**Documento de trabajo 1225: Metodología para la predicción de los precios minoristas de los alimentos localmente disponibles en las principales ciudades de Colombia**

Pontificia Universidad Javeriana, Cali

14 de diciembre de 2025

1. **Presentación**
2. **Datos**
   1. **Insumo del IPC (DANE)**

La primera fuente corresponde a los precios minoristas reportados por el DANE, utilizados como insumo para el cálculo del Índice de Precios al Consumidor (IPC) (DANE, 2024). El periodo de análisis abarca desde enero de 1999 hasta marzo de 2018 y cubre las 13 principales ciudades del país:

* Bogotá D.C.
* Medellín A.M.
* Cali A.M.
* Barranquilla A.M.
* Bucaramanga A.M.
* Manizales A.M.
* Pereira A.M.
* Cúcuta A.M.
* Pasto
* Ibagué
* Montería
* Cartagena
* Villavicencio

Una característica relevante de estos datos es su estructura de clasificación, basada en la canasta de seguimiento del IPC 2008. Esta clasificación contempla distintos niveles de agregación: Grupo, Subgrupo, Clase, Gasto Básico y Artículo. A continuación, se presenta la estructura de esta clasificación:

|  |  |  |  |
| --- | --- | --- | --- |
| **Subclase** | **IPC** | **Gasto\_basico** | **DANE** |
| 011101 | Arroz | 11101 | Arroz |
| 011102 | Pastas alimenticias | 11103 | Pastas alimenticias |
| 011103 | Avena y sus derivados | 11105 | Otros cereales |
| 011105 | Trigo y sus derivados | 11102 | Harina de maiz y otras harinas |
| 11105 | Otros cereales |
| 011106 | Maíz y sus derivados | 11102 | Harina de maiz y otras harinas |
| 11105 | Otros cereales |
| 011108 | Pan | 11201 | Pan |
| 011198 | Otras harinas, cereales y almidones | 11102 | Harina de maiz y otras harinas |
| 11104 | Cereales preparados |
| 11202 | Otros productos de panadería |
| 011199 | Otros productos de panadería | 11202 | Otros productos de panadería |
| 011201 | Carne de res y derivados | 15101 | Res |
| 011202 | Carne de Cerdo y Derivados | 15102 | Cerdo |

A continuación, se muestra la estructura de la base de datos de precios proporcionada por el DANE:

****

* 1. **Datos de SIPSA**

La información sobre los precios mayoristas corresponde a los datos del Sistema de Información de Precios y Abastecimiento del Sector Agropecuario (SIPSA), publicados por el DANE. El SIPSA no sólo informa, con frecuencia diaria, sobre los precios mayoristas de los productos agroalimentarios que se comercializan en el país; sino que, además, proporciona información, con frecuencia quincenal, sobre el nivel de abastecimiento de los alimentos en las ciudades. En este estudio se utilizarán datos de SIPSA para las tres principales ciudades de Colombia (Cali, Bogotá y Medellín) durante el período 2013:1 – 2024:1, con frecuencia mensual.

A continuación, se presenta la estructura de los datos:



1. **Metodologías**

En este estudio, se evaluarán tres metodologías para estimar los precios minoristas a partir de la información de precios mayoristas. Primero, se propone el calculo del margen mediano para cada uno de los productos reportados en SIPSA, así como el margen en el percentil 25 y 75. Segundo, se propone la estimación de un modelo SARIMAX, que permita capturar la evolución temporal de los precios minoristas, considerando los precios mayoristas como variable exógena y la variación estacional. En tercer lugar, se propone un modelo dinámico de corrección de errores (ECM), en el cual la dinámica de corto plazo de las variables está determinada por su variación respecto del equilibrio.

Para validar el ajuste de las metodologías para la predicción de los precios minoristas, se propone dividir la muestra en dos: un conjunto de datos de entrenamiento, que corresponde al 70% de la muestra total; y un conjunto de datos de validación, que corresponde al 30%. Los métodos son evaluados a partir de las métricas comúnmente empleadas: error cuadrático medio y error absoluto medio porcentual.

* 1. **Mapeo: IPC-SIPSA**

La implementación de las metodologías I-III presupone un mapeo previo entre los alimentos reportados por el DANE en la construcción del IPC y los alimentos reportados por SIPSA (en adelante, mapeo IPC-SIPSA). Es decir, para cada alimento reportado en SIPSA, existe un alimento equivalente en la canasta del IPC. A continuación, se presenta el mapeo:



Nótese que el mapeo no es 1:1, sino n: 1 (muchos: 1). Esto quiere decir que un mismo alimento de la base de datos del IPC puede tener varios alimentos equivalentes en la base de datos de SIPSA. (Un ejemplo: el arroz para seco, que es un alimento del IPC, corresponde tanto al arroz de primera como al arroz de segunda y el arroz excelso en la base de datos de SIPSA).

* 1. **Metodologías**

La siguiente tabla resume las metodologías implementadas en este documento de trabajo:

|  |  |
| --- | --- |
| No. | Metodología |
| I | Ajuste sobre los precios minoristas a partir del IPC |
| II | Estimación del margen mediano (Q2 [Q1, Q3]) |
| III | Regresión lineal en niveles, controlando por el efecto estacional. |
| IV | Regresión en primeras diferencias, controlando por el efecto estacional. |
| V | Modelo de corrección de error |
| VI | Modelo de corrección de error asimétrico |

1. **Descripción de las metodologías y resultados**
   1. **Metodología I: Aproximación a partir del IPC**

Para estimar los precios minoristas incompletos, se implementa una estrategia basada en la variación mensual del IPC a nivel de subclase. Para cada artículo en cada ciudad, se separaron los datos en dos subconjuntos:

* Conjunto de entrenamiento hasta enero de 2015 con observaciones sobre precios minoristas.
* Conjunto de validación a partir de enero de 2015, a partir del cual se estimarán los precios usando las variaciones del ÍPC.

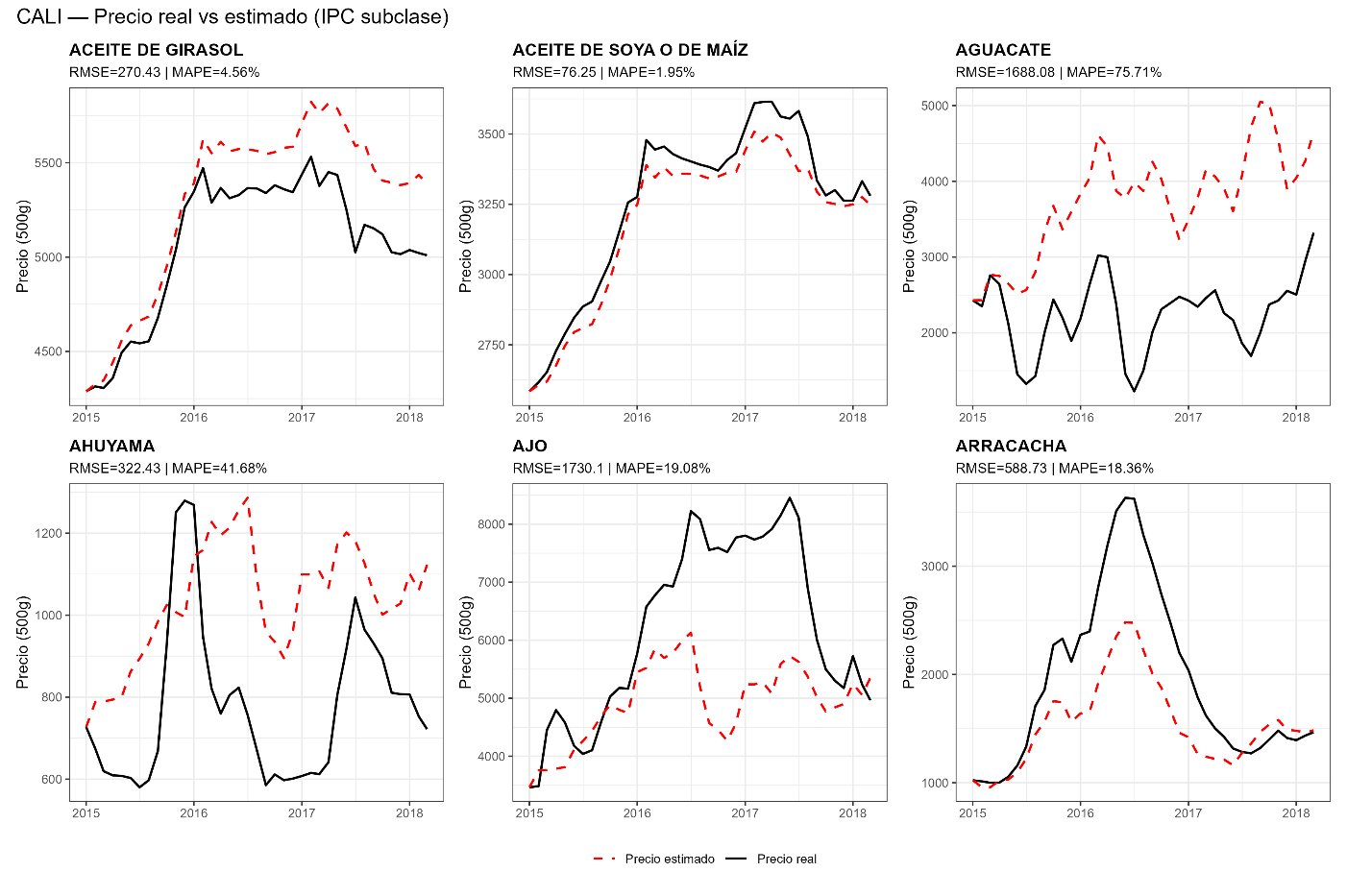
Se utilizó el último precio observado del conjunto de entrenamiento (enero de 2015) y se estimaron los precios hacia adelante multiplicando iterativamente por la variación mensual del ÍPC:

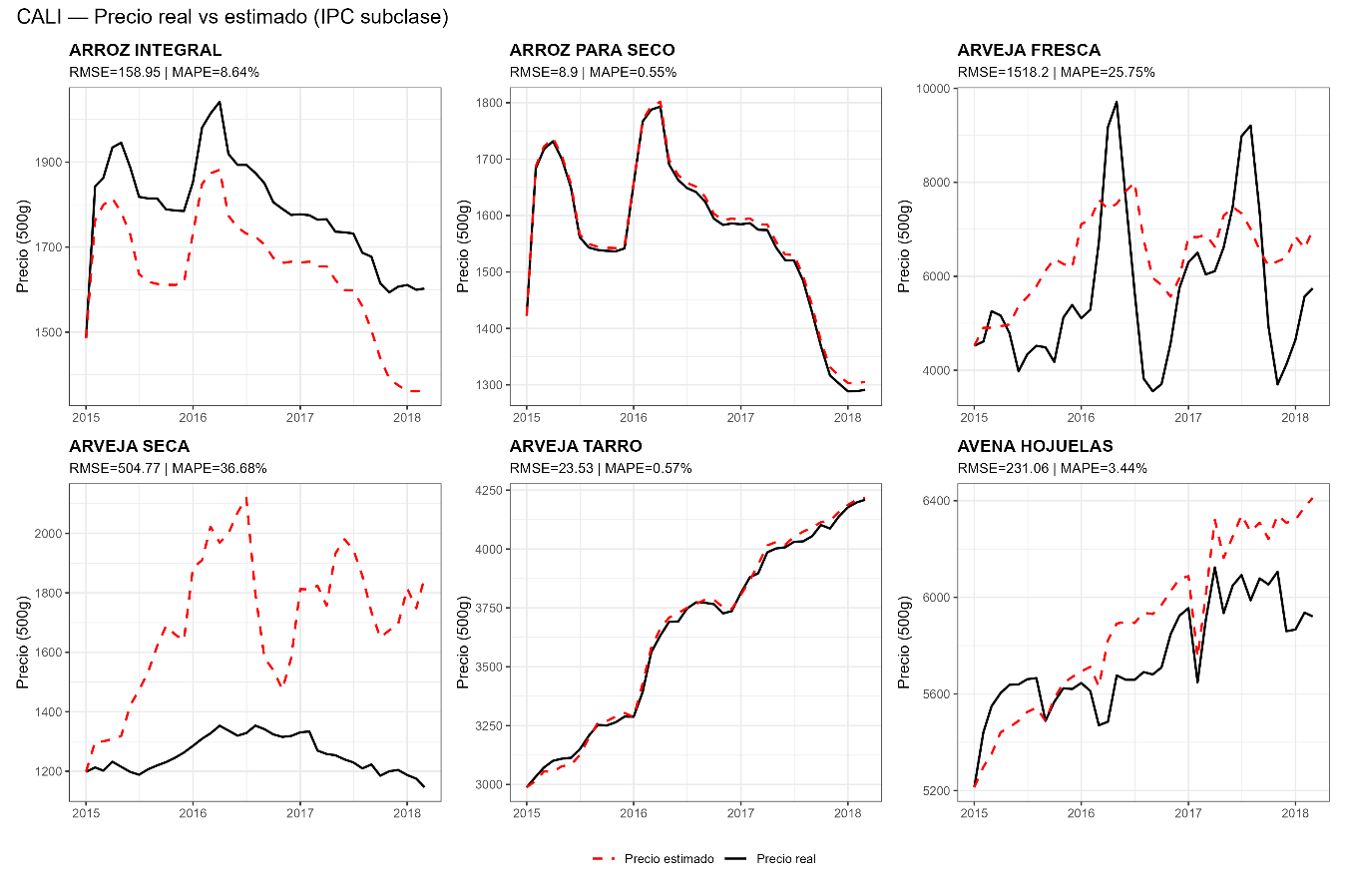
Esta operación se repite mes a mes hasta llegar al final del conjunto de validación (marzo de 2018). Esta metodología proporciona una estimación directa de los precios minorista y evita, por tanto, el tratamiento del problema de la transmisión mayorista-minorista.

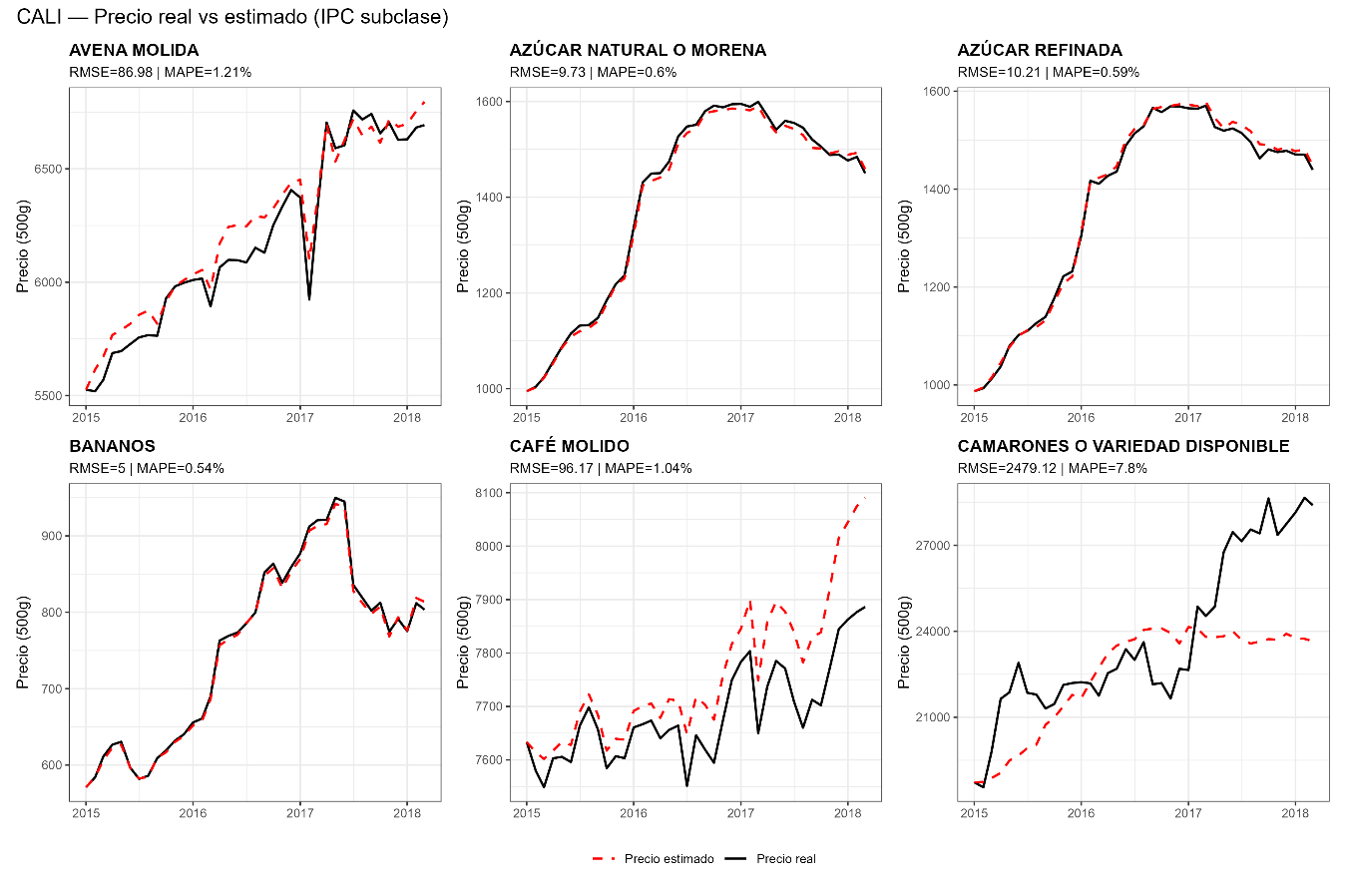
**Resultados:**

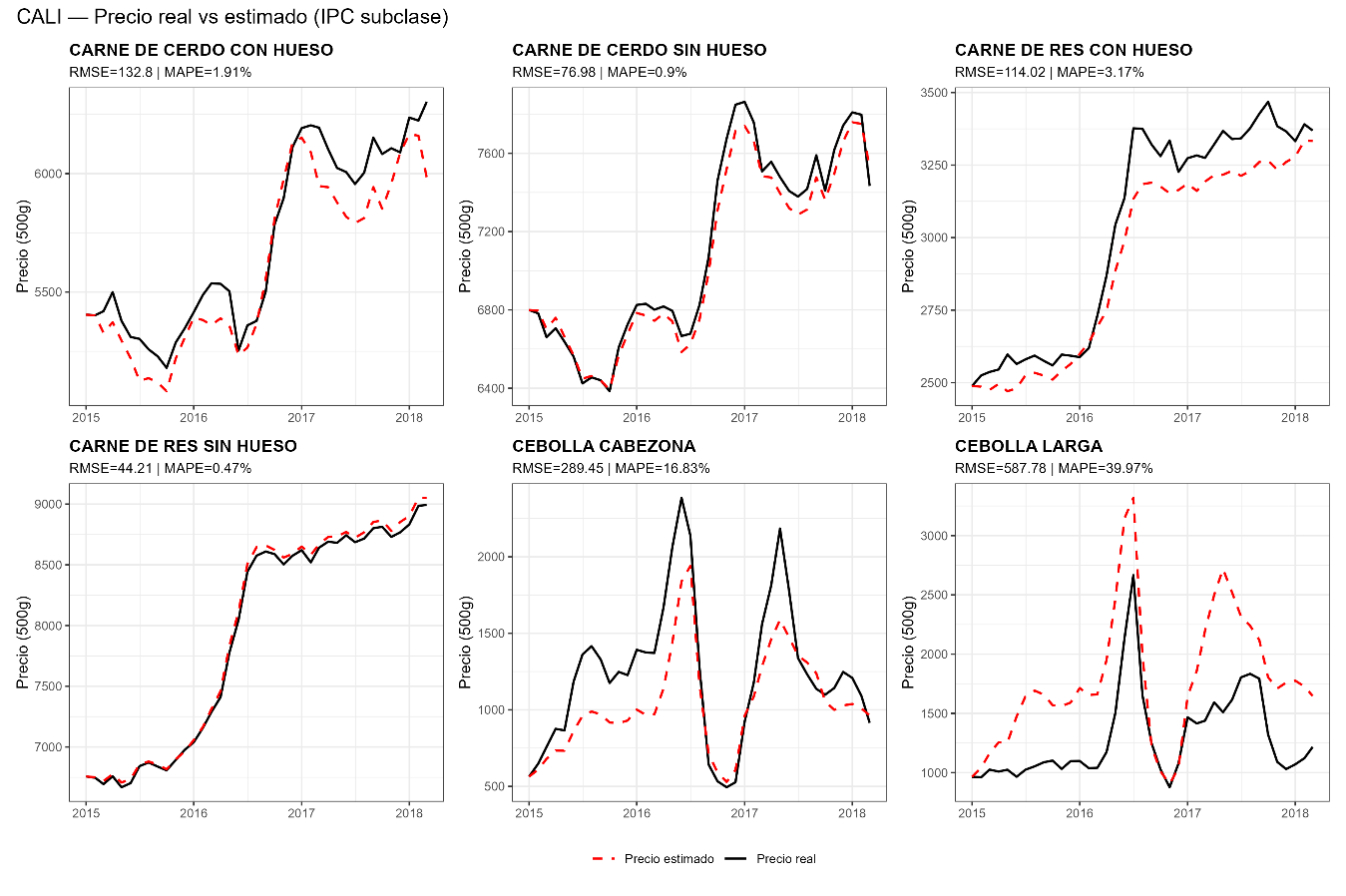
**Nota para Daniela: la ruta de los resultados es esta:**

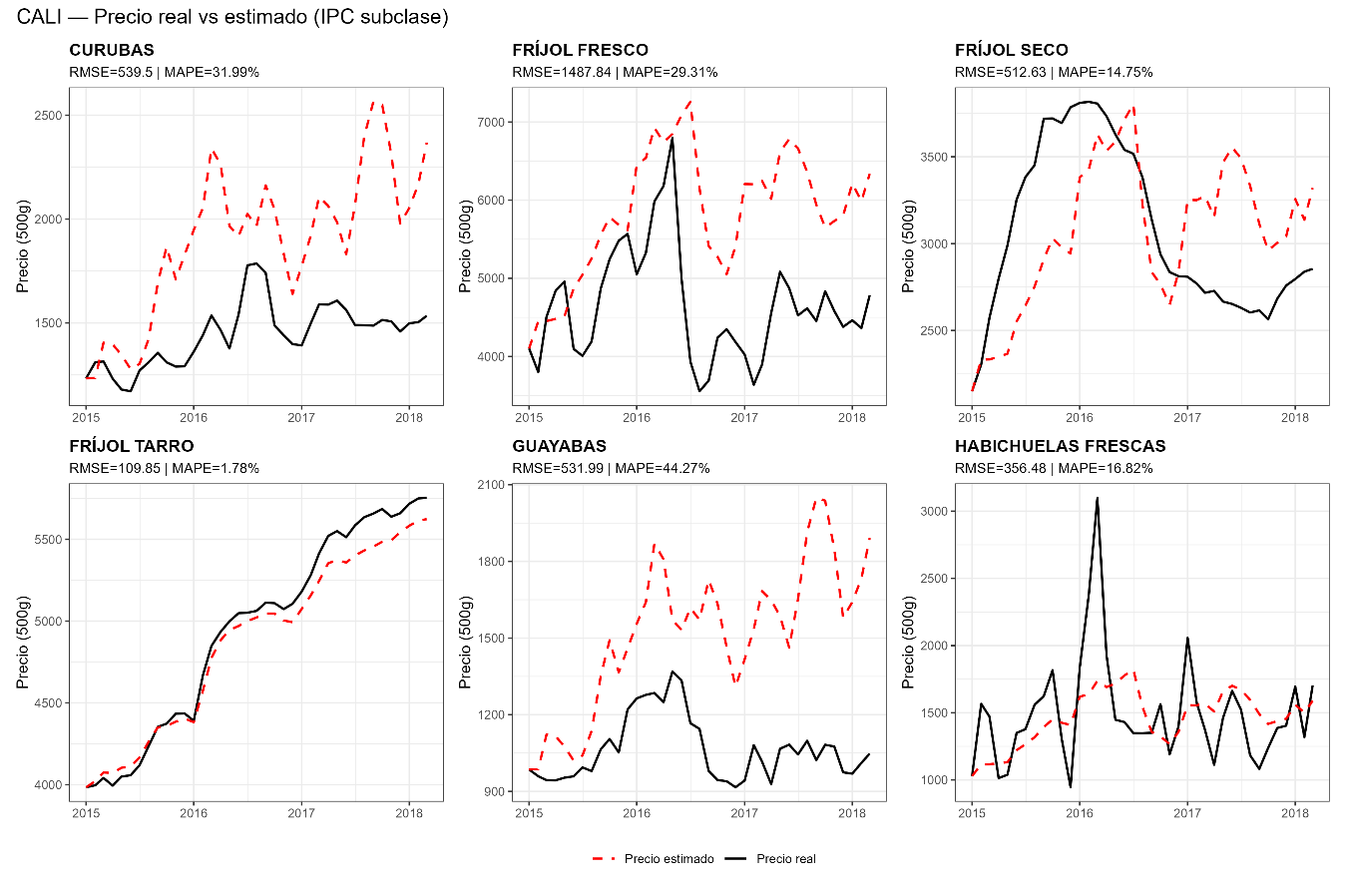
**C:\Users\Portatil\Desktop\Least-cost-diets-and-affordability\Proyecto Interno\working-papers\working-paper-1225\m0\output\_ipc\_fill**

