Documento No. 29

Economías de Escala y Concentración Bancaria
El Caso de México

Por

Héctor E. González Méndez

Las ideas contenidas en el presente ensayo son responsabilidad exclusiva de los autores y no reflejan la posición del Banco de México.

ECONOMÍAS DE ESCALA Y CONCENTRACIÓN BANCARIA

El caso de México

Por

Héctor E. González Méndez

I. Introducción

Es prerrogativa de las autoridades monetarias el permitir la entrada, fusión o expansión de los intermediarios financieros. El derecho que la ley otorga a dichas autoridades para afectar la estructura del sistema financiero se basa en su conocimiento sobre lo que es más recomendable socialmente hablando. Consecuentemente, todo aquello que contribuya a una mejor definición de los agentes financieros más eficientes habrá de permitir un mejor diseño de la política monetaria.

El propósito de este trabajo es analizar el comportamiento de la función de costos de la industria bancaria en México; resulta difícil de concebir una política efectiva de apoyos y subsidios a instituciones bancarias sin un entendimiento previo del comportamiento de las funciones de costos que caracterizan al sector. Por ejemplo, si la industria está sujeta a economías de escala, las instituciones más grandes serán las más eficientes y las que proporcionen los servicios bancarios a un costo más bajo. Por supuesto, este tipo de estudios sobre eficiencia económica son solamente un lado de la moneda, otras consideraciones tales como son los efectos de un sistema integrado por unos cuantos grandes bancos sobre la competencia; las innovaciones tecnológicas y la efectividad de la política monetaria son aspectos a tomar también muy en cuenta.

En general se reconoce que la forma de las funciones de costos y la existencia de economías de escala son consideraciones muy importantes en la teoría económica, sin embargo, su medición es difícil y los trabajos empíricos de calidad en la materia son escasos; por añadidura, la literatura específica al sector financiero es aún la más reducida en el tema y para el caso de México es prácticamente inexistente.

II. Revisión de la Bibliografía.

Los métodos empleados para medir economías de escala varían según los objetivos que persigue cada trabajo, inclusive algunos autores¹ consideran que los resultados de estos trabajos son en gran medida función de los propósitos que persiguen. Existe, sin embargo, un común denominador a la mayoría, y este es; la definición de la función de costos relevante.

George J. Benston (7) es el autor que probablemente más ha escrito sobre el tema de economías de escala en el sector financiero, sus trabajos constituyen no sólo una aportación en los resultados sino en las metodologías que ha desarrollado. Particular atención presta el autor a la definición de la unidad de cuenta, “output”, de la industria bancaria y a la clasificación de los servicios bancarios. Benston encontró que existen economías de escala en prácticamente todos los servicios bancarios; cuentas de cheques, cuentas de ahorro y operaciones de crédito.

Alhadeff (2) y Horvitz (17) analizan las economías de escala relacionando el costo total de operación por dólar en préstamo o valores con el tamaño de los bancos; previamente clasificados por diferentes tipos y montos de los depósitos; el tamaño de los bancos se mide en ese trabajo en términos de 5 a 9 clases de depósitos totales. Ambos autores encontraron que el costo como proporción del crédito y los valores declina para bancos con menos de cinco millones de dólares en depósitos totales (U.S.A. 1963), permanece prácticamente constante para bancos entre cinco y cincuenta millones de dólares y finalmente declina de nuevo para bancos mayores.

Schweiger–McGee (21) y Gramley (14) utilizan un método de regresión múltiple donde la variable dependiente fue el costo total de operación como porcentaje de los activos totales. La variable independiente más importante en ambos estudios fue los activos totales. Los trabajos reportan que el costo por dólar de activo declina considerablemente –a una tasa decreciente- hasta que los bancos alcanzan aproximadamente un total de activos de 50 millones de dólares (U.S.A. 1961), después de este punto el costo continua declinando, pero lentamente.

Briham y Grebler analizan cinco años (1962 – 1966) de información sobre tipo de asociaciones de crédito en tres mercados de los Estados Unidos. El método empleado por estos autores es muy similar al de Gramley; la variable dependiente en sus ecuaciones es el costo de operación. Los

¹ George J. Benston (8).

resultados encontrados por estos autores es que el coeficiente de regresión que mide las economías de escala aparece consistentemente negativo y estadísticamente significativo, confirmando la existencia de economías de escala.

Murray y White (19) analizan la existencia de economías de escala en las instituciones financieras de Canadá (1972 – 1975). El método empleado consistió en un ajuste de la curva de costo promedio de largo plazo por medio de una ecuación de regresión. Los autores incluyen un modelo de minimización de costo basado en los trabajos de Benston y concluyen que el sistema canadiense presenta rangos tanto de costo decreciente como creciente; economías y deseconomías de escala.

Daniel, Longbrake y Murphy (12) analizan el efecto de la tecnología en las economías de escala de los bancos en el manejo de cuentas de cheques. Aunque su enfoque se aparte del estudio propiamente de las economías de escala, la metodología que emplean los autores al incluir funciones de costos tipo Cobb-Douglas para tres tipos de tecnología bancaria con especificación amplia de variables independientes es sumamente interesante. Sus resultados probaron que los bancos que carecen de computadoras experimentan deseconomías de escala, en tanto que los que han incorporado a su operación sistemas computacionales disfrutan de costos unitarios no financieros decrecientes con la escala.

Arnaudo (3) se ocupa de estudiar las economías de operación y las economías de escala en el Sistema Bancario Comercial Argentino. Hasta donde llega nuestra información es el único trabajo en el tema para América Latina que merece atención. El autor hace una distinción importante entre lo que se puede considerar uso en la capacidad instalada y cambios en la misma. El trabajo presenta algunas deficiencias estadísticas en la especificación de la función de costos, debido básicamente a la falta de datos contenidos en cada una de las cinco categorías de bancos considerados. Arnaudo concluye que sólo los bancos muy grandes, y que tienen un rango de variación de los costos unitarios por debajo de los del resto, disfrutan de economías de escala en Argentina, señala también que éstas prácticamente no existen en los bancos oficiales provinciales y bancos regionales. El trabajo destaca por su mayor atención en las economías de operación.

