

Borradores de ECONOMÍA

Eficiencia de los establecimientos
Bancarios (EB): una aproximación
mediante modelos DEA

Por: Javier E. Pirateque
José H. Piñeros
Linda Mondragón

Núm. 798
2013



Colombia - Bogotá - Col

EFICIENCIA DE LOS ESTABLECIMIENTOS BANCARIOS (EB): UNA APROXIMACIÓN MEDIANTE MODELOS DEA[‡]

Javier E. Pirateque[§]

José H. Piñeros[♦]

Linda Mondragón^{*}

RESUMEN

Este documento analiza la eficiencia (técnica, de escala, de costos y asignativa) del sistema bancario colombiano entre 2000-2012, utilizando la metodología DEA (*Data Envelopment Analysis*). Ésta se basa en funciones de producción individuales de las entidades que componen el sistema, contraria a los métodos paramétricos. Dicha metodología tuvo en cuenta el enfoque de producción, el cual permite incorporar diversos tipos de productos y servicios financieros de carácter heterogéneo, convirtiéndose en un enfoque alternativo al presentado en otros trabajos en Colombia. Los resultados de la eficiencia, después de la crisis financiera de finales de los noventa, mostraban sus más bajos niveles del período analizado; luego, para los años siguientes se observó una mejora sistemática que se acentuó entre 2006 y 2008. Posteriormente, la eficiencia fue alterada por cambios en la dinámica de la producción de la industria bancaria, en donde variables externas pudieron haberla afectado. También se encontró que las entidades nacionales fueron más eficientes que las extranjeras durante el período de estudio. Cuando se agrupan por tamaño, no se halló evidencia clara de que las entidades grandes hayan tenido un mayor nivel de eficiencia que el de las pequeñas. Finalmente, se demostró que los determinantes de la eficiencia no solo están relacionados con la gestión de intermediación propia de cada banco, sino también con variables ambientales que trascienden la gestión individual.

Palabras claves: Eficiencia técnica y asignativa, economías de escala, frontera, modelos paramétricos y no paramétricos.

Clasificación JEL: C61; C67; D24; D61; G21; L25.

[‡] Las opiniones expresadas en este documento son responsabilidad de los autores y no comprometen al Banco de la República ni a su Junta Directiva. Los autores agradecen la colaboración y los comentarios de Miguel Sarmiento, quien fue uno de los promotores de la idea de realizar el presente documento, con la metodología señalada. Igualmente agradecemos los comentarios de José E. Gómez, Orlando Chipatecua y Andrés Murcia a una versión anterior.

[§] Profesional del Departamento de Estabilidad Financiera del Banco de la República (jpiratni@banrep.gov.co).

[♦] Profesional Líder del Departamento de Operaciones y Desarrollo de Mercados del Banco de la República (jpinergo@banrep.gov.co).

^{*} Profesional del Departamento de Operaciones y Desarrollo de Mercados del Banco de la República (lmondrrva@banrep.gov.co).

ABSTRACT

This paper analyzes the technical, scale, cost and allocative efficiency of the Colombian banking system during 2000-2012, by using the DEA methodology (Data Envelopment Analysis). This is based on individual production functions of the entities that comprise the system, in contrast with the parametric methods. This methodology took into account the production approach, which allows incorporating several types of heterogeneous products and financial services, and it becomes an alternative approach to other works presented in Colombia. The efficiency indicators reached its lowest level after the financial crisis of the late nineties when the banks were in a stage of recovery. For the following years there was a steady improvement which was more pronounced between 2006 and 2008. From that last year, the efficiency was altered by changes in the dynamic of the production, where external variables may have had some effect on it. On the other hand, it was found that domestic banks were more efficient than foreign banks during the period of analysis. When it is grouped by size, it was not found clear evidence that large institutions have had a higher level of efficiency than small ones. Finally, it was shown that the determinants of efficiency are not only related to the management of each bank itself, but also with environmental variables that transcend individual management.

Keywords: Technical and allocative efficiency, economies of scale, frontier, parametric and non-parametric models.

JEL Classification Codes: JEL C61; C67; D24; D61; G21; L25.

1. INTRODUCCIÓN

El sector financiero es de gran importancia para la economía de un país ya que facilita la canalización de recursos financieros de agentes superavitarios hacia agentes deficitarios, y sus mayores niveles de eficiencia se traducen en incrementos en bienestar para toda la sociedad. En Colombia, la liberalización financiera de comienzos de la década de los noventa, la crisis financiera de finales de esa década, la reestructuración a los entes de control y supervisión, las reformas regulatorias (modelos de riesgo crediticio, de mercado y de liquidez, entre otros) han conducido a que los establecimientos de crédito (EC)¹ exhiban cambios significativos que pueden haber repercutido en su eficiencia.

Por estas razones es interesante no sólo conocer los cambios en eficiencia del sistema bancario colombiano durante los últimos años, sino también encontrar los instrumentos idóneos para medir dicha eficiencia. En esta dirección, existen desde las metodologías más sencillas como las razones contables y financieras tradicionales, hasta las más elaboradas como las provenientes de la investigación operacional que se encuentran en la literatura económica, las cuales proveen mayor información para dicho análisis.

En efecto, una manera tradicional de medir la eficiencia de los intermediarios financieros es mediante las razones de gastos administrativos y laborales, estandarizados por los activos o por el margen financiero bruto (MFB)². Sin embargo, la volatilidad de estos rubros, especialmente la del MFB, hace que los resultados de estos indicadores pierdan consistencia. De esta manera, muchas veces los indicadores cambian de tendencia sin que sea consecuencia de una verdadera pérdida o ganancia en eficiencia económica³. Por ejemplo, si aumenta el MFB de manera importante, debido a un aumento en las tasas de interés de cartera, el indicador estaría mostrando una mejora en eficiencia, lo cual no es necesariamente cierto.

Tener una medida de eficiencia económica permite: obtener una idea del aprovechamiento y la utilización de los recursos administrativos y laborales; ver la evolución de la infraestructura operativa y la tecnología de producción utilizada por los EC; un análisis más detallado y preciso para explicar indicadores tales como la rentabilidad; ver cuáles son las entidades que tienen mejores o peores prácticas y, por último, analizar los efectos de eventos particulares (e.g. fusiones o crisis).

Para medir la eficiencia económica del sistema financiero y del sector real se cuenta básicamente con dos tipos de modelos: los paramétricos y no paramétricos. Los primeros se fundamentan en la estadística paramétrica, que es una rama de la estadística inferencial que comprende los procedimientos estadísticos y de decisión, los cuales se basan en las distribuciones de los datos reales. Estas son determinadas usando un número finito de

¹ Los EC se refieren a los bancos, las corporaciones financieras, las compañías de financiamiento tradicionales y especializadas en leasing, y a las cooperativas financieras.

² El MFB es la diferencia entre los ingresos financieros menos los egresos financieros.

³ Pese a que el cálculo de la eficiencia económica no implica el uso de variables financieras, éstas se usan debido a las características de la industria que hacen que las cantidades solo puedan ser expresadas en términos monetarios.

parámetros. Los procedimientos paramétricos requieren conocer la forma de distribución para las mediciones resultantes de la población estudiada, es decir, que se requiere fijar una forma funcional previa, haciendo que estos modelos mantengan supuestos muchas veces restrictivos.

Como se verá más adelante, los modelos no paramétricos, entre otras cosas, son menos rígidos en cuanto al requerimiento de información y están en capacidad de recoger tecnologías individuales de las firmas a partir de optimizaciones individuales.

El objetivo del presente trabajo es mostrar cuál ha sido el comportamiento de la eficiencia de los establecimientos bancarios (EB) posterior a la crisis financiera en Colombia de finales de la década de los noventa (2000-2012), período en el cual se dieron grandes cambios regulatorios en el sector financiero, reorganizaciones institucionales (fusiones, absorciones, liquidaciones y entrada de nuevas entidades financieras), estrés financiero en el mercado de deuda pública de 2006 y la reciente crisis internacional de 2008. Adicionalmente, se evalúa la eficiencia de los EB de acuerdo con su propiedad (locales versus extranjeros) y tamaño (grandes frente a pequeños, según el nivel de activos). Estos análisis se hacen aprovechando las ventajas de la metodología DEA (*Data Envelopment Analysis*), ampliamente utilizada para tal propósito en la literatura económica internacional, que por primera vez a nivel local se lleva a cabo utilizando el enfoque de producción y no de intermediación. Finalmente, con los resultados de eficiencia técnica derivados del modelo, se especifica un modelo econométrico para hallar sus determinantes y la importancia que tiene en su explicación las variables ambientales.

El presente documento consta de siete secciones, incluida la presente introducción. En la segunda se presenta el marco teórico y la elección de la metodología. En la siguiente se hace una revisión de trabajos previos, y en la cuarta se especifican las características de los modelos DEA. En quinto lugar se explica la selección de variables y los diferentes modelos a aplicar. En una sexta sección se presentan los resultados de la eficiencia a la luz del modelo DEA escogido, junto con los resultados de los determinantes de la eficiencia obtenidos mediante un modelo econométrico y, finalmente, se tienen las conclusiones de la investigación y orientación sobre futuros estudios.

2. MARCO TEÓRICO Y ELECCIÓN DE LA METODOLOGÍA

Dentro de los indicadores financieros que tradicionalmente se utilizan para diagnosticar la situación financiera de los EC se encuentran los indicadores de eficiencia. Los rubros que se usan para identificar dicha variable son los gastos administrativos y laborales; éstos generalmente son escalados por los activos, la cartera bruta, u otros activos de la hoja de balance que están relacionados con las principales actividades productivas de los EC. Asimismo, los mencionados gastos pueden ponderarse en términos del margen financiero bruto (MFB), entendiéndose que una entidad es más eficiente entre más pequeño sea el porcentaje del MFB destinado a cubrir tales gastos. Sin embargo, la volatilidad del MFB puede fácilmente cambiar los resultados de eficiencia de los EC en el corto plazo, sin que necesariamente haya cambiado la eficiencia económica.

Adicionalmente, dado que estos indicadores son razones financieras que no cuentan con un umbral de referencia, no puede establecerse con objetividad si los niveles en que opera cada una de las entidades financieras son óptimos en términos de eficiencia, ni tampoco permiten explicar dónde se originan las ineficiencias ni como éstas pueden ser superadas. De la misma forma, no consienten la combinación de varios productos, es decir, diferentes enfoques de negocio (e.g. cuando el negocio integral de un EC está representado por dos productos como cartera e inversiones), ni diversos nichos de mercado, o lo que es más importante, no tienen una fundamentación teórica fuerte. No obstante, estos indicadores tienen la ventaja de ser de fácil entendimiento y estimación.

A nivel teórico y bajo los supuestos microeconómicos habituales, se afirma que todas las firmas maximizan el beneficio, sin importar el ambiente de competencia en el que se encuentran (competencia perfecta, monopolio, oligopolio, etc.), y que este comportamiento optimizador las lleva a ser eficientes. Sin embargo, en el mundo real este proceso de optimización no es del todo completo⁴ pues no siempre se logra alcanzar el máximo beneficio, lo que conlleva a que no todas las firmas sean consideradas eficientes.

Teóricamente se considera que para que una empresa, en un ambiente de competencia perfecta, maximice el beneficio, debe alcanzar tres tipos de eficiencia que conforman la eficiencia total o económica: de escala, técnica, y asignativa (basada en la técnica y la de costos).

La eficiencia de escala, que se deriva de las economías de escala, no es otra cosa que los costos en que incurren las firmas al mantener tamaños muy grandes o pequeños de sus plantas. Por ejemplo, tal como se explica en el trabajo de Gonzales y Martínez (2008), una empresa pequeña tendría incentivos para aumentar su tamaño, siempre y cuando el producto crezca en una proporción mayor al aumento de los insumos involucrados en la producción. Así, la mencionada empresa operaría más eficientemente si su tamaño fuera superior. Por el contrario, empresas excesivamente grandes alcanzarían un incremento menos que proporcional en el producto en relación con el aumento de sus insumos. Estos ejemplos hacen referencia a los rendimientos crecientes y decrecientes a escala, respectivamente.

La eficiencia técnica puede enfocarse desde el punto de vista del producto o de los insumos. El primer caso se refiere a la destreza de una firma para maximizar el nivel de producción dado un nivel y combinación específica de insumos. El segundo caso, orientado a insumos, tiene que ver con la menor utilización de insumos posible para alcanzar un nivel determinado de producto. Cabe mencionar que los resultados en términos de eficiencia son los mismos si el enfoque es uno u otro, toda vez que se hable de rendimientos constantes a escala, es decir, cuando el incremento de los insumos es proporcional al aumento en el nivel de producción.

⁴ Esto es así porque en el mundo real no hay perfecta movilidad del capital, la función de producción no es determinística, los ambientes de competencia son diferentes a los teóricos (e.g. competencia monopolística) y, lo que es más importante, hay asimetrías de información que enfrentan los intermediarios financieros, lo que afecta sus decisiones de inversión y tolerancia al riesgo, repercutiendo a la postre en la eficiencia.

