

El Efecto de la Estructura de Mercado en las Tasas de Interés Activas en la Industria Bancaria Colombiana. Un Análisis de Panel No Estacionario

María Camila García Jiménez

Candidata a Magístra en Economía

Pontificia Universidad Javeriana

2017

Resumen

This paper conducts an empirical analysis of the two Power Market (PM) hypothesis in Colombian banking industry. The analysis explores macropanel techniques in order to determine the effect of market structure on active interest rates for two types of credit. By using this approach and considering data properties such as unit roots and cointegration, we study whether Colombian banking market's behavior supports the statements of the Structure-Conduct-Performance paradigm or the hypothesis of Relative Market Power.

Key words: Macro panel; Unit root and cointegration; Banking system; SCP paradigm; concentration.

Resumen

Este trabajo realiza un análisis empírico de las dos hipótesis bajo la teoría del Poder de Mercado en la industria bancaria colombiana. Utilizando las técnicas de macropanel se estudia el efecto de la estructura del mercado en las tasas de interés activas para dos tipos de crédito. Al emplear este enfoque y tomando en consideración las propiedades de los datos, tales como raíces unitarias y cointegración, estudiamos si el comportamiento del mercado bancario colombiano soporta los supuestos del Paradigma Estructura-Conducta-Resultado o la hipótesis del Poder de Mercado Relativo.

Palabras Clave: Macro panel; raíz unitaria y cointegración; sistema bancario; paradigma ECR; concentración.

Clasificación JEL: K21, L51, C33.

NOTA DE ADVERTENCIA

“La Universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de tesis. Solo velará por que no se publique nada contrario al dogma y a la moral católica y por qué las tesis no contengan ataques personales contra persona alguna, antes bien se vea en ellas el anhelo de buscar la verdad y la justicia”.

Índice

1	Introducción	5
2	Revisión de la literatura	6
2.1	La relación entre la estructura del mercado y los precios en el mercado bancario	7
2.1.1	El Paradigma Estructura-Conducta-Resultado	7
2.1.2	Algunas de las críticas al Paradigma ECR	8
2.2	La concentración como medida de la estructura de mercado	9
2.2.1	Las medidas de concentración empleadas en la literatura	9
2.2.2	Principales críticas a las medidas de concentración tradicionales	10
2.3	El ranking como medida del nivel de concentración	10
3	Modelo Económico	11
3.1	Selección del modelo	11
3.2	Especificación del modelo	12
3.3	Selección de las variables	13
3.3.1	Individuos y período de tiempo	13
3.3.2	Variable dependiente: tasa de interés activa	13
3.3.3	Variables explicativas	13
4	Metodología Econométrica	14
4.1	Rationale de la selección de metodología	14
4.1.1	Metodologías empleadas en la literatura sobre el Paradigma ECR	14
4.1.2	Selección de las técnicas de macropanel como metodología econométrica	16
4.2	Estimando relaciones de largo plazo con datos panel heterogéneos	18
4.3	Problema de inconsistencia al emplear métodos tradicionales de estimación	20
4.4	Sobre las pruebas de raíz unitaria	21
4.4.1	Prueba de Levin-Lin (LL)	22
4.4.2	Prueba de Im-Pesaran-Shin (IPS)	23
4.4.3	Prueba de Levin, Lin y Chu (2002)	24
4.4.4	Prueba de Maddala y Wu	25
4.4.5	La prueba de Fisher ($p\lambda$)	26
4.4.6	Criterio de selección de una prueba sobre las demás	26
4.5	Pruebas de cointegración de Pedroni	26
4.6	Estimación de la cointegración	29
4.6.1	<i>Group Mean</i> FMOLS (Pedroni - 2000)	29
4.6.2	<i>Parametric Pooled</i> DOLS (Kao y Chang - 2000)	31

5	Resultados	31
5.1	Estimación	31
5.2	Pruebas de raíz unitaria	31
5.3	Pruebas de cointegración	32
5.4	¿Cómo interpretamos H_a en las pruebas de cointegración?	32
5.5	Estimación de la cointegración	33
5.5.1	Crédito de consumo	35
5.5.2	Tarjeta de crédito	35
6	Conclusión	35
7	Bibliografía	36
8	Apéndices	38
8.1	Apéndice A - Estadísticas descriptivas	38
8.2	Apéndice B - Selección de la muestra	41
8.3	Apéndice C - Resultados de las pruebas	43
8.3.1	Pruebas de raíz unitaria	43
8.3.2	Pruebas de cointegración	43

1 Introducción

La concentración de la industria bancaria y su influencia en los niveles de rentabilidad de la misma han sido objeto de estudio durante varios años. Por una parte, se ha intentado identificar si a mayor concentración mayor rentabilidad de las firmas, y, en caso de encontrar una relación positiva entre ambas, la siguiente pregunta es si dicha relación es el resultado de un comportamiento anticompetitivo por parte de las firmas facilitado por la estructura del mercado en el que operan. De hecho, este tema fue el origen de lo que se conoce como el Paradigma Estructura-Conducta-Resultado (ERC) bajo el cual se asume que en un mercado con un alto nivel de concentración las firmas pueden llegar a adoptar comportamientos anticompetitivos o que afecten el bienestar del consumidor, y así obtener una rentabilidad más elevada.

Ha sido tan recurrente este tema en lo que se refiere a la industria bancaria, que no tardaron en salir las teorías alternativas y opuestas al Paradigma ERC. Por una parte, otra teoría que hace parte de la hipótesis del Poder de Mercado (como el Paradigma ECR), es la teoría del Poder de Mercado Relativo (PMR) que asume que es el poder de mercado de los agentes (derivado de su participación de mercado) lo que les permite fijar precios desfavorables a los consumidores.¹ No obstante, esta teoría sigue asumiendo que la explicación del comportamiento de los bancos radica en la estructura misma del mercado (se toma la estructura como dada)². Otra de las teorías que ataca los postulados del Paradigma ERC es la teoría de la Eficiencia (o Eficiencia-Estructura). Bajo este postulado, es la eficiencia de los agentes lo que genera que unos obtengan una rentabilidad más alta que los demás, o que puedan fijar precios más favorables al consumidor, y no la estructura del mercado; de hecho, asumen que los niveles de concentración se derivan de la eficiencia misma siendo esta el factor que facilita que una firma aumente su participación en el mercado al ser más eficiente que sus competidores.³

Distintos autores han abordado el tema de la concentración y la rentabilidad de la industria bancaria partiendo de estas hipótesis. Se han realizado estudios a nivel de país (especialmente para el mercado estadounidense),⁴ de región (Unión Europea), y comparaciones a través de países individualmente considerados y en grupo.⁵

Un tema en común entre estos estudios ha sido el reconocimiento que las implicaciones de los resultados pueden tener para el debate de política pública. Por un lado, el Paradigma ERC supone que el desempeño de los bancos responde a conductas anticompetitivas y que puede tratarse del resultado de una situación de colusión entre las firmas. Esto implicaría un debate en torno a la necesidad de intervención por parte del Estado mediante sanciones y regulaciones de esta industria específica dirigidas a reducir los niveles de concentración y disminuir la probabilidad de un posible acuerdo entre los competidores. La teoría de la eficiencia por otra parte propone que los resultados se derivan de la eficiencia de las firmas lo que implicaría que las autoridades sentirían la obligación de promover este tipo de conducta que al final beneficia a los consumidores y al bienestar general.

En general, las autoridades estatales, en relación con la industria bancaria, tienen la obligación de garantizar la seguridad de aquellos quienes ponen sus recursos en la industria, evitar el “contagio”, y proteger el mercado de prácticas anticompetitivas (Nabieu, 2013). Así entonces, existen dos grandes grupos de medidas regulatorias para la industria bancaria: (i) regulación estructural, que trata temas del mercado y su desempeño (e.g., barreras de entrada y salida, cargos de comisiones, topes de las tasas de interés); y (ii) regulación prudencial y de supervisión, donde se incluyen temas de solvencia y liquidez de los bancos (e.g., las reglas de Basilea). De acuerdo con esto, los hallazgos de los estudios son una parte importante en determinar qué tipos de medidas son las que debe seguir la autoridad, dado que no es lo mismo una situación en la que hay un problema de liquidez en la industria y otro en el que las barreras de salida son tan altas que los bancos adoptan un comportamiento negligente en el manejo de sus recursos (Nabieu, 2013).

¹Ver como ejemplo Rhoades (1985) quien sugiere que la relación empírica positiva entre la participación de mercado y la rentabilidad no refleja diferencias en eficiencia entre los bancos sino las ventajas que se derivan de la diferenciación de productos.

²Ver como ejemplo Smirlock (1985) quien encuentra una relación positiva entre la participación del mercado y la rentabilidad; una relación insignificante entre concentración y rentabilidad; y una relación negativa entre la interacción concentración*participación de mercado con la rentabilidad.

³Ver como ejemplo Berger (1995) quien encuentra que la rentabilidad del mercado bancario en Estados Unidos está positivamente relacionada al poder de mercado y la eficiencia X durante los años 80s.

⁴Un ejemplo es Berger et. al. (1987) y Neely y Wheelock (1997)

⁵Entre este grupo encontramos por ejemplo los análisis de Haslem (1968), Short (1979), Bourke (1989), Molyneux y Thorton (1992), Dermiguc-Kunt y Huizinga (2000), y Bikker y Hu (2002)

Ahora bien, el presente Trabajo de Grado propone explorar este tema tomando como referencia la industria bancaria colombiana. Dado que el mercado colombiano es altamente concentrado, en congruencia con lo que ocurre en la región, puede darse lugar al debate sobre la relación entre dicha estructura de mercado y el desempeño o la fijación de precios en la industria. Se parte entonces de los postulados de las dos hipótesis del Poder de Mercado (PM) - el Paradigma ERC y la teoría PMR- para analizar su efecto sobre las tasas de interés activas de dos tipos de créditos ofrecidos al público: crédito de consumo y tarjeta de crédito. Este es un tema relevante porque no se puede caer en la tentación de concluir que un alto nivel de concentración explica una relación de causalidad por sí mismo, sino que solo se limita a ser un punto de partida para posibles investigaciones (Martínez et. al., 2016). En otras palabras, no se puede condenar un alto nivel de concentración en un mercado por ese hecho solamente considerado.

El principal aporte del Trabajo de Grado es presentar el análisis de la teoría de Poder de Mercado (PM) desde la metodología de macropanel propuesta por Pedroni. Aplicar esta metodología nos permite abordar algunas de las críticas que se han señalado en la literatura a los trabajos que desarrollan estos temas. En varios de los trabajos se utilizan variables que están a nivel de banco y a nivel de mercado; uno de los ejemplos más claros de una medida que está a nivel de mercado es el nivel de concentración de la industria (usualmente medida a través del HHI o el CR_k) y una variable a nivel de banco es la participación de mercado. Se han empleado datos del tipo de series de tiempo y de corte transversal (especialmente cuando se realizan comparaciones a nivel de país o región). Finalmente, en otros estudios que cuentan con datos tipo panel se emplean técnicas como la agrupación de los datos - pooled - y técnicas de efectos fijos.

Sin embargo, en la literatura que ha empleado datos panel no se ha hecho énfasis en el concepto de estacionariedad y por lo tanto no se realizan las pruebas para determinar si existen o no raíces unitarias. De acuerdo con Pesaran y Smith (1995) las técnicas clásicas de paneles dinámicos se limitan a resolver el problema de la heterogeneidad extrayendo la dinámica y tratandola como homogénea; y otro de los problemas, que se desprende precisamente de la necesidad de que las dinámicas sean homogéneas, es que estas técnicas no son adecuadas para tratar la presencia de raíces unitarias.

Por lo anterior, analizamos la relación estructura-precio en la industria bancaria colombiana a partir de las técnicas de macropanel, empleando las pruebas de raíz unitaria y, en caso que haya evidencia de que las series son estacionarias, se aplica la prueba de cointegración de Pedroni. En último lugar, se estima el modelo para poder analizar los coeficientes obtenidos a la luz de las teorías de PM y determinar qué tipo de relación existe entre la variable de interés y la estructura de mercado.

Adicionalmente, siguiendo los planteamientos de Marlow y Right (1987) se introduce en el modelo como variable de concentración el ranking las firmas, medido por la posición de cada banco en el ranking y el cambio en el ranking de un período a otro. Esta especificación es consistente con la utilización de la técnica de macropanel dado que es una medida de concentración a nivel de individuo y no a nivel de mercado.⁶

2 Revisión de la literatura

Existe una amplia literatura sobre la industria bancaria partiendo de diferentes interrogantes, en donde se intenta encontrar respuesta a temas como la relación entre la estructura del mercado, la dependencia financiera de una industria específica y el crecimiento económico de un país (Cetorelly y Gambera, 2001);⁷ el efecto de las autorizaciones legales para establecer sucursales bancarias y su relación con la concentración y estabilidad financiera de la industria (Carlson y Mitchener, 2006);⁸

⁶“While market structure affects market performance, empirical studies have not always found a relationship between indicators of performance (e.g., profitability) and market structure, as represented by indexes of concentration. Yet such indexes play a prominent role in antitrust regulation. Antitrust policy, then, faces two problems: doubt as to whether concentrated industries can actually perform poorly, and a lack of criteria for choosing from among alternative indexes”. (Marlow and Righth, 1987, pp. 1)

⁷Parten de la hipótesis que entre mayor sea el nivel de desarrollo de la industria bancaria mayor va a ser el nivel de crecimiento de la economía. Encuentran que la estructura de la industria puede promover el crecimiento de los sectores económicos que requieren un mayor acceso a financiamiento pero a nivel general genera un efecto depresivo

⁸Dado que la actividad bancaria está regulada en aspectos tales como la constitución de nuevos bancos, los procesos de fusión o integración, las actividades que pueden realizar, Carlson y Mitchener (2006) parten del supuesto que la existencia de sucursales bancarias fomenta la estabilización de los mercados bancarios al facilitar la diversificación de la cartera de los bancos. Su hipótesis es que en escenarios en donde

el impacto de la propiedad de los bancos (extranjera o nacional) y el desempeño financiero de cada entidad (Shehzad, Haan y Scholtens, 2012);⁹ o la relación entre la competencia de los bancos con la estabilidad financiera del sistema (Schaeck, Cichak y Wolfe, 2009).¹⁰

Gran parte de esta literatura generalmente se ha enfocado en el análisis de la relación entre la concentración de la industria y las condiciones de competencia, eficiencia y desempeño del mercado. Sin embargo una de las variables de mayor interés ha sido la rentabilidad y los precios de la industria y sus determinantes respectivos. En el grupo de factores internos encontramos aquellas variables que se derivan de las cuentas individuales de cada banco (i.e. incluidas en los balances generales y estados de pérdidas resultados). Por otra parte, en el grupo de factores externos se consideran aquellos que reflejan el ambiente económico y regulatorio de determinado mercado (i.e. variables que no están relacionadas con el manejo, dirección y administración de los bancos).

En esta línea de estudio, la literatura se ha enfocado, en términos generales, en analizar la rentabilidad de sistemas bancarios (o las tasas de interés), ya sea comparando los mercados entre varios países o analizando una industria específica. Algunos ejemplos de estudios comparativos entre países son Haslem (1968), Short (1979), Bourke (1989), y Demirguc-Kunt y Huizinga (2000). En cuanto a análisis de una industria específica, gran parte se ha dedicado a analizar el mercado de Estados Unidos como en el caso de Berger (1987) y Neely y Wheelock (1997). Algunos ejemplos de otros mercados se encuentran en Athanasoglou et al (2005) para la industria griega, o Murthy y Deb (2008) para India.

2.1 La relación entre la estructura del mercado y los precios en el mercado bancario

2.1.1 El Paradigma Estructura-Conducta-Resultado

De acuerdo con lo anterior, vemos que la literatura sobre la industria bancaria se ha enfocado en determinar la mejor forma de medir la competencia en el mercado. Se observa un especial interés en estudiar la relación entre la rentabilidad o los precios de la industria y la estructura misma del mercado. Así pues, existen dos ramas que miden la competencia desde distintos puntos de partida: 1) el *enfoque estructural*, bajo el cual se encuentra el Paradigma ECR y su hipótesis opuesta, Eficiencia-Estructura;¹¹ y 2) los modelos *no estructurales* que analizan la competencia, el poder de mercado, y la conducta de los bancos en ausencia de medidas estructurales.¹² En el presente Trabajo de Grado nos vamos a centrar en el enfoque estructural.

La literatura que ha estado centrada entonces en lo que se conoce como las teorías del Poder de Mercado,¹³ y se enfoca especialmente en el paradigma ECR.¹⁴ Bajo el paradigma ECR se asume que la fijación de precios desfavorables para el consumidor en mercados más concentrados es el resultado de una imperfección en ese mercado (Berger, 1995). En otras palabras, se considera que la estructura del mercado favorece un comportamiento por parte de sus participantes que va en contra del bienestar y que el resultado que se observa no se diferencia de aquel que se obtendría en un escenario donde opera un cartel (Martínez et al, 2016).

se permite establecer sucursales, los bancos marginalmente rentables se van a ver forzados a salir del sistema bancario, ya sea a través de una fusión con otra entidad o por liquidación voluntaria. Al final encuentran que en efecto la autorización para establecer sucursales incrementa los niveles de competencia en el mercado y asimismo aumenta el número de salida de firmas en el mercado.

⁹Shehzad, Haan y Scholtens (2012) parten de la idea de que la concentración en la propiedad de los bancos contribuye al desempeño de la institución frente a la alternativa que este factor no influye en el nivel de riesgo que se adopta en las operaciones.

¹⁰Schaeck, Cichak y Wolfe (2009), haciendo uso del estadístico H del modelo de Panzar-Rosse (1987), analizan las condiciones de competencia en un grupo de 45 países, concluyendo que en sistemas más competitivos hay una mayor estabilidad financiera. Con este resultado consideran que las variables de competencia y concentración capturan diferentes características de los sistemas bancarios y descartan la concentración como una proxy del nivel de competencia en el mercado.

¹¹Estas dos hipótesis tienen sustento en la teoría de la Organización Industrial.

