
TESIS DE MAESTRÍA

MAESTRÍA EN ECONOMETRÍA

UNIVERSIDAD TORCUATO DI TELLA

**ELASTICIDADES DE LARGO PLAZO
EN TRANSPORTE INTERURBANO DE
CARGA**

Alumno:

Marcelo Pérez

Director de la Maestría:

Ph.D. Martín González Rozada

Febrero 2016

Elasticidades de largo plazo en transporte interurbano de carga.

Marcelo Pérez¹

¹ e-mail: marceloperez.econ@gmail.com

CONTENIDO

1.	Introducción.....	3
2.	Modelo Teórico	5
3.	Estrategia de estimación.....	7
A)	Modelos estáticos.....	9
B)	Modelos Estáticos con Correlación.....	12
C)	Modelos con Parámetros Heterogéneos.....	16
D)	Modelos Dinámicos	18
E)	Modelos de Corrección del Error.....	18
4.	Análisis de datos	20
4.1.	<i>Estadísticas Descriptivas</i>	21
4.2.	<i>Análisis Estadístico de las Series</i>	27
4.2.1.	<i>Pruebas de Raíz Unitaria</i>	27
4.2.2.	<i>Cointegración</i>	32
4.2.3.	<i>Dependencia</i>	37
4.2.4.	<i>Heterogeneidad</i>	38
4.3.	<i>Variables Explicativas</i>	39
5.	resultados	41
6.	Conclusiones	48
	Bibliografía.....	49
	ANEXO.....	54

1. INTRODUCCIÓN

El trabajo de investigación, tiene por objetivo disponer de estimaciones consistentes tanto de la elasticidad precio de la demanda, como de la elasticidad ingreso y la elasticidad precio de bienes complementarios para el transporte inter urbano de carga en Uruguay.

La infraestructura de transporte representa para muchos países el principal rubro de inversión pública, con una particular importancia en los niveles de productividad de la economía y por ende en las tasas de crecimiento de largo plazo del PIB. Asimismo, la provisión de infraestructura de transporte está fuertemente influenciada por el nivel de demanda futura que se espera, en la que debido al carácter discreto de la oferta² se puede incurrir en una sub provisión de capacidad generando por tanto congestión, o en una sobre provisión que implicaría la utilización en forma inadecuada de factores productivos. Cualquiera de las dos situaciones impone altos costos de ineficiencia para la sociedad.

En virtud de lo expuesto anteriormente, el estudio de la demanda de transporte ha sido un tópico de singular relevancia y en donde se han realizado desarrollos y aplicaciones que han trascendido el área de economía del transporte³. Se han utilizado tanto técnicas basadas en preferencias declaradas como relevadas. No obstante, la sistematización rigurosa de información de los últimos tiempos ha permitido incorporar modelos basados en datos históricos que resultan adecuados en contraposición con los modelos de preferencias declaradas que dependen de situaciones hipotéticas.

La teoría económica indica que la demanda de transporte es función de lo que se conoce en la literatura como precio generalizado del transporte (g).

En donde:

$$g = p + v \cdot t + \theta$$

p = los componentes monetarios del viaje

v = es el costo de oportunidad del tiempo

t = tiempo empleado

θ = valoración monetaria de los aspectos cualitativos

La anterior formulación del precio generalizado, impone una restricción al análisis econométrico debido a que dos de los aspectos centrales del precio generalizado son variables no observables. En este sentido, es frecuente la utilización de modelos en datos de panel como forma de solucionar el sesgo que se ocasiona por el problema de variables no observadas.

Además del problema de variables no observadas, se ha encontrado que las series de tráfico son procesos integrados de orden uno y que por tanto existe alta probabilidad de que estén cointegrados con las series de ingreso o PIB de la economía.

En este trabajo se propone en primer lugar, presentar un marco teórico adecuado para el análisis de la demanda de transporte carretero en Uruguay. Luego se realizará una revisión detallada y ordenada de la literatura existente en materia de paneles de series de tiempo (o paneles largos). En tercer lugar, se analizarán las propiedades estadísticas de los procesos estocásticos de las series de

² Es importante recalcar que la infraestructura de transporte no puede ajustarse marginalmente con el aumento de la demanda. En concreto, un aumento de la capacidad de un aeropuerto implicaría disponer de una pista de aterrizaje adicional, en una carretera de un carril adicional, etc.

³ En este sentido vale la pena mencionar los trabajos de Daniel McFadden en relación a los modelos de utilidad esperada.

tráfico utilizadas para estimar la demanda de transporte interurbano de carga. Cuarto, se estiman diferentes modelos en datos de panel y por último se detallarán las conclusiones obtenidas. Se utilizarán datos de tránsito de carga pasante por las rutas nacionales en Uruguay, tanto con pago por uso (peaje) como de flujo libre (sin peaje).

La base de datos construida corresponde al tránsito relevado en 23 puestos de conteo ubicados en distintas regiones y por tanto para distintos pares orígenes destino.

2. MODELO TEÓRICO

En relación a los determinantes del tráfico carretero, la teoría económica sugiere que las siguientes variables explican la demanda individual:

- Precio del transporte
- Precio de otros bienes y servicios (precio del combustible, costo de operación del vehículo, precio de modos alternativos)
- Características socioeconómicas (nivel de actividad, preferencias por tipo de viaje, factores que determinan la generación o atracción de viajes según peaje)
- Calidad del servicio (seguridad, estado de las rutas)
- Tiempo del viaje

Partiendo de la siguiente especificación funcional, nuestro modelo teórico podría tomar la siguiente forma:

$$Y_{it} = \beta_0 + \beta_{1i}P_{it} + \beta_{2i}PA_t + \beta_{3i}PC_t + \beta_{4i}CO_{it} + \beta_{5i}IVF_t + \beta_{6i}T_{it} + \beta_{7i}G_i + \beta_{8i}A_i + u_{it}$$

donde :

Y_{it} = Logaritmo del Tránsito en el peaje i durante el período t

P_{it} = Logaritmo del Precio del peaje en la estación i durante el período t

PA_t = Logaritmo del Precio de modos alternativos durante el período t

PC_t = Logaritmo del Precio del combustible durante el período t

CO_{it} = Logaritmo del Costo operativo del vehículo según peaje utilizado durante el período t

IVF_t = Logaritmo del Índice de volumen físico durante el período t

T_{it} = Logaritmo del Costo del tiempo en la estación de peaje i en el período t

G_i = Factor de generación de viajes del peaje i

A_i = Factor de atracción de viajes del peaje i

Debido a la falta de datos en relación a ciertas variables se realizarán ciertos supuestos:

1. Se supondrá que el costo del tiempo no cambia a lo largo del período utilizado:

$$T_{it} = T_i$$

2. Los costos operativos no varían en el tiempo, $CO_{it} = CO_i$

3. Podemos resumir las variables que dependen solo de los peajes en una constante, de forma que la misma recoja la heterogeneidad no observable entre ellas.

$$\alpha_i = \beta_{4i}CO_i + \beta_{6i}T_i + \beta_{7i}G_i + \beta_{8i}A_i$$

$$v_{it} = \alpha_i + u_{it}$$

Con lo que el modelo estimable nos quedaría:

$$Y_{it} = \beta_0 + \beta_{1i}P_{it} + \beta_{2i}PA_t + \beta_{3i}PC_t + \beta_{4i}IVF_t + v_{it}$$

3. ESTRATEGIA DE ESTIMACIÓN

En esta sección se presenta la metodología que sigue el trabajo para abordar el objeto de estudio. Dado los datos con que se cuenta, observaciones para diferentes puntos de conteo de tráfico en diferentes momentos del tiempo, los modelos econométricos a utilizar se enmarcan dentro de la literatura de datos de panel. El contar con este tipo de datos permite, por un lado, captar las relaciones dinámicas entre las variables al observar información en el tiempo, y por otro, controlar la posible heterogeneidad existente entre los diferentes puestos de conteo debido a variables no observables que afectan la evolución del tráfico.

De un modo general, los modelos de datos de panel pueden ser expresados de la siguiente forma,

$$y_{it} = \mathbf{x}_{it} \boldsymbol{\beta} + \alpha_i + \mu_{it} \quad (1)$$

donde $t = 1, 2, \dots, T$ representa el tiempo de la observación, e $i = 1, 2, \dots, N$ representa el punto de conteo particular. A su vez \mathbf{x}_{it} es un vector de $1 \times K$ variables explicativas, $\boldsymbol{\beta}$ es un vector de parámetros $K \times 1$ y μ_{it} es un error aleatorio. Por último, α_i es el efecto no observable, que como varía con cada puesto de conteo y no con el tiempo, usualmente es considerado como un efecto individual o heterogeneidad individual.

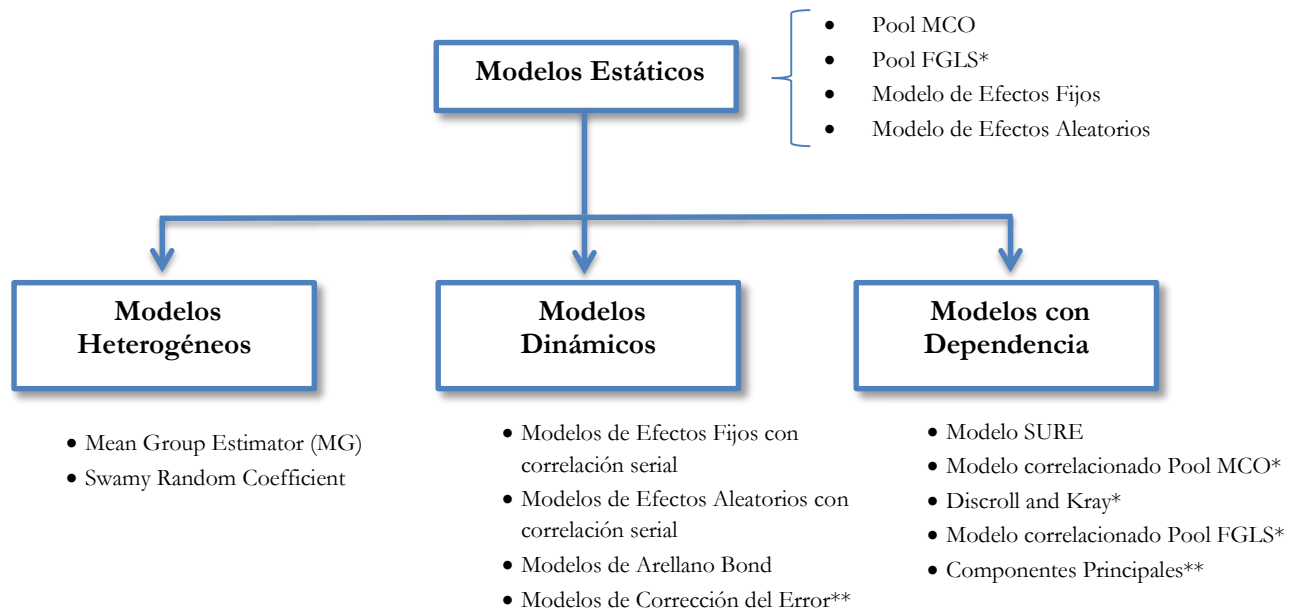
Se asume que los datos con que contamos son una muestra aleatoria extraída de la población sobre la cual se quiere inferir los resultados, es decir, las observaciones $\{(\mathbf{X}_i, \mathbf{y}_i): i = 1, 2, \dots, N\}$ con \mathbf{X}_i matriz de dimensión $T \times K$, e \mathbf{y}_i vector de $T \times 1$, son independientes e idénticamente distribuidas entre los puntos de conteo.

Debido al tamaño de la muestra con que contamos en sus dos dimensiones, N puestos de conteo, y T datos temporales, surgen tres principales aspectos tratar en los datos: heterogeneidad, dinámica y dependencia entre los puestos.

- i. Heterogeneidad: Debido a que es posible estimar una regresión separada para cada puesto de conteo al contar con una muestra considerable de información temporal, resulta natural considerar un panel heterogéneo donde los parámetros puedan diferir por puesto.
- ii. Dinámica: De la misma forma que en el caso anterior, un T largo permite la estimación de modelos dinámicos menos restrictivos. En este sentido, se debe considerar el orden de integración de las series para una correcta estimación por los métodos tradicionales. Para ello se realizan pruebas de raíz unitarias para paneles, que buscan identificar el orden de integración del panel. Sin embargo, en ocasiones, las series no estacionarias mantienen una relación de equilibrio de largo plazo, que es posible determinar estudiando la cointegración de las series.
- iii. Dependencia: Si la dependencia entre los cross-section es grande y no se la considera debidamente, no se logra una mayor eficiencia de los estimadores de paneles frente a la estimación de las series de forma individual, por lo que su tratamiento debe ser debidamente abordado.

Existen diversos modelos econométricos posibles de ser aplicados, el más apropiado dependerá del marco teórico y los supuestos que se adopten sobre las relaciones entre las variables, en particular, respecto de las tres consideraciones antes mencionadas.

El cuadro que se presenta a continuación presenta un esquema de los modelos estimados, clasificando los modelos según contemplen dentro de su estructura las tres consideraciones a tomar en cuenta. (Notar que algunos modelos permiten considerar más de un aspecto):



*Se estiman variantes del modelo que contemplan dinámica

** Se estiman variantes del modelo que contemplan heterogeneidad

A continuación se realiza una breve presentación de los modelos econométricos que se utilizaron para estimar las elasticidades, conjuntamente con los supuestos sobre los que se asientan.

A) Modelos estáticos

Estos modelos se consideran estáticos en la medida que no consideran los ajustes dinámicos, o ajustes en el tiempo, que podrían operar entre variables consideradas. Por otro lado, estos modelos asumen homogeneidad en los parámetros, es decir, se asume que las elasticidades de largo plazo del tráfico a las variables consideradas son las mismas para todos los puntos de conteo.

A.1) Pool MCO

Este modelo asume que los regresores son exógenos y no están correlacionados con el error, y por tanto, no consideran la descomposición del error en $\alpha_i + \mu_{it}$. En otras palabras, se supone que $\alpha_i=0$ y no existe heterogeneidad no observable entre las unidades consideradas. De este modo, podemos estimar de forma consistente el modelo por Mínimos Cuadrados Ordinarios (MCO) ya que se cumplen los supuestos clásicos del modelo de regresión lineal⁴.

El modelo puede ser expresado de la siguiente forma:

$$y_{it} = \alpha + x_{it} \beta + \mu_{it} \quad (2)$$

Un aspecto a considerar para realizar inferencia es la necesidad de controlar la probable correlación de los errores en el tiempo dentro de un mismo puesto de conteo (within correlation). Para contemplar esta posibilidad se realizan estimaciones de estos modelos relajando el supuesto de que los errores no están correlacionados temporalmente dentro de una misma unidad ($E(\mu_{it} \mu_{is}) \neq 0$), y permitiendo heteroscedasticidad en los errores μ_{it} .

A.2) Pool FGLS

Estos modelos al igual que en el caso anterior consideran que los regresores no están correlacionados con el efecto individual que está incluido en el término de error. Sin embargo, estos modelos pueden dar lugar a estimaciones más eficientes que en el caso anterior, ya que consideran diferentes supuestos sobre la estructura de la matriz de correlación de los errores, y por tanto, dan lugar a diferentes estimadores. Todos ellos utilizan estimadores obtenidos por el procedimiento de Mínimos Cuadrados Generalizados Factibles (FGLS).

En nuestro caso particular, se estimará el modelo (2) considerando diferentes estructuras de la matriz de correlaciones entre los errores, por ejemplo, correlaciones temporales dentro de cada puesto de la forma AR(4), AR(8) y AR(12), de forma de permitir que los errores en un determinado momento puedan estar correlacionados con los errores del mismo trimestre un año atrás, dos y tres respectivamente.

⁴Particularmente, homoscedasticidad en los residuos (varianza constante en todos los periodos y entre los individuos) y no existencia de correlación serial (covarianzas de los errores entre diferentes periodos igual a cero).

A.3) Modelo de Efecto Fijo (FE)

En muchas aplicaciones, el interés de utilizar datos de panel recae en considerar la existencia de un efecto individual no observado α_i de cada unidad, permitiendo que esté correlacionado con las variables explicativas, \mathbf{x}_{it} . El modelo de efectos fijos contempla esta posibilidad.

El modelo puede expresarse de la siguiente forma:

$$y_{it} = \alpha_i + \mathbf{x}_{it}\boldsymbol{\beta} + \mu_{it} \quad (3)$$

La idea básica de la estimación es eliminar el efecto no observado, α_i , a través de una transformación de la ecuación 3. En este sentido, existen diversas transformaciones que sirven a este propósito, a continuación se presenta la transformación más utilizada denominada transformación de efectos fijos.

La transformación de efectos fijos se obtiene promediando los valores de y_i a través del tiempo, obteniendo las siguientes ecuaciones:

$$\bar{y}_i = \bar{\mathbf{x}}_i\boldsymbol{\beta} + \alpha_i + \bar{\mu}_i, \quad \forall i$$

Restando a la ecuación 3 la anterior ecuación se obtiene:

$$(y_{it} - \bar{y}_i) = (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)\boldsymbol{\beta} + (\mu_{it} - \bar{\mu}_i) \quad (4)$$

De esta forma, como se puede apreciar, se elimina el efecto no observado α_i , y dado que la ecuación 4 cumple los supuesto requeridos de MCO, los estimadores MCO de esta ecuación serán consistentes e insesgados. Es decir, se obtienen estimadores consistentes, incluso con regresores endógenos \mathbf{x}_{it} posiblemente correlacionados con el efecto individual pero no con el error idiosincrático, μ_{it} .

Modelos de Efecto Fijo con Correlación Serial AR(1)

Adicionalmente, se realiza una estimación del modelo de efecto fijo pero suponiendo autocorrelación serial de los errores de la forma AR(1) dentro de cada puesto, lo que potencialmente podría conducir a estimaciones más eficientes que el caso anterior. De esta forma, el modelo estimado adopta la siguiente forma:

$$y_{it} = \alpha_i + \mathbf{x}_{it}\boldsymbol{\beta} + \mu_{it}$$

$$\mu_{it} = \rho_i\mu_{it-1} + \varepsilon_{it}$$

A.4) Modelo de Efecto Aleatorio (RE)

El modelo de efecto aleatorio es la estimación FGLS de la ecuación 3, bajo los supuestos que el efecto individual α_i es un efecto aleatorio i.i.d de la misma forma que el error idiosincrático, μ_{it} . Es decir, se asume que el efecto individual no está correlacionado con la los regresores a diferencia del modelo anterior.

Al igual que el modelo Pool MCO, se incluye el efecto no observado en el término de error. De esta forma, el modelo puede ser expresado como sigue:

$$y_{it} = \mathbf{x}_{it} \boldsymbol{\beta} + v_{it}$$

con los siguiente supuestos:

1. $v_{it} = \alpha_i + \mu_{it}$
2. $\alpha_i \sim (\alpha, \sigma_\alpha^2)$
3. $\mu_{it} \sim (0, \sigma_\mu^2)$

Por lo tanto, $Var(v_{it}) = \sigma_\alpha^2 + \sigma_\mu^2$ y, $Cov(v_{it}, v_{is}) = \sigma_\alpha^2 \quad \forall s \neq t$

De lo anterior se desprende que en este modelo los errores v_{it} son homoscedásticos (varianza constante) y autocorrelacionados, ya que la correlación serial entre los errores dentro de un mismo puesto adopta la siguiente forma:

$$Corr(v_{it} v_{is}) = \sigma_\alpha^2 / (\sigma_\alpha^2 + \sigma_\mu^2) \quad \forall s \neq t$$

Lo que resulta interesante de este modelo, es que utiliza información sobre la correlación serial de los residuos v_{it} en un marco de FGLS con una determinada estructura de la matriz de varianzas y covarianzas (denominada $\boldsymbol{\Omega}$) definida de la siguiente forma:

$$\boldsymbol{\Omega} = E(\mathbf{v}_i \mathbf{v}_i') = \begin{pmatrix} \sigma_c^2 + \sigma_\mu^2 & \sigma_c^2 & \dots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_\mu^2 & \dots & \sigma_c^2 \\ \vdots & & \ddots & \vdots \\ \sigma_c^2 & \dots & \dots & \sigma_c^2 + \sigma_\mu^2 \end{pmatrix}$$

donde \mathbf{v}_i es el vector del error para todos los tiempos, por lo tanto de dimensión $T \times 1$.