Por su parte, la eficiencia asignativa hace referencia al enfoque orientado a la combinación de insumos, los cuales ahora tienen en cuenta los precios respectivos dados de antemano. Cualquier cambio en el precio relativo de los insumos tiene implicaciones sobre su combinación. De esta manera, la eficiencia asignativa se define como la habilidad de una firma para utilizar los insumos en proporciones óptimas, dados los precios respectivos y un nivel determinado de producción (Worthington, 2001). Sin embargo, la concepción de esta clase de eficiencia es más abstracta, por cuanto en la práctica los insumos no son sustitutos perfectos y hay casos en los cuales pese al cambio en los precios hay que manejar una sola combinación técnica de los mismos. En el caso de la industria bancaria, podría decirse que los cambios en los precios relativos de los insumos como los gastos laborales, gastos de capital y gastos administrativos, difícilmente afectarían la proporción que se utiliza de cada uno para obtener un determinado nivel de productos financieros.

Se define la ineficiencia, de otro lado, como la brecha existente entre el nivel óptimo (medido en términos de la mejor práctica de mercado o un máximo teórico) y el nivel observado, ya sea en relación con los productos o con los insumos. En términos prácticos, González y Martínez (2008) fijan la ineficiencia técnica como la incapacidad de una firma o unidad de utilizar la mejor tecnología disponible en el mercado (e.g. una ensambladora que no utiliza las máquinas de última generación), la ineficiencia asignativa como la no elección de la combinación apropiada de insumos dado sus precios (e.g. un banco que no usa una proporción adecuada de cajeros automáticos en relación con los empleados que sirven en ventanilla, dados sus respectivos costos) y la ineficiencia de escala cuando no se maximiza el beneficio por no tener el tamaño óptimo (e.g. una empresa excesivamente grande que produce en un nivel en el cual el ingreso marginal es menor a su costo marginal).

Cabe señalar que la noción de eficiencia es diferente a la de productividad, puesto que esta última hace referencia al número de unidades de producto por unidad de factor, y son equivalentes sólo cuando se trabaja con tecnologías de coeficientes fijos⁵. Igualmente la noción de eficiencia es diferente a la de competitividad, pues esta última se refiere a ventajas por diferenciación del producto y reducción de costos sin deterioro de la calidad del mismo.

Los primeros ejercicios de eficiencia de la banca se concentraron en los conceptos de economías de escala y de alcance. El primero de ellos se refiere a la forma adoptada de U por la curva del costo promedio de la producción, que sugiere que existe un punto en el cual hay una escala de producción óptima cuyos costos serían minimizados. El concepto de economías de alcance se relaciona con las economías de producción conjuntas, donde los costos de productos conjuntos son inferiores a la suma de los costos de producción individual. En general, los resultados de esos ejercicios no han sido concluyentes acerca de si las entidades financieras gozan de economías de escala y alcance según su tamaño en las estimaciones para la banca en EE.UU. (Kwan y Eisenbeis, 1996).

Otro concepto en el que los investigadores también se han enfocado es el de la eficiencia X, la cual se refiere a las desviaciones desde la frontera eficiente de producción, la que a su

⁵ Esta tecnología hace referencia a que debe usar insumos en proporciones fijas para lograr el producto.

vez representa el máximo alcanzable de producción para un nivel dado de insumos. Estas desviaciones, según Leibenstein (1966) eran causadas por la gente y por las firmas que de ordinario no trabajaban, por diferentes circunstancias, tan duro ni tan eficazmente como ellos podrían hacerlo.

De otro lado, en la literatura económica se encuentra que los modelos de eficiencia están categorizados en dos grandes grupos: los modelos paramétricos y los no paramétricos. Dentro de los modelos paramétricos se encuentran los modelos de frontera estocástica (FE) y Frontera Gruesa (FG), y los que usan la metodología “*Distribution Free Approach*” (DFA). Dentro de los no paramétricos están los modelos que usan la metodología de “*Data Envelopment Analysis*” (DEA) y la metodología “*Free Disposal Hull*” (FDH) que es un caso particular de la anterior, entre otros.

Según Badel (2002), la principal ventaja de los modelos paramétricos sobre los no paramétricos está en que debido a que se fundamentan en procedimientos estadísticos, permiten tener en cuenta la existencia de ruido aleatorio, generado por errores de medición, el cual se complementa como aquella parte de la función de producción que no alcanza a explicar la ineficiencia de las entidades financieras. No obstante, tienen la desventaja de tomar la forma funcional de los costos o de la producción previamente, según sea el caso, de forma arbitraria y solo enfocarse en la estimación de los parámetros (en contraposición a los no paramétricos, donde la función de producción cambia de entidad a entidad), al igual que suponen las distribuciones de la ineficiencia y el error.

La fijación de una forma funcional crea otro problema que tiene que ver con expresar en términos monetarios, algo que debe ir en otras unidades; a pesar de que este es un problema que aún está sin resolver en la ciencia económica, su no consideración es más grave en los modelos paramétricos que en los no paramétricos, por la misma rigidez que imponen los del primer tipo.

A pesar de que ambos modelos admiten el uso de varios insumos y el ajuste por medio de variables categóricas (*dummies*) o externas, se puede afirmar que si la información de las diferentes variables es de calidad y las unidades de análisis son relativamente homogéneas, los modelos DEA pueden generar resultados mejores que los de los modelos paramétricos. Asimismo, cuando los errores se consideran importantes, los modelos DEA siguen siendo una buena aproximación pues existen nuevas técnicas que permiten incorporar el ruido aleatorio (Wheelock y Wilson, 2006).

La metodología propuesta en este documento para analizar la eficiencia de los EB es la de DEA tradicional. Ésta se seleccionó porque logra superar los problemas de los indicadores convencionales y porque se considera que tiene ventajas sobre las otras metodologías. En efecto, ésta es menos restrictiva en cuanto a la cantidad de datos en comparación con las metodologías paramétricas⁶, y permite capturar las tecnologías individuales de las firmas a través de la frontera, puesto que se construye con base en optimizaciones individuales. En

⁶ En las metodologías paramétricas se necesita una alta cantidad de datos para obtener estimadores con características deseables como: insesgamiento, eficiencia, consistencia y robustez.

este sentido, se puede decir que el modelo puede ser mejor aprovechado al realizar análisis individuales que agregados.

De otra parte, esta metodología permite incorporar diferentes rendimientos a escala, lo que la hace más realista ya que es muy posible que en la práctica no todas las firmas tengan el mismo tipo de rendimientos. La metodología DEA también tiene la ventaja de usar medidas radiales⁷ (que permiten la interpretación en términos de costos e ingresos) y de permitir tanto la incorporación de información externa como la opinión del investigador. En el caso de las entidades bancarias se puede incorporar información previa como es la proporción en la que se dividen los distintos insumos y productos.

A pesar de que esta metodología también posee la desventaja de ser sensible al número de insumos y productos utilizados en la estimación⁸, el análisis de las variables a incorporar (que se presenta más adelante) buscará que este hecho se reduzca al máximo posible.

El Cuadro 1, resume las ventajas y desventajas de cada uno de los modelos, según Charnes *et al.* (1994).

Cuadro 1. Ventajas y desventajas de las metodologías Paramétricas y no paramétricas.

	Metodologías Paramétricas	Metodologías No Paramétricas (DEA)
Ventajas	<ul style="list-style-type: none"> • Se pueden interpretar los coeficientes y hacer pruebas estadísticas sobre los mismos. • Representan el ruido aleatorio que captura los procesos aleatorios generadores de ineficiencias que no están bajo el control de la entidad y que son externos a ésta. • Provee medidas absolutas y no relativas⁹. 	<ul style="list-style-type: none"> • Se focaliza en las unidades de análisis no en medias poblacionales. • Cada unidad tiene una función de producción diferente. • Puede usar múltiples insumos y múltiples productos, cada uno en diferentes unidades. • No se requieren formas funcionales <i>a priori</i>. • Se puede incorporar el juicio del investigador. • Da respuestas sobre dónde nace la ineficiencia y cómo puede ser superada. • Los resultados son óptimos de Pareto. • Se focaliza en la frontera de mejores prácticas más que en las propiedades de tendencia central de la frontera.
Desventajas	<ul style="list-style-type: none"> • Se necesitan muchos datos para obtener resultados confiables. • Se debe definir una forma funcional <i>a priori</i>. • Sensibilidad a los datos atípicos. • Fuertes supuestos para separar la ineficiencia del ruido aleatorio. • No da respuestas sobre dónde nace la ineficiencia y cómo puede ser superada. 	<ul style="list-style-type: none"> • La medida no es robusta ante cambios en la cantidad de insumos o productos. • No permite la inferencia estadística ni mecanismos tradicionales como las pruebas de hipótesis¹⁰. • Es más susceptible a los errores de medición. • Provee eficiencias relativas más no absolutas. • Tiene una demanda computacional intensiva.

Fuente: Charnes *et al.* (1994).

⁷ La medida radial es la proyección radial de la firma ineficiente sobre la frontera eficiente.

⁸ El número de unidades eficientes es aproximadamente igual al número de insumos multiplicado por el número de productos, por lo cual si se incluyen demasiados productos o insumos sería difícil identificar las unidades no eficientes.

⁹ Medida absoluta en el sentido de que la entidad analizada se compara sólo con la ideal, y no tiene en cuenta la medida relativa en el sentido de que no se compara con las demás entidades.

¹⁰ En cuanto a las pruebas de hipótesis, en los últimos años se han presentado muchos avances que permiten validar los resultados del modelo. Ver Banker (1996).

3. REVISIÓN DE TRABAJOS PREVIOS

La literatura internacional presenta distintas aplicaciones de la medición de la eficiencia con modelos DEA. Ésta en general presenta aplicaciones de los modelos BCC y CCR¹¹, con descomposiciones mediante el índice de Malmquist, el cual será explicado en una sección posterior. Los trabajos más adelantados comprenden análisis con descomposición por riesgo y variables ambientales¹². Algunos ejemplos de este tipo de modelos aplicados a los establecimientos bancarios se presentan en el Cuadro 2.

Cuadro 2. Variables de insumo y producto usadas en otros trabajos sobre EB¹³.

Enfoque	Metodología	Autores	Variables Producto	Variables Insumo
Intermediación	DEA	Elyasiani y Mehdian (1990)	Cartera de Créditos e Inversiones	Depósitos, número de empleados y Capital
		Chansarn (2008)		Depósitos, Capital laboral y Costo Total
		Sarmiento, Cepeda, Mutis y Pérez (2013)		Depósitos, Capital Físico, Capital laboral y Capital Financiero
		Jemric y Vujcic (2002)		Depósitos, Capital laboral y Activos Fijos
	DEA y Frontera estocástica	Fiorentino, Karman y Koetter (2006)		Depósitos, Capital de Trabajo y Capital Físico
	Distribution Free Approach	Castro (2001)		
	Frontera estocástica	Janna (2003)		
		Cepeda, Sarmiento y Mutis (2010)		
		Estrada y Osorio (2004)	Depósitos, Capital Físico y Capital Financiero	
Producción	DEA	Jemric y Vujcic (2002)	Ingresos por intereses y Otros ingresos (operacionales y comisiones por servicios)	Egresos por intereses
		Chansarn (2008)		Comisiones por Servicios
		Banker, Chang y Lee (2010)		Gasto en personal
				Gasto en capital
		Yang (2009)	Número de transacciones para generar diferentes transacciones y procesos en las sucursales del banco	Otros egresos operacionales
				Ventas, servicios, apoyo y otros

Fuente: Documentos de Investigación de los autores señalados.

Kenjegaliev y Simper (2010), presentan un trabajo en el que explican la productividad de la banca de Europa Oriental tomando en cuenta la descomposición del riesgo y variables ambientales. Usan la metodología de rango direccional (para elementos que se encuentran fuera del balance) y estiman el índice de productividad de Luenberger para capturar el cambio en eficiencia, el cambio tecnológico, el ajuste de riesgo y el efecto de las variables de ambiente. Este documento presenta tres diferentes enfoques para describir la relación

¹¹ BCC se refiere a la metodología DEA propuesta por los autores Banker, Charnes y Cooper, y CCR a Charnes, Cooper y Rhodes.

¹² La incorporación de variables ambientales y descomposición por riesgo, se realiza con el fin de añadir las restricciones no conocidas por el investigador que pueden condicionar el proceso de maximización de la empresa. Por ejemplo, si un banco aumenta la cantidad de reservas y por ello disminuye la colocación de préstamos, no significa necesariamente que esté siendo menos eficiente, sino que está controlando por riesgo.

¹³ Para una revisión más completa de diferentes estudios de eficiencia, ver Berger, A. *et al.* (1997)

insumo/producto: el enfoque de intermediación, el enfoque de producción y un enfoque basado en los beneficios e ingresos. Los resultados del trabajo mostraron que el riesgo (por ejemplo, el creado por las condiciones cíclicas de la economía o aquel que afecta la calidad de los préstamos) se constituía en un factor significativo en cada una de las metodologías, y que las variables de ambiente (regulación, nivel de corrupción, grado de competencia, etc.) y cambio tecnológico jugaban un papel importante en la productividad de los bancos.

