluye los textos de Khandker, Koolwal y Samad (2010), Gertler *et al.* (2011) y Ravallion (2008), así como el de Bernal y Peña (2011), que incluye referentes para el caso colombiano, y el de D. Lee (2005), que desarrolla los contenidos teóricos de la práctica con mayor detalle y especialización. En relación directa con el mercado laboral, Heckman, Lalonde y Smith (1999) aportan una extensa monografía sobre la aplicación de técnicas de evaluación de impacto para el estudio de resultados en programas de política activa de empleo.

La evaluación de impacto postula el problema de medir las diferencias que sobre una variable de resultado (y) produce un tratamiento, programa o intervención (T). Preguntas de investigación que pueden ser respondidas mediante esa aproximación son, entre otras, las siguientes: ¿es el CA o alguna otra modalidad de formación dual una forma de mejorar las condiciones laborales de un individuo? ¿Son los salarios de las personas que han pasado por un *apprenticeship* mayores que los de aquellas que no estuvieron vinculadas a un proceso de formación para el trabajo? ¿Son los CA un vehículo eficaz para reducir el desempleo estructural?

Programas como el CA no son de alcance universal, tanto por razones de viabilidad como por circunstancias políticas e incluso éticas. Para el caso de Colombia, los recursos del SENA son limitados, circunstancia que restringe el número de individuos que pueden beneficiarse y que exige el diseño de un mecanismo de selección que implique algún grado de focalización. Así, la asignación de candidatos no es aleatoria y se limita a un subconjunto de la población total que puede observar algunas características comunes, pero que bien pueden ser diferentes en aspectos no observables. La no universalidad del programa y la imposibilidad de que la asignación de individuos haya surgido de un experimento aleatorio obliga la

⁹ En oposición a los *métodos estructurales* para la medición de los resultados de las iniciativas públicas, como modelos de insumo producto (IO), de equilibrio general computables multisectoriales o modelos de equilibrio general computables (CGE), dinámicos y estocásticos (DSGE). La exposición en este acápite sigue de cerca las de Bernal y Peña (2011) y Khandker, Koolwal y Samad (2010). Refiérase a esos textos para desarrollos adicionales a los tratados aquí.

estimación del impacto promedio del tratamiento sobre los tratados (*average treatment on treated*), definido como “el efecto promedio del tratamiento en el subconjunto de individuos que fueron efectivamente tratados” (Bernal y Peña, 2011):

$$\tau_{ATT} = E[\tau_i | T_i=1] \equiv E[y_{it} | T_i=1] - E[y_{i0} | T_i=1] \quad [1]$$

En el que $\forall i$ la componente $E[y_{i0} | T_i=1]$ representa el salario promedio de los aprendices del SENA que optaron por un *apprenticeship* como modalidad de fase práctica de su proceso de formación, *si no hubieran pasado por él*. La evaluación empírica de [1] exige la selección de un sustituto adecuado para los resultados contrafácticos, como, a modo de ejemplo, el valor de la variable de resultado de individuos que no participaron pero que podrían resultar elegibles para participar: jóvenes de *background* socioeconómico comparable con el de los aprendices SENA, o como el valor de la variable de resultado de aspirantes inscritos que no fueron seleccionados, o que fueron seleccionados y luego desertaron, etc.

Estas comparaciones pueden significar una fuente de sesgo, dado que participantes y no participantes pueden ser diferentes en características no observables, o incluso predefinidos en ausencia del programa. Los aprendices del SENA que entraron al programa de formación y optaron por el CA —el grupo de tratamiento— podrían ser diferentes de los que entraron al SENA y optaron por otra modalidad de desarrollo de la fase práctica de su programa (participación en proyectos productivos, monitorías, vinculación laboral o contractual *inter alia*), por razones relacionadas con cuestiones actitudinales, diferencias relativas a su capital social individual o simples preferencias profesionales, no registradas en los datos disponibles.

Al mismo tiempo, el grupo de tratamiento puede diferenciarse de individuos que, con características socioeconómicas similares, aspiraron a iniciar un programa de formación del SENA pero no fueron admitidos por cuestiones actitudinales y conductuales, de logro académico, de

competencias blandas, o cualquier combinación de ellas, para las que tampoco hay registro. Los dos grupos de comparación propuestos pueden reunirse en un único *grupo de control* por el hecho simple de ser considerados potenciales beneficiarios o individuos elegibles, y puede anticiparse que las diferencias promedio en las variables de resultados de los grupos de tratamiento y de control pueden no deberse exclusivamente a la influencia del programa.

Ravallion (2008) observa que la forma en que se asignan individuos a un programa involucra objetivos e intereses predefinidos que reflejan las preferencias de los elegibles y de los administradores del programa; de esta guisa es probable que los factores que determinan la asignación ejerzan influencia sobre el resultado contrafáctico, razón por la cual es más que adecuada la presunción de existencia de sesgo de selección en las comparaciones de los resultados de los individuos de los grupos de tratamiento y de control (Ravallion, 2008, p. 3793). La consideración de factores observables y no observables es, en consecuencia, obligada.

La ecuación de Mincer

El modelo de ingresos de Mincer (Mincer, 1974) entra en la propuesta de evaluación de impacto del CA en forma natural: una especificación lineal que exprese una relación funcional entre características del capital humano del individuo en el RHS y los retornos del activo compuesto (*compound asset*), representados en el salario, constituye una versión no restringida de la función de Mincer.

Este dispositivo analítico ha sido utilizado en estudios empíricos sobre desarrollo económico y capital humano desde hace cerca de cuarenta años. Con su estimación econométrica es posible obtener una cuantificación de los retornos económicos a la educación, a su cantidad y a su calidad. Los estudios de género modernos usan versiones ampliadas de dicho modelo para indagar por el impacto de la experiencia en las brechas salariales entre hombres y mujeres, la validación de la existencia de *glass*

ceilings, entre otros. Es frecuente encontrar aplicaciones empíricas de la ecuación en la literatura sobre crecimiento económico y su relación con la inversión en capital humano. Especificaciones particulares relacionadas con aspectos espaciales son acostumbradas, por ejemplo, en estudios de desarrollo regional (Heckman, Lochner y Todd, 2003).

La popularidad de la ecuación de Mincer no se atribuye exclusivamente a la facilidad relativa de su uso y al buen ajuste que exhibe. En ámbitos de investigación que cada vez más disponen de microdatos y de poder computacional para tratar problemas econométricos con miles de datos, estas propiedades son apreciables, si bien la robustez teórica y el poder descriptivo subyacente a su construcción son los argumentos centrales aquí: la ecuación de Mincer, además de constituir un marco empírico para el estudio de los determinantes de los salarios individuales, tiene sustento en un modelo formal de inversión en capital humano (Lemieux, 2003) que presenta dos características sobresalientes: a) es una ecuación de precios hedónicos en la que se revelan las recompensas que el mercado otorga a atributos como la educación y la experiencia y b) proporciona una medición de la tasa de retorno de la inversión en educación, que, al ser comparada con la tasa de interés de los demás activos en la economía, da lugar a la determinación de la optimalidad de las inversiones en capital humano en una economía de mercado (Heckman, Lochner y Todd, 2003, p. 1).

En su forma más común, la ecuación modela el (*log del*) ingreso laboral como la agregación de una función lineal de los años de escolaridad y de un polinomio (en general, de segundo grado) de la *experiencia potencial del individuo en el mercado de trabajo*:

$$\ln y = \ln y_0 + rS + \gamma_1 x + \beta \gamma_2 x^2 \quad [2]$$

En la expresión [2], la variable y representa los pagos que recibe el individuo por el alquiler de su fuerza de trabajo, y_0 es el nivel de ingreso de un individuo sin escolaridad ni experiencia, S

representa el número de años de escolaridad del sujeto y x es experiencia potencial en el mercado laboral del individuo¹⁰. El argumento en favor de la interpretación hedónica proviene del hecho de que la componente lineal del modelo ($\ln y = \ln y_0 + rS$) es una condición de equilibrio que caracteriza la solución de un individuo que elige invertir en capital humano para maximizar el valor presente de los retornos del activo adquirido a lo largo de la vida (Lemieux, 2003; Rosen, 1974). El parámetro r mide el retorno marginal por un año adicional de educación, que en el equilibrio debe ser igual a la tasa intertemporal de descuento. La componente no lineal en el RHS del modelo ($\gamma_1 x + \beta \gamma_2 x^2$), que es un polinomio de la experiencia, captura inversiones en capital humano posteriores a la vida escolar del trabajador (Lemieux, 2003).