¹²Algunas de las medidas no estructurales han sido desarrolladas bajo la Nueva Organización Industrial Empírica en donde se intenta observar de manera directa la conducta de las empresas. Estas metodologías se pueden clasificar en dos grupos: 1) las que se basan en la teoría del oligopolio (índice de Lerner, modelos de variables conjeturales, y el modelo de Panzar y Rosse); y 2) las que se basan en la dinámica del mercado indicador de Boone). (Martínez et al, 2016).

¹³Market Power Theories, en inglés.

¹⁴Otra hipótesis de la Teoría de Poder de Mercado es la Hipótesis del Poder Relativo de Mercado (Relative-Market-Power Hypothesis) bajo la cual solo firmas con altas participaciones de mercados y productos bien diferenciados pueden ejercer poder de mercado en el precio de los productos y obtener una rentabilidad mayor de lo usualmente observado.

El paradigma ECR tomó una especial relevancia por las implicaciones que sus resultados podían tener en temas de políticas públicas. Por una parte, al comprobarse que el resultado de un mercado es consecuencia de una conducta anticompetitiva por parte de sus agentes podría implicar la necesidad de intervención por parte del Estado para que entrara a regular ese mercado. De la misma forma, podría llegar a influenciar el debate de política en temas de derechos de los consumidores y legislación sobre competencia (e.g., en el tema de las fusiones y adquisiciones). Estas implicaciones han sido reconocidas por distintos autores que han analizado el comportamiento de la industria bancaria bajo el paradigma ECR.¹⁵

Ahora bien, como se observa, una parte esencial del Paradigma ECR radica en la estructura del mercado. El problema entonces es cómo se puede definir, o mejor, cómo se puede medir la estructura del mercado. Mientras que la literatura ha estado siempre interesada por la relación entre rentabilidad o precios y estructura, no hay un consenso sobre cuál es la forma óptima de medir o introducir una variable que dé cuenta de la estructura de mercado. De hecho, la gran mayoría de artículos han empleado medidas de concentración como proxies de la estructura del mercado por la facilidad que implica la construcción de este tipo de medidas. Sin embargo, también ha surgido una literatura encargada de señalar las deficiencias que tienen las medidas de concentración y por qué no son las más adecuadas para representar la estructura de un determinado mercado bancario.

No obstante, sí existe un punto en común en la literatura sobre el tema de la concentración: el nivel de concentración de una industria no implica un diagnóstico inmediato sobre el nivel de competencia de la misma, ya que se ha observado que en escenarios de alta concentración puede existir un elevado nivel de competencia entre las firmas, y vice versa (Bikker y Haaf, 2002).

2.1.2 Algunas de las críticas al Paradigma ECR

Los planteamientos del Paradigma ECR han estado sujetos a una serie de críticas que se derivan de sus supuestos y la forma en que los autores han modelado sus estudios. Por una parte, tenemos la hipótesis Eficiencia-Estructura que se opone a la esencia del Paradigma ECR¹⁶ al plantear que no es la estructura del mercado la que determina los precios o rentabilidad del sistema sino la eficiencia de los bancos en este mercado. En otras palabras, la hipótesis Estructura-Eficiencia sugiere que la relación entre concentración y rentabilidad obedece a la relación positiva que existe entre tamaño y eficiencia.

Distintos autores se han encargado de analizar diferentes mercados bancarios para establecer cuál de las dos hipótesis explica mejor su desempeño.¹⁷ Sin embargo todavía no existe consenso entre los autores, dado que para diferentes mercados y diferentes variables los resultados son opuestos. Esta situación en sí misma también ha sido objeto de estudio en la literatura, donde se ha intentado determinar la razón por la que se encuentran resultados contradictorios en varios mercados.

Sin embargo, en términos generales se puede decir que la mayor crítica que plantea la hipótesis Estructura-Eficiencia es que el Paradigma ECR no toma en consideración la influencia de las variables de eficiencia. Berger (1995) señala que en la mayoría de casos la variable de interés (resultado o precios) es regresada contra participación de mercado y/o concentración con el objetivo de encontrar si los precios son menores o si la participación de mercado es mayor. El problema con estas especificaciones es que omiten las variables de eficiencia que pueden estar relacionadas tanto con los precios como con la

¹⁵“The two market power (MP) hypotheses have radically contrasting implications from the two efficient-structure (ES) hypotheses for merger and antitrust policy. To the extent that the MP hypotheses are correct, mergers may be motivated by desired to set prices that are less favorable to consumers, which would decrease total consumer plus producer surplus.” (Berger, 1995, pp. 405)

¹⁶La esencia del Paradigma ECR se podría resumir en que las características estructurales de la industria son las que determinan los aspectos de competencia en la misma (Leon, 2014).

¹⁷Uno de los primeros ejemplos lo encontramos en Smirlock (1985) quien encuentra una relación positiva entre la participación de mercado y rentabilidad, y una relación insignificante entre concentración y rentabilidad. Por otra parte, Rhoadess (1985) sugiere la existencia de una relación positiva entre la participación de mercado y rentabilidad pero nota que no explica las diferencias en los niveles de eficiencia entre los bancos sino las ventajas derivadas de la diferenciación en los productos. Berger y Hannan (1989) estudian la relación entre los precios y la concentración en el mercado bancario bajo estas dos teorías. Por una parte, establecen que la hipótesis *SCP* explica la relación por un comportamiento no competitivo de precios, mientras que la hipótesis *eficiencia-estructura* la explica en razón a la mayor eficiencia de las firmas con participación dominante en el mercado. Concluyen que esta relación responde a la hipótesis *SCP*, es decir, que la rentabilidad de la industria se debe a comportamientos anticompetitivos.

estructura. Esto implicaría entonces que la estructura del mercado no sea una variable que está dada (como lo supone el Paraadigma ECR) sino que es endógena al depender del nivel de eficiencia de las firmas.¹⁸

Ahora bien, una de las críticas de carácter conceptual que se han señalado en la literatura sobre el Paradigma ECR es su postulado que establece que a un mayor nivel de concentración mayores oportunidades de colusión tienen las firmas y por lo tanto se observan mayores precios y rentabilidad. Las teorías opuestas se enfocan en desvirtuar la relación entre conducta y estructura. Martínez et. al. (2016) señalan que una industria que este altamente concentrada puede, no obstante, funcionar eficientemente en equilibrio¹⁹ y que de acuerdo con la teoría de la contestabilidad se contempla el escenario en el que un mercado monopolístico puede reflejar un comportamiento competitivo entre los agentes, siempre que los costos de entrada y salida del mercado sean bajos. Finalmente, también es posible que se presenten casos de colusión en el mercado donde hay un número plural importante de firmas (Leon, 2014).

También se han hecho varias críticas, en sentido práctico, sobre la escogencia de las medidas de concentración como reflejo de la estructura de mercado. Usualmente la concentración ha sido la variable de mayor preferencia para ilustrar la estructura de una industria debido a la facilidad en su cálculo. Dentro del grupo de las medidas tradicionales resaltan el *HHI* y las razones de concentración o CR_k . Sin embargo, uno de los problemas de emplear las medidas de concentración es que no es claro entender cuál es la implicación de los distintos niveles de concentración. De hecho, se ha criticado que la concentración puede estar dando cuenta de características de los bancos y del mercado tales como el tamaño promedio de los bancos, las complejidades a nivel de producto y actividades, el flujo de información en el mercado, entre otros (Ergungor, 2004).²⁰

De acuerdo con este último tema, la siguiente sección explica cuáles han sido las medidas tradicionales para medir la concentración, los aspectos comunes en su cálculo y las principales críticas que se han planteado en la literatura frente a su uso. También se explica una medida alternativa de concentración que ha sido planteada por distintos autores con el objetivo de reflejar mejor el nivel de competencia entre los agentes del mercado.

2.2 La concentración como medida de la estructura de mercado

2.2.1 Las medidas de concentración empleadas en la literatura

Tal como discutimos, la concentración ha sido una de las formas predilectas para medir la estructura de mercado (Bikker y Haaf, 2002).²¹ Sin embargo, hay más variables además de la concentración que definen la estructura de un mercado. Martínez et al. (2016) definen las características que toda estructura de mercado tiene: (i) número de empresas, (ii) acciones disponibles a cada empresa, (iii) expectativas acerca de las respuestas de competidores según conjunto de acciones que se crea disponga cada empresa, (iv) expectativas sobre número de empresas y posibilidad (barreras) de entrada, (v) grado de integración vertical, (vi) nivel de concentración de la producción y de participación de las empresas en el mercado, (vii) barreras de entrada y salida existentes, (viii) nivel de diferenciación de los productos, y (ix) nivel de diversificación de la producción.

Como se observa, de todas las características anteriormente listadas, la concentración resalta como una de las variables más sencillas en cuanto a la obtención de datos y su cálculo.

¹⁸Sobre este tema, podemos recordar a Martínez et. al. “(...) d) así mismo, la hipótesis de estructura eficiente (EE), avanzada por Demsetz (1973), Peltzman (1977) y otros, afirma que la estructura de un mercado puede reflejar diferencias en eficiencia entre las empresas, en vez de ser resultado de una situación competitiva, de tal modo que las empresas más eficientes tienen a ganar participación de mercado, lo cual puede conducir a una concentración de mercado más alta. Esto quiere decir que la estructura de mercado (ilustrada por los índices de concentración) no es exógena, pudiendo más bien ser el resultado de las diferencias en la eficiencia y de las estrategias empresariales.” (Martínez et. al., 2016, Pp. 14)

¹⁹ Siguiendo lo que plantea Bertrand (1883)

²⁰ A esta línea de argumentación se suma el hecho que para poder definir la concentración de un mercado primero es necesario definir ese mercado. En un primer momento, por ejemplo, sería necesario definir el mercado geográfico que puede variar de acuerdo a las actividades de los bancos. (Leon, 2014)

²¹ Los autores también señalan que las medidas de concentración dan cuenta de los bancos que entran o salen del mercado o de las fusiones que se observan entre algunas de las firmas.

Según Bikker y Haff (2002) todas las medidas de concentración planteadas en la literatura tienen en común que incluyen (i) el número de bancos, y (ii) su distribución. En la literatura se han usado las siguientes medidas de concentración: (i) el índice Herfindlah-Hirschman (HHI); (ii) la razón de concentración del banco k (CR_k); (iii) el índice Hall-Tildeman (HTI); (iv) el índice Rosenbluth (RI); (v) el índice Hannah y Kay (HKI); (vi) el índice U (U); (vii) el índice multiplicativo de Hause (H_m); (viii) el índice aditivo de Hause (H_a); (ix) el índice comprensivo de organización industrial (CCI); y (x) la medida de Entropy.^{22 23}

2.2.2 Principales críticas a las medidas de concentración tradicionales

La medida de concentración CR_k ha sido ampliamente utilizada gracias a que tiene requerimientos simples y limitados en cuanto a los datos que necesita. Sin embargo, una de las principales críticas sobre esta medida es la ausencia de un criterio claro o sustentado en alguna teoría sobre el valor del banco k , implicando que la selección del número de bancos incluidos en el cálculo es de cierta forma arbitraria. Además, como consecuencia del corte arbitrario de bancos, esta medida deja por fuera la influencia de los bancos pequeños en el mercado. Por último, se critica el hecho que todos los bancos incluidos en el cálculo reciben el mismo énfasis.

El índice HHI también es una de las medidas más populares, que ha incluso sido utilizada por autoridades de competencia en algunos países, especialmente en el análisis de fusiones entre dos entidades bancarias (Bikker y Haaf, 2002). Esta medida corrige algunas de las críticas que se plantean sobre el CR_k . Por un lado, se incluyen todos los bancos de la industria, con lo que se evita hacer cortes arbitrarios. Además, se usan ponderadores que dan cuenta de la importancia de cada banco. Ahora bien, aunque esta medida parece responder a todas las críticas que se plantean sobre la CR_k , en relación con el Paradigma ECR se ha encontrado que ambas medidas muestran una relación débil con la rentabilidad cuando también se incluye la variable de participación de mercado en la regresión (Athanasoglou, Brissimis y Delis, 2005).

Otra de las principales críticas que se plantean tanto para el CR_k como para el índice HHI es que no son las medidas adecuadas para capturar la evolución dinámica del mercado. Marlow y Right (1987) sugieren que la importancia de los cambios en la estructura de mercado se debe a que los resultados que exhiben las firmas individuales responden más bien en función de su propio mercado que de las características promedio del mercado. Finalmente, señalan que el limitarse a medir el cambio en índices estáticos seguramente no permite que se capturen los aspectos “evolutivos” de la estructura del mercado. En otras palabras, concluyen que un índice que pretenda medir el grado de potencial comportamiento anticompetitivo entre las firmas debe reflejar los aspectos dinámicos.

2.3 El ranking como medida del nivel de concentración

Siguiendo con la argumentación presentada por Marlow y Right (1987), tenemos que los niveles de competencia en un mercado pueden estar afectados por el tamaño de la participación de mercado de las principales firmas, pero que además es importante considerar la posibilidad de un potencial movimiento hacia y desde las principales posiciones del mercado.

²²Para la construcción de las medidas de concentración es importante determinar cuál es la ponderación de determinadas variables porque esto define qué tan sensible es la medida en cambios al final de la cola en la distribución del tamaño de los bancos (Marfels, 1971; Dickson, 1981).

²³Entre estas medidas están por una parte las que utilizan una ponderación de una unidad a un número (arbitrario) de bancos y de cero a los demás (e.g., CR_k), y las que utilizan la participación de mercado como ponderador (i.e., bancos más grandes tienen más peso) incluyendo todos los bancos del mercado, y otras donde el ranking de cada banco se usa como ponderador, incluyendo todos los bancos del mercado. Existe otra clasificación de las medidas de concentración que depende de si su estructura es discreta o acumulativa. Las discretas muestran la altura de la curva de concentración en un punto (arbitrario). Las acumulativas toman en consideración la distribución de la totalidad de la industria, con lo que cualquier cambio en la estructura de la misma afecta el valor del índice. Las medidas discretas han sido objeto de críticas y de apoyo por parte de diferentes sectores de la literatura. Entre sus ventajas se destaca su simplicidad y la relativamente poca información que requieren para su determinación. Aquellos que apoyan este tipo de medidas establecen que el comportamiento que realmente afecta el mercado es aquel de los bancos dominantes por lo que tomar en consideración la totalidad del mercado solo generaría cambios marginales e implicaría un desgaste innecesario. Por su lado, los críticos sostienen que todos los bancos influyen en cierta medida el comportamiento del mercado; además resaltan que las medidas discretas no toman en consideración los cambios ocurridos en la parte de la industria que no ha sido incluida en el cálculo.

De esta forma, los índices que puedan capturar esta potencialidad se consideran dinámicos y dentro de este grupo están las medidas de movilidad y rotación.²⁴

La medida de *movilidad* es el número de cambios que ocurren en el ranking entre las tres principales firmas del mercado. La medida de *rotación* es el número de firmas por fuera del grupo de los tres primeros que se mueve a este grupo. Dentro de la literatura del Paradigma ECR, se ha señalado que las medidas de estabilidad en el ranking (i.e., movilidad y rotación) son complementarias a las medidas de concentración como variable que recoge la estructura de mercado.²⁵

Hymer y Pashigian (1962) definen rotación como el cambio en el ranking de las principales firmas de una industria. Establecen que muchos autores han sugerido que además de ser un complemento a las medidas de concentración puede llegar incluso a reemplazarlas. Los autores en un principio se enfocan en la rotación que mide el cambio en los primeros tres o cinco agentes de una industria específica. Sin embargo, reconocen que ese corte arbitrario no puede tomarse como una representación general de la situación del mercado. Aun así, señalan que de esta medida se pueden extraer tres posibles resultados: (i) una alta tasa de rotación puede contribuir al cálculo de las razones de concentración señalando si una industria estuvo correctamente definida²⁶; (ii) una alta tasa de rotación puede indicar la dificultad que habría en caso que las empresas pensarán en coludir porque no sería posible llegar a un acuerdo estable; y (iii) el nivel de rotación puede señalar cuál es la dirección que está siguiendo la industria. Por otra parte Boyle y Sorensen (1971) establecen que la medida de rotación permite determinar si existen o no barreras de entrada al mercado y además si existe movimiento ascendente y descendente en la industria. En la misma línea sugieren que la rotación, ya sea un alto nivel o la falta de la misma, influye en los niveles de concentración de la industria. Este aspecto lo estudian desde el punto de vista de salida del mercado y de la estabilidad, estableciendo que en el caso en el que las tasas de salida incrementen los niveles de concentración y estabilidad deberían disminuir y viceversa.

Sin embargo, no podemos desconocer que estas medidas también han estado sujetas a críticas. Bodenhorn (1990) indica algunas deficiencias que se derivan de los cálculos de estas medidas: (i) la movilidad puede llegar a ser cero cuando dos firmas del grupo de los tres primeros salen de este, y (ii) la rotación puede tomar el valor de 1 independientemente de cuál sea la posición inicial y final de las firmas en cuestión (e.g., no diferencia si una firma pasa de la quinta a la sexta posición o de la posición 15 a la 5).

De acuerdo con estas críticas, y siguiendo la línea de lo que se propone en este Trabajo de Grado, en el modelo se incluye una variable que hace las veces de las medidas de movilidad (conjuntamente consideradas) al incluir la posición en el ranking de todos los individuos en la muestra (*rank*) y el cambio en el ranking de cada banco individualmente considerado de un período a otro (*rankchange*). Con esto se pretende corregir algunas de las críticas que se han establecido previamente: (i) son medidas que incluyen a todos los individuos y no tienen un corte arbitrario, (ii) capturan la naturaleza dinámica de la competencia, y (iii) están medidas a nivel de individuo y no de mercado. Como se explicó anteriormente, las medidas agregadas de concentración pueden ser una imagen específica de la industria en un momento determinado del tiempo pero pueden no evidenciar el real nivel de competencia entre las firmas. El hecho de emplear una medida de concentración a nivel de firma y no de mercado puede también contribuir al análisis.

3 Modelo Económico

3.1 Selección del modelo

Siguiendo a Mirzaei, Moore y Lui (2013), sabemos que se ha argumentado que la estructura de mercado es un factor que influye en el poder de cada banco, especialmente en lo que se refiere a la fijación de las tasas de interés que pueden

²⁴En inglés se conocen como *mobility* y *turnover*.