Portela *et al.* (2010), realizan una aplicación de modelos DEA con índices tipo Malmquist en presencia de datos negativos (e.g. pérdidas financieras, reducciones en el número de clientes sobre el tiempo) para sucursales de bancos. Este trabajo utiliza como variables insumo el número de empleados por establecimiento, y como producto el número de transacciones, el número de clientes, el monto de cuentas corrientes, el valor de otras cuentas, el valor de los títulos de depósito y el valor del crédito por banco y por asociados. En este trabajo se encuentra que la mayor parte del cambio en la productividad se explica por el cambio tecnológico, y que las sucursales parecen haber mantenido una posición en el tiempo relativamente estable en relación con su propia frontera (es decir, que no avanzaron mucho durante el período analizado en eficiencia). Un trabajo similar fue adelantado por Yiang (2009), en el cual se evalúan 240 sucursales de un banco canadiense. Este autor examina cómo las diferentes sucursales combinan sus recursos para apoyar la mayor cantidad de servicios financieros posibles.

Banker *et al.* (2010), realizan un trabajo para analizar el impacto diferencial de las reformas de la banca coreana en la productividad bancaria. Este trabajo se realizó en dos etapas: en la primera se usó un modelo BCC en DEA para estimar la eficiencia de cada banco por año, luego en la segunda etapa, se hizo una regresión de la eficiencia estimada contra los factores contextuales potenciales para obtener estimadores consistentes del impacto en la eficiencia de los bancos. Se encuentra que la eficiencia decreció desde 1995 hasta 1998, y que existe una correlación negativa (estadísticamente significativa) entre eficiencia técnica y el logaritmo natural de la cartera en mora, y una correlación positiva (estadísticamente significativa) entre eficiencia técnica y el capital adecuado (relación de solvencia).

Jemric y Vujcic (2002), presentan una aplicación de los modelos DEA para medir la eficiencia de los bancos en Croacia. Estos autores toman dos aproximaciones para la selección de los insumos y los productos: una aproximación desde el punto de vista de producción y otra desde intermediación. Para cada una de estas aproximaciones estiman dos modelos (CCR y BCC) y encuentran que los bancos extranjeros son más eficientes que los nacionales, que los nuevos son más eficientes que los antiguos y que los pequeños son globalmente eficientes, mientras que los grandes son localmente eficientes.

Elyasiani y Mehdián (1990), utiliza la metodología DEA para medir la eficiencia de los grandes bancos comerciales en EE.UU. En este trabajo, los bancos son tratados como intermediarios financieros que combinan depósitos (cuentas de ahorro y CDT), trabajo (número de empleados) y capital (maquinaria y equipo, y planta física adquirida por los bancos o tomados en leasing) para producir cartera e inversiones. Los préstamos se desglosan en hipotecarios, comerciales, industriales, y otros debido a que sus costos operativos son diferentes. Se encuentra que el promedio de eficiencia de los bancos en la muestra es de 89,55% en relación con la frontera I_1 (isocuanta en el momento uno) y de

77,71% relativo a la frontera I_2 (isocuanta en el momento dos). Adicionalmente, se encontró que el avance tecnológico se sesgó hacia el factor trabajo.

En Colombia, la mayoría de los trabajos sobre la eficiencia bancaria se han desarrollado bajo enfoques paramétricos (FE, FG y DFA). Los primeros análisis se orientaron hacia la evaluación de la existencia de economías de escala y de alcance (mediante aproximaciones de costos o beneficios), mientras que los más recientes se han concentrado en la cuantificación de la eficiencia económica, junto al origen de la ineficiencia y sus consecuencias¹⁴.

Entre los estudios basados en la metodología DEA sobresalen el de Martínez y Estrada (2009) y el de Sarmiento *et al.* (2013); ambos trabajos adoptan la metodología DEA y el índice de Malmquist para estudiar la eficiencia relativa y el cambio en productividad. El primero de ellos analiza las compañías aseguradoras para el período 1997 – 2007, mostrando que en efecto la eficiencia de este sector empezó a aumentar a partir de 2002, lo que le permitió recuperar parte de la caída experimentada durante los primeros años del período. Igualmente, mediante el índice de Malmquist se llegó a la conclusión de que los incrementos en la productividad total de los factores obedecieron principalmente a cambios tecnológicos positivos, los cuales contrarrestaron la tendencia decreciente de la eficiencia técnica.

El estudio de Sarmiento *et al.* (2013), utiliza la metodología DEA y el enfoque de intermediación para hallar la eficiencia de la banca colombiana y demostrar, entre otras cosas, que efectivamente las fusiones y adquisiciones ocurridas después de la crisis financiera de finales de la década de los noventa le permitieron al sistema financiero alcanzar una mayor eficiencia. Igualmente, mediante el cálculo del índice de Malmquist se reveló que el aumento en la productividad del sector bancario se explica primordialmente por el cambio tecnológico y, en menor medida, por el aumento de la eficiencia técnica.

4. MODELO Y ESPECIFICACIÓN TEÓRICA

La metodología DEA parte de la formulación tradicional microeconómica de la función de producción: $y = f(\vec{x})$, donde y representa el producto (o vector de productos) y \vec{x} el vector de insumos. Teóricamente y en el caso de la competencia perfecta, todas las firmas trabajan bajo la misma tecnología por lo que para todos los casos aplica que: $\frac{y}{f(\vec{x})} = 1 = ET$ ¹⁵, o lo que es igual $y = f(\vec{x})$, estado en el cual todas las firmas se consideran técnicamente eficientes. Puesto que en la realidad no todas las firmas alcanzan el producto que predice la función de producción, ya sea porque no todas tienen la misma función o porque surge algún tipo de ineficiencias durante el proceso productivo, se tiene que ET es en general menor a 1.

¹⁴ Para ver una descripción detallada de los trabajos que se han hecho sobre el tema, remitirse a Janna (2003).

¹⁵ ET representa eficiencia técnica.

Hasta este punto todas las formulaciones de los modelos de eficiencia coinciden: la ineficiencia técnica es la razón entre el producto factible y el observado. No obstante, éstas divergen al intentar definir el producto factible para cada vector de insumos o lo que es lo mismo la función de producción. Mientras que, por ejemplo, el modelo de frontera estocástica la define con base en un componente determinístico (generalmente una función Cobb-Douglas) y otro estocástico (el error aleatorio), el modelo DEA la define con base en las mejores prácticas observadas en la muestra analizada y una combinación convexa de las mismas. Ésta última característica hace que, por definición, siempre exista en el modelo DEA al menos una entidad relativamente eficiente.

El modelo básico de eficiencia técnica es el descrito por Charnes, Cooper y Rhodes (CCR), cuya ineficiencia se mide bajo el supuesto de rendimientos constantes a escala. Este modelo asume que existen n unidades de producción a ser comparadas. Tomando la unidad m como la unidad que maximiza su eficiencia, se puede escribir el programa matemático como sigue:

$$\max E_m = \frac{\sum_{j=1}^J v_{jm} y_{jm}}{\sum_{i=1}^I v_{im} x_{im}}$$

Sujeto a:

$$0 \leq \frac{\sum_{j=1}^J v_{jm} y_{jn}}{\sum_{i=1}^I v_{im} x_{in}} \leq 1; \quad n = 1, 2, K, N \quad v_{jm}, v_{im}; \quad i = 1, 2, K, I; \quad j = 1, 2, K, J;$$

Donde

E_m es la eficiencia de la unidad de producción m ,

y_{jm} es el producto j de la unidad m ,

v_{jm} es la participación de ese producto dentro del total de productos producidos,

x_{im} es el insumo i de la unidad de producción m ,

v_{im} es la participación de ese insumo dentro del total de insumos utilizados, y

y_{jn} y x_{in} son el producto j y el insumo i , respectivamente, de la unidad n , $n=1, 2, \dots, N$. n incluye m .

El numerador de E_m representa la combinación de productos y el denominador la combinación de insumos. El programa de maximización busca que las ponderaciones asignadas a cada producto e insumo sean las mejores para lograr que esta relación sea lo más alta posible (que se produzca lo más que se pueda con la utilización de la menor cantidad de insumos posible)¹⁶. La restricción se encarga, por una parte, de fijar que la eficiencia de la firma de referencia no sea mayor a uno ni menor a cero (por lo que n incluye a m) y, por otra, que los pesos (participaciones) encontrados mantengan las

¹⁶ El índice que se maximiza se conoce en la literatura como PTF (Productividad Total de los Factores). Esta relación se considera también de eficiencia, pues si se toma en cuenta el caso en el que se produce un único producto y las ponderaciones de los insumos son los precios de los factores (w_j), el índice se convierte en:

$$PTF = \frac{y}{\sum w_j x_j} = \frac{y}{CT} = \frac{1}{cMe}, \text{ que es la inversa del costo medio (Pinilla, 2001).}$$

eficiencias del resto de unidades en valores entre cero y uno (para hacer la medida comparable entre entidades)¹⁷. Este proceso se debe repetir para cada entidad a fin de hallar las eficiencias de un grupo de entidades. En términos prácticos, lo que se busca con las restricciones es que cada entidad tenga una función de producción propia que sea factible para las demás. Ello se hace con el propósito de que las entidades con peores prácticas puedan adoptar las de las mejores.

Este modelo puede ser reexpresado en la forma del modelo multiplicativo, que básicamente se diferencia del modelo básico en que los insumos y los productos no son agregables aditivamente sino multiplicativamente como su nombre lo indica. El modelo es especialmente útil cuando se trabaja bajo el supuesto de funciones Cobb-Douglas, puesto que el básico representa funciones de producción con insumos que son sustitutos perfectos (Ramathan, 2003).

Dado que en la práctica la forma no lineal del modelo puede generar problemas para su estimación, usualmente el problema se reescribe normalizando el numerador o el denominador (se vuelve igual a 1), para luego encontrar su forma dual que computacionalmente es menos exigente que la presentada inicialmente (Anexo 1).

A partir de los modelos con rendimientos constantes y variables a escala se puede obtener la eficiencia de escala (EE), definida en términos de la escala que permite el máximo nivel de producción por unidad de insumo. Esta se obtiene como la razón entre la eficiencia obtenida con el modelo CCR y la eficiencia obtenida del modelo BCC (que es una variación del modelo CCR, que agrega una restricción de convexidad para permitir retornos variables a escala –Anexo 1–):

$$EE = \frac{CCR \theta^*}{BCC \theta^*}$$

El modelo DEA antes descrito está referido a la eficiencia técnica; no obstante, éste se puede expandir para representar la eficiencia en costos. Este problema se escribiría como:

$$\min c_{im}x_{im}$$

Tal que:

$$\sum_{n=1}^N y_{jn}\lambda_n \geq y_{jm} ; \quad j = 1, 2, K, J$$

$$\sum_{n=1}^N x_{in}\lambda_n \leq x_{im} ; \quad i = 1, 2, K, I$$

$$\sum_{n=1}^N \lambda_n = 1 ; \quad n = 1, 2, K, N$$

$$\lambda_n \geq 0 ; \quad n = 1, 2, K, N$$

¹⁷ Dado que en algunas circunstancias es ingenuo pensar que una entidad puede dejar de producir cierto producto o producirlo en muy pequeña cantidad, se pueden fijar restricciones adicionales para que los pesos no sean menores a cierto valor, según el criterio del investigador.

De donde se encuentra la siguiente relación que define la eficiencia en costos:

$$0 \leq \frac{\sum_{i=1}^I c_{im}x_{im}^*}{\sum_{i=1}^I c_{im}x_{im}} \leq 1$$

Siendo x_{im}^* los valores obtenidos de la optimización del modelo y x_{im} los valores observados para la unidad que se evalúa. La programación se interpreta como la minimización de los costos totales, donde c_{im} es el costo del insumo i para la unidad m , sujeto a que la suma ponderada del producto de todas las firmas sea mayor o igual al producto de la unidad de análisis, y sujeto a que la combinación ponderada de insumos del resto de firmas no sea mayor al de la firma analizada. La tercera restricción corresponde al caso de retornos variables a escala. En cuanto a la relación que define la eficiencia en costos se interpreta como la distancia entre el nivel óptimo de cada insumo (de acuerdo con su costo y dado un nivel de producto) y el observado.

A partir de la eficiencia en costos y la eficiencia técnica se puede obtener la eficiencia asignativa¹⁸, la cual es definida así:

$$Eficiencia\ Asignativa = \frac{Eficiencia\ en\ Costos}{Eficiencia\ Técnica}^{19}$$

Esta relación expresa la capacidad de combinar insumos para generar productos, donde los costos de producción son minimizados para un conjunto de precios dados. Cabe anotar que para el caso de retornos variables a escala, el denominador de la anterior razón cambiaría por el producto entre la eficiencia técnica y la eficiencia de escala (Fried *et al.*, 2008).

La eficiencia en costos nos dice qué tan lejos está la entidad analizada de su nivel mínimo de costos factible, y ante cambios en los precios relativos de los insumos se mantiene su mezcla técnica. Por el contrario, en la asignativa hay cambios en dicha mezcla para alcanzar una combinación óptima, cuya pendiente es tangente a la frontera de eficiencia técnica.

Gráficamente las diferentes definiciones de eficiencia, para el caso de dos insumos y un producto, se pueden representar por medio de la Gráfica 1. Este muestra el programa de optimización CCR, cuya tecnología se representa por una isocuanta unitaria, junto a varias líneas de isocostos: i) $C_o(y, w) = wx_0$ es el costo observado para la unidad G ; ii) $C_T(y, w) = wx_T$ es el costo en el punto de eficiencia técnica de la misma unidad representado por G' ; y iii) $C_m(y, w) = wx_m$ es el costo mínimo al que podría producir

¹⁸ En éste caso dicha eficiencia está basada en costos; no obstante, ésta puede ser basada en el producto o el beneficio.