El objeto de este trabajo consiste en identificar tan cerca como sea posible la relación que existe entre los costos de operación bancarios y el tamaño o escala de operación de los intermediarios financieros en México. El enfoque que se utiliza está encaminado a dar mayor relevancia a los efectos

que las economías de escala pueden tener sobre la concentración bancaria. En base a lo anterior, conviene clarar que los resultados que se pretenden alcanzar son limitados en relación a la amplitud del tema, por lo que deberá considerarse a este trabajo solamente como una primera y parcial incursión para el caso de México.

III. Planteamiento General del Problema.

El análisis de las economías de escala en la producción se basa en la definición del concepto de “planta” o unidad básica de producción. En la mayoría de las industrias el concepto de planta está ligado al espacio físico que ocupa la unidad de producción o complejo industrial, así, la escala de producción se identifica fácilmente con el tamaño y capacidad de producción de estas unidades. En el mercado financiero, sin embargo, la escala está íntimamente relacionada con la cobertura geográfica de los servicios que se ofrecen, por lo que el concepto de planta no se puede reducir al nivel de establecimiento bancario (ventanilla al público). En el estudio que nos ocupa habremos de identificar a las instituciones bancarias como las unidades básicas de producción, dependiendo su tamaño precisamente del número de oficinas, sucursales y agencias de que dispone.

Dada la naturaleza de la legislación americana, la discusión sobre economías de escala en ese país se ha centrado en gran medida entre la conveniencia de una industria bancaria organizada a base de sucursales o a base de unidades. Consecuentemente un buen número de trabajos (Benston, Longbrake y Haslem) se concentran en el comportamiento de las funciones de costos de operación entre sucursales y unidades bancarias independientes. Este enfoque, sin embargo, no invalida el que se identifique a la “Planta” con la “Firma”; Aubrey Silberstos (22) considera que ambos conceptos Planta y Firma bien pueden considerarse como sinónimos cuando lo que se busca es la identificación de la función de costos de toda la industria bancaria.

Por otra parte y como ya antes señalábamos, la mayoría de los estudios sobre la existencia de economía o deseconomías de escala de la producción en una industria determinada, se centran generalmente en el comportamiento de las funciones de costos de las empresas que la integran. Se habla de economías de escala cuando los costos unitarios se abaten a medida que se aumenta la escala de producción, por el contrario, cuando costos unitarios y escala se mueven en la misma dirección, se dice que la industria presenta deseconomías.

La representación gráfica de las curvas de costos unitarios permite apreciar fácilmente lo que se entiende por economías o deseconomías de escala. Se dice que hay economías cuando el conjunto de curvas de costos unitarios aparece como una sucesión de imágenes que se desplazan hacia la derecha y hacia abajo en un sistema de ejes coordenadas; tramo $\overline{a_1a_2}$ en la figura 1. del mismo modo, habrá deseconomías de escala cuando dichas imágenes aparecen desplazándose hacia la derecha y hacia arriba; tramos $\overline{a_2a_3}$ de la misma figura.

Figura 1. Curvas de costos unitarios de corto y largo plazo.

Los cambios en el nivel de costos unitarios como resultado de incrementos en la capacidad instalada de producción es el aspecto que nos ocupa, por ejemplo, los puntos C, D y F en la figura 1 representan costos unitarios mínimos a tres diferentes escalas de operación, entre estos puntos existen economías de escala.

Los cambios en el nivel de costos unitarios como resultado de incrementos en la capacidad utilizada de las unidades de producción se reconocen como resultado de economías o deseconomías de operación, por ejemplo estos cambios quedarán representados por los puntos A, B y C en la figura 1. De este tipo de fluctuaciones de corto plazo no haremos mención en este trabajo.

Los cambios en el nivel de costos unitarios como resultado de variaciones simultáneas en la escala y en el nivel de operación hacen sumamente difícil un análisis sobre la forma de la función de costos. Cuando este tipo de fenómenos se presenta, generalmente los datos empíricos no permiten inferir absolutamente nada. Por ejemplo, los puntos B, E y G en la figura 1 representan un mismo nivel de costos, sin embargo, de esto no se puede obviamente inferir que la industria se caracteriza por costos constantes de la escala. La selección del año muestral es crucial a este respecto, ya que de alguna manera se supone que cada banco opera en equilibrio, de tal manera que no se presentan casos de sub o sobreutilización de la capacidad instalada; por ejemplo, el sistema financiero mexicano ha atravesado por una serie de transformaciones a partir de 1976 que nos permiten suponer que un número importante de bancos han ido alejándose de sus puntos a lo largo de la senda de expansión de la industria, lo cual hace más difícil una correcta especificación de la función de costos unitarios de operación de largo plazo, es decir, la curva envolvente que aparece en la figura 1. en el siguiente apartado se elabora sobre los criterios de selección de los años muestrales y manejo de los datos.

IV. Selección del Periodo de Análisis y Características de los Datos.

A partir de 1973 se inicia en México un periodo de desintermediación financiera, lo cual fue afectando paulatinamente la estructura de operación de los bancos, obligándolos a alejarse en mayor o menor medida de sus puntos óptimos de operación dentro de sus curvas de costos de corto plazo. Con la devaluación del peso en 1976, la inmensa mayoría de las instituciones financieras sufrieron fuertes desequilibrios, al grado que el Banco de México se vio en la necesidad de acudir en apoyo de un gran número de bancos privados y mixtos. También con la creación de la banca múltiple y la expansión de la economía, posterior a 1976, un buen número de bancos se han fusionado dando origen a instituciones financieras que aún en 1980 operan debajo de su capacidad instalada. Así también, las limitaciones al crecimiento de las instituciones más grandes y la expansión de la economía en general, han ocasionado que estos bancos se vean en la necesidad de operar muy por encima de lo que podría definirse como su punto de costo mínimo.

Ante este panorama podemos observar que el sistema bancario ha experimentado un número tal de transformaciones en los últimos cinco años que como consecuencia sitúa a un gran número de instituciones bancarias, si no es que a la mayoría, fuera de sus puntos mínimos en las curvas totales de costo unitario.

El análisis de economías de escala depende básicamente del supuesto equilibrio de operación de las instituciones sobre los cuales se traza el comportamiento de la curva de costos unitarios de largo plazo. Un análisis de economías de escala a través de cifras recientes (1979 – 1980) podría conducir a la comparación de bancos cuyos niveles de operación estarían dados por ejemplo por los puntos A, D, E y G en la figura 1, lo cual nos llevaría a conclusiones totalmente alejadas de la realidad.