²⁵Ver por ejemplo Joskow (1960).

²⁶“It may indicate that the market or industry was too broadly or too narrowly defined. An industry with rather wild fluctuations, such as leading firms in one year drop to negligible “size” the next year, or disappear (and then perhaps return), is no industry or market at all; it has been too narrowly defined” (Hymer y Pashigian, 1962, pp. 84).

afectar directamente en su desempeño. Como vimos anteriormente, Berger (1995) se apoya en las dos hipótesis de PM: 1) el paradigma ECR que, como se menciona en la sección anterior, implica que en mercados altamente concentrados las firmas pueden fijar precios que son menos favorables para los consumidores; y 2) la hipótesis de PRM que establece que en un escenario de altos niveles de participación de mercado y productos diferenciados, las firmas pueden ejercer poder de mercado en los precios de dichos productos y de esta forma obtener una alta rentabilidad.

El propósito del presente Trabajo de Grado es analizar empíricamente cuál de las dos hipótesis se comprueba de acuerdo con el comportamiento de los precios en la industria bancaria colombiana. Es decir, cuál es la relación (directa o inversa) entre la tasa de interés activa y la estructura de mercado.

Así pues, como mencionamos anteriormente, seguimos los planteamientos de Mirzaei et. al. (2013) para analizar el efecto de la estructura de mercado en las tasas de interés activas de dos tipos de créditos. Para realizar este análisis, incorporamos en el modelo los planteamientos de estas dos hipótesis junto con variables específicas de cada banco.

3.2 Especificación del modelo

La variable de interés en este Trabajo de Grado es la tasa de interés activa²⁷ para dos tipos de créditos: 1) crédito de consumo y 2) tarjeta de crédito. Estudiamos el período comprendido entre enero de 2003 y diciembre de 2014 para 13 bancos. Más adelante se explica la selección de las variables y cómo están medidas.

Desarrollamos un modelo de datos panel tomando como base los modelos empíricos existentes sobre el desempeño de los bancos y su relación con la estructura del mercado. Mirzaei et. al. (2013) siguen el enfoque tradicional para probar las hipótesis de Poder de Mercado, en donde el resultado de la industria está determinado por la estructura del mercado,²⁸

$$r_{it} = f(\text{Market Structure}) \quad (1)$$

donde i indica el i –ésimo banco, t indica el período del tiempo, r representa la tasa de interés activa del banco i en el momento t , y *Market Structure* se refiere a la medida que se emplee para determinar la estructura del mercado; para este modelo las medidas de estructura de mercado incluidas en el modelo son participación de mercado y el ranking.

De acuerdo con lo anterior, la versión extendida del modelo queda como,

$$r_{it} = \alpha + \beta_1 \text{Market Share}_{it} + \beta_2 \text{RANK}_{it} + \varepsilon_{it} \quad (2)$$

Esta formulación del modelo permite diferenciar entre las hipótesis de ECR y PRM. La combinación de $\beta_1 > 0$ y $\beta_2 = 0$ implicaría que bancos con una mayor participación de mercado son más eficientes que sus competidores por lo que tienen un mayor poder de mercado para fijar tasas de interés elevadas y obtener una rentabilidad mayor de lo usualmente esperado (“*supernormal profits*”), apoyando la hipótesis de PRM. Ahora bien, en el caso en el que $\beta_1 = 0$ y $\beta_2 > 0$ obtendríamos justificación para darle la razón al paradigma ECR, porque lo que terminaría determinando el desempeño de los bancos es la concentración que hay en el mercado.

En este modelo también incluimos variables para controlar por temas de eficiencia y características propias de los bancos. Así, la especificación final sería,

$$r_{it} = \alpha + \beta_1 \text{MS}_{it} + \beta_2 \text{RANK}_{it} + \delta_i X_{it} + \varepsilon_{it} \quad (3)$$

²⁷La elección de la tasa de interés activa como variable dependiente se debe a que en el mercado bancario el precio es la tasa de interés establecida por las entidades para cada uno de sus créditos. Sabemos que para cada tipo de crédito hay un precio diferente como resultado de las particularidades de cada uno de ellos, y que no todos los bancos ofrecen al público la misma gama de productos crediticios. De esta forma, se sigue el modelo planteado por Berger y Hannan (1989) para estudiar la relación entre el precio (tasa de interés) cobrada por los bancos por sus productos y la estructura del mercado bancario.

²⁸Dado el propósito de este Trabajo de Grado, introducimos algunas modificaciones a este modelo.

3.3 Selección de las variables

3.3.1 Individuos y período de tiempo

El criterio de selección de los individuos incluidos en la muestra es que el banco esté constituido como institución bancaria en enero de 2003 y exista disponibilidad de información mensual sobre las variables de interés en el período cubierto entre enero de 2003 y diciembre de 2014. Lo anterior, para evitar problemas de panel desbalanceado. Aplicando los anteriores criterios se tiene un total de 11 individuos en la muestra.²⁹

3.3.2 Variable dependiente: tasa de interés activa

En cuanto a la selección de la variable dependiente para hacer la estimación del modelo, se seleccionaron las tasas de interés de dos tipos de créditos: (i) de consumo y (ii) tarjetas de crédito.³⁰ Estos valores se deflataron con la inflación básica³¹ con el objetivo de evitar distorsiones en caso de encontrar cointegración entre las series y aislar los efectos reales de los nominales.

Sobre cada uno de los tipos de cartera se incluyó la información correspondiente a las tasas de interés de colocación utilizadas por cada uno de los individuos de la muestra.³² Para la elección de las tasas de interés activas como variables de interés (en vez de las tasas pasivas) se toma el supuesto que el producto sobre el cual los establecimientos bancarios tienen poder de restricción es sobre los créditos que otorgan. Este argumento se extrae del hecho que los depósitos son una de las principales fuentes de fondeo de los bancos. Así las cosas, no sería lógico esperar que una entidad decida restringir el número de depósitos que ofrece porque al hacerlo estaría comprometiendo parte de los recursos que necesita para sus varias operaciones. Por otra parte, es más común que los bancos impongan restricciones y barreras al otorgar un crédito (i.e. estudio de historia crediticia, certificados laborales, colateral) que para aceptar depósitos; en otras palabras, tiene una mayor injerencia en el momento de manejar sus carteras de crédito.

Sin embargo, la selección de estas variables no implica que futuros estudios no puedan tomar en consideración los depósitos captados por los bancos y sus respectivas tasas de interés pasivas.

3.3.3 Variables explicativas

Para la selección de las variables explicativas se siguieron los lineamientos de Mirzaei et. al. (2013). Para dar cuenta del tamaño de cada banco se incluye el valor de los activos de cada individuo, que a su vez captura los efectos de la eficiencia

²⁹La restricción del análisis a las entidades bancarias se debe a las siguientes razones. En primer lugar, entre el universo de intermediarios financieros, los bancos comerciales son los principales proveedores de crédito a las empresas del sector privado en América Latina, correspondiendo a aproximadamente el 40 % del PIB de la región (IDB, 2016; FOMIN, 2015). Segundo, el nivel de participación de los activos bancarios sobre el PIB colombiano ha mostrado un significativo y notorio aumento entre 1995 y 2014, llegando incluso a triplicar su participación (Martínez et. al., 2016). Tercero, en general cuando se estiman índices de concentración para analizar el paradigma ECR solamente se toman en consideración los bancos porque se asume que existe un bajo nivel de sustituibilidad entre los productos que estos ofrecen y los que se observan en otros sectores; en otras palabras, se parte del supuesto que los bancos y las demás entidades financieras se mueven en mercados diferentes (Martínez et. al., 2016). Sobre este último punto se han presentado varias críticas sobre cuánto se puede estar subestimando la sustituibilidad entre productos, donde la justificación que se usa es que “la participación del sector bancario sea al menos del 90 % o cercana, caso en que la sobrevaloración del índice [de concentración] sería menor si se acepta el supuesto de alta sustituibilidad entre los productos bancarios y de otros sectores financieros, o no existiría tal sobrevaloración si se cree que la sustituibilidad es muy baja” (Martínez et. al., 2016, pp. 67).

³⁰No se incluye como variable la información correspondiente a los créditos de vivienda debido a las características propias de este tipo de crédito en el que se hace la división entre los créditos de vivienda para interés social y los que no tienen dicha distinción. Tampoco se toma en consideración la información sobre el microcrédito por la alta volatilidad en sus colocaciones, el cambio en la regulación de la tasa de usura para este tipo de créditos y porque no todos los individuos de la muestra registran este tipo de información (Chavarro-Sánchez, et al, 2015).

En cuanto a los créditos preferencial, de tesorería, y sobregiros bancarios, al no haber información completa para todos los bancos ni para la totalidad del período cubierto por la evaluación, no fueron incluidos dentro del análisis.

³¹“Las medidas de inflación básica buscan eliminar los movimientos muy volátiles o que están fuera del control de la política monetaria. Por ejemplo: (i) precios de los alimentos y (ii) algunos precios regulados (combustibles, servicios públicos, transporte). Proveen una medida de tendencia de largo plazo de los precios. Esta es la inflación que es directamente afectada por las decisiones de política monetaria.” (Información obtenida en <http://www.banrep.gov.co/es/node/33142>).

³²Sobre este punto es importante tomar en consideración que el comportamiento de cada una de las tasas de interés depende de la estructura propia de cada crédito. Por ejemplo, la tasa de interés para la tarjeta de crédito tiende a ser mayor que las demás por el mayor riesgo asociado a este tipo de cartera como consecuencia de la ausencia de colateral así como de control en el destino del crédito (Chavarro-Sánchez et al, 2015).

de escala.³³ También incluimos el nivel de patrimonio de cada individuo. Se esperaría que a medida que un banco sea más grande va a tener una mayor capacidad de establecer tasas de interés más altas y hacer que el resto de los participantes del mercado lo sigan. Sin embargo, también puede tener la ventaja de ser lo suficientemente grande como para establecer precios bajos y forzar así la salida de otros competidores del mercado.³⁴

Como medida de la resiliencia de cada banco se toman en consideración su nivel de capitalización. Dado que la regulación bancaria colombiana sigue los lineamientos de Basilea es normal esperar que los bancos muestren niveles aceptables de capitalización. Un nivel adecuado de capitalización puede permitirle a un banco tener mayores fuentes de fondeo y por lo tanto tener mayor posibilidad de aumentar su oferta de crédito al público. También se asume que un banco con un buen nivel de capitalización tiene más capacidad de desarrollar nuevos negocios.

Finalmente, se incluye la razón entre los gastos y costos de cada banco sobre sus activos como medida de gastos generales, y a su vez como proxy para medir la eficiencia X .^{35 36}

La siguiente tabla resume la notación utilizada en el modelo:

Símbolo	Definición	Medición
$marketshare_{i,t}$	Participación de mercado de cada banco en el mercado	$\frac{Activos\ banco_{i,t}}{activos\ totales\ industria\ bancaria_t}$
$costratio_{i,t}$	Razón entre gastos y costos	$\frac{Gastos\ totales\ banco_{i,t}}{activos\ totales\ ibanco_{i,t}}$
$cap_{i,t}$	Capitalización de cada banco	$\frac{Patrimonio_{i,t}}{Activos\ banc}$
$Cconsumo_{i,t}$	Monto total de crédito de consumo ofrecido por cada banco en un período específico	Datos de la SF
$tasaCco_{i,t}$	Tasa de interés cobrada por el crédito de consumo por cada banco en un período específico	Datos de la SF
$Cord_{i,t}$	Monto total de crédito ordinario ofrecido por cada banco en un período específico	Datos de la SF
$tasaCord_{i,t}$	Tasa de interés cobrada por el crédito ordinario por cada banco en un período específico	Datos de la SF
$tarj_{i,j}$	Monto total de crédito de tarjeta de crédito ofrecido por cada banco en un período específico	Datos de la SF
$tasatarj_{i,t}$	Tasa de interés cobrada por el crédito de tarjeta de crédito por cada banco en un período específico	Datos de la SF
$rank_{it}$	Posición en el ranking entre los bancos de la muestra con base en los activos	Posición en un rango de 1 a 13
$rankchange_{it}$	Cambio en la posición en el ranking respecto del período anterior	$rank_{i,t} - rank_{i,t-1}$

Cuadro 1: Definición de las variables

4 Metodología Econométrica

4.1 Rationale de la selección de metodología

4.1.1 Metodologías empleadas en la literatura sobre el Paradigma ECR

Como se señaló en la sección 2, los postulados de las teorías del Poder de Mercado han sido objeto de estudio por parte de diversos autores a través de los años. Ahora bien, estos análisis no siempre emplean la misma metodología econométrica

³³Debido a la falta de datos, los autores especifican medidas indirectas de eficiencia con el tamaño de los bancos como proxy de la eficiencia de escala.

³⁴La elección de estas variables también sigue el modelo econométrico planteado por Chavarro-Sánchez et al (2015) en donde, para poder incluir características propias de las entidades financieras, se consideran las variables de tamaño y capitalización. Adicionalmente, la literatura que estudia temas de concentración, eficiencia y cartelización de la industria bancaria incluye algún tipo de variable que referencie el tamaño de las entidades para así poder determinar si existe o no una relación entre el desempeño de los bancos y el nivel de concentración en el mercado

³⁵Athanasoglou, Brissimis y Delis (2005) también señalan que los gastos de los bancos son una variable de importancia al establecer los determinantes del precio o la rentabilidad porque están estrechamente relacionados con el tema de administración eficiente (*efficient management*).

³⁶Con la inclusión de esta variable también se aborda de una de las grandes críticas que se plantean en el paradigma ECR sobre cómo generalmente no se incluyen variables que tomen en consideración características de la eficiencia de las firmas porque lo que usualmente se realiza es una regresión del precio contra medidas de concentración y/o participación de mercado sin incluir variables que den cuenta de la eficiencia.

Por esto, se cuestiona el hecho que estos ejercicios pueden estar ignorando variables que midan la eficiencia que pueden estar relacionadas tanto con los precios como con la estructura de mercado.

para realizar sus análisis. La selección de la metodología ha dependido en gran medida del tipo de datos con los que cuenta el autor y con el objetivo mismo del trabajo.³⁷

Cuando se realizan estudios para comparar entre países y/o regiones, por lo general se emplean datos del tipo de series de tiempo.³⁸ Otro tipo de datos con los que se trabajan son de corte transversal, con lo que generalmente se realizan estudios de las industrias bancarias entre países para obtener una idea general de lo que sucede en la industria en un momento específico del tiempo.

Finalmente, también se han presentado estudios que cuentan con datos tipo panel. En el mismo estudio, Goddard, Molyneux y Wilson (2004) analizan sus datos de los bancos europeos empleando las técnicas de paneles dinámicos, usando el Método Generalizado de Momentos (GMM, por sus siglas en inglés). Por su parte Athanasoglou, Brissimis y Delis (2005) estudian el efecto de los determinantes a nivel de banco, industria y contexto macroeconómico sobre la rentabilidad de los bancos en el mercado bancario griego en el período 1985-2001, usando un modelo económico en el que agrupan las variables explicativas en alguna de las tres categorías de determinantes. Los autores tienen un panel desbalanceado con el que proceden a (i) determinar si existen raíces unitarias, (ii) analizar si los efectos individuales son fijos o aleatorios, y (iii) emplear las técnicas de estimación de panel dinámico. Este es uno de los estudios que sí toma en consideración el tema de las raíces unitarias y las implicaciones que tiene sobre la estimación el hecho de ignorar su presencia. Sin embargo, en este caso se usan las técnicas de panel dinámico mientras que en el presente Trabajo de Grado se realiza un análisis desde la técnica de Pedroni, con lo que se pretende evitar las críticas sobre el empleo de las técnicas de panel dinámico ante la presencia de un macropanel.

Otro estudio en el que se emplean datos tipo panel es conducido por Jumono, Achسانی, et. Al. (2015) para la industria bancaria de Indonesia. En este caso los autores estudian los determinantes de la rentabilidad de los bancos desde el mercado de depósitos, por un lado, y el mercado de crédito, por el otro. Para dar cuenta de la estructura del mercado, se incluyen las variables de participación de mercado y CR_{10} . La metodología econométrica que emplean es la regresión de datos panel con el objetivo de determinar cuál es el mejor modelo entre (i) *Pool Least Squares* (PLS), (ii) efectos fijos, y (iii) efectos aleatorios. En este caso hacen un ejercicio parecido al de Athanasoglou, Brissimis y Delis (2005) pero no se tiene en cuenta el tema la estacionariedad.

Sobre la utilización de estos distintos métodos, Schmalensee (1989) explica las ventajas y desventajas de cada uno. En el caso de las estimaciones de corte transversal rara vez se observan estimadores consistentes de los parámetros estructurales, en parte porque los datos de corte transversal por lo general incluyen desviaciones del equilibrio de largo plazo que, a su vez, suele estar correlacionado con la variable de interés. Por su parte, usando datos panel permite modelar los ajustes del desequilibrio y el empleo de datos a nivel de industria permite minimizar el sesgo.³⁹ Finalmente, las estimaciones de corte transversal permiten obtener hechos estilizados que faciliten la construcción de modelos dinámicos.

A partir de la literatura se puede concluir que la mayoría de casos en donde se emplean datos tipo panel, cuando los autores incluyen medidas de concentración, se miden a nivel de mercado y no a nivel de individuo. Por esto, hasta el momento la técnica de Pedroni no ha sido ampliamente empleada dado que requiere que todas las variables estén a nivel de las dimensiones i y t . Por esta razón, al medir la concentración como la posición en el ranking y el cambio de un año a otro en el ranking, nos permite tener una medida a nivel de individuo que cumpla con las condiciones de las pruebas de Pedroni.

³⁷En relación con las diferentes aplicaciones de estos modelos Athanasoglou, Brissimis y Delis (2005) explican que en los casos en los que se utiliza las técnicas de series de tiempo de corte transversal se parte del supuesto que la variación entre los cortes transversales (entre bancos) en cualquiera de las variables explicativas tiene las mismas implicaciones para la variable dependiente que la variación en el tiempo de esa misma variable para un banco individual. También indican que se pueden relajar estos supuestos y aplicar las técnicas de corte transversal y panel dinámico. En el primer escenario se estudia la relación de la variable dependiente con los valores promedios de las variables explicativas; este caso supone la eliminación de cualquier variación a nivel individual de banco en el tiempo, y esto lleva a que se omita la inclusión de los rezagos. En el segundo escenario se pueden expresar todas las variables como desviaciones de los promedios de los bancos individuales.