¹⁹ En éste caso la eficiencia técnica debe ir orientada al producto.

$$M_1(X^{t+1}, Y^{t+1}, X^t, Y^t) = \left[\frac{D_I^{t+1}(X^{t+1}, Y^{t+1})}{D_I^t(X^t, Y^t)} \right] \left[\left(\frac{D_I^t(X^t, Y^t)}{D_I^{t+1}(X^t, Y^t)} \right) \left(\frac{D_I^t(X^{t+1}, Y^{t+1})}{D_I^{t+1}(X^{t+1}, Y^{t+1})} \right) \right]^{1/2}$$

Donde el primer componente de la expresión recoge lo correspondiente a los cambios en eficiencia técnica y el segundo lo correspondiente al cambio tecnológico. El numerador y el denominador del primer componente representan los niveles de eficiencia, a través de la distancia a la frontera, para los períodos t y $t + 1$. Si dicha relación es mayor que 1 se puede concluir que la unidad de producción es más eficiente en $t + 1$ que en t ; lo contrario ocurre si la relación es menor que 1. Cabe señalar que cuando se trabaja bajo el supuesto de retornos variables a escala, este componente corresponde no solo a los cambios en eficiencia técnica sino también a los cambios en eficiencia a escala.

El segundo componente recoge el cambio tecnológico mediante la comparación entre las fronteras de eficiencia de los períodos t y $t + 1$. En este caso, si el resultado es mayor que 1, se tiene que la industria presentó una mejora tecnológica; si no, se puede decir que permaneció estable u ocurrió una desmejora. Dado que el índice de Malmquist es el resultado de la multiplicación de dos componentes, un nivel mayor que uno expresa un incremento en la productividad que puede provenir de alguna de las dos fuentes mencionadas²⁰.

5. SELECCIÓN DE VARIABLES Y MODELOS A UTILIZAR

La mayoría de los estudios sobre el tema acogen el enfoque de producción o de intermediación²¹. En el enfoque de producción, los bancos son vistos como unidades de negocio que tienen el objetivo final de generar ingresos a partir de los gastos incurridos en el manejo del negocio. En el enfoque de intermediación, los bancos son tratados como intermediarios entre los proveedores de fondos (ahorradores) y los usuarios de los fondos (inversionistas), convirtiendo los depósitos en préstamos (Mester, 1997). Según Berger y Humphrey (1997), bajo el enfoque de producción se deben incluir insumos físicos como capital y trabajo, en razón a que éstos son los principales recursos que se necesitan para obtener productos como las transacciones y el procesamiento de documentos financieros (los productos propios de este enfoque). De otra parte, bajo el enfoque de intermediación es fundamental la inclusión de los depósitos dado que éstos son la “materia prima” que se transforma en el proceso de intermediación financiera.

De este modo, se considera que este último enfoque es más específico (aunque quizás más limitado) que el de producción, ya que tiene la particularidad de entender el negocio bancario como la capacidad de reasignar los recursos entre individuos superavitarios y deficitarios, a cambio de un margen con el que se queda durante el proceso. Por su parte, el enfoque de producción es más general, pues toma en cuenta dicha reasignación con varios productos que se asemejan más a los resultantes de un proceso productivo del sector real y

²⁰ Para una aproximación más completa al índice de Malmquist ver Färe *et al.* (1992).

²¹ Existen aproximaciones más específicas que no se relacionan con estos enfoques y que basan su análisis en productos como el valor agregado o los costos para los usuarios.

que comprenden las transacciones tradicionales (cartera e inversiones), otras transacciones (operaciones en el mercado monetario y cambiario), y servicios financieros como transferencias, recaudos, avales, garantías, venta de chequeras, entre otros.

El enfoque de producción utilizado en este estudio (y por primera vez a nivel local) toma relevancia toda vez que en Colombia los ingresos percibidos por los bancos se han venido diversificando, en la medida en que aquellos provenientes de rubros diferentes a los generados por el portafolio de cartera e inversiones son cada vez más importantes; es decir, que hay una alta heterogeneidad en los productos. De hecho, una parte significativa de los recursos son canalizados para realizar otras transacciones, además de proveer una gran variedad de servicios financieros. Por ejemplo, la banca colombiana ha aumentado notablemente la participación de las operaciones del mercado monetario en su hoja de balance entre 1999 y 2012 (de 1,7% a 7,6% del activo, respectivamente), y esto sin contar que después de la crisis financiera de finales de la década de los noventa ha venido expandiendo sus servicios financieros y cobrando por todos ellos. Cabe resaltar que los demás estudios elaborados sobre la eficiencia del sistema bancario colombiano se han concentrado en el enfoque de intermediación.

Las variables que se eligieron para realizar el presente estudio están basadas, entonces, en el enfoque de producción (o lo que es denominado por otros autores el enfoque operacional) que supone que los EB toman como insumos los gastos administrativos y laborales, junto con los egresos financieros, para convertirlos en ingresos por intereses y otros ingresos operacionales. Estas variables se eligieron teniendo en cuenta el análisis del negocio bancario colombiano, el estudio realizado por Jemric y Vujcic (2002) y la compilación de estudios a nivel internacional realizada por Berger y Humphrey (1997), entre otros (Cuadro 2).

Los ingresos por intereses y los otros ingresos operacionales sirven para estandarizar los distintos productos bancarios, por cuanto no todos los bancos tienen los mismos productos ni se especializan en los mismos nichos de mercado. Esto no es posible mediante el enfoque de intermediación, debido a que mientras la cartera e inversiones se pueden cuantificar por su volumen, los demás productos no²². Dentro de los productos se omitieron los ingresos operacionales, puesto que éstos responden a fuentes que no corresponden al objeto social de las entidades analizadas.

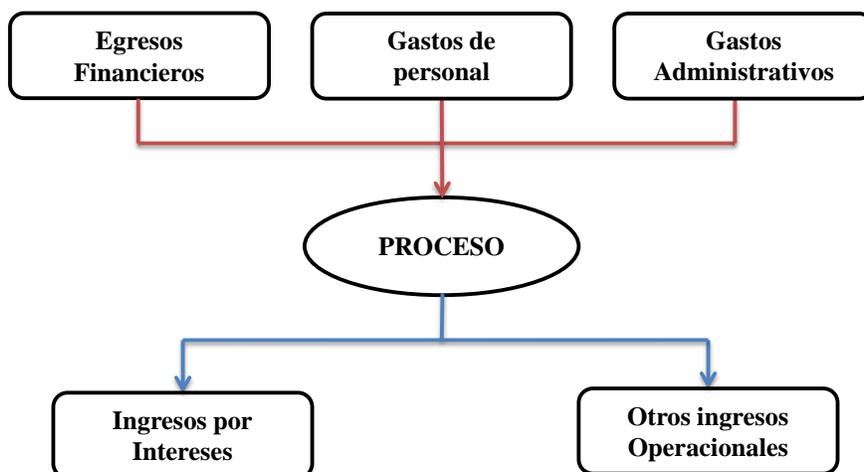
Esta selección se hizo, además, considerando que en el negocio bancario la calidad de los productos no es homogénea, debido, por ejemplo, a que la calidad de la cartera es diferente para cada intermediario financiero. En este sentido, a pesar de que dos bancos tengan un alto dinamismo en la cartera (suponiendo una utilización igual de insumos), no se podría afirmar que tienen el mismo nivel de eficiencia si la calidad no es la misma.

²² Es difícil establecer una misma unidad de medida para cuantificar los diferentes productos de los bancos de manera agregada, tal como la cartera junto con las comisiones derivadas de la prestación de servicios financieros, ya que la primera es expresada en volumen, en tanto que la segunda es expresada en ingresos (precio por volumen), por lo que en la mayoría de los trabajos realizados bajo el enfoque de intermediación tienen en cuenta exclusivamente la cartera e inversiones.

Ahora bien, el problema de calidad podría ser superado descontando las provisiones de la cartera. Sin embargo, el nivel de provisiones no representa cabalmente el deterioro de la misma, pues también está determinado por factores como la tolerancia al riesgo de la entidad, sus modelos internos de riesgo crediticio, los cambios regulatorios, entre otros. De esta manera, resulta plausible pensar que los ingresos financieros (que forman parte de los ingresos operacionales) provenientes de la cartera productiva, inversiones y de otros servicios financieros ofrecidos, son una buena forma de estandarizar el producto entre los intermediarios financieros, en el sentido de que implícitamente excluyen todos aquellos activos que no contribuyen a generar rentabilidad al negocio.

En cuanto a los insumos, se seleccionaron los gastos laborales y administrativos más los egresos financieros, ya que éstos son una *proxy* de la cantidad de trabajo y capital usados en el proceso productivo. Dentro de los gastos administrativos se encuentran rubros que tienen en cuenta los gastos de capital tales como la depreciación de activos fijos, arriendos, contribuciones y otros gastos fijos involucrados en el proceso (Gráfica 2). Los egresos financieros también se incluyeron porque, además de haber sido utilizados en la literatura, constituyen el gasto de capital más importante de la banca en su labor de intermediación (Cuadro 3).

Gráfica 2. Proceso productivo de los EB



Todas las variables provenientes de los estados financieros de los bancos, reportados a la Superintendencia Financiera de Colombia (SFC), fueron anualizadas debido a que son variables flujo²³. Ello se hizo con el fin de desestacionalizar las series y capturar los efectos del retardo que en la práctica se presentan al transformar insumos en productos: los gastos no necesariamente se reflejan de forma inmediata en el producto. Con la anualización de la información también se logra una mayor estabilidad en los resultados del modelo.

²³ Este enfoque también tiene la ventaja de que las variables escogidas son flujo únicamente. Bajo el enfoque de intermediación hay una mezcla entre variables flujo y *stock*, lo cual puede tener diferentes implicaciones en los resultados.

El Cuadro 3 resume las estadísticas descriptivas de las variables correspondientes a la muestra de EB del período estudiado que se utilizaron en el modelo.

Cuadro 3. Estadísticas descriptivas de las variables insumo y producto

Variable*	Ingresos por intereses	Otros ingresos operacionales	Gastos laborales	Gastos administrativos	Egresos financieros
Media	790,514	237,221	151,308	243,323	325,672
Mediana	536,364	141,709	96,118	163,378	224,736
Desviación estándar	889,063	338,363	177,737	286,853	360,542
Rango	5,768,191	2,006,246	1,245,196	1,809,163	2,339,235
Mínimo	12,942	1	2,880	5,541	7,299
Máximo	5,781,133	2,006,247	1,248,076	1,814,704	2,346,534

*Cifras en millones de pesos

Fuente: SFC y cálculo de los autores.

Al observar la información se encuentra una alta dispersión en los datos tanto de los insumos como de los productos de los EB, indicando de alguna manera los diferentes tamaños que exhiben los negocios de las entidades de la industria bancaria. Por ejemplo, la mayor variable producto es el ingreso por intereses, cuyo promedio asciende a \$ 790,5 mil millones (mm), su mediana es de \$536,3 mm, aunque el rango de dicha variable va desde \$12,9 mm a \$5.781 mm. En cuanto a los insumos se destacan los egresos financieros, cuyo promedio alcanza \$325,7 mm y su mediana \$224,7 mm, con un rango que oscila entre \$7,3 mm a \$2.346,5

Cabe anotar que los gastos laborales y administrativos se toman por separado, pese a estar en las mismas unidades y corresponder a rubros similares, debido a que para el análisis de potenciales mejoramientos y niveles objetivo (para transformar una entidad ineficiente en eficiente), la información discriminada puede resultar mucho más útil dadas las facilidades del modelo. Asimismo, se debe mencionar que con la metodología DEA las unidades en las que se presentan las variables no son relevantes, por cuanto lo más importante es su comportamiento. Por tal razón, la medición de variables como las laborales y las de capital pueden ser dadas en volumen o en términos monetarios, siempre y cuando estos últimos estén deflactados y así se pueda remover el efecto inflacionario de dichas variables.

Debido a que a partir de los datos no se encontró evidencia de que existan retornos constantes a escala y en razón a que intuitivamente se espera que los retornos sean variables, según el nivel en el que se produce, el modelo seleccionado para la eficiencia técnica corresponde al modelo BCC (retornos variables a escala) en su versión de maximización (enfoque orientado hacia producto). De igual forma, se eligió este enfoque por estar más acorde con las características del sector, por cuanto éste se concentra en aumentar el producto y obtener el máximo de beneficio, restando importancia a la restricción en los costos (Mohan y Ray, 2004). Además, en el enfoque de maximización, se hacen más notorias las distancias entre las entidades ineficientes y la frontera de eficiencia que cuando se usa el enfoque de minimización de costos, ya que las entidades tratan de restringirlos al máximo bajo combinaciones técnicas similares de insumos (Dasa y Ghosh, 2009). El modelo elegido para la eficiencia asignativa fue el descrito anteriormente. Para este modelo se definieron los costos unitarios de la siguiente manera:

$$\text{Costo Financiero} = \frac{\text{Egresos Financieros}}{\text{Depósitos y Exigibilidades}}$$

$$\text{Costo de Personal} = \frac{\text{Gasto de Personal}}{\text{Número de empleados}}$$

$$\text{Costo Administrativo} = \frac{\text{Gastos Administrativos}}{\text{Activos Fijos}}$$

En razón a que se quiere tener una visión global del desempeño de la eficiencia de las entidades, se hace énfasis en el análisis de la eficiencia asignativa, ya que esta se define con base en las otras nociones de eficiencia (tanto técnica como de costos).