Aunque una de las propiedades más apreciadas de la ecuación de Mincer es el buen ajuste estadístico que suele exhibir en aplicaciones empíricas, muchas investigaciones logran mejorar la eficiencia de las estimaciones incluyendo controles adicionales. Lemieux (2003) argumenta en favor de esta práctica que, a menos que la ecuación de Mincer sea una *ley de comportamiento* de aceptación universal, es difícil sostener que el modelo básico [2] sea la forma más parsimoniosa de representar las relaciones entre los niveles de capital humano de un individuo y los ingresos percibidos bajo dichas condiciones. Por ejemplo, un debate en desarrollo es si el polinomio que caracteriza la experiencia debería —como en el caso de Acemoglu y Pishke (1998)— ser sustituido por un polinomio de grado 4, pues, como es razonable suponer, los retornos de la experiencia varían a lo largo del ciclo de vida: en efecto, las bases de datos usualmente contienen información correspondiente a distintos grupos de edad y cohortes. Nuestra propuesta de evaluación de impacto consiste en estimar [2] teniendo en cuenta especificaciones funcionales que se discuten en la siguiente subsección.

¹⁰ La experiencia puede computarse como la edad menos el número de años de escolaridad menos seis, i.e. $x_i = (edad_i - S_i - 6)$, $\forall i$

Modelo

La presente propuesta de evaluación de impacto gira alrededor de la estimación de la ecuación de Mincer sobre una base de datos que reúne información proveniente de la llamada PILA y del registro de aprendices de formación del SENA. No se dispone de información útil sobre resultados operativos de las firmas, relacionables con la incorporación de aprendices a los procesos fabriles de las empresas, razón que impide desarrollar evaluaciones de los resultados del lado de la demanda. En la situación ideal, la valoración del impacto sobre las empresas implica disponer de información de costo-beneficio, como en el caso de los estudios particulares documentados en la revisión de literatura (McIntosh, 2007; Mohrenweiser, 2010).

Según se ha anticipado, el centro de interés en esta propuesta es la estimación econométrica de una ecuación *no restringida* de ingresos de Mincer como medio para identificar el impacto atribuible a los contratos de aprendizaje. El modelo de Mincer no restringido agrega a la porción básica [2] dos componentes: el primero permite distinguir a los grupos de tratamiento y de control en forma explícita mediante la inclusión del indicador de participación en el *programa*; el segundo incluye un conjunto de variables explicativas que detectan características de los individuos *antes de la asignación* al tratamiento.

Con la identificación e incorporación de regresores adicionales se logra, además de una descripción más detallada de las fuentes de variación en el ingreso (pero también en la empleabilidad) asociadas con el capital humano, reducir el sesgo de selección, que es un mal endémico de estas aproximaciones. El razonamiento es elemental: en la medida en que la ecuación [2] pueda atribuir diferencias observables en el ingreso a la formación vocacional en general y al CA en particular, es posible que parte de las diferencias en la variable de resultado pueda atribuirse a una variedad amplia de controles que pueden implicar diferencias entre los individuos, aun en ausencia del programa.

Estas diferencias son recogidas por los residuales, y aun cuando la asignación de la participación en el programa no es resultado de un experimento aleatorio, es de esperar que las decisiones de participar en un proceso de formación vocacional en el SENA y de optar por un CA como modalidad de fase práctica no sean aleatorias tampoco y puedan estar asociadas a características individuales de los individuos, como lo advierte Ravallion (2008). Considérese una versión estimable de la ecuación de Mincer:

$$\ln y = \ln y_0 + rS + \gamma_1 x + \gamma_2 x^2 + \mu \quad [3]$$

La ecuación [3] se diferencia de la [2] en el término de error estocástico que se ha supuesto $\mu \sim N(0, \sigma_\mu^2)$. En [3], $\ln y$ es el (*log* del) ingreso del individuo, $\ln y_0 = \beta_0$ es una constante que representa los ingresos promedio de un individuo sin educación ni experiencia, S y, finalmente, x es la experiencia potencial del individuo o una *proxy* de las inversiones en capital humano posteriores a la escolaridad secular¹¹. Esta aproximación es vista por algunos críticos como una caja negra que, en cualquier caso, requiere apertura. Kuckulenz y Zwick (2003, p. 3) proponen la utilización de una variable dicotómica T para significar educación vocacional continuada:

$$\ln y = \beta_0 + rS + \gamma_1 x + \gamma_2 x^2 + \beta_1 T + \mu \quad [4]$$

En [4], el indicador T distingue a los individuos con formación para el trabajo ($T = 1$) de los que no la tienen ($T = 0$). Aunque bajo algunas circunstancias esta especificación implica separabilidad fuerte entre la escolaridad secular S y las inversiones en capital humano (el polinomio de x), se deberá tener en cuenta que en el caso colombiano dicha separabilidad es más bien débil en naturaleza, dado que individuos que no han

11 Mincer (1991) señala al respecto: "La función de retornos del capital humano contiene, entre otras variables, años de experiencia (laboral), [...] que entran de modo no lineal. Los coeficientes se interpretan como parámetros de inversión en capital humano posteriores a la educación formal" (citado en Kuckulenz y Zwick, 2003, p. 3).

terminado secundaria, por ejemplo, pueden vincularse a procesos de formación vocacional. Esta aparente dificultad se sobrelleva introduciendo un término de interacción entre S y T . Considérese por ejemplo la siguiente especificación:

$$\ln y = \beta_0 + rS + \gamma_1 x + \gamma_2 x^2 + \beta_1 T + \beta_2 [S \times T] + \mu \quad [5]$$

Operando en esperanzas sobre [5], el impacto promedio del *tratamiento* sobre personas con escolaridad¹² y sin escolaridad (*ceteris paribus*, sin pérdida de continuidad, la experiencia, x) es:

$$E[y_i | T = 1, S > 0] = E[\beta_0 + rS + \gamma_1 x + \gamma_2 x^2 + \beta_1 T + \beta_2 S \times T] = \beta_0 + r + \beta_1 + \beta_2$$

$$E[y_i | T = 1, S = 0] = E[\beta_0 + rS + \gamma_1 x + \gamma_2 x^2 + \beta_1 T + \beta_2 S \times T] = \beta_0 + \beta_1$$

$$E[y_i | T = 0, S > 0] = E[\beta_0 + rS + \gamma_1 x + \gamma_2 x^2 + \beta_1 T + \beta_2 S \times T] = \beta_0 + r$$

$$E[y_i | T = 0, S = 0] = E[\beta_0 + rS + \gamma_1 x + \gamma_2 x^2 + \beta_1 T + \beta_2 S \times T] = \beta_0$$

Por tanto, el efecto de la formación vocacional para quienes tienen $S > 0$ años de escolaridad es:

$$E[y_i | T = 1, S > 0] - E[y_i | T = 0, S > 0] = (\beta_0 + r + \beta_1 + \beta_2) - (\beta_0 + r) = \beta_1 + \beta_2$$

Mientras que para aquellos que tienen $S > 0$ años de escolaridad, el impacto será:

$$E[y_i | T = 1, S = 0] - E[y_i | T = 0, S = 0] = (\beta_0 + \beta_1) - (\beta_0) = \beta_1 \blacksquare$$

El indicador T mide diferencias en el intercepto asociadas a la formación para el trabajo, pero no discrimina los efectos aislados del CA *per se* o de cualquier otra modalidad disponible para el desarrollo de la fase práctica del programa de

formación. Para aislar el efecto del CA, el procedimiento puede limitarse a incluir un *dummy* adicional para describir la situación; si $T_A = 1$ cuando el individuo elige *apprenticeship* y $T_A = 0$ en cualquier otro caso, el modelo modificado queda así:

$$\ln y = \beta_0 + rS + \gamma_1 x + \gamma_2 x^2 + \beta_1 T + \beta_2 [S \times T] + \beta_3 T_A + \beta_4 [T \times T_A] + \mu \quad [6]$$

Por supuesto, además del término $\beta_4 [T \times T_A]$, es posible introducir nuevas intracciones con, por ejemplo, la escolaridad, que como en el caso anterior contribuye a reducir en alguna medida la inercia de los coeficientes de la experiencia laboral (reduciendo la relevancia de la *caja negra* de la experiencia) y ayuda a capturar parte de la heterogeneidad individual observable que se incluye en las bases de datos disponibles (Kuckulenz y Zwick, 2003, p. 4).