Por supuesto que si existiera la posibilidad de identificar en qué punto de su curva de corto plazo se encuentra cada banco, el problema de selección del año sería trivial, sin embargo, dado que la información no permitió hacer tal inferencia, se optó por utilizar dos años donde las distorsiones que se han mencionado son probablemente de menor grado. Al mismo tiempo, se seleccionaron dos años lo suficientemente separados, de tal manera que permitan contemplar cambios en la estructura de costos.

Se seleccionó 1974 porque; 1) es el primer año con que se cuenta con información detallada de cada banco y 2) es probablemente el último año durante el cual los bancos operaron alrededor de sus puntos de equilibrio.

Se seleccionó 1978 porque; 1) la muestra de bancos –múltiples y especializados- es aún amplia; 2) los efectos de la devaluación de 1976, en sus aspectos más importantes, habían sido superados y 3) los cambios en los niveles de las tasas de interés habían empezado a dejarse sentir.

Con el fin de alcanzar una mayor aproximación de los costos mínimos a cada escala de operación de los bancos, los resultados se promediaron entre bancos respondiendo a las escalas de operación definidas en forma natural dentro del sistema, esta clasificación arrojó diez grupos; el primero lo integran las dos instituciones bancarias más grandes en ambos años, mientras que el décimo se compone de doce instituciones en 1974 y once en 1978.

V. Costos Unitarios y su Clasificación.

La información de costos disponible nos permite distinguir entre los pagos por intereses que realizan las instituciones financieras a sus depositantes y los gastos en factor trabajo y administración que resultan de su actividad financiera; los primeros representan el costo financiero, el cual –dado que está regulado por las autoridades monetarias y cuya aplicación es uniforme para todas las instituciones bancarias- no debe estar sujeto a variación dependiendo de la escala de operación de los bancos. Los conceptos de costo donde en forma más adecuada se puede fincar un análisis de economías de escala

son aquellos que habremos de denominar no financieros y que se componen de las erogaciones al factor trabajo y los gastos administrativos.

Dentro del costo no financiero existe una serie de conceptos de gastos a los cuales se les puede mejor identificar con la política de penetración de mercado de los bancos. Es decir, nos referimos a aquellos gastos que la institución no necesariamente desea minimizar y sobre los cuales es menos clara la existencia de economías de escala; por ejemplo, publicidad y propaganda, comisiones a promotores de captación, etc. Estos renglones de gasto se trasladaron del costo no financiero al costo financiero, de tal manera que el primero se aproxime en lo más posible al concepto de Costo de Operación que aparece en la mayoría de los trabajos sobre el tema.

Los datos anuales de costos se dividieron entre el promedio mensual de los saldos de captación correspondientes. En los cuadros 1 y 2 aparecen las tasas implícitas de los componentes del costo de los bancos agrupados por tamaño. La simple observación de los datos permite inferir que existe alguna relación estable en el tamaño del banco y su nivel de costos de operación (CNF).

VI. Definición de Variables.

Muy probablemente las economías de escala pueden ser medidas más fácilmente en las industrias donde el producto es sólo uno y susceptible de ser medido en unidades físicas. El caso de la industria bancaria dificulta considerablemente dicha medición, ya que ofrece al público una gran variedad de servicios. La comparación de costos y tamaño bancario debe realizarse con gran cuidado, puesto que existen otros elementos que afectan el nivel de los primeros sin estar relacionados necesariamente con la escala de operación; tal es el caso del costo administrativo entre bancos de depósito y financieras.

Cuadro 1. Indicadores de costo de las instituciones privadas de crédito en México (1974).

Tamaño de la institución bancaria			Costo Financiero (CF)	<u>Costo no Financiero (CNF)</u>		
(Millones de Pesos)			r _p	Total	r _w	r _{ad}
50,000	ó	más	6.371	4.940	2.660	2.280
13,000	a	25,000	6.898	5.070	3.360	1.710
3,300	a	12,500	8.915	3.660	2.040	1.620
1,900	a	3,200	7.381	4.760	2.090	2.670
1,200	a	1,800	5.032	4.940	2.620	2.320
530	a	1,100	4.631	4.610	2.640	1.970
230	a	525	4.724	6.560	3.670	2.890
165	a	229	5.760	6.720	3.780	2.940
120	a	160	4.915	6.690	3.670	3.020
70	a	118	4.711	8.520	4.500	4.020

Cuadro 2. Indicadores de costo de las instituciones privadas de crédito en México (1978).

Tamaño de la institución bancaria			Costo Financiero (CF)	Costo no Financiero (CNF)		
(Millones de Pesos)				Total	r _w	r _{ad}
140,000	ó	más	8.821	5.446	3.143	2.303
40,000	a	60,000	9.722	5.376	3.313	2.063
12,000	a	35,000	8.054	3.333	1.909	1.424
60,000	a	7,500	7.902	4.579	3.010	1.569
3,100	a	5,250	8.626	3.583	1.915	1.668
1,100	a	3,000	5.270	7.660	4.681	2.979
650	a	1,000	6.086	7.718	4.795	2.923
450	a	600	4.334	10.494	5.970	4.525
280	a	440	4.793	8.605	5.159	3.446
170	a	225	7.315	9.243	5.395	3.848

Este problema de la medición del producto bancario es una dificultad a la que se enfrenta cualquiera que intenta una medición estadística de las funciones de costo en el mercado financiero, una simple visión de la bibliografía permite apreciar que existe una gran variedad de criterios para especificar “adecuadamente” a la variable de escala. Dada la heterogeneidad del producto bancario, la relación entre costo y escala carecería de sentido si no contemplara simultáneamente a otros factores que afectan igualmente el nivel de costos, a la sazón los más importantes son:

1) El volumen de Recursos con que opera la institución bancaria. Independientemente del origen de los depósitos, ésta es una variable que mide la escala de operación de la institución financiera.

2) El número de cuenta-habientes que maneja cada institución y / o el saldo promedio de los depósitos. Parece claro que los bancos pequeños encuentran mayores dificultades para atraer clientes con depósitos cuantiosos; esto se debe, en parte a que la gama de servicios que pueden ofrecer es más restringida que los requerimientos de servicios de los cuenta-habientes importantes. A un mismo volumen de recursos, hay un costo asociado al mayor o menor número de cuentas con que se maneja dicho volumen.

