

Modelado Estadístico De La Función De Recaudación Del Impuesto A Los Ingresos Brutos Para Las Provincias Argentinas Entre Los Años 2002-2012*

*Mgter. Laura García**
Esp. Lic. Mario Gutiérrez***
Lic. Fernando Delahaye*****

Resumen

Este artículo tiene como propósito compartir, a modo de cierre de una investigación científica PICT-O (Proyectos de Investigación Científica y Tecnológica Orientados), un modelo que ha podido ser razonablemente contrastado con datos empíricos mediante el empleo de instrumental econométrico, para las provincias argentinas durante el periodo 2002-2012.

La pesquisa ha consistido fundamentalmente en realizar un trabajo estrictamente numérico para encontrar una función apropiada para la recaudación provincial, la que si bien no se ha obtenido para todas las variables estudiadas, se ha logrado satisfactoriamente relacionar algunas de ellas. Junto a los publicaciones parciales de la parte general de la investigación, denominados “Relevamiento de la situación de recaudación en las provincias. Determinantes de la función de Recaudación Impositiva Provincial” y “Relevamiento de la situación tributaria de las unidades subnacionales en Argentina. Determinantes de la función de recaudación impositiva en Chaco, Chubut, Misiones, Salta, San Luis, Santa Cruz, Río Negro”, conforman un relevamiento de la situación de las administraciones tributarias provinciales argentinas mencionadas.

* Trabajo recibido 01/02/2016 // aprobado 04/03/2016

** Magíster en Economía y Administración Estratégica de Negocios, egresada de la Universidad Católica de Cuyo, San Juan – Argentina. E-mail: lauraagarcia@gmail.com

*** Especialista en Docencia Universitaria (UNSJ). Especialista en Abordaje de Problemáticas Sociales (UNLa). E-mail: marioluisgutierrez@gmail.com

**** Licenciado en Comercialización (UCCu). E-mail: fernando.delahaye@hotmail.com

Palabras-clave: Recaudación de los Impuestos a los Ingresos Brutos; Base Imponible; Alícuota; Existencia de Soporte Tecnológico

Clasificación JEL: C2

Abstract

This article aims to share, as a closure of a scientific research PICT-O (Projects of Scientific Research and Technology Oriented), a model that has been reasonably contrasted with empirical data by using econometric instrumental to the Argentine provinces during the period 2002-2012. The research consisted primarily of making a strictly numerical work to find an appropriate role for the provincial revenue, which although has not been obtained for all variables studied, it has successfully managed to relate some of them. Along with the partial publication of the general part of the investigation, called "Survey of the situation in the provinces collection. Determinants of the role of Provincial Tax Collection "and" Survey of the tax situation of subnational units in Argentina. Determinants of tax collection function in Chaco, Chubut, Misiones, Salta, San Luis, Santa Cruz, Rio Black "make a survey of the situation of Argentine provincial tax administrations mentioned.

Introducción

Este artículo tiene como propósito compartir, a modo de cierre de una investigación científica PICT-O (Proyectos de Investigación Científica y Tecnológica Orientados), un modelo que ha podido ser razonablemente contrastado con datos empíricos mediante el empleo de instrumental econométrico, para las provincias argentinas durante el periodo 2002-2012.

La pesquisa ha consistido fundamentalmente en realizar un trabajo estrictamente numérico para encontrar una función apropiada para la recaudación provincial, la que si bien no se ha obtenido para todas las variables estudiadas, se ha logrado satisfactoriamente relacionar algunas de ellas. Junto a los publicaciones parciales de la parte general de la investigación, denominados "Relevamiento de la situación de recaudación en las

provincias. Determinantes de la función de Recaudación Impositiva Provincial”²⁸ y “Relevamiento de la situación tributaria de las unidades subnacionales en Argentina. Determinantes de la función de recaudación impositiva en Chaco, Chubut, Misiones, Salta, San Luis, Santa Cruz, Río Negro”²⁹, conforman un relevamiento de la situación de las administraciones tributarias provinciales argentinas mencionadas.

Se considera que estas contribuciones validan la hipótesis que sustenta al PICT-O aludido, construyendo un análisis estadístico multivariable que expresa la función de recaudación provincial como la suma de las recaudaciones de cada impuesto que la unidad tributaria administra: $R_p = f(R_i, R_a, R_{ibl}, R_{ibcm}, R_s, R_o)$. Donde, R_p : Recaudación Provincial, como variable dependiente, y como primera muestra de posibles variables independientes significativas y mensurables: R_i : Recaudación del impuesto inmobiliario; R_a : Recaudación del impuesto automotor; R_{ibl} : Recaudación del impuesto a los ingresos brutos local; R_{ibcm} : Recaudación del impuesto a los ingresos brutos convenio multilateral; R_s : Recaudación del impuesto a los sellos; R_o : Recaudación de otros impuestos. Del mismo modo, las recién mencionadas variables independientes, pueden ser consideradas como funciones de Recaudación por cada impuesto (R_t), que estarán, a su turno, determinadas por variables independientes internas y externas: $R_t = f(I, E)$. Donde las variables internas están dadas por: $I = f(N, G, y Otras)$, donde: N : Normativa: Base imponible, alícuota, gasto tributario (descuentos, exenciones, facilidades) (B, A, Ct.), entre otras; G : Gestión: Que incluye: existencia de soporte auxiliar tecnológico y

²⁸ ROSA DONATI, S; GARCÍA, L; Gutiérrez, M (2014) “Relevamiento de la situación de recaudación en las provincias. Determinantes de la función de Recaudación Impositiva Provincial”. Ministério da Educação do Brasil. Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM. Minas Gerais – Brasil. Revista Vozes dos Vales da UFVJM. Reg.: 120.2.095-2011 – PROEXC/UFVJM. ISSN: 2238-6424. Qualis – Capes / Latindex. Nº. 06. Año II. 2014, en <http://www.ufvjm.edu.br/voze>.

²⁹ GARCÍA, L; GUTIÉRREZ, M (2014) “Relevamiento de la situación tributaria de las unidades subnacionales en Argentina. Determinantes de la función de recaudación impositiva en Chaco, Chubut, Misiones, Salta, San Luis, Santa Cruz, Río Negro”. Universidad Central de Chile. Facultad de Ciencias Políticas y Administración Pública. Revista Enfoques. Volumen XIII. Nº 22. 2015.

de gestión (S), porcentaje de cobranzas, cantidad de intimaciones, Recursos Humanos (RH), atención al contribuyente (AC), realización de marketing tributario, acciones directas de cobranza, cantidad de ejecuciones fiscales, magnitud del fondo estímulo al personal, capacitación al personal, entre otras, etc.; *Otras*: Cultura tributaria y otras variables que si bien son influyentes no serán objeto de investigación en el presente trabajo.

Las variables externas consideradas fueron: el PBI nacional, PBG de cada provincia, población económicamente activa provincial, inflación, presión tributaria y otras, es decir: $E = f(PBI, PBG, NA, \pi, Pr \text{ y otras})$. Si las variables externas: PBI, PBG, población económicamente activa (NA), inflación (π) y las variables internas: la alícuota general de cada impuesto (A), y la existencia de soporte tecnológico y de gestión auxiliar (S), son representativas del resto de las variables y, por ello, consideradas eventualmente como variables Proxy de las restantes, entonces el trabajo de investigación consiste en encontrar el grado de influencia de cada una de las citadas variables en el comportamiento de la función de recaudación provincial; en ésta primera instancia para el Impuesto sobre los Ingresos Brutos, es decir que se concentró en verificar la siguiente hipótesis:

$$H_0: R_t: f(PBI, PBG, NA, \pi,$$

Los objetivos oportunamente planteados son:

General:

- Relevar la situación de Recaudación en las Provincias Argentinas, explicando la influencia de las principales variables determinantes, internas y externas de la función recaudación impositiva por provincia.