³⁸Por ejemplo, Goddard, Molyneux y Wilson (2004) estudian los determinantes de la rentabilidad de los bancos europeos con datos que tienen estructura de series de tiempo de corte transversal. Con estos datos, una de las técnicas que emplean es un modelo de series de tiempo de corte transversal *pooled* usando Mínimos Cuadrados Ordinarios (MCOs).

³⁹Porque habría presencia de variables no observadas específicas a nivel de industria.

4.1.2 Selección de las técnicas de macropanel como metodología econométrica

Al considerar los paneles de series de tiempo, existen dos tipos de paneles que tienen características individuales y a los que se les aplica técnicas diferentes. Primero, un micropanel⁴⁰ es aquel en donde la dimensión T es considerablemente corta y que además tiene las siguientes características: (i) dimensión de corte transversal extensa, (ii) se asume homogeneidad⁴¹, (iii) los dinamismos son mínimos y son tratados como homogéneos, (iv) se permite el uso de instrumentos para tratar la endogeneidad. Para este tipo de panel se aplican métodos de panel dinámico; adicionalmente algunos ejemplos de técnicas para micropanel son (i) efectos estáticos fijos o aleatorios convencionales, y (ii) variable instrumental (IV) o método generalizado de momentos (GMM) convencionales.

Segundo, un macropanel⁴² es aquel que reúne las siguientes características: (i) dimensión de serie de tiempo sustancial, (ii) dimensión de corte transversal moderada, (iii) múltiples variables por cada miembro del panel, (iv) presencia de raíces unitarias, (v) dinámicas no triviales o correlación serial, (vi) heterogeneidad sustancial entre los miembros del panel, y (vii) endogeneidad generalizada. Algunas de las herramientas empleadas al trabajar con este tipo de datos son: (i) pruebas de raíz unitaria, (ii) pruebas de cointegración, (iii) estimación de cointegración, (iv) estimación de cointegración en panel no lineal, (v) prueba de causalidad en largo plazo, y (vi) VAR estructural.

El uso de las técnicas de macropanel (no estacionario) ha resultado útil en términos econométricos porque (i) por una parte, identifica las propiedades deseables de no estacionariedad y cointegración (ej.: robusto para endogeneidad, varias formas de variable omitida, varias formas de simultaneidad, varias formas de error de medición y puede aislar las relaciones de largo plazo de las dinámicas de corto plazo); (ii) puede aplicarse con un conjunto de datos considerablemente más pequeño ya que las técnicas de paneles de series de tiempo permiten que haya variación en la dimensión de corte transversal lo que compensa la poca extensión de la dimensión temporal; (iii) las distribuciones en el límite son generalmente estándar (ej.: en las pruebas de raíz unitaria y cointegración en series de tiempo generalmente se obtiene una distribución no estándar y para las pruebas en panel se obtiene una distribución simple); y (iv) finalmente, es flexible en relación con la modelación de la heterogeneidad.

Es importante tener en cuenta que las técnicas de macropanel no estacionario son indicadas, es decir, tienen una ventaja comparativa sobre las demás, cuando los paneles presentan las siguientes características:

1. La dimensión T es:
2. Demasiado corta para hacer inferencia creíble para cualquier miembro individual.
3. No es lo suficientemente larga para poder tratar las dinámicas con flexibilidad porque en caso que sea muy corto es difícil modelar la flexibilidad de las dinámicas, o porque tradicionalmente las propiedades de la correlación serial varían entre miembros, y, además, se requiere una dimensión T suficiente para que cada miembro pueda dar cuenta de las dinámicas específicas.
4. La dimensión del corte transversal N es:
5. Muy larga para ser tratada como un sistema puro.
6. No es lo suficientemente largo como para sobrepasar la dimensión T . En caso que sea muy largo en relación con T se puede incurrir en distorsiones por tamaño y surge del hecho que los sesgos solo se eliminan en el límite cuando T va creciendo en relación con N .

Siguiendo a Pedroni y Urbain (2004), se propone emplear la metodología de paneles no estacionarios⁴³ haciendo uso de las técnicas aplicables a los macropaneles, cuyas principales ventajas se resumen a continuación:

⁴⁰También conocido como panel longitudinal

⁴¹La heterogeneidad esta usualmente limitada a efectos fijos.

⁴²También conocido como panel de series de tiempo

⁴³La no estacionariedad corresponde a las propiedades de un panel que ocurren en la dimensión temporal y son atributos de las series individuales de los miembros del panel. También se puede definir en términos de raíz unitaria como aquel panel en donde las series de tiempo tienen raíz unitaria en al menos uno de los miembros del panel.

1. Esta metodología permite abordar la no estacionariedad en sí misma, así como tratar algunos de los problemas que surgen en los casos en los que la información es potencialmente diversa entre sus miembros.
2. Representa una ventaja para el análisis de cointegración por cuanto se conservan varios aspectos del análisis de cointegración tradicional (i.e. superconsistencia - Watson, 1984- de los estimadores).
3. La flexibilidad en el tratamiento de la heterogeneidad entre los miembros del panel y por lo general ofrece distribuciones limitadas⁴⁴.
4. Las técnicas de macropanel enfatizan en las propiedades temporales de los datos, debido a que tienen una dimensión T más extensa (que el micropanel), lo que a su vez permite que sea más fácil acomodar los aspectos dinámicos de los datos.
5. Este énfasis en las propiedades temporales de los datos también permite darle más importancia a la variación de los procesos dinámicos entre los miembros del panel, en contraste a las técnicas de micropanel que limitan la heterogeneidad entre los miembros a efectos simples y asumen que las dinámicas son iguales para todos los miembros.

De esta forma, las ventajas de las técnicas de macropanel sobre las de micropanel se resumen en los siguientes argumentos:

1. Las técnicas de micropanel requieren homogeneidad estricta para la dinámica de la correlación serial, mientras que en las técnicas de macropanel se exhibe una sustancial heterogeneidad en la dinámica.
2. Las técnicas de macropanel están orientadas a manejar los problemas que se observan en los paneles que están compuestos por datos agregados, mientras que las técnicas de micropanel aplican en los casos en que se requiere trabajar con los problemas propios que se observan en los paneles que están compuestos por datos microeconómicos.
3. Los macropaneles están asociados a las dimensiones de los paneles de series de tiempo ($N < T$), mientras que los micropaneles se asocian con lo que se conoce como datos longitudinales ($N \gg T$)⁴⁵. Esta diferencia también se refleja en el hecho que normalmente los micropaneles se enfocan en el proceso autoregresivo de primer orden, mientras que en el caso de los macropaneles la atención se centra en permitir procesos autoregresivos de orden desconocido.
4. Pesaran y Smith (1995) señalan que las técnicas clásicas de paneles dinámicos se limitan a resolver el problema de la heterogeneidad que puede ser recogida como efectos fijos para N largo y T pequeño. La dinámica se extrae y se trata como homogénea. Dado que los métodos de IV y GMM no están diseñados para corregir por la endogeneidad inducida por la heterogeneidad latente, esta lleva a que la estimación empleando estimadores dinámicos (como por ejemplo de Arellano y Bond, 1991) sea inconsistente.⁴⁶
5. Es más, Pedroni y Urbain (2004) también muestran que la diferencia esencial entre un micropanel y un macropanel viene dada por la teoría asintótica, puesto que depende de la velocidad de convergencia de la dimensión temporal y la del corte transversal.

Por las razones anteriores, se concluye que los datos utilizados en la parte empírica tienen las características de un macropanel. Dado que se va a estudiar el comportamiento económico de la industria bancaria en Colombia, hay un número pequeño de individuos y un período de tiempo considerablemente largo.⁴⁷

La presente sección presenta (i) los diferentes estimadores que han sido usualmente empleados para interpretar datos macropanel, (ii) el problema de inconsistencia presente al emplear estimadores de panel clásicos, (iii) las pruebas de raíz unitaria, (iv) las pruebas de cointegración, y (v) la estimación de esa relación de cointegración.

⁴⁴Debido al promedio que ocurre en la dimensión de corte transversal

⁴⁵En contraste, los paneles longitudinales generalmente contienen largas dimensiones de corte transversal y cortas dimensiones de tiempo” (Pedroni y Urbain, 2005, Pp. 6). Traducción propia.

⁴⁶Esto implica que los estimadores no pueden ser interpretados como promedios de los coeficientes heterogéneos subyacentes.

⁴⁷Tal como se mencionó anteriormente, el número de bancos en Colombia es 22 a diciembre de 2014. Por otra parte, información sobre las tasas de interés, activos, depósitos, rentabilidad y cartera de créditos es recolectada, sistematizada y publicada mensualmente por la Superintendencia Financiera.

4.2 Estimando relaciones de largo plazo con datos panel heterogéneos

Siguiendo a Pesaran y Smith (1995), cuando en un panel la dimensión T es suficientemente grande para correr regresiones separadas para cada grupo, existen cuatro formas para calcular el efecto promedio: (i) con el estimador de medias grupales, (ii) agrupando los datos - pooled -, (iii) agregando los datos - para reducir la dimensión -, y (iv) haciendo regresiones de corte transversal. Empleando cualquiera de estas técnicas se obtiene un estimador del coeficiente promedio.

Cuando estamos frente a un modelo dinámico no se puede concluir que el estimador que se obtiene es consistente y no sesgado porque los regresores pueden estar correlacionados serialmente. De este modo, ignorar la heterogeneidad en los estimadores conlleva a una correlación serial en el error, lo que hace que los estimadores sean inconsistentes en los modelos con variable dependiente rezagada.⁴⁸

Por lo anterior, Pesaran propone aplicar modelos dinámicos para los paneles heterogéneos partiendo del supuesto de que hay coeficientes constantes en el tiempo pero que difieren aleatoriamente entre grupos y que la distribución de coeficientes es independiente de los regresores. En este caso se emplea un modelo heterogéneo dinámico simple:

$$y_{it} = \lambda_{i,t-1} + \beta_i' x_{it} + \varepsilon_{it} \quad (4)$$

con $i = 1, 2, \dots, N$ y $t = 1, 2, \dots, T$

Los coeficientes λ_i y β_i varían a través de los grupos de acuerdo a la siguiente ecuación aleatoria:

$$H_a : \lambda_i = \lambda + \eta_{1i} \quad (5)$$

$$\beta_i = \beta + \eta_{2i} \quad (6)$$

Se asume que η_{1i} y η_{2i} tienen media cero y varianza constante. También se supone que los momentos de órdenes superiores y los momentos cruzados existen y son finitos. Este es un modelo que introduce el parámetro de heterogeneidad en los coeficientes de corto plazo.⁴⁹

Siguiendo los parámetros de Pesaran, se formulan los siguientes supuestos:

Supuesto 1: x_{it} y ε_{is} están independientemente distribuidas para todo t y s , y ambas variables también están independientemente distribuidas de η_{1i} y η_{2i} . Las perturbaciones ε_{it} tienen media cero y varianza σ_i^2 , que son constantes en el tiempo.

Supuesto 2: Los x_{it} son estacionarios en covarianza y además son procesos estocásticos ergódicos continuos en el sentido cuadrático medio.⁵⁰

Supuesto 3: El soporte para los coeficientes aleatorios λ_i residen en el rango estable $(-1,1)$. Además, todos los momentos cruzados de λ_i y β_i existen y son finitos. En particular, $E(\beta_i/(1 - \lambda_i))$ y $E(1/(1 - \lambda_i))$ existen y son finitos.

Supuesto 4: Las matrices de producto cruzado $N^{-1} \sum_{i=1}^N \mu_i \mu_i'$ son no singulares para N finitos y convergen a una matriz finita no singular cuando $N \rightarrow \infty$.

Bajo estos supuestos, y con unos datos adecuados (en términos de N y T), se pueden correr regresiones para cada grupo y los estimadores que se obtienen pueden ser promediados. Los promedios no ponderados y el método de mínimos cuadrados generalizados propuestos por Swamy (1997) van a generar estimadores consistentes para λ y β , bajo H_a .

Pesaran analiza lo que ocurre al emplear los otros tres métodos: (i) *pooled*, (ii) agregado, y (iii) corte transversal. Para el primer método concluye que, bajo los supuestos 1 a 3, $y_{i,t+1}$ y x_{it} están correlacionados con el término error v_{it} , lo

⁴⁸ Aun en los casos en los que $T \rightarrow \infty$.

⁴⁹ Pesaran también incluye la explicación para el caso del largo plazo.

⁵⁰ Es decir, $E(|x_{it}|^2) < \infty$.

que genera que los estimadores *pooled* sean inconsistentes. El problema radica específicamente en que todas las variables que estén correlacionadas con $y_{i,t-1}$ o x_{it} también lo estarán con v_{it} . De esta forma, solo las variables que no estén correlacionadas con los valores latentes de ε_{it} y x_{it} van a tener una correlación cero con v_{it} , lo que a la vez implica que no estarían relacionadas con los regresores.

Suponiendo que las relaciones micro están especificadas por la ecuación (4), incluyendo un término de intercepto, con parámetros aleatorios definidos por H_a , la regresión *pooled* está dada por:

$$y_{it} = \alpha_i + \lambda y_{i,t-1} + \beta' x_{it} + v_{it} \quad (7)$$

$$v_{it} = \varepsilon_{it} + \eta_{1i} y_{i,t-1} + \eta_{2i}' x_{it} \quad (8)$$

Con $i = 1, 2, \dots, N$ y $t = 1, 2, \dots, T$.

La inconsistencia de los estimadores se observa en las siguientes ecuaciones, suponiendo que x_{it} es un escalar con media cero y función de autocovarianza $\gamma_i(s)$,

$$E(x_{it}v_{it}) = \sum_{j=0}^{\infty} E(\eta_{1i}\beta_i\lambda_i^j)\gamma_i(|j+1|)$$

y

$$E(y_{i,t-1}v_{it}) = \sum_{s=0}^{\infty} \sum_{r=0}^{\infty} E(\eta_{1i}\beta_i^2\lambda_i^{r+s})\gamma_i(|r-s|) + \sigma_i^2 \sum_{s=1}^{\infty} E(\eta_{1i}\lambda_i^{2s}) + \sum_{s=1}^{\infty} E(\eta_{2i}\beta_i\lambda_i^s)\gamma_i(|s+1|)$$

Para el segundo caso tenemos que si los x_{it} están correlacionados serialmente, entonces el error en la ecuación agregada también estará correlacionado serialmente, por lo que los estimadores serán inconsistentes en el caso en que se incluya una variable latente en el modelo.

Considerando el modelo dinámico simple RCM definido por la ecuación (6) y asumiendo que los coeficientes de corto plazo λ_i y β_i son aleatorios, es decir, se cumple H_a , y agregando la ecuación (1) en las unidades micro y empleando las ecuaciones (7) y (8), obtenemos la siguiente ecuación:

$$\bar{y}_t = \lambda \bar{y}_{t-1} + \beta' x_t + v_t \quad (9)$$

donde \bar{y}_t y x_t son las medias muestrales de y_{it} y x_{it} en i , y

$$v_t = \varepsilon_t + N^{-1} \sum_{i=1}^N (\eta_{1i} y_{i,t-1} + \eta_{2i}' x_{it}) \quad (10)$$

Con esto se observa que el término de perturbación está relacionado con los regresores en la ecuación agregada y que el método de MCOs no sería consistente aun cuando N y T puedan crecer indiscriminadamente. Para este caso, la especificación óptima agregada de series de tiempo asociada al modelo heterogéneo dinámico estaría dada por:

$$\bar{y}_t = \sum_{j=0}^{\infty} \alpha_j' x_{t-j} + \xi_t \quad (11)$$

donde los ξ_t tienen media cero y están independientemente distribuidos de los valores actuales y latentes de x_t , y con $\alpha_j = E(\beta_i \lambda_i^j)$ para todo $j = 0, 1, 2, \dots$

Finalmente, en el tercer caso también tenemos que el término error está correlacionado con los regresores, así solo el parámetro del intercepto sea lo único que varía entre los grupos, lo que generaría estimadores inconsistentes.

Al agregar en el tiempo la ecuación (1) obtenemos:

$$y_i = \beta_i' x_i + \lambda_i y_{i,-1} + \varepsilon_i \quad (12)$$

donde

$$\begin{aligned} y_{i,-1} &= T^{-1} \sum_{t=1}^T y_{i,t-1}, \\ x_i &= T^{-1} \sum_{i=1}^T x_{it} \\ \varepsilon_i &= T^{-1} \sum_{t=1}^T \varepsilon_{it} \end{aligned}$$

Ahora, bajo H_a , como definida por la ecuación (2), tenemos que la ecuación (9) se convierte a,

$$y_i = \beta' x_i + \lambda y_{i,-1} + v_i \quad (13)$$

donde $v_i = \eta_{1i} y_{i,-1} + \eta_{2i}' x_i + \varepsilon_i$.

La ecuación (14) es el punto intermedio entre la regresión para el modelo dinámico y produce estimadores inconsistentes para β y λ mientras $N, T \rightarrow \infty$, incluso cuando el parámetro en la ecuación (6) sea el único parámetro que varía entre grupos. Lo anterior responde al hecho de que v_i está correlacionado con $y_{i,-1}$ incluso para T largos.

4.3 Problema de inconsistencia al emplear métodos tradicionales de estimación

Al considerar emplear las técnicas aplicables a los macropaneles, es importante tener en cuenta las razones por las que no se pueden emplear técnicas de estimación existentes, como variable instrumental (IV) o el método generalizado de momentos (GMM). Aunque recientemente estas técnicas han sido ajustadas para tratar la presencia de raíces unitarias, el problema radica en el hecho que estas técnicas requieren que las dinámicas sean homogéneas a través de los miembros del panel. Por esto, no es adecuado emplear IV o GMM para tratar con datos macropanel ya que (i) estas técnicas se enfocan en resolver el problema de heterogeneidad de efectos fijos en los casos de un N largo y un T pequeño, (ii) las dinámicas son *pooled* y tratadas como homogéneas, (iii) las correcciones de IV o GMM no pueden corregir por la endogeneidad inducida por una heterogeneidad latente, y (iv) los coeficientes de los estimadores dinámicos se vuelven inconsistentes y ni siquiera se pueden interpretar como promedios.