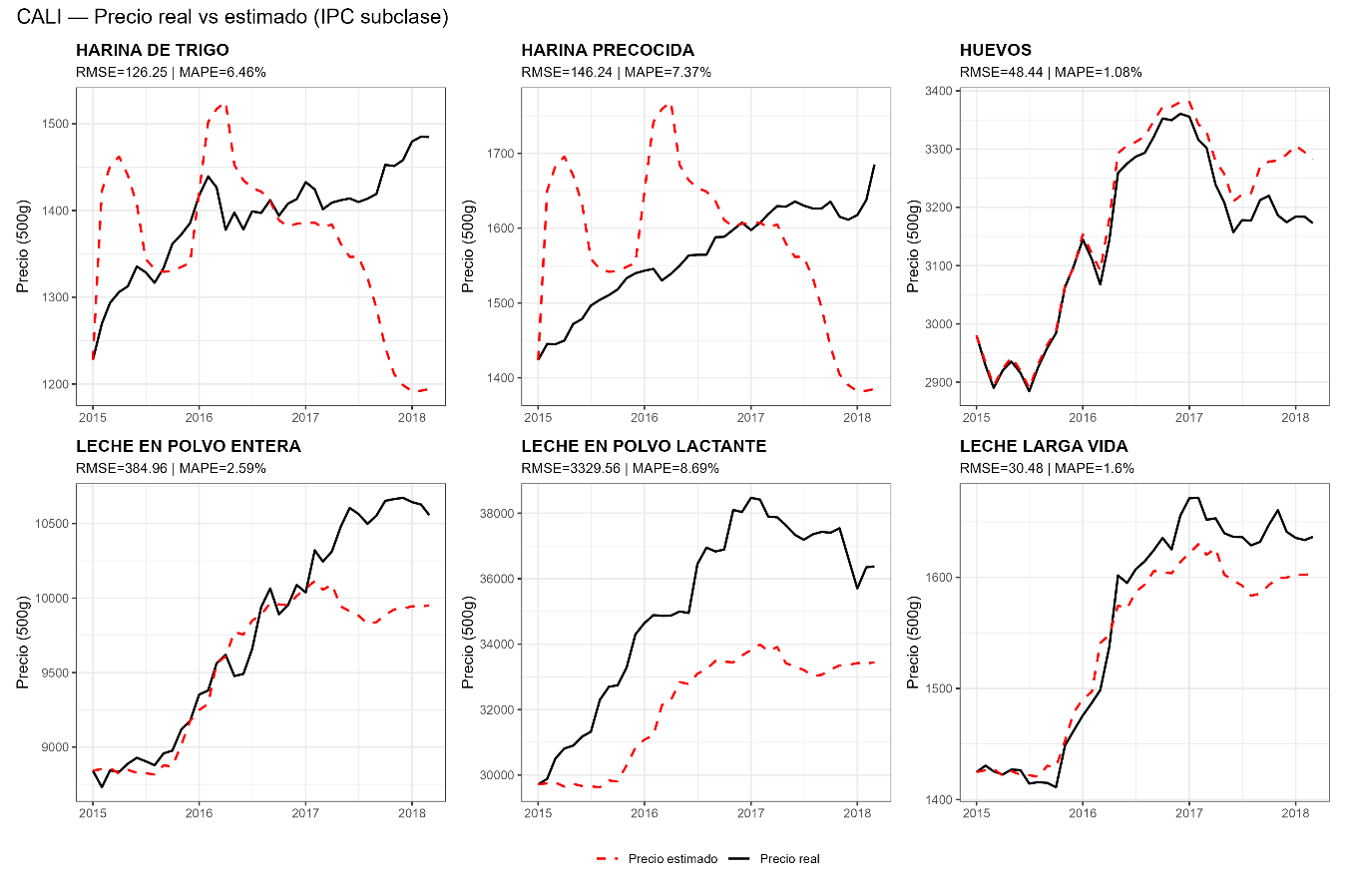
****

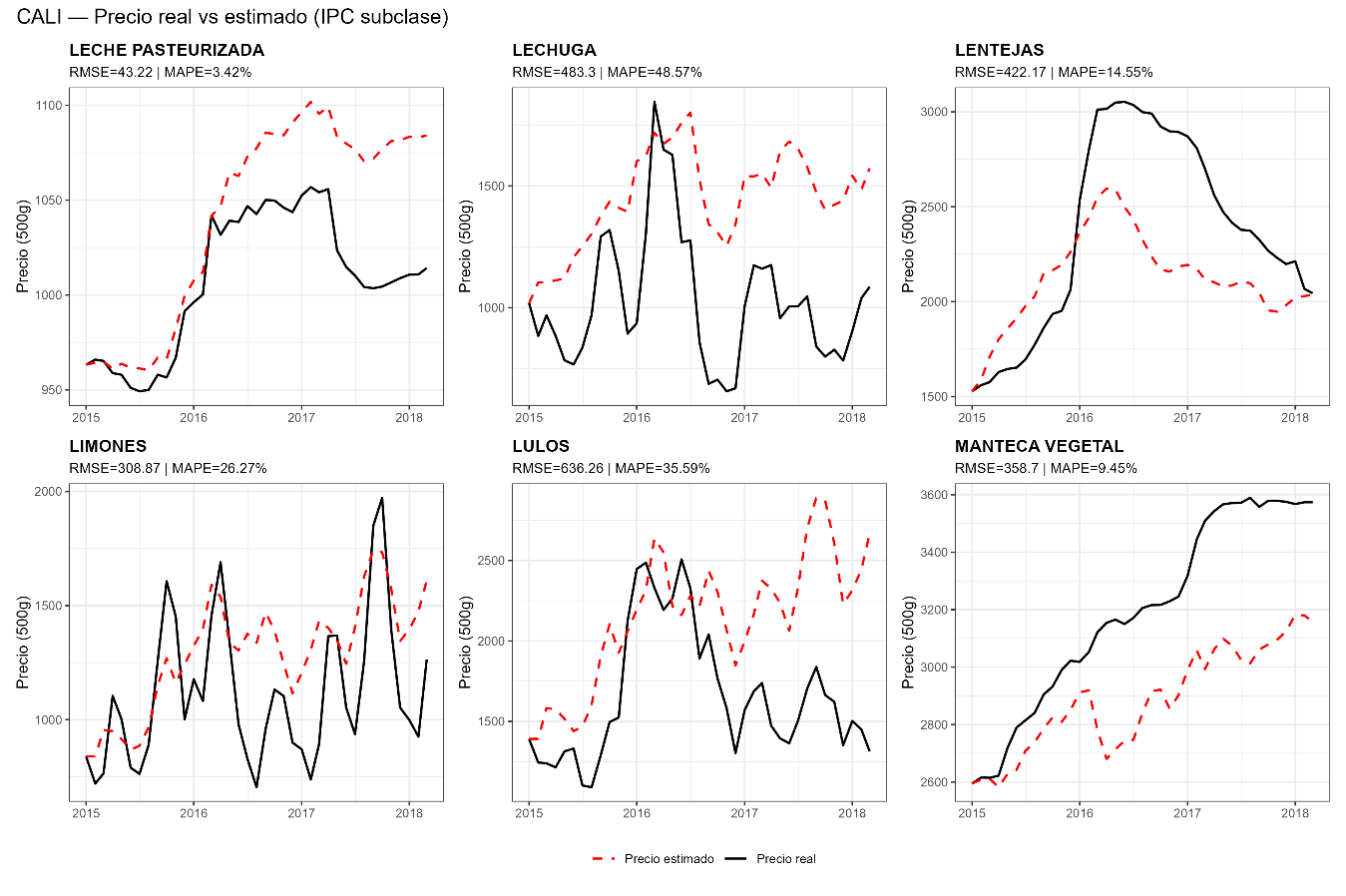
****

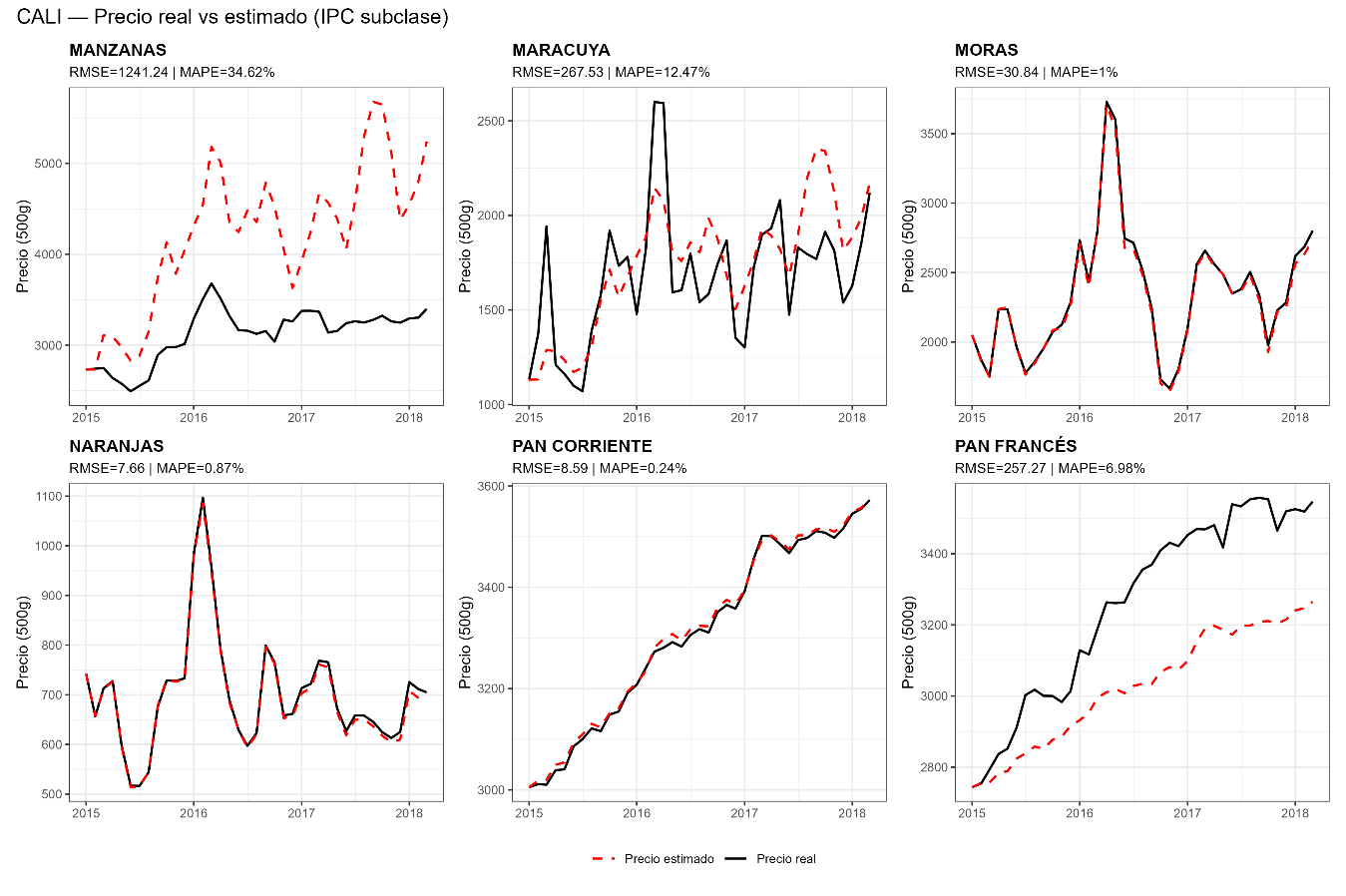
****

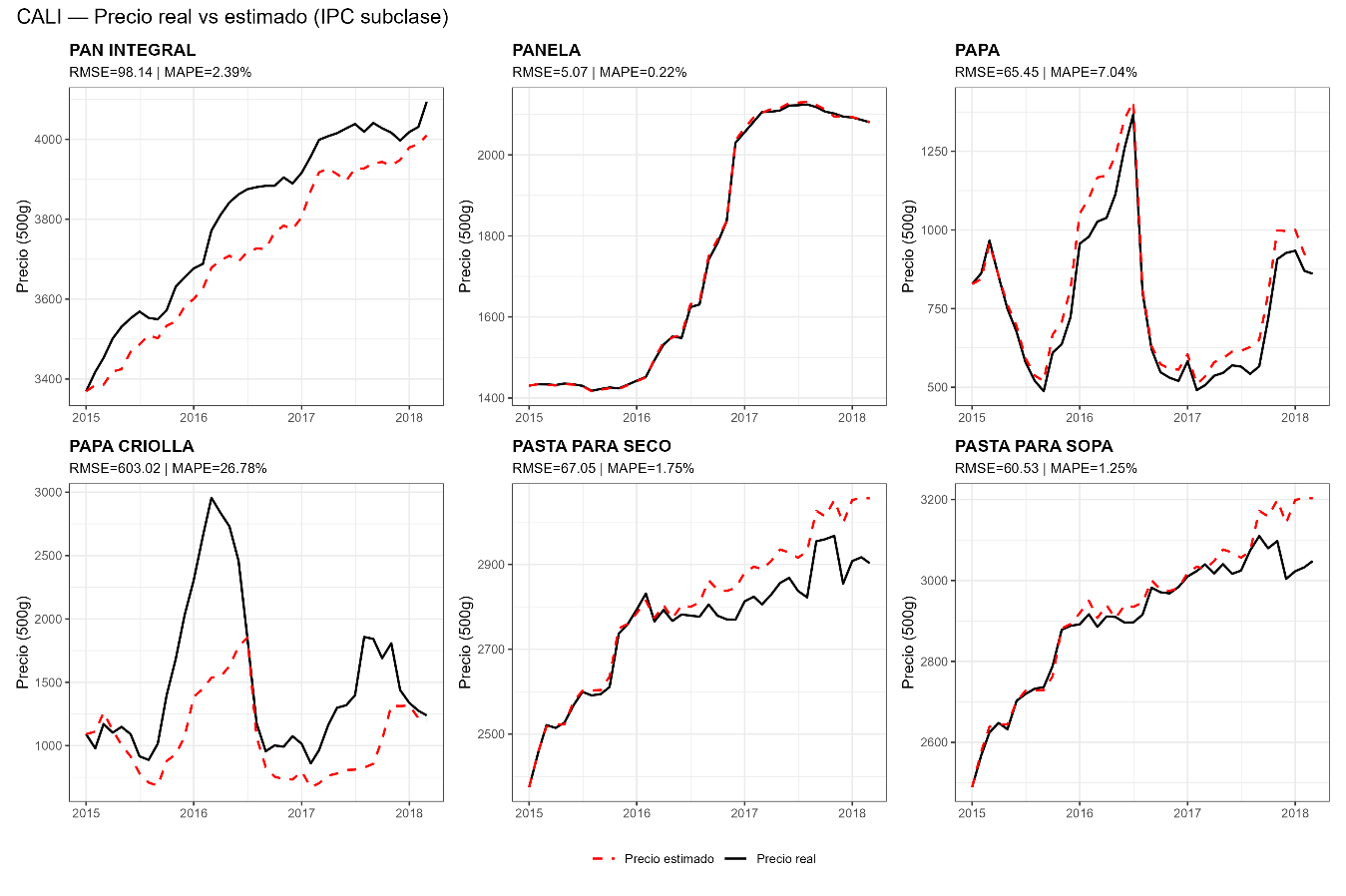
****

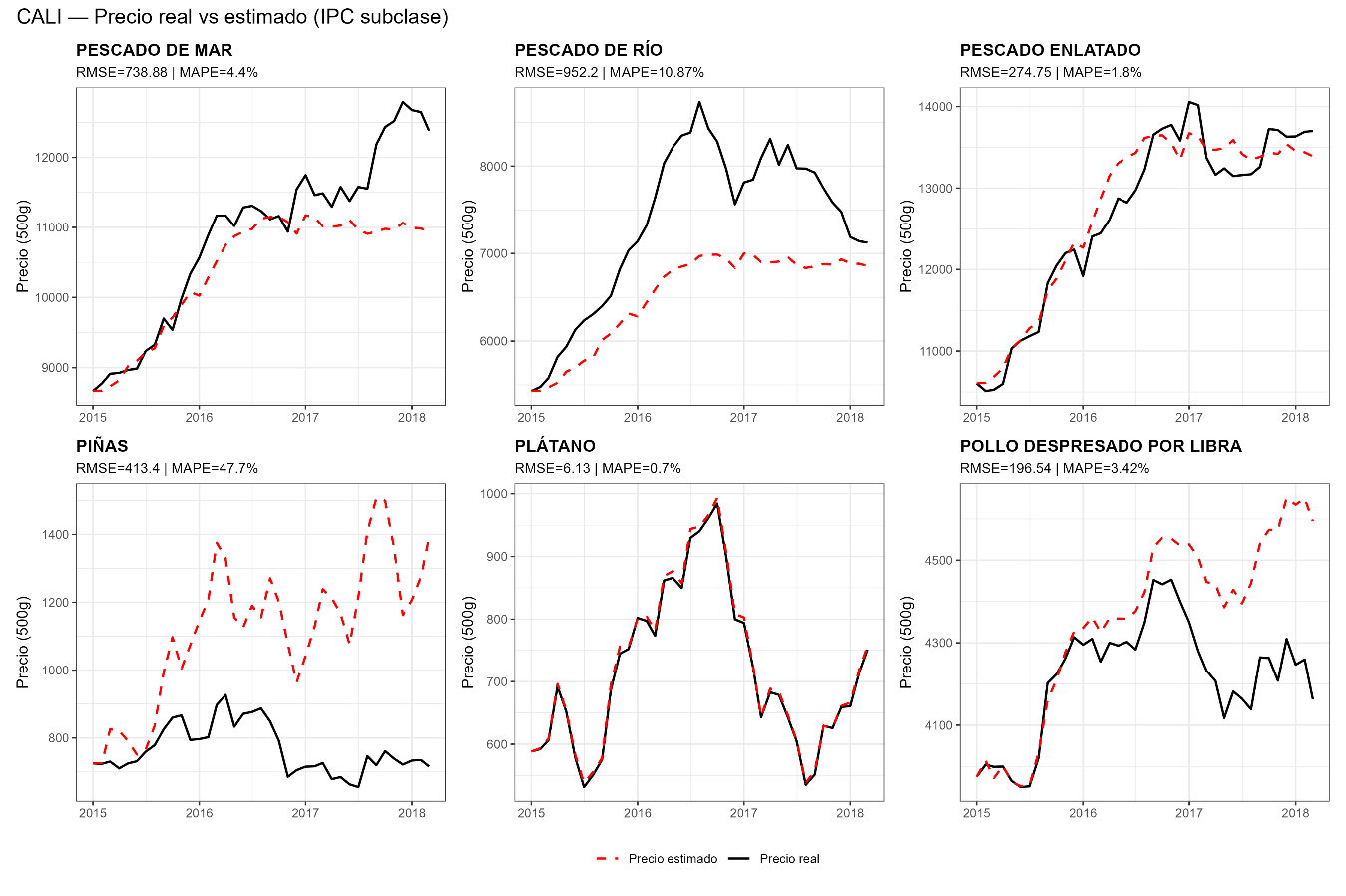
****

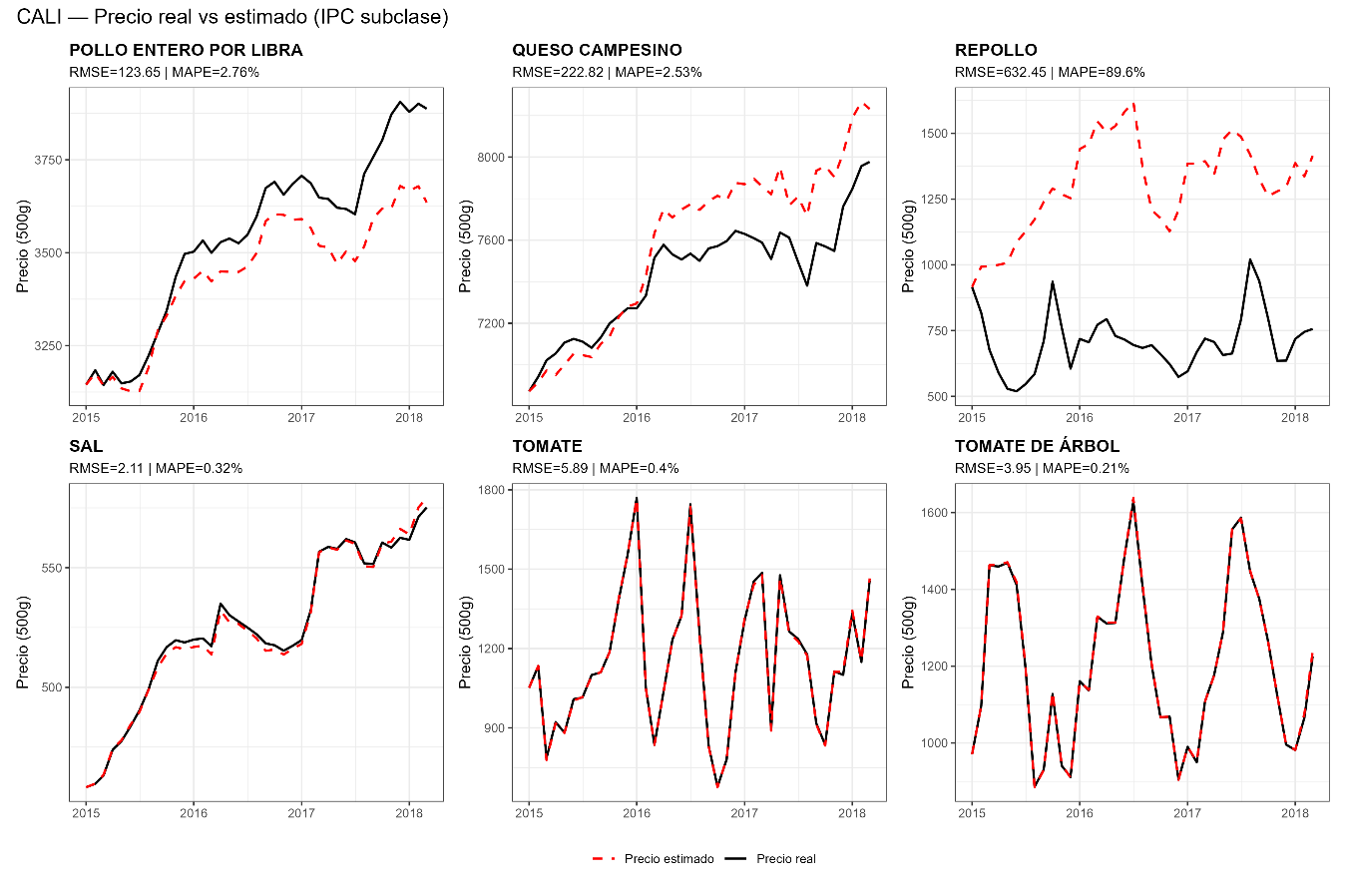
****

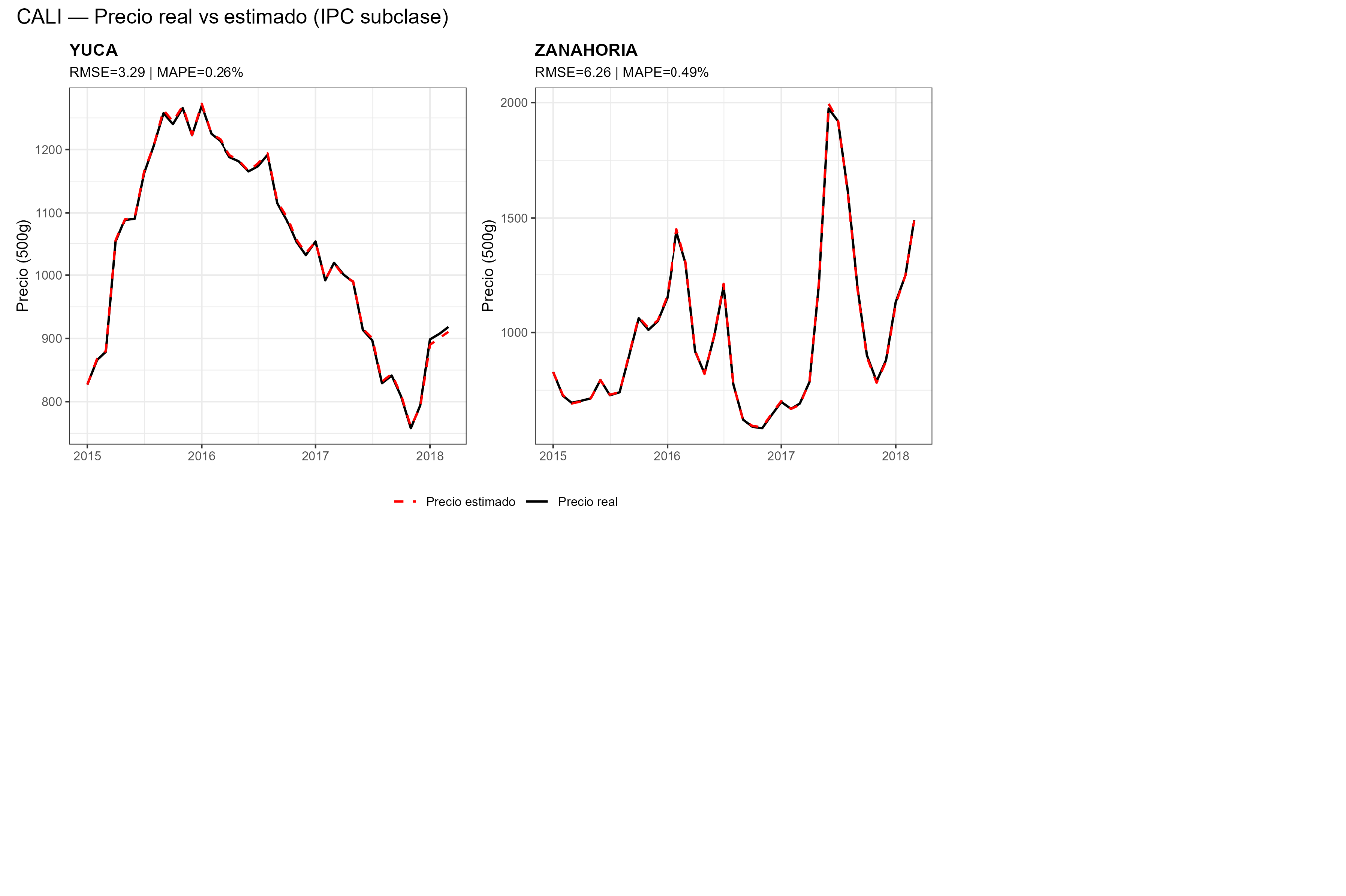
****

****

****

****

****

****

* 1. **Metodología II: Estimación a partir del margen mediano (Q1 – Q3)**

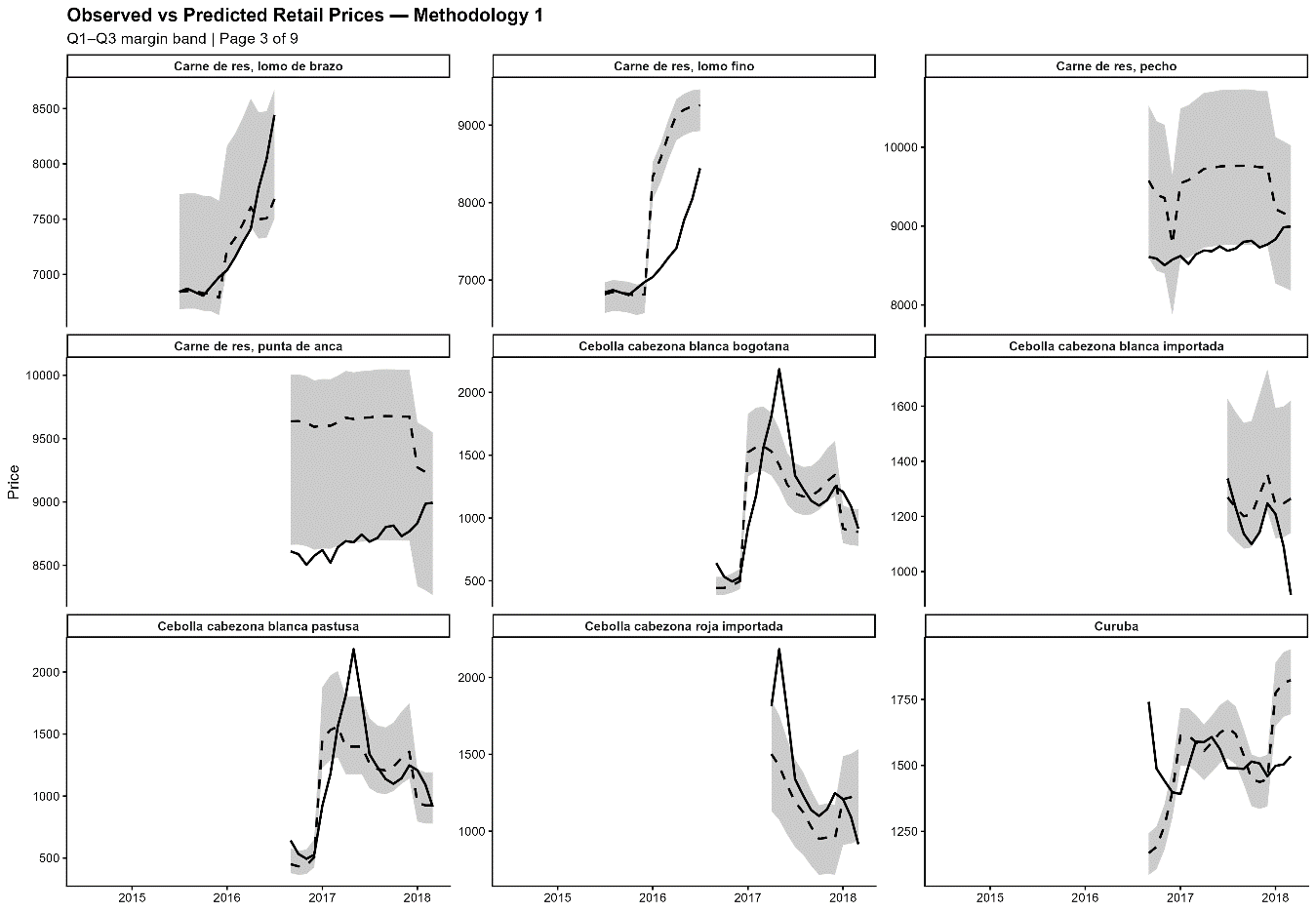
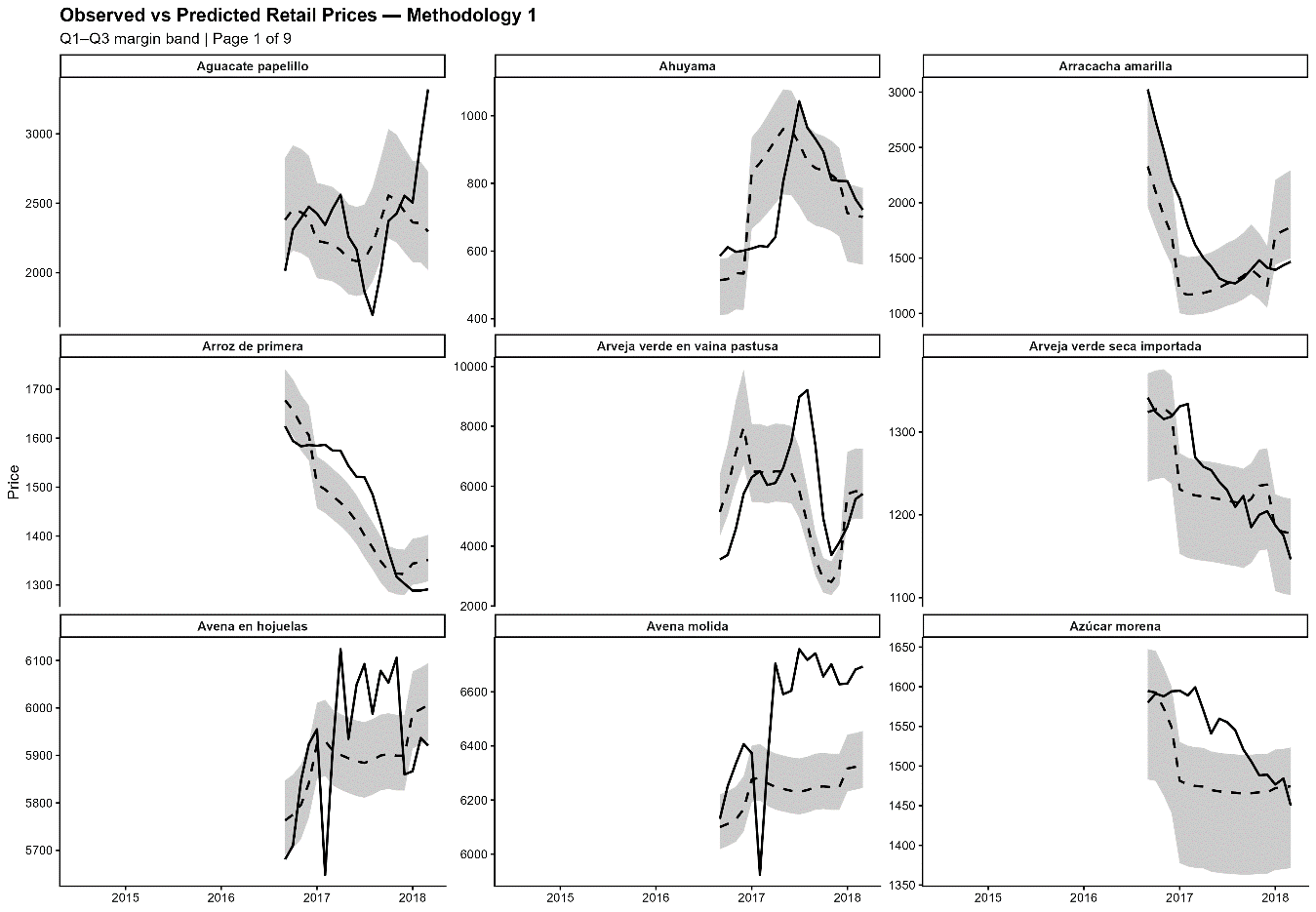
La primera metodología propuesta consiste en el cálculo del margen de comercialización para cada uno de los productos reportados por SIPSA. A partir