Particulares:

- Explicar cómo influyen en la Recaudación provincial las siguientes variables: Producto Bruto Geográfico, Población económicamente activa provincial, PBI nacional, las alícuotas generales de cada impuesto, y la existencia o no de soporte tecnológico y de gestión auxiliar.
- Conocidas las relaciones existentes entre las variables independientes y la variable dependiente, establecer los

principales indicadores estadísticos, estimadores de la performance estándar de la administración tributaria, entre ellos la media estadística recaudatoria del país,

- Cuantificar los desvíos para cada jurisdicción. Una vez determinados dichos desvíos, se realizará un análisis correlacional entre las variables bajo análisis a fin de detectar si las mismas pueden considerarse asociadas a los diferenciales de performance recaudatoria y, eventualmente, proponer la corrección de las mismas.

Correlaciones entre variables (2002 al 2012)

En lo que sigue, como modo de abordar los objetivos ya mencionados, se considera el *impuesto a los ingresos brutos*, considerando este gravamen como representativo del total de los impuestos provinciales, (variable y , en pesos) para grupos de provincias argentinas entre los años 2002 y 2012 en relación con las variables x_1 : *PBI* (en pesos), x_3 : *Población* (número de habitantes), x_4 : *Índice de inflación*, x_6 : *Asistencia técnica de terceros* (presente o no presente). Se informa cuáles son las variables independientes que aportan significativamente al *impuesto a los ingresos brutos* y se ofrecen modelos estadísticos que las vincula. El diagrama de puntos de la Figura 1 representa los datos. Se presentan pocos casos con valores muy altos del impuesto a los ingresos brutos, y otros con valores intermedios y bajos, mientras que los valores más recientes están más dispersos.

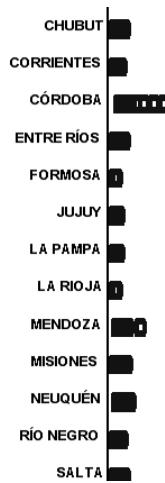


Figura 1

Una estadística descriptiva realizada por provincia, señala que los datos presentan asimetría en correspondencia con el aumento del índice de inflación. No obstante los coeficientes de asimetría y curtosis desbordan el intervalo $[-2, 2]$ lo que indica que los datos no se desvían excesivamente de la normalidad. Los datos no son lo bastante uniformes como para tratarlos en conjunto. Para su análisis se agrupan las provincias con similar recaudación promedio y dispersión.

Un agrupamiento inadecuado distorsiona los modelos estadísticos obtenidos, produciendo residuos muy alejados de la normalidad, lo que invalidaría las pruebas estadísticas realizadas, mientras que una partición excesiva reduce la cantidad de datos disponibles limitando el análisis. De acuerdo con dichas pautas se formaron los siete grupos de provincias mostrados en la Tabla 1. En la misma tabla se presenta el resumen de las correlaciones encontradas en cada grupo entre las variables involucradas.

Las correlaciones que consideran x_3 se calcularon por separado para cada provincia y las que incluyen x_6 , dentro de los grupos y en forma individual. Se considera que una correlación es significativa cuando el *valor P* de una prueba adecuada está por debajo de 0.05. En la Tabla 1 el signo positivo indica correlación

lineal positiva. La estadística descriptiva con base en la cual se formaron los grupos se presenta más adelante en el tratamiento pormenorizado de cada grupo así como también el análisis multidimensional.

Tabla 1. Correlaciones entre variables (2002 al 2012).

Grupo	Provincias	Variables significativas para Y				Correlación entre pares de variables independientes		
		x ₁	x ₃	x ₄	x ₆	(x ₁ , x ₃)	(x ₁ , x ₄)	(x ₁ , x ₆)
1	Buenos Aires		+			+		
	Ciudad Autónoma de Buenos Aires	+	+	+	+	+	+	+
2	Córdoba	+	+	+		+		
	Santa Fe		+		+	+	+	+
3	Mendoza		+			+		
	Neuquén	+	+	+		+	+	
	Tucumán		+		+	+		+
4	Chubut		+			+		
	Entre Ríos		+			+		
	Misiones		+			+		
	Río Negro	+	+	+	+	+	+	+
	Salta		+			+		
Santa Cruz		+			+			
5	Chaco		+			+		
	Corrientes		+			+		
	La Pampa	+	+	+	+	+	+	+
	San Juan		+			+		
	San Luis		+			+		
Tierra del Fuego		+			+			
6	Catamarca		+			+		
	Jujuy		+			+		
	Santiago del Estero	+	+	+	+	+	+	+
7	Formosa	+	+	+	+	+	+	+
	La Rioja		+			+		

El análisis multivariable arrojó correlaciones significativas de la variable dependiente Y con las variables independientes x_1, x_3, x_4, x_6 , entre las variables de los pares (x_1, x_3) , (x_1, x_4) , y en algunos casos (x_1, x_6) . La variable x_6 está presente o no lo está.

Puede ocurrir que la razón por la cual esta variable no resulte significativa para la regresión sea que esté presente o que no lo esté en la mayoría de los años considerados, con poca variación, recomendándose cautela en la interpretación de los resultados obtenidos. Por esta razón se describe la situación de la misma en las provincias.

✓ Aunque no es significativa en el modelo ajustado, considerada sola, la variable x_6 presenta correlación lineal positiva dentro del grupo 1. Está presente en la provincia de Buenos Aires entre los años 2003-2012 y en Ciudad Autónoma de Buenos Aires entre los años 2010 y 2012 presentando correlación lineal positiva.

✓ En las provincias de Córdoba y Santa Fé x_6 está presente entre los años 2005-2012 y entre los años 2010-2012 respectivamente. Presenta correlación lineal positiva en Santa Fé y en el grupo 2 integrado por estas provincias, aunque no es significativa en el modelo ajustado.

✓ En las provincias de Mendoza y Neuquén x_6 está presente entre los años 2003-2012 y entre los años 2002-2012 respectivamente y no tiene correlación significativa con la variable dependiente. Mientras que en la provincia de Tucumán x_6 se presenta entre los años 2006-2012 y se correlaciona positivamente con la recaudación. Dentro del grupo 3, esta variable presenta correlación lineal positiva y aporta significativamente a la variabilidad observada en y en el modelo ajustado.

✓ En las provincias de Chubut, Entre Ríos y Salta x_6 está presente entre los años 2002-2012, mientras que en Santa Cruz no lo está en el mismo período. En Misiones x_6 está presente entre los años 2004-2012 y no se correlaciona con y . En Río Negro x_6 está presente entre los años 2006-2012 y tiene correlación lineal positiva con y . No figura x_6 en el modelo ajustado correspondiente a las provincias que integran el grupo 4.

✓ Entre las provincias que componen el grupo 5 se encuentra que la variable x_6 está presente durante todo el período 2002-2012 en Chaco, Corrientes, San Juan y San Luis. En la provincia de la Pampa x_6 está presente entre los años 2006-2012 y se correlaciona positivamente con y . En Tierra del Fuego x_6 está presente entre los años 2003-2012 y no se correlaciona con la variable dependiente. Dentro del grupo la variable x_6 es significativa pero no figura en el modelo ajustado.

✓ En Catamarca y Jujuy la variable x_6 no está presente entre

los años 2002-2012, mientras que en Santiago del Estero x_6 está presente entre los años 2007-2012 presentando correlación lineal positiva, así como también dentro del grupo 5. Sin embargo no figura en el modelo ajustado.

✓ En las provincias de Catamarca y Jujuy la variable x_6 no está presente entre los años 2002-2012. En la provincia de Santiago del Estero x_6 se presenta entre los años 2007-2012 y se correlaciona positivamente con y . Lo mismo ocurre cuando las provincias se reúnen en el grupo 6, aunque no figura en el modelo ajustado.

✓ En la provincia de La Rioja la variable x_6 está presente entre los años 2002-2012. En la provincia de Formosa x_6 está presente entre los años 2007-2012 presentando correlación lineal positiva con y . Reunidas estas provincias en el grupo 7, x_6 presenta correlación lineal positiva con y aunque no se incluye en el modelo ajustado. Un análisis preliminar apoyado en la matriz de dispersión de la Figura 2, revela que si bien las correlaciones encontradas tienen un efecto lineal positivo, es más apropiado modelar la relación entre las variables en juego, mediante una función no lineal. En la diagonal de la matriz de dispersión, la variable que figura en ella es variable dependiente en la fila y se representa en el eje de ordenadas, y variable independiente en la columna y se representa en abscisas.

