

UNIVERSIDAD DE LOS ANDES

FACULTAD DE ECONOMIA

ASESOR: JUAN MAURICIO RAMIREZ

**MODELOS PARA LA INFLACIÓN BÁSICA
DE BIENES TRANSABLES Y NO TRANSABLES EN COLOMBIA**

Resumen

En este trabajo se estiman modelos de corto plazo para pronosticar la inflación de bienes transables y no transables en Colombia. Estos modelos no existían en el Banco de la República antes de 2004, y son de gran utilidad para la toma de decisiones de política monetaria. También se evalúan los beneficios, en términos de análisis y de capacidad pronóstico, de utilizar métodos que capturen la posible no linealidad de la curva de Phillips en los datos colombianos. Aunque existen diferentes razones que justifican una relación no lineal de corto plazo entre producto e inflación, cada una de ellas sugiere una forma diferente para la curva. Por esta razón, se utilizan redes neuronales artificiales (ANN) y los mínimos cuadrados flexibles (FLS), procedimientos que tienen la gran ventaja de que no imponen de antemano ninguna forma funcional que pueda sesgar los resultados. Una vez se hace la estimación la inflación básica de transables y de no transables, se realiza una prueba de no linealidad, se analizan las funciones de impulso respuesta de cada uno de los modelos y finalmente se muestra la evaluación de los pronósticos comparados con los de un típico modelo de pronóstico lineal (ARIMA). Se encuentra que la curva de Phillips en Colombia podría ser no lineal y por tanto resulta pertinente considerar modelos no lineales para su estimación. Finalmente, con estos modelos se intenta explicar el proceso de desinflación que ha vivido la economía colombiana en los últimos años tanto en la inflación de transables, como en la de no transables.

Palabras claves

Inflación, Curva de Phillips no Lineal, Redes Neuronales Artificiales, Mínimos Cuadrados Flexibles.

JEL: E31, E37, C45

José Luis Torres.

Mayo 2005

++ Los errores, las opiniones y estimaciones contenidas en este documento son responsabilidad exclusiva del autor.

+++ Se agradece especialmente la colaboración de **Martha Misas** quien es la autora de los programas con que se hicieron las estimaciones y colaboró en todo momento en la elaboración de este artículo.

1. Introducción

En la década de los ochenta, la política macroeconómica funcionaba bajo un esquema muy diferente del que se tiene ahora. En ese tiempo el Banco de la República manejaba la tasa de cambio con el llamado *crawling peg*, que consistía en una serie de mini-devaluaciones diarias con el fin de lograr una meta de mediano plazo. El modelo de desarrollo era el de *promoción de exportaciones y sustitución de importaciones*. De tal forma que para controlar la cantidad de importaciones y regular la competencia a la industria nacional, se contaba con aranceles altos para muchos bienes (en algunos casos del 100% o más), y para otros simplemente se prohibía su importación al no estar dentro de la lista de importaciones permitidas. La dirección de la política macroeconómica en su conjunto (monetaria y fiscal) la hacía el gobierno en cabeza de su ministro de hacienda y el Banco era una parte más del gobierno, que se encargaba de ejecutar los mandatos de la Junta Monetaria. En dicha Junta, se sentaban el gerente del Banco, varios ministros y representantes de los gremios, de tal forma que la política monetaria se usaba como herramienta para el fomento de la economía. La emisión para gastos corrientes del gobierno y créditos de fomento para sectores específicos era algo común. Por esta razón la economía vivió varias décadas con una inflación moderada de cerca del 25%, lo que desarrollo mecanismos formales e informales de indexación, que a la postre dificultaron el proceso de desinflación que comenzó en los noventas.

En 1991 todo este ordenamiento político-económico cambió a partir de la expedición de la nueva Constitución política, que le confirió a la Junta Directiva del Banco de la República la potestad para diseñar la política monetaria, cambiaria y crediticia¹. Pero lo más importante fue que se estableció la independencia del Banco del gobierno central, con el fin de preservar el poder adquisitivo de la moneda. Además se ordenó que a partir de ese momento la Junta debería anunciar la meta de inflación para cada año. Sin embargo esto sólo fue el comienzo del proceso para reducir la inflación y llegar al esquema de inflación objetivo con el que se cuenta ahora. Entre 1992 y 1999 la política monetaria se condujo a partir de metas monetarias intermedias y de una banda cambiaria (con una amplitud promedio de $\pm 7\%$ durante todo el período para absorber posibles choques). Con base en la meta de inflación se calculaba la pendiente de la banda cambiaria y el crecimiento esperado para los medios de pago (M1). La motivación para este esquema era que se creía que la demanda por dinero en Colombia era relativamente estable, gracias en parte al *crawling peg* que había durado varias décadas. Sin

¹ Esta sección se basa en Uribe et. al. 2002.

embargo desde el comienzo se estableció que en caso de que hubiera conflicto entre las metas, el control monetario tendría prioridad. Si bien el esquema suena muy bien, en la práctica las metas monetarias se incumplieron varias veces y la banda se movió hacia abajo en 1994 y dos veces hacia arriba en 1998 y 1999 hasta que finalmente se permitió la flotación del peso. Este hecho liberó al Banco de un objetivo de tasa de cambio y le permitió concentrarse en las metas de inflación.

En 1995 se introdujeron los primeros modelos de pronóstico de inflación y se comenzó a hacer un informe mensual de inflación de carácter interno. En diciembre de 1998 se comenzó a publicar trimestralmente el informe. En 1999 la Junta encontró que estaban dadas las condiciones para seguir una *transición* hacia el esquema de inflación objetivo. Estas condiciones eran: (i) la existencia de metas de inflación explícitas, (ii) la prioridad del control de la inflación sobre cualquier otra meta monetaria, (iii) independencia del Banco Central, y (iv) alguna capacidad técnica mínima para predecir la inflación y para entender los mecanismos de transmisión de la política monetaria. De acuerdo a este nuevo esquema, se definió que la Junta adoptaría un cambio en la postura de la política monetaria cuando una de dos cosas pasaran: cuando los modelos de pronóstico de inflación mostraran una probabilidad baja de cumplimiento de la meta o cuando hubieras desviaciones significativas de la base monetaria con respecto a su *línea de referencia* (aquel nivel de la base que se estima que es consistente con las metas de inflación). De tal forma que las consideraciones monetarias continuaron en el centro del debate, mientras se mejoraba la capacidad de pronóstico y el entendimiento del mecanismo de transmisión de la política. Este esquema ha funcionado hasta la fecha relativamente bien y en la práctica las metas monetarias se fueron abandonando gradualmente porque la demanda por dinero resultó ser mucho más inestable de lo que se pensaba (especialmente después de la introducción del impuesto a los movimientos financieros en 1998). Incluso, en varias ocasiones la base monetaria ha excedido la línea de referencia, pero la postura de política se ha mantenido invariable.

Desde el comienzo el Banco ha establecido sus metas sobre el total del Índice de Precios al Consumidor (IPC), debido a que es el índice de precios que mejor entiende el público y porque esta medida es la que se utiliza en la gran mayoría de contratos. Por esta razón su uso facilita la comunicación de los objetivos de política. Poco a poco se fue aprendiendo que algunos precios eran mucho más volátiles que otros, lo que podía potencialmente enviar señales equivocadas para las decisiones de política y que no todos los

precios respondían al mismo tipo de cosas. Esto llevó a que con el tiempo se fuera dividiendo el IPC en tres grandes grupos. En el primer grupo se encuentran los alimentos que representan el 30% de la canasta y cuyos precios se encuentran determinados fundamentalmente por factores climáticos, y de oferta y demanda en el mercado interno según la época de cosechas. Esta distinción es muy importante porque en la práctica el Banco no puede controlar estos precios a través de los instrumentos de la política monetaria. Por ejemplo, en 2003 la meta de inflación se incumplió porque en los últimos dos meses del año el precio de la papa tuvo un aumento del 100%, lo que se tradujo en un incremento de 1% en la inflación total.

Dentro del 70% restante del IPC hay otro grupo que no responde en principio a los movimientos de la política monetaria, como es la inflación de los bienes y servicios regulados (que comprende el 10% de la canasta). Entre ellos se encuentra las tarifas de agua, energía, y las tarifas de aseo y alcantarillado (que cuentan en la actualidad con comisiones de regulación que fijan sus tarifas con base en metodologías que incorporan los costos medios de las empresas prestadoras de servicios). También se incluye en este grupo la gasolina (cuyo precio lo fija el gobierno en cabeza del Ministerio de Minas y Energía con base en una fórmula que incorpora el precio internacional del petróleo y el precio interno de largo plazo de la gasolina - que es una especie de meta a la que se desea llegar-) y el transporte (sus precios los fijan a nivel local las Secretarías de Tránsito y Transporte).

Finalmente en el tercer grupo se encuentran los bienes y servicios cuyos precios puede afectar el Banco mediante su política monetaria y que excluye el precio de los alimentos y los productos regulados (corresponde al 60% de la inflación). Más recientemente se ha reconocido que dentro de ese 60% restante existen dos grupos principales cuyos precios, tanto teórica como empíricamente, responden a diferentes tipos de choques. El primero de ellos es el grupo de los bienes no transables, que representan alrededor del 35% del índice de precios. Estos productos sólo pueden ser comercializados en el medio local, no compiten en el exterior y su precio se encuentra determinado por la demanda interna. El otro grupo es el de los bienes transables que corresponde aproximadamente al 25% del IPC y que como compiten con productos extranjeros. Sus precios son afectados por factores externos como el tipo de cambio y los precios internacionales. Esta distinción es muy importante en términos de la política monetaria, porque permite estudiar con mayor detalle cómo es la fijación de precios y cómo se afectan por las condiciones locales e internacionales. Esta distinción entre inflación de transables y no transables se hizo especialmente importante en los últimos años, ya que la meta de 2002 se incumplió porque debido al *efecto* Lula (y probablemente a los escándalos

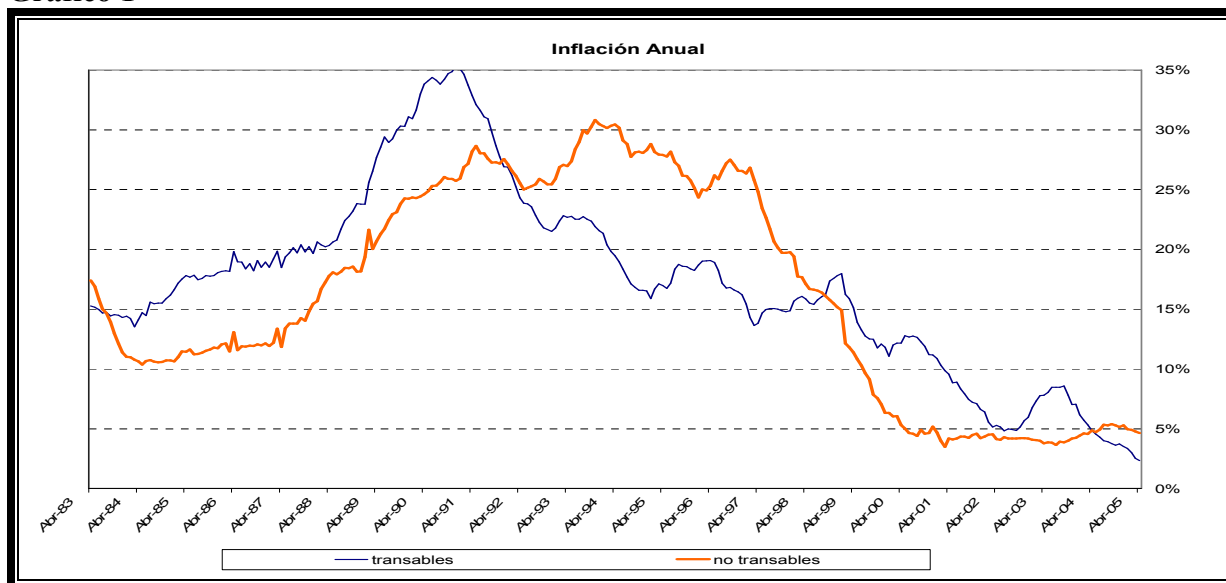
corporativos en Estados Unidos), el peso colombiano (al igual que las monedas de los países de la región) entró en una fase de fuerte depreciación desde mitad de año (con un máximo de 32% anual en abril de 2003) que aumentó el precio de los transables.

Las series con los precios de los precios de los bienes transables y no transables se encuentran a partir de 1982 y se presentan en la Gráfico 1. Podría decirse que si bien se encuentran conectados por una relación de largo plazo, en el corto plazo tienen comportamientos diferentes. En diciembre de 1990 la inflación anual de bienes transables alcanzó su máximo de 35%, mientras que la de los no transables estaba en 29%. El punto máximo de la serie de no transables se encuentra en noviembre de 1993 con una inflación anual de 31%, cuando la de bienes transables era 16%. Llama la atención el hecho de que la inflación de transables ha caído constantemente desde su punto máximo en diciembre de 1990, mientras que la serie de no transables tan sólo comienza a caer en septiembre de 1996. Es evidente que hay factores adicionales que explican el comportamiento de estas dos series.

En el Banco se cuenta con una gran cantidad de modelos de pronóstico para la inflación total, la inflación de alimentos y la inflación sin alimentos. Sin embargo hasta comienzos de 2004 no se contaba con modelos que pudieran pronosticar la inflación de los bienes transables y no transables. El objetivo principal del trabajo es entonces hacer modelos econométricos que permitan comprender y pronosticar la inflación sin alimentos de bienes transables y no transables. Adicionalmente se evalúa la ventaja en términos de error de pronóstico de utilizar un modelo que capture la posibilidad de que la curva de Phillips sea no lineal en Colombia. Para esto se utiliza un modelo de redes neuronales (ANN) y otro de mínimos cuadrados flexibles (FLS), cuya gran ventaja es que no imponen ninguna forma funcional a la relación. Los pronósticos de estos dos modelos se comparan con los de un lineal perfectamente especificado (ARIMA). Adicionalmente se analizan los efectos del tipo de cambio, la inflación en Estados Unidos y la brecha de producto, sobre la inflación de bienes transables y no transables. Por último, con base en los anteriores resultados se intenta explicar el proceso de desinflación que vivió la economía colombiana en los últimos años.

El trabajo consta de ocho partes siendo la primera de ellas esta **Introducción**. En la segunda se introduce **La Curva de Phillips**, que es el marco teórico que se utiliza para la formulación de los modelos a estimar. En la tercera se introducen **El Modelo Triangular de Inflación**, se muestran las variables que se consideraron para cada una de las series de inflación básica (transables y no transables) y se explica el procedimiento que se

Gráfico 1



Fuente: Banco de la República y cálculos del autor.

utilizó para estimar unas curvas de Phillips lineales que sirvieran de punto de comparación para los modelos no lineales. En la cuarta sección se explican brevemente cómo es **La Estimación de una Red Neuronal**, y sus principales ventajas y desventajas. En la quinta sección se ilustra la el proceso de **Selección de las arquitecturas de las dos Redes y Prueba de No Linealidad** de las series. En la sexta sección se introducen **Los Mínimos Cuadrados Flexibles y su Estimación**, sus aplicaciones, ventajas y desventajas. En la séptima se muestra la **Comparación de los Pronósticos de los Modelo Lineales y No Lineales y las Funciones Impulso Respuesta** de cada uno de los modelos ante choques a las variables exógenas y se intenta dar una explicación al proceso de desinflación en Colombia. En la octava y última sección se encuentran las **Conclusiones**.

2. La Curva de Phillips.

La relación de corto plazo entre el producto y la inflación, se encuentra en el centro de la macro economía y constituye la base para la política monetaria. Por esta razón entender la curva de Phillips de corto plazo parece ser una buena manera de comprender el comportamiento de los precios y de pronosticarlos. En su formulación original el profesor Phillips propuso que la curva de corto plazo debería ser convexa al origen, en la medida en que cuando la oferta de la economía se encuentra por encima del producto potencial, es decir cuando la brecha de producto es positiva, cualquier aumento en la demanda agregada se traduce en su mayoría en un aumento en los precios y no en un incremento de la actividad

económica. Por el contrario, cuando la brecha de producto es negativa (producto por debajo del potencial), una disminución en la demanda agregada se traduce principalmente en una disminución de la actividad económica y no en una reducción de precios.