3) Diferencias en el precio de los factores de producción entre bancos. Este trabajo habrá de suponer que dichos factores se contratan en mercados competitivos y que, por tanto, no se presentan diferencias en los costos variables unitarios entre instituciones.

4) Discrepancias en la estructura a plazos de los pasivos de los bancos. Este es un elemento que repercute considerablemente sobre el costo de la captación, ya que, tanto los costos financieros como los de operación de las cuentas de cheques difieren considerablemente de los correspondientes a las cuentas a plazo de dos años por ejemplo.

Las cuatro características anteriores no son obviamente todas las que se pueden mencionar y que hacen del producto bancario un bien heterogéneo difícil de medir. Por tal motivo sería prácticamente imposible definir la, o las, variables que satisfagan el punto de vista de la mayoría; George J. Benston (9) señala al respecto que:

“La definición de ‘output’ de las instituciones financieras es un problema que ha plagado la mayoría de estudios de la industria, así como estudios de las funciones de costo en general. Algunos investigadores han definido al producto como el total de ingresos brutos, otros como el número de depósitos y cuentas de crédito. Después de analizar, he llegado a la conclusión que el número de depósitos y cuentas de crédito es la medida de producto preferible. Sin embargo, la definición de producto depende en gran medida del método empleado para medir las economías de escala y los objetivos que se persiguen con la definición de la función de costos”^{*}.

^{*} George J. Benston: Economies of Scale of Financial Institutions. Journal of Money Credit and Banking. P312 May 1972.

Nuestro interés en el estudio de las economías de escala se orienta hacia la influencia que éstas ejercen en la concentración de la intermediación financiera; en base a estos objetivos habremos de definir nuestra unidad de cuenta o variable de escala de la operación bancaria.

En primer término nos interesa un concepto relativo del tamaño del “output” de cada banco con respecto al monto total producido por todo el sistema. Con esto se busca, además de identificar el tamaño económico viable de las instituciones bancarias, los cambios relativos que esta magnitud experimenta en el tiempo.

La escala relativa de operación de los bancos se definió en base a su participación tanto en el total de recursos del sistema bancario, como en el número total de cuenta-habientes. El primero está directamente relacionado con el tamaño del banco como intermediario financiero y el segundo con el tamaño promedio de las cuentas de captación, lo cual incide en forma muy importante sobre el costo. El tamaño del banco medido a través del número de cuenta-habientes afecta la relación escala-costos en forma adversa, ya que la mayor amortización de los depósitos tiende a incrementar el costo. La ponderación de ambos criterios de tamaño busca determinar un indicador más conservador del que podría ser la simple importancia relativa del banco en los recursos totales del sistema. La variable de escala (E); o medida relativa del “output”, se define como sigue:

$$E_j = \left(\frac{RT_j}{\sum_{j=1}^m RT_j} \right) \left(\frac{TCH_j}{\sum_{j=1}^m TCH_j} \right)$$

donde:

RT_j – Recursos totales del banco j .

TCH_j – Número total de cuenta-habientes del banco j .

m – Total de bancos del sistema

$$0 < E_j < 1$$

Al adoptar este criterio sobre la variable de escala se está suponiendo en forma implícita que los servicios bancarios (captación por ejemplo) son productos homogéneos, los cuales se pueden sumar indiscriminadamente. Este es probablemente uno de los aspectos que mayores problemas origina sobre la especificación adecuada de la función de costos de operación. Para subsanar parcialmente este error, habremos de suponer que los depósitos bancarios son productos heterogéneos en cuanto al plazo en que se encuentran constituidos, de esta manera se define una variable de liquidez o plazo de los pasivos, la cual se incluye en la especificación de la función de costos. Esto además de significar un avance respecto a la metodología empleada en otros trabajos, permite hacer una mejor distinción entre el comportamiento de los costos financieros, independientes de la escala de operación, y los costos de operación o no financieros, sobre los cuales se finca el análisis de economías.

La variable de liquidez o plazo promedio del total de pasivos de cada banco (P) se deriva a través de la importancia relativa de los depósitos en cada instrumento de captación respecto al total de la cartera pasiva. Bajo este criterio se definió a la variable de plazo de la siguiente manera:

$$P_j = \sum_{i=1}^n \left[\frac{IC_{ij}}{\sum_{i=1}^n IC_{ij}} \left(\frac{PC_i}{PM} \right) \right]$$

donde:

IC_i – Instrumento de captación i (monto de los depósitos)

PC_i – Plazo de captación del instrumento i

PM – Plazo máximo de captación

$$0 < P_j < 1 \quad \left\{ \begin{array}{l} 0 \text{ representa total liquidez} \\ 1 \text{ representa total iliquidez.} \end{array} \right.$$

VII. Especificación de las Funciones de Costos.

Nos interesa determinar la función que ajusta el comportamiento de los costos unitarios de operación a la escala o tamaño de los bancos; dependiendo del valor del coeficiente de regresión entre los datos de costos no financieros y la variable de escala de operación de los bancos podremos elaborar conclusiones sobre la existencia de economías o deseconomías de escala. Al hablar sobre la unidad de medida de los servicios bancarios, señalábamos que los depósitos a diferentes plazos constituyen en rigor productos financieros heterogéneos, de los cuales se derivan niveles de costos diferentes. Es decir, los costos totales unitarios que aparecen en los cuadros 1 y 2 varían entre instituciones como resultado, por una parte, de las diferentes mezclas de pasivos (grado de liquidez) y por la otra de la escala de operación bancaria. Por lo tanto, la función de costos totales a estimar deberá contener las características de la siguiente expresión:

$$CTU = CF + CNF \quad (6)$$

$$CF = f(E, P) \quad (7)$$

$$CNF = f(E, P) \quad (8)$$

Dado que las tasas de interés bancarias son crecientes con el plazo, es de esperarse que la parcial del costo financiero (CF) respecto a la variable (P) sea positivo. Por otra parte, el movimiento de las cuentas pasivas está directamente correlacionado con la liquidez; las cuentas de cheques –por su uso como medio de pago- acusan el mayor movimiento y manejo administrativo, en tanto que los depósitos a plazo significan un esfuerzo administrativo muy reducido. De lo anterior se desprende que el signo esperado de la parcial del costo no financiero (CNF) respecto a la variable (P) sea negativo. Una representación gráfica del comportamiento de los componentes del costo en relación al plazo de los pasivos aparece en la figura 2.