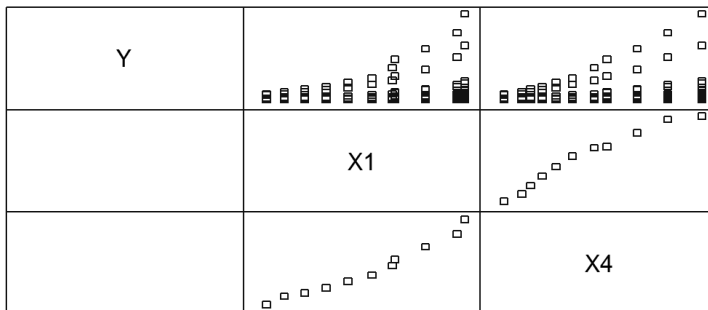


Figura 2

Basados en el análisis precedente se decide, para cada grupo de provincias, ajustar a los datos el modelo no lineal que figura a continuación, el cual se linealiza tomando logaritmo

neperiano³⁰.

$$Y = \beta_0 \cdot x_1^{\beta_1} \cdot x_3^{\beta_3} \cdot x_4^{\beta_4} \cdot x_6^{\beta_6}$$

$$\ln Y = \ln \beta_0 + \beta_1 \ln x_1 + \beta_3 \ln x_3 + \beta_4 \ln x_4 + \beta_6 \ln x_6 + \varepsilon$$

A partir del modelo linealizado se obtienen las estimaciones de los parámetros desconocidos $\beta_0, \beta_1, \beta_3, \beta_4, \beta_6$. En cada modelo estimado se retienen sólo los términos que explican significativamente la variabilidad observada en $\ln Y$, razón por la cual no todos los coeficientes figuran en el modelo ajustado. La variable x_6 está presente o no lo está. En consecuencia, para la regresión se le asignan dos valores con este significado. Acompañando al modelo propuesto se establece el valor 1 cuando no existe asistencia técnica y el número “e” en caso contrario. Más adelante se presenta el tratamiento pormenorizado para cada grupo de provincias por medio del cual se obtuvieron los modelos ajustados publicados en la Tabla 2. En la Tabla 2 se muestran los modelos obtenidos y el porcentaje de la variabilidad observada en el $\ln Y$ explicado por la variable independiente más importante. El coeficiente R^2_{Adj} (ajustado por el número de grados de libertad) expresado en porcentaje, indica el tanto por ciento de la variabilidad observada en $\ln Y$ explicada por el modelo.

Tabla 2. Modelos propuestos para representar los datos entre los años 2002 y 2012.

2002-2012				
Gr.	Provincia	Modelo linealizado	Modelo no linealizado	R^2_{aj}
1	Buenos Aires Ciudad de Buenos Aires	$\ln \hat{y}$ $= 4.51614 + 0.250582$ $\cdot \ln x_3 + 2.47896 \cdot \ln x_4$ $\ln x_4$ explica el 93.41% de la variabilidad observada en $\ln Y$	\hat{y} $= 91.481796$ $\cdot x_3^{0.250582}$ $\cdot x_4^{2.47896}$	99.35%
2	Córdoba Santa Fé	$\ln \hat{y}$ $= 6.55608$ $+ 2.56456 \ln x_4$	\hat{y} $= 703.50853$	99.52%

³⁰ GARCÍA, R. (1992) “Inferencia Estadística y Diseño de Experimentos”. Editorial Eudeba. Pg. 257.

		$\ln x_4$ explica el 99.52% de la variabilidad observada en $\ln Y$	$\cdot x_4^{2.56456}$	
3	Mendoza Neuquén Tucumán	$\ln \hat{y}$ = 4.8218 + 0.1190 $\cdot \ln x_3 + 2.3565 \cdot \ln x_4$ + 0.3118 $\cdot \ln x_6$	\hat{y} = 124.1884 $\cdot x_3^{0.1190}$ $\cdot x_4^{2.3565}$ $\cdot x_6^{0.3118}$	97%
		$\ln x_4$ explica el 95.68% de la variabilidad observada en $\ln Y$		
4	Chubut Entre Ríos Misiones Río Negro Salta Santa Cruz	$\ln \hat{y}$ = -52.2797 + 2.4265 $\cdot \ln x_1 + 0.1892 \cdot \ln x_3$ + 0.8667 $\cdot \ln x_4$	\hat{y} = 1.9734 $\cdot E^{-23}$ $\cdot x_1^{2.4265} x_3^{0.1892}$ $\cdot x_4^{0.8667}$	97.35%
		$\ln x_1$ explica el 95.37% de la variabilidad observada en $\ln Y$		
5	Chaco Corrientes La Pampa San Juan San Luis Tierra del Fuego	$\ln \hat{y}$ = 3.6429 + 0.08738 $\cdot \ln x_3 + 2.5024 \cdot \ln x_4$	\hat{y} = 38.2025 $\cdot x_3^{0.08738}$ $\cdot x_4^{2.5024}$	95.41%
		$\ln x_4$ explica el 94.88% de la variabilidad observada en $\ln Y$		
6	Catamarca Jujuy Santiago del Estero	$\ln \hat{y}$ = 0.32379 $\cdot \ln x_3$ + 2.51037 $\cdot \ln x_4$	\hat{y} = $x_3^{0.32379}$ $\cdot x_4^{2.51037}$	99.99%
		$\ln x_4$ explica el 99.98% de la variabilidad observada en $\ln Y$		
7	Formosa La Rioja	$\ln \hat{y}$ = -93.313 + 3.9886 $\cdot \ln x_1 + 0.4121 \cdot \ln x_3$	\hat{y} = 2.98 $\cdot 10^{-41}$ $\cdot x_1^{3.9886} \cdot x_3^{0.4121}$	98.23%
		$\ln x_1$ explica el 97.11% de la variabilidad observada en $\ln Y$		

Variable	Significado de la variable	Variable	Significado de la variable
y	impuesto a los ingresos brutos	x_4	Índice de inflación

x_1	PBI (en pesos)	x_6	Asistencia técnica de terceros (1 o 0)
x_3	Población (número de habitantes)		

Las variables x_3 , x_4 , figuran en el modelo ajustado a los datos correspondientes a los grupos 1, 5 y 6, abarcando las 10 provincias argentinas nombradas a continuación más Ciudad Autónoma de Buenos Aires: Buenos Aires, Chaco, Corrientes, La Pampa, San Juan, San Luis, Tierra del Fuego, Catamarca, Jujuy, Santiago del Estero. En estos modelos el coeficiente estimado para $\ln x_4$ es próximo a 2.5, explicando esta variable al menos el 93% de la variabilidad observada en la recaudación. En las provincias de Córdoba y Santa Fe que componen el grupo 2, la recaudación es función únicamente del $\ln x_4$ el que explica el 99.52% de la variabilidad observada en la recaudación. En las provincias de Mendoza, Neuquén y Tucumán que integran el Grupo 3, además de las variables x_3 y x_4 , resultó significativa para la regresión la variable x_6 .

No obstante, el 95.68% de la variabilidad observada en la recaudación se explica por la variable $\ln x_4$. En el Grupo 7 integrado por las provincias de Formosa y La Rioja un modelo para la recaudación es función de las variables x_1 y x_3 . El $\ln x_1$ explica el 97.11% de la variabilidad observada en la recaudación. Para este grupo $\ln x_4$ no aportó significativamente a la variabilidad observada en $\ln Y$, no habiendo indicios de colinealidad para este modelo.

El modelo ajustado a los datos correspondientes al grupo 4 merece una aclaración. En atención al análisis detallado de residuos, conserva las variables x_1 , x_3 , x_4 , a pesar de que se correlacionan, debilitando la estimación del efecto individual de cada una de ellas sobre la variabilidad observada en la recaudación, aun cuando representa apropiadamente los datos.