La conclusión de Phillips, aunque lógica, no tiene una justificación teórica que explique la posible no linealidad de la curva. En Dupasquier et. al. (1998) se hace una breve reseña sobre tres justificaciones teóricas que se encuentran en la literatura para que exista en el corto plazo una curva de Phillips no lineal. La primera de ellas viene del modelo de capacidad restringida, donde se supone que las empresas en el corto plazo tienen limitaciones en su capacidad para producir. De esta forma, si las firmas se esforzaran en el corto plazo por aumentar su producción para suplir un aumento en la demanda agregada, eventualmente tiene que llegar el punto en que algunas de ellas se encuentren con cuellos de botella que hagan imposible continuar incrementando su producción. En ese comienzan a haber presiones muy fuertes sobre los precios provocando la convexidad de la curva, pues estos crecen cada vez a una velocidad más grande en la medida en que la producción aumenta. La segunda justificación viene del modelo de Ball et al. (1988) con costos de ajuste, en donde se supone que cambiar precios es costoso para las empresas. Bajo este escenario con costos de menú, no todas las empresas deciden cambiar sus precios al mismo tiempo ante un incremento en la demanda agregada. Sin embargo, las empresas tienen mayores incentivos a variar sus precios en la medida en que más firmas estén cambiando precios, es decir cuando el incremento en la demanda se traduce principalmente en inflación y no en incrementos en el producto. Este modelo predice, por lo tanto, que la curva debería ser convexa. Finalmente se encuentra el modelo de competencia monopolística de Shapiro et. al. (1984), en donde las empresas tienen incentivos para bajar sus precios e intentar sacar del mercado a sus rivales. Por la misma razón no les gusta mucho aumentar sus precios, incluso si el nivel general de precios se encuentra aumentando, con el objetivo de evitar la entrada de nuevos competidores. En este caso también, la curva debería ser cóncava.

Más recientemente ha aparecido en la literatura lo que se conoce como la “nueva curva de Phillips”, que se deriva de los modelos de fijación de precios con fundamentos microeconómicos. Su gran atractivo es que se obtienen a partir del comportamiento racional de los agentes, la presencia de precios rígidos en el corto plazo con implicaciones diferentes para la dinámica de corto plazo de la inflación. Dicha literatura extiende satisfactoriamente los trabajos de precios y salarios traslapados de Taylor (1980) y Calvo (1983). La idea central es que la fijación de precios puede derivarse de un escenario de competencia monopolística donde

las firmas y los agentes están optimizando su bienestar. Al agregar sobre todos los individuos se obtiene una relación de corto plazo entre la inflación y el valor esperado del flujo de costo marginal para la empresa. El problema es que, como lo muestran Galí et al. (2000), conciliar la teoría con lo que muestran los datos no ha sido fácil. Especialmente porque estos modelos no pueden explicar la persistencia que presenta la inflación sin apelar a alguna fuente de rigidez en los precios. Por esta razón, en la actualidad se encuentran muchos autores trabajando en la manera de hacer estimaciones confiables que se ajusten a las relaciones que sugieren los modelos. La gran ventaja en términos empíricos de esta corriente de investigación, es que si bien el costo marginal no es observable (como tampoco el nivel de producto potencial), sus mediciones son en general mucho más estables que las de la brecha de producto que presentan por lo general variaciones de alta frecuencia. En el presente artículo se trabajó con la curva de Phillips tradicional, puesto que en general se ajusta mejor a los datos, y porque el objetivo del presente artículo es no sólo estimar un modelo sino utilizarlo para hacer pronósticos.

En la gran mayoría de estimaciones empíricas de la curva de Phillips se ha utilizado un modelo lineal (Elliason, 2001). Es evidente que hacer una estimación lineal es mucho más sencillo, y si el objetivo es simplemente tener una idea de la manera en que opera este mecanismo, hacer el supuesto de linealidad parece no ser muy restrictivo. Sin embargo gracias al desarrollo tecnológico y a la disponibilidad de series más largas, recientemente se ha retomado la inquietud sobre la manera de hacer estimaciones no lineales de la curva de Phillips. En cuanto a resultados empíricos se refiere, Clark et. al. (1996) muestra evidencia a favor de la hipótesis de una curva convexa, Eisner (1997) encuentra que la curva podría ser cóncava, y Filardo (1998) encuentra que la curva parece ser convexa cuando la brecha de producto es positiva y cóncava cuando la brecha es negativa. Todo esto lleva a pensar que la curva de Phillips podría ser no lineal aunque no se sabe exactamente cuál es el tipo de no linealidad que presenta. De tal manera que dependiendo de la teoría la curva podría ser lineal, convexa, cóncava o convexa-cóncava. Como en la mayoría de estimaciones empíricas sobre la curva se supone que la relación es lineal o se impone de antemano la forma de la curva antes de hacer las estimaciones econométricas del caso, los resultados tienden a estar sesgados (Elliason, 2001).

La hipótesis de la no linealidad no es nueva en el ámbito colombiano. A partir de diferentes metodologías López y Misas (1999), Gómez y Julio (2000) y Nigrinis (2003) encuentran evidencia a favor de la hipótesis de no linealidad en los datos colombianos. En Melo y Misas (2004) se prueba la existencia de cambios estructurales en la inflación

colombiana a finales de los noventa y se muestra que con un modelo de mínimos cuadrados flexibles se obtienen mejores pronósticos, en términos de menor error mínimo cuadrático, que con estimaciones tradicionales. Las principales diferencias del presente documento frente a estos trabajos son la separación entre los bienes transables y no transables, lo cual permite entender de mejor manera los determinantes de los precios, y la evaluación conjunta de dos metodologías que se han incluido por separado en trabajos anteriores.

3. El Modelo Triangular de Inflación

Para hacer los modelos de análisis y pronóstico de las inflaciones básicas se estimó una curva de Phillips para cada una de las series a partir de un típico modelo triangular de inflación. Se le llama triangular por la relación empírica que se supone presenta la inflación con tres grupos de determinantes: inercia, demanda y oferta. Siguiendo a Gordon (1996) un modelo general de inflación podría ser de la forma:

$$\Pi_t = \alpha(L)\Pi_{t-1} + b(L)D_t + c(L)O_t + \varepsilon_t .$$

Donde Π representa la inflación obtenida a partir de la diferencia de los logaritmos del índice de precios, (L) denota al polinomio autoregresivo, D es un índice del nivel de actividad económica, O es un vector de variables que capturan posibles choques de oferta y ε es un error ruido blanco. En el primer término de la ecuación se incluyen rezagos de la inflación observada para capturar el hecho de que una parte de la explicación de la inflación observada corresponde simplemente a un fenómeno de inercia. La suma de los coeficientes estimados de la inflación rezagada debería ser uno, porque como argumenta Friedman (1968) el producto en una economía se encuentra determinado en el largo plazo por factores reales como la oferta laboral, el capital físico, el capital humano y la tecnología disponible, y no por cambios en los precios. Esto significa que en el largo plazo la Curva de Phillips es vertical y sólo debería depender de fenómenos reales. Otra manera de entender esta restricción sobre los coeficientes es que debe existir un nivel natural de la variable de demanda que sea consistente con una inflación estable. Por esto en algunos trabajos se utilizan los mínimos cuadrados condicionados para garantizar que en la estimación se incluya la intuición económica, lo cual tiende a producir mejores estimaciones.

El segundo término de la ecuación es una variable que indica el nivel de actividad económica o de exceso de demanda. En este trabajo se utilizaron rezagos de la brecha de producto calculada a partir de la serie desestacionalizada del índice de producción real manufacturera sin trilla de café (IPR) que representa cerca del 30% del PIB. Se utilizó la serie

de producción manufacturera porque la frecuencia de los datos es mensual y la serie de producto de más alta frecuencia que produce el DANE es trimestral. Para desestacionalizarla se utilizó el módulo x-11 multiplicativo y para encontrar el componente permanente de la serie se usó el filtro de Hodrick y Prescott, que es la manera estándar de calcular el producto potencial. El inconveniente de este método es el sesgo del último dato, porque el filtro es muy sensible a la elección del final de la muestra, y por lo tanto, las brechas que se obtienen son demasiado volátiles. Por esta razón es común que la medición sugiera que la economía pasa de auge a recesión en tan sólo un período, lo cual no es económicamente plausible. Podrían haberse considerado otras maneras de hacer el cálculo, pero en general se ha encontrado en trabajos anteriores que la forma de hacer la medición sólo es importante para determinar el nivel de la brecha y que todos los métodos cuentan aproximadamente la misma historia. Como es tradicional, la brecha se define simplemente como la diferencia entre el producto observado y el potencial. El producto potencial corresponde al nivel de producto que es consistente con una inflación estable cuando todos los recursos de la economía se encuentran plenamente utilizados. En otros trabajos para Colombia se han considerado como medidas del nivel de actividad económica, la brecha de empleo en Gómez y Julio (2000) y la brecha de capacidad instalada en Nigrinis (2003).

Finalmente, el tercer componente de la serie es un vector de variables que capturan la presencia de choques de oferta. En Gordon (1997) se hace énfasis en la inclusión de estas variables para la estimación de una curva de Phillips, porque el no tenerlas en cuenta crea un problema de variables omitidas que sesgaría los resultados. En Balakrishnan et. al. (2002) se utiliza un índice de precios de productos importados y una medida de choques a los términos de intercambio. En el presente trabajo se incluyeron rezagos de la inflación de los Estados Unidos y rezagos de la apreciación del tipo de cambio nominal.

Para determinar el número óptimo de variables explicativas con sus respectivos rezagos, se utilizó una estrategia de búsqueda tipo *stepwise*. Lo primero que se hizo fue buscar los mejores modelos en cuanto a R^2 ajustado y al BIC (criterio de información bayesiano), tratando los rezagos de cada una de las variables como si fueran variables explicativas diferentes. Inicialmente se tomaron como posibles variables explicativas los primeros doce rezagos de la inflación que se deseaba explicar, los primeros nueve del tipo de cambio, los primeros nueve de la inflación americana y los primeros quince de la brecha de producto. Luego se hizo una búsqueda del modelo hacia adelante (*forward*), que comienza por correr

regresiones de la inflación contra una variable explicativa. La variable con el estadístico F más grande entra al modelo y luego se corren regresiones de dos variables, la que se escogió por tener la F más grande en las regresiones de una sola variable y todas las otras posibles variables. De nuevo se adiciona al modelo la variable que produzca el estadístico F más grande, para proseguir con regresiones de tres variables y así sucesivamente. El procedimiento se detiene cuando no hay ninguna variable adicional que al entrar al modelo sea estadísticamente significativa al 5%.

Posteriormente se hizo una búsqueda hacia atrás (*backwards*), en donde se comienza regresando la inflación contra todas las posibles variables explicativas y se van eliminando todas aquellas que no sean significativas al 5%. Este procedimiento se detiene cuando en el modelo sólo quedan variables que sean significativas al nivel deseado. Finalmente se hizo la búsqueda con revisión hacia adelante (*stepwise*), que es similar a la búsqueda *forward*, con la diferencia que las variables que ingresan al modelo en algún paso anterior no necesariamente van a aparecer en el modelo final, puesto que en cada uno de los pasos se revisa que la variable que va a ingresar sea significativa y que las que se queden dentro del modelo también lo sean. Obviamente con todos estos criterios de búsqueda se obtuvieron modelos AR-X diferentes para cada una de las inflaciones. Por lo tanto, se procedió a evaluar cuál de todos estos era el mejor modelo autoregresivo en cuanto a BIC, R^2 ajustado y RMSE (puesto que el objetivo de los modelos es hacer pronósticos) para cada una de las series. Los modelos AR-X que se eligieron finalmente fueron los siguientes:

$$\Pi_{tr} = -0.001 + 0.221 \Pi_{tr}(-1) + 0.136 \Pi_{tr}(-2) + 0.090 \Pi_{tr}(-7) - 0.499 \Pi_{tr}(-12) + 0.028 E(-1) + 0.041 E(-4) + 0.067 \text{GAP}(-5) - 0.018 \text{GAP}(-13) + 0.159 \Pi_{USA}(-1)$$

$$\Pi_{ntr} = 0.000 + 0.089 \Pi_{ntr}(-2) + 0.087 \Pi_{ntr}(-5) + 0.803 \Pi_{ntr}(-12) + 0.010 \text{GAP}(-4) + 0.013 \text{GAP}(-6) - 0.012 \text{GAP}(-13)$$

Donde Π_{tr} y Π_{ntr} corresponden a la inflación básica mensual de productos transables y no transables respectivamente, E es la devaluación nominal mensual calculada a partir del tipo de cambio fin de mes, Π_{USA} es la inflación (total) mensual en Estados Unidos y GAP es la brecha del producto definida anteriormente. La inflación básica de bienes no transables depende de los rezagos de sí misma y de la brecha del producto. Para esta serie también se intentó incluir la devaluación y la inflación americana, pero como era de esperarse estas series no resultaron significativas...

A partir de este modelo se comenzó a hacer la estimación de la red neuronal y del modelo de parámetros cambiantes. Esta sería una típica curva de Phillips lineal, por lo que se convierte en el punto de comparación para los modelos no lineales que se presentan en este trabajo. Adicionalmente, en la comparación se incluye un modelo ARIMA que también es lineal pero que tiene el atractivo de no tener variables exógenas, por lo que a los errores de pronóstico del modelo no se le suman los errores en el pronóstico de las variables exógenas con los modelos satélites del Banco.

En el Cuadro 1 se encuentran las medidas tradicionales de dispersión para las variables de los modelos.² En el Gráfico 2 se muestran las series de la inflación anual de bienes transables y la devaluación anual promedio del tipo de cambio nominal. En el Gráfico 3 aparece de nuevo la inflación anual de los bienes transables, pero esta vez con la inflación anual en Estados Unidos y el promedio anual de la brecha del producto.

Cuadro 1

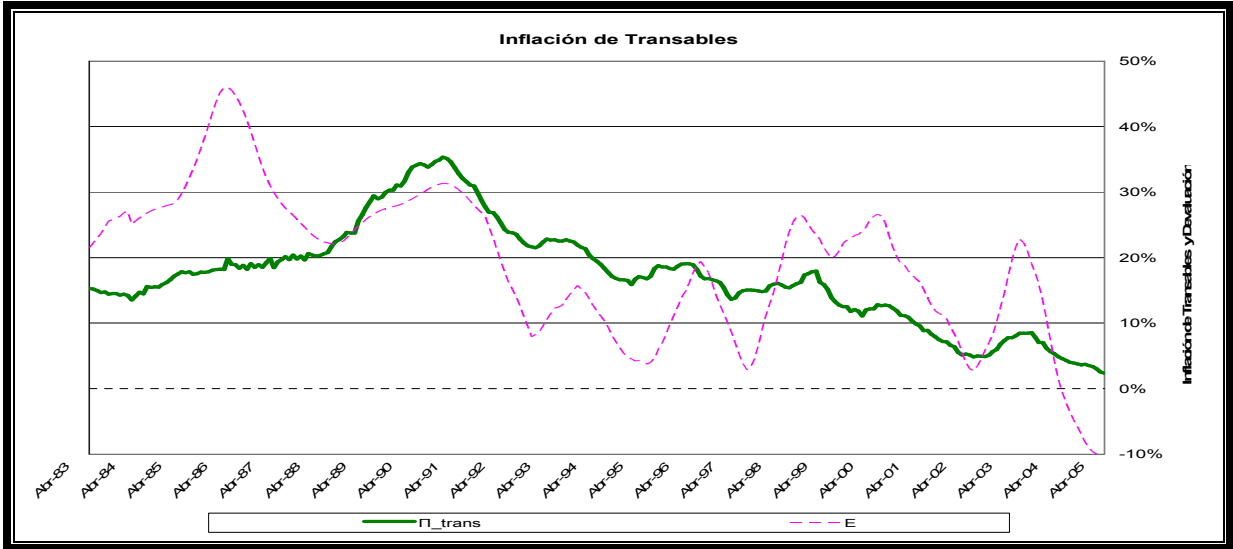
	Media	Desviación
Trans	18%	8%
No Trans	17%	9%
Deval	20%	11%
USA	3%	1%
Gap	0%	2%

Fuente: Banco de la República y cálculos del autor.

Del Cuadro 1 puede decirse que al menos con las medidas tradicionales de dispersión pareciera que la inflación básica de transables y de no transables se comportan de manera similar. También se destaca la alta volatilidad de la devaluación anual promedio. Como es de esperarse, la brecha se encuentra centrada en cero y tiene una desviación anual promedio de dos puntos porcentuales.

² El Cuadro 1, el Gráfico 2 y el Gráfico 3 se hacen a partir de los datos anuales y no de los mensuales que son con los que se estiman los modelos, porque facilita su lectura y con los datos anuales se elimina la estacionalidad.