Modelos de Efecto Aleatorio con Correlación Serial AR(1)

Al igual que en el caso de modelos de efectos fijos, se realiza una estimación del modelo de efecto aleatorio pero suponiendo correlación serial de los residuos de la forma AR(1).

A.5) First Difference Estimator (FD):

Una transformación alternativa al modelo de efectos fijos para eliminar el efecto individual, es realizar una primera diferencia las variables. Al estimador MCO de la siguiente ecuación se lo llama First Difference Estimator:

$$(y_{it} - y_{it-1}) = (\mathbf{x}_{it} - \mathbf{x}_{it-1})' \boldsymbol{\beta} + (\mu_{it} - \mu_{it-1})$$

B) Modelos Estáticos con Correlación entre los Puestos de Conteo (Dependencia)

Los modelos pertenecientes a esta categoría, a diferencia de los modelos anteriores, permiten relajar el supuesto de que los errores entre los puestos de conteo son independientes entre sí. Adicionalmente, determinados modelos permiten contemplar simultáneamente la correlación serial dentro de cada unidad y heteroscedasticidad.

Estos modelos cobran especial relevancia en nuestra aplicación ya que determinados puestos de conteo se encuentran en una misma ruta, y al no existir rutas alternativas en viajes largos, es razonable suponer que los shocks que afectan el tráfico en un punto estarán correlacionados con aquellos que afectan el otro punto ubicado a kilómetros de distancia por la misma ruta.

B.1) Modelos SURE

Como se mencionó anteriormente, se podría considerar el caso de que los errores de los diferentes puntos de conteo estén correlacionados entre sí. De este modo, el modelo SURE contempla el siguiente caso de estudio a diferencia de los modelos anteriores:

$$E(\mu_{it}, \mu_{jt}) \neq 0$$

Considérese entonces el siguiente sistema de regresiones aparentemente no relacionadas:

$$\begin{aligned} \mathbf{y}_1 &= \beta_{11} + \beta_{12}x_2 + \beta_{13}x_3 + \dots + \mu_1 \\ \mathbf{y}_2 &= \beta_{21} + \beta_{22}x_2 + \beta_{23}x_3 + \dots + \mu_2 \\ &\vdots \\ \mathbf{y}_N &= \beta_{N1} + \beta_{N2}x_2 + \beta_{N3}x_3 + \dots + \mu_N \end{aligned}$$

Donde cada \mathbf{y}_i es un vector de $t = 1, 2, \dots, T$ observaciones, con los mismos regresores \mathbf{x} para cada ecuación, vectores $\boldsymbol{\beta}$ que pueden variar entre ecuaciones y vectores de errores $\boldsymbol{\mu}_i$ de dimensión $T \times 1$.

De esta forma, se puede expresar la matriz de varianzas y covarianzas de la siguiente forma:

$$E(\mathbf{UU}') = E[(\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N)(\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N)'] ,$$

por lo que,

$$E(\mathbf{UU}') = \begin{bmatrix} E(\boldsymbol{\mu}_1\boldsymbol{\mu}_1) & E(\boldsymbol{\mu}_1\boldsymbol{\mu}_2) & \dots & E(\boldsymbol{\mu}_1\boldsymbol{\mu}_N) \\ E(\boldsymbol{\mu}_2\boldsymbol{\mu}_1) & E(\boldsymbol{\mu}_2\boldsymbol{\mu}_2) & & \\ \vdots & & \ddots & \\ E(\boldsymbol{\mu}_N\boldsymbol{\mu}_1) & E(\boldsymbol{\mu}_N\boldsymbol{\mu}_2) & \dots & E(\boldsymbol{\mu}_N\boldsymbol{\mu}_N) \end{bmatrix} \quad (5)$$

Si la correlación es contemporánea y constante entre los diferentes individuos, pero no existe correlación serial, es decir $E(\mu_{it}\mu_{js}) = \sigma_{ij}$ para $t = s$ y 0 para $t \neq s$, cada elemento de la matriz

anterior es una matriz diagonal de orden $T \times T$ que puede ser expresado como $\sigma_{ij} \cdot \mathbf{I}_T$, por lo que su representación matricial es:

$$E(UU') = \begin{bmatrix} \sigma_{11} \cdot \mathbf{I}_T & \sigma_{12} \cdot \mathbf{I}_T & \dots & \sigma_{1N} \cdot \mathbf{I}_T \\ \sigma_{21} \cdot \mathbf{I}_T & \sigma_{22} \cdot \mathbf{I}_T & & \\ \vdots & & \ddots & \\ \sigma_{N1} \cdot \mathbf{I}_T & \sigma_{N2} \cdot \mathbf{I}_T & \dots & \sigma_{NN} \cdot \mathbf{I}_T \end{bmatrix}$$

De esta forma, se puede apreciar como dentro de cada punto de conteo los errores son homoscedásticos y sin autocorrelación serial.

Utilizando el operador \otimes para definir el producto Kronecker y Σ para definir la matriz de varianzas y covarianzas particular de cada ecuación, se puede expresar lo anterior de la siguiente forma:

$$E(UU') = \Sigma \otimes \mathbf{I}_T$$

Si la matriz de varianzas y covarianzas es conocida, se puede utilizar MCG para estimar los parámetros de la regresión. En caso de que la matriz sea desconocida, se debe realizar una estimación previa de los errores por MCO a cada una de las ecuaciones de puntos de conteo, obtener la matriz de varianzas y covarianzas estimadas, y utilizar esta estimación para computar el estimador de los parámetros $\hat{\beta}$ del sistema, denominado estimador de Mínimos Cuadrados Generalizados Factible, o Mínimos Cuadrados Generalizados de Zellner.

B.2) Modelos Correlacionados Pool MCO

Se estiman modelos Pool MCO como los descritos en la parte A, pero considerando correlación en los errores μ_{it} entre los diferentes puestos de conteo. A su vez, se permite correlación serial dentro de cada puesto de conteo de la forma AR(1) y se permite heteroscedasticidad en los errores.

De esta forma, considerando el mayor grado de generalidad, los errores de la ecuación 1 adoptan la siguiente forma:

$$\mu_{it} = \rho_i \mu_{it-1} + \varepsilon_{it}$$

donde ε_{it} son serialmente incorrelacionados pero correlacionados entre puestos, $Cor(\varepsilon_{it} \varepsilon_{jt}) = \sigma_{ij}$

De esta forma para $t > 1$, se realiza la siguiente estimación por MCO:

$$(y_{it} - \hat{\rho}_i y_{it,t-1}) = (x_{it} - \hat{\rho}_i x_{it,t-1})' \beta + \hat{\rho}_i \mu_{it-1} + \varepsilon_{it}$$

Mientras que para $t = 1$, se regresa $\sqrt{(1 - \hat{\rho}_i)^2} y_{i1}$ sobre $\sqrt{(1 - \hat{\rho}_i)^2} x_{i1}$.

Adicionalmente, se realiza una estimación de este modelo considerando la siguiente restricción $\rho = \rho_i$, es decir, se supone la misma correlación serial para todos los puestos y se calcula $\hat{\rho}$ promediando los $\hat{\rho}_i$.

Modelos con matriz de covarianzas Discroll and Kray (1998)

Dentro de estos modelos, se realiza una estimación Pool MCO, utilizando el método propuesto por Discroll and Kray (1998) para estimar la matriz de covarianzas. Estos autores proponen una alternativa para tratar el problema de correlación espacial entre las diferentes unidades del panel, conjuntamente con correlación serial y heteroscedasticidad. El método se basa en una estimación no paramétrica de la matriz de covarianzas, utilizando la técnica de Newey and West (1987), lo que permite estimaciones consistentes de los errores estándar. Estos errores así estimados son robustos a una forma general de dependencia espacial y temporal entre los cross section o puestos. Para que las estimaciones de la matriz de varianzas y covarianzas sean consistentes solamente se requiere que la dependencia entre los cross section vaya decreciendo a medida que aumenta el intervalo temporal.

Adicionalmente se realiza una estimación del modelos de efectos fijos presentado anteriormente utilizando esta técnica para estimar la matriz de covarianzas.

B.3) Modelos Correlacionados Pool FGLS

Estos modelos utilizan estimadores Pool FGLS asumiendo que los supuestos sobre los errores asumidos son correctos. Las estimaciones son asintóticamente más eficientes que los modelos anteriores si el modelo está correctamente especificado.

Al igual que en el caso anterior, se supone correlación entre los errores de los puestos de conteo, heteroscedasticidad, correlación serial de la forma AR(1), con el coeficiente AR(1) ρ_i individual para cada puesto o común a todos ellos $\rho = \rho_i$.

B.4) Modelos de Principal Components

Mean Group – Principal Components

Como se ilustró en el cuadro resumen presentado anteriormente, los modelos de componentes principales, particularmente el modelo Mean Group-Principal Components (MG-PC) desarrollado por Coakley, Fuertes y Smith (2002), permiten considerar la dependencia entre los cross section y la heterogeneidad en los parámetros.

De esta forma, los autores sugieren estimar por MCO, en una primera etapa, el siguiente modelo:

$$y_{it} = \alpha_i + x_{it} \beta_i + \mu_{it}$$

y luego estimar los componentes principales de $\hat{\mu}_{it}$. Los factores estimados y seleccionados, \hat{f}_t son luego incluidos en una segunda regresión:

$$y_{it} = \alpha_i + x_{it} \beta_i + \hat{f}_t c_{it} + v_{it}.$$

Estos modelos asumen que los regresores x_{it} son generados por:

$$x_{it} = f_t \phi_{1i} + z_t \phi_{2i} + \chi_t \phi_{3i} + v_{it}$$

donde χ_t son factores comunes que influyen \mathbf{x}_{it} pero no y_{it} . Las estimaciones de β_i para T y N largos serán consistentes si se cumplen alguna de las siguientes condiciones: $\phi_{1i} = 0$ o $\phi_{3i} = 0$. De otra forma, utilizando estimaciones inconsistentes de $\hat{\mu}_{it}$, se genera un sesgo en la estimación. La demostración de los autores considera la restricción $\phi_{3i} = 0$.

Una vez obtenidos los estimadores β_i , los autores proponen el estimador β (MG-PC), realizando el promedio entre los cross section como proponen Pesaran y Smith (1995) (ver más adelante).

Fixed Effects with Principal Components

Coakley, Fuertes y Smith (2002) sugieren como alternativa al modelo anterior, considerar un modelo de efectos fijos con parámetros homogéneos, ampliado con los factores extraídos de los residuos de una primera regresión de y_{it} sobre \mathbf{x}_{it} .

$$y_{it} = \alpha_i + \mathbf{x}_{it} + \beta + \hat{\mathbf{f}}_t c_{it} + v_{it}$$

Al igual que en el caso anterior, para la consistencia debe cumplirse que los regresores \mathbf{x}_{it} no estén correlacionados con el error μ_{it} .

C) Modelos con Parámetros Heterogéneos

Los modelos presentados hasta ahora consideraban los parámetros β homogéneos y fijos⁵, en estos modelos los parámetros son supuestos heterogéneos y aleatorios. El supuesto del que parten estos modelos de parámetros aleatorios implica que (α_i, β_i) son realizaciones de una distribución de probabilidad específica con un número fijo de parámetros que la definen, distribuidos independientemente de los regresores x_{it} y del error. Adicionalmente se puede suponer que algunos coeficientes (por ejemplo α_i) son fijos mientras que otros (por ejemplo β_i) son aleatorios.

C.1) Mean Group Estimator (MG) (Pesaran y Smith (1995))

El estimador propuesto por Pesaran y Smith (1995) consiste en estimar por MCO la siguiente ecuación para cada puesto de conteo separadamente:

$$y_{it} = \alpha_i + x_{it} \beta_i + \mu_{it}$$

con $\mu_{it} \sim (0, \sigma_\varepsilon^2)$

De esta forma, el estimador MG propuesto es el promedio de las estimaciones anteriores, por lo que para cada regresor el parámetro estimado se puede expresar:

$$\hat{\beta}_{MG} = N^{-1} \sum_{i=1}^N \hat{\beta}_i$$

con varianza $V(\hat{\beta}_{MG}) = 1/N(N-1) \sum_{i=1}^N (\hat{\beta}_i - \bar{\beta})^2$

Si las variables son I(1) y cointegradas, entonces, entonces $\hat{\beta}_i$ es superconsistente (a tasa T) del coeficiente de largo plazo β_i . Sin embargo, las estimaciones serán espúreas si μ_{it} es I(1), pese a que al igual que con los estimadores Pool MCO y FE, promediar entre los puestos de conteo podría atenuar este problema para una cantidad grande de puestos (N).

C.2) Swamy Random Coefficient Model (RCM)

Swamy (1970) sugiere un estimador FGLS, que es equivalente a utilizar un promedio ponderado de los estimadores MCO estimados individualmente para cada puesto de conteo, a diferencia del estimador MG que utiliza un promedio simple.

El procedimiento es el siguiente:

Utilizando los residuos y la varianza estimada s_i^2 de cada estimador, definida de la siguiente forma

$$\hat{\mu}_i = y_i - X_i \hat{\delta}_i$$

$$s_i^2 = \hat{\mu}_i' \hat{\mu}_i / (T - k - 1)$$

⁵ A excepción del modelo Mean-Group-Principal Components presentado en el numeral B.4.

donde $\hat{\boldsymbol{\mu}}_i$ y \mathbf{y}_i son vectores de dimensión $T \times 1$; \mathbf{X}_i es una matriz de $T \times (k+1)$ ⁶ y $\hat{\boldsymbol{\delta}}_i = (\hat{\alpha}_i, \hat{\boldsymbol{\beta}}_i)$ una matriz de $(k+1) \times 1$.

La covarianza estimada $\hat{\boldsymbol{\delta}}_i$ se define como

$$V(\hat{\boldsymbol{\delta}}_i) = s_i^2 / (\mathbf{X}_i' \mathbf{X}_i)^{-1}$$

Swamy sugiere estimar $\boldsymbol{\Omega}$, la matriz de varianzas y covarianzas entre los parámetros, de la siguiente forma:

$$\hat{\boldsymbol{\Omega}} = \sum_i (\hat{\boldsymbol{\delta}}_i - \bar{\boldsymbol{\delta}}_i)(\hat{\boldsymbol{\delta}}_i - \bar{\boldsymbol{\delta}}_i)' / N - \sum_i V(\hat{\boldsymbol{\delta}}_i) / N$$

Si este estimador no es definido positivo (usualmente no lo es), el último término es definido en cero. El estimador de la media propuesto es:

$$\hat{\boldsymbol{\delta}}_{SWAMY} = \sum_i D_i \hat{\boldsymbol{\delta}}_i$$

$$\text{con } D_i = \left\{ \sum_i [\hat{\boldsymbol{\Omega}} + V(\hat{\boldsymbol{\delta}}_i)]^{-1} \right\}^{-1} [\hat{\boldsymbol{\Omega}} + V(\hat{\boldsymbol{\delta}}_i)]^{-1}$$

El estimador MG presentado anteriormente es sensible a la presencia de outliers, por lo que en principio, la ponderación que el estimador Swamy realiza reduce este problema. A medida que $T \rightarrow \infty$ la diferencia entre los dos estimadores se reduce.

⁶ Se incluye una columna de unos en la matriz \mathbf{X}_i para realizar la transformación $\hat{\boldsymbol{\delta}}_i = (\hat{\alpha}_i, \hat{\boldsymbol{\beta}}_i)$.

D) Modelos Dinámicos

Estos modelos incluyen rezagos de la variable dependiente como regresores, de forma de poder capturar la dinámica en el tiempo de las variables. Se realizan estimaciones del modelos de efectos fijos presentado anteriormente realizando una primera diferencia de las variables, sin embargo, estimaciones del modelo FE produce estimaciones inconsistentes de los parámetros vinculados a las variables rezagadas. De esta forma, estimaciones consistentes de los parámetros pueden ser obtenidas realizando estimaciones por el método de variables instrumentales (IV), utilizando como instrumentos rezagos apropiados de las variables.

El modelo general, puede ser expresado de la siguiente forma:

$$y_{it} = \gamma_1 y_{it-1} + \dots + \gamma_p y_{it-p} + x_{it} \beta_0 + x_{it-1} \beta_1 + \dots + x_{it-q} \beta_q + \alpha_i + \mu_{it} \quad (6)$$

El modelo así especificado contempla diferentes mecanismos que provocan la existencia de correlación en y a través del tiempo: 1) directamente a través de y en períodos anteriores; 2) directamente a través de los regresores x ; y 3) indirectamente a través del efecto individual no observable.

La estimación de este modelo se realiza aplicando IV a la primera diferencia de las variables, y utilizando como instrumentos, para los rezagos de las variables incluidas en la regresión, rezagos anteriores en el tiempo de la propia variable. Estos estimadores son denominados estimadores Arellano-Bond desde su trabajo Arellano-Bond (1991) donde se detalla la implementación del mismo.

E) Modelos de Corrección del Error

En años recientes la literatura de paneles dinámicos se ha empezado a enfocar en el estudio de paneles donde T y N son largos. Uno de los principales hallazgos de esta literatura es que el supuesto de homogeneidad de los parámetros es usualmente inadecuado (Pesaran and Smith (1995), Im, Pesaran and Shin (2003) y Phillips and Moon (2000)). Adicionalmente, con el incremento de las observaciones temporales (T largos), la no estacionariedad de las variables es un nuevo factor a considerar. Los modelos considerados en esta sección permiten estimar de forma consistente los parámetros para paneles de datos no estacionarios y con parámetros heterogéneos.

E.1) Pool Mean Group (Pesaran Shin y Smith (1999)) (EC-PMG)

Estos autores argumentan que en determinadas aplicaciones es más probable que los efectos de largo plazo entre las variables sean homogéneos entre los diferentes puestos, mientras que el ajuste de las variables en el corto plazo y sus varianzas sigan un comportamiento heterogéneo, es decir, particular a cada puesto.

Partiendo del modelo (6), si las variables son $I(1)$ y cointegradas, entonces el error resultante será un proceso $I(0)$ para todo i . La característica principal de variables cointegradas es su respuesta o ajuste a cualquier desviación respecto a la relación de largo plazo que las variables mantienen. Esto implica la consideración de un modelo de corrección del error en donde la dinámica de corto plazo de las variables se ve influenciada por los desvíos de la relación de largo plazo. Por lo que es común, bajo estos supuestos, reparametrizar la ecuación (6) en un modelo de corrección del error (con parámetros de corto plazo heterogéneos) que puede ser expresado de la siguiente forma:

$$\Delta y_{it} = \lambda_i(y_{it-1} - \theta_i' x_{it}) + \sum_{j=1}^{p-1} \gamma_{ij}^* \Delta y_{it-j} + \sum_{j=0}^{q-1} \beta_{ij}^* \Delta x_{it-j} + \alpha_i + \mu_{it} \quad (7)$$

donde $y_{it-1} = \theta_i' x_{it}$ es la relación de largo plazo entre las variables, y el parámetro λ_i es la corrección del error o el término de ajuste de las variables de corto plazo a la relación de equilibrio de largo plazo. Si $\lambda_i = 0$ no existe relación de equilibrio de largo plazo entre las variables. Por otro lado, es de esperarse que el parámetro sea significativamente negativo bajo el supuesto mencionado que las variables tienden a retornar al equilibrio de largo plazo. De particular importancia para nuestro caso es el vector θ_i' que contiene información sobre la relación de largo plazo entre las variables.