Análisis pormenorizado de cada grupo de provincias (2002–2012)

En lo que sigue, para cada grupo de provincias, se presenta en primer lugar una estadística descriptiva resumida de la

recaudación correspondiente al impuesto a los ingresos brutos. Cada estadística descriptiva muestra que las provincias agrupadas presentan similar coeficiente de variación y el orden de los valores no difiere demasiado. En todos los casos los coeficientes estandarizados de asimetría y curtosis caen dentro del intervalo $[-2, 2]$ señalando que los datos no se apartan excesivamente de la normalidad. Sin embargo el diagrama de dispersión que acompaña cada estadística descriptiva revela que el impuesto a los ingresos brutos presenta asimetría en correspondencia con el aumento del índice de inflación, la que se amortigua más tarde, al tomar logaritmo neperiano para ajustar a los datos el modelo propuesto.

A continuación, mediante un análisis multidimensional se detectan correlaciones entre pares de variables. Las correlaciones con la variable x_3 se calcularon por separado para cada provincia del grupo. Cada celda de las tablas presentadas, está dividida en dos partes. La parte superior contiene el coeficiente de correlación calculado y en la parte inferior figura el *valor P* de una prueba para decidir si la relación encontrada es estadísticamente significativa. El coeficiente de correlación varía entre -1 y 1 , y mide la fuerza de la relación lineal entre las variables, mientras que un *Valor P* por debajo de 0.05 indica correlación significativa. En la diagonal de la matriz de dispersión, que sigue a la tabla de correlaciones, la variable que figura en ella es variable dependiente en la fila y se representa en el eje de ordenadas, y variable independiente en la columna y se representa en abscisas.

La matriz de dispersión descubre que, si bien las correlaciones encontradas tienen un efecto lineal positivo, es más apropiado modelar la relación entre las variables en juego, mediante una función no lineal. Seguidamente se realiza el análisis de la regresión múltiple. Se ajusta a los datos el modelo el modelo no lineal que figura a continuación, el cual se linealiza tomando logaritmo neperiano.

$$Y = \beta_0 \cdot x_1^{\beta_1} \cdot x_3^{\beta_3} \cdot x_4^{\beta_4} \cdot x_6^{\beta_6}$$

$$\ln Y = \ln \beta_0 + \beta_1 \ln x_1 + \beta_3 \ln x_3 + \beta_4 \ln x_4 + \beta_6 \ln x_6 + \varepsilon$$

Se realizan pruebas de significación y se informa el coeficiente de determinación R^2_{Adj} . Una regresión por pasos descubre el aporte en la variabilidad observada en la recaudación de la variable independiente más importante. Se representa gráficamente el modelo ajustado y el efecto de cada componente. Al final de cada

análisis se incluye adicionalmente, un gráfico de residuos. Se muestra la concordancia entre el modelo realmente ajustado no linealizado y los datos, mediante una prueba para la pendiente igual a uno de la recta de regresión que contiene al origen.

GRUPO I:

BUENOS AIRES Y CIUDAD DE BUENOS AIRES (2002–2012)

1. *Estadística Descriptiva*

Tabla 3. Estadística descriptiva.

Descripción	Comunidad	
	Buenos Aires	Ciudad Autónoma de Buenos Aires
Cantidad de datos	11	11
Promedio	1.24295E10	7.90202E9
Varianza	1.05531E20	3.95736E19
Mínimo	2.08522E9	1.71659E9
Máximo	3.3357E10	2.12523E10
Coefic. estand. de asimetría	1.38	1.60
Coefic. estand. de curtosis	0.06	0.43
Coeficiente de variación	82.65%	84.89%

2. *Diagrama de Dispersión*

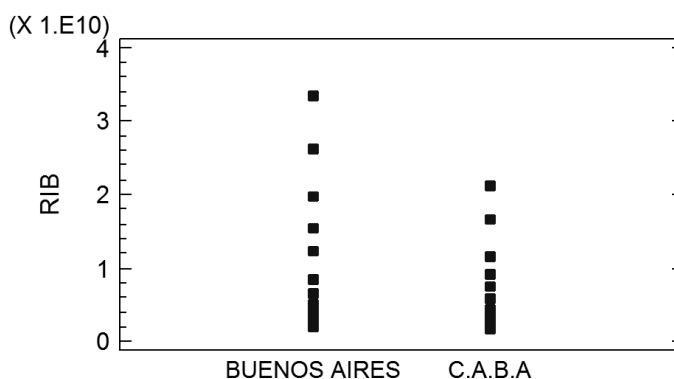


Figura 3

3. *Correlaciones*

Tabla 4. Correlaciones

Grupo I	X ₁	X ₄	X ₆	Buenos Aires	X ₁	X ₃	C. A. B. A.	X ₁	X ₃
Y	0.8773	0.9246	0.5680	Y	0.9410	0.9810	Y	0.9298	0.8815
	0.0000	0.0000	0.0058		0.0000	0.0000		0.0000	0.0000
X ₁		0.9783	0.4985	X ₁		0.9745	X ₁		0.9776
		0.0000	0.0182			0.0000			0.0000

4. Matriz de Dispersión

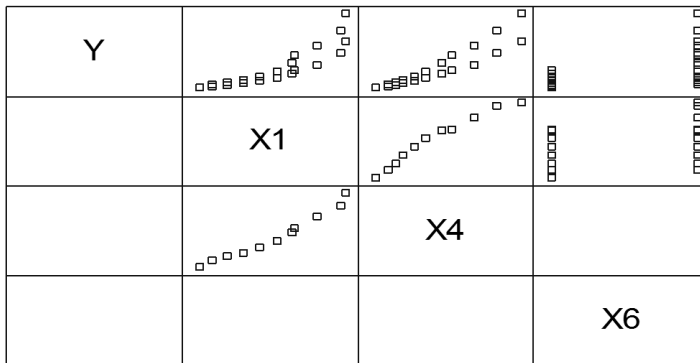


Figura 4.

5. Análisis de la Regresión Múltiple

La tabla siguiente muestra las estimaciones obtenidas, las estadísticas *t* y los valores P para pruebas acerca de la significación

de los coeficientes de regresión. Todos los valores P son menores que 0.05 señalando regresión significativa.

Tabla 5. Análisis de la Regresión Múltiple

Variable Independ.	Variable dependiente $\ln Y$				
	Parámetro	Estimación	Error estándar	Estadística t	Valor P
	β_0	4.51614	0.38507	11.728	0.0000
$\ln x_3$	β_3	0.250582	0.01846	13.5773	0.0000
$\ln x_4$	β_4	2.47896	0.04530	54.7218	0.0000

Las siguientes son las ecuaciones del modelo linealizado ajustado a los datos y el modelo original estimado respectivamente.

$$\ln \hat{y} = 4.51614 + 0.250582 \cdot \ln x_3 + 2.47896 \cdot \ln x_4$$

$$\hat{y} = 91.481796 \cdot x_3^{0.250582} \cdot x_4^{2.47896}$$

El coeficiente de determinación ajustado por el número de grados de libertad R^2_{Adj} indica que el modelo explica el 99.35% de la variabilidad observada en $\ln Y$. Una regresión por pasos hacia adelante indica que $\ln x_4$ explica el 93.72% de la variabilidad observada en $\ln Y$ y cuando se agrega $\ln x_3$ este porcentaje sube al 99.35%. La figura 5 representa gráficamente la concordancia entre los datos, o valores observados, y los valores calculados a partir del modelo linealizado estimado, o valores predichos.