Figura 2. Curva de costos bancarios en relación al grado de liquidez de los pasivos.

La pendiente de la curva de costo total unitario en relación al plazo de los pasivos puede considerarse como un indicador importante de las preferencias de los bancos y del público respecto a la liquidez de los depósitos. Actualmente, es claro, que la elevación de las tasas de interés pasivas y de la inflación han elevado las preferencias de las instituciones bancarias por mantener una mayor proporción de su cartera pasiva en instrumentos monetarios. La función de costos en relación al plazo de los pasivos, es un instrumento útil para el diseño de política monetaria ya que permite determinar con mayor precisión las preferencias del mercado por la liquidez de los instrumentos de depósito. Este es un aspecto que, aunque muy importante, escapa a los objetivos de esta nota.

Ahora bien, el comportamiento de los costos unitarios de operación (CNF) en relación a la variable de escala (E) constituye nuestro indicador de economías o deseconomías de escala, de la parcial entre estas dos variables en la ecuación (7) se desprende que habrá economías de escala si $\delta CNF / \delta E < 0$. Suponiendo una curva de costos no financieros en forma de “U” respecto a la escala y monótonicamente decreciente respecto al plazo, la representación gráfica de esta función podría asemejarse a la que aparece en la figura 3.

Figura 3. Curvas de costos de operación.

VIII. Estimación de las Curvas de Costos.

Las especificaciones utilizadas en los modelos de regresión fueron lineales y cuadráticas en los logaritmos de las variable exógenas*. Los coeficientes para el componente del costo no financiero fueron estimados en primer lugar por el método de mínimos cuadrados ordinarios para 1974 y 1978 independientemente. El mismo componente del costo se estimó también agrupando los datos de 1974 y 1978; para el caso se utilizó una versión en mínimos cuadrados ordinarios y otra a través del modelo de covarianza con una variable dicotómica. Por último, se calcularon los coeficientes de las variables independientes mediante un sistema de ecuaciones como el que aparece en la sexta sección de esta nota (Ecs. 6, 7 y 8).

* Los logaritmos de las variables exógenas se utilizan para linealizar su comportamiento, ya que ambas variables son exponenciales.

La presencia de constantes positivas en las estimaciones de costos evidencia la existencia del costo unitario fijo de operación. El hecho de que dicha constante no aparezca significativamente diferente de cero en las estimaciones del costo financiero está denotando que la curva correspondiente parte del origen, como era de esperarse. El signo positivo de la variable dicotómica en el modelo de covarianza está denotando que el costo fijo unitario se elevó a todos los niveles entre 1974 y 1978. El signo del coeficiente (o coeficientes) de regresión de la variable P está (n) denotando que la presencia de economías de escala en el sistema bancario mexicano es estadísticamente defendible. Los resultados se reportan en los cuadros 3 y 4.

En términos generales los ajustes de las ecuaciones son estadísticamente aceptables; las variables de liquidez y escala explican en todos los casos más del 80% del comportamiento de los costos de operación (CNF). No existen diferencias entre los coeficientes de las regresiones lineales estimadas a través de mínimos cuadrados ordinarios y ecuaciones simultáneas, lo cual era de esperarse. Las diferencias de coeficientes entre ambos métodos de estimación para las versiones cuadráticas no son en ninguno de los casos estadísticamente significativas, esto aparece en el cuadro 5.

Los resultados de las regresiones lineales permiten establecer lo siguiente:

1. El costo de operación es función inversa de la escala y estable entre ambos. Además, el coeficiente entre ambas variables es estadísticamente significativo.
2. El costo de operación decrece a medida que se incrementa el plazo de los pasivos bancarios, sin embargo el cambio porcentual entre las variables se ha hecho más acentuado en el tiempo. Esto puede interpretarse como un mayor encarecimiento de los factores de producción bancaria, los cuales repercuten en mayor medida sobre las cuentas de los pasivos monetarios. El coeficiente entre las variables es estadísticamente significativo.
3. Como ya antes mencionábamos el costo fijo financiero no aparece estadísticamente diferente de cero, es decir, el hiperplano de la función del costo financiero parte del origen y es creciente en ambas direcciones; plazo y escala. Esto es totalmente congruente con lo establecido en un principio; el costo variable financiero (r_p) se eleva a medida que se reduce la liquidez de la cartera pasiva de los bancos.

Cuadro 3 Funciones lineales del costo no financiero bancario.

	Clave	Costo	c	lnP	lnE	W ₂	S.E.	D.W.	R ²
<u>MCO</u>									
I -L	1974	CNF	8.9096 (7.94)	-1.6746 (-3.30)	-0.4784 (-4.53)		0.9158	1.955	0.818
II-L	1978	CNF	16.0804 (5.65)	-3.6832 (-3.54)	-0.4783 (-2.47)		1.1182	2.276	0.8415
III-L	1974-78	CNF	9.8418 (4.43)	-1.5575 (-2.01)	-0.4234 (-2.00)		1.284	2.357	0.657
IV-L	1974-78 (Modelo de covarianza)	CNF	10.583 (6.37)	-2.5028 (-3.61)	-0.5115 (-3.62)	2.3797 (4.25)	1.077	1.872	0.774
<u>SSE</u>									
V-L-I	1974	CF	2.7845 (1.76)	1.5984 (2.24)	0.4094 (2.75)		1.2878	1.152	0.644
V-L-II		CNF	8.9096 (7.94)	-1.6746 (-3.30)	-0.4784 (-4.53)		0.9158	1.955	0.8183
VI-L-I	1978	CF	1.2102 (0.47)	2.3126 (2.45)	0.4086 (2.33)		1.3415	2.453	0.7687
VI-L-II		CNF	16.0804 (5.65)	-3.6832 (-3.54)	-0.4783 (-2.47)		1.4792	2.276	0.8415

Cuadro 4 Funciones cuadráticas del costo no financiero bancario.