del mapeo IPC-SIPSA, para el alimento k en t, el margen de comercialización se calcula simplemente como:

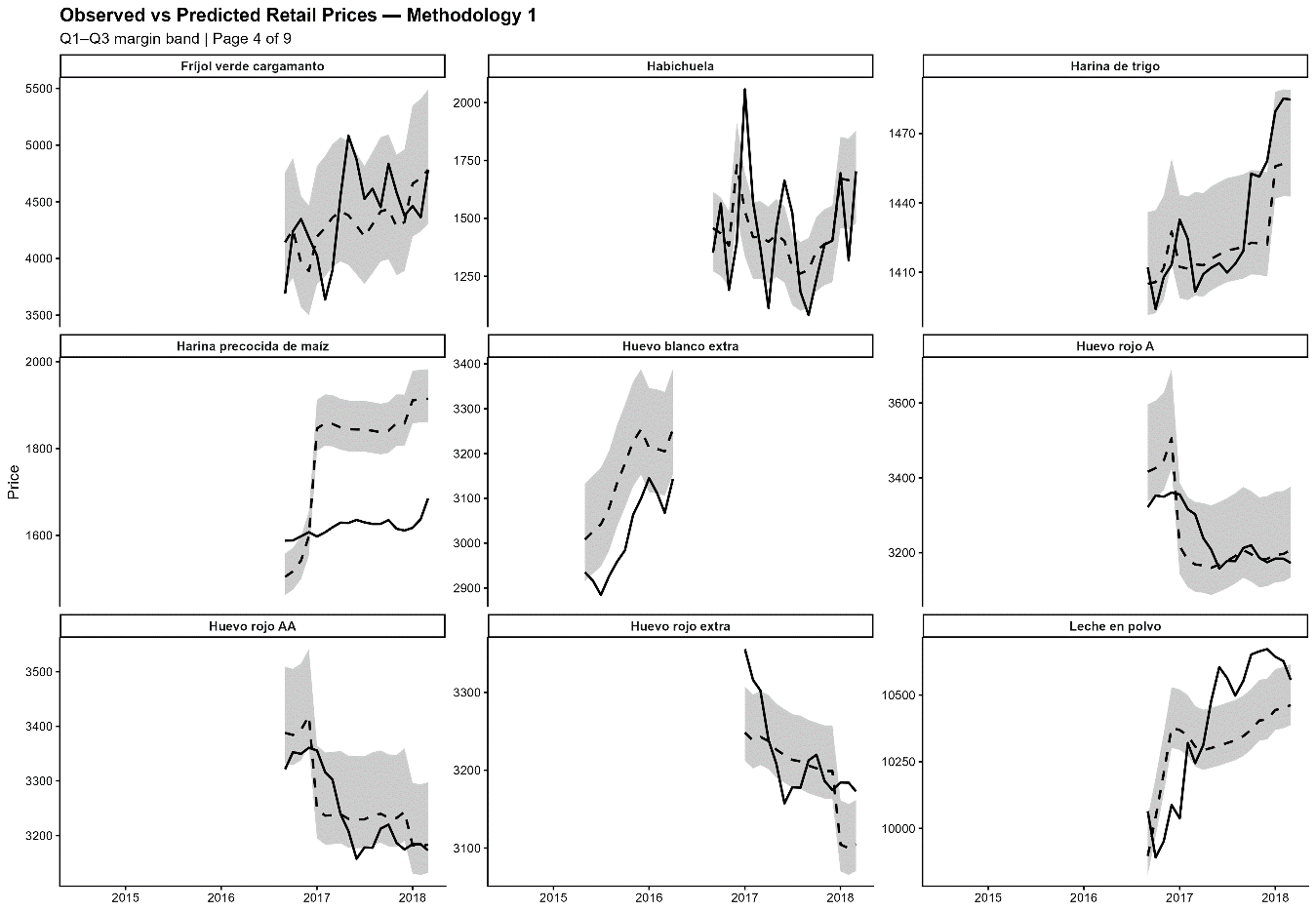
Con base en los márgenes de comercialización estimados, se obtiene la distribución para cada uno de los alimentos reportados por SIPSA. A partir de la distribución, se calcula el margen mediano (Q2), así como el margen en el percentil 25 (Q1) y 75 (Q3).

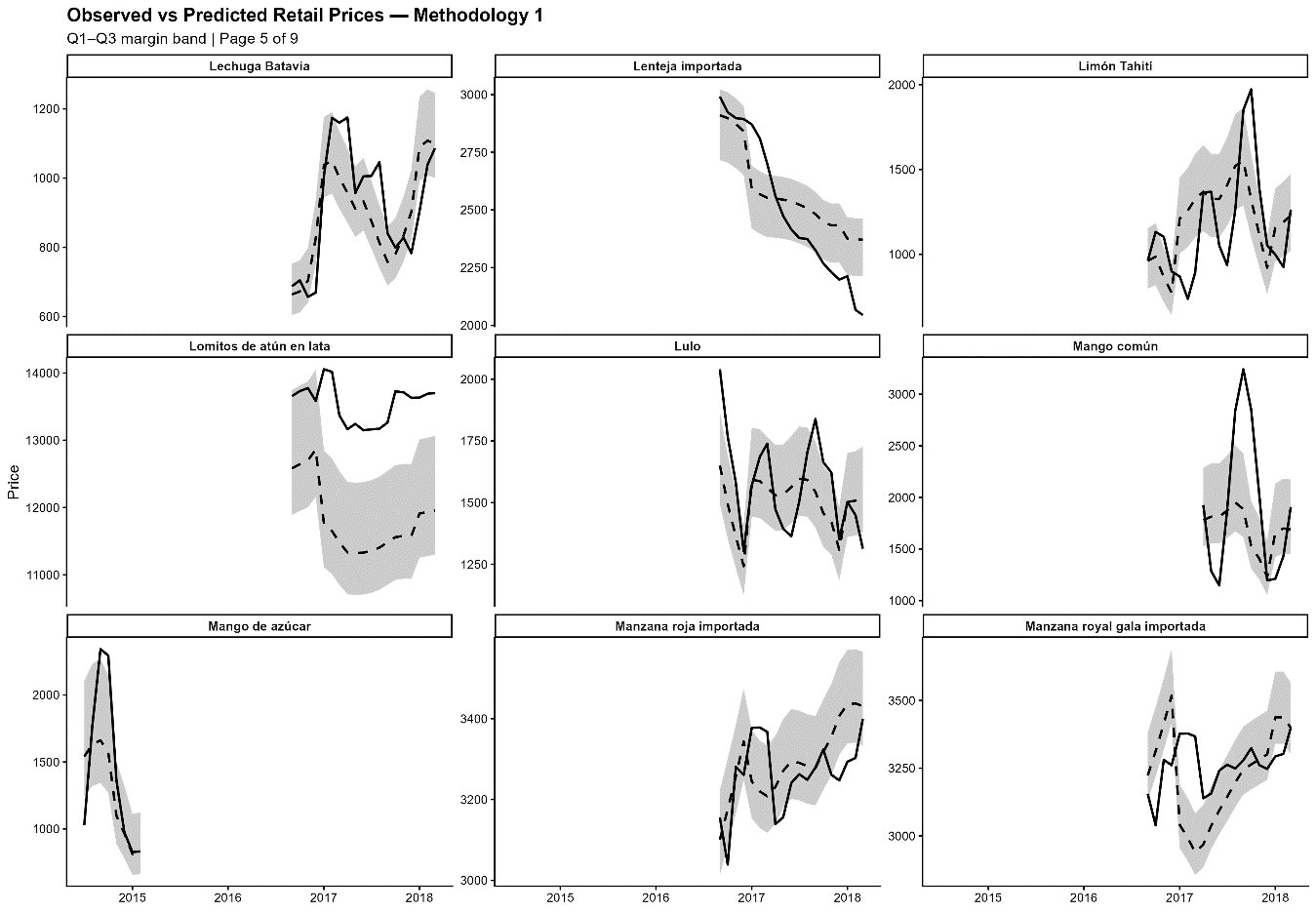
**Resultados:**

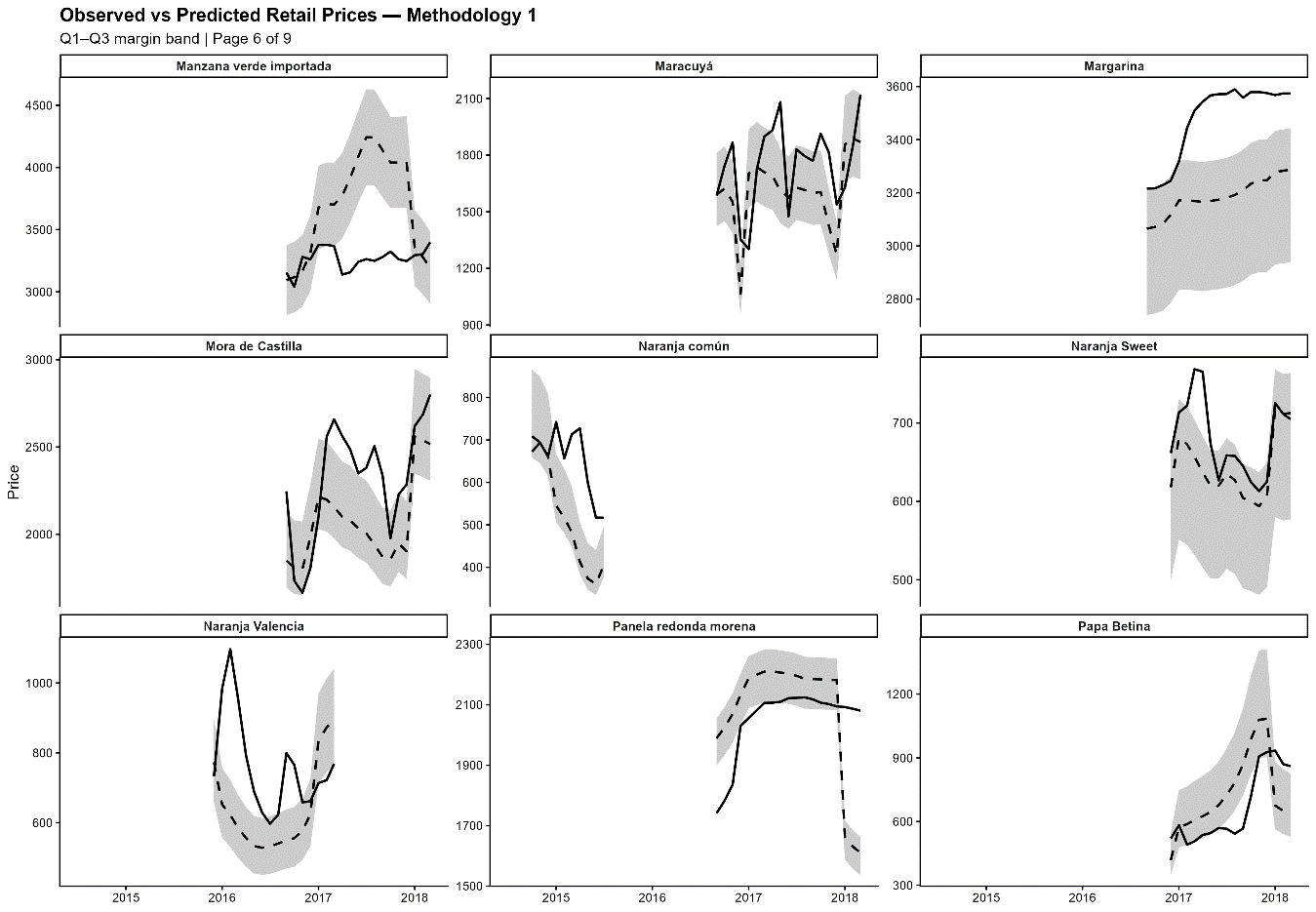
**Nota para Daniela: esta es la ruta:**

**C:\Users\Portatil\Desktop\Least-cost-diets-and-affordability\Proyecto Interno\working-papers\working-paper-1225\m1\output**