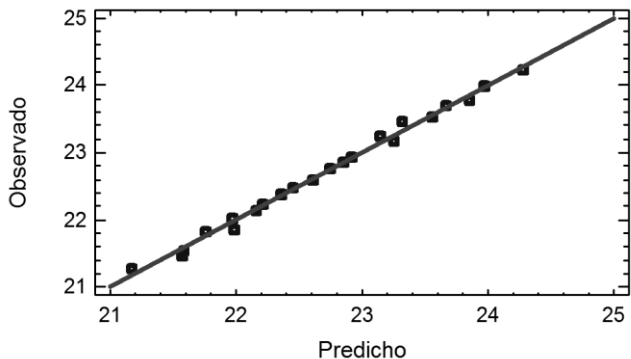


Figura 5

Las figuras 6 y 7, muestran la parte del modelo ajustado que vincula $\ln Y$ con $\ln x_3$ y con $\ln x_4$ respectivamente, ratificando que la variable $\ln x_4$ es la más importante.

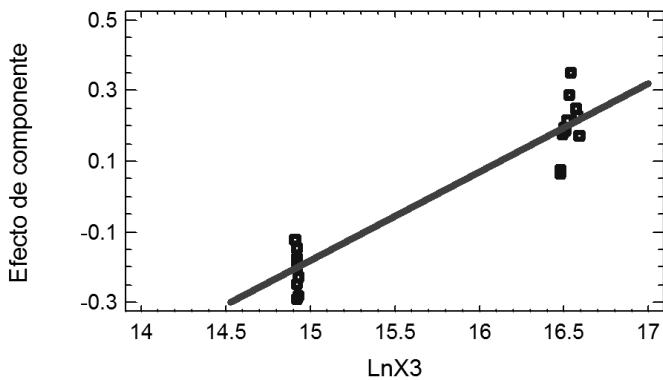
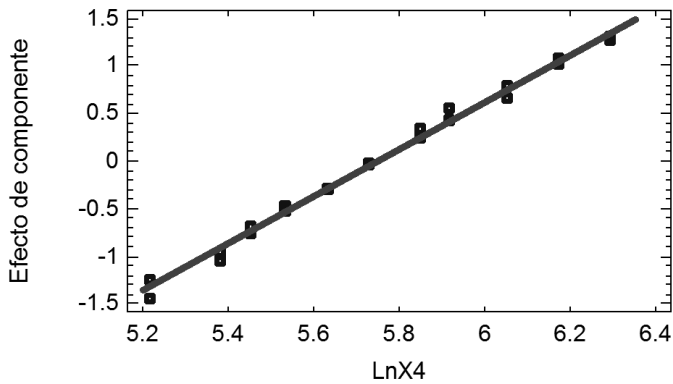


Figura 6

Figura 7



El gráfico de probabilidad normal muestra que los residuos caen aproximadamente sobre la línea recta del diagrama, señalando normalidad.

Grupo 1 - Gráfico de Probabilidad Normal

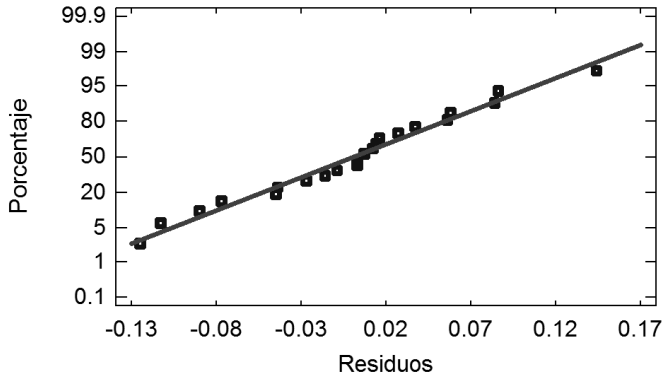
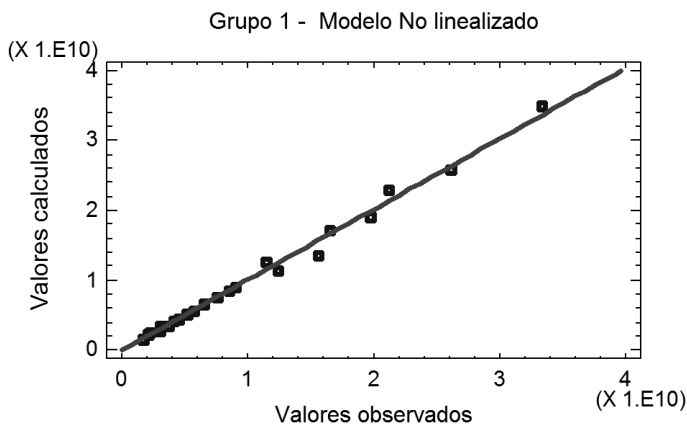


Figura 8

Análisis del Modelo no Linealizado

La concordancia entre los valores observados y los calculados a partir del modelo no linealizado se representa en la figura 9 y se sustenta en una prueba de hipótesis para la pendiente de la recta de regresión que contiene al origen ajustada a dichos valores, la que conduce a no rechazar la hipótesis de que la

pendiente es igual a uno (Valor P=0.48).



GRUPO II:

CÓRDOBA Y SANTA FÉ (2002–2012)

1. Estadística Descriptiva

Tabla 6 .Estadística descriptiva

Descripción	Provincias	
	Córdoba	Santa Fé
Cantidad de datos	11	11
Promedio	2.55572E9	2.38497E9
Varianza	4.74837E18	3.8124E18
Mínimo	4.30044E8	4.60142E8
Máximo	6.90917E9	6.4433E9
Coefic. estand. de asimetría	1.4268	1.55349
Coefic. estand. de curtosis	0.00287645	0.258088
Coeficiente de variación	85.2628%	81.8686%

2. Diagrama de Dispersión

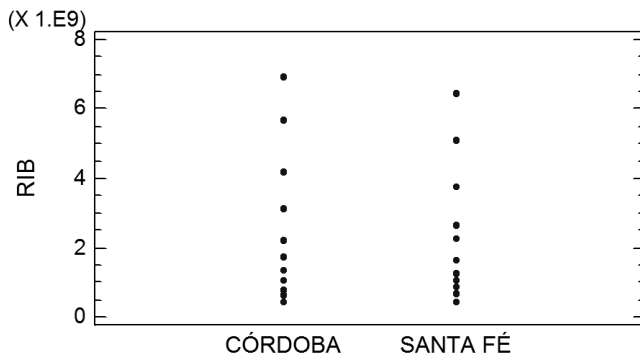


Figura 10

3. Correlaciones

Tabla 7. Correlaciones

Grupo II	X ₁	X ₃	X ₄	X ₆	Córdoba	X ₁	X ₃	Santa Fé	X ₁	X ₃
Y	0.9308	0.7997	0.9828	0.6631	Y	0.9356	0.9335	Y	0.9303	0.9165
	0.0000	0.0000	0.0000	0.0008		0.0000	0.0000		0.0000	0.0000
X ₁		0.8257	0.9783	0.6958	X ₁		0.9924	X ₁		0.9926
		0.0000	0.0000	0.0003			0.0000			0.0000

4. Matriz de Dispersión

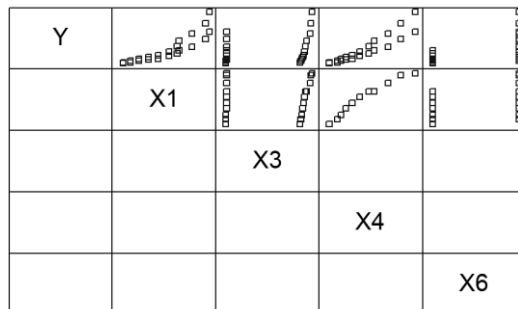


Figura 11

5. Análisis de Regresión Múltiple

La tabla siguiente muestra las estimaciones obtenidas, las estadísticas *t* y los valores P para pruebas acerca de la significación de los coeficientes de regresión. Todos los valores P son menores que 0.05 señalando regresión significativa.