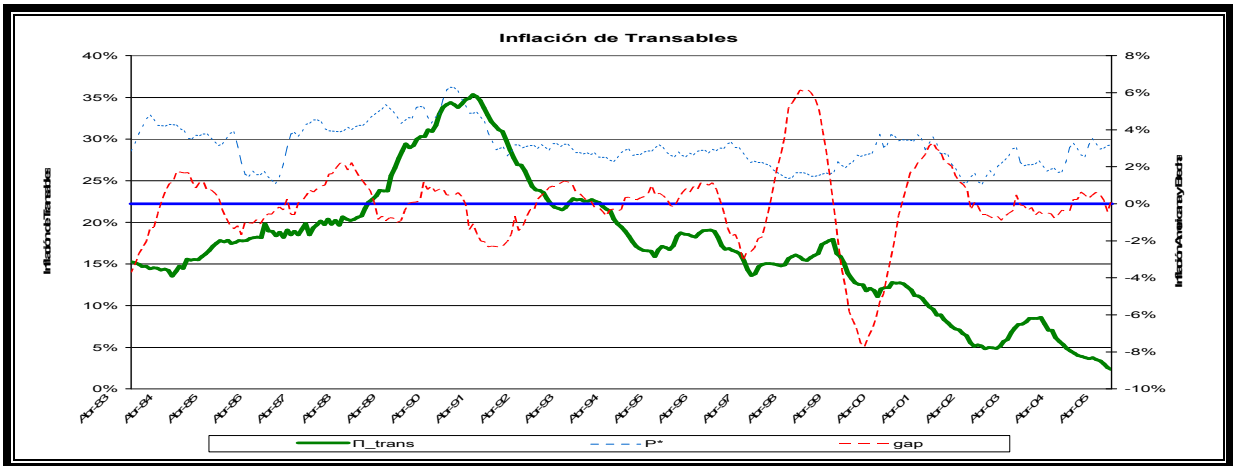
Gráfico 2



Fuente: Banco de la República y cálculos del autor.

El Gráfico 2 y el Gráfico 3 sugieren una relación más estrecha de la inflación básica de los bienes transables con la depreciación que con otras variables, aunque los movimientos del tipo de cambio son más erráticos que los de los precios. Con la brecha las cosas no son tan evidentes, pero se ve claramente que la recesión coincidió con el proceso de desinflación.

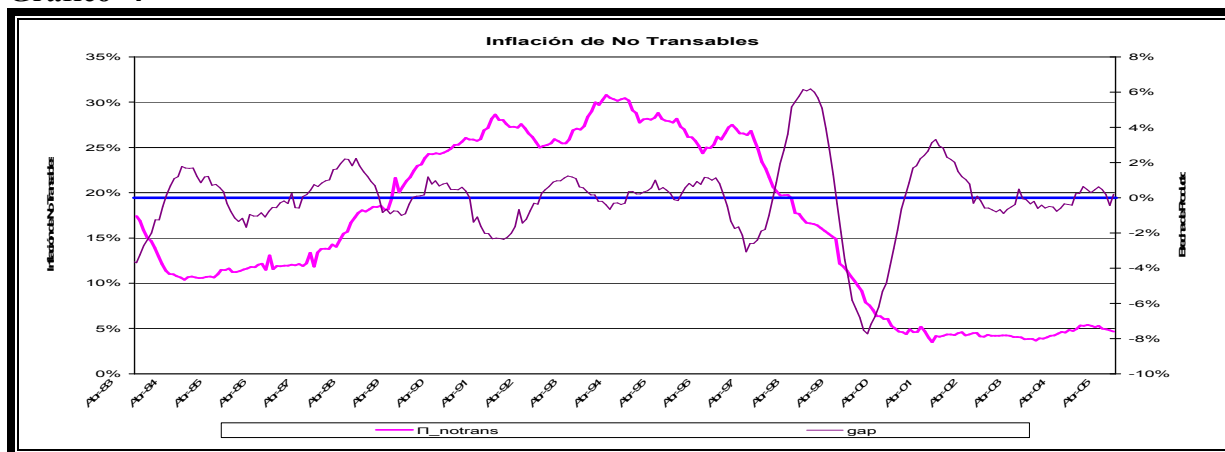
Gráfico 3



Fuente: Banco de la República y cálculos del autor.

Finalmente puede verse que Estados Unidos también ha sufrido un proceso de desinflación que coincide con el colombiano al pasar de inflaciones del orden del 5% a comienzos de los noventa a inflaciones cercanas al 2% en los últimos años, lo cual podría haber contribuido, en algún grado, a explicar nuestro proceso desinflacionario.

Gráfico 4



Fuente: Banco de la República y cálculos del autor.

En el Gráfico 4 aparece la inflación anual de bienes no transables (en el eje de la izquierda) y la serie de la brecha del producto (en el eje de la derecha). Se observa que la desinflación había comenzado antes que la crisis, al pasar de 27% en septiembre de 1996 a 16% en agosto de 1998, que es el momento en que el Gráfico muestra el inicio de la crisis con el desplome de la brecha del IPR (de 6% a -8% en el lapso de un año). Esto sugiere que aunque la crisis ayudó al proceso desinflacionario, hubo otros factores, que contribuyeron a su disminución previa (como la rebaja de aranceles que tuvo lugar desde principios de los noventa y, posiblemente, la caída en las expectativas de inflación por cuenta de una banca independiente).

4. La Estimación de una Red Neuronal.

Una red neuronal es un modelo diseñado para imitar el funcionamiento del cerebro, cuyas neuronas se activan y desactivan ante la presencia de impulsos dependiendo de la intensidad del estímulo. En el contexto de las series de tiempo puede clasificarse como un modelo no lineal entrenado, en la medida en que puede aprender unos patrones que se repiten dentro de muestra y aplicarlos fuera de muestra. La gran ventaja de una red frente a otro tipo de modelos es que tiene la capacidad de aprender perfectamente el comportamiento de una serie observada. El problema es que en la práctica una serie se compone de la suma de comportamientos fundamentados y de choques. Una buena red es entonces aquella que aísla los choques que son completamente aleatorios, mientras que aprende el componente que se repite. Por eso al escoger la red se debe tener en cuenta la calidad de los pronósticos del modelo dentro y fuera de muestra. Por pertenecer al tipo de modelos que se denominan de caja

negra, los parámetros estimados no tienen ninguna interpretación económica. Por esta razón la red no sirve para determinar elasticidades, ni entender patrones y generalmente se emplea cuando el objetivo del modelo es simplemente pronosticar de la mejor manera posible alguna serie.

En el trabajo se incluyeron las estimaciones de una red neuronal porque de cierta manera es el modelo no lineal más general, ya que como lo comenta Franses et. al. (1999) una red permite aproximarse a casi cualquier función no lineal sin necesidad de construir una parametrización específica y puede entenderse como una generalización del modelo de regresión suave (STR). Tkacz y Hu (1999), muestran como la red es capaz de modelar comportamientos atípicos como la presencia de *outliers* o cambios de nivel. De todas formas, en Franses et. al. (1999) se muestra que la red no es capaz de capturar la no linealidad tipo cambio de régimen de Markov (MRS), ni tampoco la no linealidad tipo GARCH³.

Una red neuronal en general se compone de dos partes, de una lineal que es simplemente un modelo autoregresivo (AR-X) con intercepto y de una parte no lineal:

$$y_t = \phi x_t + \sum_{j=1}^q B_j G(\gamma_j z_t) \quad ; \quad G(x) = \frac{1}{1 + e^{-x}} \quad ; \quad z_t \subseteq x_t .$$

Donde el primer término de y_t corresponde a la parte lineal del modelo y el segundo a la no lineal. En el primer término, ϕ es el vector de parámetros estimados mediante OLS del modelo AR-X y x_t es el vector que contiene el conjunto de información disponible para la parte lineal que contiene rezagos de la variable y_t además de rezagos de variables adicionales explicativas. En general se suelen normalizar las series al restarles su media y dividir las por la desviación estándar. Esta transformación se hace para que todas las variables queden con magnitudes comparables independientemente de las unidades y de la escala en la que se encuentren.

La parte no lineal contiene en primer lugar unas funciones de activación que se denotan por q . Estas son simplemente funciones logísticas $G(x)$ que se utilizan para modelar el paso de una variable de un estado a otro. Durante el proceso de selección de la red se debe elegir el número óptimo de q 's, pero por lo pronto vale la pena tener en cuenta que entre más haya mejor será el ajuste del modelo dentro de muestra y muy seguramente peor fuera de ella. En

³ Para una introducción más detallada a las redes puede consultarse Franses et. al. (1999) y algunas aplicaciones que se han hecho en Colombia como las de López et. al. (2002) y Arango et. al. (2003) para la inflación total y la demanda de efectivo respectivamente.

este artículo se trabajaron redes con un máximo de cuatro q 's, porque con cuatro el ajuste de la red ya es muy bueno dentro de muestra y aumentar el número de funciones de activación simplemente lleva a que la red se aprenda los errores aleatorios en la serie (fenómeno que se conoce como *overfitting*), que como ya se mencionó no es algo deseable. Los γ 's son los parámetros de las funciones de activación, que provienen inicialmente de una distribución uniforme aleatoria entre $[-1,1]$ y determinan la respuesta cuando se activa cada una de las funciones $G(x)$. Dependiendo de su valor pueden atenuar o amplificar la señal original de la serie. Idealmente deberían estimarse de manera directa, pero como la red es un modelo tan complejo los γ 's iniciales se escogen mediante un método indirecto de selección a partir de simulaciones. En general $\gamma_{i,j}$ corresponde al i -ésimo parámetro de la función de activación j -ésima. En este caso i depende del número de parámetros de z_t que es el conjunto de información de la parte no lineal del modelo y j del número de q 's. Adicionalmente se requiere de un γ adicional por cada q , que hace las veces del intercepto en la parte no lineal y que puede entenderse como el umbral que determina cuándo hay cambio de régimen. En este trabajo se asumió que z_t es un subconjunto de x_t , esto no necesariamente tendría que ser así pero se hace para que la identificación de la mejor red sea más sencilla. El número de elementos que contiene z_t , que obviamente tiene que ser menor o igual que el de x_t , se denota por p . De tal forma que cuando en la literatura sobre redes se menciona que una red tiene arquitectura (p,q) , simplemente se hace referencia al número de elementos de z_t y al número de funciones de activación que se utilizan.

Un detalle adicional que vale la pena mencionar es que en la parte lineal no importa el orden en que las variables entran al modelo, porque en últimas todas ellas entran a la estimación. Mientras que en la parte no lineal el orden de las variables dentro de z_t cambia considerablemente los resultados de la red. Para evitar introducir variables innecesarias al modelo en la parte no lineal, se desearía introducir primero aquellas variables que son más importantes para dicha estimación. Desafortunadamente no existen criterios de selección de modelo no lineales, por tal razón podría pensarse que utilizar los criterios lineales podría ser la mejor opción. Sin embargo, al realizar las evaluaciones de las redes obtenidas con esta aproximación se ha encontrado que su desempeño es inferior al que se obtiene al ordenar las variables primero de acuerdo a la intuición económica y luego por orden de rezagos.

Como los γ 's vienen de una selección aleatoria, se debe considerar un número grande de conjuntos de γ 's para cada una de las arquitecturas posibles. En este trabajo se hicieron cincuenta selecciones diferentes de γ 's, porque de nuevo se ha aprendido con la práctica que trabajar con más no reporta beneficios importantes en la elección final de la red y sí es computacionalmente muy exigente. Posteriormente cada uno de estos grupos de γ 's se multiplica por cada uno de los diferentes z_t 's con los que se trabaja. En realidad se multiplica cada uno de los γ_i 's por z_t de tal manera que se obtienen j vectores con estos productos. Luego se estiman unos B 's por mínimos cuadrados ordinarios (OLS) al regresar y_t contra los j vectores $(\gamma_i' z_t)$. Una vez se ha hecho todo esto, con el conjunto inicial de parámetros $\theta_{ini} = (\gamma, B, \phi)$ se hace un pronóstico preliminar de la red \hat{y}_t , que es la suma del pronóstico lineal $(\phi' x_t)$ y del pronóstico no lineal $\sum_{j=1}^q B_j G(\gamma_j z_t)$. Luego se diseña una función objetivo que es simplemente la suma de los errores al cuadrado del pronóstico preliminar más un término, que se conoce en la literatura como *weight decay*, que se encarga de penalizar la presencia de parámetros muy grandes en θ . El último paso es minimizar con respecto a (γ, B, ϕ) el valor de la función objetivo, con el procedimiento de optimización no lineal de Quasi-Newton. Dicho procedimiento simplemente se encarga de buscar un mínimo dentro del hiperplano que determinan los parámetros del vector θ , de tal manera que en el óptimo la función objetivo sea mínima y los gradientes del vector de parámetros resultante θ_{res} sean aproximadamente cero.

Con todo esto se obtiene un valor de la función objetivo optimizado para cada uno de los cincuenta juegos iniciales de parámetros, en cada una de las arquitecturas (p, q) . Obviamente para cada una de las arquitecturas se escoge el vector θ_{res} que produzca el menor valor de la función objetivo. El paso siguiente es entonces, decidir a partir de medidas de evaluación de pronóstico dentro y fuera de muestra, cuál de todas las arquitecturas es la mejor para pronosticar la serie en cuestión. Dentro de los criterios de evaluación de pronóstico que se utilizaron en este trabajo se encuentran el AIC, BIC, R2, RMSE, RMSPE, MAE, MAPE, SR y la UTHEIL, que se encuentran definidas en el Anexo 1. Para escoger la mejor arquitectura (p, q) a partir de estas medidas se puede utilizar una función objetivo, cuyo valor se desea minimizar, a partir de ponderaciones de las medidas en cada una de las arquitecturas. Una vez se define la arquitectura, se procede a determinar si hacer desconexiones de los

parámetros de z_t y x_t mejora los pronósticos. Es decir se mira qué pasa con las medidas de evaluación si el parámetro ϕ asociado a alguno de los componentes de x_t se fija en cero, o si los γ_i 's asociados a alguno de los componentes de z_t se vuelve cero o si los $(\phi, \gamma$'s) asociados a la misma variable en x_t y z_t son simultáneamente cero.

Después de todo finalmente se puede escoger la mejor red, que tiene una arquitectura (p,q) , posiblemente con alguna desconexión y con un vector de parámetros resultantes θ . En la Tabla 1 se encuentra resumido el proceso de estimación de una red. Es claro que la estimación de estos modelos es un proceso complejo en varias etapas. De ahí que se debe probar que su desempeño es muy superior al de un simple modelo lineal para que valga la pena utilizar este tipo de procedimientos para hacer pronósticos. Esto es precisamente lo que se hace en la séptima sección del documento.

Pasos para la Estimación de una Red Neuronal

Tabla 1

1. Se ajusta un modelo AR-X a la serie normalizada, utilizando criterios de selección como el R2-ajustado y el BIC, además de métodos de búsqueda tipo **stepwise**. En este paso se escogen las variables explicativas de la parte lineal y se estiman mediante OLS los parámetros ϕ .

2. Se selecciona de manera aleatoria cincuenta grupos de γ 's para cada una de las arquitecturas (p, q), a partir de una distribución uniforme [-1,1].

3. Se estiman los β 's mediante OLS, regresando y_t contra los j vectores ($\gamma_i' z_t$). De tal forma que el número de β 's debe ser igual al número de funciones de activación q.

4. Con el vector de parámetros iniciales $\theta_{ini} = (\gamma, \beta, \phi)$ se pronostica y_t como la suma del suma del pronóstico lineal ($\phi' x_t$) y del pronóstico no lineal $\sum_{i=1}^q B_j G(\gamma_j z_t)$

5. Se diseña una función objetivo que sea la suma de los errores de pronóstico al cuadrado, más un término que penalice por valores muy grandes de los parámetros de θ .

6. Se utiliza el método de optimización no lineal Quasi-Newton, para minimizar con respecto a (γ, β, ϕ) el valor de la función objetivo. De donde se obtiene el vector de parámetros optimizados $\theta_{res} = (\gamma, \beta, \phi)$. El gradiente de este vector debería ser casi cero, para que el punto sea efectivamente un mínimo. En este punto se escoge para cada una de las arquitecturas (p, q) el θ_{res} que minimice el valor de la función objetivo.

7. Con el θ_{res} escogido para cada una de las arquitecturas, se procede a hacer pronósticos dentro y fuera de muestra. Con base en criterios de evaluación de pronóstico como el BIC, R2, RMSE, RMSPE, MAE, MAPE y la UTHEEL, se elige cual de todas las arquitecturas es la mejor para pronosticar.

8. Para la arquitectura (p,q) que se eligió como la mejor en el paso anterior, se procede a hacer pruebas de desconexión de variables de la parte lineal y no lineal. La idea es fijar en cero el ϕ de una de las variables de la parte lineal, o los γ 's asociadas a alguna variable en la parte no lineal y evaluar si los pronósticos mejoran. Se deben tratar todas las posibles desconexiones lineales y no lineales, lo que da una idea de las variables más importantes del modelo.

Nota: El número de parámetros que componen al vector θ no siempre es el mismo, puesto que depende de la arquitectura (p, q) con la que se trabaje. En general el número de ϕ 's depende de cuántas variables explicativas haya en el modelo AR-X, el número de β 's depende de q y el número de γ 's es igual $q*(p+1)$.