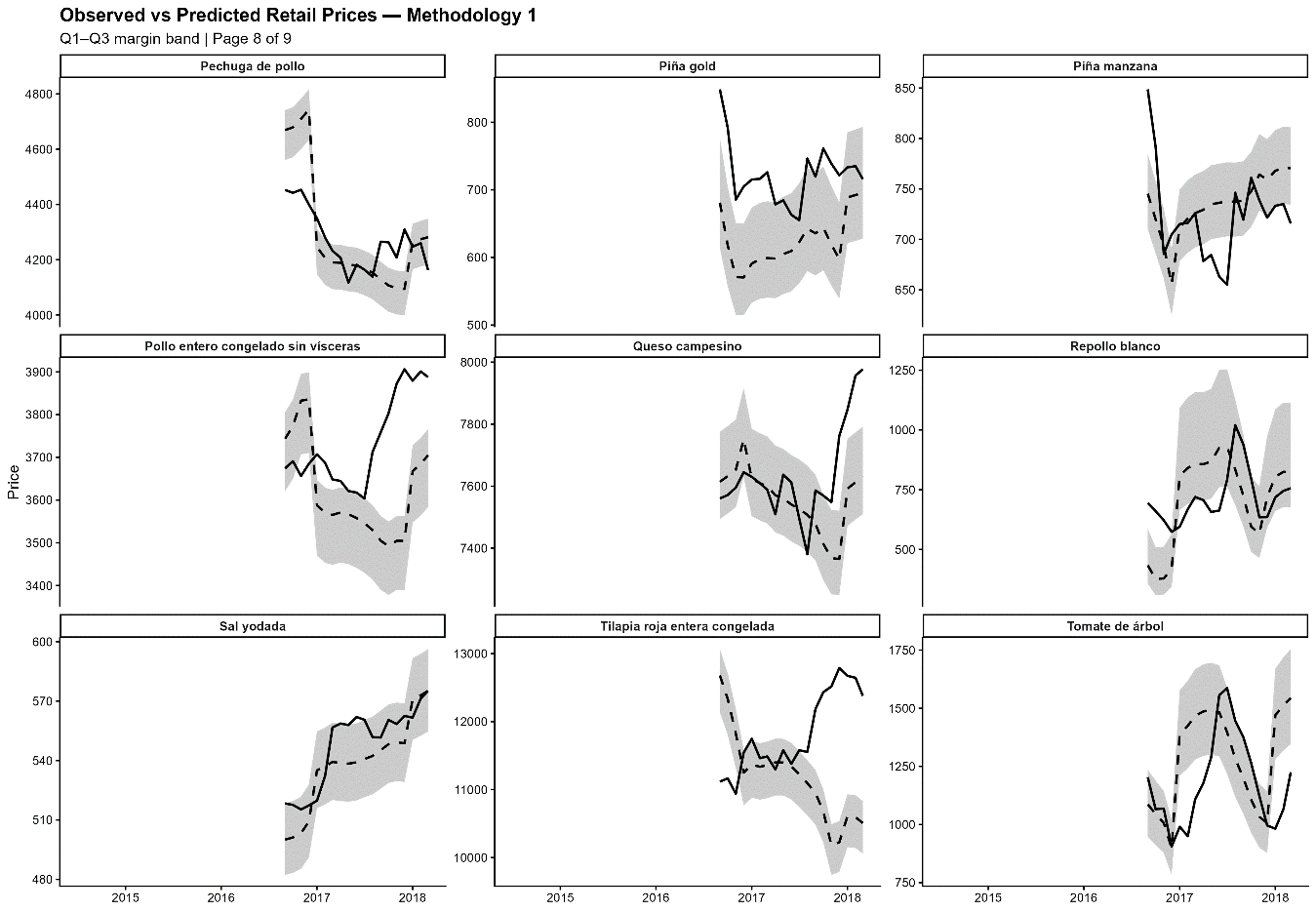
****

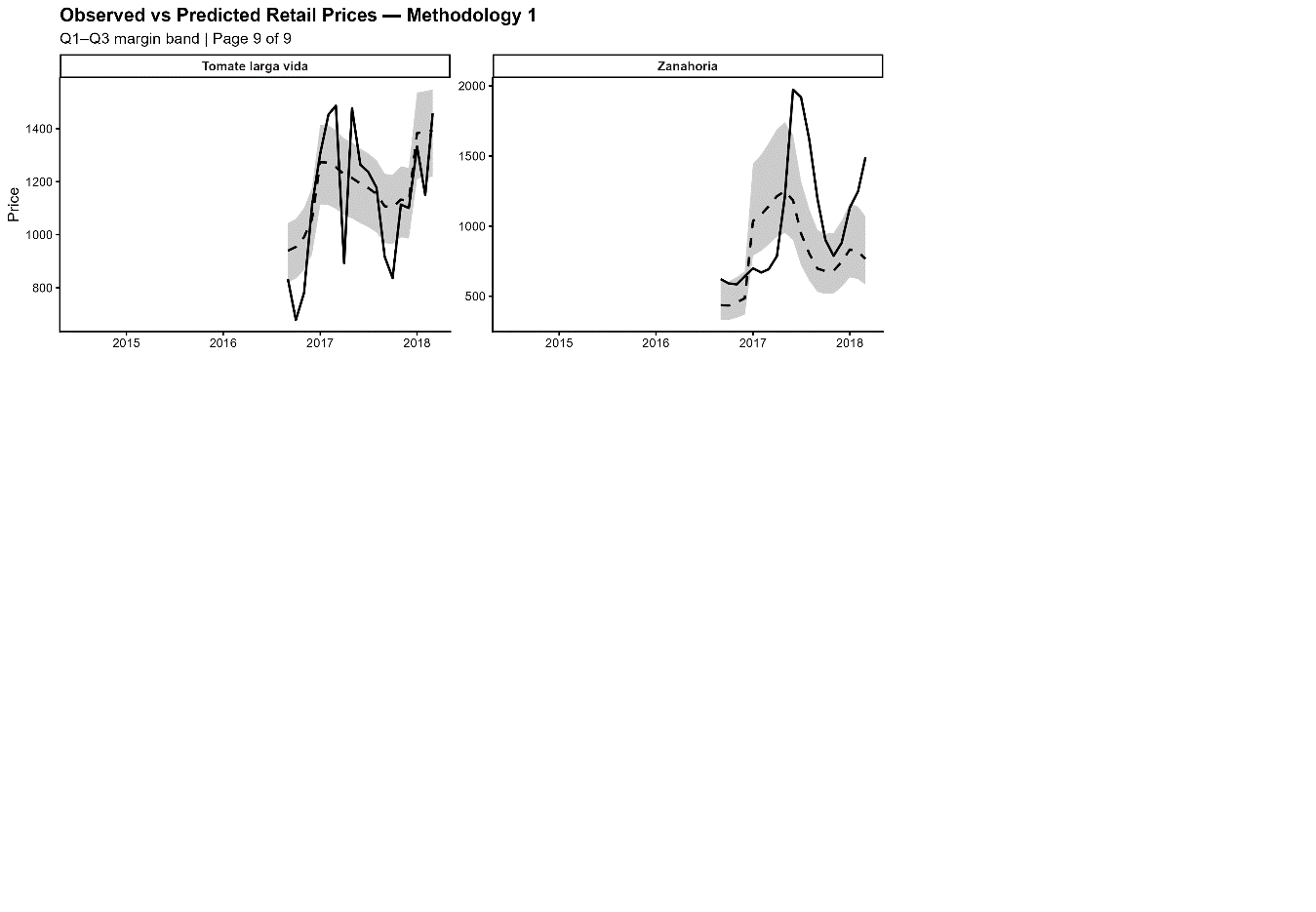
****

****

****

****

****

****

* 1. **Metodología III: Regresión lineal en niveles y primeras diferencias**

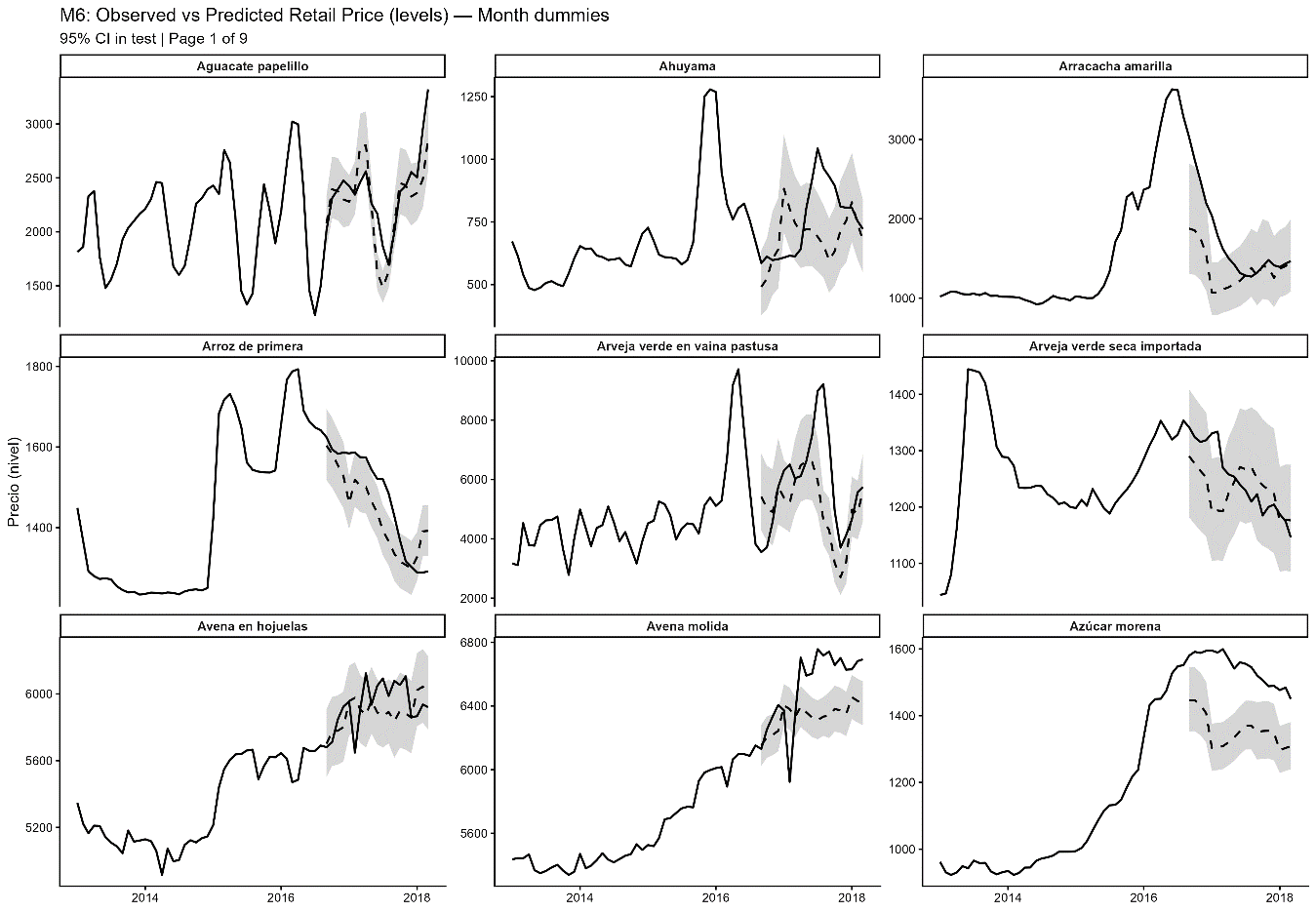
Sea el logaritmo natural del precio minorista, y sea el logaritmo natural del precio mayorista del alimento en el período . Para cada alimento , se estima el siguiente modelo de regresión lineal en niveles logarítmicos:

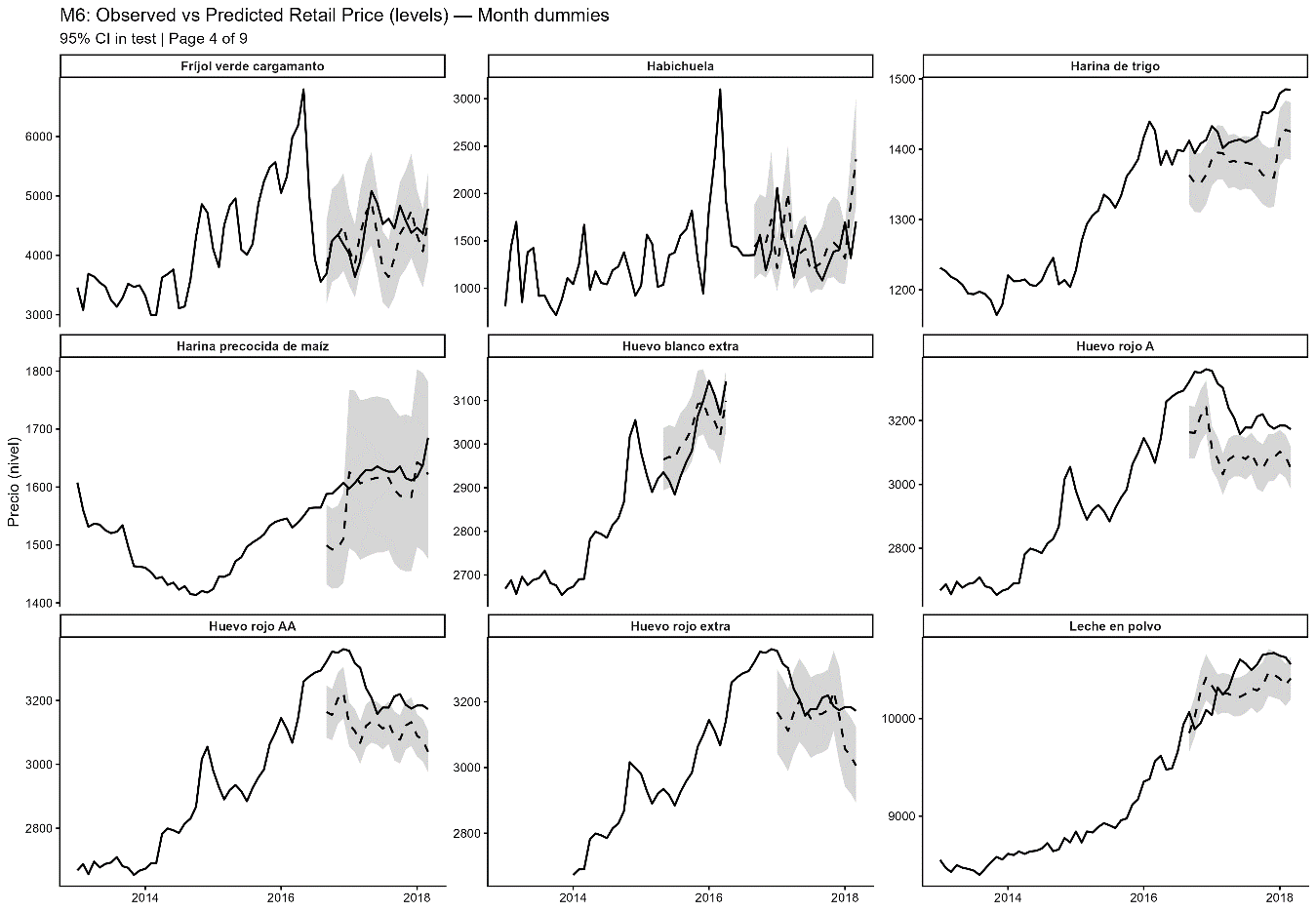
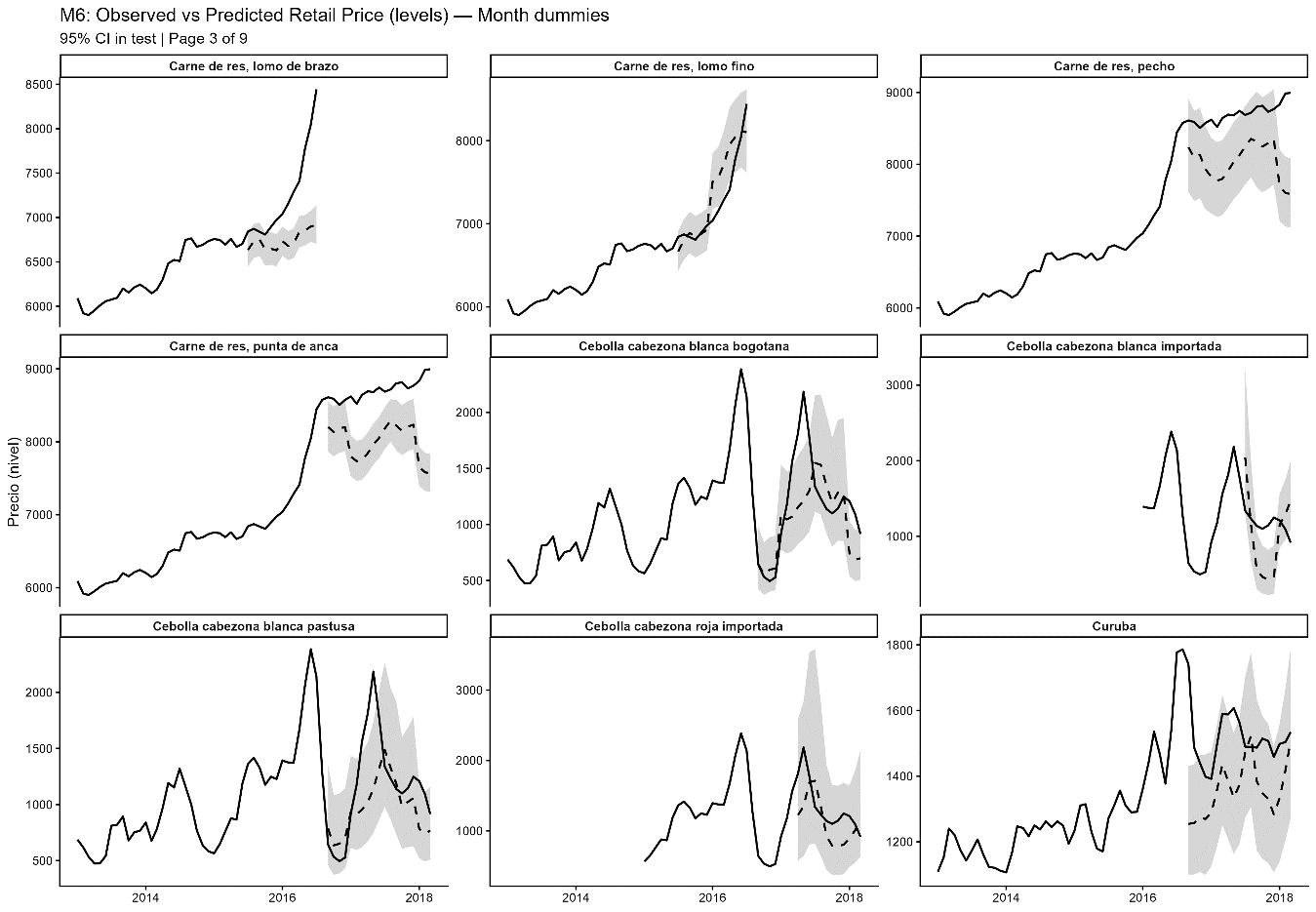
donde

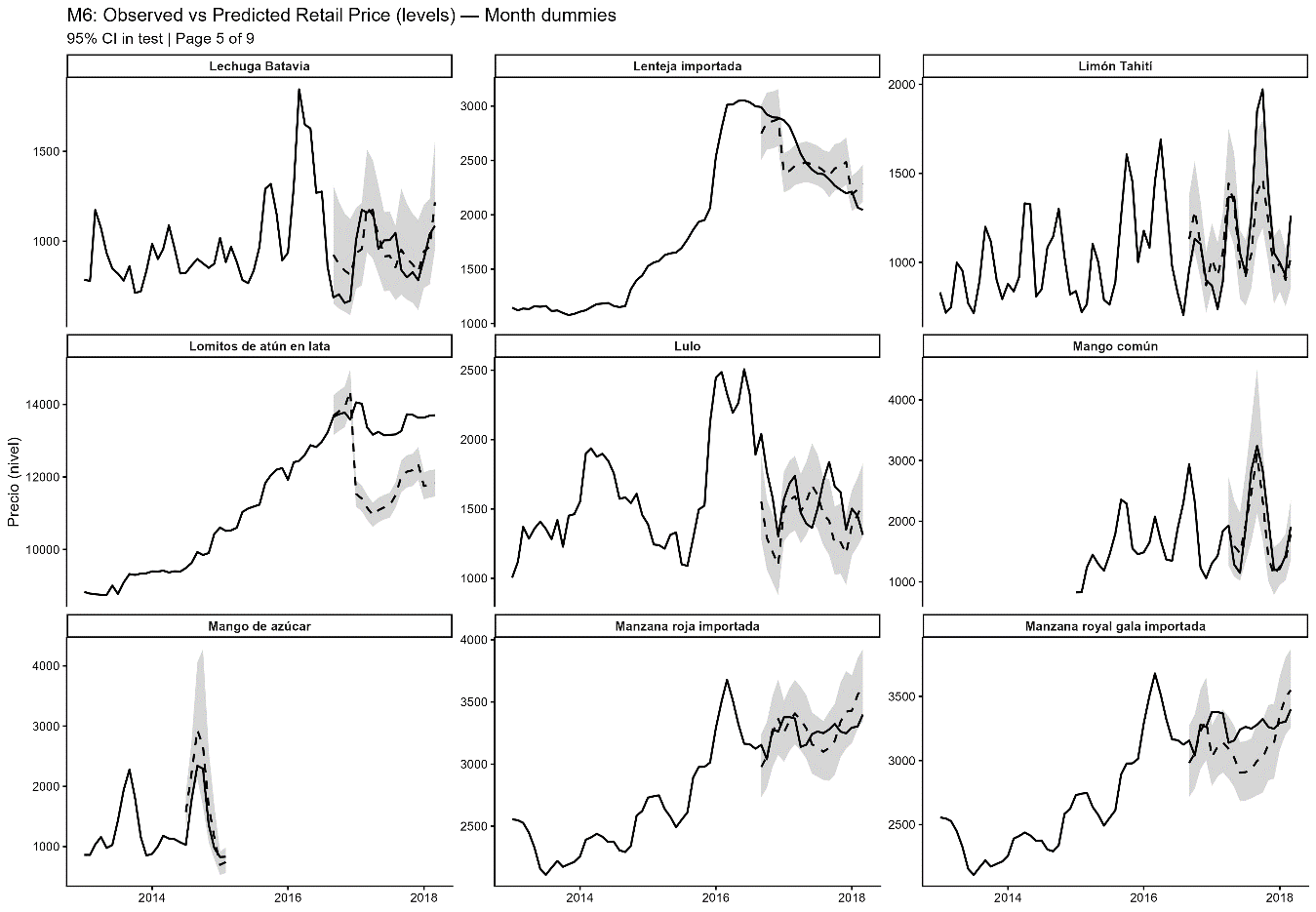
Nótese que, para cada alimento i, mide la elasticidad (o pass-through contemporáneo) dl precio minorista respecto del precio mayorista. Esto implica que captura el margen de comercialización implícito.

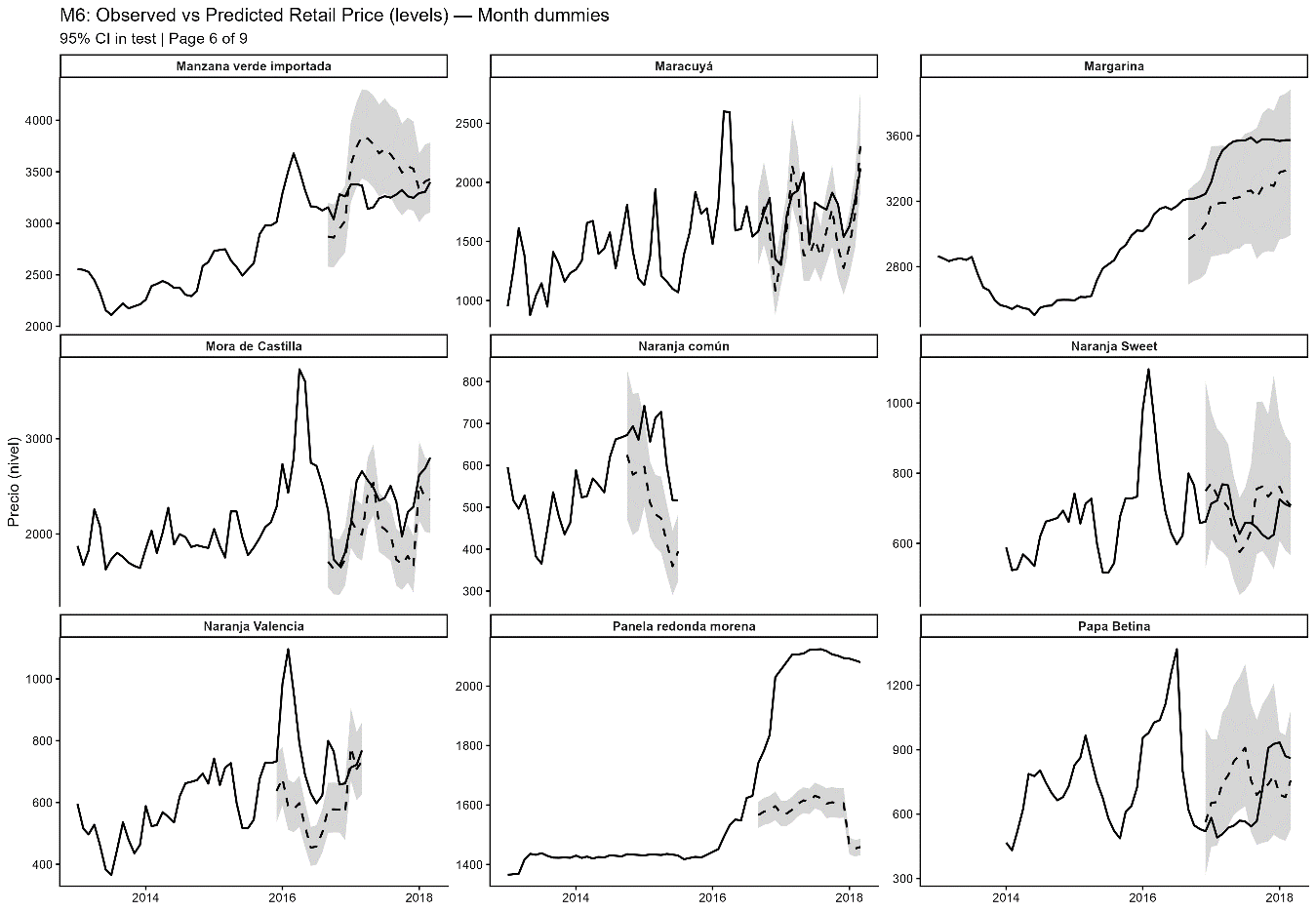
**Resultados:**

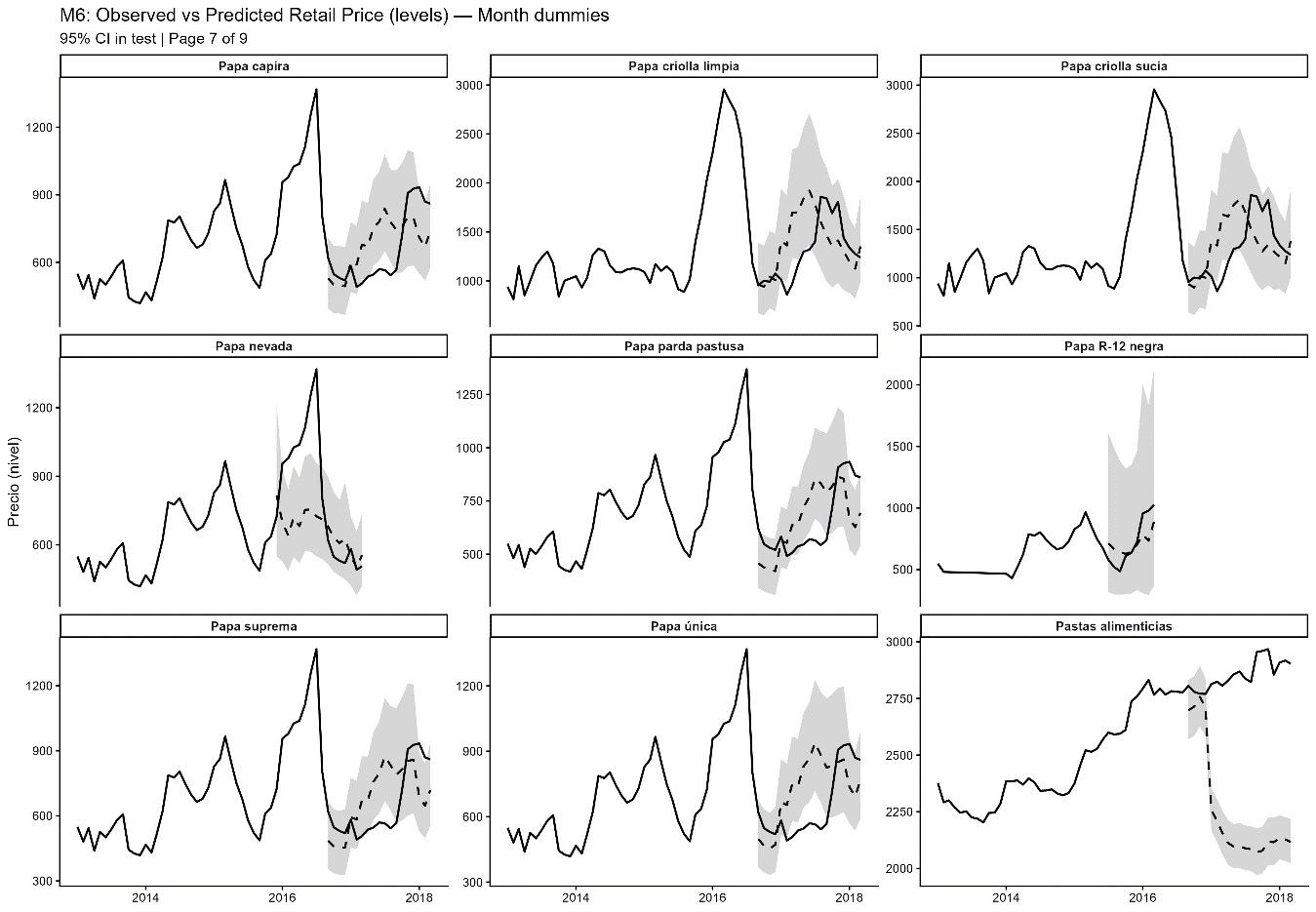
**C:\Users\Portatil\Desktop\Least-cost-diets-and-affordability\Proyecto Interno\working-papers\working-paper-1225\m6\** **output\_dummies**