Tabla 8. Análisis de Regresión Múltiple

Variable Independ.	Variable dependiente <i>lnY</i>				
	Parámetro	Estimación	Error estándar	Estadística t	Valor P
	β_0	6.55608	0.225558	29.0661	0.0000
lnx_4	β_4	2.56456	0.0391883	65.442	0.0000

Las siguientes son las ecuaciones del modelo linealizado ajustado a los datos y el modelo original estimado respectivamente.

$$ln\hat{y} = 6.55608 + 2.56456lnx_4$$

$$\hat{y} = 703.50853 \cdot x_4^{2.56456}$$

El coeficiente de determinación ajustado por el número de grados de libertad R^2_{Adj} indica que el modelo explica el 99.51% de la variabilidad observada en *lnY*. La figura 12 representa gráficamente la concordancia entre los datos, o valores observados, y los valores calculados a partir del modelo linealizado estimado, o valores predichos.

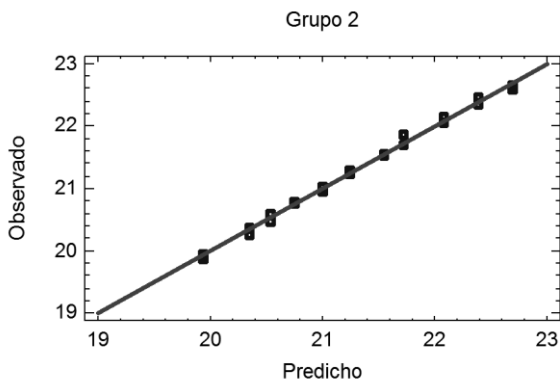


Figura 12

La figura 13, representa gráficamente el efecto lineal positivo de $\ln x_4$ sobre la recaudación.

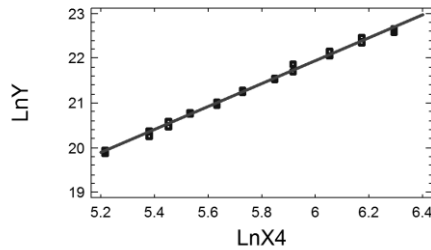


Figura 13

El gráfico de probabilidad normal muestra que los residuos caen aproximadamente sobre la línea recta del diagrama, señalando normalidad.

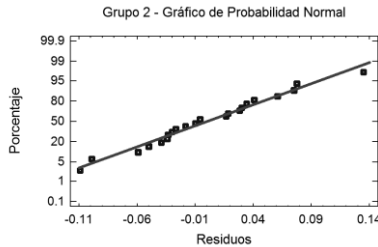


Figura 14

6. *Análisis del Modelo no Linealizado*

La concordancia entre los valores observados y los calculados a partir del modelo no linealizado se representa en la figura 15 y se sustenta en una prueba de hipótesis para la pendiente de la recta de regresión que contiene al origen ajustada a dichos valores, la que conduce a no rechazar la hipótesis de que la pendiente es igual a uno (Valor $P=0.33$) y en el coeficiente $R^2_{Adj}=0.99$.

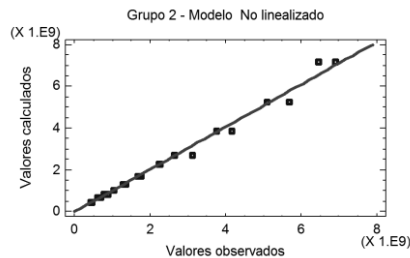


Figura 15

GRUPO III

MENDOZA, NEUQUÉN Y TUCUMÁN (2002–2012)

1. Estadística Descriptiva

Tabla 9. Estadística descriptiva

Descripción	Provincias		
	Mendoza	Neuquén	Tucumán
Cantidad de datos	11	11	11
Promedio	1.08893E9	8.2378E8	8.08142E8
Varianza	9.69657E17	4.36154E17	5.12497E17
Mínimo	1.8242E8	1.64275E8	1.06256E8
Máximo	3.53709E9	2.24036E9	2.39682E9
Coefic. estand. De asimetría	2.37418	1.67494	1.70791
Coefic. estand. De curtosis	2.25476	0.560884	0.74388
Coefficiente de variación	90.4292%	80.1694%	88.5845%

2. Diagrama de Dispersión

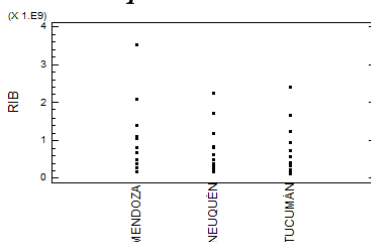


Figura 16

Mendoza no presenta normalidad. Como se puede ver en la figuras16 y17, la recaudación del impuesto a los ingresos brutos en al año 2012 se aleja mucho del centro de la correspondiente distribución. A pesar de ello una prueba de Cochran para la igualdad de varianza permite concluir que no difieren significativamente entre sí ($C=0.5055$, $Valor P=0.2498$).

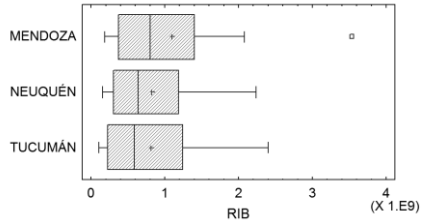


Figura 17

3. Correlaciones

Tabla 10

Gr up o III	Mendoza				Neuquén		Tucumán	
	X ₁	X ₃	X ₄	X ₆	X ₁	X ₃	X ₁	X ₃
Y	0.878	0.019	0.093	0.043	0.087	0.090	0.092	0.094
	0.000	0.028	0.000	0.001	0.000	0.000	0.000	0.000
X ₁		0.009	0.097	0.059		0.099		0.099
		0.060	0.000	0.000		0.000		0.000
		0.008	0.000	0.000		0.000		0.000

4. Matriz de Dispersión

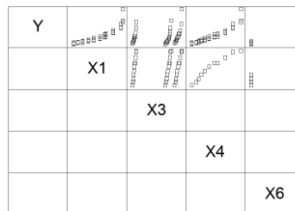


Figura 18

5. Análisis de Regresión Múltiple

La tabla siguiente muestra las estimaciones obtenidas, las estadísticas *t* y los valores P para pruebas acerca de la significación de los coeficientes de regresión. Todos los valores P son menores que 0.05 señalando regresión significativa.

Tabla 11

Variable Independ.	Variable dependiente <i>lnY</i>				
	Parámetro	Estimación	Error estándar	Estadística t	Valor P
	$ln\beta_0$	4.8218	0.8425	5.7230	0.0000
lnx_3	β_3	0.1190	0.0525	2.2666	0.0311
lnx_4	β_4	2.3565	0.1017	23.1606	0.0000
lnx_6	β_6	0.3118	0.0863	3.6140	0.0011

Las siguientes son las ecuaciones del modelo linealizado ajustado a los datos y el modelo original estimado respectivamente.

$$ln\hat{y} = 4.8218 + 0.1190 \cdot lnx_3 + 2.3565 \cdot lnx_4 + 0.3118 \cdot lnx_6$$

$$\hat{y} = 124.1884 \cdot x_3^{0.1190} \cdot x_4^{2.3565} \cdot x_6^{0.3118}$$

La estadística R² ajustada por el número de grados de libertad indica que el modelo explica el 97.02% de la variabilidad observada en Y. Una regresión por pasos hacia adelante arroja que lnx_4 explica 95.68% de la variabilidad observada en Y. La figura 19 representa gráficamente la concordancia entre los datos, o valores observados, y los valores calculados a partir del modelo linealizado estimado, o valores predichos.

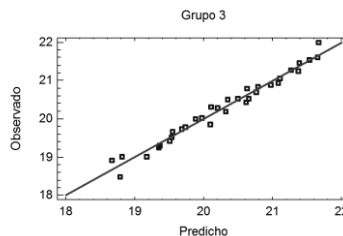


Figura 19

Las gráficas de las figuras 20, 21 y 22 muestran la parte del modelo ajustado que vincula $\ln Y$ con $\ln x_3$ y con $\ln x_4$ respectivamente, así como también el efecto positivo de la variable x_6 , ratificando que la variable $\ln x_4$ es la más importante.