5. Selección de las arquitecturas de las dos Redes y Prueba de No Linealidad

A partir de los dos modelos lineales que se presentaron en la tercera sección, se comenzó a hacer la estimación de las respectivas redes, de la forma que se describió en la sección anterior. El orden de las variables dentro de z_t es el mismo en el que se presentan los resultados del AR-X. En la Tabla 2 se muestran las evaluaciones de pronóstico dentro de muestra de la red para la inflación sin alimentos de los bienes transables. Como se calculan varias medidas, la mejor arquitectura según una de las medidas no necesariamente es la mejor de acuerdo a otra. Por esta razón se diseñó una función objetivo que se desea minimizar a partir de ponderaciones de las medidas de evaluación de pronóstico. En la Tabla 2 se reportan además de las medidas de evaluación, las ponderaciones que se utilizaron para calcular la función objetivo y su respectivo valor para cada una de las arquitecturas. Analizando los resultados queda claro, que el mejor ajuste dentro de muestra se logra con la red (7,4) y uno de los peores con la red (0,0), el modelo AR-X de la sección anterior. Esto simplemente confirma la idea del comienzo de la segunda sección, que la red puede lograr un ajuste casi perfecto cuando se utiliza una combinación de (p, q) suficientemente grande. Sin embargo llama la atención el buen ajuste que se logra con la red (4,2), porque como en todo modelo además del ajuste la parsimonia es muy importante.

En la Tabla 3 se presenta la evaluación fuera de muestra de esta misma red y las ponderaciones utilizadas para calcular la respectiva función objetivo. No pueden ser las mismas ponderaciones que se utilizaron para la evaluación dentro de muestra, puesto que cuando se hace evaluación fuera de muestra se utilizan algunas nuevas medidas. En especial son de vital importancia las U de Theil a diferentes horizontes de pronóstico. La idea cuando el horizonte es uno, es evaluar si la razón entre la raíz de los errores al cuadrado de un modelo particular y los de la predicción de una caminata aleatoria (el dato inmediatamente anterior) es menor que uno. Para horizontes más amplios, la idea es ver por ejemplo si el mejor pronóstico seis meses adelante es lo que acaba de ocurrir o lo que predice el modelo. En este caso queda

Tabla 2

P	Q	ITER	AIC	BIC	R2	RMSE	RMSPE	MAE	MAPE	SR	FO
0	0	50	-11.17	-11.13	0.73	0.0037	0.38	0.0028	0.25744	63.52	-0.624058143
1	1	50	-11.19	-11.15	0.74	0.0037	0.38	0.0028	0.25521	64.38	-0.628654666
1	2	50	-11.20	-11.12	0.75	0.0036	0.37	0.0027	0.24847	63.09	-0.625634854
1	3	50	-11.20	-11.06	0.75	0.0036	0.37	0.0027	0.24519	63.95	-0.626090659
1	4	50	-11.18	-11.00	0.75	0.0036	0.36	0.0027	0.24623	64.38	-0.62414087
2	1	50	-11.20	-11.14	0.74	0.0036	0.37	0.0027	0.25199	63.52	-0.626709687
2	2	50	-11.28	-11.16	0.77	0.0034	0.37	0.0027	0.24784	63.95	-0.629510299
2	3	50	-11.28	-11.10	0.78	0.0034	0.36	0.0026	0.23827	62.66	-0.62726624
2	4	50	-11.32	-11.08	0.79	0.0033	0.35	0.0025	0.23153	64.38	-0.631648704
3	1	50	-11.23	-11.16	0.75	0.0036	0.38	0.0026	0.24394	59.66	-0.617422302
3	2	50	-11.27	-11.12	0.77	0.0034	0.37	0.0026	0.24011	61.37	-0.622931105
3	3	50	-11.24	-11.02	0.78	0.0034	0.35	0.0025	0.23007	65.24	-0.63125832
3	4	50	-11.28	-10.99	0.79	0.0033	0.38	0.0024	0.23426	61.37	-0.61448872
4	1	50	-11.24	-11.15	0.76	0.0035	0.38	0.0026	0.24272	59.66	-0.618082167
4	2	50	-11.2879	-11.1107	0.7793	0.0034	0.3768	0.0025	0.2370	66.9528	-0.634276761
4	3	50	-11.35	-11.08	0.80	0.0032	0.33	0.0024	0.21627	64.38	-0.636907541
4	4	50	-11.51	-11.15	0.84	0.0029	0.33	0.0022	0.21851	65.67	-0.643331763
5	1	50	-11.26	-11.16	0.76	0.0035	0.37	0.0026	0.23818	64.38	-0.631713818
5	2	50	-11.28	-11.08	0.78	0.0033	0.36	0.0025	0.23112	65.24	-0.631665501
5	3	50	-11.37	-11.06	0.81	0.0031	0.34	0.0023	0.21877	63.95	-0.632291923
5	4	50	-11.45	-11.04	0.84	0.0029	0.33	0.0021	0.20213	65.67	-0.638880933
6	1	50	-11.40	-11.28	0.80	0.0032	0.39	0.0024	0.23559	57.94	-0.617668261
6	2	50	-11.43	-11.19	0.81	0.0031	0.40	0.0023	0.23320	61.80	-0.620948945
6	3	50	-11.49	-11.13	0.84	0.0029	0.40	0.0022	0.22823	67.81	-0.633495713
6	4	50	-11.50	-11.03	0.85	0.0028	0.36	0.0021	0.20526	66.09	-0.633138846
7	1	50	-11.45	-11.31	0.81	0.0031	0.40	0.0024	0.23351	60.94	-0.62451308
7	2	50	-11.50	-11.24	0.83	0.0029	0.35	0.0022	0.21607	66.52	-0.645338334
7	3	50	-11.57	-11.17	0.85	0.0027	0.35	0.0021	0.20957	66.95	-0.644947915
7	4	50	-11.61	-11.08	0.87	0.0026	0.31	0.0020	0.18995	68.67	-0.653196002

Inflación

BIC	0.05
RMSE	0.35
RMSPE	0.2
MAE	0.1
MAPE	0.05
SR	0.25

Fuente: Cálculos del autor.

Tabla 3

P	Q	ITER	RMSE	RMSPE	MAE	MAPE	SR	UTHEIL	UTHEIL3	UTHEIL6	FO
0	0	50	0.0031	1.5385	0.0028	0.9441	45.4545	1.0588	0.9161	0.4284	0.43988165
1	1	50	0.0035	0.9064	0.0030	0.7405	45.4545	0.6238	0.5397	0.2524	0.227823429
1	2	50	0.0034	1.2951	0.0030	0.8972	45.4545	0.8913	0.7711	0.3606	0.35919743
1	3	50	0.0034	1.5410	0.0030	0.9664	54.5455	1.0606	0.9176	0.4291	0.42328265
1	4	50	0.0034	1.4010	0.0031	0.9476	54.5455	0.9642	0.8342	0.3901	0.377086703
2	1	50	0.0033	1.0963	0.0029	0.8057	45.4545	0.7545	0.6527	0.3053	0.291616193
2	2	50	0.0032	1.4833	0.0027	0.8922	54.5455	1.0208	0.8832	0.4130	0.402140414
2	3	50	0.0035	1.4472	0.0031	0.9462	54.5455	0.9960	0.8617	0.4030	0.392110894
2	4	50	0.0066	1.8643	0.0046	1.2802	36.3636	1.2830	1.1100	0.5191	0.575347628
3	1	50	0.0030	1.5557	0.0024	0.8418	45.4545	1.0707	0.9263	0.4332	0.442352639
3	2	50	0.0038	1.3213	0.0034	0.9651	45.4545	0.9094	0.7868	0.3679	0.369920423
3	3	50	0.0033	1.2858	0.0031	0.8937	45.4545	0.8849	0.7656	0.3580	0.356082505
3	4	50	0.0033	1.2666	0.0027	0.7832	54.5455	0.8717	0.7542	0.3527	0.328297958
4	1	50	0.0030	1.5676	0.0024	0.8491	45.4545	1.0789	0.9334	0.4365	0.446457514
4	2	50	0.0034	0.9707	0.0028	0.7287	63.6364	0.6680	0.5780	0.2703	0.212037369
4	3	50	0.0034	1.5169	0.0029	0.9249	54.5455	1.0440	0.9032	0.4224	0.414165523
4	4	50	0.0065	3.0827	0.0052	1.9605	54.5455	2.1216	1.8355	0.8584	0.956548043
5	1	50	0.0031	1.3102	0.0026	0.8316	54.5455	0.9017	0.7801	0.3648	0.343883743
5	2	50	0.0034	1.9979	0.0028	1.0881	54.5455	1.3750	1.1896	0.5563	0.575836684
5	3	50	0.0032	1.3710	0.0025	0.7898	45.4545	0.9436	0.8164	0.3818	0.380655499
5	4	50	0.0033	1.5869	0.0028	0.9014	36.3636	1.0921	0.9449	0.4419	0.472601599
6	1	50	0.0033	1.8163	0.0025	0.9281	63.6364	1.2500	1.0815	0.5058	0.493619537
6	2	50	0.0037	1.6800	0.0029	0.9585	63.6364	1.1562	1.0003	0.4678	0.450217348
6	3	50	0.0037	2.0369	0.0029	1.0947	54.5455	1.4018	1.2128	0.5672	0.588835435
6	4	50	0.0042	2.2196	0.0034	1.1858	72.7273	1.5276	1.3216	0.6181	0.614909317
7	1	50	0.0033	1.8052	0.0024	0.9297	45.4545	1.2423	1.0748	0.5027	0.526388878
7	2	50	0.0028	1.8855	0.0023	0.9842	36.3636	1.2977	1.1227	0.5250	0.572278951
7	3	50	0.0044	1.3683	0.0030	0.8465	45.4545	0.9417	0.8147	0.3810	0.381793615
7	4	50	0.0042	1.4269	0.0036	0.9948	45.4545	0.9820	0.8496	0.3973	0.405346734

Inflacion

RMSE	0.25
RMSPE	0.1
MAE	0.07
MAPE	0.03
SR	0.2
UTHEIL1	0.15
UTHEIL3	0.12
UTHEIL6	0.08

Fuente: Cálculos del autor.

Tabla 4

Desconecta	ITER	AIC	BIC	R2	RMSE	RMSPE	MAE	MAPE	SR	FO
0	50	-11.29	-11.11	0.78	0.0034	0.377	0.00254	0.23704	66.95	-0.634
L1	50	-10.00	-9.82	0.20	0.0064	0.720	0.00554	0.51161	66.95	-0.486
L2	50	-11.11	-10.94	0.74	0.0037	0.503	0.00277	0.28252	73.39	-0.614
L3	50	-10.84	-10.66	0.65	0.0042	0.613	0.00333	0.34558	63.09	-0.549
L4	50	-10.47	-10.30	0.50	0.0051	0.754	0.00380	0.40057	59.23	-0.490
L5	50	-11.20	-11.03	0.76	0.0035	0.432	0.00263	0.25289	64.81	-0.613
L6	50	-11.23	-11.05	0.77	0.0035	0.378	0.00262	0.24257	65.24	-0.626
L7	50	-11.22	-11.04	0.76	0.0035	0.405	0.00252	0.23623	61.37	-0.611
NL1	50	-9.89	-9.72	0.11	0.0068	1.212	0.00553	0.66243	59.66	-0.356
NL2	50	-11.06	-10.89	0.72	0.0038	0.394	0.00275	0.25182	63.09	-0.609
NL3	50	-10.80	-10.63	0.64	0.0043	0.445	0.00332	0.30541	59.66	-0.574
NL4	50	-11.06	-10.88	0.72	0.0038	0.475	0.00282	0.27329	66.09	-0.599
L,NL,1	50	-9.95	-9.78	0.16	0.0066	0.719	0.00557	0.51790	59.66	-0.465
L,NL,2	50	-10.90	-10.72	0.67	0.0041	0.517	0.00304	0.29761	69.10	-0.589
L,NL,3	50	-10.94	-10.76	0.69	0.0040	0.507	0.00308	0.30426	61.37	-0.573
L,NL,4	50	-10.31	-10.14	0.42	0.0055	0.894	0.00395	0.43917	58.37	-0.450

Fuente: Cálculos del autor.

Tabla 5

Desconecta	ITER	RMSE	RMSPE	MAE	MAPE	SR	UTHEIL	UTHEIL3	UTHEIL6	FO
0	50	0.00343	0.9707	0.00280	0.7287	63.6364	0.6680	0.5780	0.2703	0.2120
L1	50	0.00967	2.5182	0.00914	2.2337	27.2727	1.7331	1.4994	0.7012	0.8364
L2	50	0.00351	2.1105	0.00305	1.1958	63.6364	1.4525	1.2566	0.5877	0.5976
L3	50	0.00432	3.0112	0.00363	1.6524	54.5455	2.0724	1.7930	0.8385	0.9234
L4	50	0.00687	4.3719	0.00616	2.7435	45.4545	3.0088	2.6031	1.2174	1.4186
L5	50	0.00391	0.9324	0.00349	0.8048	54.5455	0.6417	0.5552	0.2596	0.2202
L6	50	0.00352	1.4719	0.00278	0.8580	54.5455	1.0130	0.8764	0.4099	0.3975
L7	50	0.00396	1.3195	0.00288	0.8337	54.5455	0.9081	0.7857	0.3674	0.3472
NL1	50	0.00563	4.9659	0.00418	2.4011	45.4545	3.4176	2.9568	1.3828	1.6015
NL2	50	0.00342	1.0197	0.00273	0.7319	63.6364	0.7018	0.6072	0.2839	0.2281
NL3	50	0.00566	1.7955	0.00511	1.4019	45.4545	1.2357	1.0691	0.5000	0.5382
NL4	50	0.00420	1.8238	0.00367	1.1819	36.3636	1.2552	1.0859	0.5078	0.5585
L,NL,1	50	0.00610	1.9896	0.00476	1.5015	54.5455	1.3692	1.1846	0.5540	0.5863
L,NL,2	50	0.00352	2.2182	0.00298	1.2082	63.6364	1.5266	1.3208	0.6177	0.6331
L,NL,3	50	0.00354	2.2821	0.00286	1.1848	36.3636	1.5706	1.3588	0.6355	0.7078
L,NL,4	50	0.00588	4.0219	0.00516	2.3783	54.5455	2.7679	2.3947	1.1199	1.2751

Fuente: Cálculos del autor.

claro a partir de la Tabla 3 que la mejor red fuera de muestra es la (4,2) y como su desempeño también es muy bueno dentro de muestra se decidió que esta era la arquitectura más conveniente para esta serie. Finalmente, en la Tabla 4 y en la Tabla 5 se presentan las evaluaciones dentro y fuera de muestra de las desconexiones. En la primera columna de las Tablas se indica cuál es la variable que se está desconectando. Primero se desconectan una a una las variables del conjunto de información lineal, luego las del conjunto no lineal y finalmente se desconecta de manera simultánea la misma variable en la parte lineal y en la no lineal. Queda claro que no hay ninguna desconexión que mejore los resultados de las evaluaciones simultáneamente dentro y fuera de muestra. Por lo que se decidió que la red (4,2) sin ninguna desconexión es el mejor modelo para la inflación sin alimentos de los bienes transables.

De la misma forma se procedió a escoger la mejor red para la inflación de los bienes no transables. En la Tabla 6 y en la Tabla 7 se presentan las evaluaciones dentro y fuera de muestra de esta red. De la Tabla 6 no es claro identificar cuál es la mejor red dentro de muestra. Si bien la red (7,4) gana en algunas de las medidas, en general parece que la red (5,3) es muy buena al igual que la red (5,2). Sin embargo, al observar la Tabla 7 las redes (6,2), (6,3) y (6,4) parecen ser las mejores. La red (5,2) que había sido de las mejores dentro de muestra, también tiene un desempeño muy bueno fuera de muestra. Por estas razones se decidió que la mejor arquitectura para pronosticar inflación básica de los productos no transables es la (5,2). Por último, al igual que en el caso anterior, se procedió a evaluar la conveniencia de la desconexión de alguna de las variables explicativas, tanto en su parte lineal como en la no lineal. En la Tabla 8 y en la Tabla 9 se presentan respectivamente las evaluaciones dentro y fuera de muestra de las desconexiones. En la Tabla 8 no se puede ver con claridad alguna desconexión que mejore definitivamente la evaluación de pronóstico dentro de muestra. Sin embargo en la Tabla 9 se ve con claridad que desconectar la variable 8 de la parte lineal, que corresponde al rezago 15 de la brecha de producto, mejora considerablemente la evaluación fuera de muestra. Por lo tanto, la red para la inflación básica de productos no transables presenta una arquitectura (5,2) con una desconexión de la octava variable de la parte lineal.