De esta forma, la crítica de Pesaran y Smith (1995) explica el porqué estos métodos no son adecuados para trabajar con datos macropanel. Para demostrar la inconsistencia que surge al usar los métodos erróneos, consideran la siguiente regresión *pooled* dinámica,

$$\Delta y_{it} = \alpha_i + \rho y_{it-1} + X_{it}' \beta + \mu_{it} \quad (14)$$

donde los verdaderos coeficientes son heterogéneos de tal forma que,

$$\rho_i = \bar{\rho} + \eta_{\rho,i}, \quad \beta_i = \bar{\beta} + \eta_{\beta,i}, \quad (\eta_{\rho,i}, \eta_{\beta,i})' \sim i.i.d$$

donde los valores ρ_i están en la región estable de (1,-1), $X_{it} \sim I(0)$. Lo anterior implica que los residuales son de la siguiente forma,

$$\mu_{it} = \eta_{\rho,i} y_{it-1} + X_{it}' \eta_{\beta,i} + \nu_{it} \quad (15)$$

con lo que se puede concluir que: (i) los residuales están correlacionados serialmente y a la vez correlacionados con la variable dependiente latente y otros regresores; (ii) los métodos de IV o GMM no son adecuados por cuanto cualquier variable relacionada con las variables que debe ser instrumentada también estará correlacionada con los errores; de la misma forma, cualquier variable no relacionada con los errores tampoco estará correlacionada con regresores.

Bajo estas condiciones, Pesaran y Smith (1995) prueban que,

$$\hat{\rho} \rightarrow 1, \hat{\beta} \rightarrow 0 \text{ independientemente de sus valores}$$

este problema existe aun cuando la única fuente de heterogeneidad esté en β_i y que $\hat{\rho} < 1$ dado que la persistencia en X_{it} aumenta.

De acuerdo con todo lo anterior, lo que se realiza en este Trabajo de Grado es (i) las pruebas de raíz unitaria para determinar si las series contienen o no raíces unitarias; (ii) en caso en el que las pruebas nos lleven a la conclusión que no hay evidencia para rechazar el supuesto de existencia de raíces unitarias, se realizan las pruebas de cointegración de Pedroni para determinar si las series están cointegradas; y (iii) finalmente, en caso que sí haya evidencia para decir que las series están cointegradas, se hace la estimación del modelo.

En este último aspecto es importante señalar que las técnicas de estimación que propone Pedroni tienen una serie de ventajas sobre las demás estimaciones. En cuando a las propiedades de muestras pequeñas, la estimación por medias grupales FMOLS presentan una menor distorsión relativa; tienen mejores propiedades de pequeñas muestras que en los casos puros de series de tiempo; y en los casos en los que FMOLS presentan problemas en series de tiempo los FMOLS de media grupal tienen un buen desempeño.

4.4 Sobre las pruebas de raíz unitaria

Los usos de las pruebas de raíz unitaria se resumen a continuación:

1. Las anteriores pruebas de homogeneidad pueden emplearse como pruebas previas para la dinámica de las técnicas de micropanel.
2. Las pruebas básicas de heterogeneidad pueden emplearse como pruebas previas de las condiciones para el análisis de cointegración en el panel.
3. Algunas hipótesis de interés para datos macro se pueden plantear en términos de pruebas de raíz unitaria para paneles.

Existen las llamadas pruebas de raíz unitaria de “primera generación” (ej.: Breitung & Meyer - 1994, Quah - 1994, y Harris & Tzavalis - 1999) que se identifican por tratar con datos en donde hay un gran N y un T finito. Estas pruebas son más aproximadas a las pruebas de micropanel y asumen, necesariamente, homogeneidad en la dinámica. Tienen la ventaja de que se pueden aplicar a paneles muy cortos. La desventaja radica en que es poco probable que el supuesto de dinámica homogénea $\varphi_{ik} = \varphi_k, \forall i$ sobreviva para las aplicaciones agregadas o macro. De esta forma, no son pruebas óptimas para efectuar previas pruebas de cointegración en un panel u otras técnicas de macropanel.

Las pruebas de raíz unitaria de “segunda generación” (ej.: Levin, Lin y Chu - 2002; Im, Pesaran & Shin - 2003; Maddala y Wu -1999) se identifican por tratar con datos en donde hay un gran N y T . Estas pruebas permiten que haya una dinámica heterogénea y, en los casos con un T grande, las dinámicas específicas a cada miembro pueden estimarse. Sus ventajas es que son útiles para las aplicaciones de macropanel y para efectuar pruebas previas de raíz unitaria para paneles cointegrados dado que involucran la dinámica heterogénea de los datos. Algunas de sus desventajas son que requieren que T crezca a mayor velocidad que N , es decir $(N/T) \rightarrow 0$, y también requiere que $T > N$ para evitar distorsiones por tamaño.

De acuerdo con Maddala y Wu (1999), las pruebas de raíz unitaria clásicas - Dickey-Fuller (DF), Dickey-Fuller aumentado (ADF)⁵¹, y Phillips-Perron (PP)- no tienen la potencia suficiente para distinguir un escenario en el que la raíz unitaria es nula de aquel de las alternativas estacionarias. Por esto, reconocen que al emplear pruebas de raíz unitaria para datos panel es una medida que incrementa la potencia de las pruebas de raíz unitaria cuando se realizan sobre una serie de tiempo singular.

Bajo las pruebas de raíz unitaria para datos panel, la hipótesis que se pretende demostrar es la siguiente: $\Delta y_{i,t} = \rho_i y_{i,t-1} + \varepsilon_{it}$, para $i = 1, 2, \dots, N$ y $t = 1, 2, \dots, T$. Para las pruebas de raíz unitaria de los datos panel evalúan la siguiente hipótesis:

$$H_0 : \rho_i = 0, \text{ vs. } H_1 : \rho_i < 0$$

⁵²para todo $i = 1, 2, \dots, N$.

4.4.1 Prueba de Levin-Lin (LL)

La primera forma de especificación de la prueba es examinar si el coeficiente que puede representar la existencia de una raíz unitaria es estadísticamente igual a cero, en cuyo caso habría evidencia para concluir que existe una raíz unitaria. Así, la forma básica de esta prueba es,

$$\Delta y_{it} = \alpha_i + \delta_i t + \theta_t + \rho_i y_{i,t-1} + \xi_{i,t}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T.$$

Esta especificación permite la existencia de efectos fijos, tendencias de tiempo específicas en adición a los efectos temporales comunes. En este caso, la fuente de heterogeneidad son los efectos fijos de cada unidad por cuanto el coeficiente de la variable dependiente latente está restringido a ser homogéneo entre todos los miembros del panel. Para estos casos se manejan las siguientes hipótesis,

$$H_0 : \rho_i = 0, \quad H_a : \rho_i = \rho < 0.$$

Ahora bien, Levin y Lin establecen que en las pruebas de raíz unitaria para datos panel, los estadísticos tienen función de distribución truncada. Por esta razón, proponen una nueva prueba en la cual se abordan los temas de heteroscedasticidad y autocorrelación que está compuesta de los siguientes pasos:

1. Sustraer los promedios de corte transversal de los datos para de esta forma lograr eliminar la influencia de los efectos agregados;
2. Aplicar la prueba Dickey-Fuller aumentada para probar cada serie individual y así normalizar las perturbaciones;
3. Estimar la razón de largo y corto plazo de la desviación estándar para cada una de las series individuales y después calcular la razón promedio del panel;
4. Computar el estadístico de la prueba de panel. Se obtiene el estadístico T ,

$$t_{\rho=0} = \frac{\hat{\rho}}{RSE(\hat{\rho})}$$

donde

$$RSE(\hat{\rho}) = \hat{\sigma}_\varepsilon \left[\sum_{i=1}^N \sum_{t=2+p_i}^T \hat{V}_{i,t-1}^2 \right]^{-1/2}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=2+p_i}^T (e_{i,t} - \hat{\rho} \tilde{V}_{i,t-1})^2$$

⁵¹Por sus siglas en inglés (augmented Dickey-Fuller).

⁵²Cuando se emplean las pruebas de raíz unitaria que no son de datos panel, la hipótesis que se busca probar es $H_0 : \rho_1=0, H_1 : \rho_1 < 0$.

con $T = T - p - 1$ y $p = \frac{1}{N} \sum_{i=1}^N p_i$. Sin embargo, como el estadístico no está centrado en cero, los autores proponen utilizar el estadístico T ajustado,

$$t_{\rho}^* = \frac{t_{\rho=0} - NTS_{NT}\sigma_{\varepsilon}^{-2}RSE(\rho)\mu_T^*}{\sigma_T^*}$$

donde μ_T^* y σ_T^* son la media y desviación estándar ajustados que se obtienen por medio de una simulación Monte Carlo.

Bajo esta prueba, se plantea la siguiente hipótesis,

$$H_0 : \rho = 0,$$

$$t_{\rho}^* \rightarrow N(0, 1)$$

Es importante señalar que la mayor limitación de esta prueba, según planteado por Pesaran, es que ρ tiene el mismo valor para todas las observaciones. En caso que se quisiera emplear este estadístico permitiendo la variación de ρ_i , tendríamos las siguientes hipótesis,

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_N = \rho = 0 \quad (16)$$

$$H_1 : \rho_1 = \rho_2 = \dots = \rho_N = \rho < 0 \quad (17)$$

4.4.2 Prueba de Im-Pesaran-Shin (IPS)

Por otra parte, Im, Pesaran y Shin (1997) parten de uno de los modelos planteados por Levin y Lin⁵³ para permitir la variación de ρ (i.e. ρ_i). Bajo este esquema, se tiene un modelo con tendencia lineal para cada uno de los individuos de las N secciones transversales. En este caso, se emplean pruebas de raíz unitaria separadas para cada una de las N secciones transversales.

Algunas de las principales características de esta prueba son que (i) modela las dinámicas de la correlación serial empleando la especificación AR(K) en las diferencias latentes; (ii) la prueba de la media grupal “entre dimensiones” trata los parámetros de interés como heterogéneos entre miembros, se basa en la estimación de la media grupal, y permite que otros parámetros sean particulares a cada miembro.

La construcción de esta prueba consiste en los pasos que se describen a continuación. Primero, se computa la siguiente ecuación para extraer los efectos temporales de la dependencia del corte transversal,

$$\tilde{y}_{it} = y_{it} - \bar{y}_t, \quad \text{con} \quad \bar{y} = N^{-1} \sum_{i=1}^N y_{it}$$

Después se estima la regresión ADF para cada miembro i

$$\Delta y_{it} = \tilde{\alpha}_i + \rho_i y_{it-1} + \sum_{k=1}^{K_i} \varphi_{ik} \Delta y_{it-k} + \eta_{it}$$

usando \tilde{y}_{it} y $\Delta \tilde{y}_{it}$ obtenidos en el primer paso y sabiendo que $\tilde{\alpha}_i$ contiene los aspectos determinísticos particulares a cada miembro (i.e. efectos fijos). Se estima por MCOs para cada miembro y se selecciona un K_i lo suficientemente grande para asegurar que haya ruido blanco η_{it} .

Luego se computa el estadístico t (t_{ρ_i}) para la hipótesis nula $H_0 : \rho_i = 0$ para cada uno de los miembros. Una vez se realiza lo anterior, se procede a computar

$$\bar{t}_{\rho} = N^{-1} \sum_{i=1}^N t_{\rho_i}$$

⁵³Se basan en el modelo 5 planteado por Levin y Lin.

que es la media grupal del valor del estadístico t para el panel.

Finalmente se construye el estadístico para la media grupal del panel a través de la prueba de raíz unitaria,

$$Z_{NT}^{IPSS} = \sqrt{N/\nu} (\bar{t}_\rho - \mu)$$

donde $\mu = E[t_{\rho_i}]$ y $\nu = Var[t_{\rho_i}]$ bajo la hipótesis nula.

Algunas de las críticas sobre esta prueba son:

1. Se requiere hacer simulaciones para trabajar con paneles no balanceados en tanto se asume que para todo i se tienen un mismo T .
2. La distribución es muy sensible a la selección del número de rezagos en la prueba de raíz unitaria del tipo ADF.
3. Los resultados solo aplican para una clase específica de pruebas de raíz unitaria lo que imposibilita, por ejemplo, el uso de ERS.
4. Se supone que existen unas cuantas correlaciones cruzadas entre los miembros del panel, pero se exige que estas se recojan con efectos comunes en el tiempo. Maddala y Wu señalan que este no es un supuesto muy realista dado que si existieran tales correlaciones, sería muy difícil que tuvieran esta forma.

4.4.3 Prueba de Levin, Lin y Chu (2002)

La idea básica es realizar una prueba paramétrica que sea análoga a ADF. De esta forma se puede modelar las dinámicas de la correlación serial empleando la especificación AR(K) en las diferencias latentes.

Para construir esta prueba primero se computa la siguiente ecuación para poder extraer los efectos temporales de la dependencia de corte transversal,

$$\tilde{y}_{it} = y_{it} - \bar{y}_t, \quad \text{con} \quad \bar{y}_t = N^{-1} \sum_{i=1}^N y_{it}$$

1. Estimar la regresión ADF para cada miembro i ,

$$\Delta y_{it} = \tilde{\alpha}_i + \rho_i y_{it-1} + \sum_{k=1}^{K_i} \varphi_{ik} \Delta y_{it-k} + \eta_{it}$$

2. En caso que se extraigan efectos temporales, se deben usar los \tilde{y}_{it} y $\Delta \tilde{y}_{it}$ del paso (1)
3. Estimar por MCOs individualmente para cada miembro.
4. Seleccionar un K_i lo suficientemente grande para asegurar que haya ruido blanco η_{it} .
5. Utilizar los residuos $\hat{\eta}_{it}$ del paso (2) para computar,

$$\hat{S}_i^2 = T^{-1} \sum_{t=1+K_i}^T \hat{\eta}_{it}^2$$

Esta última ecuación es la varianza residual estimada para cada miembro i y sirve para ponderar a los individuos.

Luego, para cada i se estima

$$\Delta y_{it} = \tilde{\alpha}_i + \sum_{k=1}^{K_i} \varphi_{ik} \Delta y_{it-k} + \varepsilon_{1,it},$$

$$y_{it} = \tilde{\alpha}_i + \sum_{k=1}^{K_i} \varphi_{ik} \Delta y_{it-k} + \varepsilon_{2,it}$$

y establecer

$$\Delta \hat{y}_{it}^* = \hat{\varepsilon}_{1,it}, \quad \hat{y}_{it}^* = \varepsilon_{2,it}.$$

De esta forma se extraen las dinámicas específicas de cada miembro. $\hat{\alpha}_i$ contiene los aspectos determinísticos específicos de cada miembro (i.e. efectos fijos). En este caso también se realiza una estimación por MCOs y se vuelve a elegir un K_i lo suficientemente grande para asegurar que haya ruido blanco η_{it} .

Para cada i se calcula

$$\Delta \hat{y}_{it}^{**} = \hat{S}_i^{-1} \Delta \hat{y}_{it}^*, \quad \hat{y}_{it}^{**} = \hat{S}_i^{-1} \hat{y}_{it}^*$$

donde cada miembro es ponderado por la varianza de los residuales.

Después se estima la regresión del panel,

$$\Delta \hat{y}_{it}^{**} = \rho \hat{y}_{it}^{**} + \mu_{it}^{**}$$

y se calcula el estadístico t correspondiente para $H_0 : \rho = 0$ de acuerdo con la siguiente ecuación,

$$t_{\rho NT} = \frac{\sum_{i=1}^N \sum_{t=2+K_i}^T \hat{y}_{it-1}^{**} \hat{\mu}_{it}^{**}}{\sqrt{\hat{S}_{NT}^{2**} \sum_{i=1}^N \sum_{t=2+K_i}^T \hat{y}_{it-1}^{2**}}},$$

$$\text{con } \hat{S}_{NT}^{2**} = N^{-1} \sum_{i=1}^N \mu_{it}^{2**}$$

Luego se construye

$$Z_{NT}^{LLC} = \nu^{-1/2} (t_{\rho NT} - \mu \sqrt{N})$$

donde

$$\mu = \frac{E[num(t_{\rho_i})]}{E[den(t_{\rho_i})]}, \quad \nu = \frac{Var[num(t_{\rho_i})]}{(E[num(t_{\rho_i})])^2}$$

bajo la hipótesis nula de $\rho_i = 0$. Este es el estadístico de la prueba de raíz unitaria.

4.4.4 Prueba de Maddala y Wu

La idea básica de esta prueba consiste en realizar una prueba paramétrica análoga a ADF. Al igual que IPS, modela la dinámica de la correlación serial empleando la especificación AR(K) para diferencias latentes. Adicionalmente, hace una prueba de significancia marginal acumulada con la que (i) *pools* los valores significativos para las pruebas t de ADF entre los miembros del panel; (ii) trata todos los parámetros como heterogéneos entre los miembros; y (iii) se basa en el estadístico de *Pearson lambda* (también conocido como el estadístico de *Fisher*).

La prueba comienza, en caso que se desee, computando

$$\tilde{y}_{it} = y_{it} - \bar{y}_t, \quad \text{con } \bar{y}_t = N^{-1} \sum_{i=1}^N y_{it}.$$

Luego se estima la regresión ADF para cada miembro i

$$\Delta y_{it} = \tilde{\alpha}_i + \rho_i y_{it-1} + \sum_{k=1}^{K_i} \varphi_{ik} \Delta y_{it-k} + \eta_{it}$$

donde, al igual que en anteriores pruebas, se usan \tilde{y}_{it} y $\Delta\tilde{y}_{it}$ obtenidos en el primer paso, se sabe que $\tilde{\alpha}_i$ contiene los aspectos determinísticos particulares a cada miembro (i.e. efectos fijos), se estima por MCOs para cada miembro y se selecciona un K_i lo suficientemente grande para asegurar que haya ruido blanco η_{it} .