	Clave	Costo	c	lnP	lnE	(lnE) ²	W ₂	S.E.	D.W.	R ^{2/}
<u>MCO</u> ^{1/}										
I-C	1974	CNF	5.5058 (4.08)	-0.4302 (-0.82)	-0.3385 (-6.40)	0.1673 (3.03)		0.424	3.094	0.949
II-C	1978	CNF	15.4819 (7.45)	-3.659 (-5.10)	-0.3312 (-2.64)	0.1284 (2.35)		0.7912	2.260	0.942
III-C	1974-78	CNF	8.8516 (4.41)	-1.5099 (-2.13)	-0.3129 (-1.60)	0.1965 (2.19)		1.124	2.397	0.753
IV-C	1974-78 ^{2/} (Modelo de covarianza)	CNF	9.3788 (5.82)	-2.1130 (-3.30)	-0.4351 (-3.70)	0.1218 (2.07)	1.9647 (4.17)	0.9111	1.942	0.849
<u>SSE</u> ^{3/}										
V-C-I	1974	CF	2.7845 (1.74)	1.5984 (2.21)	0.4094 (3.72)			1.487	1.15	0.644
V-C-II		CNF	6.0194 (3.62)	-0.6414 (-0.99)	-0.4016 (-4.28)	0.1475 (2.08)		0.8098	2.73	0.893
VI-C-I	1978	CF	1.2102 (0.45)	2.3126 (2.37)	0.4086 (2.25)			1.549	2.45	0.769
VI-C-II		CNF	13.1509 (4.97)	-2.9293 (-3.15)	-0.4233 (-2.53)	0.1564 (2.91)		1.398	2.86	0.894

^{1/} MCO – Mínimos cuadrados ordinarios.^{2/} Conjugando series de tiempo y corte transversal.^{3/} SSE – Sistema simultáneo de ecuaciones.

Cuadro 5 Prueba t sobre la igualdad de los coeficientes de regresión de las versiones cuadráticas ^{*/}.

		^t CAL				^t TAB.
		c	lnP	lnE	(lnE) ²	
I-C	Vs. V-C-II	-0.380 ^A	0.404 ^A	1.194 ^A	0.359 ^A	± 2.447
		0.309 ^A	-0.327 ^A	-0.673 ^A	-0.257 ^A	
II-C	Vs. VI-C-II	1.122 ^A	-1.018 ^A	.0734 ^A	-1.243 ^A	± 2.447
		-0.882 ^A	0.785 ^A	-0.551 ^A	1.007 ^A	

^{*/} Ho : $\beta_0 = \beta_1$

Ha : $\beta_0 \neq \beta_1$

4. El costo financiero se eleva con la escala de operación de los bancos, este es un echo al que no se le encontró una razón convincente. En principio, el valor esperado del coeficiente entre ambas variables debiera ser cero, ya que las tasas de interés pasivas se encuentran oficialmente reguladas y no existe trato preferencial a instituciones al respecto. Al parecer el valor positivo y significativo del coeficiente se está debiendo a que en este renglón de costo se incluyeron gastos relacionados con la política de captación (Ejm. Publicidad y propaganda). De ser cierta esta hipótesis, el signo del coeficiente está indicando que las instituciones bancarias más grandes encarecen el costo social de la intermediación financiera a través de una desviación de sus objetivos de maximización de ganancias (se busca incrementar o mantener su participación en el mercado, por ejemplo). De cualquier manera, este es un punto de los resultados que deberá investigarse con mayor detenimiento.

Cuadro 6. Pruebas F sobre la estructura de las ecuaciones.

	^F CAL:	^F TAB:
<u>Del modelo de covarianza</u>		
1974 vs. 1978 (cuadráticas)	4.7015 ^R	1.4591
1974 vs. 1978 (lineales)	3.8353 ^R	1.4591
<u>Tipo Chow</u>		
I-L vs. I-C (1974)	2.0576 ^A	2.0921
II-L vs. II-C (1978)	1.4371 ^A	2.0921

Por su parte, los resultados del modelo de covarianza confirman que existe un cambio en la estructura de la función de costos de operación entre 1974 y 1978; las pruebas F que aparecen en la

parte superior del cuadro 6 señalan que se rechaza la hipótesis de igualdad entre ambas regresiones (1974 vs. 1978).

Al realizar pruebas t sobre los coeficientes de las regresiones se encontró que el cambio estructural entre ambos años proviene básicamente del costo fijo (ver cuadro 7). Al comparar los coeficientes relacionados al costo variable, tanto en las ecuaciones lineales como en las cuadráticas, aparece que estos no resultan estadísticamente diferentes entre ambos años, mientras que la hipótesis de igualdad del término constante entre 1974 y 1978 se rechaza en todos los casos.

Cuadro 7. Pruebas t sobre la igualdad de los coeficientes de regresión ^{*/}.

			^t CAL				^t TAB.
			c	lnP	lnE	(lnE) ²	
I-L	vs.	II-L	-6.394 ^R	3.958 ^R	0.0095 ^A	-	± 2.365
II-L	vs.	I-L	2.522 ^R	-1.933 ^A	0.0052 ^A	-	± 2.365
I-C	vs.	II-C	-7.398 ^R	6.170 ^R	-0.1380 ^A	0.704 ^A	± 2.447
II-C	vs.	I-C	4.802 ^R	-4.504 ^R	0.5820 ^A	-0.711 ^A	± 2.447

Las conclusiones que se derivan de los coeficientes de las ecuaciones cuadráticas son prácticamente los mismos que se señalaron en su caso con las ecuaciones lineales, a excepción de la relación entre el costo de operación y la variable de escala. Las ecuaciones cuadráticas definen una función de costos en forma de “U”, es decir, costos de operación decrecientes y crecientes de la escala, lo cual conduce a un tamaño óptimo. En términos generales los ajustes de las funciones cuadráticas del costo de operación son superiores a los lineales; los coeficientes de determinación son en todos los casos más elevados y los errores estándar de los coeficientes de regresión son inferiores. Sin embargo, como se puede apreciar en la parte inferior del cuadro 6, las diferencias en la suma de los residuales entre los ajustes lineales y cuadráticos no son estadísticamente significativas. Esto es importante en relación a los resultados, ya que nos está indicando que, aunque el mejor ajuste define un tamaño óptimo de escala de operación bancaria, la insuficiencia de datos no permite concluir que esto es estadísticamente cierto.

En términos de la curva de costo de largo plazo dependiente de la escala, lo anterior significa que la forma de la misma no varió considerablemente entre 1974 y 1978, pero que, sin embargo, si experimentó un desplazamiento importante hacia la derecha y hacia arriba.