De esta forma, los autores desarrollan un método de Máxima Verosimilitud⁷ (MV) para estimar los parámetros de la anterior ecuación pero imponiendo la restricción de que los parámetros de largo plazo sean iguales entre los puestos, $\theta_i = \theta$.

E.2) Modelo de Corrección del Error Mean Group (EC-MG)

Adicionalmente, se realiza una estimación de la ecuación (7) sin imponer la restricción mencionada en el modelo anterior ($\theta_i = \theta$). Es decir, se realiza una estimación del modelo de corrección del error suponiendo que tanto los parámetros de corto plazo como los de largo plazo son heterogéneos, de forma análoga a la estimación MG presentada anteriormente para modelos estáticos.

E.3) Modelo de Corrección del Error con Efectos Fijos (EF-FE)

Por último se realiza una estimación del modelo (7) pero a la restricción impuesta por el PMG sobre homogeneidad de los parámetros de largo plazo, se impone restricciones de homogeneidad sobre el coeficiente de ajuste ($\lambda_i = \lambda$) y sobre los parámetros de corto plazo. En decir, se realiza una estimación del modelo de Efectos Fijos presentado anteriormente pero aplicada sobre una especificación de modelo de corrección del error.

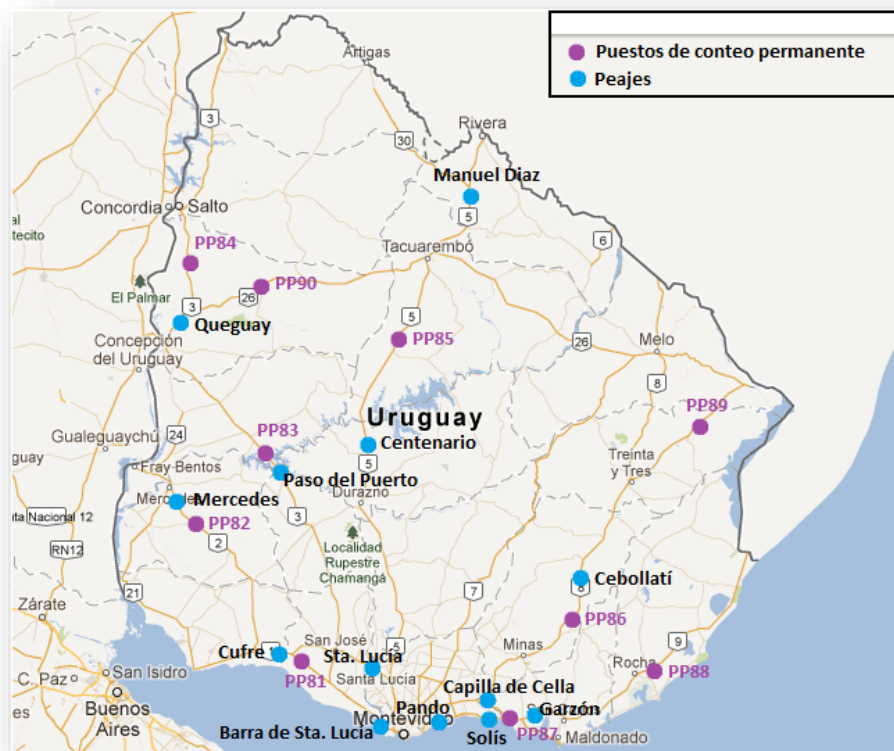
⁷ Se implementa una estimación por MV dado que la ecuación (7) es no lineal en los parámetros, debido a que:

$$\begin{aligned} \lambda_i &= -(1 - \sum_{j=1}^p \gamma_{ij}); \\ \theta_i &= \sum_{j=0}^q \beta_{ij} / (1 - \sum_k \lambda_{ik}); \\ \gamma_{ij}^* &= -\sum_{m=j+1}^p \gamma_{im} \text{ con } j = 1, 2, \dots, p-1; \text{ y} \\ \beta_{ij}^* &= -\sum_{m=j+1}^q \beta_{im} \text{ con } j = 1, 2, \dots, q-1. \end{aligned}$$

4. ANÁLISIS DE DATOS

En este capítulo se realizará un análisis de los datos a utilizar en el estudio. Los datos de tráfico surgen de dos fuentes principales:

- Puestos de conteo permanente: Se cuenta con información de 10 puestos de conteo permanente, ubicados en diferentes puntos de las Rutas Nacionales de Uruguay, que registran, entre otros datos, el volumen de tráfico que circula en ambos sentidos por estos puestos, todos los días del año en las 24 horas del día. Esta información es recopilada en los Anuarios de Vialidad, por la Dirección Nacional de Vialidad (Ministerio de Transporte y Obras Públicas)
- Peajes: Se cuenta con datos del tráfico que circula en ambos sentidos por los 13 peajes de Uruguay. Esta información es aportada por la Corporación Vial del Uruguay, organismo que tiene a su cargo la concesión de 1.600 km de Rutas Nacionales (aproximadamente la mitad de la Red Primaria de Uruguay)

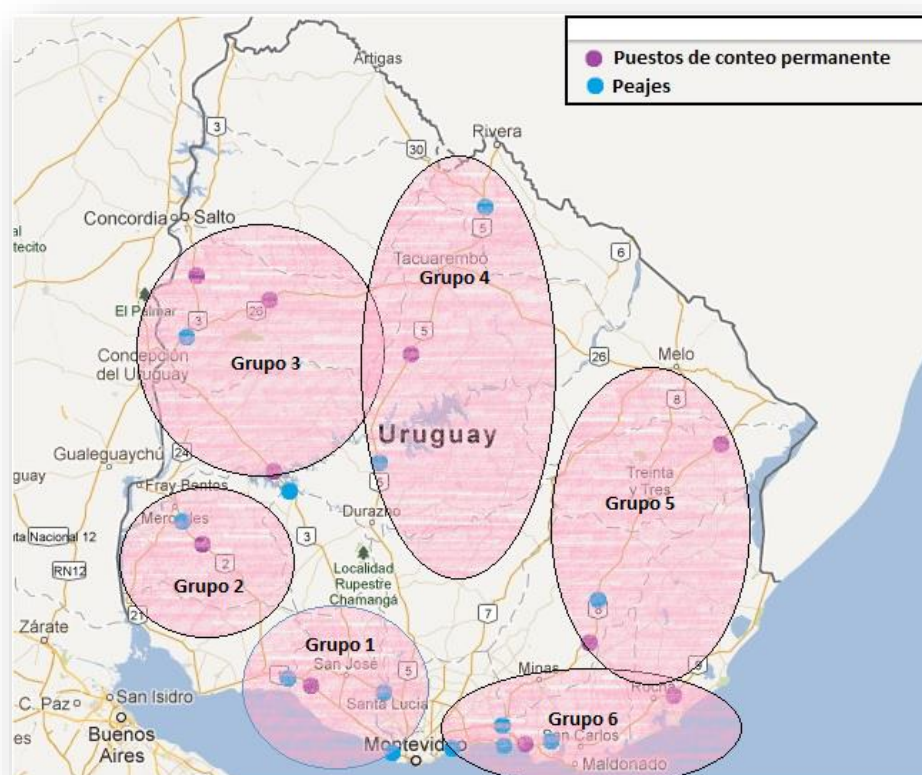


De este modo, se cuenta con datos del volumen de tráfico que circuló en ambos sentidos en 23 puntos de las Rutas Nacionales del país, para autos, ómnibus, camiones medianos y camiones pesados. Se utilizarán datos trimestrales y en el período comprendido entre el tercer trimestre de 2003 y el último trimestre de 2012.

Este capítulo se estructura de la siguiente forma: comienza con un sección donde se describen las series de tráfico consideradas; seguidamente se presenta una sección donde se realiza un análisis estadístico de las series, en particular, se presentan diferentes pruebas de raíz unitaria que permitan concluir sobre la estacionariedad de las mismas; se presentan pruebas de cointegración para las series; pruebas de dependencia y pruebas de heterogeneidad. Finalmente se presentan las variables explicativas incluidas en el estudio.

4.1. Estadísticas Descriptivas

En esta sección se realiza una descripción estadística de las series, presentando las series y sus principales características. Para facilitar la presentación de las series con fines expositivos, y debido a la cantidad de series, es conveniente dividir las series por grupos en función de la ubicación de los puestos de conteo en las Rutas Nacionales y proximidad entre ellos. En el mapa presentado a continuación se explicita la agrupación considerada.



A continuación se presenta una tabla con las estadísticas descriptivas de cada puesto de conteo, agrupados por grupos.

Tabla 1: Estadísticas descriptivas del Tráfico de Camiones Pesados en los distintos puntos de conteo

Grupo	Puesto	Ruta	Km	Mediana	Desv. Estándar	Máximo	Mínimo	TPD A (2011)	Obs
1	Peaje Barra Sta Lucía	1	23.500	90.783	19.864	122.352	57.412	1.283	38
	PP81	1	105.140	24.698	5.530	35.701	14.067	345	38
	Peaje Cufre	1	107.350	26.604	5.518	37.012	16.261	388	38
	Peaje Sta Lucía	11	81.000	17.643	3.227	25.382	12.884	234	38
2	PP82	2	240.600	26.247	8.433	43.515	13.461	418	38
	Peaje Mercedes	2	284.400	46.004	14.741	80.664	22.836	755	38
3	Peaje Paso del Puerto	3	245.200	34.354	8.547	49.536	19.845	409	38
	PP83	3	267.610	29.070	9.618	44.419	12.008	388	38
	Peaje Queguay	3	392.750	39.303	13.749	62.137	17.089	635	38
	PP84	3	436.470	19.875	5.121	30.401	9.597	288	38
	PP90	26	96.710	7.183	3.929	14.797	2.153	122	38
4	Peaje Centenario	5	246.350	25.107	7.135	40.359	13.640	396	38
	PP85	5	313.300	17.427	5.594	31.562	10.743	288	38
	Peaje Manuel Díaz	5	423.200	20.786	8.219	34.834	9.230	372	38
5	PP86	8	171.240	15.425	4.344	22.928	7.294	223	38
	Peaje Cebollatí	8	206.250	16.924	4.675	26.018	8.075	253	38
	PP89	18	360.640	8.064	2.895	15.003	4.620	129	38
6	Peaje Pando	Interb.	33.000	1.028	1.253	4.802	23	27	38
	Peaje Solís	Interb.	81.000	651	713	2.939	84	21	38
	Peaje Capilla de Cella	9	79.500	23.225	6.507	35.441	4.529	293	38
	PP87	9	87.710	22.285	5.106	31.832	13.623	317	38
	Peaje Garzón	9	177.650	12.907	3.331	21.912	7.567	165	38
	PP88	9	215.510	7.987	1.277	11.405	5.214	99	38

El puesto de conteo que registra mayor tráfico de camiones pesados es el peaje de Barra de Santa Lucía, con una media de tráfico de casi el doble del puesto que lo sigue, Peaje Mercedes.

Luego de estos puestos, son los conteos ubicados en el Grupo 3 los que registran mayores valores, junto con el resto de los puestos del Grupo 1 y 2.

Gráfico 1: Tráfico de Camiones Pesados Grupo 1

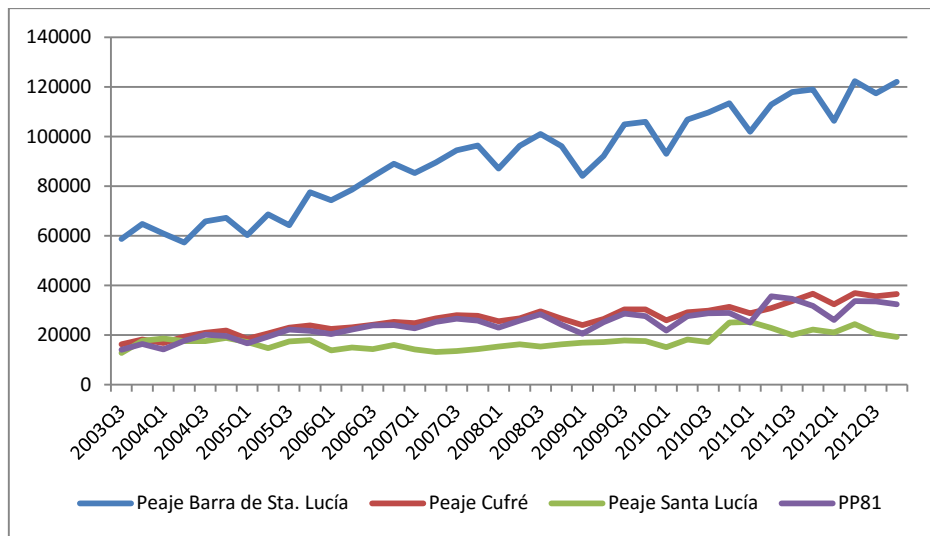
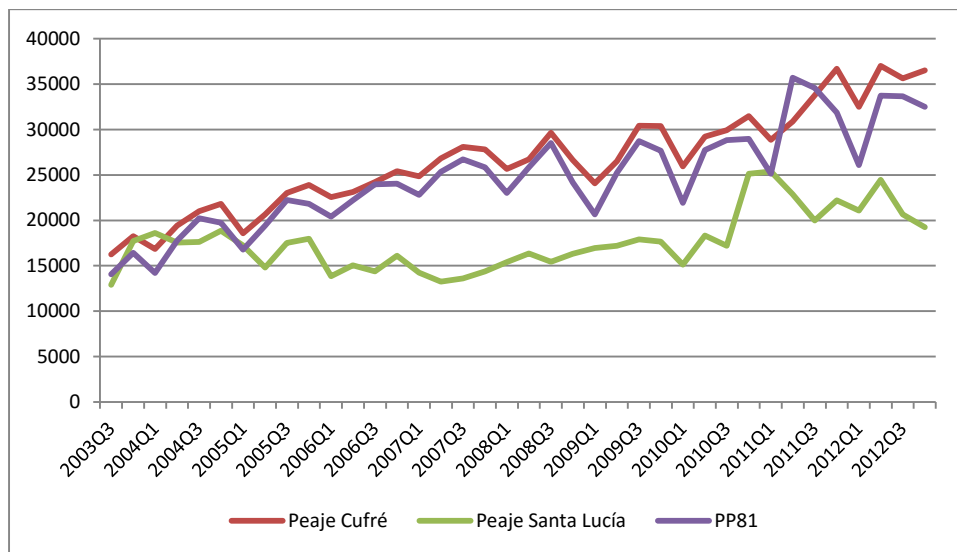
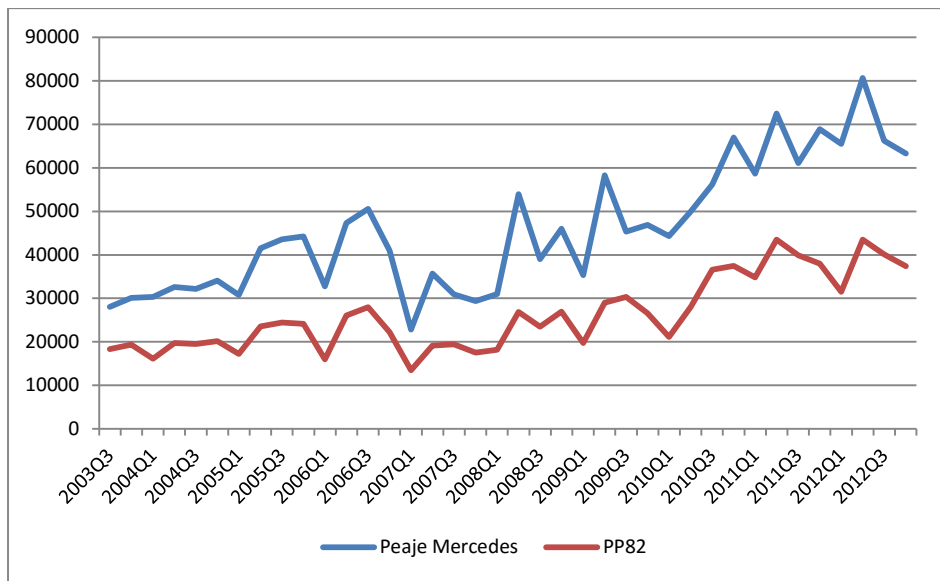


Gráfico 2: Tráfico de Camiones Pesados Grupo 1(b)



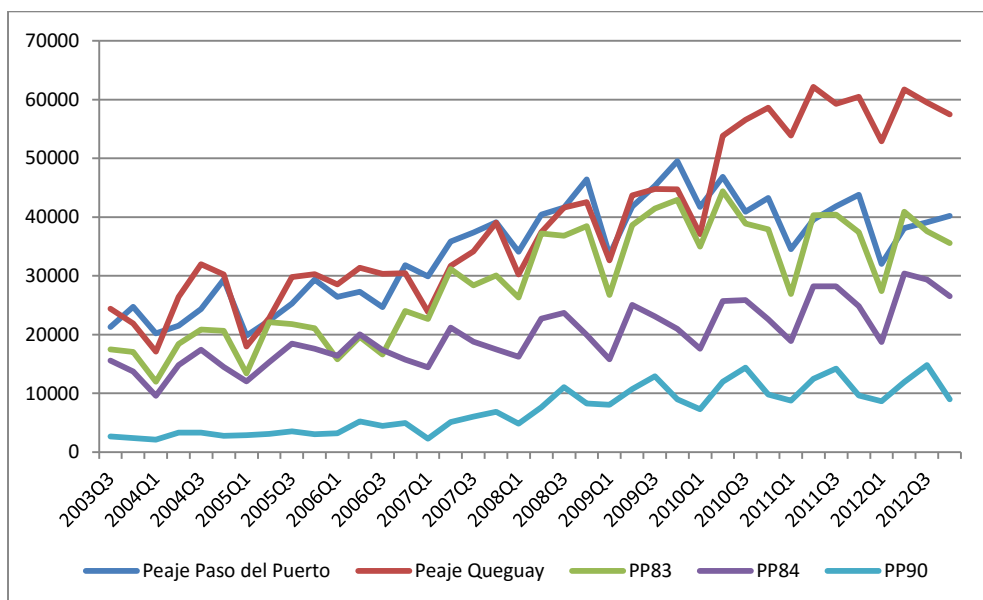
El Grupo 1 lo componen los puestos situados en Ruta 1, y el Peaje Santa Lucía ubicado en Ruta 11. Se destaca por su nivel de tráfico el Peaje Barra de Sta. Lucía, puesto más próximo a Montevideo. Por otro lado, los puestos PP81 y Peaje Cufre presentan niveles de tráfico y evolución en el tiempo muy similares, como era esperable dado su proximidad (se encuentran ubicados a menos de 2 km de distancia). Para una mejor apreciación de la similitud de estas series se presenta el gráfico 10 que no contiene la serie Barra de Santa Lucía, eliminando su efecto de escala y permitiendo una mejor visualización del resto.