A partir de esta especificación, el modelo admite una serie amplia de términos de interacción entre las modalidades observables de formación para el trabajo y las características de los sujetos considerados para la identificación de retornos del capital humano atribuibles a estas inversiones entre individuos heterogéneos. Como resultado, un atractivo concreto de la propuesta es que no se limita a evaluar el CA exclusivamente, sino que puede aportar información de impacto sobre la formación vocacional, sea esta titulada o complementaria, sea esta impartida por el SENA o por otras entidades reconocidas por el Estado, en los términos de la Ley 789 de 2002.

Los efectos de la formación vocacional varían entre trabajadores con diferentes carreras profesionales o sin ellas, las características de los lugares de trabajo, el logro educacional, el estatus profesional, etc. Para estas cualidades, considérese un componente adicional (vectorial), $\sum_{j=1}^n \delta_j X_j$, donde $\forall j=1, \dots, n, X_j$ es una matriz de variables explicativas adicionales que controlan por género, región, estado civil, estrato socioeconómico, nivel del Sisbén, del lado de las características del individuo, pero también la rama productiva de la empresa donde el individuo está actualmente empleado, el tamaño de la firma en la que el individuo está vinculado, el lugar

¹² S es continua, por lo cual se especifican dos casos: $S > 0$ y $S = 0$.

(ciudad o región) donde la persona está empleada o el *inter alia*. Desde estas consideraciones, el modelo de evaluación queda así:

$$\ln y = \beta_0 + rS + \gamma_1 X + \gamma_2 X^2 + \beta_1 T + \beta_2 [S \times T] + \beta_3 TA + \beta_4 [T \times TA] + \sum_{j=1}^n 1\delta_j X_j + \mu \quad [7]$$

Estrategias de estimación y contrafácticos

En ausencia de regresores adicionales, la estimación de [3] da lugar a estimadores menos eficientes que los que podrían obtenerse de [7], aun cuando el problema del resultado contrafáctico pudiera haber sido resuelto de manera exitosa, en observación de las recomendaciones habituales sobre este particular (Bernal y Peña, 2011, pp. 59 y ss.; Ravallion, 2008, p. 1154; Khandker, 2010, pp. 33 y ss.). Cuando el contrafáctico es adecuado, la estimación de [3] puede ser mejorada en términos de eficiencia y potencia, al adoptar una especificación como [7], que introduce una variedad amplia de regresores para incrementar el poder explicativo del modelo básico de Mincer.

La asignación de individuos en el caso de la formación vocacional y del CA no es aleatoria. De este modo, la conformación de un grupo de control adecuado no es requerimiento inocuo. Gertler *et al.* (2011) identifican esta búsqueda con la de encontrar un *clon perfecto* para los individuos en el grupo de tratamiento, habida cuenta de que una selección inadecuada redundaría en que los estimadores sean probablemente sesgados y los resultados de la evaluación sean inútiles.

Trabajos como el de Kuckulenz y Zwick (2003), o los de Estacio *et al.* (2010) y de Medina y Núñez (2005) para el caso colombiano, adoptan procedimientos estándares para resolver el problema. Estrategias como la estimación dietápica vía MCO (variable instrumental), o bien, aquella de corrección de Heckman, que pretenden minimizar la influencia de no observables en la generación de sesgo, son procedimientos típicos que tienen que formar parte del *toolkit* del evaluador y deben

considerarse complementos antes que sustitutos. La aplicación de técnicas de emparejamiento, como los de la familia de los *propensity scores*, se usan en forma extensiva en combinación con diversas estrategias computacionales en la mayoría de los ejercicios de evaluación de resultados de un tratamiento.

Intuitivamente, considere la estimación de [7] condicionada a un conjunto de variables de tratamiento T^* (o T_A^* , según sea el caso) que se obtiene de la estimación del (los) indicador(es) de tratamiento en regresores que pueden explicar la variable de resultado, así como el (los) indicador(es) de tratamiento. Considere entonces la siguiente relación:

$$T^* = \varphi(Z\gamma) + u > 0 \quad [8]$$

La pieza central en la expresión [8] es una función lineal (no lineal) que relaciona Z , que es una matriz de características individuales que no (necesariamente) están en X y que contribuyen a determinar si un individuo entra o no en el programa; por su parte, γ es un vector de parámetros. Si $\varphi(\cdot)$ es combinación lineal del vector γ y las columnas de Z , una aproximación estadística apropiada puede obtenerse vía mínimos cuadrados; si, por contraste $\varphi(\cdot)$, es una función de distribución acumulativa normal típica¹³, la función [8] puede usarse como un predictor de la propensión a entrar al *tratamiento* y corregir preventivamente el sesgo a través de la estimación de *propensity scores*. La elección de un instrumento adecuado (Z) es difícil, dada la precariedad del registro del individuo en la institución que ofrece los programas de formación y la ausencia total de información socioeconómica en la PILA. Se tendrá en cuenta, sin embargo, que los individuos bajo análisis son típicamente de bajos ingresos *antes* del tratamiento, que es un periodo en el que estas personas no

13 Es posible obtener versiones semiparamétricas del emparejamiento si la imposición de condiciones distribucionales (como, precisamente, la normalidad) resulta muy fuerte o no deseable. Ver, por ejemplo, el trabajo de Hahn (1998).

deben reportar pagos a la seguridad social, dado además el grupo etario al que deben pertenecer.

La condición económica pretratamiento puede definirse a partir de la vinculación al Sistema General de Seguridad Social en Salud (SGSSS), que debe corresponder al régimen subsidiado en el periodo antes del programa; es posible que haya individuos que pertenezcan al régimen contributivo e incluso que, estando en este régimen, aparezcan a título de aportante —caso en el cual el individuo se entenderá como *obligado*, dado que tiene capacidad de pago—. Sin embargo, la condición de estar en el régimen contributivo, que implica no pertenecer al Sisbén I o II y no tener capacidad de pago, está hablando del estatus socioeconómico del hogar del beneficiario y está necesariamente incorrelacionado con el error en [7]. Note que el puntaje de Sisbén no es requisito suficiente, pues se puede estar en los menores niveles y aun tener capacidad de pago, caso en el cual no es posible encontrar al individuo en el régimen subsidiado. El registro del SENA no contiene información confiable para construir esta variable, por lo que la alternativa consiste en obtener la variable “tipo de afiliación al SGSSS” directamente de la Base de Datus Única de Afiliados (BDUA) del Ministerio de Salud.

Con estas consideraciones, sea \hat{T}^* la predicción MCO no restringida de T^* bajo [8]; este resultado, que es variable continua, requiere ser transformada para reemplazar T (y/o T_A), que es binaria, en [7]. Como el valor no restringido de $\hat{T}^* \in [-\infty, \infty]$, una transformación adecuada del resultado es la aplicación de la regla¹⁴:

$$[T = 1 \leftrightarrow \hat{T}^* > 0] \wedge [T = 0 \leftrightarrow \hat{T}^* \leq 0]$$

14 Esto, obviamente, si no se imponen restricciones sobre la probabilidad de T tales que $T \in [0, 1]$; esto es, si se estima un modelo de probabilidad lineal básico, por ejemplo. A diferencia del modelo lineal no restringido [24], es fácil obtener directamente las probabilidades de que $T = 1$, o bien, de que $T = 0$ mediante procedimientos como la aplicación de Heckman (Heckit), o bien, mediante técnicas de emparejamiento estándar o *propensity scores*. Las opciones son amplias y, en presencia de una base de datos rica y diversa, deberán ser exploradas.