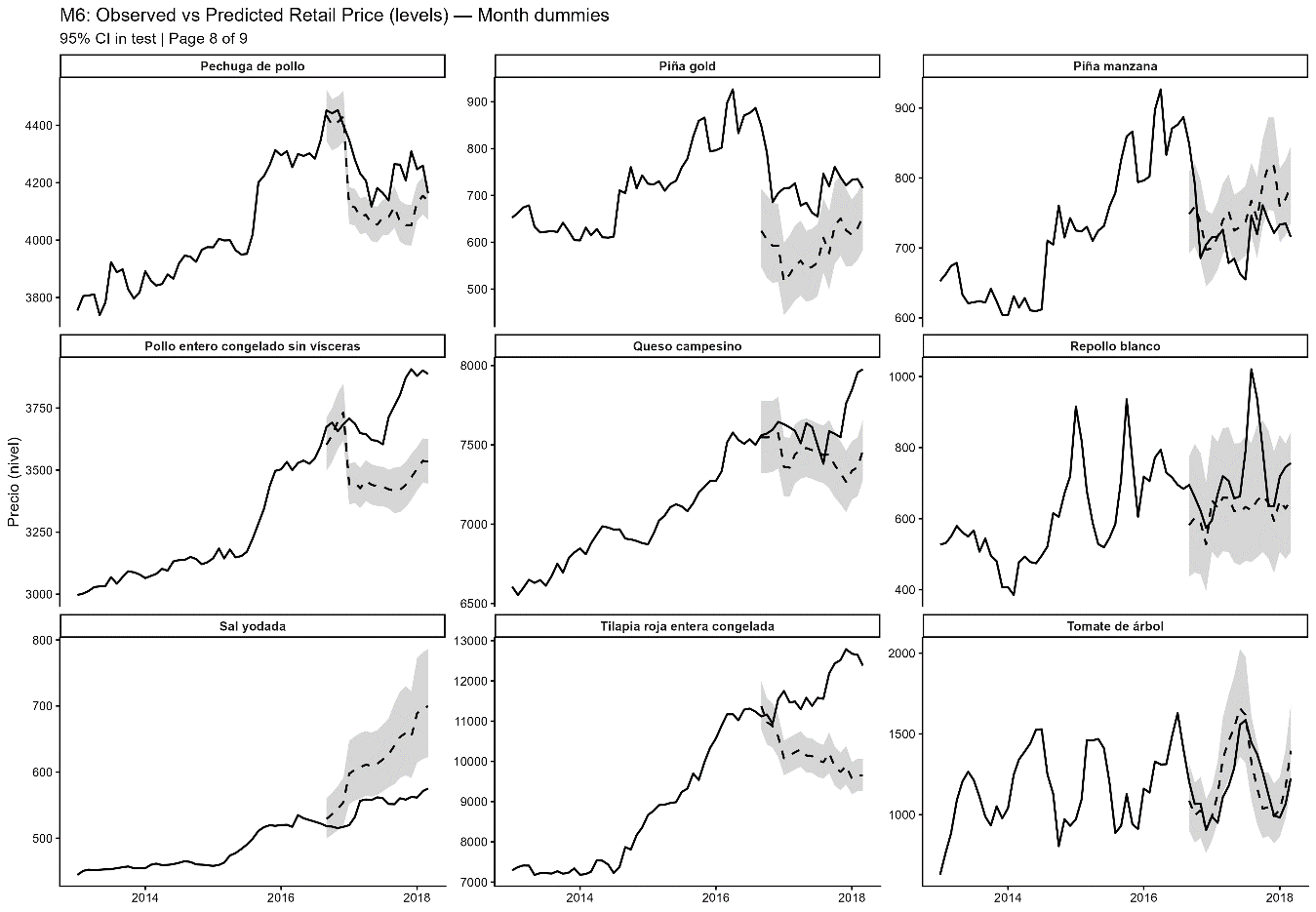
****

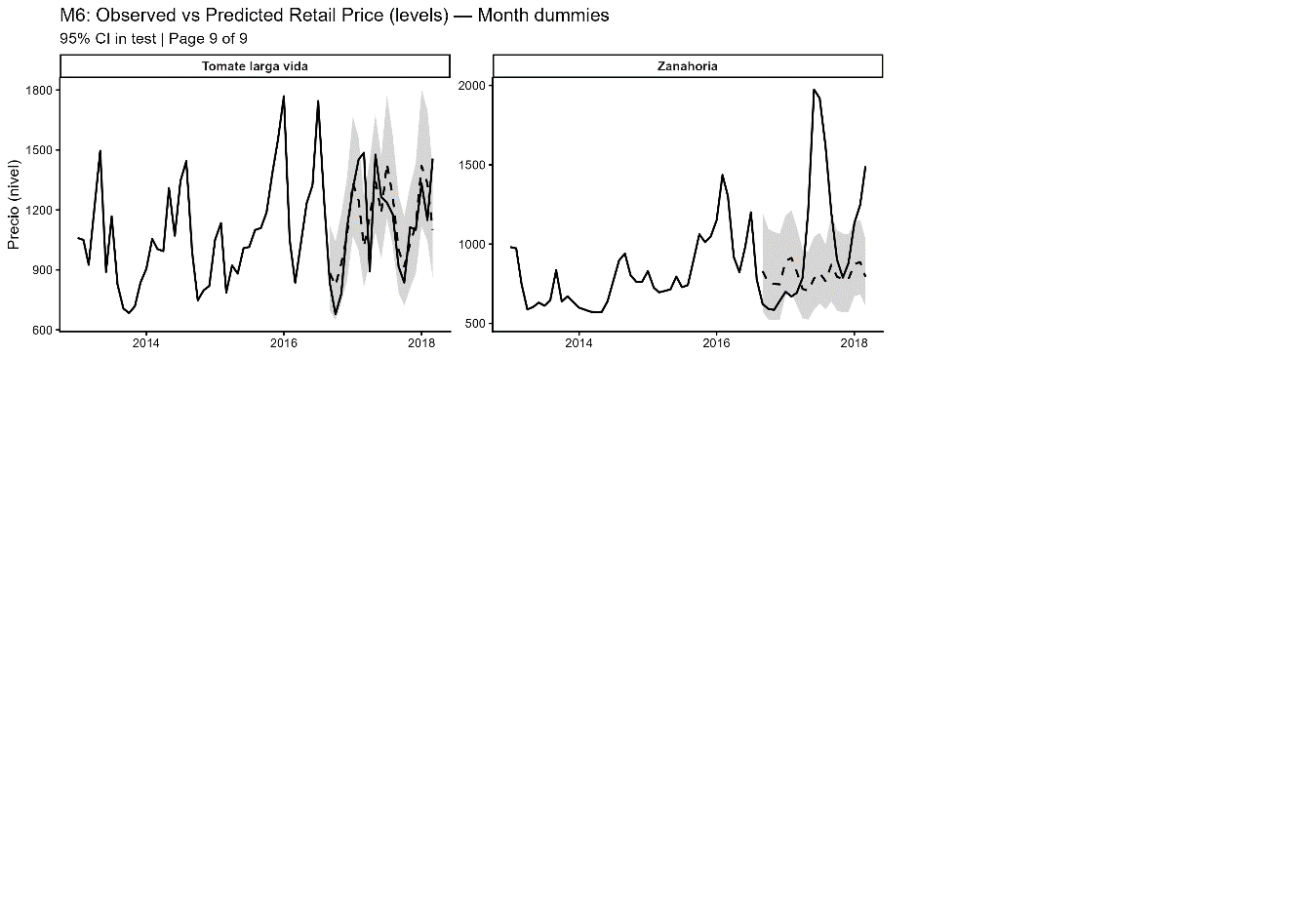
****

****

****

****

****

****

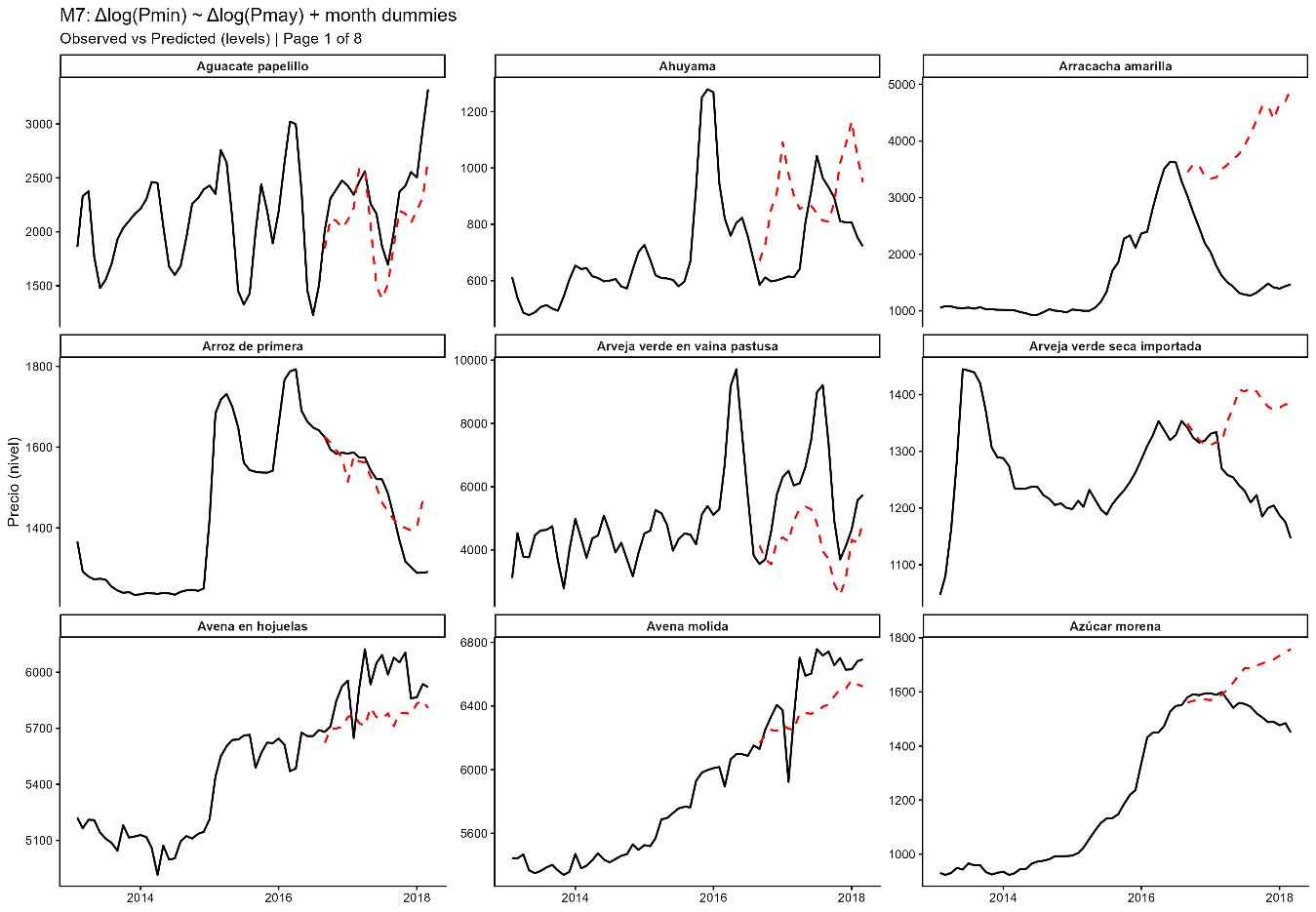
**Metodología IV: Regresión lineal en niveles y primeras diferencias**

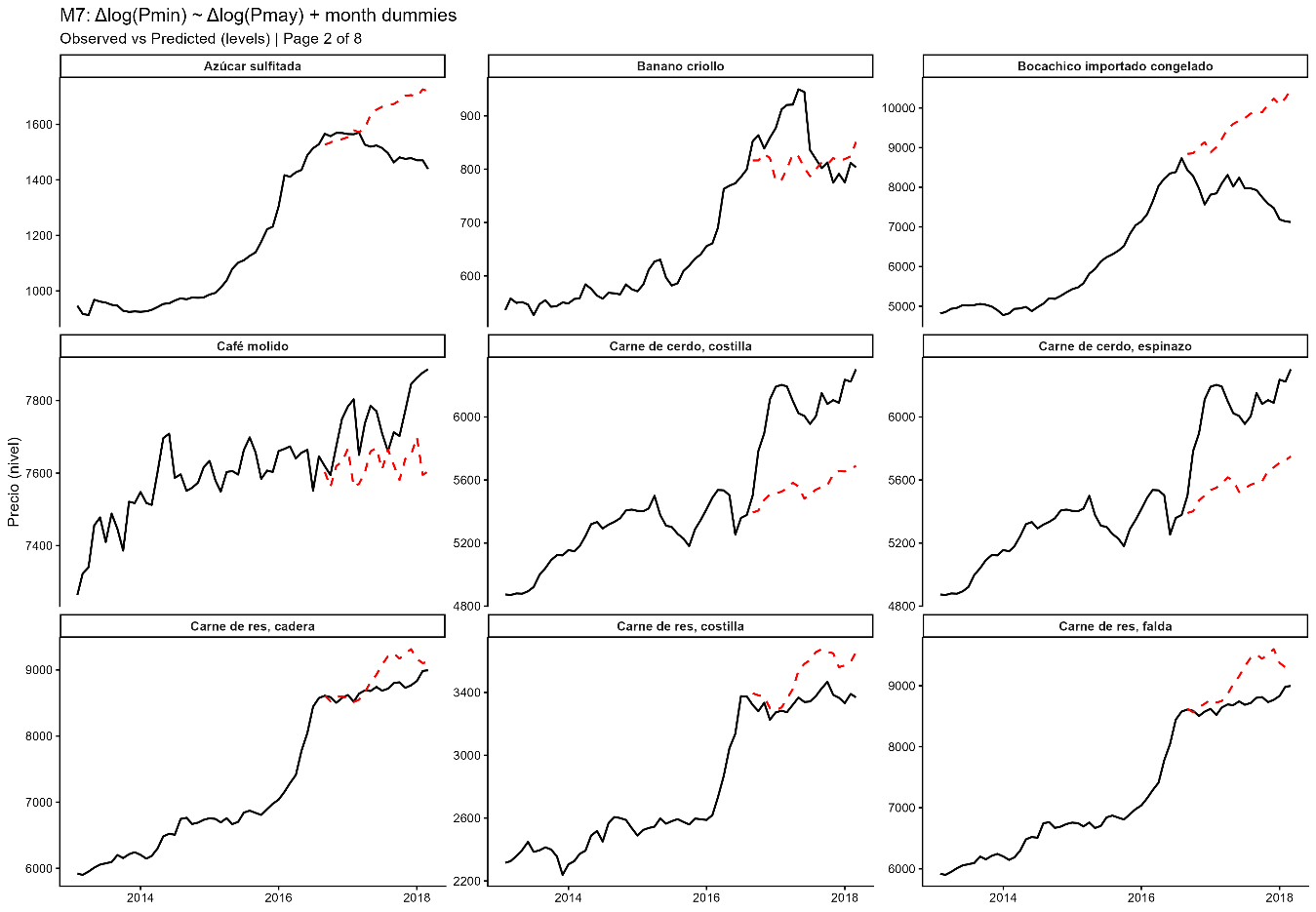
A partir de pruebas ADF, se verificó que los precios de los alimentos corresponden a series I(1). La estimación en niveles puede verse afectado por problemas de no estacionariedad y correlaciones espurias. Para garantizar series covarianza-estacionarias, se estima la siguiente especificación:

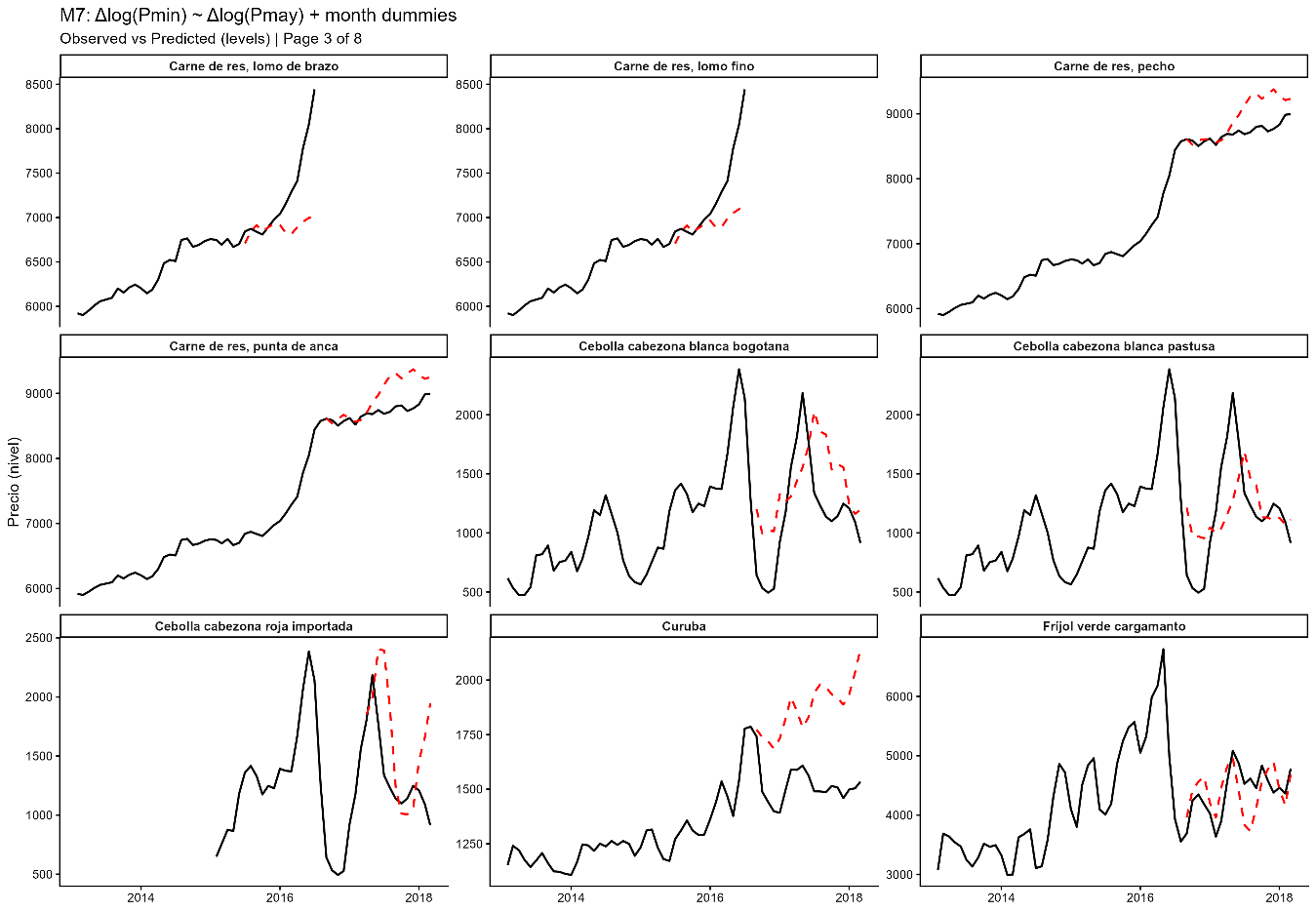
En este caso, captura la respuesta inmediata del crecimiento del precio minorista ante cambios en el precio mayorista, es decir, captura el *pass-through* de corto plazo.

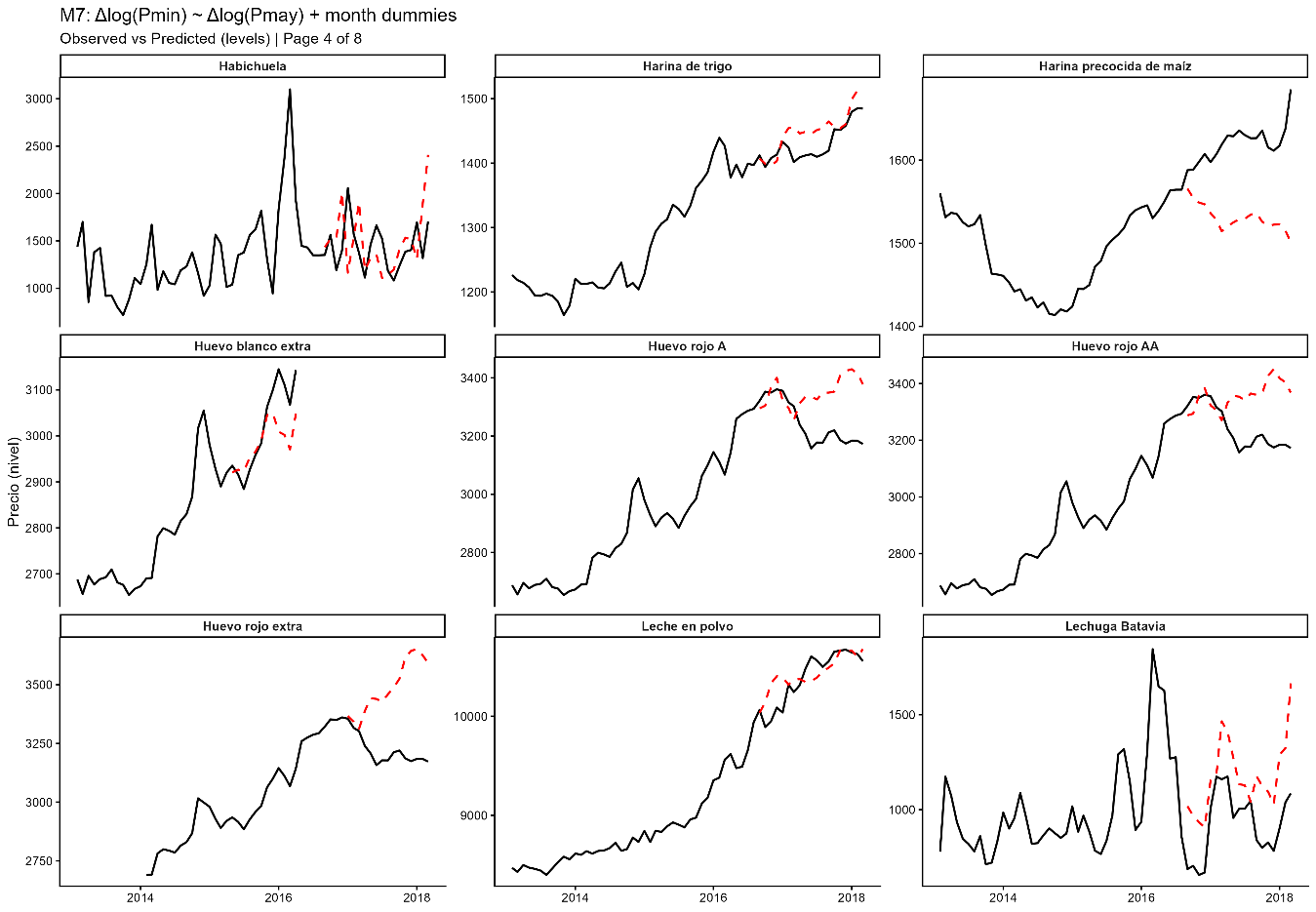
**Resultados:**

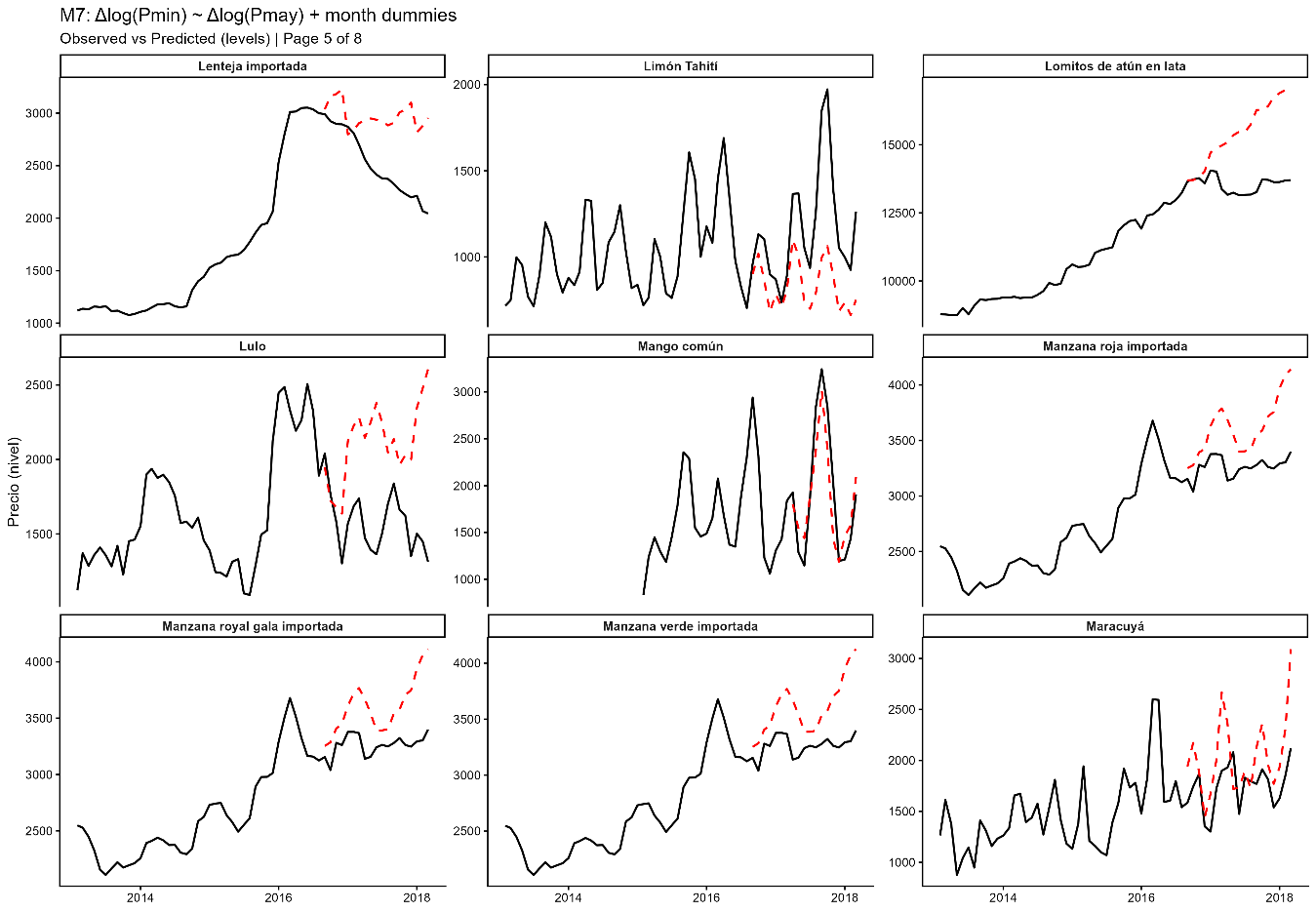
**C:\Users\Portatil\Desktop\Least-cost-diets-and-affordability\Proyecto Interno\working-papers\working-paper-1225\m7\** **output\_dummies**