Gráfico 3: Tráfico de Camiones Pesados Grupo 2



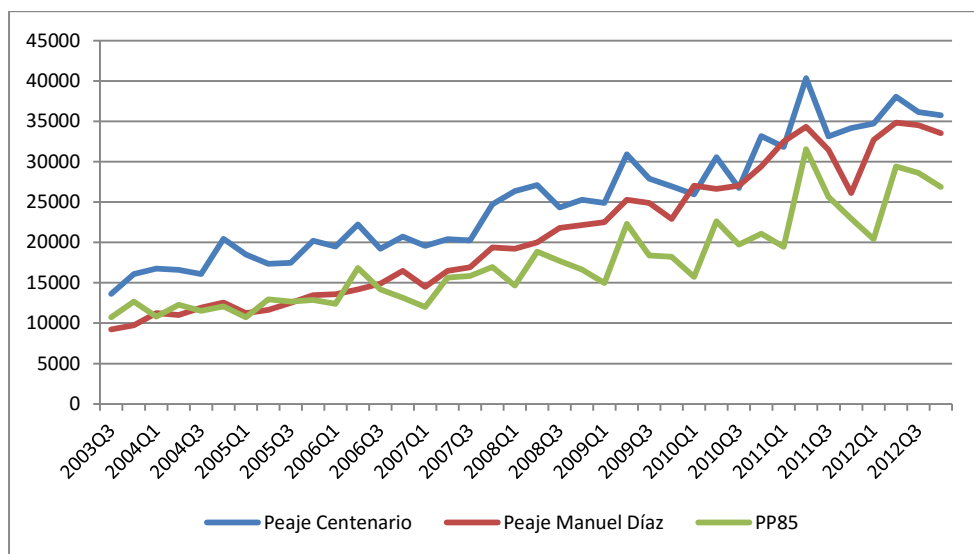
El Grupo 2, compuesto por los puestos ubicados en Ruta 2 son los que siguen en importancia en cuanto al volumen de tráfico a los puesto Barra de Santa Lucía del Grupo 1 y puestos del Grupo 3. Estas series se ven afectadas por el corte del puente de Fray Bentos en casi todo el período de estudio, por lo que los niveles de tráfico se mantienen estables hasta la reapertura del puente a fines del año 2010, donde comienza una tendencia al alza. Ambas series presentan un comportamiento similar en cuanto a su evolución temporal, lo que es esperable dado la poca distancia entre ellas, menos de 44 km.

Gráfico 4: Tráfico de Camiones Pesados Grupo 3



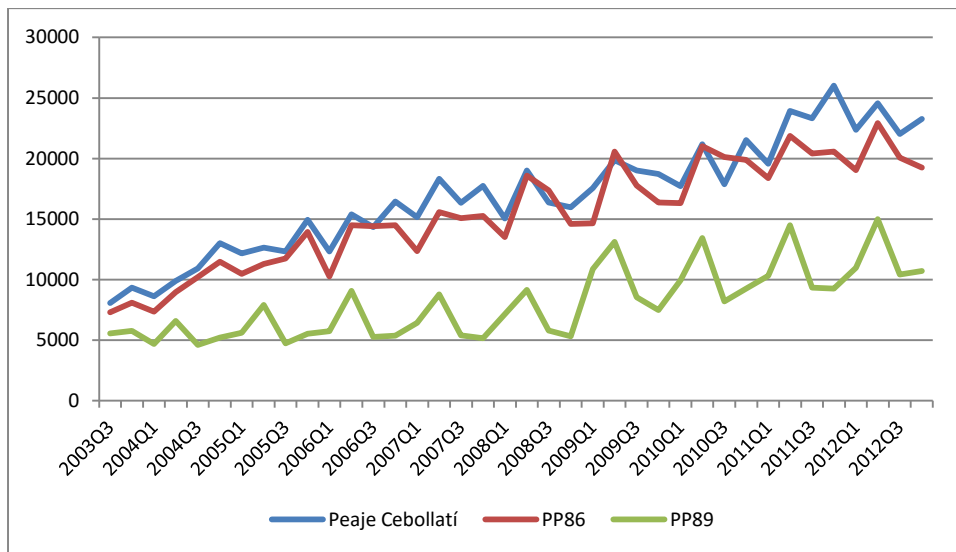
En referencia al Grupo 3, éste está compuesto por los puestos de conteo ubicados en Ruta 3, al que se le agregó adicionalmente el puesto PP90, ubicado en Ruta 26. Este último se consideró apropiado incorporarlo por la proximidad en distancia respecto al Peaje Queguay, menos de 110 km, y ya que capta el tráfico que circula entre Paysandú- Salto con Tacuarembó-Rivera. Con respecto al volumen, el puesto Peaje Queguay presenta los mayores niveles, seguido por los puestos Peaje Paso del Puerto, PP83, PP84 y finalmente el PP90. Las series Peaje Paso del Puerto y PP83 siguen un comportamiento similar en el tiempo dado su proximidad (menos de 23 km), y comienzan a diferenciarse a partir del año 2010. Todas las series presentan un componente estacional, siendo este más acentuado en los puestos Peaje Queguay, PP83, y PP84.

Gráfico 5: Tráfico de Camiones Pesados Grupo 4



El Grupo 4 está compuesto por los puestos ubicados en Ruta 5. Los tres puestos siguen una similar evolución temporal en cuanto al crecimiento, con diferencias en los niveles de tráfico (siendo el Peaje Centenario quien presenta los registros más elevados), con una estacionalidad más marcada en caso del Peaje Centenario y PP85.

Gráfico 6: Tráfico de Camiones Pesados Grupo 5



Con referencia al Grupo 5, este está compuesto por los puestos ubicados en Ruta 8 al que se incorpora el puesto PP89 ubicado en Ruta 18 (dado su proximidad a la Ruta 8 y su ubicación entre el corredor vial que une a Montevideo con Río Branco. En lo que respecta al volumen de tráfico, el puesto Peaje Cebollatí presenta los registros más elevados, seguido por el puesto PP86, pese a encontrarse este último más próximo a Montevideo.

Gráfico 7: Tráfico de Camiones Pesados Grupo 6

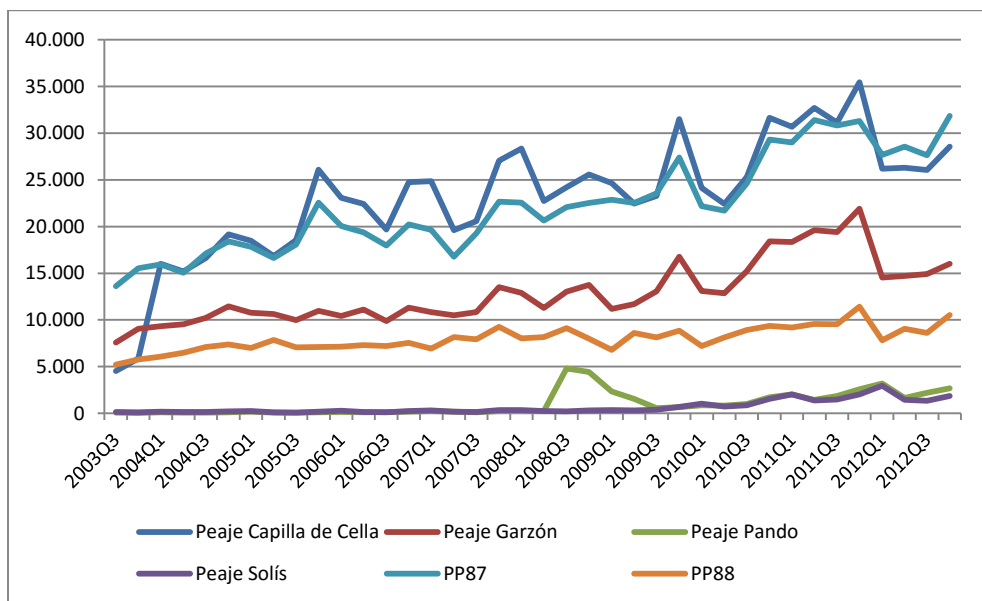
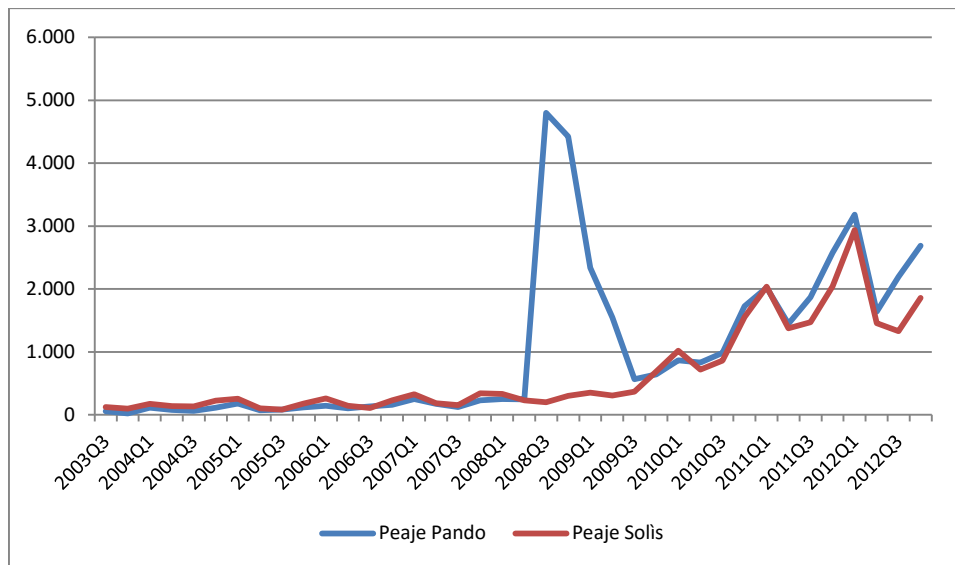


Gráfico 8: Tráfico de Camiones Pesados Grupo 6(b)



Finalmente, el Grupo 6 lo componen los puestos ubicados en el corredor Ruta Interbalnearia-Ruta 9, que como se mencionó anteriormente, presentan los menores volúmenes de tráfico, así como los desvíos estándar. La excepción a lo anterior son el Peaje Capilla de Cella y PP87, que presentan niveles de tráfico considerables. Las series muestran una estacionalidad débil, siendo más pronunciada para estos últimos dos puestos. Debido a que el volumen de tráfico de Peaje Pando, Peaje Solís es mucho menor al de las restantes series, se presentan estas últimas en otro gráfico (ver gráfico grupo 6(b)). Se puede apreciar como la serie Peaje Pando tuvo un crecimiento inusual durante el año 2008, multiplicando su tráfico por cinco.

4.2. Análisis Estadístico de las Series

4.2.1. Pruebas de Raíz Unitaria

Una serie es estacionaria (débil) si la media y las autocovarianzas de la serie no dependen del tiempo. Cuando a una variable es necesario realizarle una primera diferencia para transformarla en estacionaria, se dice que es integrada de orden 1, expresando $I(1)$, si es necesario diferenciarla dos veces, será $I(2)$.

Los procedimientos usuales de inferencia estadística no aplican a regresiones donde la variable dependiente o los regresores son variables integradas (orden 1 o mayor)⁸. De este modo, resulta necesario realizar un análisis de las series para ver si estas son estacionarias o no antes de realizar el estudio econométrico. Los métodos formales para estudiar la estacionariedad de las series son las pruebas de raíz unitaria.

⁸ Una excepción a lo mencionado se da en el caso de que las series sean integradas pero exista entre ellas una relación de largo plazo, por lo que se dice que las series están cointegradas. Ver Engler y Granger (1987).

En esta sección se presentan las pruebas de raíz unitaria para paneles de datos. Es reconocido en la literatura (Trapani (2004), Smith y Fuertes (2010), Breitung y Pesaran (2005)) que las pruebas de raíz de unitaria considerando las series de forma individual no tienen la potencia necesaria para rechazar la presencia de raíz unitaria cuando la serie parecería no ser $I(1)$, y en este sentido, se ha sugerido utilizar las pruebas en paneles para abordar el problema. Considerando estas últimas, se suele hacer una división entre las pruebas de primera generación, que no permiten dependencia entre los cross section y las denominadas de segunda generación que contemplan estas alternativas.

Pruebas de Raíz Unitaria en paneles de Primera Generación:

Dentro de las pruebas de primera generación, se puede realizar una separación entre aquellas pruebas que suponen una raíz unitaria común entre los cross section (Levin, Lin y Chu (2002), Breitung (2000), Hadri (2000)) y las que permiten que el coeficiente autorregresivo varíe entre las diferentes unidades (Im, Pesaran y Shin (2003), Maddala y Wu (1999) y Choi (2001)).

Para lograr una mejor comprensión de estas pruebas supongamos el siguiente proceso autorregresivo:

$$y_{it} = \rho_i y_{it-1} + \delta_i x_{it} + \varepsilon_{it}$$

Donde $i = 1 \dots N$, $t = 1 \dots T$, x_{it} representa un vector con variables exógenas que puede incluir efecto fijo o tendencia individual, ε_{it} se asume que es independiente e idénticamente distribuido entre las i y los t , y ρ_i es el coeficiente de interés para detectar la presencia de raíz unitaria, si $|\rho_i| = 1$ se dice que la serie presenta una raíz unitaria. Alternativamente la anterior expresión puede transformarse de la siguiente forma:

$$\Delta y_{it} = \alpha_i y_{it-1} + \gamma_i x_{it} + \varepsilon_{it} \quad \text{con } \alpha_i = \rho_i - 1$$

De esta forma, las pruebas que suponen raíz unitaria común, $\rho = \rho_i$ ($\alpha_i = \alpha$) $\forall i$, realizan la siguiente prueba de hipótesis:

$$\begin{aligned} H_0) \alpha &= 0 \\ H_1) \alpha &< 0 \end{aligned}$$

Mientras que las que permiten raíces unitarias individuales testean las siguientes hipótesis:

$$\begin{aligned} H_0) \alpha_i &= 0 \quad \forall i \\ H_1) \alpha_1 &< 0, \dots, \alpha_{N_0} < 0 \quad \text{con } N_0 \leq N \end{aligned}$$

Las primeras pruebas apilan las observaciones de todos los cross section y estiman los parámetros antes de construir el estadístico de la prueba, en tanto que las que permiten raíces individuales operan directamente sobre los estadísticos individuales usando promedios simples de los estadísticos u otras transformaciones adecuadas.

De este modo, se realizaron pruebas de raíces unitarias para paneles sobre las series de tráfico de autos considerando los dos tipos de pruebas. Más específicamente, se realizaron las pruebas de Levin, Lin y Chu (2002), Breitung (2000) y Hadri (2000) para la alternativa de homogeneidad entre los puestos; y las pruebas de Im, Pesaran y Shin (2003), ADF Fisher y PP Fisher (estas últimas propuestas por Maddala y Wu (1999) y Choi (2001)), para la alternativa de heterogeneidad.

A continuación se presentan los resultados de las pruebas:

Tabla 2: Pruebas de Raíces Unitarias para Paneles

Prueba de Raíz Común	Niveles (P-Valor)	Primer Diferencia (P-Valor)	Especificaciones (1)*	Resultado de Prueba en Niveles (3)
Levin, Lin and Chu	0.0000	0.0000	Constante	Rechazo Ho (Ho: No existencia de raíz unitaria común)
	1.0000	0.0000	Constante y tendencia	No Rechazo Ho (Ho: Existencia de raíz unitaria común)
	1.0000	0.0000	Sin constante ni tendencia	No Rechazo Ho (Ho: Existencia de raíz unitaria común)
Breitung	0.0208	0.0000	En la única especificación que permite el test (Const. y tend.)	No rechazo Ho (Ho: Existencia de raíz unitaria común)
Hadri (2)	0.0000	0.8241	Constante	Rechazo Ho (Ho: Existencia de raíz unitaria común)
	0.0000	0.0013	Constante y tendencia	Rechazo Ho (Ho: Existencia de raíz unitaria común)
Pruebas de Raíz Individual	Niveles (P-Valor)	Primer Diferencia (P-Valor)	Especificaciones (1)*	Resultado de Prueba en Niveles (3)
Im Pesaran Shin	0.1774	0.0000	Constante	No Rechazo Ho (Ho: Existencia de raíces unitarias individuales)
	0.0013	0.0000	Constante y tendencia	Rechazo Ho (Ho: No existencia de raíces unitarias individuales)
ADF Fisher	0.0109	0.0000	Constante	Rechazo Ho (Ho: No existencia de raíces unitarias individuales)
	0.0018	0.0000	Constante y tendencia	Rechazo Ho (Ho: No existencia de raíces unitarias individuales)
PP Fisher	0.0002	0.0000	Constante	Rechazo Ho (Ho: No existencia de raíces unitarias individuales)
	0.0000	0.0000	Constante y tendencia	Rechazo Ho (Ho: No existencia de raíces unitarias individuales)

*En las pruebas donde se requieren especificar lags de la variable dependiente en diferencias, se permitió un máximo de 12 lags, seleccionando el modelo por el criterio Schwarz. (Se probaron con otros lags máximos y criterios de información y los resultados se mantienen)

- (1) Las especificaciones consideradas son: sin constante ni tendencia, con constante, con constante y tendencia
- (2) Esta prueba considera como Ho) la no existencia de raíz unitaria común, de forma inversa a las anteriores
- (3) Se considera rechazada la hipótesis nula si el p-valor en la prueba en niveles es inferior a 0.10

Del cuadro anterior se concluye que las pruebas de raíz común señalan la existencia de raíz unitaria común en todas las pruebas consideradas, excepto para la prueba de Levin, Lin y Chu en la especificación con constante. Por el contrario, los resultados de las pruebas de raíces heterogéneas

arrojan el resultado inverso, se rechaza la existencia de raíz unitaria excepto para la prueba de Im, Pesaran y Shin, en la especificación con constante.

Adicionalmente, cuando se realizan las pruebas en la primera diferencia de la serie, todas las pruebas son coincidentes en rechazar la existencia de raíz unitaria, excepto en la prueba de Hadri en la especificación con constante y tendencia.

Sin embargo, Maddala y Wu (1999) a través de simulaciones presentan evidencia sobre el pobre desempeño de las pruebas de Levin y Lin (1992) e Im, Pesaran y Shin (2003), cuando existe dependencia entre las unidades del panel, siendo las pruebas Fisher las que muestran un mejor desempeño en una variedad de situaciones. Debido a esto, resulta conveniente realizar una prueba de raíz unitaria de segunda generación, es decir, pruebas que permiten que las unidades del panel estén correlacionadas. En este sentido, se presenta a continuación la prueba de raíz unitaria de Pesaran (2007).

Pruebas de Raíz Unitaria en paneles de Segunda Generación: Pesaran (2007)

En las pruebas anteriores se asume que las observaciones y_{it} , son independientes entre las i , sin embargo, en nuestra aplicación, es necesario considerar que las unidades del panel estén contemporáneamente correlacionadas. Esta correlación puede deberse a una variedad de factores, como por ejemplo, factores comunes omitidos y observables, factores comunes no observables, efectos de derrame espacial, o interdependencia entre los residuos que pueden permanecer incluso luego de considerar todos los efectos comunes observables y no observables. Dejando de lado los factores comunes observables y la correlación serial de los residuos, Pesaran (2007) parte de la siguiente especificación, considerando nuevamente la siguiente ecuación:

$$\Delta y_{it} = \alpha_i y_{it-1} + \gamma_i x_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \lambda_i f_t + \omega_{it}$$

Donde f_t es un factor común no observado, serialmente no correlacionado con media cero y varianza constante e igual a uno; y ω_{it} es el error individual específico independientemente distribuido para t e i . Las variables λ_i y ω_{it} se asume que son independientes para todo i . Combinando las dos expresiones:

$$\Delta y_{it} = \alpha_i y_{it-1} + \gamma_i x_{it} + \lambda_i f_t + \omega_{it}$$

De este modo, Pesaran (2007), plantea la prueba de hipótesis de forma análoga a la prueba de coeficientes heterogéneos presentada antes:

$$\begin{aligned} H_0) \alpha_i &= 0 \quad \forall i \\ H_1) \alpha_1 &< 0, \dots, \alpha_{N_0} < 0 \quad \text{con } N_0 \leq N \end{aligned}$$

La prueba de Pesaran (2007) consiste en realizar una prueba de raíz unitaria ADF estándar para cada i a la regresión aumentada con las medias de las cross section en niveles y en diferencias rezagadas (CADF). Estas últimas, se puede demostrar, ofician de proxies de los factores comunes no observables f_t , y su inclusión en la regresión filtra asintóticamente el efecto del factor. De esta forma, la prueba de hipótesis se base en el t ratio de la estimación del coeficiente α_i por MCO de la siguiente regresión:

$$\Delta y_{it} = \alpha_i y_{it-1} + \gamma_i x_{it} + c_i \bar{y}_{t-1} + d_i \Delta \bar{y}_t + \omega_{it}$$

Donde $\bar{y}_t = N^{-1} \sum_{i=1}^N y_{it}$ y $\Delta \bar{y}_t = N^{-1} \sum_{i=1}^N \Delta y_{it}$.