Como se mencionó en la sección anterior, la estimación de una red es un proceso complejo que demanda mucho tiempo y cuidado. Por esta razón sólo debería emplearse para proyectar series que definitivamente son no lineales y donde asumir la

Tabla 6

P	Q	ITER	AIC	BIC	R2	RMSE	RMSPE	MAE	MAPE	SR	FO	Inflación
0	0	50	-10.85	-10.81	0.8293	0.0043389	1.499	0.00314	0.5691	70.815451	-0.387368019	
1	1	50	-10.88	-10.83	0.8333	0.0042881	1.489	0.00311	0.5749	71.673820	-0.392489094	BIC 0.05
1	2	50	-10.85	-10.76	0.8325	0.0042984	1.494	0.00314	0.5839	70.815451	-0.385214451	RMSE 0.35
1	3	50	-10.84	-10.70	0.8347	0.0042697	2.200	0.00315	0.6885	72.961373	-0.241218826	RMSPE 0.2
1	4	50	-10.88	-10.70	0.8456	0.0041261	1.435	0.00309	0.5486	70.815451	-0.395987709	MAE 0.1
2	1	50	-10.88	-10.82	0.8343	0.0042755	1.479	0.00311	0.5744	71.244635	-0.392515731	MAPE 0.05
2	2	50	-10.91	-10.79	0.8452	0.0041315	1.727	0.00307	0.6185	68.240343	-0.332031476	SR 0.25
2	3	50	-10.91	-10.73	0.8499	0.0040686	1.632	0.00302	0.5989	69.098712	-0.351151257	
2	4	50	-10.91	-10.68	0.8557	0.0039895	1.908	0.00300	0.6073	70.386266	-0.295983052	
3	1	50	-10.90	-10.82	0.8391	0.0042129	1.767	0.00313	0.5815	69.098712	-0.329547836	
3	2	50	-10.99	-10.84	0.8597	0.0039331	1.617	0.00294	0.5397	70.815451	-0.367203987	
3	3	50	-10.98	-10.76	0.8639	0.0038746	1.926	0.00289	0.5631	73.390558	-0.306245462	
3	4	50	-10.99	-10.70	0.8716	0.0037629	1.443	0.00288	0.5382	70.815451	-0.394816231	
4	1	50	-10.98	-10.89	0.8530	0.0040260	1.941	0.00287	0.5720	70.815451	-0.303103802	
4	2	50	-11.13	-10.95	0.8802	0.0036348	1.487	0.00269	0.4779	72.103004	-0.405126293	
4	3	50	-11.15	-10.88	0.8878	0.0035182	1.620	0.00265	0.5323	70.386266	-0.367861202	
4	4	50	-11.22	-10.86	0.9007	0.0033090	2.664	0.00245	0.6321	73.819742	-0.161888225	
5	1	50	-10.99	-10.89	0.8565	0.0039790	1.951	0.00278	0.5364	69.957082	-0.30072599	
5	2	50	-11.17	-10.96	0.8868	0.0035331	0.939	0.00270	0.4272	69.957082	-0.512414101	
5	3	50	-11.24	-10.93	0.9002	0.0033172	0.870	0.00257	0.3979	75.965665	-0.540927835	
5	4	50	-11.26	-10.85	0.9083	0.0031809	1.086	0.00236	0.4298	74.248927	-0.487923045	
6	1	50	-10.99	-10.87	0.8569	0.0039733	1.696	0.00281	0.5143	72.103004	-0.357106282	
6	2	50	-11.31	-11.07	0.9027	0.0032755	1.214	0.00244	0.4384	72.532189	-0.468719487	
6	3	50	-11.34	-10.99	0.9122	0.0031119	1.805	0.00233	0.4942	73.819742	-0.346808469	
6	4	50	-11.40	-10.93	0.9226	0.0029216	1.535	0.00214	0.4743	72.961373	-0.396709046	
7	1	50	-11.02	-10.89	0.8627	0.0038918	1.458	0.00283	0.4879	72.532189	-0.408018326	
7	2	50	-11.20	-10.93	0.8933	0.0034306	1.300	0.00251	0.4669	71.244635	-0.439915867	
7	3	50	-11.31	-10.91	0.9121	0.0031141	2.158	0.00237	0.5805	71.673820	-0.262970268	
7	4	50	-11.43	-10.90	0.9277	0.0028232	1.224	0.00223	0.4405	71.673820	-0.456226595	
8	1	50	-11.20	-11.06	0.8865	0.0035383	1.285	0.00266	0.4876	71.244635	-0.448028799	
8	2	50	-11.29	-10.99	0.9041	0.0032524	0.966	0.00248	0.3907	72.103004	-0.515590108	
8	3	50	-11.37	-10.93	0.9194	0.0029819	2.875	0.00231	0.5729	71.244635	-0.119774204	
8	4	50	-11.42	-10.83	0.9292	0.0027940	1.277	0.00214	0.4145	73.819742	-0.448627432	

Fuente: Cálculos del autor.

Tabla 7

P	Q	ITER	RMSE	RMSPE	MAE	MAPE	SR	UTHEIL1	UTHEIL3	UTHEIL6	FO
0	0	50	0.0025	1.1612	0.0021	0.8994	63.6364	0.5391	0.4398	0.2672	0.195514235
1	1	50	0.0026	1.2182	0.0022	0.9379	45.4545	0.5656	0.4614	0.2803	0.247532274
1	2	50	0.0027	1.2039	0.0023	0.9410	45.4545	0.5589	0.4560	0.2770	0.24402363
1	3	50	0.0028	1.4406	0.0023	1.0748	63.6364	0.6688	0.5457	0.3315	0.271831559
1	4	50	0.0030	1.4964	0.0026	1.1403	63.6364	0.6947	0.5668	0.3443	0.288050949
2	1	50	0.0026	1.2485	0.0022	0.9493	45.4545	0.5797	0.4729	0.2873	0.255587207
2	2	50	0.0031	1.6044	0.0025	1.2163	63.6364	0.7448	0.6077	0.3692	0.317779227
2	3	50	0.0025	1.2115	0.0022	0.9770	63.6364	0.5625	0.4589	0.2788	0.210626861
2	4	50	0.0030	1.1794	0.0027	1.0429	45.4545	0.5475	0.4467	0.2714	0.240954457
3	1	50	0.0025	0.9824	0.0021	0.8238	54.5455	0.4561	0.3721	0.2261	0.165976923
3	2	50	0.0023	0.9730	0.0020	0.8036	54.5455	0.4517	0.3685	0.2239	0.162930977
3	3	50	0.0023	0.8414	0.0020	0.7272	54.5455	0.3906	0.3187	0.1936	0.127197644
3	4	50	0.0024	1.0658	0.0021	0.8759	54.5455	0.4948	0.4037	0.2453	0.188724505
4	1	50	0.0019	1.7474	0.0015	0.9385	72.7273	0.8112	0.6619	0.4021	0.327232643
4	2	50	0.0019	1.9867	0.0016	1.0233	72.7273	0.9223	0.7525	0.4572	0.390595774
4	3	50	0.0016	0.7005	0.0009	0.4161	54.5455	0.3252	0.2653	0.1612	0.081819801
4	4	50	0.0020	0.9219	0.0015	0.7009	63.6364	0.4280	0.3492	0.2121	0.128583255
5	1	50	0.0018	1.8626	0.0014	0.9639	81.8182	0.8647	0.7055	0.4286	0.339080357
5	2	50	0.0023	0.8691	0.0017	0.6393	63.6364	0.4035	0.3292	0.2000	0.11342537
5	3	50	0.0048	5.5452	0.0029	2.4406	72.7273	2.5744	2.1005	1.2761	1.338058706
5	4	50	0.0029	2.9608	0.0023	1.5577	36.3636	1.3746	1.1215	0.6814	0.727158429
6	1	50	0.0018	2.0384	0.0014	0.9866	72.7273	0.9464	0.7721	0.4691	0.402609717
6	2	50	0.0015	0.5828	0.0011	0.4384	45.4545	0.2706	0.2208	0.1341	0.070777833
6	3	50	0.0016	0.6223	0.0013	0.5254	54.5455	0.2889	0.2357	0.1432	0.065257802
6	4	50	0.0013	0.8918	0.0011	0.6021	63.6364	0.4140	0.3378	0.2052	0.117786363
7	1	50	0.0018	1.0088	0.0015	0.7250	72.7273	0.4684	0.3821	0.2321	0.133165657
7	2	50	0.0020	1.3643	0.0015	0.7960	63.6364	0.6334	0.5168	0.3139	0.243826536
7	3	50	0.0026	1.4218	0.0023	1.0525	63.6364	0.6601	0.5386	0.3272	0.266355807
7	4	50	0.0017	2.5775	0.0013	1.1026	54.5455	1.1966	0.9763	0.5931	0.579371282
8	1	50	0.0018	1.3429	0.0014	0.7903	72.7273	0.6234	0.5087	0.3090	0.219991228
8	2	50	0.0051	2.0926	0.0030	1.2347	63.6364	0.9715	0.7927	0.4816	0.442932795
8	3	50	0.0023	0.9262	0.0019	0.7817	54.5455	0.4300	0.3508	0.2132	0.150408428
8	4	50	0.0020	2.0622	0.0015	0.9401	63.6364	0.9574	0.7811	0.4746	0.42549081

Inflacion	
RMSE	0.25
RMSPE	0.1
MAE	0.07
MAPE	0.03
SR	0.2
UTHEIL1	0.15
UTHEIL3	0.12
UTHEIL6	0.08

Fuente: Cálculos del autor.

Tabla 8

Desconecta	ITER	AIC	BIC	R2	RMSE	RMSPE	MAE	MAPE	SR	FO
0	50	-11.17	-10.96	0.89	0.0035	0.939	0.00270	0.42718	69.96	-0.5124141
L1	50	-10.71	-10.50	0.82	0.0045	2.722	0.00346	0.78355	69.96	-0.114
L2	50	-10.68	-10.47	0.81	0.0045	4.418	0.00354	0.98173	66.95	0.244
L3	50	-11.13	-10.92	0.88	0.0036	0.884	0.00274	0.40928	70.82	-0.524
L4	50	-11.16	-10.95	0.89	0.0036	0.846	0.00270	0.38884	69.10	-0.530
L5	50	-10.53	-10.33	0.79	0.0049	3.821	0.00362	1.01950	68.67	0.129
L6	50	-11.16	-10.96	0.89	0.0035	0.913	0.00272	0.41306	71.67	-0.522
L7	50	-11.05	-10.84	0.87	0.0038	1.871	0.00283	0.54827	66.09	-0.304
L8	50	-11.09	-10.88	0.88	0.0037	1.315	0.00281	0.50698	70.82	-0.431
NL1	50	-10.90	-10.69	0.85	0.0040	1.259	0.00293	0.45526	68.24	-0.429
NL2	50	-10.86	-10.65	0.85	0.0041	1.160	0.00289	0.43458	68.67	-0.449
NL3	50	-10.38	-10.18	0.75	0.0052	2.728	0.00379	0.86411	66.95	-0.085
NL4	50	-10.77	-10.56	0.83	0.0043	2.419	0.00328	0.67867	69.96	-0.183
NL5	50	-10.08	-9.88	0.66	0.0061	1.976	0.00389	0.63596	69.10	-0.237
L,NL,1	50	-10.38	-10.17	0.75	0.0052	3.130	0.00405	0.88442	68.24	-0.007
L,NL,2	50	-10.80	-10.59	0.84	0.0043	3.950	0.00325	0.88435	69.53	0.133
L,NL,3	50	-10.42	-10.22	0.76	0.0051	3.241	0.00366	0.94760	66.09	0.022
L,NL,4	50	-10.81	-10.61	0.84	0.0042	1.999	0.00319	0.62135	69.53	-0.272
L,NL,5	50	-9.38	-9.17	0.32	0.0087	4.708	0.00536	1.36939	62.23	0.399

Fuente: Cálculos del autor.

Tabla 9

Desconecta	ITER	RMSE	RMSPE	MAE	MAPE	SR	UTHEIL	UTHEIL3	UTHEIL6	FO
0	50	0.00233	0.8691	0.00171	0.6393	63.6364	0.4035	0.3292	0.2000	0.113
L1	50	0.00384	2.0527	0.00344	1.6331	72.7273	0.9530	0.7775	0.4724	0.426
L2	50	0.00288	2.0915	0.00248	1.2979	72.7273	0.9710	0.7922	0.4813	0.426
L3	50	0.00238	0.9425	0.00174	0.7103	72.7273	0.4376	0.3570	0.2169	0.116
L4	50	0.00231	0.8920	0.00171	0.6836	63.6364	0.4141	0.3379	0.2053	0.121
L5	50	0.00357	2.5379	0.00302	1.7129	54.5455	1.1782	0.9613	0.5840	0.588
L6	50	0.00241	0.9833	0.00180	0.7570	45.4545	0.4565	0.3725	0.2263	0.182
L7	50	0.00154	1.6515	0.00124	0.8261	54.5455	0.7667	0.6255	0.3800	0.336
L8	50	0.00159	0.5575	0.00110	0.3906	63.6364	0.2588	0.2112	0.1283	0.027
NL1	50	0.00192	1.0271	0.00153	0.7449	72.7273	0.4768	0.3890	0.2364	0.138
NL2	50	0.00188	0.9748	0.00157	0.7200	81.8182	0.4526	0.3693	0.2243	0.106
NL3	50	0.00333	2.5414	0.00274	1.6657	90.9091	1.1799	0.9626	0.5848	0.515
NL4	50	0.00388	1.7687	0.00309	1.2393	72.7273	0.8211	0.6700	0.4070	0.342
NL5	50	0.00278	3.6998	0.00213	1.7073	72.7273	1.7177	1.4014	0.8514	0.847
L,NL,1	50	0.00426	2.5231	0.00392	1.9081	81.8182	1.1714	0.9557	0.5806	0.536
L,NL,2	50	0.00226	1.3212	0.00184	0.9082	72.7273	0.6134	0.5004	0.3040	0.218
L,NL,3	50	0.00408	3.1532	0.00325	2.0260	81.8182	1.4639	1.1944	0.7256	0.700
L,NL,4	50	0.00371	1.9862	0.00311	1.3402	72.7273	0.9221	0.7523	0.4571	0.401
L,NL,5	50	0.00651	7.4531	0.00593	4.1809	45.4545	3.4602	2.8231	1.7151	1.930

Fuente: Cálculos del autor.

linealidad resulta muy costoso en términos de errores de pronóstico. Por eso se procedió a realizar la prueba de no linealidad de Lee et. al. (1993). Como la red es un modelo paramétrico no lineal, cuya función es aproximarse lo mejor posible al proceso generador de datos, Franses et. al. (1999) resaltan que no tiene sentido buscar la no linealidad tipo redes. Por esta razón el procedimiento adecuado es buscar una no linealidad general. Recordando que una red se denota por:

$$y_t = \phi x_t + \sum_{j=1}^q B_j G(\gamma_j z_t)$$

probar la hipótesis nula de linealidad equivale a decir que $H_0 = B_1 = \dots = B_q = 0$ cuando se conocen q y γ_j . Es decir que no hay ninguna ganancia en utilizar un componente no lineal para explicar la dinámica de la serie. El problema es que bajo la hipótesis nula, los parámetros γ_j no se conocen, por lo que se deben fijar de manera aleatoria a partir de alguna distribución. La prueba se realiza en tres pasos: el primero es correr un modelo AR (P)-X de donde se obtienen los residuos del modelo lineal u_t . En segundo lugar se toman valores para $\gamma_{i,j}$ donde $i = 0, \dots, p$ y $j = 1, \dots, q$ a partir de una distribución uniforme entre $[-1, 1]$ y a partir de ellos se estima el valor de las funciones de activación $G(\gamma_j' z_t)$. El último paso es calcular el estadístico de la prueba como el nR^2 de la regresión auxiliar: $u_t = \phi' x_t + \sum_{j=1}^q B_j G(\gamma_j z_t)$, que se distribuye χ^2 con q grados de libertad. Es evidente que el valor del estadístico de prueba depende de los parámetros $\gamma_{i,j}$ elegidos de manera aleatoria, luego el rechazo o no rechazo de la hipótesis nula de linealidad depende en cierta forma del azar. Por tal razón cada vez que se corre la prueba el resultado será diferente. Franses et. al. (1999) sugieren que la manera de proceder es entonces correr la prueba un número suficientemente grande de veces y ordenar los p-valores de las pruebas de menor a mayor. De tal forma que si al menos con el menor de ellos se alcanza a rechazar la hipótesis de linealidad, podría ser un indicio de que hay una posible no linealidad. En todo caso esta prueba no tiene una potencia adecuada como la mayoría de las pruebas de no linealidad y no es muy agradable el hecho de que cada vez que se corra, el menor p-valor sea diferente. Sin embargo, se presentan los resultados de las pruebas para las series de inflaciones con las que se trabajaron a manera de ilustración.

Para cada una de las series se trabajó con mil replicaciones de la prueba y a continuación se ordenaron de menor a mayor los mil p-valores obtenidos. Para cada una de las series se utilizó la misma arquitectura que se encontró como la mejor en la sección

anterior. Es decir para los transables se trabajó con una arquitectura (4,2) y para los no transables con una (5,2). El menor p-valor obtenido en las mil replicaciones de la prueba para la serie de transables fue de **0.0495** y para la serie de no transables fue de **0.0615**. Si bien es cierto que la prueba no captura con absoluta certeza la presencia de una no linealidad en las series todavía falta ver el desempeño relativo de la red en términos de error de pronóstico frente a los otros modelos, que se presenta en la séptima sección.