IX. Economías de Escala – Resultados.

Los resultados obtenidos en el análisis estadístico indican que; 1) existen economías de escala, por lo menos hasta que los bancos alcanzan un tamaño (óptimo) de alrededor del 3.0% de los recursos totales del sistema, 2) que este tamaño óptimo se elevó entre 1974 y 1978, es decir, que la curva de costos de operación de largo plazo se desplazó hacia la derecha (y hacia arriba).

En 1974 el tamaño óptimo que resultó fue de 6,206.86 millones de pesos (2.75% de los recursos) y en 1978 de 20,056.78 millones de pesos (3.63%), es decir, a pesos constantes de 1974 el tamaño de costo mínimo de los establecimientos bancarios aumentó de 6.2 a 9.2 millones de pesos (48.8%). Los requerimientos de mayor escala entre ambos años se debieron fundamentalmente al incremento en el costo fijo, el cual se elevó durante el período.

La estructura a plazo de los pasivos afectó en forma importante a las funciones de costos; tanto financieros como no financieros. La importancia relativa del índice de plazos de los pasivos en la función del costo financiero (CF) pasó de 4.94% de la captación en 1974 a 10.10% de la misma en 1978. Es muy probable que la mayor carga financiera sumada al incremento en el CNF a todos los niveles haya afectado en mayor grado a los bancos pequeños –por lo que a escala se refiere- y aquellos cuya cartera pasiva se encontraba a un plazo mayor. Es comprensible que junto a estos movimientos en los requerimientos de escala y liquidez, el sistema haya experimentado una mayor presión hacia la concentración de los recursos en un número más reducido de bancos y cuyas preferencias por los pasivos monetarios sea mayor.

El ahorro en gastos no financieros imputable al cambio tecnológico es prácticamente insignificante durante el cuatrienio en consideración; a una misma escala de operación bancaria, el ahorro en dichos costos no financieros, dependientes del plazo, se abatieron en sólo 0.5% respecto a la captación de los bancos. Esta reducción, como se señalaba antes, contrasta con un incremento de más de cinco puntos porcentuales, respecto a la captación, en el costo financiero dependiente del plazo.

Yendo hacia la conclusión del tema central de este trabajo, podemos establecer que los resultados empíricos señalan que existen economías de escala en la operación bancaria. Por otra parte, los cambios recientes (por ejemplo en las tasas de interés e inflación) han ocasionado un desplazamiento del tamaño óptimo hacia una mayor participación dentro de los recursos totales del sistema. La elevación del tamaño económico viable de los bancos viene a favorecer la concentración

bancaria ya que coloca a los bancos pequeños en una posición de rentabilidad más precaria, obviamente esto no significa que existan elementos para aseverar que actualmente se esté observado solamente un efecto causal entre ambos fenómenos.

X. Economías de Escala y Concentración.

La presencia de economías de escala que se deriva del análisis estadístico anterior carecería de relevancia si no resultara congruente con el comportamiento del sistema bancario a través del tiempo. Ante la ausencia de una regulación específica que impida el crecimiento de los bancos después de cierto límite y ante la existencia de costos decrecientes de la escala, es de esperarse que se presente una mayor concentración de los recursos en instituciones de “tamaño óptimo”. Sobre este particular habremos de enfocar nuestra atención como último punto de este trabajo.

En términos absolutos, la concentración bancaria en México presenta uno de los índices más elevados; las dos instituciones bancarias más grandes poseen aproximadamente el 46.6% del total de recursos del sistema. En 1978 de un total de 56 instituciones privadas de crédito, el 28.5% de ellas (16 en números absolutos) captaron el 89.9% de los recursos del público ahorrador.

Por otra parte, las condiciones competitivas del mercado financiero mexicano se han venido reduciendo a lo largo de los últimos treinta años; el control del 75% de los recursos bancarios ha pasado de 42 instituciones en 1950 a solamente 6 de ellas en 1979, al mismo tiempo el total de bancos privados se redujo de 248 a menos de la mitad. Como se puede apreciar en el cuadro 8, la baja más sustancial que se presentó en el número de instituciones corresponde al período 1970-1978. Al respecto cabe hacer la salvedad, sin embargo, que el número de bancos se redujo a la mitad entre 1976 y 1978 con la integración de la banca múltiple; en algunos casos esta nueva fórmula institucional dio marco legal a fusiones que en la práctica se habían dado con anterioridad, por lo que la concentración acelerada del período sobre-estima en alguna medida lo que sucedió en la realidad.

Cuadro 8. Número de instituciones que integran el Sistema Bancario Mexicano

Por ciento del total de recursos *	Número de instituciones bancarias					
	1950	1960	1970	1975	1978	1979
20%	1	1	1	1	1	1
40%	4	3	2	2	2	2
60%	14	7	5	4	4	4
75%	42	26	18	10	9	6
85%	-	-	-	21	14	12
Total	248	244	240	139	106	100

* o proporción mayor de estos.

La posición de liderazgo que mantenía una sola institución bancaria ante una atomización de las restantes en 1950 (estas últimas controlaban en su conjunto poco más del 80% de los recursos financieros totales), ha sido sustituida por una oligopolización del mercado, donde poco más de 15 instituciones controlan aproximadamente el 90% de los activos bancarios. La concentración bancaria, medida a través del coeficiente de GINI, ha sido particularmente acelerada durante la década de los setentas; dicho indicador pasó de 0.675 en 1950 a 0.764 en 1978 (1.0 representa concentración absoluta). En el cuadro 9 se reporta este coeficiente t su tasa de cambio.

Cuadro 9. Índice de concentración bancaria.

Años	Coeficiente de GINI	
	Valor	Tasa de cambio
1950	0.675	
		2.9%
1960	0.695	
		3.6%
1970	0.720	
		6.1%
1978	0.764	

XI. Conclusiones.

El sistema bancario mexicano se caracteriza por la existencia de economías de escala, aunque éstas parecen agotarse cuando los bancos alcanzan un tamaño de operación de aproximadamente 3.0% de los recursos totales del sistema.

Durante el período de estabilidad en la inflación y las tasas de interés (1958-1974), las economías de escala permitieron que un número reducido de bancos cobrara un papel preponderante dentro del sector. Sin embargo, esto no fue en detrimento de los bancos más pequeños que sostuvieron una posición competitiva.

Al iniciarse el periodo de inflación y movimientos en las tasas pasivas de interés el tamaño óptimo de operación bancaria se ha venido desplazando hacia niveles superiores de participación en los recursos totales del sistema, lo cual ha introducido una presión adicional hacia una rápida concentración de recursos. Como ya antes se apuntaba, el número de bancos se ha reducido notablemente durante los últimos ocho años.