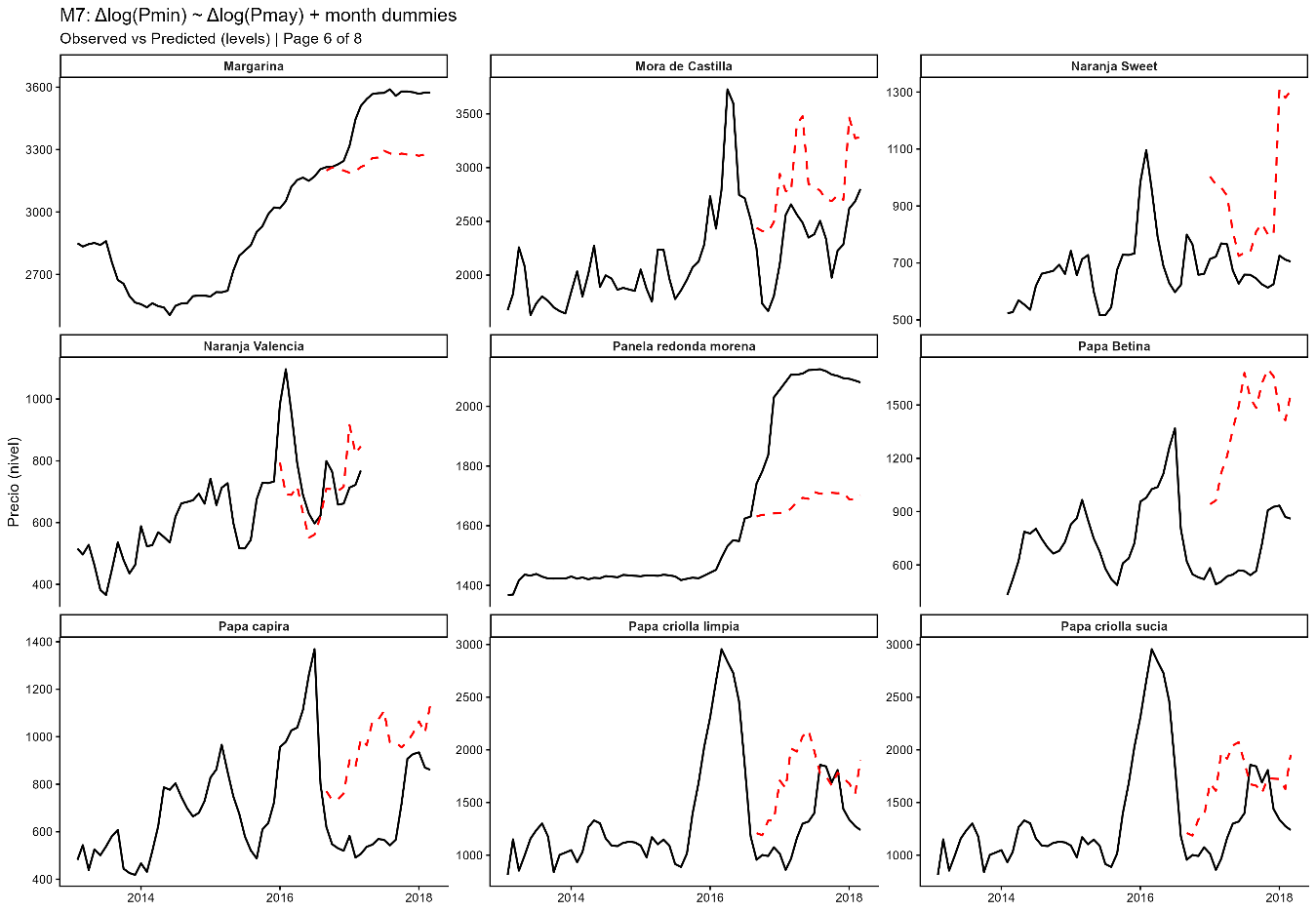
****

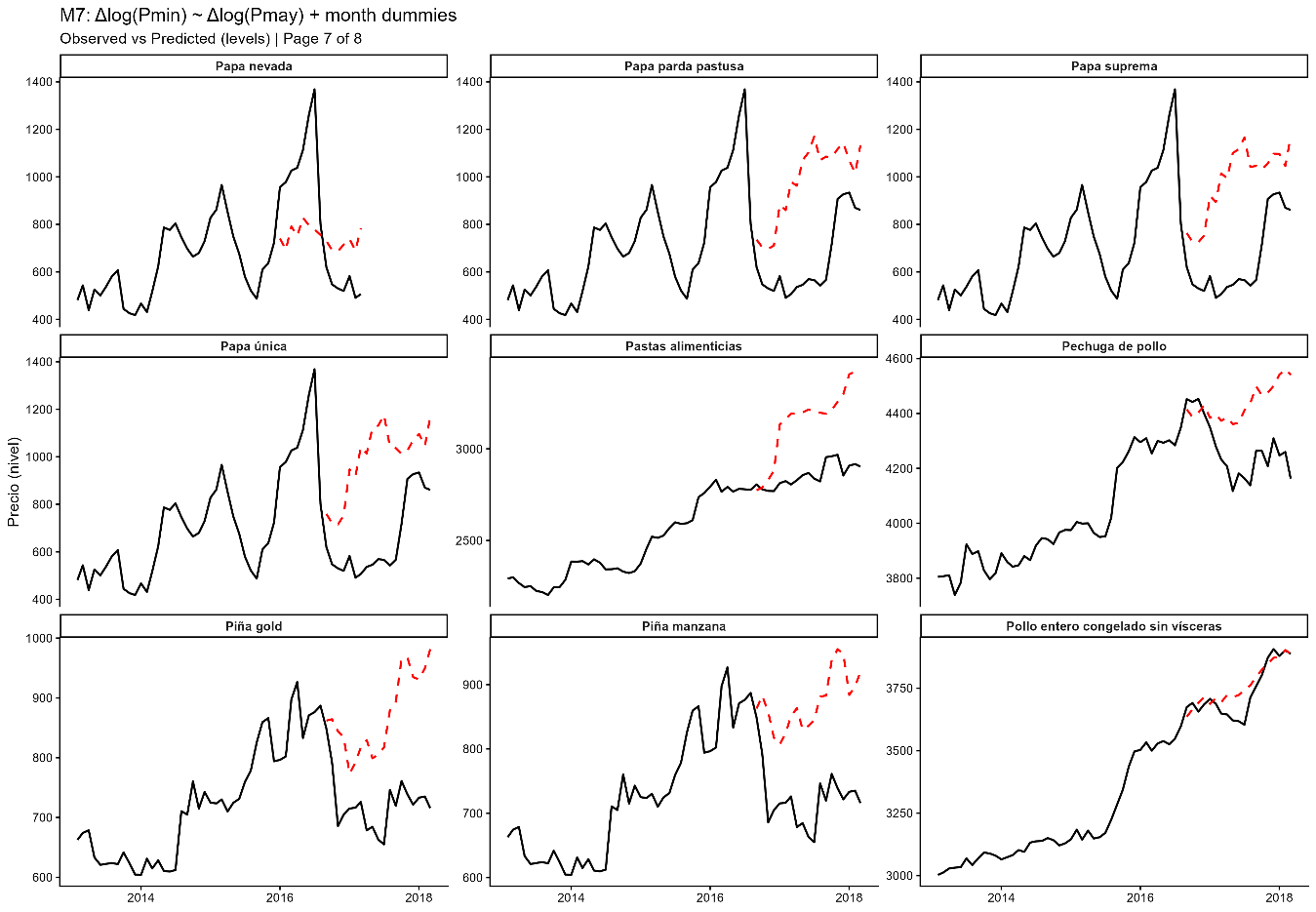
****

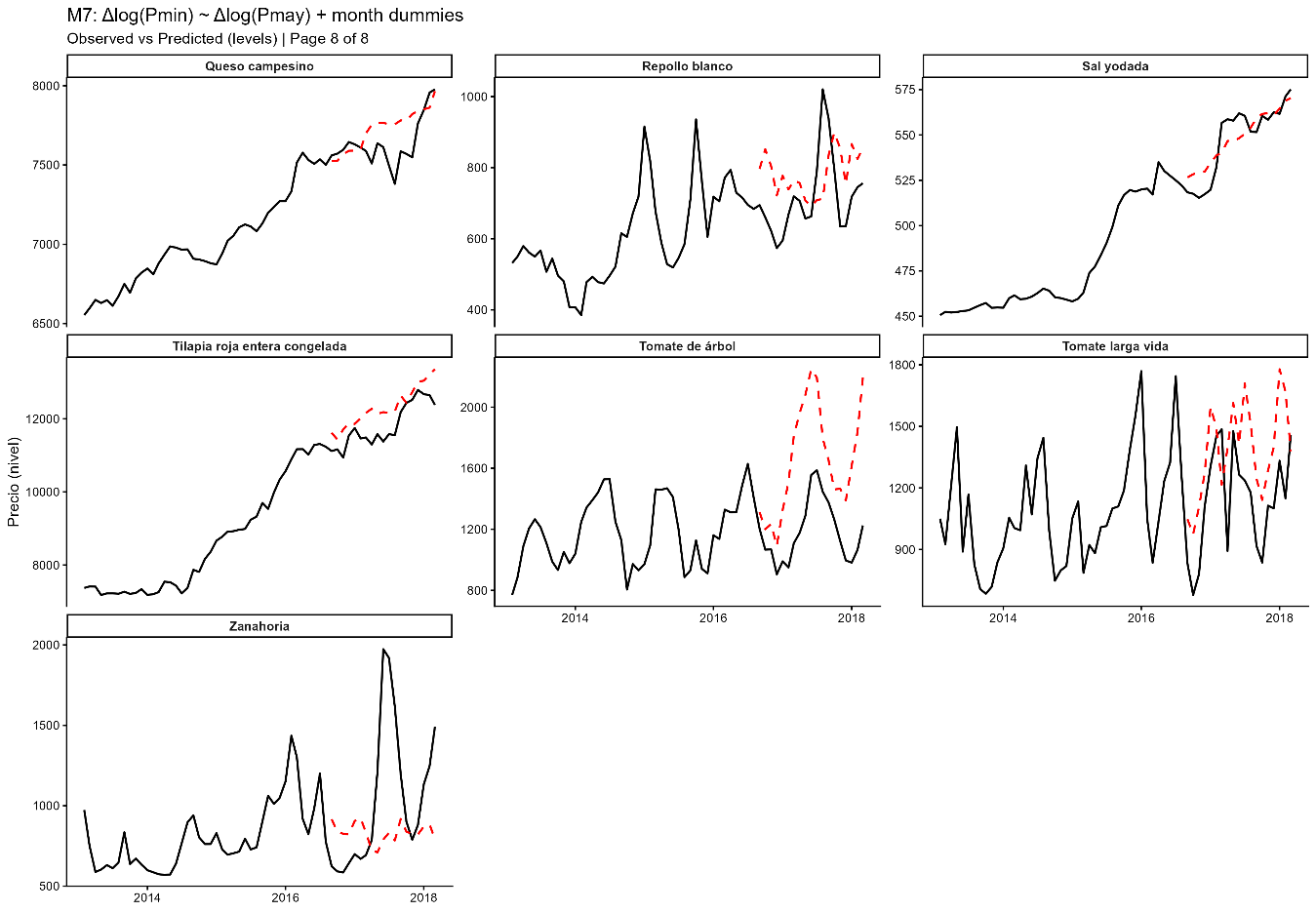
****

****

****

****

****

****

* 1. **Metodología V: Modelo de Corrección de Error (ECM)**

Como señala Enders (2014), una característica de las variables cointegradas es que su trayectoria temporal está determinada por la desviación respecto del equilibrio de largo plazo. Esto implica que las dinámicas de corto plazo deben estar influenciadas por su desviación respecto de la relación de largo plazo. El modelo dinámico utilizado para capturar simultáneamente la relación de largo plazo y los ajustes dinámicos de corto plazo es el modelo de corrección de error (ECM).

La relación de equilibrio de largo plazo entre ambos precios viene dada por:

¿Se incorpora el componente estacional?

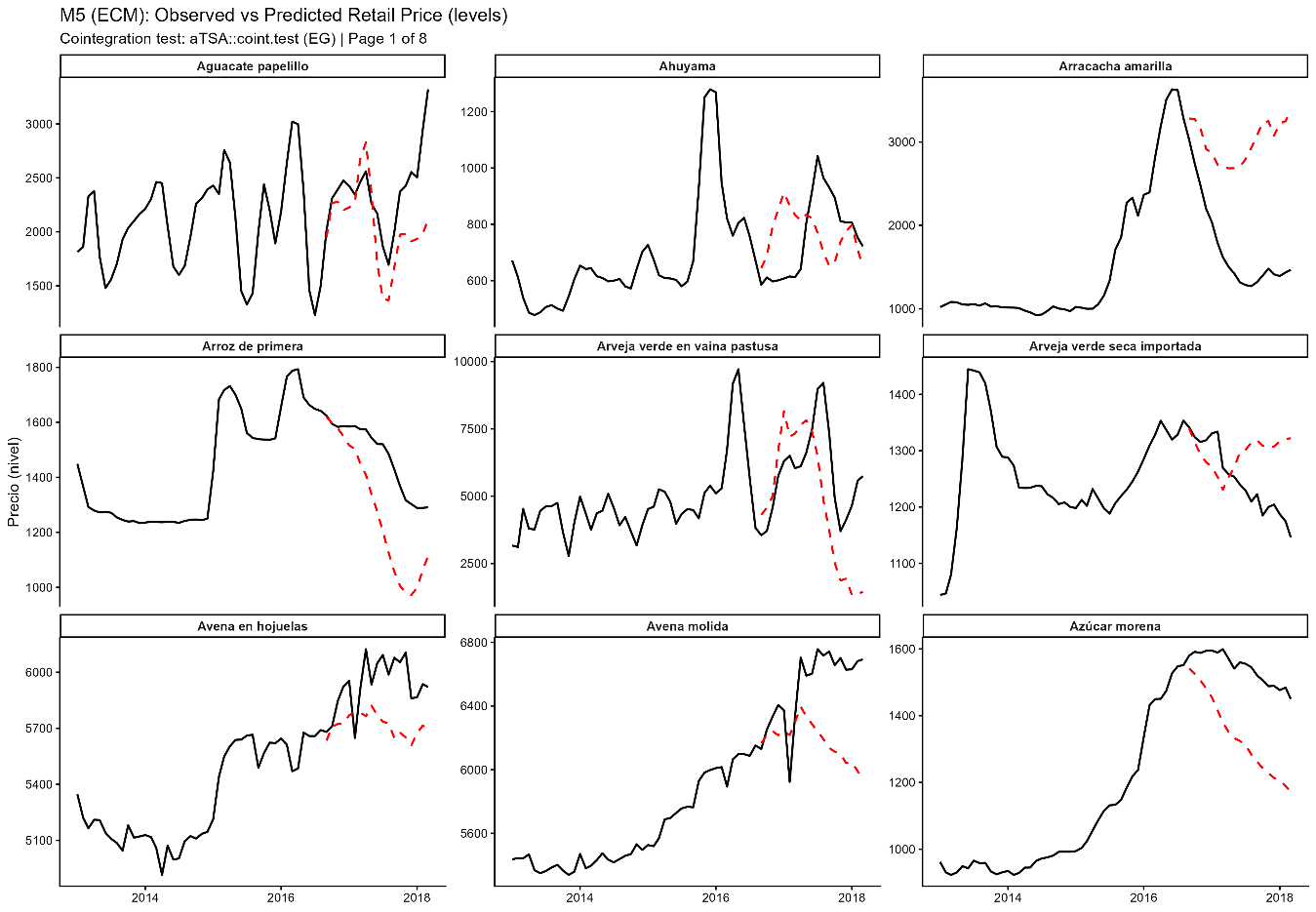
Así, el término de corrección de error corresponde al residuo rezagado de la relación de cointegración:

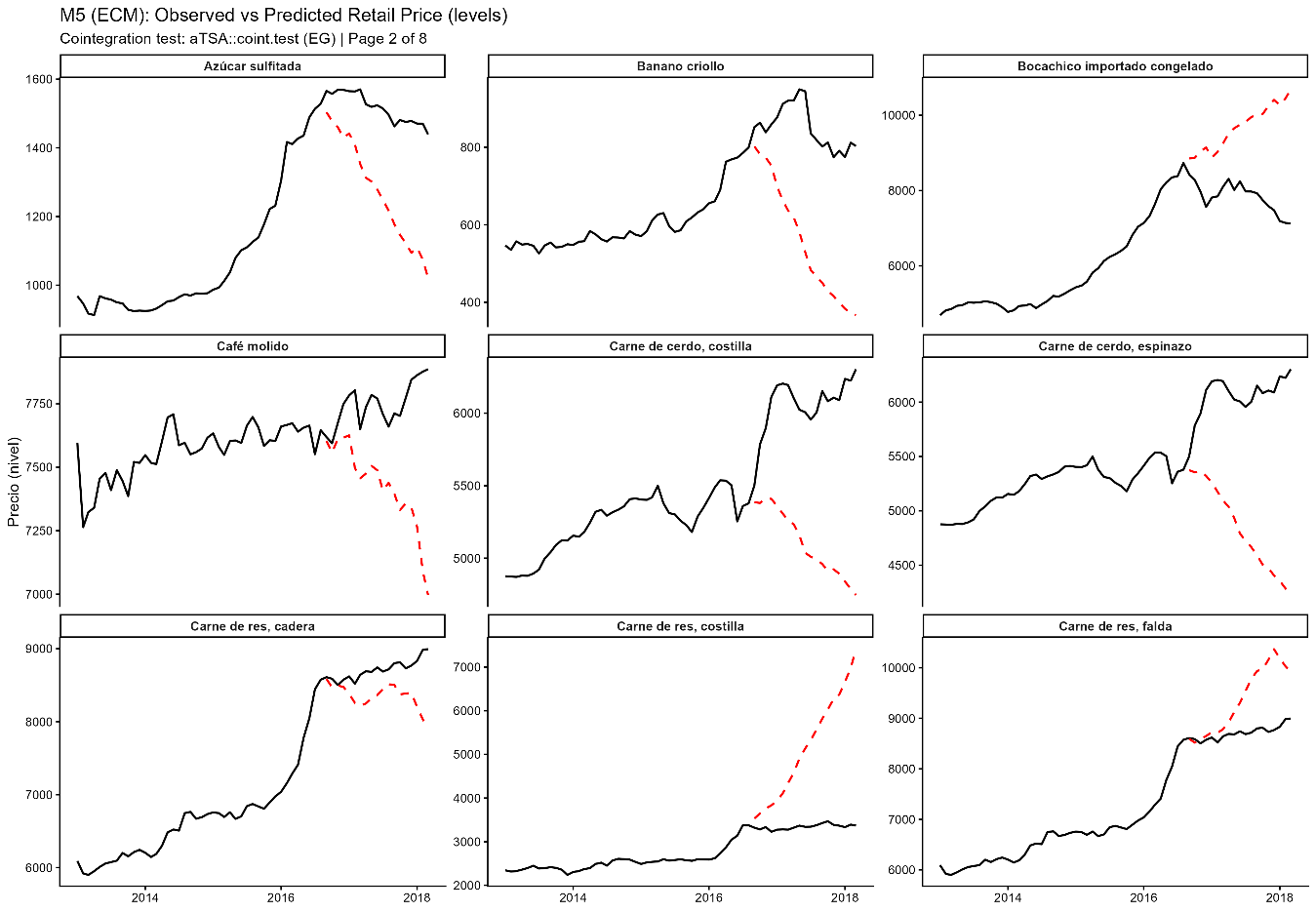
Note que la expresión anterior captura el desequilibrio en t-1. Si , el precio minorista está por encima del nivel consistente con el equilibrio de largo plazo (dado el precio mayorista). En consecuencia, la dinámica de corto plazo se captura mediante la siguiente ecuación ECM:

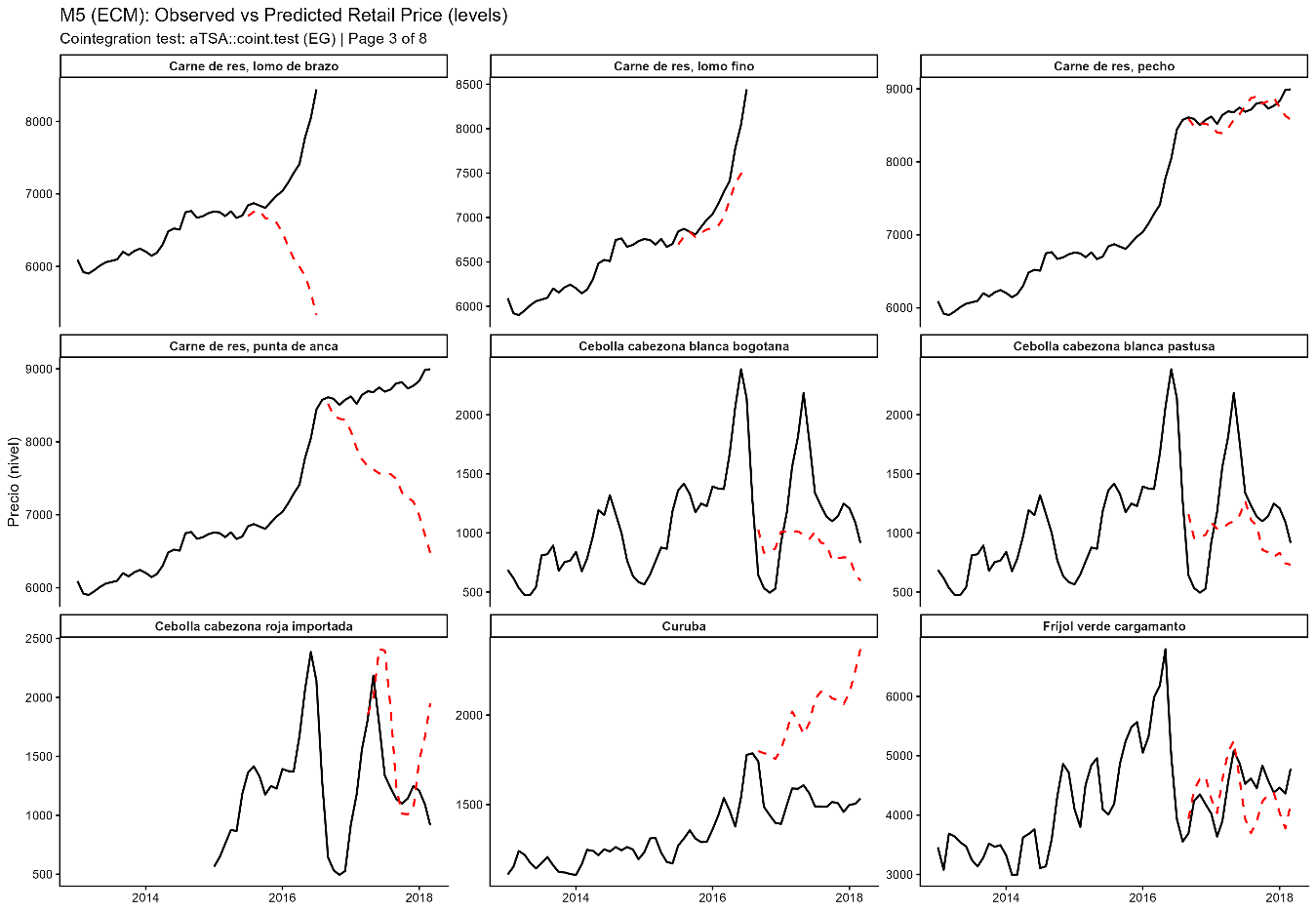
Naturalmente, captura la dinámica autorregresiva de corto plazo del precio minorista, y captura el impacto de corto plazo de variaciones en el precio mayorista. En general, es una medida de la velocidad de ajuste hacia el equilibrio de largo plazo. (Se espera que , es decir, que desviaciones positivas sean corregidas mediante reducciones en el crecimiento del precio minorista).