6. Los Mínimos Cuadrados Flexibles y su Estimación

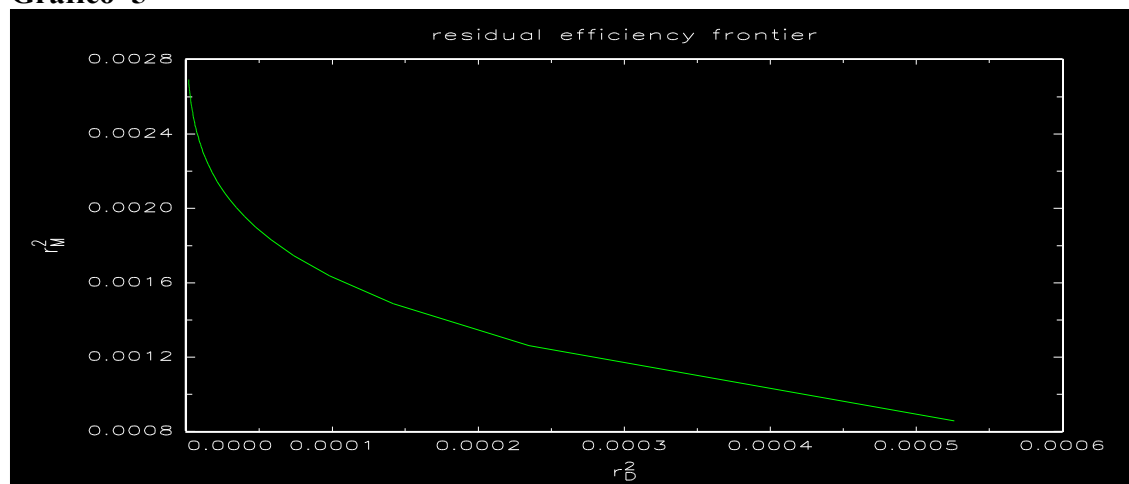
En la última década un gran número de países alcanzaron una inflación baja y estable. Como se mencionó anteriormente Colombia no fue la excepción, ya que su inflación total pasó de 32% en 1990 a niveles del 6% que se observan en la actualidad. Esto ha llevado a pensar que es posible que haya ocurrido un cambio estructural en el proceso generador de la inflación. De ser así, la aproximación de los mínimos cuadrados ordinarios (OLS) resulta inadecuada porque supone la constancia de los parámetros en el tiempo. La ventaja de los modelos con parámetros cambiantes es que consideran dicha inestabilidad en las relaciones (y por lo tanto la posibilidad de cambios estructurales), aunque no entran a explicar el cambio porque no poseen una ecuación dinámica. La metodología de los mínimos cuadrados flexibles, que fue propuesta por Kalaba y Tesfatsion (1989,1990), es en esencia, una regresión lineal con parámetros cambiantes en el tiempo. Se utilizan para capturar cambios suaves en los coeficientes durante el período de análisis, pero como ya se mencionó la ecuación dinámica es desconocida.

Para cada selección de parámetros $b = (b_1 \dots b_t)$ se pueden definir dos tipos de errores. El error de medida que se define como la diferencia entre el valor observado de una serie y el estimado $y - \hat{y} = y - x'b$ que corresponde al error tradicional de una regresión lineal. El error dinámico es la diferencia entre los coeficientes estimados en dos períodos del tiempo $b_{t+1} - b_t$. La solución de los mínimos cuadrados flexibles es la secuencia de los betas que minimiza la suma vectorial de los errores dinámicos y de medida al cuadrado, que determinan la frontera eficiente. Para hallar la solución se define una función de costos con ponderaciones cambiantes entre el costo de medida $r_M^2 = \sum_{t=1}^T [y_t - x'_t b_t]$ y el costo dinámico $r_D^2 = \sum_{t=1}^T [b_{t+1} - b_t]' [b_{t+1} + b_t]$. El objetivo de la estimación de la frontera eficiente es entonces minimizar la función de incompatibilidad o de pérdidas que se define por $C = \mu r_D^2 + r_M^2$. La

frontera eficiente es entonces la colección de puntos (r_D^2, r_M^2) que minimizan la suma ponderada de los errores al cuadrado, dependiendo del valor que tome μ . Dicha frontera debe ser convexa al origen e indica el costo en términos del error de medida que se debe pagar por disminuir el error dinámico. Lo interesante de esta parametrización es que cuando $\mu \rightarrow \infty$ la función de costos le otorga absoluta prioridad en la minimización al error dinámico, lo que implica que en la práctica se busca minimizar r_M^2 s.a. $r_D^2 = 0$ que es el problema que minimiza la solución de OLS.

En el Gráfico 5 se muestra una frontera eficiente que es la solución de FLS. La forma de la curva dice mucho sobre la utilidad de la solución de mínimos cuadrados flexibles para un problema en particular. Si los parámetros del modelo fueran en realidad constantes, en la vecindad del origen que corresponde a la solución de OLS (que es donde el error dinámico es cero) la pendiente de la curva debería ser poco empinada. Esto sugeriría que sólo se pueden obtener pequeñas disminuciones en el error de medida incluso si se permiten grandes errores dinámicos.

Gráfico 5



Fuente: Cálculos del autor.

Si en realidad el proceso generador de datos tiene parámetros cambiantes, entonces podrían obtenerse disminuciones importantes en el error de medida incluso con pequeños aumentos en el error dinámico. Es decir que si la curva es muy empinada cerca del origen la solución de FLS parece ser la adecuada, como se muestra en el Gráfico 5.

La parametrización de la solución no es para nada sencilla y no vale la pena entrar en detalles para el propósito de este documento. Basta decir que la solución al problema es única

y que en la práctica lo realmente importante es encontrar cuál es el μ adecuado para el horizonte de pronóstico que se desea.⁴ En este trabajo se buscó el μ adecuado para la serie de la inflación de transables y otra para los bienes no transables. De la discusión anterior es obvio que entre más pequeño sea el μ , mayor será la variabilidad que se permite en los parámetros. Por el contrario como ya se mencionó, cuando $\mu \rightarrow \infty$ los parámetros no cambian y se obtiene la solución de OLS. Al igual que en el caso de la red, se debe tener cuidado de no incurrir en un problema de *overfitting*. De tal manera que el μ óptimo para cada una de las series se obtiene a partir de la evaluación de los pronósticos fuera de muestra con las medidas tradicionales, utilizadas para escoger la arquitectura de las redes, que se definen en el anexo 1. Como en principio el μ podría tomar cualquier valor entre cero e infinito, en primer lugar se hicieron pronósticos con 25 μ 's diferentes a partir de una serie geométrica que comenzaba en 1 y se incrementaba en la raíz del número anterior multiplicado por cinco (es decir entre 1 y 545'915,034). Con base en una evaluación tradicional de pronósticos, se decidió que un rango entre el que se encontraba el μ para cada una de las series. Luego se hicieron nuevos pronósticos con 100 μ 's que se encontraran en el rango que se había establecido con los primeros 25 μ 's.

Todo esto para finalmente elegir un μ óptimo para cada una de las series, que de cierta manera indica el grado de variabilidad que se le permite a los parámetros en la estimación. Por esta razón, para series muy volátiles, el μ debería ser muy pequeño y, en general, si se busca un μ diferente para cada horizonte de pronóstico, se suele encontrar que entre mayor sea el horizonte de pronóstico mayor será el valor de μ . Esto es esperable ya que las relaciones económicas tienden a ser más fuerte en el largo plazo y aunque presentan desviaciones en el corto plazo. En este trabajo se eligió un único μ para cada una de la series y no un μ para cada horizonte de pronóstico para cada una de las series, para hacer más justa la comparación entre todos los modelos. De lo contrario también habría que elegir una red posiblemente con arquitecturas diferentes para cada una de las series para cada horizonte de pronóstico. Esto sería ideal para hacer pronósticos muy confiables, pero no es el objetivo de este trabajo y requeriría el trabajo de varias personas durante un período prolongado de tiempo.

⁴ Para mayores detalles ver Kalaba y Tesfatsion (1989,1990).

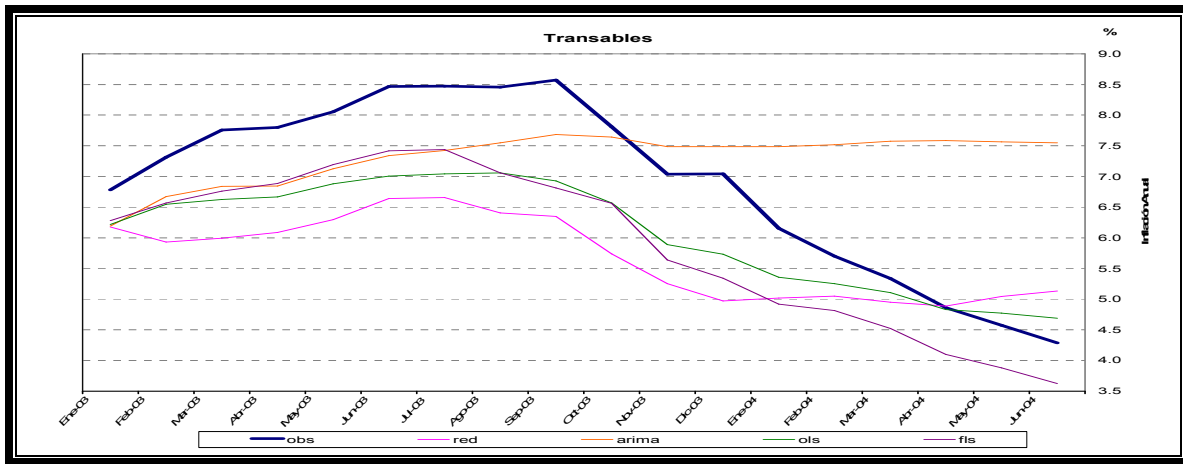
7. Comparación de los Pronósticos de los Modelo Lineales y No Lineales y las Funciones Impulso Respuesta

Luego de haber estimado una curva de Phillips por mínimos cuadrados ordinarios, con la red y con mínimos cuadrados flexibles para cada una de las series de inflación básica, se procedió a ajustar un modelo ARIMA tradicional y sin variables exógenas⁵. De esta manera se estimaron cuatro modelos para cada una de las series, dos que suponen linealidad en las estimaciones y dos que capturan una posible no linealidad de la curva en la inflación para Colombia. La evaluación de los pronósticos se hizo con dos tipos de ejercicios fuera de muestra. El primero de ellos es un análisis estático. Para esto se toman datos hasta diciembre de 2002 y se hacen 18 pronósticos (fuera de muestra) con cada uno de los cuatro modelos⁶. En el Gráfico 6 aparecen los pronósticos de cada uno de los cuatro modelos para la inflación anual de bienes transables, además del dato finalmente observado. En el Gráfico 7 aparecen los resultados para la serie de no transables. A primera vista parece que los modelos para la inflación de transables se comportan mejor que los de la inflación de no transables. Para cuantificar la bondad de ajuste en la Tabla 10 y en la Tabla 11 se reportan los criterios de evaluación de pronóstico para cada una de las series, respectivamente. Aquí es mucho más claro que las estimaciones no lineales hacen un mejor trabajo, especialmente porque a diferencia de los modelos lineales producen mejores pronósticos que los que haría una caminata aleatoria, según lo muestra la U de Theil. Sin embargo, en las demás medidas también puede verse la mejoría que se tiene con estas metodologías.

⁵ Los modelos ARIMA estimados para cada una de las series se encuentran en el anexo 2.

⁶ Se hicieron pronósticos hasta junio de 2004, como es tradicional en este tipo de análisis se tomaron los datos exógenos observados para evitar sumarle al error del modelo los errores en la predicción de las variables exógenas, aunque en la práctica al utilizar los modelos en tiempo real estas observaciones obviamente no se conocen.

Gráfico 6



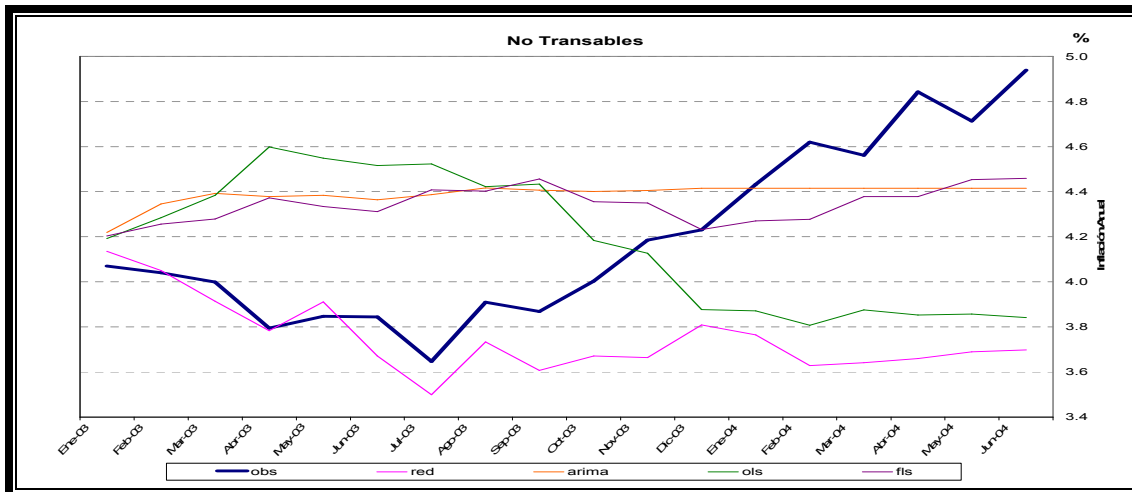
Fuente: Cálculos del autor.

Tabla 10

Evaluación de Pronósticos de la Inflación Anual de Transables						
Modelo	BIC	RMSE	RMSPE	MAE	MAPE	UTHEIL
ANN	-2.64	0.19	0.92	0.26	0.41	0.70
FLS	-2.43	0.22	1.00	0.37	0.52	0.76
OLS	-2.23	0.24	1.36	0.50	0.72	1.04
ARIMA	-1.69	0.31	2.09	0.63	1.00	1.59

Fuente: Cálculos del autor.

Gráfico 7



Fuente: Cálculos del autor.

Tabla 11

Evaluación de Pronósticos de la Inflación Anual de No Transables						
Modelo	BIC	RMSE	RMSPE	MAE	MAPE	UTHEIL
FLS	-1.11	0.42	0.10	0.37	0.09	0.69
ANN	-0.84	0.48	0.12	0.37	0.09	0.90
ARIMA	-0.30	0.62	0.13	0.46	0.12	1.32
OLS	-0.21	0.65	0.15	0.58	0.14	1.54

Fuente: Cálculos del autor.

Tabla 12

Evaluación del Rolling para la Inflación Anual de Transables								
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
1 mes	0.02	0.20	0.04	0.17	0.03	0.49	18	Rolling_FLS_Trans
1 mes	0.07	0.29	0.05	0.23	0.04	0.66	18	Rolling_Red_Trans
1 mes	0.04	0.29	0.04	0.20	0.03	0.71	18	Rolling_ARIMA_Trans
1 mes	0.47	0.75	0.13	0.62	0.11	1.80	18	Rolling_OLS_Trans
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
3 meses	0.13	0.35	0.05	0.25	0.04	0.35	18	Rolling_FLS_Trans
3 meses	0.24	0.53	0.12	0.42	0.09	0.52	18	Rolling_Red_Trans
3 meses	0.40	0.73	0.11	0.47	0.08	0.73	18	Rolling_ARIMA_Trans
3 meses	0.68	1.00	0.16	0.81	0.14	0.96	18	Rolling_OLS_Trans
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
6 meses	0.33	0.90	0.16	0.62	0.12	0.47	18	Rolling_FLS_Trans
6 meses	0.61	0.98	0.24	0.76	0.19	0.51	18	Rolling_Red_Trans
6 meses	1.12	1.62	0.32	1.20	0.25	0.84	18	Rolling_ARIMA_Trans
6 meses	1.28	1.83	0.32	1.52	0.28	0.88	18	Rolling_OLS_Trans
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
9 meses	1.01	1.38	0.35	1.12	0.28	0.53	18	Rolling_Red_Trans
9 meses	1.11	2.04	0.36	1.49	0.29	0.78	18	Rolling_FLS_Trans
9 meses	1.84	2.42	0.51	1.88	0.41	0.92	18	Rolling_ARIMA_Trans
9 meses	2.09	2.86	0.54	2.33	0.46	1.05	18	Rolling_OLS_Trans

Fuente: Cálculos del autor.