Los valores críticos del estadístico t bajo la hipótesis nula son presentados en Pesaran (2007) en las tres especificaciones usuales (sin constante ni tendencia, con constante, y con constante y tendencia). Una vez obtenidos los estadísticos t de cada cross section, Pesaran (2007) presenta la prueba de raíz unitaria para paneles de forma similar a la prueba de Im, Pesaran y Shin (2003), realizando el promedio de los estadísticos t computando sus valores críticos. A su vez, propone alternativamente otro estadístico que se distribuye normal bajo la hipótesis nula de la forma:

$$Z(N, T) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \Phi^{-1}(p_{iT})$$

Donde p_{iT} es el valor-p del individuo i resultante de la prueba CADF particular.

A continuación se presenta la prueba de raíz unitaria de Pesaran (2007):

Tabla 3: Prueba de Raíz Unitaria para paneles correlacionados, Pesaran (2007)

Especificación con constante							
lags	Z	valor-p	Resultado	t-bar	Resultado		
					valor crítico 10%= -2.080	valor crítico 5%= -2.160	valor crítico 1%= -2.300
0	-11.900	0.000	Rechazo Ho	-4.103	Rechazo Ho	Rechazo Ho	Rechazo Ho
1	-11.288	0.000	Rechazo Ho	-3.982	Rechazo Ho	Rechazo Ho	Rechazo Ho
2	-4.885	0.000	Rechazo Ho	-2.727	Rechazo Ho	Rechazo Ho	Rechazo Ho
3	-1.863	0.031	Rechazo Ho	-2.135	Rechazo Ho	No Rechazo Ho	No Rechazo Ho
4	-0.763	0.223	No Rechazo Ho	-1.919	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
5	0.426	0.665	No Rechazo Ho	-1.687	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
6	2.234	0.987	No Rechazo Ho	-1.332	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
7	2.609	0.995	No Rechazo Ho	-1.259	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
8	1.410	0.921	No Rechazo Ho	-1.494	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
Especificación con constante y tendencia							
lags	Z	valor-p	Resultado	t-bar	Resultado		
					valor crítico 10%= -2.580	valor crítico 5%= -2.650	valor crítico 1%= -2.780
0	-10.315	0.000	Rechazo Ho	-4.233	Rechazo Ho	Rechazo Ho	Rechazo Ho
1	-9.795	0.000	Rechazo Ho	-4.137	Rechazo Ho	Rechazo Ho	Rechazo Ho
2	-2.667	0.004	Rechazo Ho	-2.829	Rechazo Ho	Rechazo Ho	Rechazo Ho
3	-2.094	0.910	No Rechazo Ho	1.341	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho

4	-1.847	0.996	No Rechazo Ho	2.687	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
5	-1.550	1.000	No Rechazo Ho	4.306	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
6	-1.147	1.000	No Rechazo Ho	6.500	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
7	-0.837	1.000	No Rechazo Ho	8.192	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho
8	-1.267	1.000	No Rechazo Ho	5.846	No Rechazo Ho	No Rechazo Ho	No Rechazo Ho

Como se puede apreciar de la tabla, una vez se incorporan 3 o 4 lags a las regresiones de cada cross section para tomar en cuenta la correlación serial de las series, las pruebas de raíz unitaria en ambas especificaciones consideradas arrojan los mismos resultados: no se rechaza la presencia de una raíz unitaria en el panel de datos. De acuerdo a las simulaciones presentadas por el autor en su trabajo, esta prueba es la que presenta un mejor comportamiento en presencia de correlación entre las unidades del panel.

De esta forma, se considera que la serie camiones pesados presenta raíz unitaria, es integrada de orden uno.

4.2.2. Cointegración

El propósito de esta sección es determinar la existencia de cointegración entre el tráfico de camiones y las variables explicativas que se consideran según nuestro modelo teórico. Mediante el estudio de cointegración podremos detectar y analizar la existencia de una relación de largo plazo en el tiempo entre las variables consideradas. Dado nuestro estudio en la sección previa, donde la mayoría de las series parecería ser $I(1)$, se requiere diferenciar las series (o quitarles la tendencia temporal) de forma de transformarlas en estacionarias y evitar el problema de regresión espurias, entre otros. Sin embargo, al transformar de esta forma la serie se podría perder información valiosa sobre la relación de largo plazo entre las series, por lo que el presente análisis provee un marco para retener la información de corto y largo plazo, enriqueciendo el estudio.

Se dice que dos variables están cointegradas, si dado dos series $I(1)$, la combinación lineal de ellas es $I(0)$, por lo que existe una relación de equilibrio de largo plazo entre las variables, denominada relación de cointegración. En caso de considerar más de dos variables, puede existir más de una relación de cointegración entre ellas.

La literatura que trata cointegración en datos de panel, al igual que en el caso del estudio de raíz unitaria, puede ser dividida en enfoques de primera generación donde no se consideran los posibles efectos de factores comunes no observables, y los enfoques de segunda generación donde se comienzan a tener en cuenta la dependencia entre las unidades del panel. Sin embargo, como señalan Breitung y Pesaran (2005) este último enfoque no está desarrollado completamente y es sujeto de actuales investigaciones.

A su vez, dentro del enfoque de primera generación existen dos aproximaciones al estudio de cointegración, el enfoque que considera los residuos de una regresión por un lado (Kao (1999), Pedroni (1999,2001,2004) y Westerlund (2005)), y la utilización de sistemas por otro (Larsson, Lyhagen y Lothgren (2001)).

El enfoque residual parte de la idea de Engler y Granger (1987) para series temporales, se estima una regresión por MCO y luego se realizan pruebas de raíz unitaria ADF en los residuos (considerando valores críticos apropiados), si los residuos son $I(0)$ las variables están cointegradas.

Este es esencialmente el procedimiento sugerido por Pedroni (1999,2004), y el que se aplicará en el presente estudio.

Pruebas de cointegración de primera generación: Pedroni (1999,2004)

El autor parte de la siguiente especificación:

$$y_{it} = \alpha_i + \delta_i t + \beta_{1i} x_{1i,t} + \beta_{2i} x_{2i,t} + \dots + \beta_{Mi} x_{Mi,t} + e_{i,t}$$

Para $t = 1, \dots, T$; $i = 1, \dots, N$; $m = 1, \dots, M$; donde y y x se asume que son integradas de orden 1. Los parámetros α_i y δ_i son efectos individuales y tendencias individuales que pueden ser definidas en cero.

Bajo la hipótesis nula de no cointegración, los residuos $e_{i,t}$ serían I(1). De este modo, el procedimiento consiste en obtener los residuos de la regresión anterior y luego realizar un test sobre estos para definir su orden de integración. Para ello se considera la siguiente regresión auxiliar para cada cross section:

$$e_{i,t} = \rho_i e_{i,t-1} + \sum_{j=1}^n \psi_{ij} \Delta e_{i,t-j} + v_{ij}$$

El autor describe varios métodos para construir estadísticos para testear la hipótesis nula de no cointegración ($\rho_i = 1$). El autor presenta siete estadísticos, cuatro de estos se utilizan para testear la hipótesis alternativa homogénea, ($\rho = \rho_i < 1$), apilando las observaciones; mientras que los restantes se consideran para testear la hipótesis alternativa heterogénea, $\rho_i < 1$ para todo i .

En el primer conjunto de pruebas, tres requieren el uso de correcciones no paramétricas, como el caso de Phillips y Perron (1988), y la cuarta es una pruebas paramétrica ADF. En el segundo conjunto de pruebas, dos usan correcciones no paramétricas, y la restante una prueba ADF. Los estadísticos generados tiene diferentes propiedades, tamaño y fuerza, para diferentes valores de N y T .

Las series incluidas en la prueba fueron las presentadas en la sección 4.2.3. conjuntamente con las series de tráfico de camiones expresadas en logaritmos. A continuación se presentan los resultados de las pruebas de Pedroni para especificaciones con constante y con constante y tendencia.

Tabla 4: Test de Cointegración de Pedroni (1999,2004), especificación con constante; y constante y tendencia

Especificación con constante			
Hipótesis Alternativa: coeficiente AR común.			
Estadístico	Valor	Prob.	Resultado
Panel v-Statistic	5.936835	0.0000	Rechazo Ho
Panel rho-Statistic	-6.780674	0.0000	Rechazo Ho
Panel PP-Statistic	-10.07851	0.0000	Rechazo Ho
Panel ADF-Statistic	-7.501429	0.0000	Rechazo Ho
Hipótesis Alternativa: coeficientes AR individuales			
Estadístico	Valor	Prob.	Resultado
Group rho-Statistic	-3.683405	0.0001	Rechazo Ho

Group PP-Statistic	-12.70342	0.0000	Rechazo Ho
Group ADF-Statistic	-4.756923	0.0000	Rechazo Ho
Especificación con constante y tendencia			
Hipótesis Alternativa: coeficiente AR común.			
<u>Estadístico</u>	<u>Valor</u>	<u>Prob.</u>	<u>Resultado</u>
Panel v-Statistic	3.790323	0.0001	Rechazo Ho
Panel rho-Statistic	-3.207143	0.0007	Rechazo Ho
Panel PP-Statistic	-8.885718	0.0000	Rechazo Ho
Panel ADF-Statistic	-4.361025	0.0000	Rechazo Ho
Hipótesis Alternativa: coeficientes AR individuales			
<u>Estadístico</u>	<u>Valor</u>	<u>Prob.</u>	<u>Resultado</u>
Group rho-Statistic	-1.917976	0.0276	Rechazo Ho
Group PP-Statistic	-16.31368	0.0000	Rechazo Ho
Group ADF-Statistic	-3.100027	0.0010	Rechazo Ho

Como se puede apreciar del cuadro, todos los test propuestos rechazan la hipótesis nula de no cointegración entre las series.

Pruebas de cointegración de segunda generación: Westerlund (2007)

Una debilidad de los test anteriores, basados en el estudio de los residuos, es que requieren que los parámetros de largo plazo de las variables en niveles sean iguales a los parámetros de corto plazo para las variables en diferencia. Esta restricción puede causar una significativa pérdida de poder en estas aproximaciones. En este sentido, Westerlund (2007) desarrolla cuatro pruebas de cointegración para paneles basados en dinámicas estructurales, y no en residuales, salvando la anterior restricción.

La idea es testear la hipótesis nula de no cointegración infiriendo si el término de corrección del error en un modelo de corrección del error para paneles es igual a cero. Los nuevos test se distribuyen normales, permiten considerar dinámicas de corto plazo por unidad, tendencia individual y coeficientes específicos en cada cross section, y dependencia entre los cross section. De este modo, dos test son diseñados para testear la hipótesis alternativa homogénea de cointegración para todo el panel, mientras que los otros dos testean la alternativa de al menos una unidad cointegrada.

El modelo de corrección de error asume el siguiente proceso de generación de los datos:

$$\Delta y_{it} = C_i + \delta_i t + \alpha_i (y_{i,t-1} - \beta_i' x_{i,t-1}) + \sum_{j=1}^{p_i} \alpha_{ij} \Delta y_{i,t-j} + \sum_{j=-q_i}^{p_i} \gamma_{ij} \Delta x_{i,t-j} + e_{it}$$

Para $t = 1, \dots, T$; $i = 1, \dots, N$; refiriendo las unidades de cross section y las series de tiempo respectivamente. Por simplicidad, se modela el vector K dimensional $x_{i,t}$ como un random walk puro de forma que $\Delta x_{i,t}$ es independiente de e_{it} , y se asume que los errores e_{it} son independientes entre los cross section y las series de tiempo.

La dependencia entre los puestos de conteo, las unidades de cross section i , se contempla utilizando técnicas de bootstrapping y computando errores estándar robustos.

La anterior expresión puede ser rescrita de la siguiente forma:

$$\Delta y_{it} = C_i + \delta_i t + \alpha_i y_{i,t-1} + \lambda_i' x_{i,t-1} + \sum_{j=1}^{p_i} \alpha_{ij} \Delta y_{i,t-j} + \sum_{j=-q_i}^{p_i} \gamma_{ij} \Delta x_{i,t-j} + e_{it} \quad (1)$$

Donde $\lambda_i' = -\alpha_i \beta_i'$. El parámetro α_i determina la velocidad en que el sistema se corrige hacia la relación de equilibrio $(y_{i,t-1} - \beta_i' x_{i,t-1})$ luego de recibir shocks.

Si $\alpha_i < 0$, existe una corrección del error y por lo tanto las variables están cointegradas.

Si $\alpha_i = 0$, no existe corrección del error, y por tanto las variables no están cointegradas.

De esta forma, el autor plantea cuatro test de cointegración, dos de ellos asumiendo homogeneidad del parámetro de corrección ($\alpha_i = \alpha$) para todos los puntos de conteo; y otros dos no considerando este supuesto.

Estos últimos se basan en los siguientes estadísticos, construidos a partir de estimaciones MCO del parámetro α en (1):

$$G_{\tau} = \frac{1}{N} \sum_{i=1}^N \frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)}, \quad G_{\alpha} = \frac{1}{N} \sum_{i=1}^N \frac{T \hat{\alpha}_i}{\hat{\alpha}_i(1)}$$

Donde $SE(\hat{\alpha}_i)$ es el error estándar de $\hat{\alpha}_i$, y $\hat{\alpha}_i(1) = 1 - \sum_{j=1}^{p_i} \alpha_{ij}$, es estimado a partir de un estimador kernel.

Luego se realiza la prueba:

$$\begin{aligned} H_0) \alpha_i &= 0 \quad \forall i \\ H_A) \alpha_i &< 0 \quad \text{para al menos un } i \end{aligned}$$

Para el caso de los estadísticos que agrupan toda la información, y por tanto asumen homogeneidad del parámetro de corrección del error, se realiza la siguiente prueba de hipótesis:

$$\begin{aligned} H_0) \alpha_i &= 0 \quad \forall i \\ H_A) \alpha_i &= \alpha < 0 \quad \forall i \end{aligned}$$

Utilizando los estadísticos:

$$P_{\tau} = \frac{\hat{\alpha}}{SE(\hat{\alpha})}, \quad P_{\alpha} = T \hat{\alpha}$$

A continuación se presentan los resultados de las pruebas propuestas por Westerlund (2007) para nuestro caso.

Tabla 5: Test de Cointegración de Westerlund (2007), especificación con constante; y constante y tendencia

Especificación con constante				
Hipótesis Alternativa: coeficiente de corrección del error homogéneo				
Estadístico	Valor	p-valor	p-valor robusto	Resultado
P_{τ}	-3.658	0.996	0.470	No Rechazo Ho
P_{α}	-1.962	1.000	0.896	No Rechazo Ho
Hipótesis Alternativa: coeficiente de corrección del error heterogéneo				
Estadístico	Valor	p-valor	p-valor robusto	Resultado
G_{τ}	-1.705	1.000	0.704	No Rechazo Ho
G_{α}	-1.814	1.000	0.814	No Rechazo Ho
Especificación con constante y tendencia				
Hipótesis Alternativa: coeficiente de corrección del error homogéneo				
Estadístico	Valor	p-valor	p-valor robusto	Resultado
P_{τ}	-5.200	1.000	0.070	No Rechazo Ho
P_{α}	-1.165	1.000	0.862	No Rechazo Ho
Hipótesis Alternativa: coeficiente de corrección del error heterogéneo				
Estadístico	Valor	p-valor	p-valor robusto	Resultado
G_{τ}	-1.585	1.000	0.930	No Rechazo Ho
G_{α}	-0.607	1.000	0.774	No Rechazo Ho

Contrariamente a las pruebas de primera generación, los test propuestos rechazan la existencia de cointegración de las series.

4.2.3. Dependencia

Para examinar la dependencia de las series de tráfico de camiones pesados, se analiza en una primera instancia el coeficiente de correlación promedio en valor absoluto y el porcentaje de variación explicado por los primeros 4 componentes principales. A su vez, se presentan estos coeficientes para los residuos de las series obtenidas de la estimación MCO de cada serie utilizando como regresores el IVF, el precio de la nafta, el precio del peaje, y dummies estacionales y efecto pascua. Los resultados se presentan en la siguiente tabla:

Tabla 6: Coeficientes de correlación y componentes principales

	Tráfico de camiones	Residuos MCO
Correlación promedio	0,7561	0,2748
V(PC 1)	0.7755	0.2945
V(PC 2)	0.0638	0.1929
V(PC 3)	0.0470	0.0937
V(PC 4)	0.0298	0.0803

Como puede observarse las series de tráfico de camiones tienen un coeficiente de correlación promedio elevado. Por otra parte, los primeros cuatro componentes principales dan cuenta del casi 92% de la variación de las series, siendo el primer componente muy importante para explicar la variación, 77%. En cuanto a los residuos, el coeficiente de correlación promedio cae a 0,2748, y los primeros cuatro factores explican el 66% de la variación.

Pesaran (2004) propone un test formal para testear la dependencia en el panel, que recoge las debilidades del test LM de Breusch y Pagan cuando N es grande y T pequeño. En este sentido, propone el siguiente estadístico:

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right) \sim N(0,1)$$

donde $\hat{\rho}_{ij}$ es el coeficiente de correlación muestral de los residuos, específicamente,

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^T e_{it} e_{jt}}{(\sum_{t=1}^T e_{it}^2)^{1/2} (\sum_{t=1}^T e_{jt}^2)^{1/2}}$$

y e_{it} son los residuos resultantes de la siguiente regresión estimada por MCO considerando las series de forma individual:

$$e_{it} = y_{it} - \hat{\alpha}_i - \mathbf{x}_{it} \hat{\boldsymbol{\beta}}_i$$

El autor prueba que la distribución asintótica es válida para N y T tendiendo a infinito en cualquier orden, es robusto para series no estacionarias, parámetros heterogéneos y cambios estructurales, y prueba que tiene buenas propiedades en pequeñas muestras.

En nuestro caso particular, el test de dependencia arroja los siguientes resultados para la hipótesis nula de independencia de las series:

Tabla 7: Test de Dependencia entre cross section de Pesaran (2004)

Variable	CD-test	p-valor	Resultado
Tráfico de camiones	76,14	0.0000	Rechazo Ho

De acuerdo a estos resultados, se deberá considerar una estructura de dependencia entre los puestos de conteo para las estimaciones.

4.2.4. Heterogeneidad

En los modelos clásicos de datos de panel se asume que la heterogeneidad no observada en cada unidad es capturada por la constante individual, considerándola fija o aleatoria. Sin embargo, en numerosas aplicaciones, no se puede descartar que las elasticidades respecto a las variables de interés también sean particulares de cada cross section o individuo (o puesto de conteo en nuestro caso). Ignorar esta heterogeneidad podría resultar en sesgos de estimación e inferencia, y en este sentido, es importante testear la homogeneidad de los parámetros.

De esta forma, Pesaran y Yamagata (2008) proponen diversos test cuando N y T tienen conjuntamente a infinito⁹, partiendo del estadístico propuesto por Swamy (1970) para paneles con N pequeño.