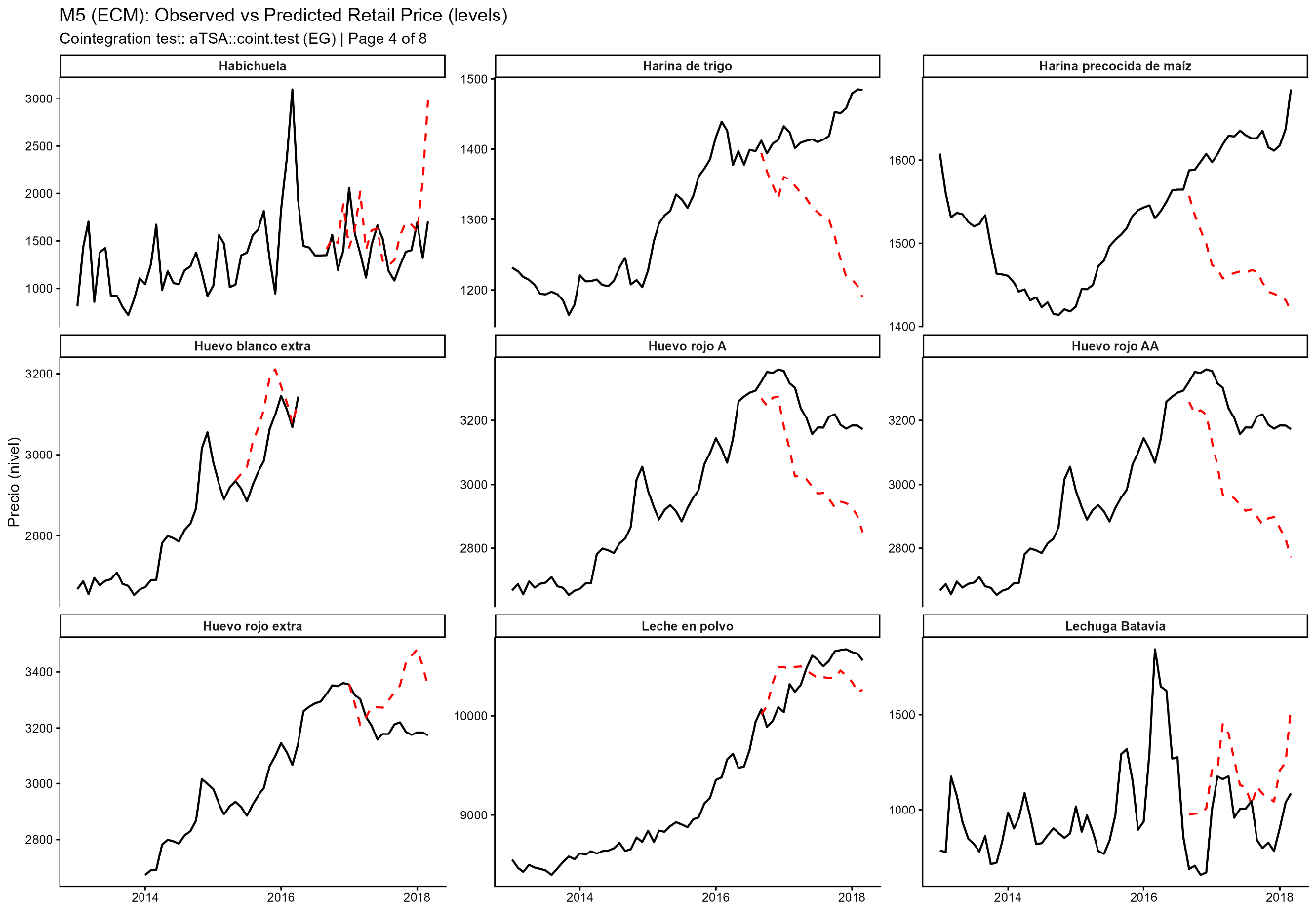
**Resultados:**

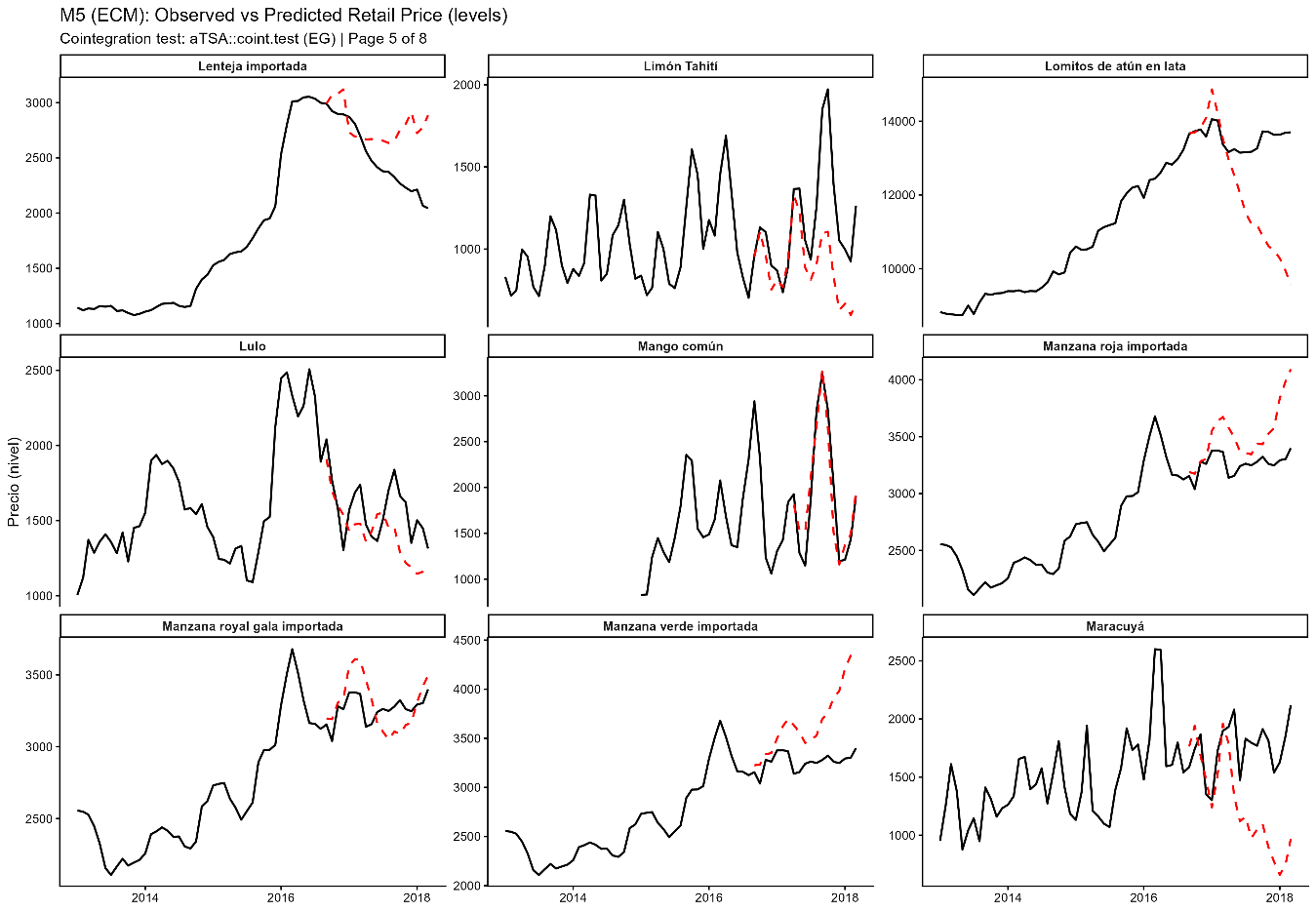
**C:\Users\Portatil\Desktop\Least-cost-diets-and-affordability\Proyecto Interno\working-papers\working-paper-1225\m5\** **output\_dummies**

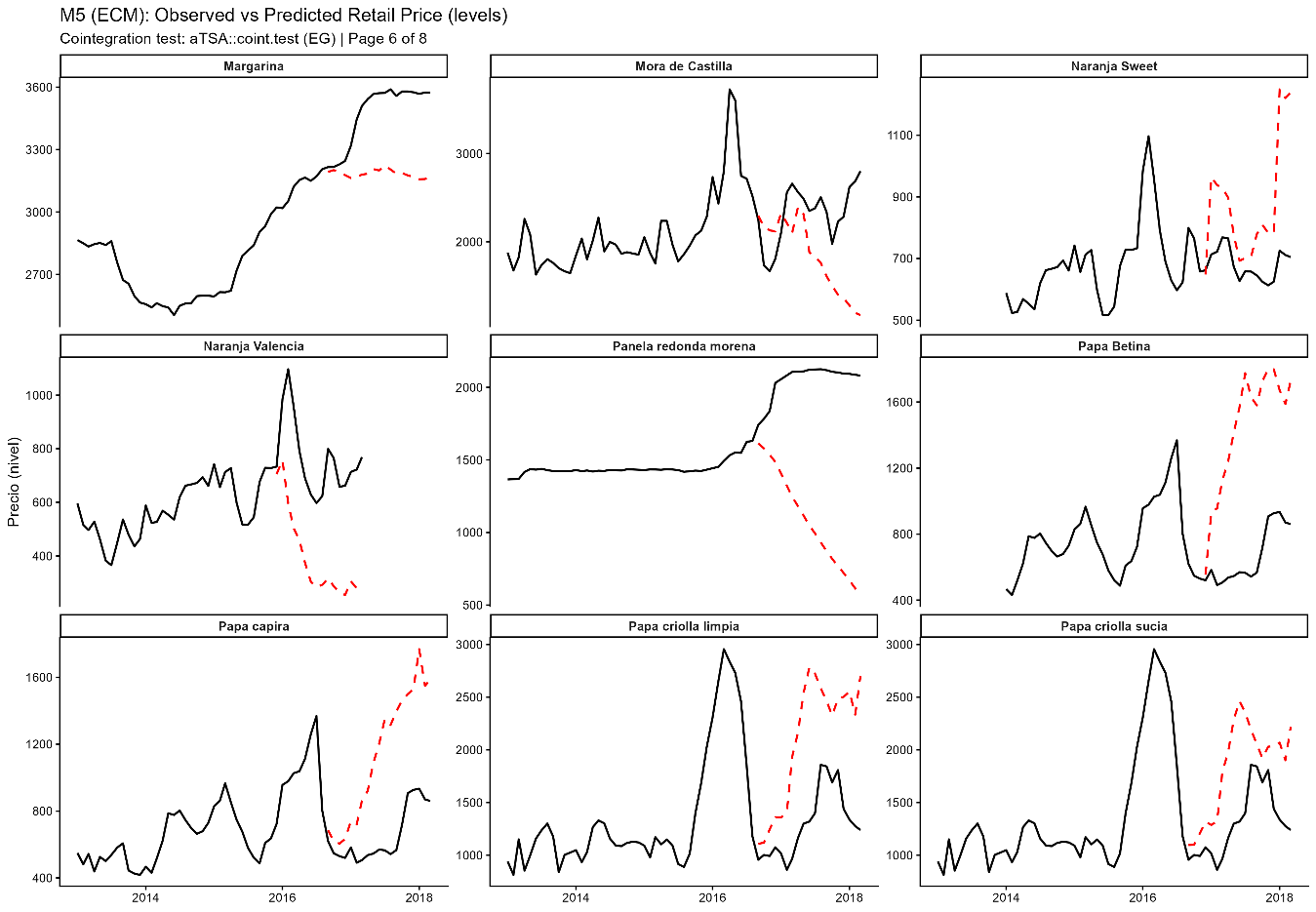
****

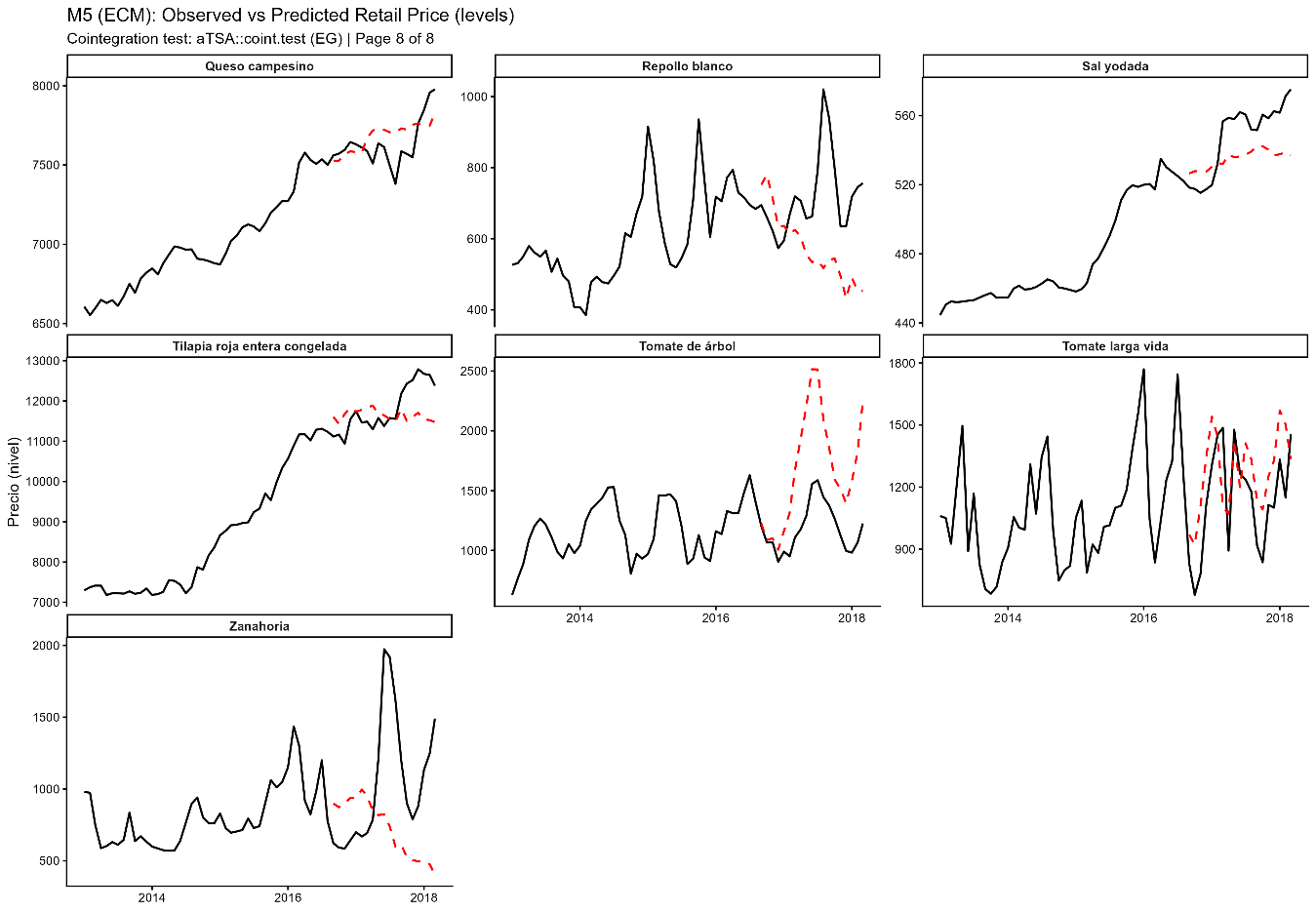
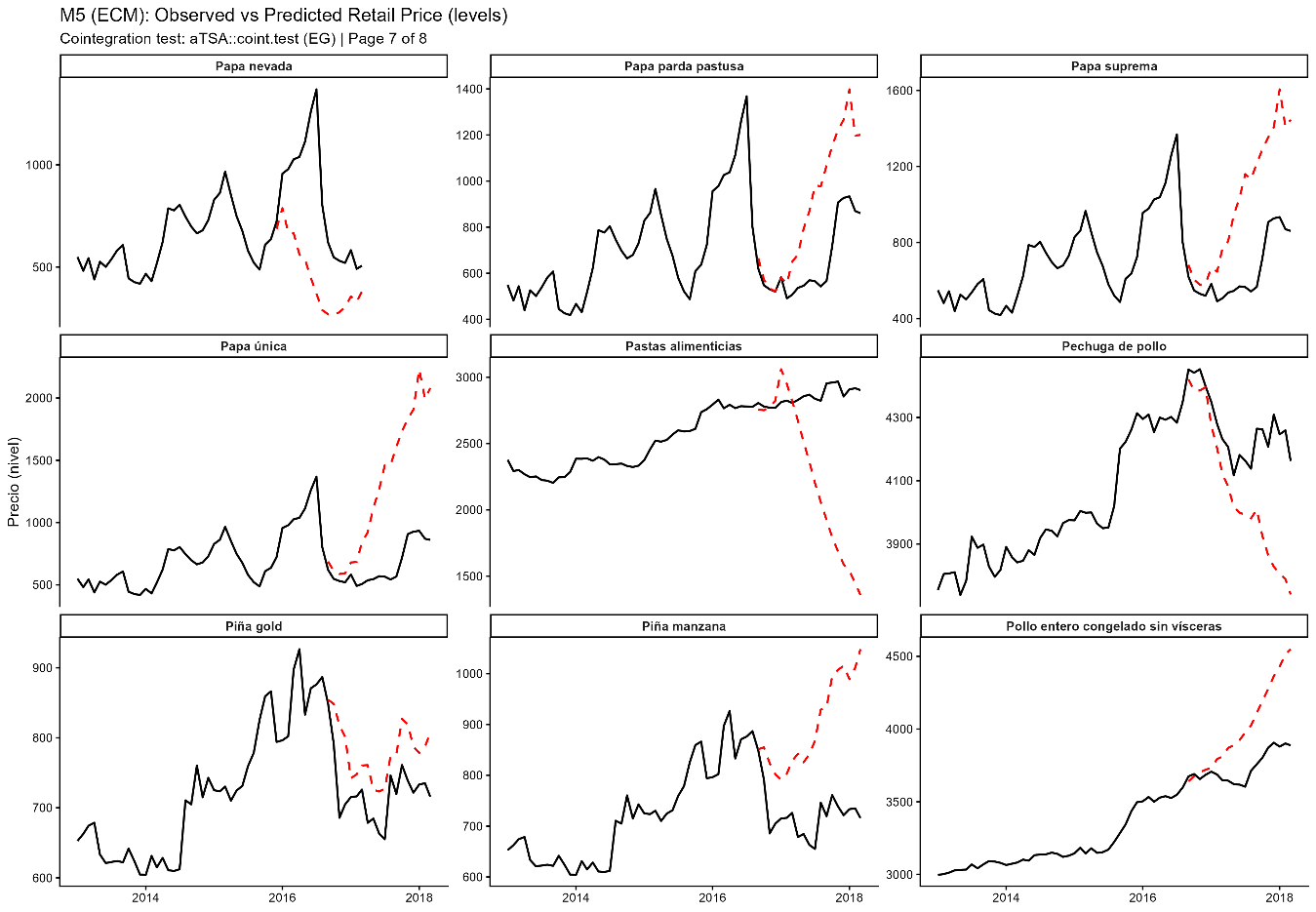
****

****

****

****

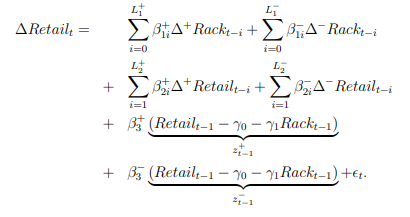
****

****

* 1. **Metodología VI: Modelo de Corrección de Error (A-ECM) Asimétrico**

En tercer lugar, se estima una versión del modelo ECM, de tal forma que se estima un modelo individual para cada ciudad principal. Siguiendo el estudio de Chesnes (2010), se propone la implementación de un modelo ECM para capturar las posibles asimetrías en la velocidad y el patrón de ajuste ante aumento y reducciones en el precio mayorista.

Sea el precio minorista observado en la ciudad de análisis en el período t, y sea el precio mayorista correspondiente. El modelo sigue la forma:



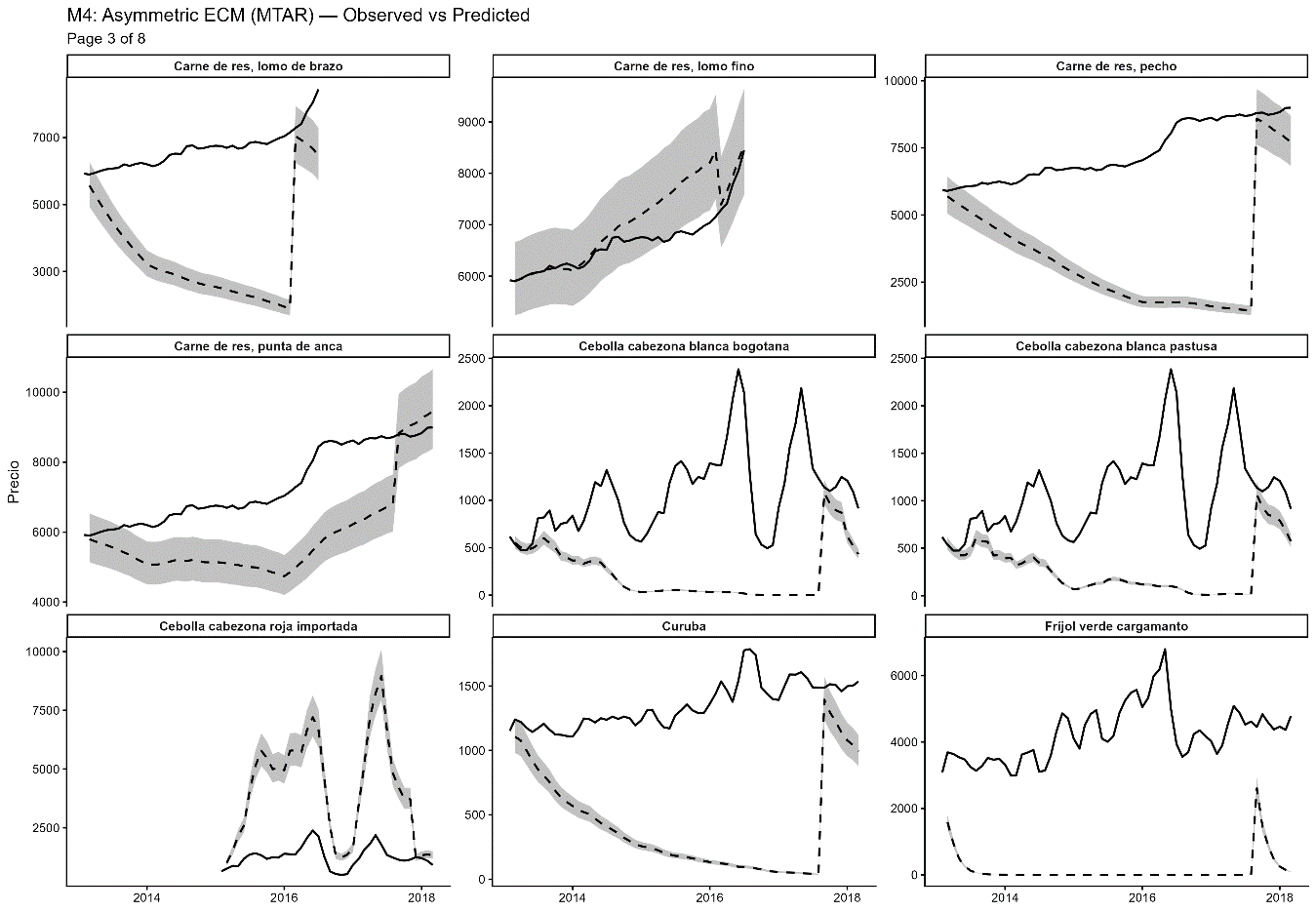
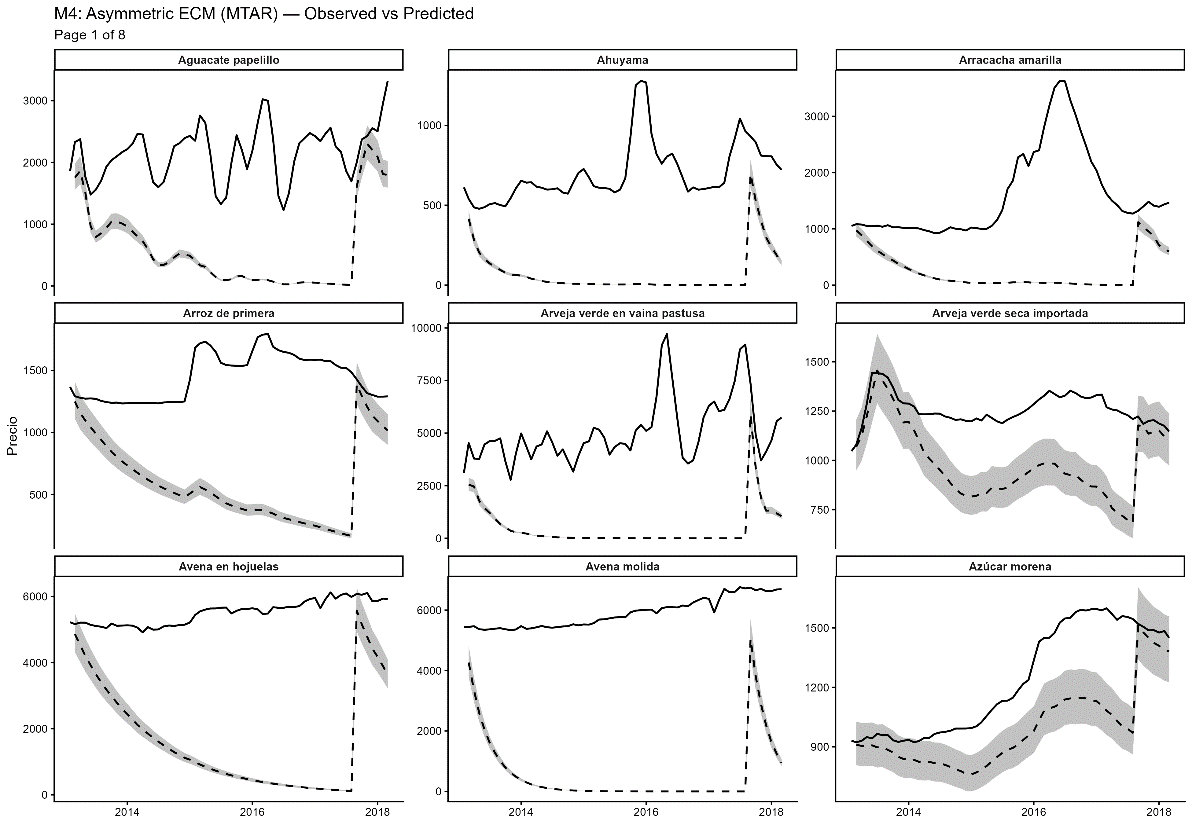
La expresión corresponde al término de corrección de error y captura la relación de largo plazo entre el precio minorista y el precio mayorista. Así, se espera que ambos sean negativos: si el precio minorista está por encima () (por debajo [) del equilibrio, los precios minoristas deberían disminuir (aumentar).