6. RESULTADOS

Para evaluar la eficiencia del sistema financiero colombiano y su evolución, se tomaron las series anuales provenientes de los estados financieros mensuales reportados a la SFC tanto de los productos como de los insumos que se escogieron (junto a su costo unitario) para cada entidad, desde el año 2000 hasta 2012. Se escogió para el análisis un panel balanceado de 18 EB, los cuales han permanecido en el mercado durante el período señalado.

Adicionalmente, un banco en cada momento del tiempo (año a año) representa una unidad de producción diferente, por lo que puede ser comparado consigo mismo en diferentes fechas, bajo una misma frontera de eficiencia obtenida durante todo el período de análisis que representa las mejores prácticas del mercado alcanzadas por dichas unidades. En este sentido, un banco logrará una eficiencia mayor (ineficiencia menor) entre más se aproxime a las prácticas mencionadas.

Existe la posibilidad de que las entidades bancarias eficientes en este trabajo sean ineficientes en otros, cuando nuevos bancos son incluidos en el período de estudio. Por lo tanto, es necesario que los análisis sobre eficiencia bajo esta metodología sean permanentemente revisados, a fin de que el estudio sea más confiable.

Como ya se explicó en la sección anterior, este trabajo se concentra en la eficiencia orientada a maximizar los productos, dado un nivel determinado de insumos bajo el enfoque de producción. Igualmente, se realizó un ejercicio incorporando variables exógenas, de la forma propuesta por Banker y Morey (1986)²⁴. Las variables que se incluyeron fueron el índice de estabilidad financiera (IEFI), el cual es un indicador de

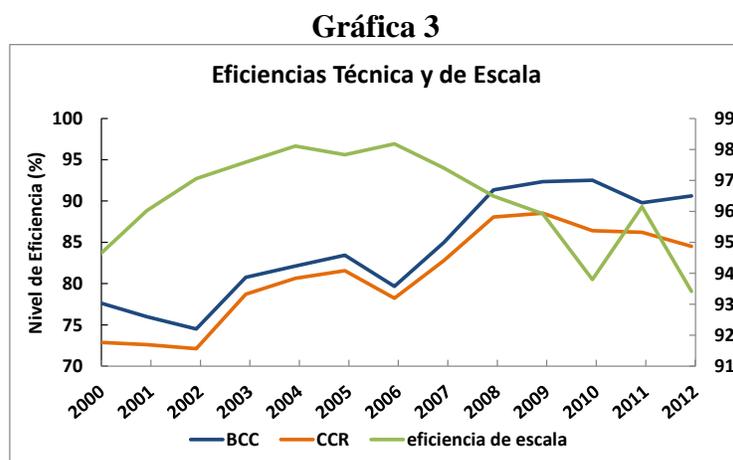
²⁴ Estos autores ampliaron el análisis de los modelos tradicionales (BCC y CCR) mediante la inclusión de variables ambientales, que anteriormente no se consideraban por el hecho de que las mismas no estaban bajo el control de la administración de la unidad de producción. Sin embargo, los autores encontraron que estas variables tienen impacto sobre los niveles de eficiencia.

riesgo, y la tasa de colocación del portafolio de cartera ex ante, que captura la postura efectiva de política monetaria. Dichas variables pueden afectar la eficiencia de la banca de manera positiva o negativa.

6.1. Eficiencias técnica, de escala, de costos y asignativa

La **eficiencia técnica** obtenida por los establecimientos bancarios durante el lapso de estudio (2000- 2012) bajo el enfoque BCC orientado a productos (maximización) fue estimada en promedio en 84,29%, en tanto que para CCR fue de 81,02% (Gráfica 3). De esta manera, dado que el enfoque de BCC asume retornos variables y el de CCR incorpora retornos constantes, los resultados son consistentes con la teoría, donde la eficiencia del primer enfoque es superior a la del segundo.

Al analizar por sub-períodos se encuentra que durante los primeros años de la década del 2000, los bancos, que se encontraban en una fase de recuperación, mostraban los más bajos niveles de eficiencia de todo el período de estudio, lo que coincidió con la finalización de la crisis financiera más profunda que haya experimentado el sector financiero colombiano en el pasado reciente. En efecto, entre 2000 y 2002 la eficiencia técnica se mantuvo en niveles del 75% en comparación con la mejor práctica del sistema bancario²⁵. Ya a partir de 2002, el indicador de eficiencia empezó a mejorar y entre 2006 y 2008 se acentuó hasta estabilizarse en los siguientes años alrededor del 91% para BCC y del 87% para CCR (Gráfica 3).



Fuente: Superintendencia Financiera de Colombia (SFC) y cálculo de los autores.

El comportamiento descrito de la eficiencia de la banca colombiana es consistente con los resultados que se obtuvieron en el estudio de Sarmiento *et al.* (2013), aunque aquellos

²⁵ Es importante destacar que al evaluar entre entidades, si el indicador de eficiencia técnica de un banco es, por ejemplo, de 33%, mientras que ese mismo para otro banco es de 100%, no se puede interpretar como que la eficiencia técnica del primero es la tercera parte de la del segundo; más bien se debe considerar que si la entidad de menor puntaje llegase a adoptar las prácticas de la más eficiente, mejoraría su desempeño. En este sentido, el mencionado indicador permite a los bancos organizarlos de manera ordinal.

presentaron niveles más bajos de eficiencia técnica, lo cual podría explicarse, entre otras cosas, por el enfoque de intermediación que no contempla varios de los productos bancarios que han tomado trascendencia dentro de los ingresos operacionales de la banca colombiana en los años recientes. Igualmente, estos resultados son coherentes con los obtenidos en el estudio de Chansarn (2008), cuya eficiencia de la banca tailandesa, considerando el enfoque de producción, era más alta que bajo la perspectiva de intermediación.

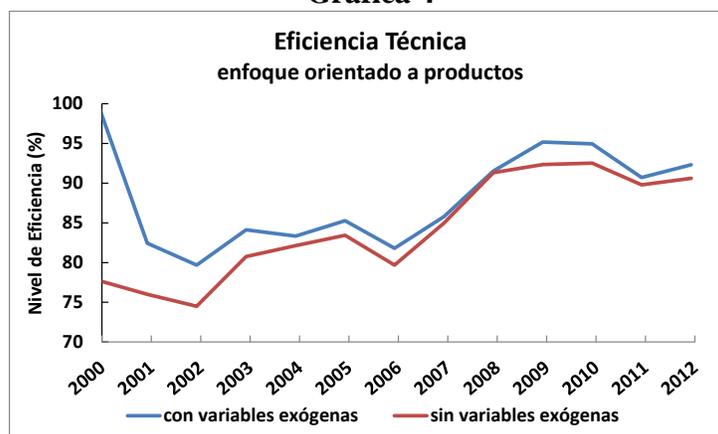
Cabe mencionar que durante la década del 2000 sucedieron en Colombia procesos de reorganización institucional que incluyeron aumentos de capital extranjero (junto con la entrada de algunos establecimientos de crédito), el fortalecimiento de la regulación financiera (siguiendo medidas macroprudenciales propuestas por Basilea) y un ambiente de liquidez que facilitó el crecimiento del sector financiero colombiano. No obstante, a partir del año 2007 se registraron algunos hechos que contribuyeron probablemente a alterar la eficiencia del conjunto de bancos: cambios en la postura de política monetaria y repercusiones en Latinoamérica de la crisis financiera internacional de 2008.

Haciendo un análisis más minucioso de los resultados obtenidos del modelo, y teniendo en cuenta otra de las virtudes del mismo respecto a la eficiencia técnica, se advierte que los bancos ineficientes tienen dos opciones para llegar a ser eficientes: disminuir la cantidad de insumos que están siendo utilizados (enfoque de minimización) o aumentar el nivel de producción (enfoque de maximización). En el primer caso, se encontró que en promedio los bancos ineficientes deberían reducir sus recursos orientados al área laboral en una mayor proporción que para el área administrativa y para sus egresos financieros, a fin de alcanzar la frontera de eficiencia. En el segundo caso, los ingresos por intereses y los otros ingresos operacionales deberían aumentar en proporciones similares para que se vuelvan eficientes.

De otro lado, **la eficiencia de escala**, que relaciona el enfoque CCR frente al de BCC, alcanzó en promedio durante todo el lapso estudiado 96,34%, lo que significa que la banca ha venido operando en una escala próxima a la óptima. Entre 2000 y 2004, la eficiencia de escala venía en aumento hasta situarse en los siguientes dos años en 98%; desde ese entonces se revirtió la tendencia ubicándose en diciembre de 2012 en 93% (Gráfica 3). En estas circunstancias, durante buena parte de la década del 2000 los bancos sacaron provecho de los retornos crecientes para incrementar su eficiencia; es decir, que después de mantener bajos índices de eficiencia había espacio para reducir el costo marginal promedio de los insumos mediante aumentos en el volumen de los productos bancarios, lo que se llevó a la práctica en algunos casos a través de procesos de fusiones y absorciones durante ese período.

Las variables ambientales, a su turno, tuvieron un papel positivo en el desempeño agregado de la eficiencia técnica. En otras palabras, la disminución de la aversión al riesgo y los niveles relativamente estables de las tasas de interés, pudieron haber afianzado la mejora de los resultados de la eficiencia técnica de las entidades. Así, al incluir en el modelo DEA las mencionadas variables, los resultados muestran un nivel promedio de eficiencia técnica superior en relación con el modelo que no las considera (88,24% vs. 84,29%, respectivamente), aunque las tendencias observadas al interior del período de estudio son similares (Gráfica 4).

Gráfica 4



Fuente: SFC y cálculo de los autores.

En cuanto a los resultados por entidades, el Cuadro 4 presenta los principales indicadores de aquellos bancos que se destacan por su elevada eficiencia técnica. La segunda columna señala el número de veces en que una entidad resultó eficiente durante el período analizado, con lo cual se muestra la persistencia en los niveles de eficiencia. La tercera columna presenta el número de veces en que una entidad, además de ser eficiente durante el período, fue usada como referente²⁶ para encontrar los niveles de eficiencia de las demás.

Cuadro 4. Entidades eficientes y referentes

Banco	No de veces que fue		Participación %	
	Eficiente	Referente	Eficiente	Referente
B1	6	46	8.2%	5.3%
B2	1	1	1.4%	0.1%
B3	4	49	5.5%	5.7%
B4	4	31	5.5%	3.6%
B5	1	1	1.4%	0.1%
B6	3	65	4.1%	7.5%
B7	6	68	8.2%	7.8%
B8	3	15	4.1%	1.7%
B9	5	26	6.8%	3.0%
B10	4	15	5.5%	1.7%
B11	10	229	13.7%	26.4%
B12	1	1	1.4%	0.1%
B13	8	180	11.0%	20.8%
B14	1	17	1.4%	2.0%
B15	1	1	1.4%	0.1%
B16	6	27	8.2%	3.1%
B17	7	92	9.6%	10.6%
B18	2	3	2.7%	0.3%
Total	73	867	100.0%	100.0%
Coef. Corr	85%			

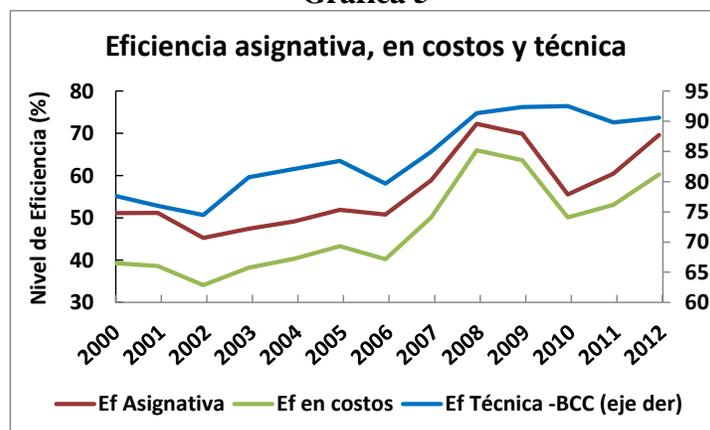
Fuente: SFC y cálculo de los autores.

²⁶ Para que una entidad sea referente debe ser eficiente y tener una estructura productiva similar a la de las entidades que se comparan frente a ella.

Los resultados obtenidos individualmente muestran de manera consistente (con un coeficiente de correlación del 85%), que el banco que más veces fue eficiente es el mismo que más fue utilizado como referente. Esta última característica tiene una especial connotación porque además de señalar que una entidad fue varias veces eficiente, también revela que aquella posee una mezcla similar de insumos y productos en relación con las demás. Por ejemplo, dos entidades se pueden mantener sobre la frontera eficiente, pero con una mezcla diferente de productos: una orientada hacia cartera y otra hacia otros productos (inversiones, transacciones en el mercado monetario, etc.). En ese caso, si la mayor parte de entidades se encuentran orientadas hacia cartera, se compararán con la primera y no con la segunda. Este tratamiento, que resalta una de las ventajas del modelo, permite la comparación de los bancos según su especialización por producto, y no los considera a todos homogéneos.