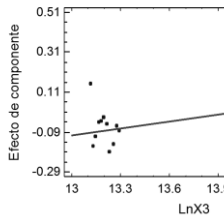


Figura 20

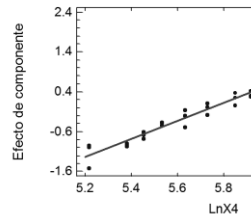


Figura 21

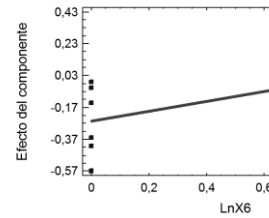


Figura 22

El gráfico de probabilidad normal muestra que los residuos caen aproximadamente sobre la línea recta del diagrama, señalando normalidad.

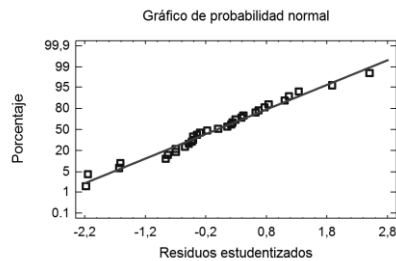


Figura 23

6. *Análisis del Modelo no Linealizado*

La concordancia entre los valores observados y los calculados a partir del modelo no linealizado se representa en la figura 24 y se sustenta en una prueba de hipótesis para la pendiente de la recta de regresión que contiene al origen ajustada a dichos valores, la que conduce a no rechazar la hipótesis de que la pendiente es igual a uno (Valor $P=0.17$) y en el coeficiente $R^2_{Adj} = 0.99$.

Se excluyó el dato correspondiente a Mendoza, año 2012.

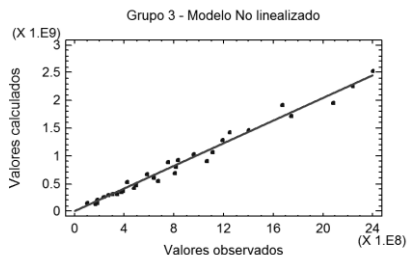


Figura 24

El coeficiente de correlación lineal ajustado por el número de grados de libertad es igual a 99.28 y una prueba para la pendiente igual a 1 de la recta de regresión conduce a no rechazar esta hipótesis con un valor P igual a 0.1618. Se excluyó el dato correspondiente a Mendoza, año 2012.

GRUPO IV

CHUBUT, ENTRE RÍOS, MISIONES, RÍO NEGRO, SALTA Y SANTA CRUZ (2002–2012)

1. Estadística Descriptiva

Tabla 12. Estadística descriptiva para la recaudación por el impuesto a los ingresos brutos.

Descripción	Provincias					
	Chubut	Entre Ríos	Misiones	Río Negro	Salta	Santa Cruz
Cantidad de datos	11	11	11	11	11	11
Promedio	5.0570 2E8	5.425E 8	5.4473 E8	4.1970 6E8	5.5352 E8	4.3115 4E8
Varianza	1.6416 7E17	1.7313 7E17	2.2654 6E17	1.0292 1E17	1.6312 E17	1.2126 7E17
Mínimo	8.4846 7E7	1.2192 3E8	8.7613 5E7	7.4864 9E7	1.2566 E8	3.7477 5E7
Máximo	1.3439 3E9	1.4267 7E9	1.6125 7E9	1.0892 8E9	1.3542 2E9	1.2054 9E9
Coefic.	1.2054	1.5494	1.7432	1.3898	1.2695	1.6123

estand. de asimetría	9E9	3	7	9	6	3
Coefic. estand. de curtosis	0.234688	0.431264	0.87936	0.252108	0.00435053	0.830957
Coeficiente de variación	80.1215%	76.6999%	87.377%	76.4377%	72.9659%	80.768%

2. Diagrama de Dispersión

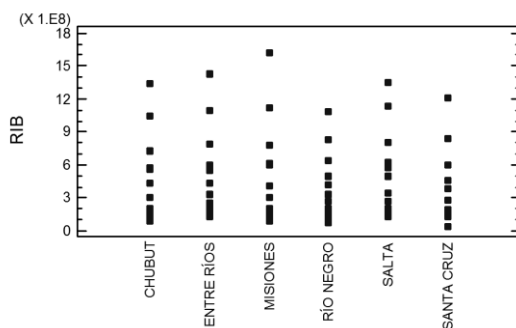


Figura 25

3. Correlaciones

Tabla 13

Grupo IV	X ₁	X ₃	X ₄	X ₆	Chubut	X ₁	X ₃	Entre Ríos	X ₁	X ₃	Misiones	X ₁	X ₃
	Y	0.924	0.212	0.968		0.198	Y		0.950	0.9612		Y	0.937
	0.00	0.08	0.00	0.11		0.00	0.00		0.00	0.00		0.00	0.00

	0	9	0	1		0			0	0		0	0
X ₁		0.	0.	0.	X ₁		0.9	X ₁		0.	X ₁		0.
		09	97	22			29			98			99
		9	8	5			0.0			0.			0.
		12	00	06			00			00			00
		8	0	9			00			0			0

Río Negro	X ₁	X ₃	Salta	X ₁	X ₃	Santa Cruz	X ₁	X ₃
Y	0.948	0.972	Y	0.956	0.934	Y	0.930	0.930
	0.000	0.000			0.000		0.000	0.000
X ₁		0.931	X ₁		0.994	X ₁		0.992
		0.000			0.000			0.000

4. Matriz de Dispersión

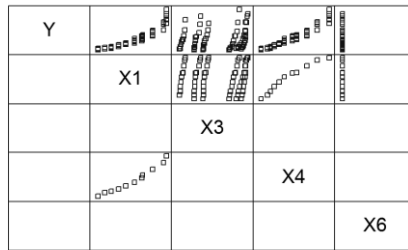


Figura 26

5. Análisis de la Regresión Múltiple

La tabla siguiente muestra las estimaciones obtenidas, las estadísticas *t* y los valores P para pruebas acerca de la significación de los coeficientes de regresión. Todos los valores P son menores que 0.05 señalando regresión significativa.

Tabla 14

Variable Independ.	Variable dependiente <i>lnY</i>				
	Parámetro	Estimación	Error estándar	Estadística <i>t</i>	Valor P
	β_0	-52.2797	11.2062	-4.6652	0.0000
lnx_1	β_1	2.4265	0.4927	4.9249	0.0000
lnx_3	β_3	0.1892	0.0290	6.5136	0.0000
lnx_4	β_4	0.8666	0.3337	2.5968	0.0117

Las siguientes son las ecuaciones del modelo linealizado ajustado a los datos y el modelo original estimado respectivamente

$$\ln \hat{y} = -52.2797 + 2.4265 \cdot \ln x_1 + 0.1892 \cdot \ln x_3 + 0.8666 \cdot \ln x_4$$

$$\hat{y} = 1.9734 \cdot E^{-23} \cdot x_1^{2.4265} x_3^{0.1892} \cdot x_4^{0.8666}$$

El coeficiente de determinación ajustado por el número de grados de libertad R^2_{Adj} indica que el modelo explica el 97.35% de la variabilidad observada en $\ln Y$. Una regresión por pasos hacia adelante indica que $\ln x_1$ explica el 95.37% de la variabilidad observada en $\ln Y$ y cuando se agrega $\ln x_3$ este porcentaje sube al 97.11%. La figura 27 representa gráficamente la concordancia entre los datos, o valores observados, y los valores calculados a partir del modelo linealizado estimado, o valores predichos.

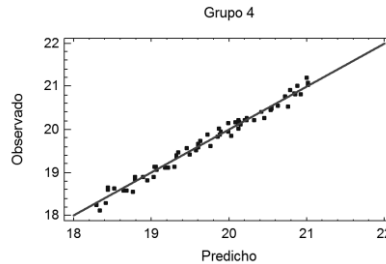


Figura 27

Las figuras 28, 29 y 30 muestran el efecto de cada componente del modelo ajustado sobre $\ln Y$.