El mercado financiero acusa una dinámica propia que en el mejor de los casos dejará a ocho bancos con el 85% del total de recursos de la banca privada y mixta. La mayor carga financiera actual, no compensada (al menos en el corto plazo) por reducciones importantes en los componentes del costo no financiero, ha conducido a una elevación de los requerimientos de escala de operación, colocando en una posición más difícil de competitividad a los bancos pequeños y/o cuyos pasivos a largo plazo son considerables. Bajo el supuesto de que el sistema termine por integrarse solamente por bancos de tamaño económico viable, el número máximo de instituciones a que se puede aspirar es de doce.

Esta situación de competitividad hacia la que se encamina el mercado financiero bancario debe contrastarse con los costos sociales de mantener un mayor número de bancos en el sistema. Existen actualmente un número determinado de instituciones pequeñas, cuyo papel a favor de la competitividad del mercado es dudoso, pero que, sin embargo, si encarecen considerablemente el costo de los recursos intermediados.

Cuadro 9. Costos de operación que se desprenden de la escala y grado de liquidez de los bancos.

E = 1%		E = .%	
P	CNF	P	CNF
0.3	10.5788	0.5	8.0240
0.4	9.9702	0.6	7.6386
0.5	9.4981	0.7	7.3125
0.6	<u>9.1124</u>	0.8	<u>7.0300</u>
Prom.	9.7899		7.5013

Así, por ejemplo, se pueden tomar a los 29 bancos más pequeños cuya participación no supera el 1.0% de los recursos totales en 1979 y estimar el ahorro de fusionarlos en un número inferior de bancos de tamaño óptimo. En el cuadro 10 aparecen los diferentes niveles de costo bajo supuestos de liquidez y tamaño de operación. Utilizando la captación a marzo de 1980 y los resultados del cuadro anterior se estima que el ahorro anual en el costo no financiero sería de aproximadamente 5,107.13 millones de pesos. Sería deseable cuantificar la pérdida de competitividad, si la hay, al eliminar 29 bancos cuyo tamaño económico no es viable y comparar dicha cifra con el ahorro en costos de sustituir a esas instituciones por 9, que desde un punto de vista meramente económico serían “eficiente”.

BIBLIOGRAFÍA

1. Aschheim, J., "Commercial Banks and Financial Intermediaries; Fallacies and Policy Implications". *Journal of Political Economy* (1959), pags. 59-71.
2. Alhadeff, D.A., "Monopoly and Competition in Banking". Berkeley: University of California Press, 1954.
3. Arnaudo, A.A., "Economías de Operación y Economías de Escala en el Sistema Bancario Comercial Argentino, 1960-72". *Ensayos Económicos* No. 2 (junio 1977). Banco Central de la República Argentina.
4. Barts, G.H., "The Benston Paper: Some Comments". *Journal of Money, Credit and Banking* (mayo 1972), pag. 419.
5. Bell, F.W. y N.B. Murphy. "Costs in Commercial Banking: A Quantitative Analysis of Bank Behavior, and its Relation to Bank Regulation", Research Report No. 41. Boston: Federal Reserve Bank of Boston. 1968.
6. Bell, F.W. y N.B. Murphy, "Economies of Scale and Division of Labor in Commercial Banking". *Southern Economic Journal*, Vol. XXXV (Oct. 1968) pags. 131-139.
7. Benston, George J., "Commercial Bank Price Discrimination Against Small Loans: An Empirical Study". *Journal of Finance* XIX (Dic. 1964), pags. 631-643.
8. Benston, George J., "Economies of Scale and Marginal Cost in Banking Operations", *National Banking Review* II, No. 4 (Jun. 1965), pags. 507-549.
9. Benston, George J., "Economies of Scale of Financial Institutions", *Journal of Money, Credit and Banking*, May 1972, pags. 312-341.
10. Benston, George J., "Branch Banking and Economies of Scale", *The Journal of Finance*, (May 1972), pags. 312-331.
11. Carson, D., "Banking Earnings and the Competition for Saving Deposits". *Journal of Political Economy* (1959), pags. 580-588.
12. Daniel, D.L., W.A. Longbrake y N.B. Murphy, "The Effect of Technology on Bank Economies of Scale for Demand Deposits". *Journal of Finance*, (1973), pags. 131-146.
13. Edwards, F.R., "concentration and Competition in Commercial Banking: A Statistical Study". Research Report to Federal Reserve Bank of Boston. No. 26 (1964).

14. Gramley, L.E., "A Study of Scale Economies in Banking". Kansas City, Mo. Federal Reserve Bank of Kansas City, 1962.
15. Grebler, L. y E.F. Brugham, "Saving and Mortgage Markets in California". Pasadena California Saving and Loan League, 1963.
16. Grennbaum, S.I., "A Study of Bank Cost", The National Banking Review, (Jun. 1976), pags. 415-434.
17. Horvitz, P.M., "Economies of Scale in Banking"; in Private Financial Institutions. Englewood Cliff, N.J. Prentice-Hall, 1963.
18. Longbrake, W.A. y J.A., Haslem, "Productive Efficiency in Commercial Banking". Journal of Money, Credit and Banking (Ago 1975), pags. 317-330.
19. Murray, J.D. y R.W. White, "Economies of Scale and Deposit-Taking Financial Institutions in Canada", Journal of Money, Credit and Banking, Vol. 12 No. 1 (Feb 1980). Ohio State University Press.
20. Powers, J.A., "Branch Versus Unit Banking: Bank Output and Cost Economies", Southern Economic Journal, (Oct 1969), pags. 153-164.
21. Schweiger, I., J.S. McGee, "Choicago Banking: The Estructure and Performance at Banks and Realted Financial Institutions in Chicago and other Areas", The Journal of Business, XXXIV, (Jul 1961), pags. 201-336.
22. Silberstos, A., "Economies of Scale in Theory Practice", St. Joh's College, Cambridge. (Speech).
23. Spellman, L., "Competition in the Financial Sector". Workshop on Capital MarketDevelopment in L.A., (Jul 1974) O.A.S.
24. Taylor, R.A., "Economies of Scale in Large Credit Unions". Applied Economies (1972). Pags. 33-40.