Caso en el cual el modelo transformado resultará ser:

$$\ln y = \beta_0 + rS + \gamma_1 X + \gamma_2 X^2 + \beta_1 \hat{T}^* + \beta_2 [S \times \hat{T}^*] + \beta_3 \hat{T}_A^* + \beta_4 [\hat{T}^* \times \hat{T}_A^*] + \sum_{j=1}^n \delta_j X_j + \mu \quad [9]$$

El acervo de información de que se puede hacer uso en el plazo inmediato comprende las tablas de aportes al Sistema de Seguridad Social, la PILA y el registro de individuos inscritos a programas de formación titulada del SENA. Con estos datos se abre la posibilidad de reemplazar $[y_{i1} | T = 0]$ por $[y_{i0} | T = 0]$, tomando como contrafácticos los resultados de los aspirantes no admitidos, o bien, construir resultados contrafácticos con base en información de los tratados antes del tratamiento. Bajo la primera modalidad de construcción se utiliza como resultado contrafáctico el conjunto de individuos elegibles que no fueron efectivamente elegidos y, por tanto, no fueron tratados.

El registro del SENA distingue a los *aprendices* —o individuos inscritos en todo caso— según su “estado” en relación con un programa de formación al cual el sujeto ha aplicado en algún momento de su vida. De acuerdo con esta categorización, los individuos registrados en esta tabla pueden ser tenidos en cuenta como *admitidos* o *no admitidos*. A su vez, los individuos admitidos pueden exhibir el estado de “certificados” o “titulados”, caso en el cual $T = 1$; los individuos para los cuales $T = 1$ tendrán a su vez $T_A = 1$ si su modalidad de fase práctica ha sido CA, y $T_A = 0$ en caso contrario; para los no admitidos, $T = T_A = 0$ y la incorporación de la componente $\beta_4 [T \times T_A]$ en [25] resulta apenas obvia.

La no aleatoriedad del proceso de asignación significa que la intención de los individuos de vincularse a un programa de formación y de optar eventualmente por un contrato de aprendizaje puede relacionarse con un conjunto de características que pueden ser compartidas entre los individuos de los grupos de tratamiento y de control; al mismo tiempo, el resultado de su asignación definitiva puede corresponder a otra serie de eventos (pruebas de acceso, resultados

de políticas institucionales de priorización, etc.) que determinarán la asignación de los individuos ya sea al grupo de tratamiento o al de control. Se enfatiza que el ignorar estas realidades conlleva resultados inciertos e incluso inútiles, y que es tarea de investigación resolverlo en relación con dicho *counterfeit counterfactual*, para limitar o eliminar la influencia negativa del sesgo (Khandker, Koolwal y Samad, 2010, p. 47).

De los resultados de la selección del grupo de control depende la calidad del resultado. Además, el procedimiento de estimación habrá de variar según la estrategia de medición de impacto asociada, porque la construcción inicial del contrafáctico no elimina el riesgo de sesgo, en la medida en que los residuales siempre podrán contener información no registrada en las bases de datos. Se entiende que la estimación no se limita a ejecutar una rutina computacional disponible en el *software*, sino que obliga a efectuar el conjunto de procedimientos de indagación necesario para comprobar las hipótesis relevantes, no únicamente en relación con el valor del impacto *per se*, sino, en particular, en lo atinente a la verificación de postulados relativos a la pertinencia de la selección de los elementos que se contienen en la tabla de datos de los individuos del grupo de control.

En efecto, se puede estimar el impacto del tratamiento cuando se tiene un contrafáctico estructurado con base en el uso de los datos de los no admitidos, que es una de las alternativas con que se puede contar; en este caso, el modelo básico de diferencias hace posible estudiar la relación *con-sin* entre los individuos observados. No es improbable, sin embargo, que por razones de diversa índole —entre ellas, la deficiencia metodológica en el registro— la elección del contrafáctico propuesto no sea útil o no contenga mayor variabilidad. Este caso, determinado por la naturaleza de los datos, obliga a revisar la estrategia de construcción del contrafáctico y, quizás, a adoptar otro enfoque, como el de efectuar comparaciones entre los resultados *pre- y posprograma* de los participantes, que

en la literatura se denomina con frecuencia *análisis reflexivo* (Khandker, Koolwal y Samad, 2010, p. 23). Este es el segundo tipo de contrafáctico que se propone. La vinculación de las tablas del SENA y del registro de la PILA abren la posibilidad de contar con líneas base a partir de las cuales se puede construir el panel de datos que se requiere para esta clase de análisis.

Sea $(y_i^t | T_i)$ el valor de la variable de resultado del individuo i dado su estatus de participación T_i , para $t = 1, 2$ periodos que representan los momentos *previo* y *posterior* al proceso de formación. Así, bajo la segunda modalidad de construcción de contrafáctico, la variable de resultado podrá adoptar los valores que se presentan en la siguiente tabla.

Tabla 1.

Valores de la variable de resultado (segundo tipo de contrafáctico)

	Tratamiento	Control
$t - 1$	y_{t2}^1	y_{t0}^1
$t - 2$	y_{t1}^2	y_{t0}^2

Fuente: elaboración del autor a partir de Bernal y Peña (2011, p. 72).

El modelo de diferencias en diferencias (DD, o *double difference*), utilizado, por ejemplo, en la investigación de Estacio *et al.* (2010) para Fedesarrollo, estima el impacto del tratamiento, entendido como “el cambio esperado en la variable de resultado entre el periodo posterior y el periodo anterior a la implementación del tratamiento, menos la diferencia esperada en la variable de resultado en el grupo de control durante el mismo periodo” (Bernal y Peña, 2011, p. 72). Note que en el modelo DD, el grupo de control es el grupo de tratamiento *antes del tratamiento*. Bajo el modelo DD, y en términos esquemáticos, el impacto del programa es:

$$\tau[DD] = (E[y_{t1}^2] - E[y_{t1}^1]) - (E[y_{t0}^2] - E[y_{t0}^1])$$

Mientras que en términos muestrales es este:

$$\hat{t}_{[DD]} = \left(\left[\frac{1}{m} \sum_{i=1}^m y_i^2 | T = 1 \right] - \left[\frac{1}{m} \sum_{i=1}^m y_i^1 | T = 1 \right] \right) - \left(\left[\frac{1}{m} \sum_{i=1}^m y_i^2 | T = 0 \right] - \left[\frac{1}{m} \sum_{i=1}^m y_i^1 | T = 0 \right] \right)$$

$$\hat{t}[DD] = [\Delta \bar{y} | T=1] - [\Delta \bar{y} | T=0] \quad [10]$$

Si para la media condicional de y_i se postula un modelo lineal, se tiene:

$$y_i = \beta_0 + \beta_1 T_i + \mu_i \quad [11]$$

El estimador de diferencias en diferencias $\hat{t}[DD]$ es β_1 en:

$$y_i^2 - y_i^1 = \beta_0 + \beta_1 T_i + (u_i^2 - u_i^1)$$

O sea:

$$\Delta y_i^t = \beta_0 + \beta_1 T_i + v_i \quad [12]$$

El estimador de β_1 será insesgado siempre que pueda mantenerse la hipótesis de independencia condicional¹⁵. No debe dejarse de notar que con los datos actuales se puede incluir dentro del grupo de control el conjunto de individuos con los cuales se construye el grupo de control del primer tipo de contrafactual propuesto; por consiguiente, en lugar de dos, hay en realidad tres esquemas de medición por explorar.