En cuanto a **la eficiencia en costos** se observa una tendencia similar a la de la eficiencia técnica (Gráfica 5); sin embargo, los niveles de la eficiencia en costos alcanzados por el conjunto de bancos son inferiores a los de la técnica. Ello es consistente con la teoría, en el sentido en que esta última no siempre está reflejando una minimización de costos, dada la mezcla utilizada de insumos. En otras palabras, el uso de determinadas proporciones de insumos puede llevar a una firma a ser eficiente técnicamente, pero no necesariamente a minimizar sus costos de producción.

Gráfica 5



Fuente: SFC y cálculo de los autores.

Así, en el periodo inicial hasta 2006, las entidades venían presentando una estructura de costos, que distaba de la óptima, entendida en términos de aquellas realizadas por las entidades con mejores prácticas. El importante crecimiento de los activos que experimentaron los bancos, originados en el aumento de la cartera, inversiones y otros productos y servicios financieros, desde finales del mencionado año, hicieron que la eficiencia en costos aumentara ostensiblemente hasta el año 2008, ubicándose en 66%. De allí en adelante, dicha eficiencia fue alterada por cambios en la dinámica de la producción

de la industria bancaria, en donde variables externas acaecidas en ese momento pudieron haberla afectado.

La **eficiencia asignativa**, a diferencia de la eficiencia técnica y la de costos, asume que ante cambios en los precios de los insumos, la firma puede modificar su mezcla para obtener nuevamente una combinación óptima de los mismos. De este modo, dicha eficiencia, que relaciona la eficiencia en costos sobre la técnica, si bien sigue un comportamiento similar a éstas, es superior a la primera pero inferior a la técnica (Gráfica 5).

Sin embargo, se podría esperar que teniendo en cuenta los productos y servicios que ofrece la banca colombiana, los insumos utilizados no son sustitutos perfectos, y en esa medida no se puede modificar su mezcla técnica para alcanzar una combinación óptima, por lo menos en el corto plazo, una vez se dan cambios en los precios relativos de tales insumos.

Después de la crisis financiera de finales de la década de los noventa, la eficiencia de asignación distó de alcanzar las mejores prácticas desarrolladas por los bancos que trazaron su frontera. Luego en la etapa de recuperación de la banca, se presentaron mejoras en la elección adecuada de insumos, de acuerdo con su costo para la generación de productos. Desde finales de 2006, dicha eficiencia también se vio afectada por cambios en la oferta de productos y servicios bancarios que necesariamente repercutieron en la combinación óptima de insumos.

6.2. Índice de Malmquist

Tal como se mencionó anteriormente, el índice de Malmquist permite la separación de los cambios de la productividad total de los factores entre eficiencia técnica y el cambio tecnológico, a través, de una parte, de aumentos o reducciones en la eficiencia de las unidades de producción en relación con la frontera de eficiencia estimada para todo el período de análisis y, de otra, del desplazamiento anual de dicha frontera, respectivamente.

Los resultados del índice de Malmquist muestran que el incremento de la productividad se debió fundamentalmente a cambios tecnológicos y, en menor medida, a cambios en eficiencia entre un año y otro para cada unidad de producción. En efecto, se encuentra que la variación promedio anual de la eficiencia técnica aumentó en 0,31% durante todo el período. Esta mejora estuvo acompañada de un cambio tecnológico de 3,16% en promedio (Cuadro 5).

Cuadro 5. Índice de Malmquist

Fecha	Índice de Malmquist	Cambio Eficiencia Técnica	Cambio Tecnológico
2000	0.9840	1.0311	0.9595
2001	0.9498	0.9874	0.9631
2002	1.0600	1.0121	1.0441
2003	0.9835	0.9734	1.0159
2004	1.0510	1.0207	1.0326
2005	1.0169	0.9826	1.0363
2006	1.0478	1.0036	1.0448
2007	1.0475	1.0322	1.0163
2008	1.0843	0.9857	1.0993
2009	1.1375	0.9888	1.1476
2010	1.0741	1.0208	1.0524
2011	0.9429	0.9776	0.9670
2012	1.0546	1.0242	1.0312

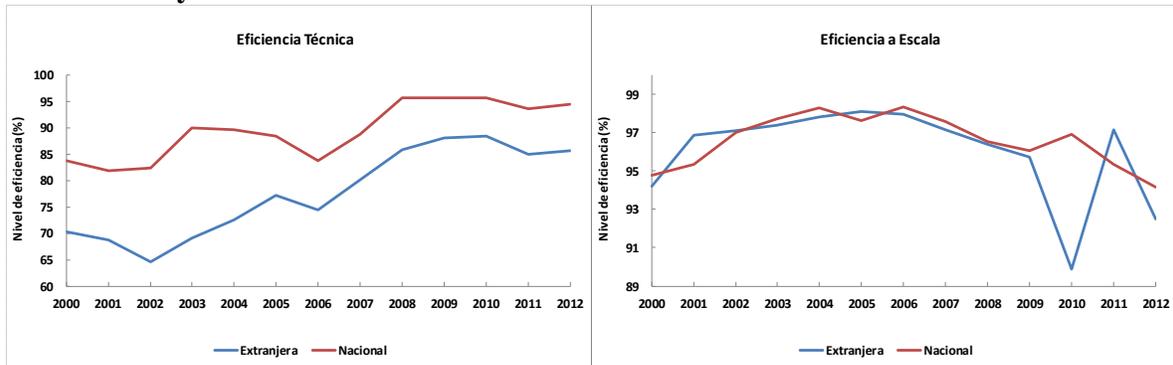
Fuente: SFC y cálculo de los autores.

Se resalta el avance que presentó el sistema bancario en lo que corresponde al cambio tecnológico entre 2008 y 2009, en tanto que para ese lapso se observó una ligera reducción de la eficiencia técnica.

6.3. Eficiencia y propiedad de los EB

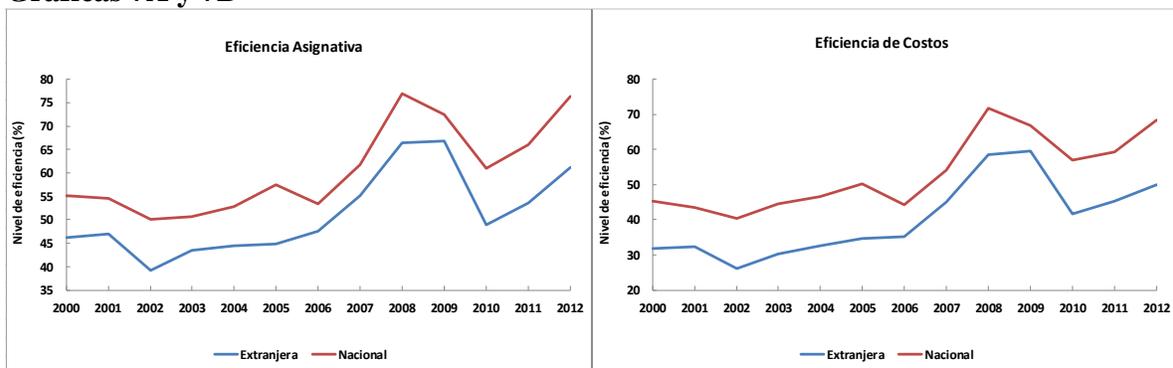
Con la presente metodología se realizó un ejercicio, separando los bancos según su propiedad (nacional o extranjera) para el lapso comprendido entre 2000 y 2012, bajo el enfoque BCC orientado a productos. Allí se encontró que el conjunto de bancos locales alcanza mayores niveles de eficiencia (técnica, asignativa y de costos) en comparación con la banca extranjera, aun cuando el comportamiento (tendencias) es muy similar. Por su parte, la eficiencia de escala para ambas clases de bancos se mantiene en niveles similares para todo el período de estudio (Gráficas 6 A y B, y 7 A y B). De este modo, mientras el promedio de eficiencia técnica del agregado de los bancos locales es de 89,6%, el de la banca con capital extranjero es de 77,7%; asimismo, ocurre con la eficiencia asignativa (60,7% frente a 51,1%) y la eficiencia de costos (53,3% vs. 40,3%). La eficiencia de escala, por su lado, fue muy parecida para ambos grupos hasta el año 2009, fecha a partir de la cual dicha eficiencia se hizo muy volátil para el segundo grupo, donde probablemente la crisis internacional pudo haberla afectado más que a la banca local.

Gráficas 6A y 6B



Fuente: SFC y cálculo de los autores.

Gráficas 7A y 7B



Fuente: SFC y cálculo de los autores.

6.4. Eficiencia y tamaño de los EB

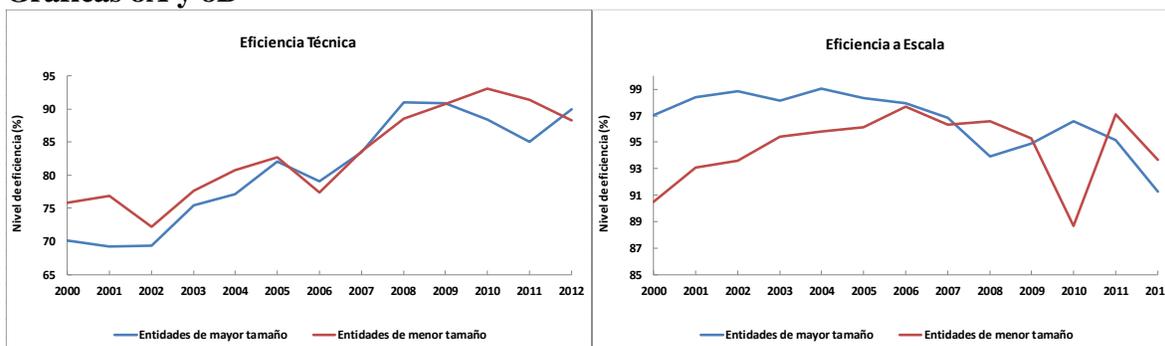
A fin de analizar el efecto del tamaño de las entidades sobre la eficiencia, se realizó un ejercicio que tuvo en cuenta los seis bancos más grandes y los seis más pequeños con el fin de comparar el comportamiento de ambos grupos. Para el efecto, se ordenaron los EB jerárquicamente según el promedio de los activos durante el período de estudio.

Contrario a los postulados de la teoría microeconómica, donde se afirma que los bancos grandes son más eficientes que los chicos (gracias a que los costos fijos promedio disminuyen en la medida que aumenta el tamaño de la firma), en este caso no hay clara evidencia de que ello sea así. Los EB grandes obtuvieron en promedio durante todo el lapso de análisis un nivel de eficiencia técnica inferior al grupo de los pequeños (80,9% vs. 83,0%, respectivamente). Este mismo comportamiento se reflejó en las eficiencias de costos y asignativa. Tan solo los resultados de la eficiencia de escala mostraron que los EB grandes alcanzaban un nivel más alto en promedio frente a los pequeños (96,7% vs. 94,6%, respectivamente). De hecho, los mayores crecimientos en este tipo de eficiencia fue logrado por el grupo de los grandes entre 2000 y 2006; de allí en adelante ambos grupos se alejaron del tamaño óptimo de la industria bancaria (Gráficas 8 A y B, y 9 A y B). De otra parte, la utilización de insumos fue mucho más alta en los bancos grandes frente a las mejores prácticas que en los pequeños para el primer sub-período (2000-2006), lo cual es coherente

con los efectos de la poscrisis sobre los recursos utilizados por los bancos de mayor tamaño al tener que frenar abruptamente la dinámica en los niveles de productos y servicios financieros ofrecidos para esa época.

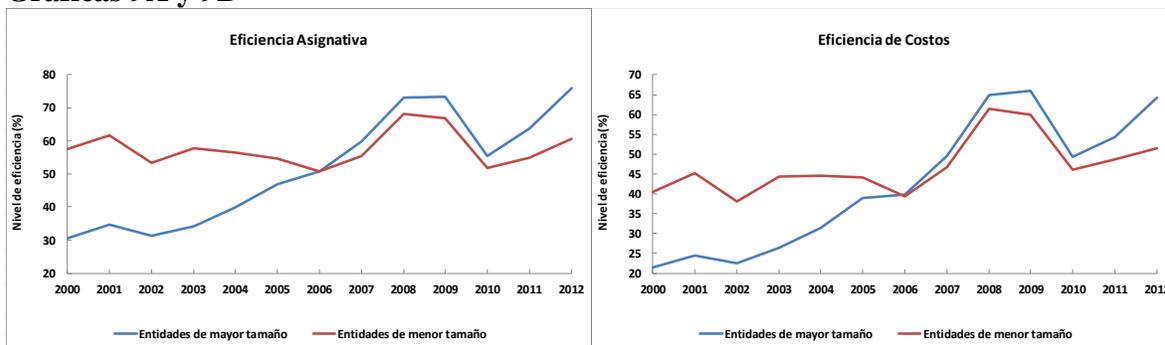
Una explicación de por qué no es claro que los bancos grandes sean más eficientes que los pequeños reside en que los primeros tienen un buen porcentaje del mercado asegurado, en tanto que los pequeños están permanentemente tratando de alcanzar una mayor participación del mismo. Esta ventaja comparativa de los EB de gran tamaño hace que se relajen para mantener altos niveles de eficiencia, mientras que los chicos buscan de manera persistente aumentarla en todo momento (Fernández y Estrada, 2013). Esta disquisición es complementaria con la que afirma que la mayor eficiencia de los EB grandes no es aprovechada para sacar a los bancos pequeños del mercado, sino que prefieren convivir con ellos asegurando ganancias extraordinarias derivadas de dicha ventaja.