Tabla 13

Evaluación del Rolling para la Inflación Anual de No Transables								
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
1 mes	0.00	0.11	0.02	0.09	0.02	0.59	18	Rolling_OLS_NoTrans
1 mes	0.00	0.11	0.03	0.11	0.02	0.60	18	Rolling_FLS_NoTrans
1 mes	-0.02	0.14	0.04	0.15	0.03	0.95	18	Rolling_ARIMA_NoTrans
1 mes	-0.08	0.23	0.05	0.18	0.04	1.27	18	Rolling_Red_NoTrans
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
3 meses	-0.01	0.22	0.04	0.17	0.04	0.64	18	Rolling_OLS_NoTrans
3 meses	-0.01	0.22	0.04	0.17	0.04	0.65	18	Rolling_FLS_NoTrans
3 meses	-0.09	0.30	0.07	0.27	0.06	0.90	18	Rolling_ARIMA_NoTrans
3 meses	-0.28	0.56	0.12	0.44	0.09	1.67	18	Rolling_Red_NoTrans
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
6 meses	-0.02	0.47	0.10	0.41	0.09	0.76	18	Rolling_FLS_NoTrans
6 meses	0.00	0.53	0.11	0.47	0.10	0.86	18	Rolling_Red_NoTrans
6 meses	-0.14	0.60	0.13	0.51	0.11	0.97	18	Rolling_ARIMA_NoTrans
6 meses	-0.74	0.93	0.20	0.74	0.16	1.52	18	Rolling_OLS_NoTrans
a	ME	RMSE	RMSPE	MAE	MAPE	U-THEIL	No.	Modelo
9 meses	-0.27	0.63	0.13	0.53	0.11	0.80	18	Rolling_FLS_NoTrans
9 meses	-0.41	0.83	0.17	0.72	0.15	1.05	18	Rolling_Red_NoTrans
9 meses	-0.12	0.91	0.19	0.78	0.16	1.15	18	Rolling_OLS_NoTrans
9 meses	-0.90	1.17	0.24	1.00	0.20	1.48	18	Rolling_ARIMA_NoTrans

Fuente: Cálculos del autor.

El segundo ejercicio consiste en hacer un rolling, que es simplemente un pronóstico dinámico fuera de muestra. En este caso, partiendo en diciembre de 2002 se hacen nueve pronósticos, luego se agrega un nuevo dato y se hacen nueve pronósticos y así sucesivamente. El ejercicio concluye cuando se agregan datos hasta octubre de 2004. La evaluación se hace para cada uno de los horizontes de pronóstico (1 a 9) y se incluyen 18 observaciones para cada horizonte. Los resultados se presentan en la Tabla 12 y en la Tabla 13 y son muy parecidos al resultado que se obtuvo para el caso estático. En todo caso llama la atención que para la serie de los no transables, en el corto plazo la estimación por OLS es muy buena, aunque como era de esperarse en el mediano plazo pierde los primeros lugares de la evaluación con los modelos no lineales. Con este par de ejercicios se concluye entonces que definitivamente hay ganancias significativas al utilizar modelos que puedan incorporar la no linealidad de las Curvas de Phillips implícitas en los dos modelos.

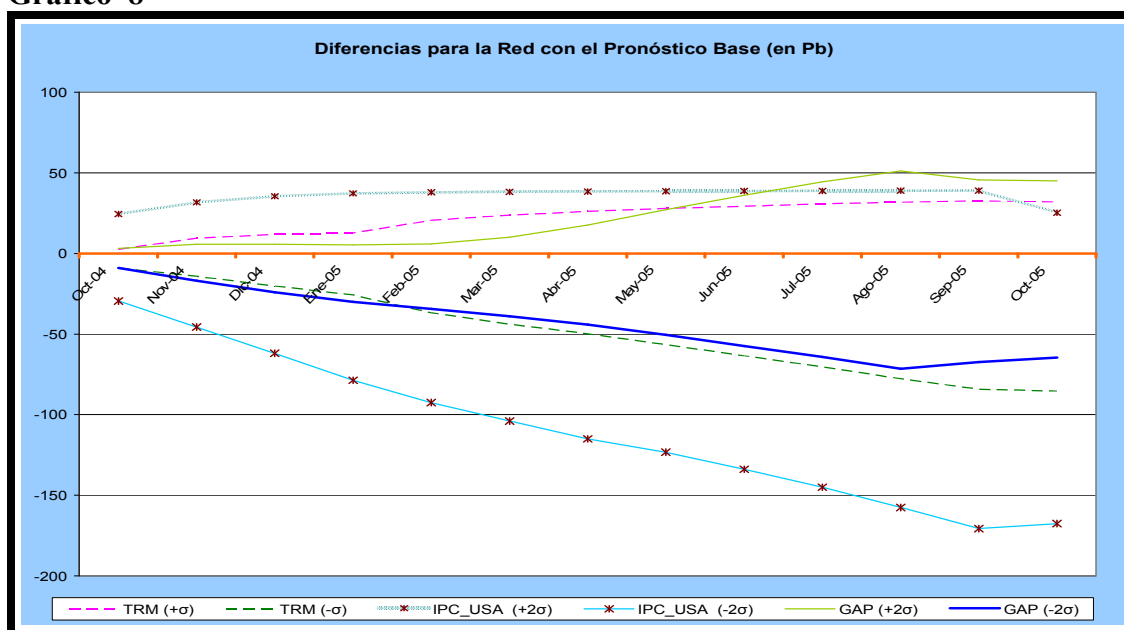
Finalmente se decidió utilizar estas estimaciones para medir es el impacto de cada una de las variables exógenas sobre la inflación y para intentar explicar el proceso de desinflación que ha vivido la economía colombiana durante los últimos trece años.⁷ De acuerdo a los datos con los que se cuenta entre mayo de 1982 hasta la fecha se calcularon las desviaciones estándar para cada una de las variables exógenas en los modelos (devaluación mensual, inflación mensual americana y brecha mensual) que se presentan en el Cuadro 2. De esta manera se estimó que una desviación estándar para la devaluación mensual equivale a una depreciación o apreciación de 2% en un mes, para la inflación americana a un cambio de 0.5% en un mes y para la brecha equivale a una variación de 7% en un mes.

Cuadro 2

	Media	Desviación
Trans	1.3%	0.7%
No Trans	1.2%	1.0%
Deval	1.4%	2%
USA	0.3%	0.5%
Gap	0%	7%

Fuente: Banco de la República, *Bureau of Economic Analysis* y cálculos del autor.

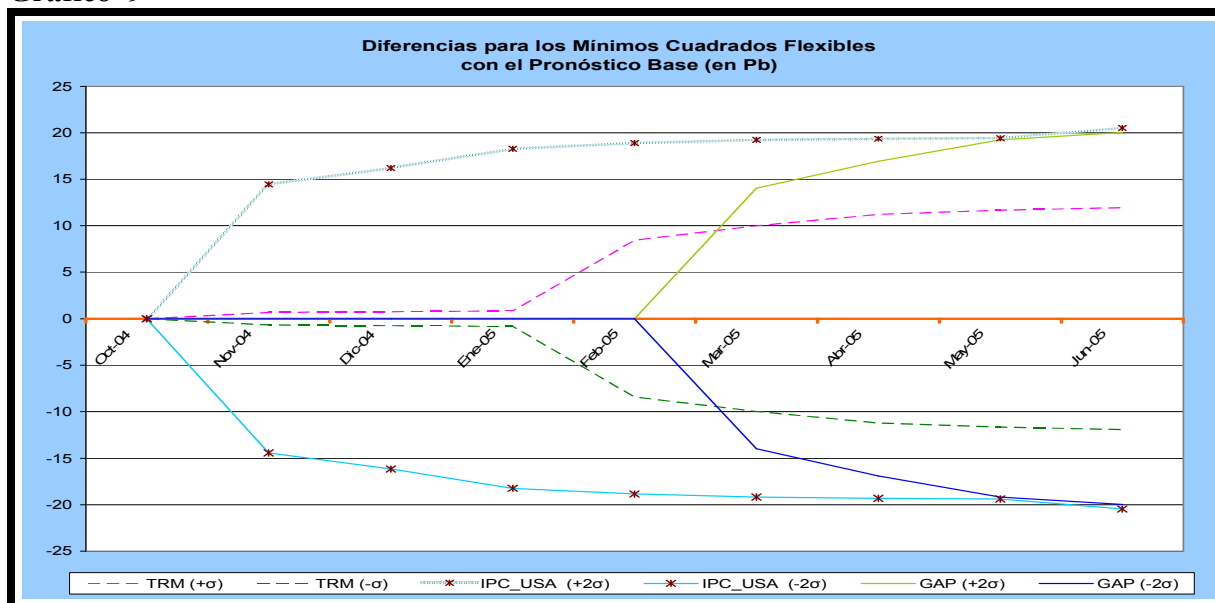
Gráfico 8



Fuente: Cálculos del autor.

⁷ Por obvias razones este ejercicio no puede hacerse con el ARIMA ya que no cuenta con variables exógenas que puedan tener choques.

Gráfico 9



Fuente: Cálculos del autor.

Con base en lo anterior se procedió a evaluar con cada uno de los modelos, el impacto de una ó dos desviaciones estándar en cada una de estas variables sobre la inflación de transables y de no transables.⁸ El ejercicio se hizo con datos hasta septiembre de 2004. Primero se hizo un pronóstico base para los siguientes doce meses en el que se suponía que la brecha estaba cerrada (producto igual al potencial), que no había depreciación o apreciación y que no había cambios en precios en la inflación americana.⁹ Luego se hacen los mismos doce pronósticos pero suponiendo que en octubre de 2004 ocurría un choque de una o de dos desviaciones estándar por una sola vez (en noviembre de 2004 ya no hay variables exógenas como en el caso base). Esto se hizo para la red, los mínimos cuadrados flexibles y para la estimación por OLS.

En el Gráfico 8 aparecen los resultados para la red de bienes transables. Lo primero que llama la atención es que dada la no linealidad del modelo, no es lo mismo un choque positivo que uno negativo.¹⁰ De acuerdo a esto, una depreciación de 2% en un mes aumenta en cerca de 30 pb la inflación de transables al siguiente año, mientras que una apreciación de la misma magnitud reduce la inflación en cerca de 80 pb en un año. Una inflación de 1% en un mes en Estados Unidos incrementa la inflación de transables en 40 pb luego de doce meses,

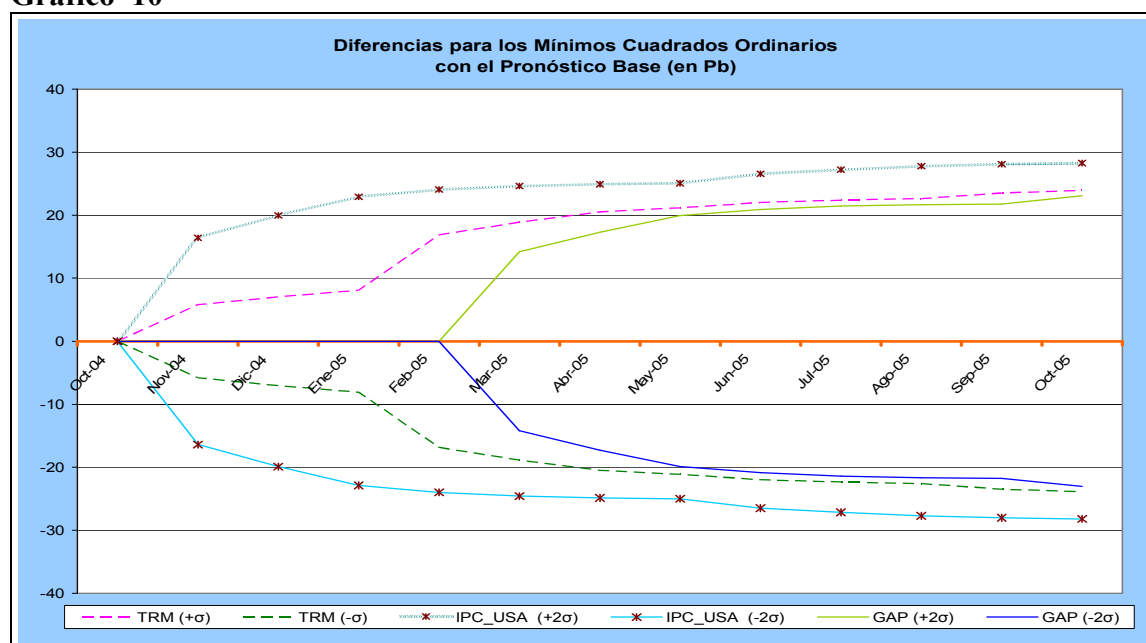
⁸ Para la brecha y la inflación de Estados Unidos se utilizaron dos desviaciones y para la devaluación sólo una porque el efecto sobre la inflación en pb es parecido para estos tres tipos de choques.

⁹ Es decir sin incluir a las variables exógenas.

¹⁰ También es diferente un choque grande de uno pequeño y los choques son diferentes dependiendo del nivel de la serie antes del choque y de su tendencia, sin embargo por limitaciones de espacio esto no se muestra. De ahí que la red se presente como el modelo no lineal más general que existe.

mientras que una deflación de la misma magnitud reduciría la inflación de transables en 170 pb en un año. Finalmente, si la brecha de producto se abre 14% en un mes, esto se traduce en 50 pb de inflación de transables al cabo de un año, pero si en cambio se cerrara en la misma magnitud esto reduciría la inflación de transables en 70 pb luego de un año. En el Gráfico 9 aparecen los resultados para el modelo de mínimos cuadrados flexibles de la serie de transables.

Gráfico 10



Fuente: Cálculos del autor.

Es claro que este modelo es lineal en cuanto a los choques, ya que uno positivo es igual a uno negativo, pero lo novedoso es que cambia el valor de los parámetros estimados en el tiempo. Una variación de 2% en el tipo de cambio se traduce en una variación de la inflación de transables de 12 pb nueve meses después. Un cambio de 1% mensual en la inflación americana, implica un cambio de 50 pb luego de nueve meses. Finalmente si la producción se ubica 14 puntos por encima o por debajo del potencial durante un mes, al cabo de nueve meses la inflación tendría un cambio de 20 pb al cabo de un año.

En el Gráfico 10 se encuentra el mismo ejercicio pero para la estimación por OLS (lineal) de la curva de transables. Una variación de 2% en el tipo de cambio se traduce en una variación de la inflación de transables de 24 pb doce meses después. Un cambio de 1% mensual en la inflación americana, implica un cambio de 28 pb luego de doce meses. Finalmente si la producción se ubica 14 puntos por encima o por debajo del potencial durante

un mes, al cabo de nueve meses la inflación tendría un cambio de 22 pb al cabo de un año. En la **Tabla 14** aparece el resumen de los choques. Es evidente que cada uno de los modelos cuenta algo diferente, pero es probable que cada uno de ellos capture una parte importante de la historia. Idealmente debería utilizarse algún método de combinación de los choques, que condicione el promedio dependiendo de la efectividad de los pronósticos. Sin embargo por simplicidad, se decidió tomar para el caso de los transables el promedio de las respuestas de la red y el modelo de mínimos cuadrados flexibles. Para el caso de los no transables se toma el promedio de todos los modelos, ya que como se mostró al parecer en el corto plazo la estimación por OLS hace un buen trabajo.

Tabla 14

Máxima Diferencia en Pb en la Inflación de Transables							Máximo Efecto
2%	-2%	1%	-1%	14%	-14%		
	TRM (+ σ)	TRM (- σ)	IPC_USA (+2 σ)	IPC_USA (-2 σ)	GAP (+2 σ)	GAP (-2 σ)	
ANN	30	-80	40	-170	50	-70	1 año
FLS	12	-12	50	-50	20	-20	9 meses
OLS	24	-24	28	-28	22	22	1 año
Promedio (1-2)	21	-46	45	-110	35	-45	11

Fuente: Cálculos del autor.

Con base en la información de la Tabla se intentará identificar los factores que explican la desinflación de los bienes transables de 35% anual en enero de 1991 a 3.6% anual en octubre de 2004. Un supuesto adicional es que los parámetros que se encontraron con datos hasta septiembre de 2004 han sido parecidos durante los últimos quince años. Aunque esto puede no ser muy plausible, es necesario para poder realizar el ejercicio.