Desde la estrategia DD, la ecuación de evaluación puede ser estimada convenientemente bajo OLS, siempre que se pueda contar con pares ordenados en $t = 0$ (antes del tratamiento) y en $t = 1$ (luego del tratamiento), lo cual es una posibilidad objetiva, porque la PILA contiene información sobre ingresos laborales a lo largo del tiempo; además, ofrece la posibilidad de obtener controles como la actividad económica del aportante (CIU) y una

medición de su tamaño, en la medida en que es posible conocer el número de cotizantes que tiene a su cargo.

Si, por contraste, como resulta en el caso de Westergård-Nielsen y Rasmussen (1997), el impacto se mide sobre la ocupabilidad, la variable de impacto y_i es sustituida por el estatus de ocupación del individuo (ocupado/no ocupado) y la relación propuesta se estima mediante un *probit* truncado o *Tobit*.

En efecto, aun cuando la mayor parte de la discusión se ha referido al ingreso laboral como la variable de impacto, esta puede ser reemplazada por algún indicador de ocupabilidad, o bien, por algún indicador como el ICE sobre calidad del empleo, utilizado por Estacio *et al.* (2010) y otros investigadores. Por ejemplo, tomando datos PILA, la ocupabilidad puede introducirse como el *ratio* de periodos en los que el sujeto no ha aportado, a número total de periodos transcurridos desde que obtuvo su primer trabajo formal. Esta última variable es incluida en la investigación de Medina y Núñez (2005).

En cualquiera de los casos posibles, el tratamiento principal es el CA, dado que el sujeto aparece vinculado a un proceso de formación titulada. Una serie de regresores adicionales deberá ser incorporada, esto es, características observadas/no observadas que determinan la participación de algunos individuos en el programa. La hipótesis que se quiere comprobar es si el (los) tratamiento(s) tiene(n) efectos positivos sobre la variable de resultado, cualquiera que sea su construcción. Se hace de nuevo hincapié en que las manifestaciones empíricas del tratamiento están limitadas por los datos disponibles y que, dados los datos compilados, siempre será necesario depurar los contenidos disponibles para concretar el contrafactual correcto.

15 El tratamiento del modelo DD con regresores adicionales es tratado en detalle en Bernal y Peña (2011, pp. 77 y ss.) y en Wooldridge (2010, 603 y ss.).

DATOS

Toda estrategia de estimación del impacto del CA está limitada por los datos disponibles. Se presenta la estrategia para la construcción de la base de datos mínima necesaria para la estimación de funciones de retorno entre los ingresos y la empleabilidad del aprendiz bajo CA, así como una serie de regresores de interés que obrarán como controles para distinguir impactos atribuibles a otras causas distintas al CA.

Bernal y Peña (2011) plantean como paso previo a la evaluación de impacto la necesidad de adelantar análisis factuales y de contexto; esto es, el análisis y estudio de las condiciones socioeconómicas e institucionales en las que el programa tiene lugar y se desarrolla, que incluye estimaciones del número de beneficiarios y la cobertura del programa. Se han venido adelantando estudios de seguimiento sobre este particular en el interior de la DMFT que puntualizan no únicamente sobre estas particularidades cuantitativas, sino que compilan el conjunto de disposiciones jurídicas que dan forma al CA en Colombia.

En relación con el análisis de contexto, se ha venido avanzando en la compilación sobre la experiencia mundial y nacional en este tipo de programas; la revisión de literatura al comienzo de esta propuesta tiene también ese propósito. Es en todo caso razonable analizar otras fuentes de datos que puedan servir a la caracterización factual y de contexto del programa, y que aprovechan la riqueza informacional de investigaciones continuadas, como las de las encuestas de calidad de vida y de hogares del Departamento Administrativo Nacional de Estadística (DANE).

Justamente, información complementaria y de contexto es la que puede ser obtenida de las encuestas de hogares actuales. La Gran Encuesta Integrada de Hogares (GEIH) es la última evolución de las encuestas que el DANE utiliza para la producción de información estadística sobre el mercado laboral en Colombia.

La GEIH sirve al propósito de la evaluación de impacto del CA de manera general. Contiene dos

componentes de interés: el capítulo de educación (Sección "G" en el formulario actual), que incluye preguntas sobre el mayor nivel educativo alcanzado (pregunta 1), y el título o diploma de mayor nivel obtenido (pregunta 5), en la cual aparece la opción "técnico" o "tecnológico". La GEIH también contiene un módulo sobre empleo (Sección "H"), que, además de indagar acerca del estatus ocupacional (ocupado/no ocupado) (pregunta 2), lo hace en torno las razones de la no ocupación (pregunta 9), que incluyen como respuestas posibles "b.- No existe empleo para mi oficio o profesión", "f.- Carece de la experiencia necesaria" y "h.- Es muy joven o muy viejo". La sección "I" que está dedicada a obtener información sobre ocupados e indaga (pregunta 14) acerca de cuánto ganó (antes de impuestos) en el trabajo reportado.

Esta información habilita la estimación de una ecuación de ingresos de Mincer, en la que si bien, de acuerdo con la información supra, es posible incluir como variable de tratamiento el hecho de que el encuestado alcanzó un nivel de educación técnico o tecnológico, no es posible introducir información sobre CA en forma explícita. La variable de ingresos está disponible en la GEIH y quizás pueda hacerse algún esfuerzo para caracterizar la ocupabilidad con base en la información relativa a la pregunta sobre búsqueda de empleo en la semana previa a la encuesta, así como con base en los datos relativos a la estructura económica de la población, en especial en lo que se refiere a ocupación. Sin embargo, para responder a la pregunta de si el CA supone un *premium* sobre las demás modalidades de formación para el trabajo, se requiere contar con otras fuentes cuya estructura y naturaleza se reseñan a continuación.

La principal fuente de información individual sobre formación para el trabajo en Colombia es el Sistema Optimizado para la Formación Integral del Aprendizaje (SOFÍA), que viene siendo operando desde 2010 por el SENA. Dicho sistema contiene información sobre todos los individuos que han pasado por el SENA, tanto al nivel mínimo de la inscripción como al nivel de egresado. Todos aquellos individuos que aspiraron a formarse en el

SENA tienen una entrada en el sistema, así como el estatus del individuo respecto de su aspiración, es decir, si culminó el proceso, si desertó o si definitivamente no fue admitido. Además de información socioeconómica mínima como el género, el estrato socioeconómico y la calificación en Sisbén, hay información sobre la modalidad elegida por el aprendiz para culminar la fase práctica de su proceso de formación, de acuerdo con el *Manual de procesos y procedimientos* de dicha institución. Los análisis adelantados muestran que estas variables, sin embargo, presentan alta dispersión y ponen de manifiesto alguna clase de problemas metodológicos en el registro.

SOFÍA, por otra parte, no registra ni contiene información de resultado o de impacto. No hace posible el seguimiento de la carrera laboral del individuo y, por tanto, no permite validar la bondad relativa de los programas de formación en relación con los objetivos generales propuestos en el marco jurídico relevante. Los datos sobre la empresa en la cual el aprendiz desarrolla la fase práctica del CA se registran en un sistema aparte, a cargo de la División de Relaciones Corporativas del SENA. Así, se consolida una situación en la cual se cuenta con dos sistemas de datos que, si bien no conversan entre ellos, sí tienen un identificador común: el aprendiz.

Se llama la atención sobre el hecho de que en SOFÍA no se incluyen las variables de resultado propuestas ni regresores claves de la ecuación de ingreso de Mincer, como la experiencia que habría que construir, según la regla provista, con base en la escolaridad del individuo *al momento de la radicación de su aplicación* a un programa de formación titulada. SOFÍA no tiene identificadores de firma con los cuales sea posible aproximarse al tamaño de esta y a la rama de la producción en la que opera; sin embargo, hay un registro a cargo de la División de Promoción y Relaciones Institucionales del SENA que incluye la identificación de la firma donde el CA tuvo lugar.