Luego se computa el estadístico t (t_{ρ_i}) para la hipótesis nula $H_0 : \rho_i = 0$ para cada uno de los miembros y se computa el p-value correspondiente (P_i). Estos p-value deben ser obtenidos por simulación dado que t_{ρ_i} tiene una distribución no estándar. Finalmente se utilizan los p-values para computar el estadístico *Pearson lambda* para el panel.

$$P_\lambda = -2 \sum_{i=1}^N \ln P_i$$

Sobre esta prueba es importante señalar algunos de los problemas prácticos que se presentan. En primer lugar la obtención del p-value es no trivial. Además, el estadístico t ADF para $H_0 : \rho_i = 0$ tiene una distribución no estándar; por esto, en la práctica se debe simular la distribución de t_{ρ_i} para obtener el mapa entre los valores t_{ρ_i} de los miembros y sus correspondientes p-value.

4.4.5 La prueba de Fisher ($p\lambda$)

La prueba propuesta por Fisher (1932) parte del supuesto de que si los estadísticos de las pruebas son continuos, los niveles de significancia $\pi_i (i = 1, 2, \dots, N)$ son variables uniformes independientes y que $-2\log_e \pi_i$ tiene una distribución χ^2 con $2N$ grados de libertad.

4.4.6 Criterio de selección de una prueba sobre las demás

Maddala y Wu al realizar una simulación incluyendo alternativas estacionarias y no estacionarias, a partir de un proceso generador de datos (DGP)⁵⁴, llegan a la conclusión que la prueba de Fisher es la más adecuada por cuanto:

1. En aquellos casos en los que no exista correlación en los cortes transversales de los errores, la prueba IPS es más poderosa que la de Fisher por cuanto la primera tiene más poder cuando en ambos casos se tienen el mismo tamaño y, en todo caso, estas pruebas son más poderosas que la LL.
2. Cuando existe heteroscedasticidad y correlación serial en los errores, todas las pruebas tienen la capacidad de encargarse de esos errores. Ninguna, sin embargo, es completamente capaz de manejar problemas de correlación cruzada en los errores. En todo caso, la prueba Fisher es la que mejor maneja este tema, específicamente en los casos en que T es grande y N no lo es tanto.
3. En casos en los que se incluye una mezcla de series estacionarias y no estacionarias en el grupo como hipótesis alterna, la prueba Fisher es la más adecuada porque tiene el mayor poder de distinción entre la hipótesis nula y la alternativa.

4.5 Pruebas de cointegración de Pedroni

Las pruebas de cointegración son útiles por dos razones en particular. Primero, sirven como una prueba previa para los modelos de estimación de cointegración en un panel; segundo, probar la existencia de cointegración ya es en sí misma una prueba valiosa para el análisis. Sobre este último argumento, se debe tener en cuenta que la cointegración puede reflejar relaciones estables a largo plazo entre variables no estacionarias, y también puede ser valioso poder determinar si existe una relación entre las variables, independientemente del momento en que dicha relación empezó.

⁵⁴Por sus siglas en inglés (data-generating process).

Ahora bien, una definición de cointegración, de acuerdo con las notas de clase de Pedroni (2013), es *si y_t , $t=1, \dots, T$, tiene raíz unitaria, tal que $y_t \sim I(1)$, y x_t , $t = 1, \dots, T$, tiene raíz unitaria, tal que $x_t \sim I(1)$, entonces y_t y x_t están cointegradas bajo una combinación lineal, $e_t = y_t - \alpha - \beta x_t$ que sea estacionaria, tal que $e_t \sim I(0)$. De la misma forma, Pedroni también define qué se debe entender por cointegración en un panel como *si y_{it} , $t = 1, \dots, T$, tiene una raíz unitaria para cada miembro $i = 1, \dots, N$, y x_{it} , $t = 1, \dots, T$, tiene una raíz unitaria para cada miembro $i = 1, \dots, N$, entonces y_{it} y x_{it} forman un panel cointegrado si alguna combinación lineal $e_{it} = y_{it} - \alpha_i - \beta_i x_{it}$, es estacionaria tal que $e_{it} \sim I(0)$ para cada miembro $i = 1, \dots, N$.**

En sus publicaciones, Pedroni (1999, 2004) introduce siete estadísticos para probar la hipótesis nula de no cointegración en paneles no estacionarios. El método propuesto por Pedroni utiliza los residuales de la regresión de cointegración dada por

$$y_{i,t} = \alpha_i + \delta_i t + x'_{i,t} \beta_i + e_{i,t}$$

donde $i = 1, 2, \dots, N$, $t = 1, 2, \dots, T$, $\beta_i = (\beta_{1i}, \beta_{2i}, \dots, \beta_{Mi})'$, $x_{i,t} = (x_{1i,t}, x_{2i,t}, \dots, x_{Mi,t})'$, y M siendo el número de regresores. Esta especificación permite la existencia de heterogeneidad en el panel.

De esta forma, se tiene que bajo H_0 y definiendo $z_{i,t} = (y_{i,t}, x'_{i,t})$, $\xi'_{i,t} = (\xi^y_{i,t}, \xi^x_{i,t})$,

$$z_{i,t} = z_{i,t-1} + \xi_{i,t}$$

con una condición para $\xi'_{i,t}$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \xi_{i,t} \Rightarrow \beta_i (\Omega_i) \text{ para cada } i \text{ mientras } T \rightarrow \infty$$

donde $\beta_i (\Omega_i)$ es un vector de Brown con covarianza asintótica dada por Ω_i y asumiendo que Ω_{22i} es positivo finito. La especificación del proceso para $\xi_{i,t}$ impone independencia del corte transversal pero al mismo tiempo permite que exista un amplio rango de dependencia temporal en los datos. Además, no se imponen requerimientos sobre exogeneidad en los regresores.

Bajo estos supuestos Pedroni propone la construcción de siete estadísticos de cointegración en un panel; cuatro de ellos están basados en lo que él describe como *pooling* “intra dimensiones” y tres basados en *pooling* “inter dimensiones”. En la primera categoría se usan tres correcciones no paramétricas y una cuarta que es paramétrica basada en la prueba ADF. Para la segunda categoría dos de las pruebas usan correcciones no paramétricas mientras que la tercera es nuevamente paramétrica basada en la prueba ADF.

Siendo γ_i el coeficiente autorregresivo de los residuales en el i -ésimo individuo o corte transversal, tenemos las siguientes hipótesis,

$$H_0 : \gamma_i = 1, \quad H_a : \gamma_i < \gamma \quad \text{Primera categoría}$$

$$H_0 : \gamma_i = 1, \quad H_a : \gamma_i < 1 \quad \text{Segunda categoría.}$$

Tomando en consideración las regresiones antes presentadas, se calculan diferentes series y parámetros

$$\hat{S}_i^{*2} = \frac{1}{T} \sum_{t=1}^T \hat{\mu}_{i,t}^{*2},$$

$$\hat{S}_{N,T}^{*2} = \frac{1}{N} \sum_{n=1}^N \hat{S}_i^{*2},$$

$$\hat{L}_{11i}^{-2} = \frac{1}{T} \sum_{t=1}^T \hat{\eta}_{i,t}^2 + \frac{2}{T} \sum_{s=1}^{k_t} \left(1 - \frac{s}{k_i + 1}\right) \sum_{t=s+1}^T \hat{\eta}_{i,t} \hat{\eta}_{i,t-s},$$

$$\begin{aligned}\hat{\lambda}_i &= \frac{1}{T} \sum_{s=1}^{k_t} \left(1 - \frac{s}{k_i + 1}\right) \sum_{t=s+1}^T \hat{\mu}_{i,t} \hat{\mu}_{i,t-s}, \\ \hat{S}_i^2 &= \frac{1}{T} \sum_{t=1}^T \hat{\mu}_{i,t}^2, \\ \hat{\sigma}_i^2 &= \hat{s}_i^2 + 2\hat{\lambda}_i, \\ \tilde{\sigma}_{N,T}^2 &= \frac{1}{N} \sum_{n=1}^N \hat{L}_{11i}^2 \hat{\sigma}_i^2.\end{aligned}$$

Ahora bien, para la construcción de los siete estadísticos de Pedroni, se parte de la siguientes ecuaciones:

$$\begin{aligned}\text{panel } \nu &: T^2 N^{\frac{3}{2}} \left(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^2 \right)^{-1}; \\ \text{panel } \rho &: T\sqrt{N} \left(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^2 \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} (\hat{e}_{i,t-1} \Delta \hat{e}_{i,t} - \hat{\lambda}_i); \\ \text{panel } t &: \left(\tilde{\sigma}_{N,T}^2 \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^2 \right)^{-\frac{1}{2}} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} (\hat{e}_{i,t} \Delta \hat{e}_{i,t} - \hat{\lambda}_i); \\ \text{panel ADF} &: \left(\tilde{s}_{N,T}^{*2} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^{*2} \right)^{-\frac{1}{2}} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^* \Delta \hat{e}_{i,t}^*; \\ \text{grupo } \rho &: T \frac{1}{\sqrt{N}} \sum_{i=1}^N \left(\sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^2 \right)^{-1} \sum_{t=1}^T (\hat{e}_{i,t-1} \Delta \hat{e}_{i,t} - \hat{\lambda}_i); \\ \text{grupo } t &: \frac{1}{\sqrt{N}} \sum_{i=1}^N \left(\hat{\sigma}_i^2 \sum_{t=1}^T \hat{e}_{i,t-1}^2 \right)^{-\frac{1}{2}} \sum_{t=1}^T (\hat{e}_{i,t} \Delta \hat{e}_{i,t} - \hat{\lambda}_i); \\ \text{grupo ADF} &: \frac{1}{\sqrt{N}} \sum_{i=1}^N \left(\sum_{t=1}^T \hat{s}_i^{*2} \hat{e}_{i,t-1}^{*2} \right)^{-\frac{1}{2}} \sum_{t=1}^T \hat{e}_{i,t-1} \Delta \hat{e}_{i,t}.\end{aligned}$$

Los estadísticos son ajustados para que estén distribuidos $N(0, 1)$ bajo la hipótesis nula.

Para interpretar los resultados, tenemos que las siguientes hipótesis:

- Para todos los estadísticos

H_0 : No hay cointegración (para todo i , para algunos i)

$$\frac{z_{N,T} - \mu\sqrt{N}}{\sqrt{\nu}} \rightarrow N(0, 1)$$

- Para todos los estadísticos, excepto los estadísticos de varianza *pooled*

H_a : Hay cointegración (para todo i , para algunos i)

$$\frac{z_{N,T} - \mu\sqrt{N}}{\sqrt{\nu}} \rightarrow -\infty$$

En este caso, grandes valores negativos implican el rechazo de H_0 . Como la prueba es de una sola cola los valores críticos son: (i) al 10 %: -1.28, y (ii) al 5 %: -1.64. En caso que se obtenga un valor que esté ubicado a la izquierda de los valores críticos, estaríamos ante un situación en la que se rechaza H_0 .

- En el caso de los estadísticos de varianza *pooled*

$$\frac{z_{N,T} - \mu\sqrt{N}}{\sqrt{\nu}} \rightarrow +\infty$$

Para este caso grandes valores positivos implican el rechazo de H_0 . Igual que en el anterior, la prueba es de una sola cola pero los valores críticos son positivos: (i) al 10 %: 1.28, y (ii) al 5 %: 1.64. Y a diferencia del caso anterior, si se obtiene un valor que esté ubicado a la derecha de los valores críticos, estaríamos ante un situación en la que se rechaza H_0 .

Ahora bien, se pueden hacer ajustes a los estadísticos dependiendo del número de regresores, si se incluyó o no tendencias temporales, y sobre el tipo del estadístico. Y para la interpretación del rechazo de la hipótesis nula, se puede tomar en consideración la explicación de Baltagi (2013), “*el rechazo de la hipótesis nula significa que suficientes de los cortes transversales individuales tienen estadísticos alejados de las medias estimadas por la teoría cuando debían ser generadas bajo la hipótesis nula*”.

4.6 Estimación de la cointegración

4.6.1 Group Mean FMOLS (Pedroni - 2000)

Bajo este escenario se tratan todos los parámetros como heterogéneos entre los miembros del panel. Con esto, se permite (i) por un lado la presencia tanto de dinámicas heterogéneas como de vectores de cointegración exógenos, y (ii) por otro se permite endogeneidad completa.

En pocas palabras, esta técnica es la aplicación en panel del modelo FMOLS para series de tiempo.⁵⁵ Así, tenemos una media grupal de los estimadores FM individuales y cada estimador FMOLS individual corrige por endogeneidad y correlación serial por medio de la estimación directa de la covarianza de largo plazo.

Los pasos para crear el estimador son:

1. Computar

$$\tilde{y}_{it} = y_{it} - \bar{y}_t \quad \text{donde } \bar{y}_t = N^{-1} \sum_{i=1}^N y_{it}$$

$$\tilde{x}_{m,it} = x_{m,it} - \bar{x}_{m,t} \quad \text{donde } \bar{x}_{m,t} = N^{-1} \sum_{i=1}^N x_{m,t}$$

y se extraen los efectos temporales de la dependencia de corte transversal.

2. Estimar la regresión de panel cointegrado

$$y_{i,t} = \alpha_i + \delta_i t + \beta_1 x_{1i,t} + \beta_2 x_{2i,t} + \dots + \beta_M x_{Mi,t} + e_{i,t} \quad \text{para } t = 1, \dots, T; \quad i = 1, \dots, N; \quad m = 1, \dots, M$$

⁵⁵Desarrollado por Phillips y Hansen.

en donde se hace la estimación por MCOs para cada miembro individualmente. En este paso además se incluyen \tilde{y}_{it} y \tilde{x}_{it} en caso que se hayan extraído efectos temporales, se incluye el efecto temporal y las heterogeneidades en las tendencias de tiempo, y se guardan los residuales estimados ($\hat{e}_{i,t}$) para uso posterior.

3. Para cada miembro i se usan los $\hat{e}_{i,t}$ y se forman los vectores de series de tiempo $\xi_{it} = (\hat{e}_{i,t}, \Delta X_{it})$ donde $\Delta X_{i,t}$ es un vector $Mx1$, y se usan para computar la matriz de covarianza de largo plazo para cada miembro $\Omega_i = \sum_{j=-\infty}^{\infty} \Psi_{ij}$ donde Ψ_{ij} es la j -ésima autocovarianza para ξ_i .
4. Para cada miembro i se computan los términos de ajuste

$$\hat{\gamma}_i \equiv \hat{\Gamma}_{21i} + \hat{\Omega}_{21i}^o - \frac{\hat{\Omega}_{21i}}{\hat{\Omega}_{22i}} (\hat{\Gamma}_{22i} + \hat{\Omega}_{22i}^o)$$

que se usan para (i) corregir por las dinámicas de las correlaciones seriales específicas de cada miembro, (ii) corregir la endogeneidad específica de cada miembro, y (iii) se usan las diferencias en regresores como “instrumentos internos”.

5. Usar los términos de ajuste del punto (4) y calcular los estimadores FMOLs específicos de cada miembro y su estadístico t

$$\hat{\beta}_{FM,i}^* = \left(\sum_{t=1}^T (x_{it} - \bar{x}_i)^2 \right)^{-1} \left(\sum_{t=1}^T (x_{it} - \bar{x}_i) y_{it}^* - T \hat{\gamma}_i \right)$$

$$t_{\hat{\beta}_{FM,i}^*} = \left(\hat{\beta}_{FM,i}^* - \beta_{o,i} \right) \left(\Omega_{11i}^{-1} \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 \right)^{\frac{1}{2}}$$

6. Computar el estimador FMOLs de media grupal y el estadístico t

$$\hat{\beta}_{GFM}^* = N^{-1} \sum_{i=1}^N \hat{\beta}_{FM,i}^*$$

$$t_{\hat{\beta}_{GFM}^*} = N^{-1/2} \sum_{i=1}^N t_{\hat{\beta}_{FM,i}^*}^*$$

7. Comparar el estadístico de panel del paso (6) con los valores críticos de las colas de la distribución $N(0,1)$ para rechazar.

En este caso, las hipótesis con las que se trabaja son:

- $H_0 : \beta_i = \beta_0$ (para todos/algunos i); entonces $t_{\hat{\beta}_{GFM}^*} \implies N(0, 1)$
- $H_0 = \beta_i \neq \beta_0$ (para todos/algunos i); entonces $t_{\hat{\beta}_{GFM}^*} \rightarrow \pm\infty$

Esto implica que con valores absolutos grandes se rechaza H_0 . Es importante tener en cuenta que la prueba tiene dos colas.

4.6.2 Parametric Pooled DOLS (Kao y Chang - 2000)

Con este método se parte del hecho de que se usan MCOs en el primer paso para poder obtener los términos de corrección. El estimador ponderado FMOLS entre dimensiones se construye como

$$\hat{\beta}_{NT}^* = \left(\sum_{i=1}^N \hat{L}_{22i}^{-2} \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 \right)^{-1} \sum_{i=1}^N \hat{L}_{11i}^{-1} \hat{L}_{22i}^{-1} \left(\sum_{t=1}^T (x_{it} - \bar{x}_i) y_{it}^* - T \hat{\gamma}_i \right)$$

$$\text{donde } y_{it}^* = (y_{it} - \bar{y}_i) - \frac{\hat{L}_{21i}}{\hat{L}_{22i}} \Delta x_{it} + \frac{\hat{L}_{11i} - \hat{L}_{22i}}{\hat{L}_{22i}} \beta (x_{it} - \bar{x}_i)$$

Las ventajas de este estimador frente al *Group Mean* FMOLS es que aunque tiene las mismas propiedades asintóticas que los FMOLS, al comparar por las propiedades de pequeñas muestras se observa que las distorsiones por grandes muestras son peores en los *pooled* FMOLS que en los *pooled* DOLS. Esta situación se debe precisamente a la distinción entre los métodos paramétricos y no paramétricos.

Pedroni (2013) señala que el estimador group DOLS es preferido a los demás cuando se quiere tener una aproximación paramétrica que permita endogeneidad total, aun cuando para muestras pequeñas no es tan bueno como group FMOLS cuando se tienen componentes MA.