A partir del Gráfico 3 puede verse la desinflación que se desea explicar y el nivel de las demás variables explicativas. A continuación se presentan los hechos que colaboraron mayormente para dicha desinflación (se excluyen los factores que hicieron repuntar temporalmente la inflación ya que lo que se desea explicar es la desinflación):

- Entre noviembre de 1990 y febrero de 1992 hubo una caída en la inflación anual de Estados Unidos de 3.5 p.p., al pasar de 6.3% a 2.8%. Esto equivale a una desinflación de 0.2% mensual. A partir de la Tabla 11 se puede decir que en promedio, esto aportó a la desinflación 22 pb durante quince meses, lo que daría una desinflación aproximada en la serie de transables de 330 pb.
- El tipo de cambio sufrió una fuerte apreciación, al pasar de una depreciación anual de 31% en enero de 1991 a una depreciación de 8% anual en diciembre de 1992. Esto equivale a una depreciación mensual de -1% durante dos años. De

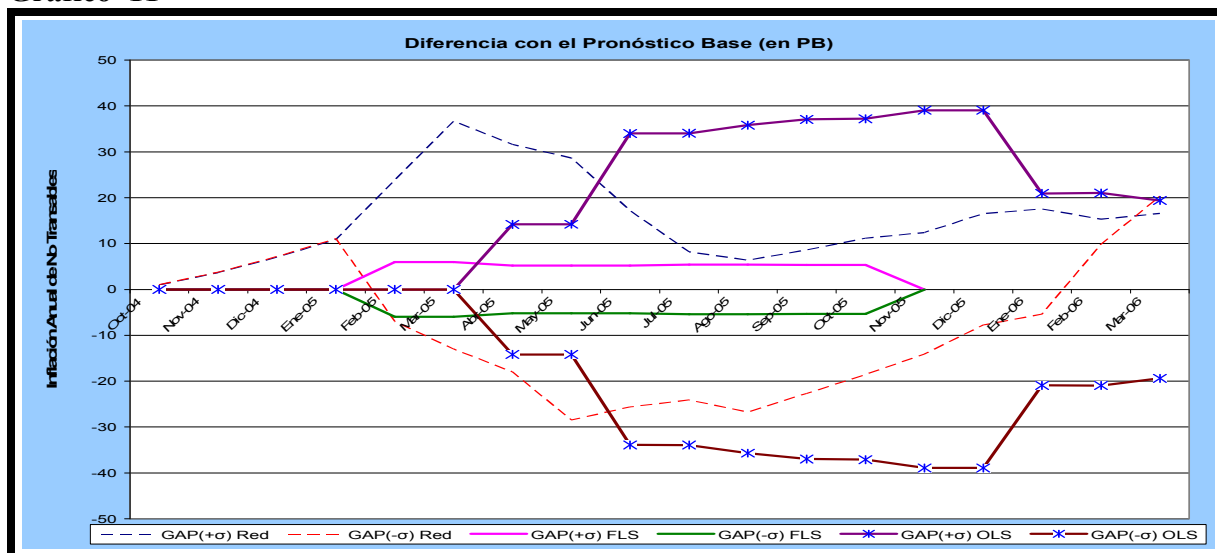
nuevo a partir de la Tabla 11, esto se traduce en una desinflación mensual de 23 pb lo que en dos años acumula 552 pb.

- El tipo de cambio sufrió otra fuerte apreciación al pasar de una depreciación anual de 26% en mayo de 2000 a 3% en junio de 2002. Esto es una depreciación de -1% mensual durante dos años, que implica otra desinflación de 552 pb.
- La última apreciación fuerte del tipo de cambio ocurrió entre julio de 2003 y septiembre de 2004, al pasar de 22% a -1% anual en septiembre de 2004. Esto es una depreciación de -1.6% durante catorce meses, que se traduce en una desinflación mensual de 37 pb y en total de 515 pb.
- Finalmente la variación más fuerte y prolongada en la brecha se dio entre junio de 1998 cuando la brecha era cero y noviembre de 1999 cuando la brecha llegó a ser -8%. Esta recesión implicó una caída en el producto en cerca de ocho puntos del PIB en 17 meses ó -0.8% mensual. Este hecho, de acuerdo con la Tabla 11, se traduce en una desinflación en la serie de transables de tan sólo 3 pb por mes y en total de 44pb.
- Todo lo anterior explica una caída en la inflación de bienes transables de cerca de 20 pb Como en total la caída fue de 31 pb, esto quiere decir que con el modelo se puede explicar cerca de dos terceras partes de la desinflación de esta serie.

El mismo ejercicio se hizo para la inflación básica de los bienes no transables, con resultados diferentes al del caso de los bienes transables. Los resultados se presentan en el Gráfico 11. Para facilitar la comparación con el caso anterior, se decidió utilizar el mismo tipo de choque pero sólo de una desviación estándar (7 puntos del PIB por encima o por debajo en octubre de 2004 y luego la brecha cerrada de ahí en adelante). De nuevo, queda en evidencia la no linealidad de la red, ya que el choque positivo tiene su mayor efecto seis meses después del choque al aumentar la inflación de los no transables en 37pb. En cambio, cuando el choque es negativo el mayor efecto se presenta ocho meses después con una caída en la inflación de los no transables de 28 pb. De acuerdo al modelo de los mínimos cuadrados flexibles, el mayor efecto aparece cinco meses después del choque y sólo afecta la inflación en cinco pb. Finalmente de acuerdo al modelo estimado con mínimos cuadrados tradicionales el mayor efecto a parece catorce meses después del choque con una variación en la inflación de 39 pb. Una última diferencia que queda en evidencia a partir de la Gráfica es que de acuerdo a la red

el choque se demora dieciocho meses en desaparecer, de acuerdo a los mínimos cuadrados flexibles el choque desaparece al cabo de trece meses y de acuerdo al modelo de OLS el choque tarda más de año y medio en desaparecer. Los resultados aparecen resumidos en la Tabla 15.

Gráfico 11



Fuente: Cálculos del autor.

Con base en esta información se hizo el mismo ejercicio de descomposición de la historia de la desinflación durante los noventa para la serie de los no transables. En el Gráfico 4 puede verse la desinflación que se desea explicar y el nivel de la brecha en el mismo período de tiempo. A continuación se presentan los períodos en que la brecha colaboró mayormente para la desinflación de los bienes no transables (una vez más se excluyen los factores que hicieron repuntar temporalmente la inflación):

- Entre abril de 1996 y marzo de 1997 la brecha pasó de 1.5% a -3.4%. Esto equivale a una caída de 0.4% mensual durante doce meses. Con un simple cálculo a partir de la Tabla 15 puede decirse que esto se traduce aproximadamente en una desinflación de 2 pb mensuales y en total en 24 pb.
- Entre junio de 1998 y noviembre de 1999 la brecha pasó de 6.5% a -8%, lo que equivale a una caída de 0.85% durante diecisiete meses. Esto implica de acuerdo a los modelos una desinflación mensual de 3 pb y en total de aproximadamente 60 pb.

- Finalmente entre mayo de 2001 y noviembre de 2002, la brecha pasó de 3.3% a -1%. Esto es una caída de 0.24% durante dieciocho meses, que finalmente se traduce en una desinflación de 1 pb por mes y en total de 18 pb.
- Todo lo anterior explica una desinflación de 100 pb cuando el total fue de 2750 pb entre noviembre de 1993 (31%) y marzo de 2001 (3.5%). Es decir que con los modelos se puede explicar solamente el 4% de la desinflación.

Tabla 15

Máxima Diferencia en Pb en la Inflación de No Transables					
		7%	-7%		
		GAP (+σ)	GAP (-σ)	Máximo Efecto	Duración del Efecto
ANN		37	-28	7* meses	18 meses
FLS		5	-5	5 meses	13 meses
OLS		39	-39	14 meses	> 18 meses
Promedio		27	-24	9 meses	19** meses
* Para la red se promediaron los resultados del choque positivo y negativo					
** Para el promedio se asumió que el choque de OLS dura 24 meses					

Fuente: Cálculos del autor.

Llama mucho la atención que aparentemente, los cambios en la brecha no tienen un gran impacto sobre la inflación de no transables como se esperaría a partir de la teoría de la curva de Phillips que se presentó anteriormente. Puede ser que la brecha que se utilizó en este documento no sea la más adecuada para este ejercicio, porque hablar de una producción siete puntos por encima o por debajo del potencial no es algo que parezca muy plausible. Pero para la serie de brechas mensuales del IPR esto es algo perfectamente normal (es su desviación estándar), de ahí que la relación entre esta brecha y la inflación parece no ser tan fuerte.¹¹ Desde este punto de vista el modelo fracasó en el objetivo de contar una historia coherente para el proceso de desinflación en la serie de los no transables que se vio entre 1993 y 2001. Se debe concluir, por lo tanto, que al parecer hay otras variables que pudieron influir en el comportamiento de la inflación de no transables. El candidato más obvio son las expectativas de inflación, las cuales no son consideradas en este trabajo.

8. Conclusiones

En este trabajo se evaluó la ganancia en términos de error de pronóstico de considerar un modelo que pudiera capturar la posible no linealidad de la curva de Phillips. Existen varias teorías que sugieren una posible no linealidad en la curva, pero en la práctica no es muy claro

¹¹ En todo caso, ejercicios similares realizados en el Banco de la República con un modelo más estructural (Modelo de Mecanismos de Transmisión), y con brecha del producto estimada a partir del PIB trimestral también muestran un impacto relativamente bajo de la brecha del producto sobre la inflación sin alimentos.

cuál es el tipo de no linealidad que se presenta. Por esto se utilizaron dos modelos que no impusieran de antemano ningún tipo de relación a las estimaciones. El objetivo fue entonces el de ajustar algunos modelos para pronosticar la inflación básica de productos transables y de no transables. Para cada una de las series se estimó una curva de Phillips con mínimos cuadrados ordinarios (OLS), con mínimos cuadrados flexibles (FLS) y con una red neuronal (ANN), además de un modelo ARIMA. Una vez estimados los modelos, se utilizaron las redes para realizar la prueba de no linealidad de Lee et. al. (1993) para corroborar si en verdad los datos colombianos presentan este tipo de comportamientos y debería siempre utilizarse modelos de pronóstico que capturen este hecho. Los resultados de la prueba no fueron muy concluyentes a favor de la hipótesis de no linealidad en contra de lo que se esperaba. Luego se procedió a evaluar qué tan buenos eran los pronósticos de cada uno de los cuatro modelos para cada una de las series. Los resultados en términos de error de pronóstico resultaron ampliamente favorables a los modelos no lineales. En particular, se encontró que los pronósticos de las redes presentaban la mitad del error de pronóstico que los de los modelos ARIMA, y que las U de Theil eran menores. Por tal razón, no se toman en cuenta los resultados de la prueba de no linealidad y se mantiene la hipótesis de que las series sí presentan no linealidades, por lo que se considera que las redes y los mínimos cuadrados flexibles son mejores modelos de pronóstico que los típicos modelos lineales.

Finalmente se intentó utilizar estos modelos, que son bastante buenos para hacer pronósticos, para explicar el proceso desinflacionario que vivió la economía colombiana en los últimos doce años. Los modelos para la inflación de transables permitieron explicar dos terceras partes de la desinflación observada, pero los modelos de la inflación de no transables fracasaron completamente. Para futuros trabajos sería interesante ampliar el horizonte de evaluación de pronóstico a más de un año, explorar los resultados de otras pruebas de no linealidad que tengan una potencia más adecuada, y finalmente intentar trabajar con otras brechas que sean mucho menos estacionales y que permitan tener mayor intuición económica. Aunque estos modelos no fueron adecuados para explicar la desinflación en Colombia, sí son muy buenos para pronosticar la inflación, que era el otro objetivo de este trabajo, y por esa misma razón contribuyen a mejorar la calidad de las decisiones de la política monetaria. En efecto, los modelos de pronóstico de inflación de bienes transables y no transables apenas se comenzaron a hacer el año pasado en el Banco de la República y es de esperar que, al igual que ocurrió con los otros modelos de pronóstico (inflación total, sin alimentos y de alimentos), con el tiempo se irán refinando cada vez más.

Bibliografía

Arango, C., J. N. Hernández, E. López. y M. Misas. (2003). “La Demanda de Efectivo en Colombia: Una Caja Negra a la Luz de las Redes Neuronales”. Borradores de Economía No. 268. Banco de la República.

Balakrishnan, R. y J. López-Salido. (2002). “Understanding UK inflation: the role of openness”. Working Paper No. 164, Bank of England.

Ball, L., N. Mankiw. and D. Romer. (1988). “The New Keynesian Economics and the Output, Inflation Trade-off”. *Brookings Papers on Economic Activity*, 1: 1-65.

Calvo, G. (1983). “Staggered Prices in a Utility Maximizing Framework”. *Journal of Monetary Economics*, 12: 393-398.

Clark, P., D. Laxton and D. Rose. (1996). “Asymmetry in the US output inflation nexus”. IMF Staff Papers, 43: 216-251.

Dupasquier, C. and N. Ricketts. (1998). “Nonlinearities in the output-inflation relationship: some empirical results for Canada”. Working Paper 98-14. Bank of Canada.

Eisner, R. (1997). “New View of the NAIRU”. In *Improving the Global Economy: Keynesian and the Growth in Output and Employment*, Edward Elgar Publishing, UK.

Eliasson, A. (2001). “Is the Short-run Phillips Curve Nonlinear? Empirical Evidence for Australia, Sweden and the United States”. Sveriges Riksbank Working Paper Series, 124.

Filardo, A. J. (1998). “New Evidence on the Output Cost of Fighting Inflation”. *Economic Review*, third quarter: 33-61.

Franses, P. H. y D. Van Dijk. (2000). *Non-linear Time Series Models in Empirical Finance*, Cambridge University Press.

Friedman, M. (1968). “The Role of Monetary Policy”. *American Economic Review*, 58: 1-17.

Galí, J. y M. Gertler. (1999). “Inflation Dynamics: A Structural Econometric Analysis”. *Journal of Monetary Economics*, 44: 195-222.

Gómez, J. y J. M. Julio. (2000). “An estimation of the nonlinear Phillips curve in Colombia”. Borradores de Economía No. 160. Banco de la República.

- Gordon, R. (1997). "The time-varying NAIRU and its implications for economic policy". *Journal of Economic Perspectives*, 11, 11-32.
- Kalaba, R. y L. Tesfatsion. (1989). "Time-Varying Linear Regression via Flexible Least Squares", *Computers Math Applied*, Vol. 17, No. 89, pg. 1215-1244.
- Ibid. (1990). "Flexible Least Squares for Approximately Linear Systems", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 20, No.5, pg.978-989.
- Lee, T, H. White and C. Granger. (1993). "Testing for Neglected Nonlinearity in Time Series Models". *Journal of Econometrics*, 56: 269- 290.
- López, E. y Misas, M. (1999). "Un examen empírico de la curva de Phillips en Colombia". Borradores de Economía No. 117. Banco de la República.
- López, E., M. Misas y P. Querubín. (1999). "La Inflación en Colombia: una aproximación desde las redes neuronales". Borradores de Economía No. 199. Banco de la República.
- Melo, L. F. y M. Misas. (2004). "Modelos Estructurales de Inflación en Colombia: Estimación a través de Mínimos Cuadrados Flexibles". Borradores de Economía No. 283. Banco de la República.
- Nigrinis, M. (2003). "¿Es lineal la Curva de Phillips en Colombia?". Borradores de Economía No. 282. Banco de la República.
- Phillips, A. W. (1958). "The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957". *Economica* 25: 283-299.
- Shapiro, C. and J. Stiglitz. (1984). "Equilibrium Unemployment as a Worker Discipline Device". *American Economic Review*, 74: 433-444.
- Taylor, J. (1980). "Aggregate Dynamics and Staggered Contracts". *Journal of Political Economy*, 88:1-23.
- Tkacz, G. y S. Hu. (1999). "Forecasting GDP Growth Using Artificial Neural Networks". Working Paper 99-3, Bank of Canada.
- Uribe, J., Vargas, H. y Gómez, J. "The Implementation of Inflation Targeting in Colombia". Borradores de Economía No. 202. Banco de la República.

Anexo 1

Medidas de Evaluación de Pronóstico

$$AIC(k) = \ln(\hat{\sigma}^2) + \frac{2k}{n}$$

$$BIC(k) = \ln(\hat{\sigma}^2) + \frac{k \ln(n)}{n}$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^n \hat{\varepsilon}_t^2}{n}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2}$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left(\frac{\hat{y}_t - y_t}{y_t} \right)^2}$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |\hat{y}_t - y_t|$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t - y_t}{y_t} \right|$$

$$U - THEIL = \sqrt{\frac{\sum_{t=1}^n \hat{\varepsilon}_t^2}{(y_t - y_{t-1})}}$$

$$SRP = \frac{1}{n} \sum_{t=1}^n I_t[\{(y_t - y_{t-1}) \cdot (\hat{y}_t - \hat{y}_{t-1})\} > 0]$$

$$SRN = \frac{1}{n} \sum_{t=1}^n I_t[\{(y_t - y_{t-1}) \cdot (\hat{y}_t - \hat{y}_{t-1})\} < 0]$$

$$SR = SRP + SRN$$

donde:

n es el número de observaciones consideradas

$\hat{\varepsilon}_t$ residuales obtenidos a partir de la red neuronal

k número de parámetros de la red neuronal

\hat{y}_t valor estimado por la red neuronal

Anexo 2

Modelos ARIMA

Transables

Parametro	Valor	P_value
AR(1)	0.37726	<.0001
AR(2)	0.22853	0.0005
AR(6)	0.22669	<.0001
MA(12)	0.48099	<.0001
d37	0.01565	<.0001
d71	0.0052679	0.0502
d90	0.0046503	0.08
d191	-0.01125	<.0001

No Transables

Parametro	Valor	P_value
MA(9)	-0.18356	0.006
MA(12)	0.16347	0.0159
AR(1)	0.26652	<.0001
AR(2)	0.23877	0.0003
AR(5)	0.13841	0.0276
d37	0.01799	<.0001
d72	-0.0095818	0.0011
d191	-0.01156	<.0001

Definición de las Dummies

d37	Abr-86
d70	Ene-89
d71	Feb-89
d72	Mar-89
d90	Feb-90
d191	Feb-99