En resumen, las fuentes de datos del SENA permiten identificar aprendices con y sin CA que pueden caracterizarse por su edad, género, nivel

de educación y alguna otra caracterización para la priorización de acuerdo con la política de asignación del SENA. Una variable como “Tipo de población”, que tiene entradas como “Jóvenes emprendedores”, “Población negra desplazada” o “Jóvenes participantes en conflicto armado”, constituye categorías por tener en cuenta como identificadores concretos en la estimación de [8] y [9].

La información sobre ingresos o empleabilidad no está disponible en SOFÍA, que no prevé el seguimiento de sus aprendices. Por otra parte, aun cuando el Observatorio Laboral del Ministerio de Educación tenga previsto un módulo para el seguimiento de individuos con máximo nivel de educación alcanzado, sea técnico o tecnólogo, hasta la fecha es un proyecto sin resultados utilizables.

La alternativa contemplada para construir la variable de impacto en el LHS de la ecuación de Mincer consiste en usar la información contenida en el Sistema Integrado de Pagos de Aportes a la Seguridad Social, cuyo instrumento principal es la PILA. El sistema registra los pagos que cada trabajador en cada organización empleadora hace por concepto de pensiones, salud y riesgos profesionales o laborales. Como cada cotizante (trabajador) es responsable junto con su empleador de estos aportes (que deben efectuarse mensualmente), el resultado es una base de datos de varios millones de registros, en los cuales, para cada trabajador, se registran los aportes de Ley.

La Resolución 1747 de 2008, del Ministerio de la Protección Social, posteriormente modificada por la Resolución 643 del 2010, de este mismo Ministerio, adopta el diseño y contenido de la PILA. De acuerdo con estas disposiciones, hay tres tipos de registro: el tipo I, que contiene los datos generales del aportante; el tipo II, que “incluye toda la información completa del formulario para todo el Sistema de Seguridad Social integral y aportes parafiscales asociados a la nómina”; y el tipo III, que contiene la liquidación detallada de los aportes al sistema.

Asociado al registro del tipo III, hay en PILA una tabla en la cual, para cada *cotizante* —esto es,

para cada individuo—, hay información mensual sobre los aportes a pensión, salud, sistema de riesgos profesionales, SENA e ICBF. Estos aportes son pecuniarios en naturaleza, y dado el ingreso base de cotización y el tipo de cotizante (dependiente o independiente), se puede obtener el valor concreto del ingreso del trabajador por mes.

Se plantea enlazar la base de información a nivel individual sobre aprendices del SENA con los registros existentes para los mismos individuos en la PILA, que deberían tener un resultado concreto y efectivo, pues, de acuerdo con las disposiciones en la Ley 789 de 2002, el aprendiz deberá estar afiliado en la fase lectiva al Sistema General de Seguridad Social en Salud (SGSSS) y en la fase práctica al Sistema de Riesgos Laborales (en adición). En el Sistema Integrado de Pagos de Aportes a la Seguridad Social se distinguen distintos tipos de cotizantes; entre ellos hay dos categorías del mayor interés: “Aprendiz del SENA en etapa lectiva” y “Aprendiz del SENA en etapa productiva”, que actúan como identificadores adicionales a los básicos para cada aprendiz, haya este optado o no por el CA. Con la llave *tipo de documento* y *número de documento*, la tabla resultante de la unión entre el registro del SENA y los aportes a seguridad social debe ser adicionada con información sobre estatus de afiliación al SGSSS proveniente de la BDUA, según se discute en la sección precedente.

La ocupabilidad del individuo y la movilidad pueden ser objeto de seguimiento, porque el sistema incluye variables que señalan la movilidad intra- o interfirma del individuo, que son: a) las novedades relacionadas con variaciones del salario, sean estas permanentes o transitorias, y b) las novedades asociadas con el cambio de aportante. Un reto de mayor nivel tiene que ver con la síntesis de la información que para cada individuo aparece en PILA, pues la historia de cada sujeto es variable. Cada individuo tiene una tabla de datos que registra por cada aportante—incluido él mismo, si ha tenido periodos de trabajador por cuenta propia—, que debe ser resumida de acuerdo con algún criterio asociado a la estrategia de estimación que la investigación requiera poner en ejercicio.

La estrategia involucra la construcción de un registro individual de características, incluidos ingreso y otras posibles variables de impacto como la ocupabilidad. La construcción se debe concebir como la combinación de la información de los individuos registrados en la base de datos de inscritos del SENA-SOFIA, con la información correspondiente a cada uno de los individuos detectados en la base PILA.

Las dimensiones de las tablas por relacionar son de interés. En primer lugar, las tablas del SENA contienen un total de 6821 millones de entradas entre 2004 y 2015¹⁶; por su parte, a abril de 2016 se cuenta con un total de 11,6 millones de cotizantes únicos para los cuales es posible computar los ingresos agregados por año, hasta 2015, así como los niveles mediano, máximo y mínimo, y el número de periodos cotizados por año. La PILA comienza a institucionalizarse en 2008, y la información que aparece en ella hasta 2011 es regular y debería ser completada con otras fuentes de pago que no parece útil reunir.

En consecuencia, se plantea construir la unión PILA-SENA para los años 2012 a 2015, con frecuencia anual, con lo cual podrían compilarse cerca de 1,7 millones de registros, es decir, cerca de 450.000 individuos por año. De este total, cerca del 24% son “certificados”, en tanto que el resto pueden ser tenidos en cuenta como “no admitidos”. Al tener en cuenta que el trabajo de Estacio *et al.* (2010) se desarrolla con base en una encuesta diseñada (100 encuestas en total) y levantada para el propósito específico del estudio, o que los trabajos de Jiménez y Kugler (1988) y de Medina y Núñez (2005) usan como insumo los registros de las encuestas de hogares—que se limitan a algunos miles de registros en el menor de los casos—, resulta más que evidente el potencial de la base estadística que se proyecta consolidar.

En todo caso, es necesario tomar en forma conservadora esta perspectiva. En primer lugar, la forma como el SENA registra la información,

16 Preliminar, según División de Relaciones Corporativas del SENA (abril del 2016).

su política de registro y mantenimiento de información, que parece girar alrededor de SOFÍA y SOFÍA-Plus, a pesar de la prensa y divulgación que disfruta, puede y debe ser objeto de revisión: esta es un diagnóstico pertinente a la fecha (abril de 2016). En segundo lugar, los cruces planteados entre las fuentes propuestas (SENA-PILA-RUA-BDUA) requieren de un aparato institucional que involucra al Ministerio de Salud, del cual es responsabilidad la información sobre aportes a las Seguridad Social; al DNP, que es responsable del Sisbén y de las calificaciones que determinan los cortes para los niveles I y II y, por consiguiente, de los contenidos de la BDUA; del Ministerio del Trabajo, que es corresponsable de la política de formación para el trabajo; y, naturalmente, del SENA, que debe proporcionar seguridad y veracidad sobre la información aportada en el caso de los aprendices.

Las restricciones institucionales y técnicas al uso de las tablas señaladas deberá ser, desde luego, temporal, porque la información debe reconstruirse como condición necesaria no solamente para adelantar la evaluación presente, sino para el seguimiento de los resultados de la política de formación para el trabajo en general y del CA en particular.

COMENTARIOS FINALES

A lo largo de las secciones previas se ha documentado una propuesta metodológica para la evaluación de impacto del CA en Colombia. La propuesta se estructura alrededor de la construcción de una base de datos masiva que comprende información sobre variables de resultado como el ingreso salarial y la ocupabilidad, y características individuales de un conjunto apreciable de personas que, entre 2004 y 2015, han aplicado, con éxito o sin él, a programas de formación del SENA, y que han concluido estos programas, o que por razones diversas no terminaron su proceso de formación.

Con la evaluación de impacto propuesta se quiere indagar si la formación vocacional en general y si el CA en particular contribuye al mejoramiento de las condiciones laborales de los individuos considerados, que incluyen mejores ingresos

y una vida laboral más estable, es decir, una vida laboral con menos episodios de desempleo que en aquellos casos en los que la formación vocacional no tuvo lugar.