5 Resultados

5.1 Estimación

En esta sección se presentan las tablas que resumen los resultados de la estimación de las diferentes pruebas para los distintos tipos de créditos ofrecidos por los bancos en el mercado colombiano.

Los resultados están basados en una muestra de 11 bancos colombianos (independientemente si son de propiedad extranjera o nacional) que han reportado información a la Superintendencia Financiera desde enero de 2003 hasta diciembre de 2014 (Ver Apéndice B sobre los criterios empleados para la muestra). El tipo de información reportada por estas instituciones consiste en las tasas de interés que ofrecen al público en los diferentes tipos de crédito así como el monto total de cada crédito para un período de tiempo. Esta información es suministrada por la Superintendencia Financiera en formato mensual⁵⁶.

Adicionalmente, los bancos están en la obligación de reportar otro tipo de información que también es suministrada por la Superintendencia Financiera de manera mensual. En este caso se trata de información relativa a los balances generales de los bancos, con datos sobre su nivel de activos, pasivos, patrimonio, entre otros.

5.2 Pruebas de raíz unitaria

De conformidad con lo explicado en la sección (4), el primer paso para poder determinar la relación entre la estructura del mercado bancario sobre las tasas de interés es identificar si existen o no raíces unitarias en las series. Para estos efectos, se realizaron las pruebas de Levin, Lin y Chu (LLC), Im, Pesaran y Smith (IPS), Fisher y Hadri LM. Las dos primeras pruebas (LLC, IPS) se realizaron buscando la parsimonia del modelo bajo los criterios de Akaike (AIC) y Schwarz (BIC), en adición a la regla *rule-of-thumb* de $\frac{T}{4}$.⁵⁷

⁵⁶Los valores vacíos presentes en algunos de los datos se completaron de acuerdo a un promedio trimestral del trimestre inmediatamente anterior al dato vacío.

⁵⁷Los rezagos por criterios de información no aseguran que la prueba esté adecuadamente realizada, por lo que se realizaron otras pruebas con el propósito de determinar que los resultados fueran consistentes.

Las pruebas de raíz unitaria muestran que para las variables del crédito de consumo (monto y tasa de interés), del crédito de tarjeta de crédito (monto y tasa de interés) y para la posición en el ranking no hay evidencia suficiente para rechazar H_0 , es decir, para rechazar la idea de que las series contengan raíces unitarias. Por esto, para estas variables tenemos que el panel no es estacionario, es decir, que sus propiedades estadísticas cambian en el tiempo (ver Apéndice C).

5.3 Pruebas de cointegración

Ahora que ya hemos comprobado que hay evidencia suficiente para concluir que tenemos series no estacionarias en el panel, es importante determinar si existe o no cointegración. En este momento es importante recordar que estamos tratando de encontrar evidencia que sustente la relación que tiene la estructura de mercado sobre los precios (tasas de interés activas). Recordamos entonces que bajo la teoría del Poder Relativo existe el Paradigma ECR para el que la concentración (en este caso medida a través de la posición en el ranking de cada banco) es la variable clave, mientras que para la teoría del Poder Relativo de Mercado es la participación de mercado de cada banco lo que determina que el precio sea favorable o desfavorable para el consumidor.

Ahora bien, al realizar las pruebas de cointegración queremos determinar si existen o no relaciones de cointegración.

Se realizan siete diferentes pruebas de las cuales cuatro son *pooled* “intra dimensional”, tres de media grupal “inter dimensional”. En ambos casos se trata de pruebas paramétricas y semi-paramétricas y hay versiones ponderadas y no ponderadas.

Ahora bien, como se estableció en la sección 4.6, sabemos que valores negativos (o positivos para el caso del estadístico de varianza *pooled*) implican el rechazo de H_0 . Al comparar los resultados con el valor crítico al 5% de significancia vemos que existe evidencia suficiente para rechazar la hipótesis nula, es decir, para concluir que sí hay cointegración.

Para el crédito de consumo y el de tarjeta de crédito vemos que se rechaza H_0 con los siete estadísticos.

Así las cosas, aunque con los siete estadísticos podemos rechazar H_0 , vale la pena igual preguntarse ¿cuál de los estadísticos es más idóneo que los otros? ¿Cuál de todos debemos tomar en cuenta para rechazar o no H_0 ? De acuerdo con Neal (2014) la potencia relativa de cada uno de los estadísticos no está claramente definida e incluso puede haber contradicción entre ellos, tal y como se observa en este caso. Sin embargo, Pedroni (2004) realiza una serie de pruebas para medir la potencia de los estadísticos bajo diferentes escenarios. Después de evaluar la potencia de cada uno al permitir incrementos en N y dejando T fijo concluye que para los casos de datos mensuales contar con un período superior a 20 años permite diferenciar casi por completo los casos más extremos para H_0 de no cointegración. De la misma forma establece que para paneles pequeños el estadístico con mayor potencial es el *rho - grupo* al ser el más conservador en términos de tamaño.

Sin embargo, en este caso llegamos a la misma conclusión utilizando cualquiera de los siete estadísticos: sí existe una relación de cointegración.

5.4 ¿Cómo interpretamos H_a en las pruebas de cointegración?

Una vez revisados los resultados y llegar a la conclusión que hay evidencia suficiente para concluir que hay cointegración, es importante discutir sobre cuál es la hipótesis que se evalúa con estas pruebas. Pedroni señala que H_0 puede tener dos interpretaciones diferentes; por un lado, puede probar si existe o no cointegración para todos los miembros del panel; por el otro, se prueba si existe cointegración para algunos de los miembros. ¿Por qué es esta distinción importante? La forma en que interpretamos el rechazo de H_0 genera implicaciones de diferente naturaleza para el caso que tenemos bajo estudio:

1. Al rechazar H_0 podemos concluir que la relación entre las tasas de interés de los créditos de consumo y de tarjeta de crédito con sus respectivos montos y su posición en el ranking, para *todos* los bancos, está explicada por una tendencia estocástica común de largo plazo.

2. Al rechazar H_0 podemos concluir que la relación entre las tasas de interés de los créditos de consumo y de tarjeta de crédito con sus respectivos montos y su posición en el ranking, para *algunos* bancos, está explicada por una tendencia estocástica común de largo plazo.

En este punto es importante recordar que el propósito de la cointegración es contrastar relaciones tendenciales de equilibrio a largo plazo. Para este caso específico entonces la cointegración entre la serie nos muestra si existe o no una tendencia común de largo plazo entre la tasa de interés activa y algunas variables explicativas. Mirado desde los supuestos de las teorías analizadas en este Trabajo de Grado, lo que se observaría es que en efecto las tasas de interés responden a los cambios en las variables de monto y posición en el ranking porque tienen una tendencia común a largo plazo que no permite que las variables se alejen mucho entre sí.

Ahora bien, al tomar en consideración las dos H_0 que tenemos, es importante tener cuidado a analizar las consecuencias de una y otra. Si tomamos el primer caso, tendríamos que para los todos los individuos de la muestra sus tasas de interés tienen una tendencia común a largo plazo con las variables explicativas y que por esto nunca van a alejarse mucho entre ellas ya que en el largo plazo llegan a un equilibrio (estadístico). Por el contrario, analizamos la segunda opción estamos diciendo que lo anterior es verdad pero solo para algunos de los individuos de la muestra, es decir, que existiría un grupo de individuos para los que las tasas de interés no tienen una relación a largo plazo con las variables explicativas. En otras palabras, si suponemos que estamos en el segundo escenario, podría presentarse el caso que haya bancos para los que la tasa de interés y las demás variables no tienen una relación de largo plazo que explica su relación sino que simplemente las propiedades estadísticas de las series estarían cambiando en el tiempo (porque aunque no estén cointegradas seguirían teniendo raíz unitaria).

La selección de una u otra opción no es caprichosa sino que responde a las propiedades del proceso generador de datos (DGP, por sus siglas en inglés). En el caso en el que el DGP requiera que todos individuos de un panel estén uniformemente cointegrados o no cointegrados, H_a se interpreta como “**todos** los individuos están cointegrados”. En cambio, cuando el DGP permite que algunos de los miembros del panel estén cointegrados y otros no, H_a se interpreta como “una **porción** significativa de los individuos están cointegrados” (Pedroni, 2004).

Ahora bien, dado que DGP no es observable, intentamos inferir con la información disponible en qué escenario nos encontramos. Como bien se planteó al comienzo, estamos utilizando las técnicas de macropanel que permiten que los procesos dinámicos varíen entre los miembros del panel dado que por lo general las propiedades de las series varían entre los miembros. Por lo anterior, podemos hacer un primer intento y concluir que el DGP de este caso es de aquellos que permiten que los *slope coefficients* sean heterogéneos entre los individuos.

De esta forma, podemos concluir que en el presente caso estamos en el segundo escenario en donde el rechazo de H_0 indica que hay evidencia suficiente para considerar que la relación entre las tasas de interés de los créditos de consumo y de tarjeta de crédito con sus respectivos montos y su posición en el ranking, para *algunos* bancos, está explicada por una tendencia estocástica común de largo plazo.

5.5 Estimación de la cointegración

Después de determinar que sí existe cointegración entre las series, pasamos a estimar dicha relación de cointegración. De acuerdo con los planteamientos de la sección 4.6. se realizaron las estimaciones con los dos métodos discutidos,

2(a) Crédito de consumo

Variable	FMOLS				DOLS			
	Grouped	Pooled	Grouped	Pooled	Grouped	Pooled	Grouped	Pooled
Participación de mercado	(1) -66.27 (44.35)	(2) -33.55 (42.45)	(3) 76.61** (21.53)	(4) 72.05** (17.99)	(5) -96.20** (55.71)	(6) -56.04 (53.94)	(7) 71.15** (20.95)	(8) 68.08** (17.28)
Monto del crédito de consumo	(1) -2.27E-05 (8.12E-06)	(2) -2.13E-05 (8.20E-06)	(3) -3.56E-05 (2.30E-06)	(4) -3.55E-05 (2.29E-06)	(5) -2.43E-05 (1.12E-05)	(6) -2.12E-05 (1.12E-05)	(7) -3.54E-05 (2.42E-06)	(8) -3.56E-05 (2.40E-06)
Razón de costos a activos	(1) 7.19** (2.35)	(2) 7.53** (2.41)	(3) 9.95** (3.39)	(4) 9.95** (3.39)	(5) 7.46** (3.93)	(6) 7.90** (4.01)	(7) 9.42** (4.72)	(8) 10.02** (4.73)
Capitalización	(1) -21.04** (11.06)	(2) -26.83** (10.37)	(3) 0.21 (5.21)	(4) 0.42 (5.10)	(5) -24.86** (14.25)	(6) -27.48** (12.55)	(7) 1.53 (5.51)	(8) 2.05 (5.36)
Posición en el ranking	(1) -0.61** (0.32)	(2) -	(3) 0.11 (0.29)	(4) -	(5) -0.76** (0.45)	(6) -	(7) 0.12 (0.29)	(8) -
Cambio en la posición	(1) -	(2) 0.67 (0.68)	(3) -	(4) 0.05 (0.52)	(5) -	(6) 2.37 (1.18)	(7) -	(8) 0.36 (1.74)
NT	1287	1287	1287	1287	1269	1269	1269	1269
R ²	-27.95	-10.11	0.49	0.49	-37.62	-11.44	0.59	0.58
R ² ajustado	-28.29	-10.24	0.49	0.49	-42.72	-13.08	0.53	0.53

2(b) Tarjeta de crédito

Variable	FMOLS				DOLS			
	Grouped	Pooled	Grouped	Pooled	Grouped	Pooled	Grouped	Pooled
Participación de mercado	(9) -233.57** (67.19)	(10) -185.72** (64.17)	(11) -16.34 (26.29)	(12) -7.18 (23.25)	(13) -279.32** (55.71)	(14) -217.07** (84.16)	(15) -27.16 (20.95)	(16) -19.83 (24.00)
Monto del crédito de tarjeta	(9) -5.86E-05 (3.53E-05)	(10) -8.00E-05 (3.47E-05)	(11) 2.59E-06 (4.01E-06)	(12) 2.61E-06 (4.02E-06)	(13) -3.12E-06 (1.12E-05)	(14) -9.34E-05 (6.98E-05)	(15) 6.70E-06 (2.42E-6)	(16) 6.79E-06 (4.33E-06)
Razón de costos a activos	(9) 6.03** (3.30)	(10) 6.31** (3.40)	(11) 4.59 (3.90)	(12) 4.44 (3.91)	(13) 8.79 (3.93)	(14) 8.94 (5.76)	(15) 7.84 (4.72)	(16) 7.85 (5.64)
Capitalización	(9) -23.74 (15.65)	(10) -16.00 (14.88)	(11) 2.86 (6.03)	(12) 1.53 (5.88)	(13) -26.43 (14.25)	(14) -19.20 (18.89)	(15) 1.21 (5.51)	(16) 0.16 (6.52)
Posición en el ranking	(9) -0.43 (0.44)	(10) -	(11) -0.24 (0.33)	(12) -	(13) -0.66 (0.71)	(14) -	(15) -0.20 (0.35)	(16) -
Cambio en la posición	(9) -	(10) 0.10 (0.83)	(11) -	(12) 0.004 (2.48)	(13) -	(14) 0.63 (2.48)	(15) -	(16) 0.10 (2.15)
NT	1287	1287	1287	1287	1269	1269	1269	1269
R ²	-46.45	-44.74	0.05	0.05	-44.04	-54.40	0.10	0.09
R ² ajustado	-47.01	-45.28	0.04	0.04	-49.99	-61.72	-0.01	-0.01

Errores estándar entre paréntesis. El periodo de estimación es de 2003M02-2014M12. La covarianza de largo plazo se estima con un kernel Bartlett y el estimador robusto Newey-West con ancho de banda fijo.

Table 2: Estimaciones de las relaciones de cointegración

5.5.1 Crédito de consumo

Para determinar cuál de todas las especificaciones puede ser más robusta a las condiciones específicas de los datos usados en la estimación, seguimos los planteamientos de Pedroni (2014). De acuerdo con sus notas de clase, para los estimadores FMOLS es preferible utilizar el *grouped* porque (i) tiene una distorsión relativamente menor para muestras pequeñas, (ii) es mucho mejor en las propiedades de muestra pequeña que en los casos de series de tiempo puro, (iii) la potencia de las pruebas tiende a ser baja. Para el caso de los estimadores DOLS, Pedroni (2014) también prefiere utilizar *grouped* porque para muestras pequeñas tiene un mejor desempeño que los *pooled*.

De acuerdo con lo anterior, tenemos que para el caso del crédito de consumo los resultados (1) y (5) nos muestran que un aumento en la participación de mercado genera una disminución en la tasa de interés (situación que favorece a los consumidores). Esta situación puede implicar la industria colombiana no se encuentra en un caso en el que una mayor participación de mercado genere efecto negativos para el bienestar general.

Si además observamos lo que ocurre con las demás variables del modelo notamos que los resultados siguen la intuición económica. Por un lado, ante mayores costos mayor es la tasa de interés que se cobra; esto es un reflejo de que ante un mayor costo una firma va a cobrar un mayor precio. Por otra parte, ante un aumento en la oferta del crédito de consumo (en el monto ofrecido por cada banco) se evidenciaría una disminución en la tasa de interés.

Sin embargo, es importante notar que el resultado para la variable de concentración varía de acuerdo con la especificación empleada. Para el caso (1) y (5) tenemos que ante un aumento en la posición del ranking se genera una disminución de la tasa de interés. Esto va de la mano con lo expuesto anteriormente respecto de cómo es posible tener un alto nivel de concentración y un alto nivel de competencia en un mismo mercado. Para los casos (2) y (6) por el contrario, tenemos que un aumento en el cambio de la posición del ranking genera un mayor precio, es decir, una tasa de interés más elevada. Esta situación apoyaría los supuestos del paradigma ECR que ante una mayor concentración se observan mayores precios, dado que las firmas tienen la capacidad de fijar precios desfavorables para el consumidor.

5.5.2 Tarjeta de crédito

De acuerdo con los criterios señalados en la sección anterior respecto de las ventajas de un estimador sobre el otro, para este caso también vamos a considerar los resultados presentados por los estimadores *grouped*. Así, tenemos que para las dos especificaciones del modelo un aumento en la participación de mercado genera una disminución en la tasa de interés lo que, al igual que en el caso del crédito de consumo, estaría justificando las críticas realizadas al paradigma ECR al considerar que altos niveles de participación de mercado en las firmas no genera necesariamente condiciones desfavorables para los consumidores.

En relación con las variables de concentración empleadas, observamos los mismos resultados que para el crédito de consumo: en la especificación con la posición en el ranking, un aumento en la posición en el ranking genera una disminución en la tasa de interés; en el caso del cambio en la posición, un aumento en el cambio de posición en el ranking genera un aumento en la tasa de interés. Por esto, aplican las mismas consideraciones que en el caso anterior.

6 Conclusión

La concentración de la industria bancaria ha sido uno de los temas que más interés ha generado en la literatura. Diversos autores se han encargado de estudiar el efecto de la concentración en distintas variables y ver cómo afecta en general la estabilidad y funcionamiento del mercado bancario.

Uno de los temas más populares al analizar la relación entre la concentración y el desempeño de la industria y sus precios es el paradigma ECR, que supone que es la concentración del mercado lo que favorece la fijación de precios desfavorables al consumidor y que puede a su vez facilitar la existencia de colusión entre los agentes. Otra hipótesis de la misma línea,

la teoría PMR, supone que es la participación de mercado lo que explica que los precios sean desfavorables pues ante una mayor participación en el mercado, mayor poder de mercado tiene el agente y así impone sus precios.

Lo que hicimos en este Trabajo de Grado fue probar empíricamente cuál de las dos hipótesis se comprobaba a partir de estudiar el comportamiento de las tasas de interés activas para dos tipos de crédito. Empleamos las medidas de rotación y movilidad identificadas por la literatura como complementarias e incluso alternativas a los índices tradicionales de concentración y las incluimos en nuestro estudio al medirlas como la posición del banco i en el ranking en el momento t y como el cambio en esa posición de un período a otro.