Así partiendo del estadístico de Swamy:

$$\hat{S} = \sum_{i=1}^N (\hat{\beta}_i - \hat{\beta}_{WFE})' \frac{X_i' M_{\tau} X_i}{\hat{\sigma}_i^2} (\hat{\beta}_i - \hat{\beta}_{WFE})$$

donde

$$\hat{\sigma}_i^2 = \frac{(y_i - X_i \hat{\beta}_i)' M_{\tau} (y_i - X_i \hat{\beta}_i)}{(T - k - 1)}$$

con k el número de regresores estrictamente exógenos y $M_{\tau} = I_T - \tau_T (\tau_T' \tau_T)^{-1} \tau_T'$ siendo τ_T un vector de unos de dimensión $T \times 1$.

y $\hat{\beta}_{WFE}$ es estimador pool de efecto fijo ponderado definido de la siguiente forma:

$$\hat{\beta}_{WFE} = \left(\sum_{i=1}^N \frac{X_i' M_{\tau} X_i}{\hat{\sigma}_i^2} \right)^{-1} \sum_{i=1}^N \frac{X_i' M_{\tau} y_i}{\hat{\sigma}_i^2}$$

⁹ Con condiciones de expansión en algunos casos, como ser $\frac{\sqrt{N}}{T}$ o $\frac{\sqrt{N}}{T^2}$.

Pesaran y Yamagata proponen el siguiente estadístico:

$$\hat{\Delta} = \sqrt{N} \left(\frac{N^{-1}S - k}{\sqrt{2k}} \right) \rightarrow_d N(0,1) \text{ cuando } (N, T) \rightarrow \infty \text{ conjuntamente y } \frac{\sqrt{N}}{T}$$

A su vez, proponen el siguiente estadístico ajustado que muestra mejores propiedades en muestras pequeñas bajo el supuesto adicional de errores distribuidos normales:

$$\hat{\Delta}_{adj} = \sqrt{N} \left(\frac{N^{-1}\hat{S} - E(\hat{z}_{iT})}{\sqrt{Var(\hat{z}_{iT})}} \right)$$

$$E(\hat{z}_{iT}) = \frac{k(T - k - 1)}{T - K - 3}, \quad Var(\hat{z}_{iT}) = \frac{2k(T - k - 1)^2(T - 3)}{(T - k - 3)^2(T - k - 5)}$$

En nuestro caso particular, el test de heterogeneidad arroja los siguientes resultados para la hipótesis nula de homogeneidad de los parámetros β ($H_0: \beta_i = \beta$ para todo i ; y $H_1: \beta_i \neq \beta$ para una fracción no nula de pares de parámetros para $i \neq j$)

Tabla 8: Test de Heterogeneidad de los parámetros, Pesaran Yamagata (2008)

Estadístico Pesaran- Yamagata	Valor	p-valor	Resultado
$\hat{\Delta}$	51,9934	0.0000	Rechazo H_0
$\hat{\Delta}_{adj}$	41,8252	0.0000	Rechazo H_0

Considerando los resultados, es necesario considerar modelos que contemplen la heterogeneidad de los parámetros.

4.3. Variables Explicativas

Por último se presentan los datos de las variables explicativas consideradas, tomando en cuenta las variables que determinan la evolución la evolución del tráfico de carga interurbano. En este sentido, se presentan las tres variables sobre las que se pretende hallar la elasticidad del tráfico:

Nivel de Actividad de la Economía: para esta variable se utilizó el Índice de Volumen Físico del Producto Bruto Nacional calculada por el Banco Central del Uruguay.

Precio de la Nafta: para esta variable se consideró el precio de la nafta en Uruguay durante el período de estudio. Este precio es fijado por el organismo estatal denominado Administración Nacional de Combustibles Alcohol y Portland.

Precio del Peaje: los puestos de conteo considerados están ubicados en Rutas Nacionales, las cuales tienen un costo para los usuarios de las mismas. Este precio es fijado por el Poder Ejecutivo.

Respecto a este precio, durante el período de estudio considerado, casi todos los peajes existentes tienen la misma evolución de precios, con excepción del puesto de conteo Manuel Díaz que cobra un precio un 80% menor al vigente en cada momento del tiempo en el resto de los peajes durante el periodo de estudio, Peaje Queguay que cobra un 80% menor al precio hasta 06/2006, y Peaje Garzón que cobra un precio un 80% menor hasta 05/2009.

En cuanto al orden de integración de las series, a continuación se presenta el siguiente cuadro con los resultados de las pruebas de raíz unitaria:

Tabla 9: Pruebas de Raíz Unitaria de las variables explicativas

Variable	Especificación	Test*	p-valor	Resultado**
IVF	Constante	ADF	0.8621	No Rechazo Ho
	Constante y tendencia	ADF	0.0003	Rechazo Ho
	Constante	Phillips y Perron	0.9784	No Rechazo Ho
	Constante y tendencia	Phillips y Perron	0.0003	Rechazo Ho
Precio Nafta	Constante	ADF	0.4981	No Rechazo Ho
	Constante y tendencia	ADF	0.6250	No Rechazo Ho
	Constante	Phillips y Perron	0.4558	No Rechazo Ho
	Constante y tendencia	Phillips y Perron	0.5632	No Rechazo Ho
Precio Peaje***	Constante	ADF	0.9988	No Rechazo Ho
	Constante y tendencia	ADF	0.9473	No Rechazo Ho
	Constante	Phillips y Perron	0.9033	No Rechazo Ho
	Constante y tendencia	Phillips y Perron	0.8144	No Rechazo Ho

*Para los test ADF se seleccionaron los lags del modelo de acuerdo al criterio de Schwartz.

** Ho: Existencia de Raíz Unitaria en ambos test

*** Las series correspondientes a Peaje Manuel Díaz, Peaje Queguay y Peaje Garzón exhiben los mismos resultados.

5. RESULTADOS

En esta sección se presentan los resultados de las estimaciones de los diferentes modelos descritos en el capítulo 3¹⁰.

La especificación finalmente estimada en todos los modelos (a menos que se aclare posteriormente) considera a las variables en logaritmos, y adopta la siguiente expresión:

$$\log(y_{it}) = \alpha_i + \beta_{i1} \log(IVF) + \beta_{i2} \log(\text{Precio Nafta}) + \beta_{i3} \log(\text{Precio Peaje}) \\ + \beta_{i4} \text{Trim1} + \beta_{i5} \text{Trim2} + \beta_{i6} \text{Trim3} + \beta_{i7} \text{Efecto Pascua} + u_{it}$$

donde y_{it} es el tráfico de camiones pesados, con $i=1 \dots 23$ representando los 23 puestos de conteo, $t = 1 \dots 38$ las observaciones temporales, y Trim_i y Efecto Pascua son variables dummies estacionales que intentan captar el efecto de pertenecer de la observación a determinado trimestre y aislar el efecto pascua respectivamente.

Todos los modelos no consideran Peaje Pando ni Peaje Solís, ya que las observaciones de estos puestos resultan en datos atípicos en las estimaciones, según diferentes medidas. Estos resultados parecerían ser razonables dado que los peajes mencionados pertenecen a la Ruta Interbalnearia donde no está permitido circular camiones pesados, y el volumen de tráfico captado en los datos debe responder a otros determinantes que los considerados para las Rutas Nacionales libres de restricciones de circulación.

¹⁰ Ver Anexo para los modelos completos

5.1. Modelos Estáticos

Las estimaciones de los modelos estáticos se presentan a continuación en la siguiente tabla. Cabe mencionar que los modelos Pool FGLS consideran una estructura dinámica en los errores, así como los modelos FE-AR(1) y RE-AR(1):

Tabla 10: Resultados de Elasticidades Modelos Estáticos

	POOL MCO	POOL FGLS AR(4)	POOL FGLS AR(8)	POOL FGLS AR(12)	FE	FE AR(1)	RE	RE AR(1)	FD
log(IVF)	1.5767 (0.3003)**	1.3404 (0.1429)**	1.3387 (0.1500)**	1.3585 (0.1517)**	1.6546 (0.1893)**	1.5221 (0.1003)**	1.6543 (0.1893)**	1.5221 (0.1003)**	0.9459 (0.2014)**
log(Precio Peaje)	0.2172 (0.6576)	0.2689 (0.1083)**	0.3277 (0.1281)**	0.2563 (0.1299)**	-0.4427 (0.1950)	0.2221 (0.1718)	-0.4313 (0.1950)	0.2221 (0.1718)	0.5398 (0.1645)
log(Precio Nafta)	-0.0080 (0.1368)	0.1261 (0.1115)	0.0985 (0.1158)	0.0538 (0.1110)	-0.2332 (0.1232)	0.0636 (0.0935)	-0.2325 (0.1232)	0.0636 (0.0935)	0.1767 (0.1020)

Nota: Variable dependiente log(Tráfico de camiones). N=21, T=38.

*,** indican significación al 5% y 1% respectivamente

En paréntesis se presentan los errores estándar robustos a heteroscedasticidad de White.

Se omiten los reportes de la constante, de las dummies estacionales y efecto pascua

Los resultados de los diferentes modelos arrojan resultados similares en cuanto a la significación de las variables. Mientras la variable log(IVF) es significativa al 1% en todos los modelos, las restantes parecerían no tener un efecto significativo sobre el tráfico, excepto log(Precio Peaje) en los modelos Pool-FGLS AR. En cuanto

al valor de la elasticidad ingreso, el mismo se encuentra entre 0.9459 (en el modelo en diferencias (FD)) y 1.6546 (en el modelo de Efectos Fijos). Es posible que los resultados del modelo FD estén mostrando la dinámica de corto plazo entre las variables, mientras que el resto reflejen la relación de largo plazo.

5.1. Modelos Estáticos con Correlación entre Cross Section (Dependencia)

Los modelos presentados a continuación contemplan la dependencia entre los cross section, permitiendo que los errores de los puestos de conteo estén correlacionados en un mismo tiempo. A su vez, estos modelos permiten incorporar la dinámica de las variables considerando la correlación serial dentro de cada puesto de conteo, a excepción del modelo SURE.

Tabla 11: Resultados de Elasticidades Modelos con Correlación

	SURE	POOL FGLS Corr - AR(1)	POOL FGLS 21 Corr AR(1)	POOL MCO Corr - AR(1)	POOL MCO 21 Corr AR(1)	<u>DISROLL</u> <u>AND KRAY</u> <i>(1998)</i>	<u>DISROLL</u> <u>AND KRAY</u> <i>(1998) -FE</i>
log(IVF)	1.5701 (0.2756)**	1.4256 (0.1059)**	1.3843 (0.0582)**	1.2907 (0.1743)**	1.3214 (0.1355)**	1.5767 (0.0854)**	1.6545 (0.0792)**
log(Precio Peaje)	0.2213 (0.0299)**	0.4683 (0.0918)**	0.3963 (0.0909)**	0.4897 (0.1955)*	0.4295 (0.1723)*	0.2172 (0.1723)	-0.0443 (0.1685)
log(Precio Nafta)	0.0111 (0.0374)	0.1540 (0.0698)*	0.1797 (0.0567)**	0.1424 (0.1133)	0.1641 (0.1075)	-0.0080 (0.1590)	-0.0233 (0.1553)

Nota: Variable dependiente log(Tráfico de camiones). N=21, T=38.

*,** indican significación al 5% y 1% respectivamente

Se omiten los reportes de la constante y de las dummies estacionales y efecto pascua.

Corr - AR(1): Modelo FGLS/MCO con correlación contemporánea entre los cross section y la misma autocorrelación serial estimada para todos los puestos (valor estimado 0.8473)

21 Corr – AR(1): Modelo FGLS/MCO con correlación contemporánea entre los cross section y 21 autocorrelaciones seriales estimadas específicas para cada puesto.

Discroll and Kraay (1998): Se aplica el método de Discroll and Kraay (1998) para obtener errores estándar que permiten autocorrelación de una forma general hasta m lags y correlación. (Spatial correlation)

Discroll and Kraay (1998) – FE: Se estima el modelo de efectos fijos aplicando el método de Discroll and Kraay para calcular los errores robustos a correlación entre cross section y autocorrelación serial.

Los resultados de los modelos que contemplan dependencia arrojan resultados similares a los modelos anteriores en cuanto a la significación y el valor de la elasticidad ingreso.

Respecto de la elasticidad precio del peaje, en la mayoría de los modelos resulta significativa y positiva, al 5% y al 1%, y el valor se encuentra entre 0.2213 y 0.4897. El signo del coeficiente no coincide con lo que predice la teoría, por lo que se podría suponer que esta elasticidad no está captando el efecto buscado.

5.2. Modelos Dinámicos Arellano-Bond:

Estos modelos son esencialmente de efectos fijos eliminando el efecto fijo de cada puesto haciendo una primera diferencia, e incluyendo lags de la variable dependiente o de los regresores como variables explicativas para contemplar los ajustes temporales. Dado que el estimador FE es inconsistente en la estimación de los coeficientes de las variables rezagadas, se incluyen lags apropiados de la variable dependiente y de los regresores como instrumentos para la estimación. Estos modelos suponen que no existe la correlación serial, y esta es captada por los rezagos.

El siguiente modelo considera 8 rezagos de la variable dependiente con un máximo de 4 lags utilizados como instrumentos y errores calculados robustos a heteroscedasticidad.

Tabla 12: Resultados de Elasticidades Modelos Dinámicos

	ARELLANO- BOND
log(IVF)	1.2439 (0.2421)**
log(Precio Peaje)	-0.0741 (0.2461)
log(Precio Nafta)	-0.1539 (0.0995)

Nota: Variable dependiente log(Tráfico de camiones). N=21, T=38.

*,** indican significación al 5% y 1% respectivamente

En paréntesis se presentan los errores estándar robustos a heteroscedasticidad

Se omiten los reportes de la constante, de las dummies estacionales y efecto pascua

Nuevamente se obtienen resultados similares a los anteriores. La elasticidad ingreso es significativa, con un valor más bajo al hallado con anterioridad, y el resto de las variables no resultan significativas.

5.3. Modelos con Parámetros Heterogéneos y Dependencia

En esta sección se presentan los modelos que consideran heterogeneidad de los parámetros estimados, y dos modelos de componentes principales que contemplan dependencia entre los cross section, uno de estos simultáneamente con heterogeneidad (MG-PC).

Tabla 13: Resultados de Elasticidades Modelos con Parámetros Heterogéneos

	MG	RCM	MG-PC	FE-PC
log(IVF)	1.5442 (0.1271)**	1.6626 (0.1809)**	1.5477 (0.1266)**	1.6609 (0.1920)**
log(Precio Peaje)	-0.1563 (0.1993)	-0.0933 (0.2113)	-0.1929 (0.1874)	-0.0653 (0.2003)
log(Precio Nafta)	0.0753 (0.0903)	-0.0393 (0.1289)	0.0749 (0.0780)	-0.0245 (0.1238)

Nota: Variable dependiente log(Tráfico de camiones). N=21, T=38.

*,** indican significación al 5% y 1% respectivamente

Se omiten los reportes de la constante, de las dummies estacionales y efecto pascua

En paréntesis se presentan los errores estándar robustos a heteroscedasticidad

MG: Mean Group Estimator (Pesaran y Smith (1995))

RCM: Swamy Random Coefficient Model (1970)

MG-PC: Se estima el modelo de componentes principales utilizando los primeros 4 componentes principales de los residuos MCO de cada serie individual.

FE-PC: Modelos de Efecto Fijo que incorpora los primeros 4 componentes principales de los residuos MCO de cada serie individual

Como se aprecia en la tabla, los resultados se mantienen en cuanto a la significación de las variables respecto a las anteriores estimaciones. En relación a los valores de la elasticidad ingreso estos son más elevados que el modelo Arellano-Bond estimado, y se asimilan a los modelos Pool MCO estimados.

5.4. Modelos de Corrección del Error

Finalmente se presentan los modelos de corrección del error que consideran la posible cointegración entre la variable dependiente y los regresores, el modelo EC-MG considera heterogeneidad en las elasticidades de largo plazo.

Tabla 14: Resultados de Elasticidades Modelos con Corrección del Error

	EC- PMG	EC-FE	EC-MG
log(IVF)	1.4105 (0.0457)**	1.7008 (0.3131)**	1.3172 (0.2752)**
log(Precio Peaje)	0.2544 (0.2359)	0.2956 (0.4024)	0.2962 (0.5318)
log(Precio Nafta)	-0.0485 (0.1117)	-0.1014 (0.1764)	0.8645 (0.2217)

Nota: Variable dependiente log(Tráfico de camiones). N=21, T=38.

**, ** indican significación al 5% y 1% respectivamente

Se omiten los reportes de la constante, de las dummies estacionales y efecto pascua

EC-PMG: Pool Mean Group (Pesaran Shin y Smith (1999)). Modelo que considera heterogeneidad en el ajuste de corto plazo pero homogeneidad en el coeficiente de largo plazo, en variable cointegradas.

EC-FE: El siguiente modelo realiza una estimación de la especificación del modelo con corrección del error considerando efectos fijos por puesto de conteo

EC-MG: Estos modelos realizan la estimación de corrección del error en cada punto de conteo individual y luego lo promedian (son estimadores que consideran parámetros heterogéneos)

Las estimaciones arrojan nuevamente los mismos resultados en cuanto a la significación de la elasticidad ingreso y su valor. Las variables de elasticidad precio no resultan significativas. Cabe mencionar, que los resultados de estos modelos deben ser tomados con cautela ya que el coeficiente de corrección del error si bien significativo en los tres modelos, tiene un valor positivo indicando la ausencia de ajuste a la relación de largo plazo entre las variables.

6. CONCLUSIONES

Se realizaron estimaciones de la elasticidad ingreso, elasticidad precio peaje, y elasticidad precio de la nafta, a través de diferentes modelos enmarcados en la literatura de paneles largos. Los resultados son concluyentes en cuanto a la significación del IVF como determinante del volumen del tráfico de carga, resultando la estimación puntual entre 0.9 en el modelo en diferencias y 1.7 en el moldeo de corrección del error con efectos fijos. Sin embargo, debido a los diferentes test que se realizaron a las variables, donde se constata dependencia entre los puestos de conteo, heterogeneidad en los parámetros, series integradas de orden 1 y no cointegración; parecería ser recomendable la utilización del modelo MG-PC que arroja una elasticidad puntual de 1.5477, a pesar de no incorporar la dinámica entre las variables.

Respecto a la elasticidad precio del peaje y la elasticidad del precio de la nafta, la mayoría de los modelos estimados rechazan la significación de las variables utilizadas como proxy para medir estas elasticidades. Los modelos FGLS (con y sin dependencias) arrojan resultados significativos de la variable Precio del Peaje, pero el signo no está en línea con la teoría, por lo que se entiende que están recogiendo otro efecto sobre el tráfico.

Una posible extensión al presente estudio es realizar estimaciones considerando modelos de Econometría Espacial, donde se considera la distancia física entre los diferentes cross section como medida para considerar la dependencia entre las unidades.

BIBLIOGRAFÍA

- Aschauer, D. (1989a), Is public expenditure productive? *Journal of Monetary Economics*, 23, 177-200.
- Alvarez, J., and M. Arellano (2003), The Times Series and Cross Section Asymptotics of Dynamic Panel Data Estimators, *Econometrica*, 71, 1121-1159.
- Arellano, M (2003) *Panel Data Econometrics*, Oxford University Press.
- Baltagi, B.H. (2008) *Econometrics Analysis of Panel Data*, 4th edition New York: Wiley.
- Baltagi, B.H. and J. M. Griffin (1984), Short and long-run effects in pooled models, *International Economic Review*, 25, 631-645.
- Baltagi, B.H. and J. M. Griffin (1984), Pooled Estimators versus their Heterogeneous Counterparts in the context of dynamic demand for gasoline, *Journal of Econometrics*, 77, 303-327.
- Baltagi B.H., G. Bresson and A. Pirotte (2003), Fixed effects, random effects or Hausman-Taylor? A pre-test estimator, *Economics Letters*, 79, 361-369.
- Banerjee, A. (1999) Panel Data, Unit Roots and Cointegration: An Overview, *Oxford Bulletin of Economics and Statistics*, Special Issue on Testing for Unit Roots and Cointegration using Panel Data, Theory and Applications, 61, November, 607-629.
- Banerjee, A. and M. Marcellino (2008) Factor-Augmented Error Correction Models, CEPR Discussion Paper Series, No 6707.
- Breitung, J. M. and Meyer (1994), Testing for unit roots using panel data: are wages on different bargaining levels cointegrated, *Applied Economics*, 26, 353-361.
- Breitung, J. and S. Das (2008), Testing for unit roots in panels with factor structure, *Econometric Theory*, 24, 88-106.
- Chudik, A. M.H Pesaran and E. Tosetti (2009), Weak and strong cross-section dependence and estimation of large panels, Cambridge.
- Coakley, J., A-M. Fuertes and R.P. Smith (2002), A Principal Components approach to cross-section dependence in panels, mimeo Birkbeck College.
- Coakley, J., A-M. Fuertes and R.P. Smith (2006), Unobserved heterogeneity in panel time series models, *Computational Statistics and Data Analysis*, 50 (9) 2361-2380.
- De Silva, S., K. Hadri and R. Tremayne (2009) Panel unit root tests in the presence of cross-sectional dependence: finite sample performance and an application, *The Econometrics Journal*, 12, 340-366.
- Driscoll, J.C. and A.C. Kraay (1995) Spatial Correlation in Panel Data, The World Bank Policy Research Department Macroeconomics and Growth Division, December 1995.