De la revisión de literatura que se ha adelantado, resulta que invariablemente, dentro del conjunto de estrategias de formación vocacional, los programas de formación *dual* representan ventajas para los individuos que pueden disfrutar del *tratamiento*. La razón es que si bien la formación para el trabajo implica retornos al capital humano mayores para quienes han disfrutado de este tipo de educación, el hecho de que los *apprenticeships* sometan al individuo al contacto directo de la experiencia a través de la práctica supone un *premium* adicional asociado a la práctica en ambientes de aprendizaje que enfrentan al aprendiz con el día a día de las actividades fabriles de los empleadores.

La revisión de literatura no se limita a la descripción de resultados generales de políticas nacionales de formación dual en un conjunto de países seleccionados por su experiencia en programas de esta clase. Las fuentes revisadas tienen además en común que desarrollan ejercicios de medición que, en forma más que característica, acuden a la estimación econométrica de ecuaciones de ingreso *à la* Mincer. Dicho dispositivo econométrico, además de constituir un instrumento de medición flexible y eficaz desde el punto de vista computacional, también disfruta de un sólido fundamento lógico en toda la teoría neoclásica del capital humano. No sorprende, en consecuencia, que trabajos que desarrollan postulados abstractos sobre las ventajas de las prácticas, como los de Ashenfelter (1978), Acemoglu y Pishke (1998) o aquel más reciente de Heckman, Lalonde y Smith (1999), contengan contrapartes empíricas que acuden a la mencionada ecuación.

Las evaluaciones de impacto constituyen una subdisciplina de medición socioeconómica que parece atractivo poner en práctica con el problema que se nos presenta. Sin embargo, hay una serie de requerimientos por satisfacer para lograr que los resultados sean válidos, esto es, para lograr que no adolezcan de sesgo de selección. Se ha mostrado cómo el reto principal: la construcción

de los resultados contrafactuales, no es una labor inocua que pueda tomarse con ligereza y que obliga a echar mano de las diversas técnicas que se han desarrollado en aquel marco metodológico. Los datos disponibles abren la posibilidad de estudiar alternativas como la estimación del modelo básico de diferencias, al igual que el modelo de las diferencias en diferencias, es decir, la posibilidad de verificar los resultados “con-sin” y “antes-después” habituales.

No obstante, aunque los datos puedan permitir la construcción de bases de información con registros para individuos tratados y no tratados, se advierte que la caracterización de los sujetos en cada grupo no se limita a seleccionar individuos admitidos o no admitidos, sino que exige implementar procesos de depuración adicionales como, entre otros, estimar la probabilidad de ser tratado con base en los regresores a mano, así como a partir de un conjunto de instrumentos que deben satisfacer requerimientos específicos bien definidos.

Los registros del SENA son de gran amplitud, pero contienen errores atribuibles a la disciplina del registro en esa institución. La PILA, si bien es un documento público que recoge pagos al Sistema de Seguridad Social y que, dadas las condiciones jurídicas, implican un incentivo para que quienes diligencian el documento introduzcan información veraz y correcta, también puede contener errores atribuibles probablemente a cuestiones de digitación y, en alguna medida, a conductas estratégicas de parte del contribuyente: inexactitud y mora en las contribuciones sociales.

Las dos fuentes se articulan a través del documento de identificación del individuo. Hay problemas potenciales por considerar: en ausencia de depuración adicional, es de esperar que haya un subconjunto de individuos para los cuales pueden presentarse duplicidades. Estas pueden surgir cuando un sujeto pasa a ser mayor de edad —y, por tanto, su documento de identidad se modifica—, o bien, cuando para el mismo individuo resulten varias entradas, al pasar del estatus de *no admitido* en un primer momento al de *admitido* en un periodo posterior. Hay múltiples posibilidades, desde luego: el *no admitido* pudo haber dejado de lado

su intención o pudo haber insistido y ser admitido finalmente; el *admitido* pudo haberse certificado o pudo haber desertado o aplazado.

La ecuación de Mincer, según se ha argüido, es un resultado que considera el trabajo como un bien compuesto de características hedónicas, esto es, hay consideraciones de equilibrio general competitivo por tener en cuenta: incluir en la ecuación de Mincer características de los empleadores como los que aparecen en PILA, que registran la rama de actividad y que permiten construir un indicador de tamaño de la firma, hace posible contrastar la hipótesis de que el impacto de las prácticas en los lugares de producción de las firmas depende de características del aprendiz y de la naturaleza de su entrenamiento, dadas las condiciones del lugar de trabajo.

A pesar de dicho resultado, el beneficio neto de los programas duales no puede ser estimado en forma explícita, como sí resulta en la investigación de Mohrenweiser (2010). Este tipo de iniciativas solo es posible si se dispone de datos privados sobre costo/beneficio; sin embargo, Colombia no los tiene. Se puede argüir que el registro de la Superintendencia de Sociedades puede ser útil; sin embargo, aun cuando este pueda incluir pagos por concepto de salarios, no discrimina el tipo de trabajadores ni identifica o discrimina entre aprendices y trabajadores seculares.

Al mismo tiempo, puede resultar de interés tratar de aproximar el impacto macroeconómico del CA. Sin embargo, si se tiene en cuenta que el número de CA que agencia el SENA no supera los 200.000 individuos por año, en un escenario en el cual en el ámbito nacional el número de ocupados asciende a unos 22 millones de individuos, puede resultar que el impacto macro sea de orden menor, en la medida en que probablemente no contribuye de modo sustancial a la reducción del desempleo agregado, por ejemplo.

La pregunta que hay que responder, en nuestra opinión, no es si una estrategia de política activa de empleo como el CA contribuye a reducir el desempleo global, sino, más bien, si con el CA, que tiene un alcance limitado, se mejoran las condiciones de las personas que son sujeto de este

tratamiento y si, dado su éxito o fracaso relativo, es socialmente adecuado, responsable y rentable mantener y ampliar su alcance, o si, por el contrario, resulta procedente reevaluar su pertinencia,

en consideración de otras alternativas de formación posibles, que también deberán ser sujeto de evaluación *ex ante* y *ex post*, como un requerimiento para afinar la política general de formación para el trabajo en el país.

REFERENCIAS

1. Acemoğlu, D. y Pishke, J.-S. (1998). Why do firms train? Theory and evidence. *The Quarterly Journal of Economics*, 113, 79-119.
2. Acemoğlu, D. y Robinson, J. (2013). *Why nations fail*. Boston: Crown Business.
3. Arango, L. E., Gómez, M. y Posada, C. E. (2010). La demanda de trabajo formal en Colombia: determinantes e implicaciones de política. En M. Jalil, L. Mahadeva, M. Jalil y L. Mahadeva (Eds.), *Mecanismos de transmisión de la política monetaria en Colombia* (pp. 607-642). Bogotá: Banco de la República.
4. Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *The Review of Economics and Statistics*, 60(1), 47-57.
5. Bagehot (2011, 16 de noviembre). Britains labour market. The british obsession with apprenticeships. *The Economist*. Recuperado de <http://www.economist.com/blogs/bagehot/2011/11/britains-labour-market>
6. Banco Mundial (2012). *World development report 2013: Jobs*. Washington: Autor.
7. Becker, G. S. (1964). *Human capital. A theoretical and empirical analysis, with special reference to education*. Chicago: Universidad de Chicago.
8. Bernal, R. y Peña, X. (2011). *Guía práctica para la evaluación de Impacto*. Bogotá: Uniandes.
9. Borjas, G. (2012). *Labor economics*. Nueva York: McGraw-Hill.
10. Brunello, G. (2009). *The effect of economic downturns on apprenticeships and initial workplace training: a review of the evidence*. Bonn: IZA.
11. Cahuc, P. y Zylberberg, A. (2004). *Labor economics*. Nueva York: MIT Press.
12. Congreso de la República de Colombia (2002, 27 de diciembre). Ley 789 de 2002, Por la cual se dictan normas para apoyar el empleo y ampliar la protección social y se modifican algunos artículos del Código Sustantivo de Trabajo. *Diario Oficial* 45.046.
13. Departamento Administrativo Nacional de Estadística (DANE) (2012, 11 de marzo). Fuerza laboral y educación (boletín de prensa). Recuperado de http://www.dane.gov.co/files/investigaciones/boletines/especiales/educacion/Bol_edu_2012.pdf
14. Departamento Administrativo Nacional de Estadística (DANE) (2013a, 10 de julio). Mercado Laboral de la Juventud (14 a 28 años), trimestre móvil marzo-mayo de 2013 (Resumen ejecutivo). Recuperado de http://www.dane.gov.co/files/investigaciones/boletines/ech/juventud/RE_trim_mar_may13.pdf
15. Departamento Administrativo Nacional de Estadística (DANE) (2013b). *Ficha de metodología Gran Encuesta Integrada de Hogares*. Bogotá: Autor.
16. Departamento Administrativo Nacional de Estadística (DANE) (2013c). *Resumen Ejecutivo. Mercado Laboral de la Juventud (14 a 28 años). trimestre móvil marzo - mayo de 2013*. Recuperado de <http://www.dane.gov.co>: http://www.dane.gov.co/files/investigaciones/boletines/ech/juventud/RE_trim_mar_may13.pdf