En el primer paso de nuestra metodología determinamos si los paneles contenían raíces unitarias. Observamos que en el crédito de consumo para las variables de la tasa de interés, el monto y la posición en el ranking no había evidencia suficiente para rechazar H_0 , con lo que no podíamos concluir que los paneles eran estacionarios. Lo mismo ocurrió en el caso del crédito de tarjeta de crédito.

Con la prueba de cointegración de Pedroni estudiamos si los paneles que previamente habíamos determinado como no estacionarios estaban relacionados por una tendencia estocástica común de largo plazo para el crédito de consumo y de tarjeta de crédito. A partir de los resultados de esta prueba se encontraron relaciones de cointegración entre las variables consideradas.

Ahora bien, de acuerdo con los resultados de la especificación que utiliza un estimador tipo FMOLS *grouped* y DOLS *grouped* se encontró que una mayor concentración no se traduce en mayores precios para los consumidores. Sin embargo, vemos que en las especificaciones alternativas, basadas en los estimadores tipo *pooled*, los resultados difieren de los discutidos. Por esto, es importante reconocer que los resultados son sensibles a la selección de uno u otro estimador, haciendo la salvedad que dicha discrepancia se puede explicar por las propiedades de cada estimador en muestra pequeña.

Finalmente, seguimos en la misma línea de la literatura que si bien la concentración en general puede levantar una sospecha sobre el comportamiento de una industria en específico, no es correcto asimilar que un mercado altamente concentrado va a presentar signos de colusión. Como bien lo establecen Martínez et. al. (2016) un alto nivel de concentración solo puede generar como efecto un punto de partida para una sospecha, no una conclusión en sí misma.

7 Bibliografía

- Athanasoglou, Panayiotis P. Brissimis, Sophocles N. (2008). Bank-specific, industry-specific and macroeconomics determinants of bank profitability. *Journal of International Financial Markets, Institutions and Money*, Vol. 18, 121-136.
- Bajari, Patrick. Ye, Lixin. (2002). Deciding between competition and collusion. *The Review of Economics and Statistics*, 85 (4), 971-989.
- Beck, Thorsten. Demirguc-Kunt, Asli. Maksimovic, Vojislav. (2004). Bank competition and access to finance: International evidence. *Journal of Money, Credit and Banking*, Vol. 36, No. 3, Part. 2: Bank concentration and competition: An evolution in the making. A conference sponsored by the Federal Reserve Bank of Cleveland, 627-648.
- Beck, Thorsten. Demirguc-Kunt, Asli. Levine, Ross. (2005). Bank concentration, competition and crises: First results. *Journal of Banking & Finance* 30 (2006), 1581-1603.
- Beck, Thorsten. (2008). Bank competition and financial stability: friends or foes?. *The World Banks Development Research Group, Finance and Private Sector Team. Policy Research Working Paper* 4656.
- Berger, Allen N. Hannan, Timothy H. (1989). The price-concentration relationship in banking. *The Review of Economics and Statistics*, vol. 71, No. 2, 291-299.
- Berger, Allen N. (1995). The profit-structure relationship in banking - Tests of market-power and efficient-structure hypothesis. *Journal of Money, Credit and Banking*, Vol. 27, No. 2, 404-431.

- Berger, Alan N. Demircug-Kunt, Asli. Levine, Ross. Haunrich, Joseph G. (2004). Bank concentration and competition: An evolution in the making. *Journal of Money, Credit and Banking*, Vol. 36, No. 3, Part. 2: Bank concentration and competition: An evolution in the making. A conference sponsored by the Federal Reserve Bank of Cleveland, 433-451.
- Bikker, Jacob A. Haaf, Katharina. (2002). Competition, concentration and their relationship: An empirical analysis of the banking industry. *Journal of Banking & Finance* 26 (2002), 2191-2214.
- Bourke, P. (1989). Concentration and other determinants of bank profitability in Europe, North America, and Australia. *Journal of Banking and Finance*, Vol. 13, No. 1, 65-79.
- Boyle, Stanley E. Sorensen, Robert L. (1971). Concentration and mobility: Alternative measures of industry structure. *The Journal of Industrial Economics*, Vol. 19, No. 2, 118-132.
- Calem, Paul S. Carlino, Gerald A. (1991). The concentration/conduct relationship in bank deposit markets. *The Review of Economics and Statistics*, Vol. 73, No. 2, 268-276.
- Carlson, Mark. Mitchener, Kris J. (2006). Branch banking, bank competition, and financial stability. *Journal of Money, Credit and Banking*, Vol. 38, No. 5, 1293-1328.
- Cetorelli, Nicola. (2003). Real effects of bank competition. Federal Reserve Bank of Chicago. WP 2004-03.
- Cetorelli, Nicola. Gambera, Michele. (2001). Banking market structure, financial dependence and growth: International evidence from industry data. *The Journal of Finance*, vol. 56, No. 2, 617-648.
- Chavarro-Sanchez, Ximena. Cristiano-Botia, Jose E. González-Molano, Eliana. Huertas-Campos, Carlos. (2015). Evaluación de la transmisión de la tasa de interés de referencia a las tasas de interés del sistema financiero. *Borradores de Economía*, Núm. 874.
- Claessens, Stijn. Laeven, Luc. (2004). What drives bank competition? Some international evidence. *Journal of Money, Credit and Banking*, Vol. 36, No. 3, Part. 2: Bank concentration and competition: An evolution in the making. A conference sponsored by the Federal Reserve Bank of Cleveland, 563-583.
- Deb, Ashis T. Murthy, K.V. Bhanu. (2008). Operationalizing and measuring competition: Determinants of competition in private banking industry in India. AT, Operationalizing and measuring competition: Determinants of competition in private banking industry in India (April 2008).
- Dermircug-Kunt, Asli. Huizinga, Harry. (2000). Determinants of commercial banks interest margins and profitability: some international evidence. *The World Bank Review*, Vol. 13, No. 2, 379-408.
- Engle, Robert F. Granger, W.J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, Vol. 55, No. 2, 251-276.
- Ergungor, O.E. (2004). Comment on “bank competition and access to finance: International evidence “ by Thorsten Beck, Asli Dermircug-Kunt, and Vojislav Maksimovic. *Journal of Money, Credit and Banking*, Vol. 36, No. 3, 21-23.
- Goddard, John. Molyneux, Phil. Wilson, John O.S. (2004). Dynamics of growth and profitability in banking. *Journal of Money, Credit and Banking*, Vol. 36, 1069-1090.
- Heggstad, Arnold A. Rhoades, Stephen A. (1976). Concentration and firm stability in commercial banking. *The Review of Economics and Statistics*, Vol. 58, No. 4, 443-452.
- Hymer, Stephen. Pashigian, Peter. (1962). Turnover of firms as a measure of market behavior. *The Review of Economics and Statistics*, Vol. 44, No. 1, 82-87.
- Jumono, Supto. Achسانی, Noer A. Hakim, Dedi B. Fidaus, Muhamad. (2015) Market concentration, market share, and profitability (study at Indonesian commercial banking in the period of 2001-2012). *Asian Social Science*, 11(27), 18.
- Leon, Florian. (2014). Measuring competition in banking: a critical review of methods. *CERDI Etudes et Documents*, 12: 1-44.

- Levy, Eduardo. Micco, Alejandro. (2003). Banking competition in Latin America. Organization for Economic Cooperation and Development and Inter-American Development Bank, First Meeting of the Latin American Competition Forum.
- Maddala, G.S. Wu, Shaowen. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, Special issue (1999), 0305-90049.
- Marlow, Michael L. Wright, George E. (1987). Measuring market power as competition over time. *Journal of Economics and Business*, Vol. 39, 171-183.
- Martinez, Astril. Misas, Martha. Zuleta, Luis Alberto. Jaramillo, Lino. (2016). La competencia y eficiencia en la banca colombiana. Fedesarrollo, Asobancaria.
- Mirzaei, Ali. Moore, Tomoe. Liu, Guy (2013). Does market structure matter on banks profitability and stability? Emerging vs. advanced economies. *Journal of Banking & Finance*, Vol. 37, 2929 - 2937.
- Nabieu, Gladys (2013). The structure, conduct and performance of commercial banks in Ghana. *European Journal of Business and Innovation Research*, Vol. 1, No. 4, 34-47.
- Pedroni, Peter. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics*, Special issue (1999), 0305-90049.
- Pedroni, Peter. Urbain, Jean-Pierre. (2005). *The Econometrics of Nonstationary Panels*. Clarendon Press. Oxford.
- Pesaran, M. Hashem. Smith, Ronald. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, Vol. 68, Issue 1, 79-113.
- Pesaran, M. Hashem. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Economics*, Vol. 22, No. 2, Heterogeneity and Cross Section Dependence in Panel Data Models: Theory and Applications, 265-312.
- Rose, Peter S. Fraser, Donald R. (1976). The relationships between stability and change in market structure: An analysis of bank prices. *The Journal of Industrial Economics*, Vol. 24, No. 4, 251-266.
- Rubin, Jonathan L. (2004). Cointegration and antitrust: A primer. *Economics Committee Newsletter*, Vol. 4, No. 1.
- Shaek, Klaus. Cihak, Martin. Wolfe, Simon. (2009). Are competitive banking systems more stable? *Journal of Money, Credit and Banking*, Vol. 41, No. 4, 711-734.
- Shehzad, Choudhry Tanveer. De Haan, Jakob. Scholtens, Bert. (2012). The impact of bank ownership concentration on impaired loans and capital adequacy. Elsevier B.V. *Journal of Banking & Finance*. 34 (2010), 399-408.
- Schmalensee, Richard. (1989). Inter-industry studies of structure and performance. R. Schmalensee and R.D. Willig (eds), *Handbook of Industrial Organization*, Amsterdam, Elsevier Science Publishers, Vol. 2, 959-1009.

8 Apéndices

8.1 Apéndice A - Estadísticas descriptivas

La siguiente tabla resume las estadísticas descriptivas para las variables del modelo:

Participación de Mercado				
	Percentiles	Más pequeño		
1%	.01	0		
5%	.03	.01		
10%	.03	.01		
25%	.04	.01	Obs	1584
50%	.06		Media	.0746
		Más grande	Des.Est	.0512
75%	.1	.23		
90%	.15	.23	Varianza	.0026
95%	.2	.23	Asimetría	1.307
99%	.22	.23	Curtosis	3.951

Cap				
	Percentiles	Más pequeño		
1%	.01	0		
5%	.06	0		
10%	.07	0		
25%	.09	0	Obs	1584
50%	.11		Media	.1090
		Más grande	Des.Est	.0413
75%	.13	.32		
90%	.15	.57	Varianza	.0017
95%	.16	.58	Asimetría	3.105
99%	.2	.59	Curtosis	35.800

Crédito de consumo				
	Percentiles	Más pequeño		
1%	1264	661		
5%	3611	691		
10%	8120	713		
25%	38053.5	792	Obs	1584
50%	92305.5		Media	129548
		Más grande	Des.Est	126905.3
75%	172786	662728		
90%	302230	683083	Varianza	1.61e+10
95%	407575	733144	Asimetría	1.5971
99%	538529	762282	Curtosis	5.6371

Tasa de interés Crédito Consumo				
	Percentiles	Más pequeño		
1%	14.73	13.12		
5%	15.52	14.24		
10%	16.26	14.37		
25%	18.61	14.42	Obs	1584
50%	21.07		Media	21.655
		Más grande	Des.Est	4.023
75%	25.15	30.35		
90%	27.27	30.37	Varianza	16.187
95%	28.29	30.42	Asimetría	.157
99%	29.45	30.52	Curtosis	1.981

Tarjeta de crédito				
	Percentiles	Más pequeño		
1%	1891	1490		
5%	4607	1508		
10%	5798	1631		
25%	12697.5	1651	Obs	1584
50%	48946		Media	99563.21
		Más grande	Des.Est	125438
75%	115255.5	597293		
90%	298980	616117	Varianza	1.57e+10
95%	393403	643847	Asimetría	1.8343
99%	515445	670241	Curtosis	5.8204

Tasa de interés tarjeta de crédito				
	Percentiles	Más pequeño		
1%	19.99	16.35		
5%	21.18	16.48		
10%	22.34	16.68		
25%	24.99	17.77	Obs	1584
50%	28.23		Media	27.197
		Más grande	Des.Est	3.190
75%	29.38	32.75		
90%	30.6	32.77	Varianza	10.177
95%	31.4	32.77	Asimetría	-.6674
99%	32.6	32.77	Curtosis	2.6563

Rank				
	Percentiles	Más pequeño		
1%	1	1		
5%	1	1		
10%	2	1		
25%	3	1	Obs	1584
50%	6		Media	6.002
		Más grande	Des.Est	3.167
75%	9	11		
90%	10	11	Varianza	10.034
95%	11	12	Asimetría	.00496
99%	11	13	Curtosis	1.7878

Rankchange				
	Percentiles	Más pequeño		
1%	-1	-2		
5%	0	-2		
10%	0	-2		
25%	0	-2	Obs	1584
50%	0		Media	0
		Más grande	Des.Est	.3078
75%	0	2		
90%	0	2	Varianza	.0947
95%	0	2	Asimetría	-.2599
99%	1	2	Curtosis	22.387

costratio				
	Percentiles	Más pequeño		
1%	.01	.01		
5%	.02	.01		
10%	.03	.01		
25%	.055	.01	Obs	1584
50%	.1		Media	.1039
		Más grande	Des.Est	.0618
75%	.15	.33		
90%	.19	.34	Varianza	.0038
95%	.21	.36	Asimetría	.6156
99%	.27	.36	Curtosis	3.1726

8.2 Apéndice B - Selección de la muestra

El período considerado para el análisis empieza en enero de 2003 hasta diciembre de 2014. Para diciembre de 2014 había 22 bancos reportando información a la Superintendencia Financiera⁵⁸:

⁵⁸El año de cambio de nombre se establece de acuerdo con el cambio registrado en las bases de la Superintendencia Financiera

<i>Nombre Banco</i>	<i>Nombre anterior</i>	<i>Año cambio nombre</i>
<i>Banco de Bogota</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Popular</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Corpbanca</i>	<i>Banco Santander Colombia</i>	<i>Agosto - 2012</i>
<i>Bancolombia</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Citibank</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco GNB Sudameris</i>	<i>GNB Sudameris</i>	<i>Julio - 2004</i>
<i>BBVA Colombia</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco de Occidente</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Caja Social BSCS</i>	<i>Banco Caja Social</i>	<i>Enero - 2005</i>
<i>Davivienda</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Colpatría</i>	<i>Red Multibanca Colpatría</i>	<i>Julio - 2014</i>
<i>Banco Agrario</i>	<i>N.A.</i>	<i>N.A.</i>
<i>AV Villas</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Procredit</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Bancamía</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco WWB</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Coomeva</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Finandina</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Falabella</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Pichincha</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Cooperativo Coopcentral</i>	<i>N.A.</i>	<i>N.A.</i>
<i>Banco Santander</i>	<i>N.A.</i>	<i>N.A.</i>

A partir de esta lista de entidades, se plantearon diferentes criterios para conformar la muestra final. Es importante recordar que para el análisis realizado era necesario contar con un panel balanceado. De esta forma, los criterios para delimitar la muestra fueron:

1. Entidades que hayan estado conformadas y reportando información a la Superintendencia Financiera desde enero de 2003 y hasta diciembre de 2014.
2. En caso que una entidad haya cambiado de nombre, se siguió el código asignado por la Superintendencia para determinar la evolución de la entidad.
3. En caso que una entidad haya entrado al mercado bancario después del año 2003 no se incluyen
4. Cumpliendo el criterio anterior, que la entidad haya reportado información sobre los créditos de consumo y de tarjeta de crédito durante el período analizado.

Aplicando estos criterios la lista de bancos incluidos en la muestra es la siguiente:

<i>Bancos incluidos en el análisis</i>
<i>Banco de Bogotá</i>
<i>Banco Popular</i>
<i>Bancolombia</i>
<i>BBVA Colombia</i>
<i>Banco de Occidente</i>
<i>Banco Caja Social HSBC</i>
<i>GNB Sudameris</i>
<i>Davivienda</i>
<i>Banco Agrario</i>
<i>Red Colpatría Multibanca</i>
<i>Banco AV Villas</i>

8.3 Apéndice C - Resultados de las pruebas

8.3.1 Pruebas de raíz unitaria

Después de realizar todas las pruebas de raíz unitaria, se obtuvieron los siguientes grados de integración para cada variable:

Variable	Grado de integración
Participación de mercado	$I(0)$
Capitalización	$I(0)$
Monto del crédito de consumo	$I(1)$
Tasa de interés crédito de consumo	$I(1)$
Monto del crédito de tarjeta de crédito	$I(1)$
Tasa de interés crédito tarjeta de crédito	$I(1)$
Posición en el ranking	$I(1)$
Cambio en la posición	$I(0)$
Razón de gastos a activos totales	$I(0)$

Es importante recordar que al momento de realizar las pruebas de raíz unitaria se identificaron aquellas que no eran estacionarias; estas fueron diferenciadas y se volvía a aplicar la prueba de raíz unitaria tantas veces como fuera necesario hasta que la serie fuera estacionaria. Sin embargo, en este caso todas las series que contenían raíces unitarias son $I(1)$.

8.3.2 Pruebas de cointegración

Después de haber definido cuáles eran las series que contenían raíces unitarias, realizamos las pruebas de cointegración para determinar si existe o no una relación de cointegración entre ellas. Es decir, queremos saber si existe una tendencia estocástica común de largo plazo que explique su relación y que las lleve a un equilibrio (estadístico).

Al realizar las pruebas, obtenemos los siguientes resultados,

Estadístico	Tipo	Crédito Consumo [cCo, tasaCco, rank]	Tarjeta de Crédito [tarj, tasatarj, rank]
<i>v</i>	-	3.59**	6.75**
<i>Rho</i>	<i>Pooled</i>	-6.51**	-12.88**
	<i>Grupo</i>	-8.64**	-16.73**
<i>T</i>	<i>Pooled</i>	-5.32**	-9.15**
	<i>Grupo</i>	-7.19**	-11.40**
<i>ADF</i>	<i>Pooled</i>	-1.57**	-2.29**
	<i>Grupo</i>	-2.673**	-2.26**