Gráficas 8A y 8B



Fuente: SFC y cálculo de los autores.

Gráficas 9A y 9B



Fuente: SFC y cálculo de los autores.

6. 5. Determinantes de la Eficiencia

Dado que la eficiencia está sujeta a potenciales influencias de variables relacionadas con la estructura de operación de las firmas y de algunas variables ambientales como lo demuestra Pinilla (2001), se realizó un análisis econométrico con datos panel para estimar el aporte de algunas de éstas en los cambios de la eficiencia de los EB. Las variables seleccionadas

fueron: la solvencia, que es la razón entre patrimonio y activos (SOLV), la participación de la cartera comercial en el total de la cartera bruta (CCIAL), el índice de Herfindahl-Hirschman²⁷ (IHH), y el índice de confianza del consumidor (ICC). Cabe anotar que estas variables se eligieron con base en los trabajos previos realizados para Colombia por Badel (2002) y Castro (2001), y el de Kenjegalieva y Simper (2010) para otros países (Anexo 2).

De una parte, la solvencia y la eficiencia de la banca son elementos que muestran causalidad entre sí. En este sentido, en teoría los bancos menos eficientes generan menores utilidades y, por lo tanto, tienden a ser menos capitalizados; por consiguiente, los bancos más eficientes serían los más solventes. No obstante, la causalidad puede ser inversa al considerar el “riesgo moral”, que sugiere que los bancos que van inminentemente a la quiebra y sus administradores, pueden perseguir sus propios intereses sin considerar los efectos sobre la eficiencia, de manera que la baja capitalización puede llevar a períodos de alta ineficiencia (Berger y Humphrey, 1997). Dada la mencionada característica de causalidad en dos sentidos que puede tener esta variable y los posibles problemas consecuentes de endogeneidad, esta variable se incluye en el modelo de forma rezagada.

La especialización o el nicho de mercado es otra de las variables del negocio bancario que afecta la eficiencia de cada EB. La banca concentrada en préstamos de consumo o de microcrédito tiene una cantidad muy elevada de deudores con bajos montos de crédito, lo que exige mayores costos promedio para administrar este tipo de cartera. Por otro lado, los bancos especializados en banca corporativa requieren menos recursos (administrativos y laborales) por crédito administrado. La variable utilizada es la proporción de cartera comercial sobre el total de cartera, que además captura la heterogeneidad del producto principal de la banca. Si este indicador es bajo, significa que el banco está concentrado en otras modalidades de cartera (consumo, microcrédito y vivienda), lo que requiere una estructura de costos más elevada, dado que se trata de un negocio más atomizado. Si por el contrario, el indicador es alto, su nicho de mercado está en el crédito comercial y/o corporativo, lo que implica, *ceteris paribus*, menos clientes, menores costos y, por lo tanto, una menor ineficiencia.

Las otras dos variables exógenas o ambientales son el ICC y un indicador que caracteriza la concentración en el sector bancario (IHH). Es importante tener en cuenta el comportamiento del ICC para explicar la eficiencia, dado que es un factor que afecta el proceso de decisión de la oferta de productos y servicios financieros de los bancos. Por ejemplo, éstos ajustarán su oferta si el ICC está cayendo para reflejar las condiciones de la demanda potencial (McWhinney, 2013). Otra forma de relacionar la eficiencia con el ICC es que en la medida en que haya un aumento generalizado en la confianza de la economía, se reduce la incertidumbre en el futuro cercano y, en consecuencia, aumenta el crecimiento de los diferentes sectores económicos y su eficiencia.

En el caso del IHH, existen teorías en la economía de la información que muestran que las presiones competitivas son la forma más efectiva de promover la eficiencia productiva. Empíricamente se ha demostrado como por ejemplo en el trabajo de Jemric y Vujcic (2002), que el desempeño productivo está relacionado positivamente con el grado de

²⁷ Este índice cuantifica el nivel de concentración del mercado.

concentración, encontrando incluso que su efecto es más importante que la misma naturaleza de la propiedad (privada o pública). Sin embargo, hay también evidencia empírica que demuestra que entre más concentrado se encuentran algunos sectores como el sector financiero, mayores niveles de eficiencia se obtienen.

Ahora bien, para este ejercicio se utilizó el mismo panel balanceado de 18 EB que se utilizó en el cálculo de la eficiencia con una periodicidad anual e información para cada una de las variables propuestas (en total 1170 observaciones). Con base en los resultados estimados del modelo DEA para las diferentes eficiencias, la variable dependiente seleccionada fue la eficiencia técnica (ET), debido a que ésta recopila las tendencias de la eficiencia asignativa y de costos gracias a la relación directa que tiene con éstas²⁸. La siguiente es la regresión estimada:

$$ET_{i,t} = CONST + ICC_{i,t} + SOLV_{i,t-1} + IHH_{i,t} + CCIAL_{i,t} + \mu_{i,t}$$

Donde $\mu_{i,t}$ corresponde al término de perturbación.

La estimación se hizo usando el modelo de regresión truncada en el que el término de error tiene una distribución normal truncada²⁹. Este modelo se eligió tomando en cuenta que la variable dependiente se encuentra acotada entre 0 y 100%, y que no existe evidencia de la existencia de efectos fijos. En efecto, al comparar los resultados de los modelos de efectos fijos y aleatorios a través de la prueba de Hausman, que tiene como hipótesis nula que no hay diferencia entre los coeficientes de los modelos, y dado un p-valor de 0,8351 mayor a un nivel de significancia de 0,05 se concluye que no se rechaza la hipótesis nula. Los resultados de la estimación se presentan en el Cuadro 5³⁰.

Como se observa en el Cuadro 5 todas las variables resultan ser estadísticamente significativas, usando diferentes niveles de confianza. En particular, se observa que el nivel de solvencia es la variable que más aporta en los cambios de la eficiencia técnica. Por su parte, la variable de especialización del negocio bancario (CCIAL) tiene un efecto negativo sobre la eficiencia, en tanto que el ICC y el IHH tienen efectos positivos pero marginales sobre la misma. Todo lo anterior permite afirmar que no sólo las variables que están bajo el control de la administración de las entidades bancarias determinan los cambios en eficiencia, sino también las variables ambientales ajenas al control de la gestión de intermediación de los EB.

²⁸ Cabe anotar que al realizar ejercicios tomando como variables dependientes la eficiencia de costos o la asignativa, se mantienen las relaciones (Anexo 3).

²⁹ Está es una distribución que ha sido escalada de tal forma que tiene una integral igual a uno sobre el área restringida.

³⁰ Se resalta que al realizar las pruebas sobre los residuales se encontró que éstos no tenían problemas de autocorrelación ni de heteroscedasticidad.

Cuadro 6. Resultados de la estimación de los determinantes de la eficiencia

Variable Dependiente: <i>ET</i>	
Variables Independientes:	
<i>ICC</i>	0.0020379*** (0.0007561)
<i>SOLV(-1)</i>	1.332451*** (0.3707102)
<i>IHH</i>	0.0000563*** (0.0000123)
<i>CCIAL</i>	-0.2730919*** (0.0489456)
<i>CONST</i>	0.7750042*** (0.0462989)
Efectos Fijos	
	NO
Número de obs	1170
P valor Hausman	0.8351
R ²	0.2245228

Nota: Errores estándar en paréntesis

*** p<0.01, ** p<0.05, * p<0.1

7. CONCLUSIONES

La metodología elegida tiene importantes ventajas sobre otras, en la medida en que los resultados brindan una mayor información para el análisis de la eficiencia de la industria bancaria. El modelo aquí propuesto presenta resultados en términos relativos, a diferencia de los modelos paramétricos cuyas técnicas exigen funciones de distribución predeterminadas, haciendo que cada entidad tenga su propia función de costos y de producción, y por lo tanto sea más flexible, con lo cual se tienen menos supuestos y unos resultados que se acercan más a la realidad del mercado.

El estudio se concentró en el enfoque de producción, con retornos variables (BCC), orientado a la maximización de los productos dado un nivel de insumos, a través de la utilización de la metodología DEA. Después de hacer un cuidadoso análisis de las variables a elegir, fueron tomados como insumos los gastos administrativos y laborales, junto con los egresos financieros, para convertirlos en ingresos por intereses y otros ingresos operacionales. Ambos ingresos representan la amplia heterogeneidad de productos del negocio bancario colombiano.

Mediante técnicas que proporcionan los modelos no paramétricos se obtuvieron resultados de las mejores prácticas de los bancos (frontera de eficiencia), analizando específicamente la eficiencia técnica, de escala, de costos y asignativa. Se encontró, de este modo, que en los primeros años de la década del 2000, al finalizar la crisis financiera más profunda

exhibida por el sector financiero colombiano, los EB alcanzaron sus más bajos índices de eficiencia de todo el período de estudio, mejorando entre 2006 y 2008. De allí en adelante, los niveles de eficiencia se alejaron de su tendencia creciente, con algunas fluctuaciones hasta el año 2012. Cabe señalar que dicho comportamiento de la eficiencia de la banca es consistente con los resultados que se obtuvieron en el estudio de Sarmiento *et al.* (2013).

La eficiencia asignativa reveló índices más bajos que la técnica, en razón a que los recursos utilizados por la banca para su transformación en productos no son sustitutos perfectos, y en esa medida se dificulta alcanzar una combinación óptima en el corto plazo, una vez se han dado cambios en sus precios relativos. Así, la eficiencia de asignación distó de alcanzar las mejores prácticas desarrolladas por los bancos que trazaron su frontera.

Igualmente se encontró que los orígenes de la ineficiencia técnica por parte de las entidades analizadas se centraban más en la sobreutilización de los gastos laborales que en los administrativos, y éstos a su vez más que en los egresos financieros.

También se analizó la eficiencia de los bancos según propiedad (nacional o extranjera). Se encontró que el conjunto de bancos locales alcanza mayores niveles de eficiencia (técnica, asignativa y de costos) en comparación con la banca extranjera, aun cuando siguen comportamientos (tendencias) similares. De la eficiencia de escala se resalta que se hizo muy volátil para el grupo de la banca extranjera desde 2007, donde es posible que la crisis internacional pudiera haberla afectado más que a la banca local.

De otra parte, no hubo clara evidencia de que la banca grande fuera más eficiente que la pequeña. Solamente los resultados de la eficiencia de escala mostraron que los EB grandes alcanzaban un nivel más alto en promedio frente a los segundos. Por otro lado, la utilización de insumos fue mucho más alta en los bancos grandes, dado sus mayores índices de ineficiencia, para el sub-período (2000-2006), lo cual parece ser coherente con los efectos de la poscrisis sobre los recursos utilizados por los bancos grandes al tener que frenar abruptamente la dinámica en los niveles de producción y servicios financieros ofrecidos.

Finalmente, mediante un modelo de regresión se concluyó que no solamente las variables que están bajo el control de gestión de los EB determinan los cambios en eficiencia, sino también las variables ambientales ajenas al control de la gestión de intermediación de estas entidades. Este hallazgo adquiere una especial connotación para los hacedores de política monetaria, crediticia y cambiaria, dado que los cambios en la postura de aquellas repercuten en la eficiencia del sistema financiero.

Los futuros trabajos locales dentro de la metodología DEA deberían enfocarse en la incorporación del error aleatorio en las funciones de producción y costos como explicación adicional a la ineficiencia de las entidades. De igual forma, sería importante realizar ejercicios para analizar cómo la especialización y el nicho de mercado en que se desarrolla la labor de intermediación de los diferentes bancos colombianos afectan su eficiencia.

7. REFERENCIAS

Badel, A. (2002). "Sistema bancario colombiano: ¿Somos eficientes a nivel internacional?", Archivos de Economía, Departamento Nacional de Planeación, Documento 190.

Banker, R., A. Charnes y W. Cooper (1984). "Some models for estimating technical and scale inefficiencies in data envelopment analysis", Management Science 30, 1078–1092

Banker, R. (1996). "Hypothesis tests using data envelopment analysis", The Journal of Productivity Analysis 7, 139-157.

Banker, R., H. Chang y S. Lee (2010). "Differential impact of Korean banking system reforms on bank productivity", Journal of Banking & finance 34, 1450-1460.

Berger, A. y D. Humphrey (1997). "Efficiency of financial institutions: international survey and directions for future research", European Journal of Operational Research 98, 175-212.

Banker, R. y R. Morey (1986). "Efficiency analysis for exogenously fixed inputs and outputs", Operations Research 34(4), 513 – 521.

Bernal, O. y S. Herrera (1983). "Producción, costos y economías de escala en el sistema bancario colombiano", Revista ESPE 3, 7-36.

Castro, C. (2001). "Eficiencia-X en el sector bancario colombiano", Desarrollo y Sociedad, 48, 1-52.

Chansarn, S. (2008). "The relative efficiency of commercial banks in Thailand: DEA Approach", International Research Journal of Finance and Economics, 18, 53-68.

Charnes, A., W. Cooper, A. Lewin y L. Seiford (1994). "DEA: theory, methodology, and application", Editorial Kluwer, Boston.