- Driscoll, J.C. and A.C. Kraay (1997) Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data, *Review of Economics and Statistics*, forthcoming.
- Eberhardt, M. and F. Teal (2010), Econometrics for Grumblers: A new look at cross-country growth empirics, *Journal of Economic Survey*, forthcoming.
- Engle, R.F. and C.W.J. Granger (1987) Cointegration and Error Correction: representations, estimation and testing, *Econometrica*, 64, 813-816.
- Gourieroux, C., P.C.B. Phillips and J. Yu (2006) Indirect Inference for Dynamic Panel Models, Cowles Foundation Discussion Paper No 1550.
- Granger, C.W.J. and P. Newbold (1974), Spurious Regression in Econometrics, *Journal of Econometrics*, 2, 111-120.
- Gutierrez, Luciano (2003), On the power of panel cointegration tests: a monte carlo comparison. *Economics Letters* 80, 105-111.
- Hadri, K (2000) Testing for stationarity in heterogeneous panel data. *The Econometrics Journal*, 3,2, 148-161.
- Harris D., D.I. Harvey, S.J. Leybourne and N.D. Sakkas (2010), Local Asymptotic Power of the Im-Pesaran-Shin panel unit root test and the impact of initial observations, *Econometric Theory*, 26, 311-324.
- Harris, R.D.F. and E. Tzavalis (1999b), Inference for unit roots in dynamic panels with heteroscedastic and serially correlated errors, mimeo University of Exeter.
- Hausman, J.A. (1978), Specification tests in econometrics, *Econometrica*, 49, 1377-1398.
- Hsiao, C. (1986) *Analysis of Panel Data*, Cambridge: Cambridge University Press.
- Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu (1999), Bayes Estimation of short-run Coefficients in Dynamic Panel Data Models, in *Analysis of Panels and Limited Dependent Variable Models: A Volume in Honour of G. S. Maddala* edited by C. Hsiao, K. Lahiri, L-F. Lee and M.H. Pesaran, Cambridge: Cambridge University Press.
- Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu (2002) Maximum Likelihood estimation of fixed effect dynamic panel data models covering short time-periods, *Journal of Econometrics*, 109, 107-150.
- Im, K.S., M.H. Pesaran and Y. Shin (2003), Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, 115, 53-74.
- Jang, M.J. and D.W. Shin (2005) Comparison of panel unit root tests under cross sectional dependence, *Economics Letters*, 89, 12-17.
- Johansen, S. (1988) Statistical Analysis of Cointegrating Vectors, *Journal of Economic Dynamics and Control*, 12, 231-254.

- Kao, C (1999), Spurious Regression and Residuals based test for cointegration in panel data, *Journal of Econometrics*, 90, 1-44.
- Kapetanios, G (2007), Dynamic Factor Extraction of Cross-sectional Dependence in panel unit root test, *Journal of Applied Econometrics*, 22, 313-338.
- Kapetanios, G., M.H. Pesaran and T. Yamagata (2008) Panels with nonstationary multifactor error structures, mimeo Cambridge.
- Karlsson, S. and M. Lothgren (2000), On the power and interpretation of panel unit root tests, *Economics Letters*, 66, 249-255.
- Kwiatkowski, D., P.C. Phillips, P. Schmidt and Y. Shin (1992), Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, 54, 159-178.
- Lazarova, S. L. Trapani and G. Urgan (2007), Common Stochastic Trends and Aggregation in Heterogeneous Panels, *Econometric Theory*, 23, 89-105.
- Levin, A. and C.-F. Lin and C.-S.J. Chu (2002), Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties, *Journal of Econometrics*, 108, 1-24.
- Maddala, G.S. and I.-M. Kim (1998), *Unit Roots, Cointegration and Structural Change*, Cambridge: Cambridge University Press.
- Maddala, G.S. and S. Wu (1999) A Comparative Study of Unit Root Tests with panel data and a new simple test, *Oxford Bulletin of Economics and Statistics*, Special Issue on Testing for Unit Roots and Cointegration using Panel Data, Theory and Applications, 61, 631-652.
- Moon H.R. and B. Perron (2007), An empirical analysis of nonstationarity in a panel of interest rates with factors, *Journal of Applied Econometrics*, 22, 383-400.
- Pedroni, P (1999) Critical values for cointegration tests in heterogenous panels with multiple regressors, *Oxford Bulletin of Economics and Statistics*, Special Issue on Testing for Unit Roots and Cointegration using Panel Data, Theory and Applications, 61, November, 653-670.
- Pedroni, P. (2004) Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time-series Tests with applications to the PPP hypothesis, *Econometric Theory*, 3, 579-625.
- Pesaran, M.H. (2006) Estimation and Inference in Large Heterogeneous Panels with a multifactor error structure, *Econometrica*, 74(4) 967-1012.
- Pesaran, M.H. (2007) A simple panel unit root test in the presence of cross section dependence, *Journal of Applied Econometrics*, 22(2).p265-312.
- Pesaran, M.H. Y. Shin and R.J. Smith (2001) Bounds Testing approaches to the Analysis of Level Relationships, *Journal of Applied Econometrics*, 16, 289-326.
- Pesaran, M.H., Y. Shin and R.P. Smith (1999) Pooled Mean Group Estimation of Dynamic Heterogenous Panels, *Journal of the American Statistical Association*, 94, 621-634.

Pesaran, M.H. and R.P. Smith (1995), Estimating Long-run relationships from Dynamic Heterogenous Panels, *Journal of Econometrics*, 68, 79-113.

Pesaran, M.H. and R.P. Smith (1998) Structural Analysis of Cointegrating VARs, *Journal of Economic Surveys*, 12, 471-506, reprinted in Oxley and McAleer (1999).

Pesaran, M.H. L.V. Smith and T. Yamagata (2008) Panel Unit Root Tests in the Presence of a Multifactor Error Structure, mimeo Cambridge.

Pesaran, M.H. R.P. Smith and K.S. Im (1996) Dynamic Linear Models for Heterogenous Panels p145-195 of Matyas and Sevestre (1996).

Pesaran, M.H. and E. Tosetti (2009) Large Panels with Common Factors and Spatial Correlations, mimeo Cambridge.

Pesaran M.H., A. Ullah and T. Yamagata (2006) A bias adjusted LM test of error cross-section independence, forthcoming *Econometrics Journal*.

Pesaran M.H. and T. Yamagata (2008) Testing slope homogeneity in large panels, *Journal of Econometrics*, 142, 50-93.

Phillips, P.C.B. (1986) Understanding Spurious Regressions in Economics, *Journal of Econometrics*, 33, 311-340.

Phillips, P.C.B. (2004) Challenges of trending time-series econometrics, Cowles Foundation Discussion Paper 1472.

Phillips, P.C.B. and Z. Xiao (1998) A Primer on Unit Root Testing, *Journal of Economic Surveys*, 12, 423-470, reprinted in Oxley & McAleer (1999).

Phillips, P.C.B., and H.R. Moon (1999) Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica*, 67,5, 1057-1112.

Phillips, P.C.B., and H.R. Moon (2000) Nonstationary Panel Data Analysis: an overview of some recent developments, *Econometric Reviews* 19, 3, 263-286.

Phillips, P.C.B. and D. Sul (2003) Dynamic Panel Testing and Homogeneity Testing under Cross section dependence, *The Econometrics Journal*, 6, 217-259.

Phillips, P.C.B. and D. Sul (2006) Bias in Dynamic Panel Estimation with Fixed Effects incidental trends and cross-section dependence, *Journal of Econometrics*, 137, p162-188.

Quah, D (1994) Exploiting Cross-section Variation for Unit Root Inference in Dynamic Data, *Economics Letters*, 44, 9-19.

Shin, Y. and A. Snell (2006) Mean group tests for stationarity in heterogeneous panels, *Econometrics Journal*, 9, 123-158.

Stock, J.H. and M. W. Watson (2006) Heteroskedasticity Robust Standard Errors for Fixed Effects Panel Data Regression, NBER Working Paper T0323.

Swamy, P.A.V.B. (1970) Efficient Inference in a Random Coefficient Regression Model, *Econometrica*, 38, 311-323.

Westerlund, J. (2007) Testing for error correction in panel data, *Oxford Bulletin of Economics and Statistics*, 69(6) p709-748.

Westerlund, J. (2005) New Simple Tests for Panel Cointegration, *Econometric Review*, 24, p297-316.

Westerlund, J. and S.A. Basher (2008) Mixed Signals among tests for panel cointegration, *Economic Modelling*, 25, p128-136.

Westerlund J. and R. Larsson (2009) A note on the pooling of individual PANIC unit root tests, *Econometric Theory*, 25, 1851-1868.

Wooldridge, J.M. (2002) *Econometric Analysis of Cross-section and Panel Data*, MIT press.

Zellner, A. (1962) An Efficient Method of Estimating Seemingly Unrelated Regressions and Test for Aggregation Bias, *Journal of the American Statistical Association*, 57, 348-368.

Zellner, A (1969) On the aggregation problem; A new approach to a troublesome problem, in K.A. Fox et al. eds *Economic models, estimation and risk programming; Essays in honor of Gerhard Tintner* (Springer-Verlag, Berlin) 365-378.

ANEXO

Pool MCO

Linear regression

Number of obs = 798
 F(7, 20) = 83.30
 Prob > F = 0.0000
 R-squared = 0.1540
 Root MSE = .61985

(Std. Err. adjusted for 21 clusters in id)

lcam	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
livf	1.576759	.300341	5.25	0.000	.9502587	2.203259
lprpeaje	.217245	.6576545	0.33	0.745	-1.154598	1.589088
lprnafta	-.0080099	.1368069	-0.06	0.954	-.293384	.2773643
xtrim1	-.0052325	.0243075	-0.22	0.832	-.0559371	.0454721
xtrim2	.1265306	.0356448	3.55	0.002	.0521768	.2008844
xtrim3	.0779907	.0375878	2.07	0.051	-.000416	.1563974
xpascua	.0045845	.0026495	1.73	0.099	-.0009422	.0101111
_cons	1.277199	2.354261	0.54	0.593	-3.633702	6.188101

Pool FGLS AR(4)

GEE population-averaged model

Group and time vars:

Link: id fecha2
 Family: identity
 Correlation: Gaussian
 AR(4)

Number of obs = 798
 Number of groups = 21
 Obs per group: min = 38
 avg = 38.0
 max = 38
 wald chi2(7) = 580.26
 Prob > chi2 = 0.0000

Scale parameter: .3867063

(Std. Err. adjusted for clustering on id)

lcam	Coef.	Semirobust Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.340482	.1429605	9.38	0.000	1.060284	1.620679
lprpeaje	.2689608	.1083556	2.48	0.013	.0565877	.4813338
lprnafta	.1261326	.1115833	1.13	0.258	-.0925667	.3448318
xtrim1	-.0234664	.0227817	-1.03	0.303	-.0681176	.0211849
xtrim2	.1042244	.0344132	3.03	0.002	.0367757	.1716732
xtrim3	.059799	.0296561	2.02	0.044	.0016741	.1179239
xpascua	.0044375	.0008896	4.99	0.000	.002694	.006181
_cons	1.616725	.7899061	2.05	0.041	.068538	3.164913

Pool FGLS AR(8)

```

GEE population-averaged model
Group and time vars:      id fecha2
Link:                     identity
Family:                   Gaussian
Correlation:              AR(8)
Scale parameter:         .3857258
Number of obs            =      798
Number of groups         =      21
Obs per group: min      =      38
                      avg      =     38.0
                      max      =      38
wald chi2(7)            =     656.45
Prob > chi2              =     0.0000
  
```

(Std. Err. adjusted for clustering on id)

lcam	Coef.	Semirobust Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.338756	.1500887	8.92	0.000	1.044587	1.632924
lprpeaje	.3277927	.1281596	2.56	0.011	.0766045	.5789809
lprnafta	.0985103	.1158423	0.85	0.395	-.1285363	.325557
xtrim1	-.0217081	.0239105	-0.91	0.364	-.0685719	.0251557
xtrim2	.1089312	.0314416	3.46	0.001	.0473068	.1705556
xtrim3	.0589049	.0287367	2.05	0.040	.0025821	.1152278
xpascua	.0026913	.0008763	3.07	0.002	.0009738	.0044087
_cons	1.425458	.837006	1.70	0.089	-.2150435	3.06596

Pool FGLS AR(12)

```

GEE population-averaged model
Group and time vars:      id fecha2
Link:                     identity
Family:                   Gaussian
Correlation:              AR(12)
Scale parameter:         .3860021
Number of obs            =      798
Number of groups         =      21
Obs per group: min      =      38
                      avg      =     38.0
                      max      =      38
wald chi2(7)            =     585.54
Prob > chi2              =     0.0000
  
```

(Std. Err. adjusted for clustering on id)

lcam	Coef.	Semirobust Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.358536	.1516825	8.96	0.000	1.061244	1.655829
lprpeaje	.2563601	.1299984	1.97	0.049	.0015679	.5111524
lprnafta	.0538345	.1110395	0.48	0.628	-.163799	.271468
xtrim1	-.0212956	.0244578	-0.87	0.384	-.069232	.0266408
xtrim2	.1076472	.0307661	3.50	0.000	.0473467	.1679477
xtrim3	.0563557	.0281304	2.00	0.045	.001221	.1114903
xpascua	.0021838	.0007958	2.74	0.006	.0006241	.0037435
_cons	1.851177	.8292884	2.23	0.026	.2258014	3.476552

FE

Fixed-effects (within) regression
 Group variable: id

Number of obs = 798
 Number of groups = 21

R-sq: within = 0.7061
 between = 0.0015
 overall = 0.1532

Obs per group: min = 38
 avg = 38.0
 max = 38

corr(u_i, xb) = -0.0004

F(7,20) = 87.81
 Prob > F = 0.0000

(Std. Err. adjusted for 21 clusters in id)

lcam	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
livf	1.654692	.1893614	8.74	0.000	1.259691	2.049693
lprpeaje	-.0442773	.195049	-0.23	0.823	-.4511423	.3625877
lprnafta	-.0233227	.1232805	-0.19	0.852	-.2804814	.233836
xtrim1	-.0037268	.0232614	-0.16	0.874	-.0522494	.0447957
xtrim2	.1273361	.0344885	3.69	0.001	.0553943	.1992778
xtrim3	.0830905	.030742	2.70	0.014	.0189638	.1472173
xpascua	.0054108	.0014527	3.72	0.001	.0023806	.008441
_cons	2.299268	.767584	3.00	0.007	.6981163	3.90042
sigma_u	.60795665					
sigma_e	.17248516					
rho	.92550343	(fraction of variance due to u_i)				

FE-AR

FE (within) regression with AR(1) disturbances
 Group variable: id

Number of obs = 777
 Number of groups = 21

R-sq: within = 0.4579
 between = 0.0016
 overall = 0.1441

Obs per group: min = 37
 avg = 37.0
 max = 37

corr(u_i, xb) = 0.0010

F(7,749) = 90.37
 Prob > F = 0.0000

lcam	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
livf	1.428374	.1001894	14.26	0.000	1.231689	1.62506
lprpeaje	.2238961	.1696347	1.32	0.187	-.1091199	.5569121
lprnafta	-.068203	.0971779	-0.70	0.483	-.2589764	.1225704
xtrim1	-.0270735	.0129991	-2.08	0.038	-.0525925	-.0015545
xtrim2	.1090062	.0166046	6.56	0.000	.0764092	.1416033
xtrim3	.0765841	.0128753	5.95	0.000	.0513081	.1018602
xpascua	.0042316	.00177	2.39	0.017	.0007567	.0077064
_cons	2.174047	.3205433	6.78	0.000	1.544777	2.803317
rho_ar	.5922941					
sigma_u	.6066483					
sigma_e	.13751164					
rho_fov	.95112973	(fraction of variance because of u_i)				

F test that all u_i=0: F(20,749) = 119.31 Prob > F = 0.0000

RE

```

Random-effects GLS regression           Number of obs   =       798
Group variable: id                     Number of groups =        21

R-sq:  within = 0.7061                 obs per group: min =       38
        between = 0.0015                 avg =             38.0
        overall = 0.1532                 max =             38

Random effects u_i ~ Gaussian          Wald chi2(7)    =       617.09
corr(u_i, X) = 0 (assumed)             Prob > chi2     =        0.0000
theta = .95509922

```

(Std. Err. adjusted for 21 clusters in id)

lcam	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.654345	.1893799	8.74	0.000	1.283167	2.025523
lprpeaje	-.043113	.193253	-0.22	0.823	-.4218819	.3356559
lprnafta	-.0232545	.1233036	-0.19	0.850	-.2649251	.218416
xtrim1	-.0037335	.0232632	-0.16	0.872	-.0493286	.0418615
xtrim2	.1273325	.0344932	3.69	0.000	.0597271	.1949379
xtrim3	.0830678	.030759	2.70	0.007	.0227814	.1433543
xpascua	.0054071	.0014517	3.72	0.000	.0025619	.0082523
_cons	2.294718	.7568363	3.03	0.002	.8113461	3.77809
sigma_u	.62254081					
sigma_e	.17248516					
rho	.92870705	(fraction of variance due to u_i)				

RE-AR(1)

```

RE GLS regression with AR(1) disturbances   Number of obs   =       798
Group variable: id                         Number of groups =        21

R-sq:  within = 0.7038                 obs per group: min =       38
        between = 0.0015                 avg =             38.0
        overall = 0.1538                 max =             38

corr(u_i, Xb) = 0 (assumed)             Wald chi2(8)    =       786.82
                                           Prob > chi2     =        0.0000

```

lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.522079	.1003545	15.17	0.000	1.325388	1.71877
lprpeaje	.2220932	.1717919	1.29	0.196	-.1146128	.5587993
lprnafta	.0636765	.0935579	0.68	0.496	-.1196936	.2470466
xtrim1	-.0153456	.0130619	-1.17	0.240	-.0409464	.0102552
xtrim2	.1139994	.017098	6.67	0.000	.080488	.1475107
xtrim3	.0729061	.0132299	5.51	0.000	.0469759	.0988363
xpascua	.0045389	.0018265	2.49	0.013	.000959	.0081188
_cons	1.266657	.7632267	1.66	0.097	-.2292398	2.762554
rho_ar	.5922941	(estimated autocorrelation coefficient)				
sigma_u	.59165816					
sigma_e	.14134129					
rho_fov	.9460125	(fraction of variance due to u_i)				
theta	.90876837					

SURE

Cross-sectional time-series FGLS regression

Coefficients: generalized least squares
 Panels: heteroskedastic with cross-sectional correlation
 Correlation: no autocorrelation