17. Estacio, A., Millán, N., Olivera, M., Parra Torrado, M. y Santa María, M. (2010). *Evaluación de impacto de dos programas de formación del SENA: formación de técnicos y tecnólogos y formación especializada del recurso humano vinculado a las empresas*. Bogotá: Fedesarrollo.
18. Gertler, P., Martínez, S., Premand, P., Rawlings, L. B. y Vermeersch, C. M. (2011). *La Evaluación de impacto en la práctica*. Washington: Banco Mundial.
19. Ginsburgh, V. y Keyzer, M. (1997). *The structure of applied general equilibrium models*. Cambridge: MIT Press.
20. Gunderson, M. y Harry, K. (2011). *Returns to apprenticeship: analysis based on the 2006 census*. Toronto: University of Toronto.
21. Hahn, J. (1998). On the role of the Propensity Scores in efficient semiparametric estimation of average treatment effects. *Econometrica*, 66(2), 315-331.
22. Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, 54(1), 3-56.
23. Heckman, J. J., Lalonde, R. J. y Smith, J. A. (1999). The economics and econometrics of active labor market Programs. En O. C. Ashenfelter y D. Card (Eds.), *Handbook of labor economics* (Vol. 3A, pp. 1865-2097). Amsterdam: Elsevier.
24. Heckman, J. J., Lochner, L. J. y Todd, P. E. (2003). *Fifty years of Mincer earnings regressions*. Cambridge, MA: NBER Working Paper 9732.
25. Heckman, J., Matzkin, R. y Nesheim, L. (2005). Simulation and estimation of hedonic models. En T. J. Kehoe, T. N. Srinivasan y J. Whalley (Eds.), *Frontiers in applied general equilibrium modeling* (pp. 277-339). Cambridge: Cambridge University Press.
26. Jehle, C., G. A. y Reny, P. J. (2011). *Advanced microeconomic theory* (3.^a ed.). Upper Saddle River (NJ): Prentice-Hall.
27. Jiménez, E. y Kugler, B. (1987). The earnings impact of training duration in a development Country. An ordered probit selection model for Colombia's Servicio Nacional de Aprendizaje (SENA). *Journal of Human Resources*, 22(2), 228-247.
28. Khander, S. R., Koolwal, G. B. y Samad, G. A. (2010). *Handbook on impact evaluation quantitative methods and practices*. Washington: Banco Mundial.
29. Korpi, T. y Antje, M. (2003). Training systems and labor mobility: a comparison between Germany and Sweden. *The Scandinavian Journal of Economics*, 105(4), 597-617.
30. Kuckulenz, A. y Zwick, T. (2003). *The impact of training on earnings: differences between participant groups and training forms*. Mannheim: ZEW Discussion Paper 03-57.
31. Laing, D. (2011). *Labor economics: introduction to classic and the new labor economics*. Nueva York: W. W. Norton & Company.
32. Lee, D. (2005). An estimable dynamic general equilibrium model of work, schooling and occupational choice. *International Economic Review*, 46(1), 1-34.
33. Lee, M.-J. (2005). *Micro-econometrics for policy, program and treatment effects*. Oxford: Oxford University Press.
34. Lemieux, T. (2003). *The "mincer equation" thirty years after 'schooling, experience and earnings'*. Berkeley: Universidad de California.
35. Lucas Jr., R. E. (2009). In defense of dismal science. *The Economist*, 41. Recuperado de <http://www.economist.com/node/14165405>

36. Mas-Colell, A., Whinston, M. D. y Green, J. G. (1995). *Microeconomic theory*. Oxford: Oxford University Press.
37. McIntosh, S. (2007). *A cost-benefit analysis of apprenticeships and other vocational qualifications*. Sheffield: Universidad de Sheffield.
38. Medina, C. y Núñez, J. (2005). *The impact of public and private job training in Colombia*. Bogotá: Banco Interamericano de Desarrollo.
39. Mincer, J. (1975). Education, experience, and the distribution of earnings and employment: an overview. En F. T. Juster (Ed.), *Education, income, and human behavior* (pp. 71-94). Cambridge (MA): NBER.
40. Ministerio de la Protección Social de Colombia (2008). Resolución 1747 de 2008, por la cual se modifica la Resolución 634 de 2006. Diario Oficial 46.998. Ministerio de la Protección Social (2010). Resolución 643 de 2010, por la cual se modifica parcialmente la Resolución 1615 del 12 de julio de 2006.
41. Mohrenweiser, J. (2010). *The causes and consequences of apprenticeship training. an economic and econometric analysis*. Recuperado de <http://edudoc.ch/record/96567/files/zu11047.pdf>
42. Munk, K. J. (2005). *Introduction to construction and use of CGE Models for Policy Analysis*. Aarhus: University of Aarhus.
43. Organización Internacional del Trabajo (2012). *Overview of apprenticeship systems and issues*. Ginebra: Autor.
44. Peña, X. y Bernal, R. (2011). *Guía práctica para la evaluación de impacto*. Bogotá: Universidad de los Andes.
45. Quintini, G. (15 de May de 2012). *Going separate ways? Differences in school-to-work pathways between Europe and the US*. Recuperado de <http://www.voxeu.org/article/school-work-pathways-europe-and-us>
46. Ravallion, M. (2008). Evaluating Anti-Poverty Programs. En T. P. Schultz y J. Strauss (Eds.), *Handbook of development economics* (vol. 4, pp. 3788-3846). Ámsterdam: Elsevier.
47. Robinson, J. A. y Ragnar, T. (2005). White elephants. *Journal of Public Economics*, 89, 197 - 2010.
48. Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.
49. Ryan, P. (2011). *Apprenticeship: between theory and practice, school and workplace*. Zürich: Leading House.
50. Ryan, P., Wagner, K., Teuber, S. y Backes-Gellner, U. (2010). Trainee Pay in Britain, Germany and Switzerland: Markets and Institutions. *SKOPE Research Paper*, 96.
51. Sen, A. K. (1988). The Concept of Development. En H. D. Chenery y T. N. Srinivasan (Eds.), *Handbook of development economics* (vol. 1, pp. 9-26). Ámsterdam: Elsevier.
52. Servicio Nacional de Aprendizaje (SENA). Resolución 02367 de 2010, por la cual se modifica parcialmente el Manual de Procesos y Procedimientos del SENA, Código: MPP-002 de diciembre de 2008, adoptado mediante la Resolución N° 00.751 de 2008.
53. Varian, H. (1993). *Análisis microeconómico*. Barcelona: Antoni Bosch.
54. Westergaard-Nielsen, N. y Rasmussen, A. R. (1997). *Apprenticeship training in Denmark: The impact of subsidies*. Aarhus: Universiad de Aarhus.
55. Wolter, S. C. (2012). *Apprenticeship training can be profitable*. Berna: European Expert Network on Economics of Education (EENEE).
56. Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambdridge (MA): MIT Press. Anexo 1. Resultados en el sistema de gestión ambiental