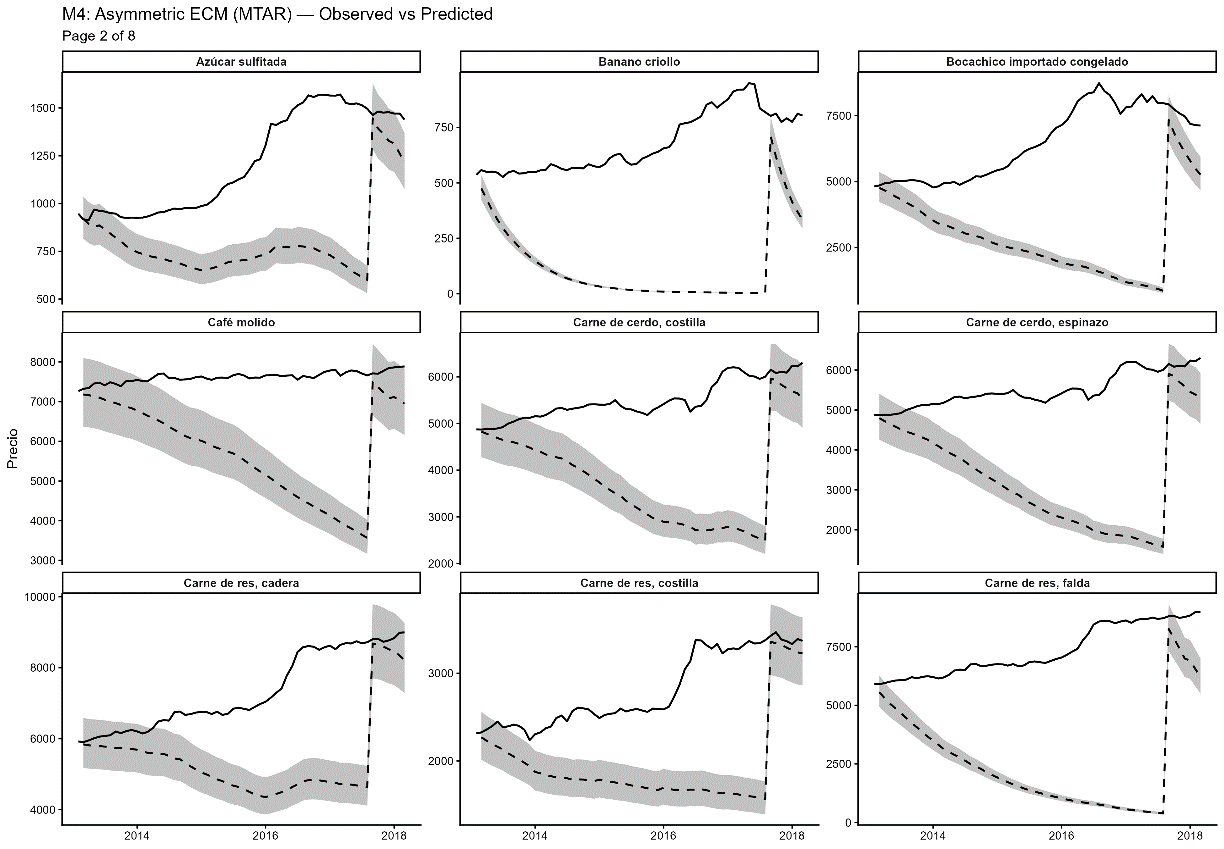
Siguiendo la metodología en dos etapas de Engle & Granger (1987), se estima y a partir de la siguiente ecuación de largo plazo:

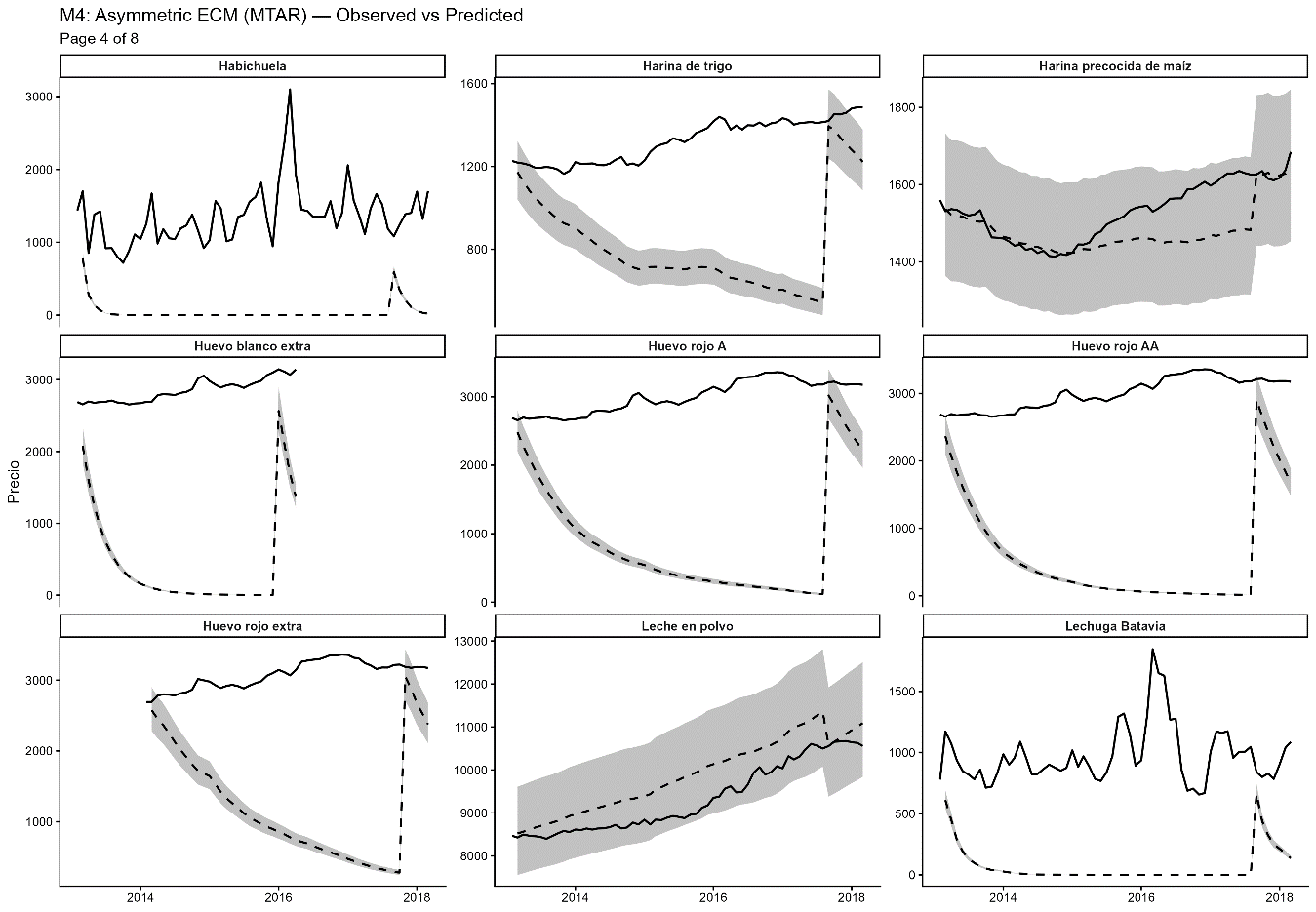


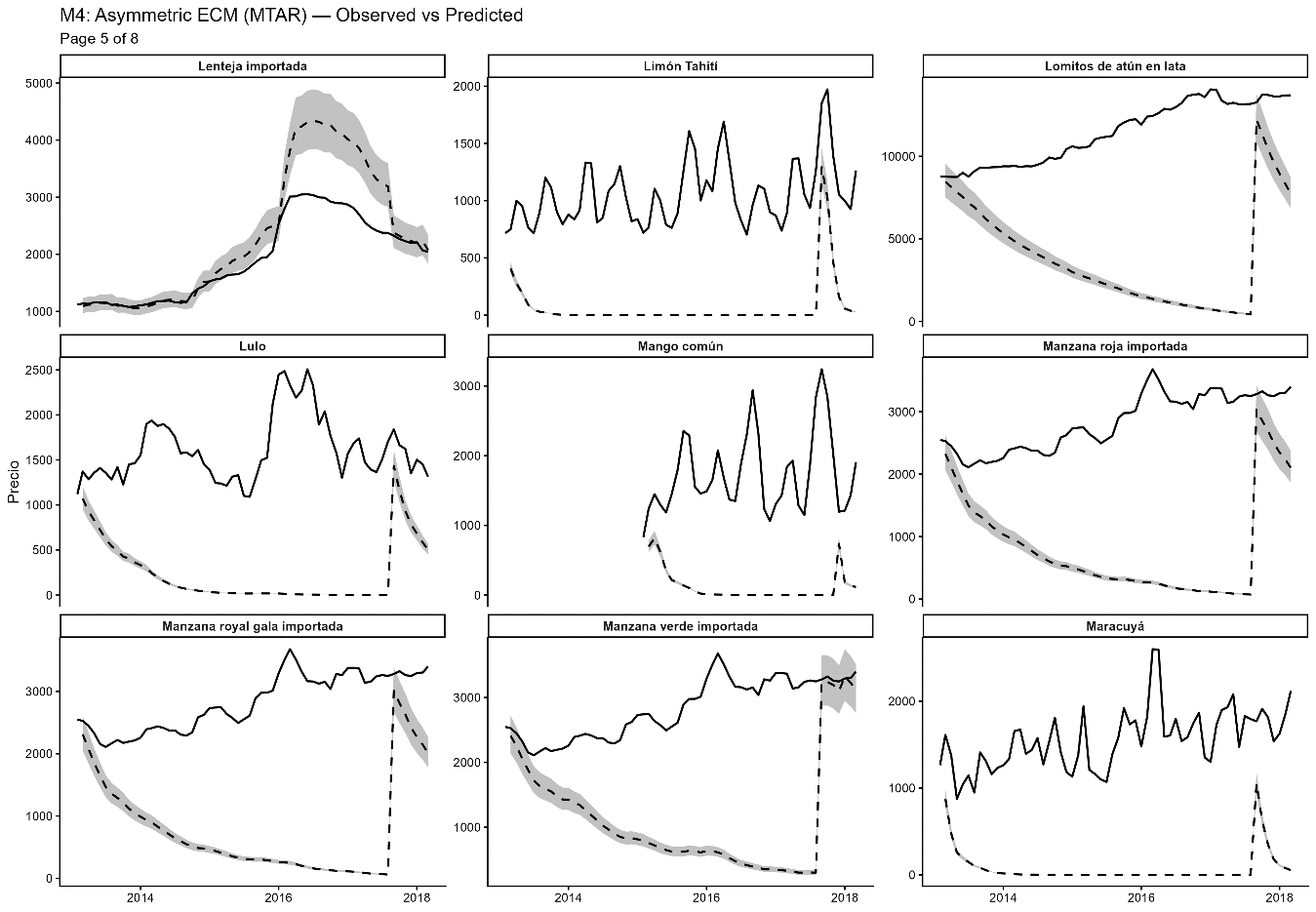
Los residuales son incluidos en el modelo ECM.

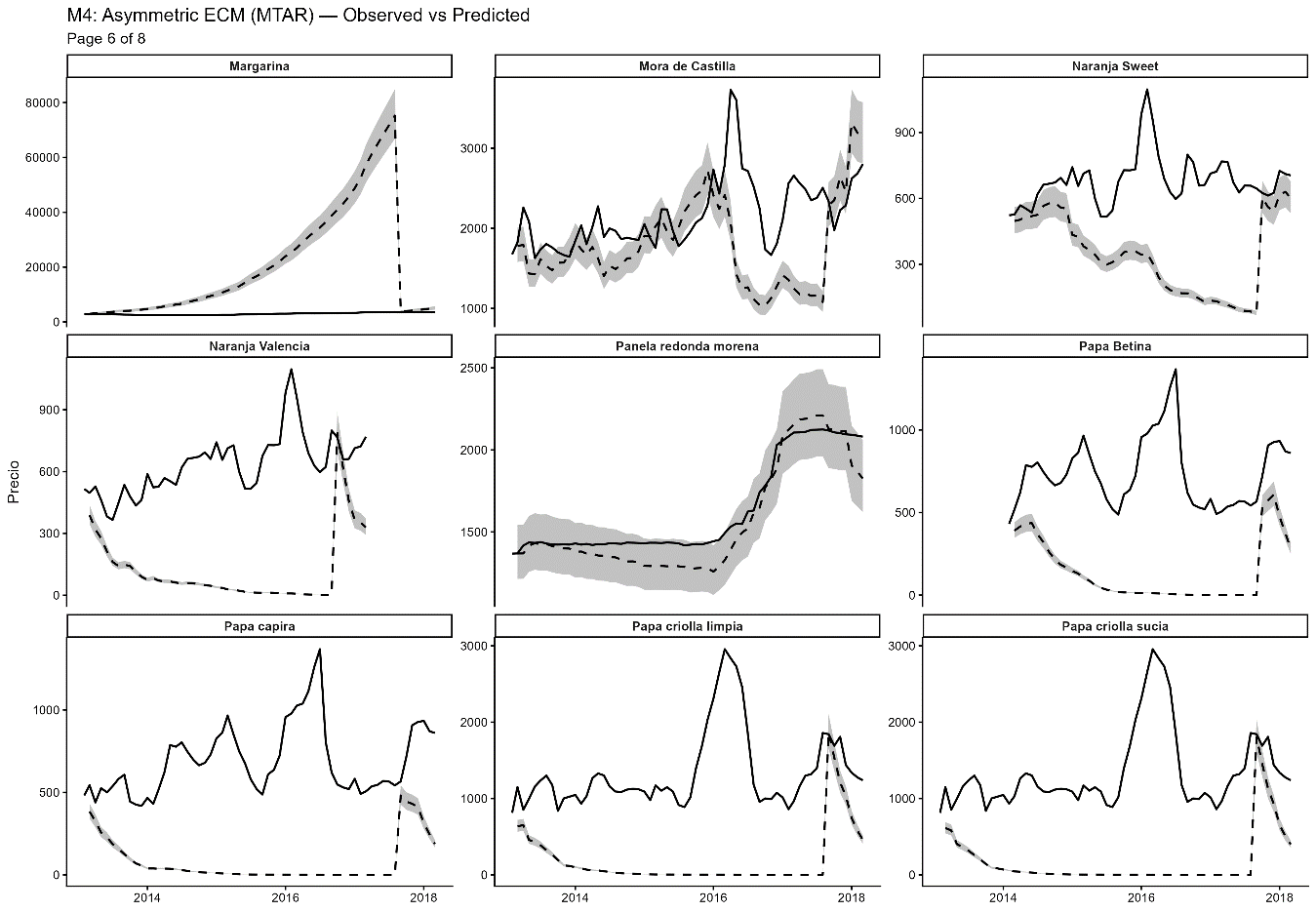
**Resultados:**

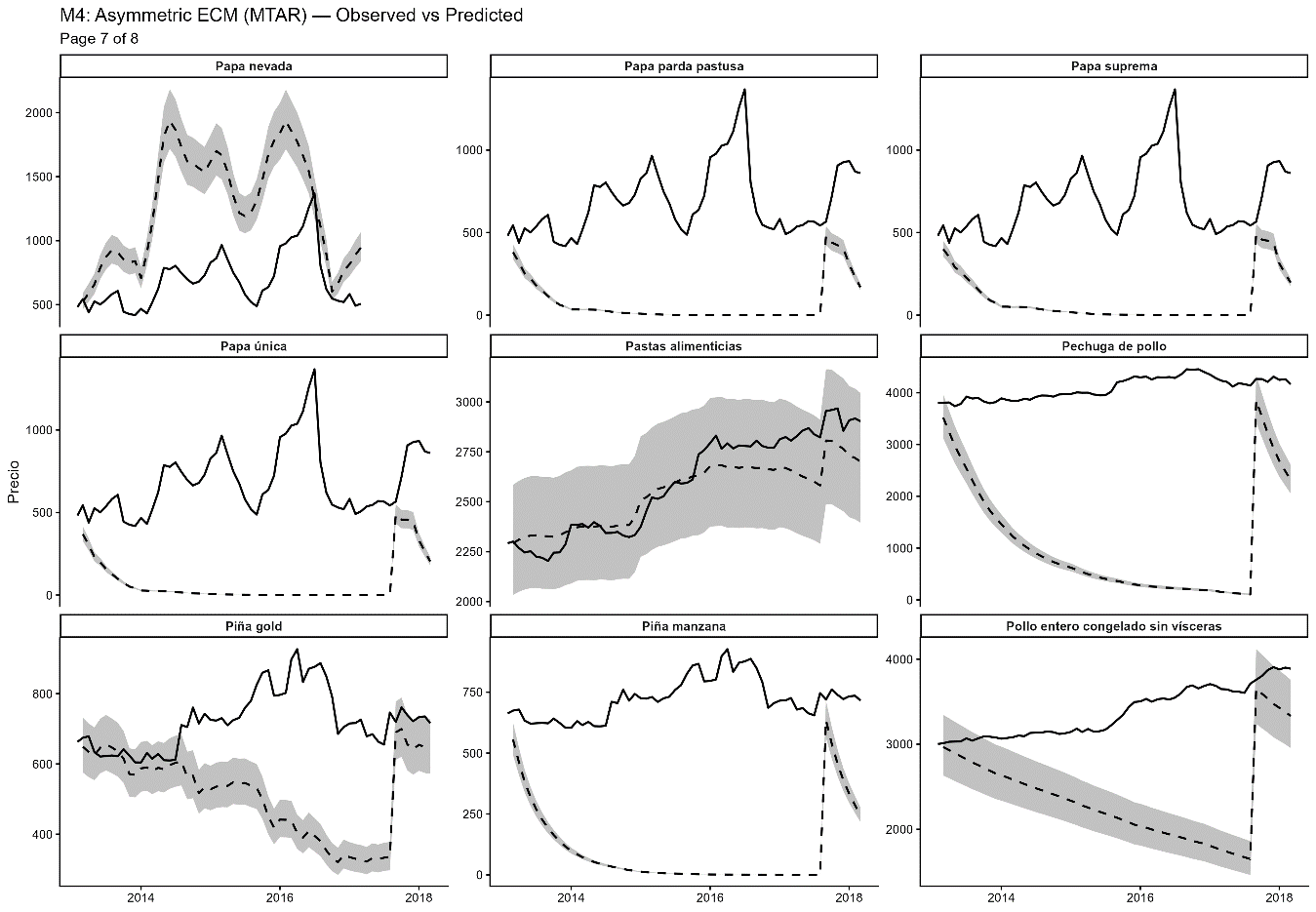
**C:\Users\Portatil\Desktop\Least-cost-diets-and-affordability\Proyecto Interno\working-papers\working-paper-1225\m4\** **output**

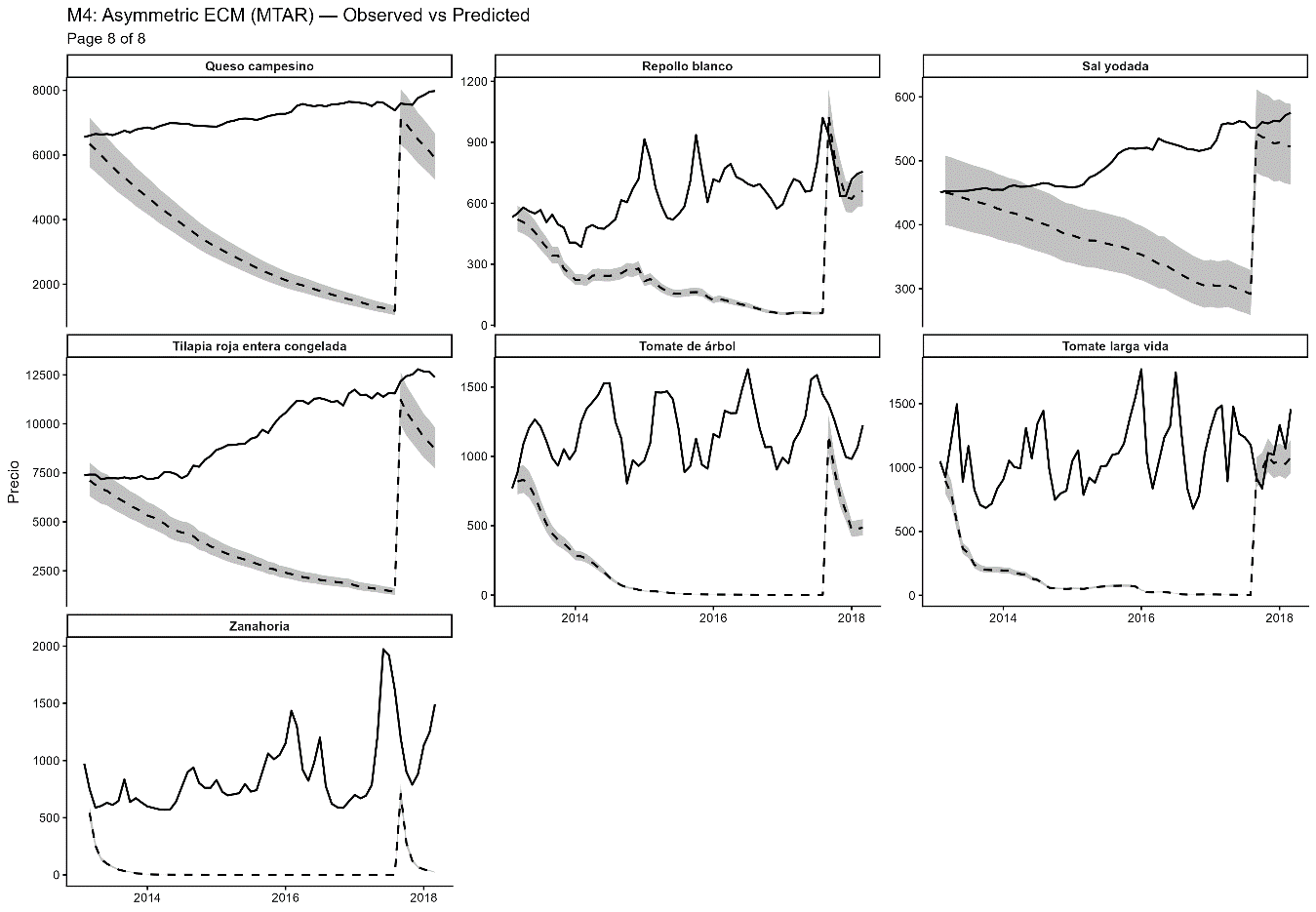












1. **Discusión**
   1. **Métricas de validación**

La siguiente tabla muestra las métricas obtenidas a partir de cada metodología:

* 1. **Márgenes de comercialización implícitos**

Note que algunas de las metodologías implementadas permiten obtener un margen de comercialización implícita

* 1. **Validación con márgenes implícitos**

1. **Conclusiones**