Estimated covariances	=	231	Number of obs	=	798
Estimated autocorrelations	=	0	Number of groups	=	21
Estimated coefficients	=	8	Time periods	=	38
			wald chi2(7)	=	6715.34
			Prob > chi2	=	0.0000

lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.570088	.0275618	56.97	0.000	1.516068	1.624108
lprpeaje	.2212712	.0298939	7.40	0.000	.1626803	.2798621
lprnafta	.0109915	.0374555	0.29	0.769	-.0624201	.084403
xtrim1	.0007908	.0097982	0.08	0.936	-.0184133	.019995
xtrim2	.1333039	.0124958	10.67	0.000	.1088127	.1577952
xtrim3	.085455	.0092175	9.27	0.000	.067389	.103521
xpascua	.0039308	.0016277	2.41	0.016	.0007406	.007121
_cons	1.219644	.1615615	7.55	0.000	.9029889	1.536298

Pool FGLS Corr AR(1)

Cross-sectional time-series FGLS regression

Coefficients: generalized least squares
 Panels: heteroskedastic with cross-sectional correlation
 Correlation: common AR(1) coefficient for all panels (0.8473)

Estimated covariances	=	231	Number of obs	=	798
Estimated autocorrelations	=	1	Number of groups	=	21
Estimated coefficients	=	8	Time periods	=	38
			wald chi2(7)	=	1093.28
			Prob > chi2	=	0.0000

lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.425676	.1058869	13.46	0.000	1.218142	1.633211
lprpeaje	.468335	.0918155	5.10	0.000	.2883799	.64829
lprnafta	.1540509	.0698138	2.21	0.027	.0172185	.2908834
xtrim1	-.0202421	.0099019	-2.04	0.041	-.0396494	-.0008348
xtrim2	.1074642	.0119096	9.02	0.000	.0841217	.1308066
xtrim3	.0680191	.0099394	6.84	0.000	.0485382	.0875001
xpascua	.0039406	.001061	3.71	0.000	.001861	.0060202
_cons	.1383383	.5798886	0.24	0.811	-.9982226	1.274899

Pool FGLS 21 Corr AR(1)

Cross-sectional time-series FGLS regression

Coefficients: generalized least squares
 Panels: heteroskedastic with cross-sectional correlation
 Correlation: panel-specific AR(1)

Estimated covariances	=	231	Number of obs	=	798
Estimated autocorrelations	=	21	Number of groups	=	21
Estimated coefficients	=	8	Time periods	=	38
			wald chi2(7)	=	1910.19
			Prob > chi2	=	0.0000

lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.384367	.0582019	23.79	0.000	1.270294	1.498441
lprpeaje	.3962617	.090912	4.36	0.000	.2180775	.5744458
lprnafta	.1796747	.0567134	3.17	0.002	.0685185	.290831
xtrim1	-.0261211	.0074188	-3.52	0.000	-.0406616	-.0115805
xtrim2	.1117938	.0096743	11.56	0.000	.0928325	.130755
xtrim3	.0625506	.007574	8.26	0.000	.0477057	.0773954
xpascua	.0023395	.0010259	2.28	0.023	.0003288	.0043502
_cons	.6523524	.4701871	1.39	0.165	-.2691973	1.573902

Pool MCO Corr AR(1)

Prais-winsten regression, correlated panels corrected standard errors (PCSEs)

Group variable: id	Number of obs	=	798		
Time variable: fecha2	Number of groups	=	21		
Panels: correlated (balanced)	Obs per group: min	=	38		
Autocorrelation: common AR(1)	avg	=	38		
	max	=	38		
Estimated covariances	=	231	R-squared	=	0.9082
Estimated autocorrelations	=	1	wald chi2(7)	=	399.98
Estimated coefficients	=	8	Prob > chi2	=	0.0000

lcam	Panel-corrected		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
livf	1.29074	.1743125	7.40	0.000	.9490934	1.632386
lprpeaje	.4897543	.1955195	2.50	0.012	.1065431	.8729655
lprnafta	.142457	.1133135	1.26	0.209	-.0796333	.3645473
xtrim1	-.0325608	.0161188	-2.02	0.043	-.0641531	-.0009685
xtrim2	.0950843	.0194051	4.90	0.000	.057051	.1331177
xtrim3	.0560883	.0162648	3.45	0.001	.0242099	.0879667
xpascua	.0042819	.0017146	2.50	0.013	.0009214	.0076423
_cons	.7160522	1.079788	0.66	0.507	-1.400293	2.832397
rho	.8473095					

Pool MCO 21 Corr AR(1)

Prais-winsten regression, correlated panels corrected standard errors (PCSEs)

Group variable:	id	Number of obs	=	798	
Time variable:	fecha2	Number of groups	=	21	
Panels:	correlated (balanced)	Obs per group: min	=	38	
Autocorrelation:	panel-specific AR(1)	avg	=	38	
		max	=	38	
Estimated covariances	=	231	R-squared	=	0.9939
Estimated autocorrelations	=	21	wald chi2(7)	=	467.95
Estimated coefficients	=	8	Prob > chi2	=	0.0000

lcam	Panel-corrected			z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.					
livf	1.32141	.1355491	9.75	0.000	1.055738	1.587081	
lprpeaje	.4295582	.1722798	2.49	0.013	.091896	.7672204	
lprnafta	.1641417	.1074854	1.53	0.127	-.0465258	.3748092	
xtrim1	-.0318747	.0142423	-2.24	0.025	-.0597892	-.0039602	
xtrim2	.1011781	.0177806	5.69	0.000	.0663289	.1360274	
xtrim3	.0572789	.0143781	3.98	0.000	.0290983	.0854594	
xpascua	.004098	.0017275	2.37	0.018	.0007123	.0074838	
_cons	.8341534	.9566305	0.87	0.383	-1.040808	2.709115	
rhos = .8969016 .8683551 .9484658 .9670515 .9965777015866							

Discroll and Kray (1998)

Regression with Driscoll-Kraay standard errors	Number of obs	=	798
Method: Pooled OLS	Number of groups	=	21
Group variable (i): id	F(7, 20)	=	130.49
maximum lag: 3	Prob > F	=	0.0000
	R-squared	=	0.1540
	Root MSE	=	0.6198

lcam	Drisc/Kraay			t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.					
livf	1.576759	.0854173	18.46	0.000	1.398582	1.754936	
lprpeaje	.217245	.1731584	1.25	0.224	-.143957	.5784471	
lprnafta	-.0080099	.1590292	-0.05	0.960	-.339739	.3237192	
xtrim1	-.0052325	.0212888	-0.25	0.808	-.0496402	.0391752	
xtrim2	.1265306	.0246914	5.12	0.000	.0750252	.178036	
xtrim3	.0779907	.019033	4.10	0.001	.0382885	.1176929	
xpascua	.0045845	.0016853	2.72	0.013	.001069	.0080999	
_cons	1.277199	1.019659	1.25	0.225	-.8497726	3.404172	

Driscoll and Kraay (1998) – FE

Regression with Driscoll-Kraay standard errors Number of obs = 798
 Method: Fixed-effects regression Number of groups = 21
 Group variable (i): id F(7, 20) = 267.25
 maximum lag: 3 Prob > F = 0.0000
 within R-squared = 0.7061

lcam	Coef.	Drisc/Kraay Std. Err.	t	P> t	[95% Conf. Interval]	
livf	1.654692	.0792379	20.88	0.000	1.489405	1.819979
lprpeaje	-.0442773	.1685079	-0.26	0.795	-.3957787	.307224
lprnafta	-.0233227	.1543973	-0.15	0.881	-.3453898	.2987443
xtrim1	-.0037268	.0227364	-0.16	0.871	-.0511541	.0437005
xtrim2	.1273361	.025094	5.07	0.000	.074991	.1796812
xtrim3	.0830905	.0189503	4.38	0.000	.0435609	.1226201
xpascua	.0054108	.0014593	3.71	0.001	.0023667	.0084549
_cons	2.299268	.9714256	2.37	0.028	.2729101	4.325627

ARELLANO-BOND

Arellano-Bond dynamic panel-data estimation
 Group variable: id
 Time variable: fecha2

Number of obs	=	609
Number of groups	=	21
Obs per group:	min =	29
	avg =	29
	max =	29

Number of instruments = 124	wald chi2(15)	= 2702.43
	Prob > chi2	= 0.0000

One-step results

(Std. Err. adjusted for clustering on id)

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
lcam						
L1.	.3604141	.059576	6.05	0.000	.2436473	.477181
L2.	-.1668522	.0552389	-3.02	0.003	-.2751184	-.058586
L3.	-.1540172	.0596549	-2.58	0.010	-.2709387	-.0370956
L4.	.4630449	.0555131	8.34	0.000	.3542412	.5718486
L5.	-.3181803	.0777918	-4.09	0.000	-.4706495	-.1657111
L6.	-.0798165	.058608	-1.36	0.173	-.1946861	.0350532
L7.	.1448769	.1181353	1.23	0.220	-.086664	.3764178
L8.	-.0267833	.0667941	-0.40	0.688	-.1576974	.1041308
livf	1.243949	.2420967	5.14	0.000	.7694484	1.71845
lprpeaje	-.0740541	.2461907	-0.30	0.764	-.556579	.4084707
lprnafta	-.153861	.0994938	-1.55	0.122	-.3488654	.0411433
xtrim1	.0107733	.0281852	0.38	0.702	-.0444686	.0660152
xtrim2	.091259	.020653	4.42	0.000	.0507798	.1317382
xtrim3	.0283474	.0150158	1.89	0.059	-.0010831	.0577779
xpscua	.0052734	.0018952	2.78	0.005	.0015589	.008988
_cons	2.674314	.8971884	2.98	0.003	.9158569	4.432771

Instruments for differenced equation

GMM-type: L(2/5).lcam

Standard: D.livf D.lprpeaje D.lprnafta D.xtrim1 D.xtrim2 D.xtrim3 D.xpscua

Instruments for level equation

Standard: _cons

MG

Pesaran & Smith (1995) Mean Group estimator

All coefficients present represent averages across groups (id)
 Coefficient averages computed as outlier-robust means (using rreg)

```

Mean Group type estimation
Group variable: id

Number of obs   =   798
Number of groups =   21

Obs per group: min =   38
                avg  =  38.0
                max  =   38

wald chi2(7)    =   189.26
Prob > chi2     =   0.0000
    
```

lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.54425	.127118	12.15	0.000	1.295104	1.793397
lprpeaje	-.1563404	.1993509	-0.78	0.433	-.5470611	.2343803
lprnafta	.0753092	.0903069	0.83	0.404	-.101689	.2523074
xtrim1	-.0261669	.0202916	-1.29	0.197	-.0659378	.0136039
xtrim2	.1123644	.0319741	3.51	0.000	.0496963	.1750324
xtrim3	.0630871	.0225125	2.80	0.005	.0189635	.1072107
xpascua	.005522	.0012837	4.30	0.000	.0030059	.0080381
_cons	2.591602	.9228133	2.81	0.005	.7829215	4.400283

Root Mean Squared Error (sigma): 0.1099
 (RMSE uses residuals from group-specific regressions: unaffected by 'robust').

RCM

```

Random-coefficients regression
Group variable: id

Number of obs   =   798
Number of groups =   21

Obs per group: min =   38
                avg  =  38.0
                max  =   38

wald chi2(7)    =   348.50
Prob > chi2     =   0.0000
    
```

lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.662616	.1808621	9.19	0.000	1.308133	2.0171
lprpeaje	-.0933305	.2113672	-0.44	0.659	-.5076026	.3209415
lprnafta	-.0393225	.1289061	-0.31	0.760	-.2919737	.2133288
xtrim1	-.0008961	.0255914	-0.04	0.972	-.0510543	.0492621
xtrim2	.127678	.0373296	3.42	0.001	.0545133	.2008428
xtrim3	.0911456	.0319881	2.85	0.004	.02845	.1538411
xpascua	.0061105	.0021293	2.87	0.004	.0019372	.0102839
_cons	2.568291	.9617113	2.67	0.008	.6833715	4.45321

Test of parameter constancy: $\chi^2(160) = 36820.79$ Prob > $\chi^2 = 0.0000$

MG-PC

Pesaran & Smith (1995) Mean Group estimator

All coefficients present represent averages across groups (id)
 Coefficient averages computed as outlier-robust means (using rreg)

Mean Group type estimation	Number of obs	=	874
Group variable: id	Number of groups	=	23
	Obs per group: min	=	38
	avg	=	38.0
	max	=	38
	wald chi2(11)	=	211.46
	Prob > chi2	=	0.0000

lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
livf	1.547757	.126621	12.22	0.000	1.299585	1.79593
lprpeaje	-.1928769	.1873909	-1.03	0.303	-.5601563	.1744024
lprnafta	.0748918	.0780725	0.96	0.337	-.0781275	.2279112
PC1	.0112175	.0056623	1.98	0.048	.0001196	.0223155
PC2	.0188959	.0050721	3.73	0.000	.0089547	.0288371
PC3	.005367	.0048302	1.11	0.267	-.0041001	.014834
PC4	.003	.00344	0.87	0.383	-.0037422	.0097422
xtrim1	-.0163195	.0214916	-0.76	0.448	-.0584422	.0258032
xtrim2	.1161159	.0323212	3.59	0.000	.0527676	.1794642
xtrim3	.0625161	.0260623	2.40	0.016	.0114349	.1135974
xpascua	.0060327	.0013155	4.59	0.000	.0034545	.008611
_cons	2.633532	.9159329	2.88	0.004	.8383366	4.428727

Root Mean Squared Error (sigma): 0.1190

FE-PC

```

Fixed-effects (within) regression      Number of obs   =   798
Group variable: id                    Number of groups =   21

R-sq:  within = 0.7269                Obs per group:  min =   38
        between = 0.0015                avg   =   38.0
        overall = 0.1575                max   =   38

                                         F(11,20)       =   70.83
corr(u_i, xb) = -0.0006                Prob > F        =   0.0000

```

(Std. Err. adjusted for 21 clusters in id)

lcam	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
livf	1.66095	.1920397	8.65	0.000	1.260362	2.061537
lprpeaje	-.065276	.2002925	-0.33	0.748	-.4830789	.3525269
lprnafta	-.0245522	.1238439	-0.20	0.845	-.2828861	.2337816
PC1	.0110044	.0052015	2.12	0.047	.0001543	.0218544
PC2	.0168587	.0041835	4.03	0.001	.008132	.0255853
PC3	.0015366	.0049488	0.31	0.759	-.0087865	.0118597
PC4	.0009852	.0046382	0.21	0.834	-.0086898	.0106603
xtrim1	-.0036059	.0233206	-0.15	0.879	-.0522518	.04504
xtrim2	.1274007	.0345681	3.69	0.001	.0552929	.1995086
xtrim3	.0835	.0309097	2.70	0.014	.0190234	.1479766
xpascua	.0054771	.0014721	3.72	0.001	.0024063	.0085479
_cons	2.381334	.7693896	3.10	0.006	.7764156	3.986253
sigma_u	.60800759					
sigma_e	.16670234					
rho	.93008236	(fraction of variance due to u_i)				

EC-PMG

Pooled Mean Group Regression
(Estimate results saved as pmg)

Panel Variable (i): id
Time Variable (t): fecha2

Number of obs = 693
Number of groups = 21
Obs per group: min = 33
avg = 33.0
max = 33
Log Likelihood = 952.7203

LD.lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ec						
livf	1.410479	.1214698	11.61	0.000	1.172403	1.648556
lprpeaje	.2543883	.2358995	1.08	0.281	-.2079661	.7167427
lprnafta	-.0485233	.1117171	-0.43	0.664	-.2674848	.1704383
SR						
ec	.1826921	.0418265	4.37	0.000	.1007137	.2646705
lcam						
L2D.	-.3210907	.0506699	-6.34	0.000	-.4204019	-.2217795
L3D.	-.1820135	.0314254	-5.79	0.000	-.2436062	-.1204209
L4D.	-.1299843	.0421958	-3.08	0.002	-.2126864	-.0472821
livf						
LD.	.5274989	.2326131	2.27	0.023	.0715855	.9834122
L2D.	-.2207115	.2893708	-0.76	0.446	-.7878679	.3464448
L3D.	-.2400705	.1934985	-1.24	0.215	-.6193206	.1391796
L4D.	.8777805	.2613642	3.36	0.001	.3655161	1.390045
lprpeaje						
D1.	-.5958611	.183521	-3.25	0.001	-.9555557	-.2361666
lprnafta						
D1.	.0626057	.0760464	0.82	0.410	-.0864425	.211654
xtrim1						
D1.	.0775975	.0282257	2.75	0.006	.0222761	.1329189
xtrim2						
D1.	.0149295	.040351	0.37	0.711	-.0641571	.094016
xtrim3						
D1.	.0949064	.0226337	4.19	0.000	.0505451	.1392676
xpascua						
D1.	-.0040243	.0013657	-2.95	0.003	-.006701	-.0013475
_cons	-.3754236	.0906989	-4.14	0.000	-.5531902	-.1976571

EC-MG

Mean Group Estimation: Error Correction Form
(Estimate results saved as mg)

LD.lcam	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ec						
livf	1.317203	.2751893	4.79	0.000	.777842	1.856564
lprpeaje	.2962012	.5318252	0.56	0.578	-.7461571	1.338559
lprnafta	.0864483	.2216656	0.39	0.697	-.3480083	.5209048
SR						
ec	.2828454	.0546124	5.18	0.000	.1758071	.3898836
lcam						
L2D.	-.4419794	.0520788	-8.49	0.000	-.544052	-.3399067
L3D.	-.2972725	.0346554	-8.58	0.000	-.3651958	-.2293493
L4D.	-.1927793	.0393588	-4.90	0.000	-.2699211	-.1156376
livf						
L4D.	.4338334	.2266848	1.91	0.056	-.0104607	.8781275
lprpeaje						
L2D.	.4893185	.2568672	1.90	0.057	-.0141319	.992769
lprnafta						
LD.	.2147442	.1066838	2.01	0.044	.0056479	.4238405
xtrim1						
D1.	.0861889	.0303206	2.84	0.004	.0267616	.1456161
xtrim2						
D1.	-.023178	.0304508	-0.76	0.447	-.0828604	.0365044
xtrim3						
D1.	.0452426	.0275649	1.64	0.101	-.0087836	.0992688
xpascua						
D1.	-.0038159	.0011427	-3.34	0.001	-.0060557	-.0015762
_cons	.3458404	.7126512	0.49	0.627	-1.05093	1.742611

EC-FE

Dynamic Fixed Effects Regression: Estimated Error Correction Form
(Estimate results saved as DFE)

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ec						
livf	1.7008	.3131463	5.43	0.000	1.087045	2.314556
lprpeaje	.2956194	.4024406	0.73	0.463	-.4931496	1.084388
lprnafta	-.1014329	.17646	-0.57	0.565	-.4472881	.2444223
SR						
ec	.2411886	.0636177	3.79	0.000	.1165003	.365877
lcam						
L2D.	-.4679505	.0596892	-7.84	0.000	-.5849392	-.3509617
L3D.	-.5138879	.0769376	-6.68	0.000	-.6646828	-.363093
L4D.	-.4672155	.0790972	-5.91	0.000	-.6222432	-.3121878
livf						
D1.	.4429107	.2034239	2.18	0.029	.0442072	.8416141
lprpeaje						
D1.	-.5870162	.260281	-2.26	0.024	-1.097158	-.0768747
lprnafta						
D1.	-.0424206	.0568255	-0.75	0.455	-.1537966	.0689553
xtrim1						
D1.	.0628863	.0247374	2.54	0.011	.0144018	.1113707
xtrim2						
D1.	-.0310622	.0182197	-1.70	0.088	-.0667721	.0046478
xtrim3						
D1.	.0368006	.0174944	2.10	0.035	.0025123	.071089
xpascua						
D1.	-.0014336	.0009994	-1.43	0.151	-.0033924	.0005252
_cons	-.1066549	.5110129	-0.21	0.835	-1.108222	.894912