Dasa, A. y S. Ghosh (2009). "Financial deregulation and profit efficiency: A nonparametric analysis of Indian banks", Journal of Economics and Business 61, 509-528.

Elyasiani, E. y S. Mehdián (1990). "A nonparametric approach to measurement of efficiency and technological change: the case of large U.S. commercial banks", Journal of Financial Services Research 4, 157-168.

Estrada, D. y P. Osorio. (2004). "Effects of financial capital on Colombian banking efficiency", Borradores de Economía, 292.

Färe, R., S. Grosskopf, B. Lindgren y P. Ross (1992). "Productivity changes in Swedish pharmacies 1980-1989: a nonparametric Malmquist approach", Journal of Productivity Analysis 3, 85-101.

Farrell, M. J. (1957). "The measurement of productive efficiency", *Journal of the Royal Statistical Society, Series A (General)*, 120, pp. 253–289.

Fernández, D. y D. Estrada (2013). "Colombian bank efficiency and the role of market structure", *Temas de Estabilidad Financiera, Banco de la República*, 77

Forsund, F., K. Lovell y P. Schmidt (1980). "A survey of frontier production functions and of their relationship to efficiency measurement", *Journal of Econometrics* 13, 5-25.

Fried, H., K. Lovell y S. Schmidt (2008). "The measurement of productive efficiency and productivity growth", Editorial Oxford University Press.

González, A. y J. Martínez (2008). "Estimación de la frontera eficiente para el sector asegurador", *Revista Fasecolda*, 127, 28-36.

Janna, M. (2003). "Eficiencia en la banca: un recuento de la literatura para el caso colombiano", *Temas de Estabilidad Financiera, Reporte de Estabilidad financiera*.

Jemric, I. y B. Vujcic (2002). "Efficiency of banks in Croatia: a DEA approach", *Irving Fisher Committee on Central-Bank Statistics, IFC Bulletin (December)*, 106-202.

Kenjegalieva, K. y R. Simper (2010). "A Productivity analysis of Eastern European banking taking into account risk decomposition and environmental variables", *Discussion Paper Series, University of Loughborough (Julio/Agosto)*.

Koetter, M., A. Karman y E. Fiorentino (2006). "The Cost Efficiency of German Banks: A comparison of SFA and DEA", *Deutsche Bundesbank, Discussion Paper Series 2: Banking and Financial Studies*, 10.

Kwan, S. y R. Eisenbeis (1996). "An analysis of inefficiencies in banking: a stochastic cost frontier approach", *Federal Reserve Bank of San Francisco, Economic Review* 2, 16-26.

Leibenstein, H. (1966). "Allocative Efficiency vs. X-Efficiency", *The American Economic Review* 56(3), Junio, 392-415.

Martínez, J. y D. Estrada. (2009). "Análisis de eficiencia y productividad en la industria de seguros colombiana", *Documento de trabajo Fasecolda*.

McWhinney, J. (2013). "Understanding The Consumer Confidence Index", *Investopedia*, obtenido el 21 de Noviembre de 2013 desde <http://www.investopedia.com/articles/05/010604.asp>

Mester, L. (1997). "Measuring efficiency at US banks: accounting for heterogeneity is important", *European Journal of Operational Research* 92, 230-242.

Mohan, R. y S. Ray (2004). "Comparing Performance of Public and Private Sector Banks: A Revenue Maximisation Efficiency Approach", *Economic and Political Weekly*. 39 (12), 1271-1272+1274-1276.

Pinilla, A. (2001). *La medición de la eficiencia y la productividad*, 1ª Ed., Editorial Pirámide.

Portela, M. y E. Thanassoulis (2010). "Malmquist-type indices in the presence of negative data: An application to bank branches", *Journal of Banking & Finance* 34, 1472–1483.

Ramathan, R. (2003). *An introduction to Data Envelopment Analysis*", 1ª ed., Editorial Sage.

Sarmiento, M, A. Cepeda, H. Mutis y J. F. Pérez (2013). "Nueva Evidencia sobre la Eficiencia de la Banca Colombiana: Una medición con modelos de frontera no-paramétricos", *Archivos de Economía*, Documento 392, Departamento Nacional de Planeación.

Sarmiento, M. y J. Galán (2007). "Banknote printing at modern central banking: trends, costs and efficiency", *Borradores de Economía* 476, Banco de la República.

Wheelock, D. y P. Wilson (2006). "Robust Non-parametric Quantile Estimation of Efficiency and Productivity Change in U.S. Commercial Banking, 1985-2004", Federal Reserve Bank of St. Louis, Working Paper Series, 42.

Worthington, Andrew (2001). "An Empirical Survey of Frontier Efficiency Measurement Techniques in Education", *Education Economics* 9 (3): pp. 245-268.

Yang, Z. (2009). "Bank Branch Operating Efficiency: A DEA Approach", *Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009*, Vol II, March 18 - 20, Hong Kong.

ANEXO 1. Programas matemáticos de las formulaciones empleadas

Al normalizar el denominador del programa básico inicial de la eficiencia técnica, el programa queda de la siguiente manera:

$$\max z = \sum_{j=1}^J v_{jm} y_{jm}$$

Sujeto a:

$$\sum_{i=1}^I v_{im} x_{in} = 1$$

$$\sum_{j=1}^J v_{jm} y_{jn} - \sum_{i=1}^I v_{im} x_{in} \leq 0; \quad n = 1, 2, K, N$$

$$v_{jm}, v_{im} \geq 0^{33}; \quad i = 1, 2, K, I; \quad j = 1, 2, K, J;$$

La anterior constituye la formulación conocida como de maximización de producto. No obstante, si se normaliza el numerador y se minimiza la suma ponderada de insumos se puede obtener el programa de minimización de insumos. Éste puede ser representado de la siguiente manera:

$$\min z' = \sum_{i=1}^I v'_{im} x_{im}$$

Sujeto a:

$$\sum_{j=1}^J v_{jm} y_{jm} = 1$$

$$\sum_{j=1}^J v'_{jm} y_{jn} - \sum_{i=1}^I v'_{im} x_{in} \leq 0; \quad n = 1, 2, K, N$$

$$v'_{jm}, v'_{im} \geq 0; \quad i = 1, 2, K, I; \quad j = 1, 2, K, J;$$

El programa de maximización de producto puede ser reescrito en su forma dual de la siguiente manera³⁴:

$$\min \theta_m$$

Tal que:

³³ La formulación original fijaba que estos pesos fueran mayores o iguales a cero. Sin embargo los autores más adelante hicieron una modificación que cambiaba este valor a ε , el cual se define como una constante infinitesimal.

³⁴ Se puede escribir el dual de cualquier programa de programación lineal siguiendo ciertas reglas. La idea es que si es un problema de maximización se convierta en uno de minimización y viceversa; en ambos casos los valores óptimos de las funciones objetivo serán los mismos. Igualmente se espera que el número de coeficientes de la función objetivo se vuelva igual al número de restricciones y viceversa.

$$\sum_{n=1}^N y_{jn} \lambda_n \geq y_{jn}; \quad j = 1, 2, K, J$$

$$\sum_{n=1}^N x_{in} \lambda_n \leq \theta_m x_{im}; \quad i = 1, 2, K, I$$

$$\lambda_n \geq 0; \quad n = 1, 2, K, N$$

θ_m no restringido

Donde

θ_m es la variable dual correspondiente a la restricción de igualdad que normaliza la suma ponderada de insumos (normalización del denominador),

λ_n son las variables duales correspondientes al resto de restricciones de desigualdad del programa primal u original³⁵.

Esta forma dual es importante, ya que computacionalmente es menos exigente que la inicialmente presentada, pues mientras la primera tenía restricciones que dependían del número de unidades de producción analizadas, las restricciones de la forma dual dependen del número de insumos y productos (lo mismo ocurre con la forma dual del programa que corresponde a minimización de insumos). De igual forma esta formulación también es importante porque permite identificar las unidades que sirven de referencia³⁶.

Ahora bien, el anterior modelo se trabaja bajo el supuesto de rendimientos constantes a escala; sin embargo, si se agrega una restricción de convexidad sobre λ_n , este modelo permite retornos variables a escala (modelo BCC³⁷). El programa en este caso se escribiría como:

$$\min \theta_m$$

Tal que:

$$\sum_{n=1}^N y_{jn} \lambda_n \geq y_{jn}; \quad j = 1, 2, K, J$$

$$\sum_{n=1}^N x_{in} \lambda_n \leq \theta_m x_{im}; \quad i = 1, 2, K, I$$

³⁵ Donde $\sum_{n=1}^N \lambda_n$ indica la escala en la que la unidad de producción se desempeña. Cuando es mayor que uno se asocia con retornos decrecientes a escala, cuando es igual a uno se asocia con retornos constantes y cuando es menor que uno se asocia con retornos crecientes a escala.

³⁶ Cabe anotar que algunas veces el programa se define con variables de holgura para obtener restricciones de igualdad. De cualquier forma, en el óptimo el valor de θ_m debe ser igual a uno y las variables de holgura asociadas a las restricciones deben ser iguales a cero.

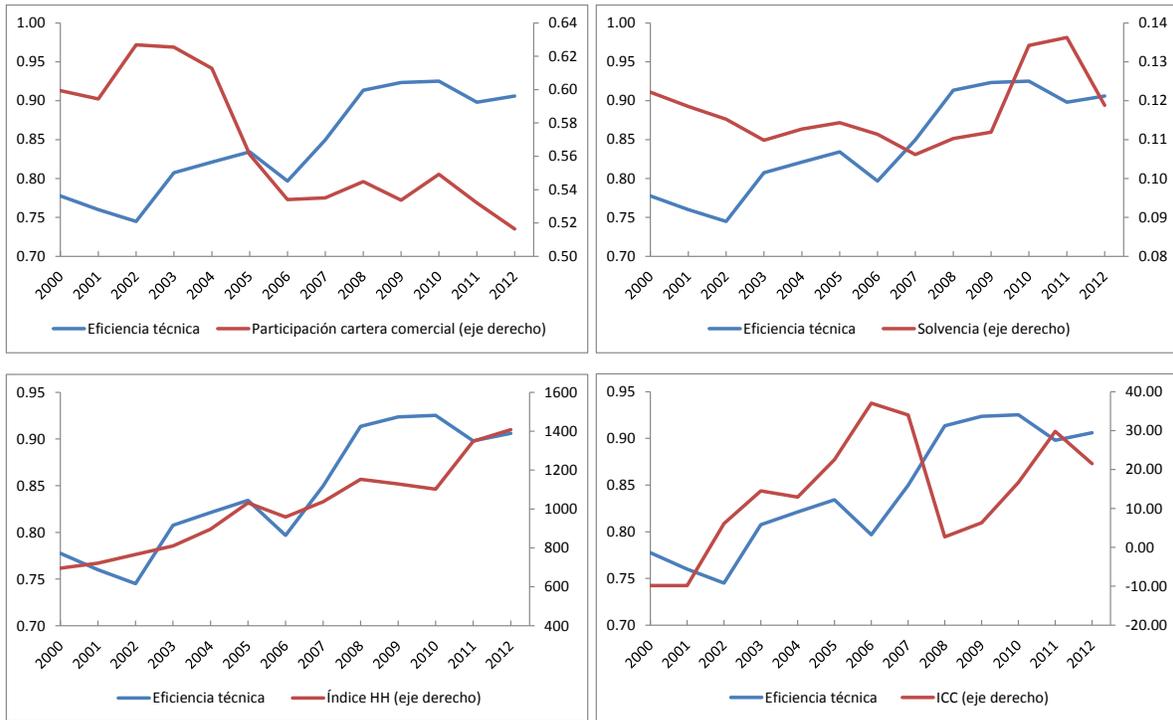
³⁷ Este modelo fue propuesto por Banker, Charnes y Cooper en 1984.

$$\sum_{n=1}^N \lambda_n = 1$$

$$\lambda_n \geq 0; \quad n = 1, 2, \dots, N$$

θ_m no restringido

ANEXO 2. Evolución promedio de las variables ambientales y la eficiencia técnica



ANEXO 3. Resultados de los modelos usando la eficiencia asignativa y de costos

Variable Dependiente: *EA*

Variables Independientes:

<i>ICC</i>	0.0000696 (0.0009415)
<i>SOLV(-1)</i>	0.1574546 (0.3434982)
<i>IHH</i>	0.000053*** (0.00000979)
<i>CCIAL</i>	-0.3862221*** (0.0573622)
<i>CONST</i>	0.7055444*** (0.052198)

Efectos Fijos	NO
Número de obs	1170
P valor Hausman	0.0726
R ²	0.23675795

Nota: Errores estándar en paréntesis

*** p<0.01, ** p<0.05, * p<0.1

Variable Dependiente: *EC*

Variables Independientes:

<i>ICC</i>	0.0005294 (0.0010619)
<i>SOLV(-1)</i>	0.0561033 (0.3839317)
<i>IHH</i>	0.0000587*** (0.0000104)
<i>CCIAL</i>	-0.4158045*** (0.0663216)
<i>CONST</i>	0.6225923*** (0.0600793)

Efectos Fijos	NO
Número de obs	1170
P valor Hausman	0.2815
R ²	0.2123067

Nota: Errores estándar en paréntesis

*** p<0.01, ** p<0.05, * p<0.1