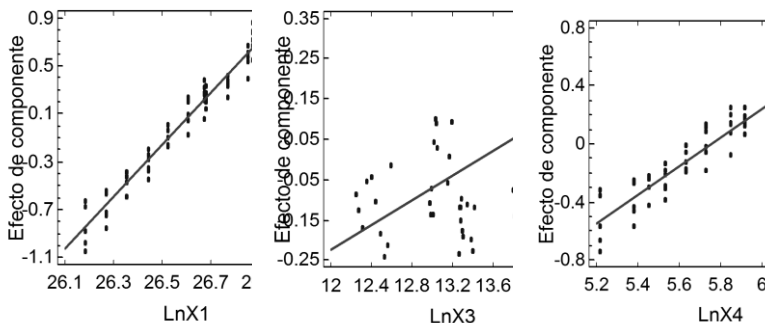


Figura 28

Figura 29

Figura 30

El gráfico de probabilidad normal muestra que los residuos caen aproximadamente sobre la línea recta del diagrama, señalando normalidad.

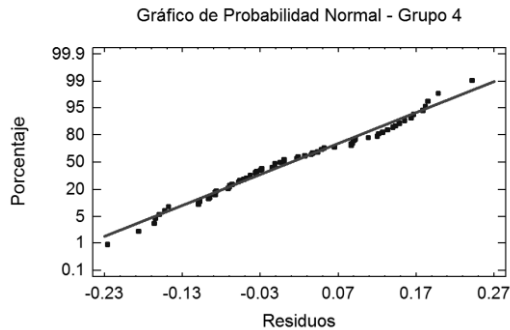


Figura 31

6. *Análisis del Modelo no Linealizado*

La concordancia entre los valores observados y los calculados a partir del modelo no linealizado se representa en la figura 32 y se sustenta en una prueba de hipótesis para la pendiente de la recta de regresión que contiene al origen ajustada a dichos valores, la que conduce a no rechazar la hipótesis de que la pendiente es igual a uno (*Valor P*=0.10, R^2_{Adj} =0.98).

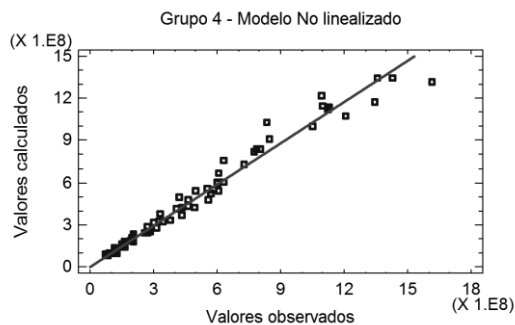


Figura 32

(Sin el dato correspondiente a Santa Cruz en el año 2002 que produce una fuerte distorsión en los residuos)

GRUPO V**CHACO – CORRIENTES – LA PAMPA – SAN JUAN –
SAN LUIS – TIERRA DEL FUEGO (2002–2012)****1. Estadística Descriptiva**

Tabla 15. Estadística descriptiva para la recaudación por el impuesto a los ingresos brutos.

Descripción	Provincias					
	Chaco	Corrientes	La Pampa	San Juan	San Luis	Tierra del Fuego
Cantidad de datos	11	11	11	11	11	11
Promedio	3.6615 8E8	2.7937 3E8	2.5094 5E8	2.6971 7E8	3.3323 7E8	2.6114 E8
Varianza	8.2053 E16	6.4735 E16	3.9152 E16	5.2380 E16	6.9052 E16	4.5248 E16
Mínimo	8.2639 2E7	5.8273 9E7	4.3924 8E7	4.3113 1E7	6.4945 6E7	4.2738 9E7
Máximo	9.2964 8E8	8.2287 E8	6.6606 1E8	7.6980 9E8	8.8284 1E8	7.1321 5E8
Coefic. estand. De asimetría	1.5194	1.6714 5	1.5856 7	1.6914 6	1.4677 8	1.5377 6
Coefic. estand. De curtosis	0.1390 64	0.4768 57	0.4449 87	0.6570 09	0.2538 06	0.4500 76
Coeficiente de variación	78%	91%	79%	85%	79%	81%

2. Diagrama de Dispersión

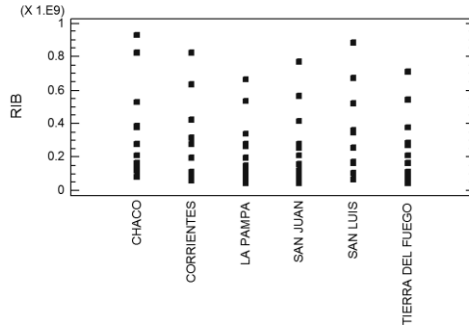


Figura 33

3. Correlaciones

Tabla 16

Grupo V	X ₁	X ₃	X ₄	X ₆	Chaco	X ₁	X ₃	Corrientes	X ₁	X ₃
Y	0.9098	0.1510	0.9582	0.2649	Y	0.9352	0.9625	Y	0.9175	0.8775
	0.0000	0.2262	0.0000	0.0316		0.0000	0.0000		0.0001	0.0004
X ₁		0.0606	0.9783	0.3444	X ₁		0.9907	X ₁		0.9895
		0.6290	0.0000	0.0046			0.0000			0.0000

La Pampa	X ₁	X ₃	San Juan	X ₁	X ₃	San Luis	X ₁	X ₃	Tierra del Fuego	X ₁	X ₃
Y	0.9345	0.8984	Y	0.9268	0.8639	Y	0.9390	0.8929	Y	0.9389	0.9328
	0.0000	0.0002		0.0000	0.0006		0.0000	0.0002		0.0000	0.0000
X ₁		0.9695	X ₁		0.9803	X ₁		0.9892	X ₁		0.9949
		0.0000			0.0000			0.0000			0.0000

		000			000			000			000
--	--	-----	--	--	-----	--	--	-----	--	--	-----

4. Matriz de Dispersión

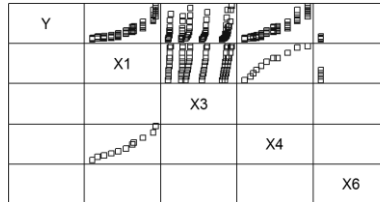


Figura 34

5. Análisis de la Regresión Múltiple

La tabla siguiente muestra las estimaciones obtenidas, las estadísticas t y los valores P para pruebas acerca de la significación de los coeficientes de regresión. Todos los valores P son menores que 0.05 señalando regresión significativa.

Tabla 17

Variable Independ.	Variable dependiente $\ln Y$				
	Parámetro	Estimación	Error estándar	Estadística t	Valor P
	β_0	3.6429	0.5416	6.72648	0.0000
$\ln x_3$	β_3	0.08738	0.0301	2.9002	0.0051
$\ln x_4$	β_4	2.5024	0.0686	36.4565	0.0000

Las siguientes son las ecuaciones del modelo linealizado ajustado a los datos y el modelo original estimado respectivamente.

$$\ln \hat{y} = 3.6429 + 0.08738 \cdot \ln x_3 + 2.5024 \cdot \ln x_4$$

$$\hat{y} = 38.2025 \cdot x_3^{0.08738} \cdot x_4^{2.5024}$$

El coeficiente de determinación ajustado por el número de grados de libertad R^2_{Adj} indica que el modelo explica el 95.41% de la variabilidad observada en $\ln Y$. Una regresión por pasos hacia adelante indica que $\ln x_4$ explica el 94.88% de la variabilidad observada en $\ln Y$ y cuando se agrega $\ln x_3$ este porcentaje sube al

95.41%. La figura 35 representa gráficamente la concordancia entre los datos, o valores observados, y los valores calculados a partir del modelo linealizado estimado, o valores predichos.

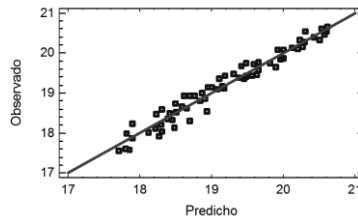


Figura 35

Las figuras 36 y 37, muestran la parte del modelo ajustado que vincula $\ln Y$ con $\ln x_4$ y con $\ln x_3$ respectivamente, ratificando que la variable $\ln x_4$ es la más importante.

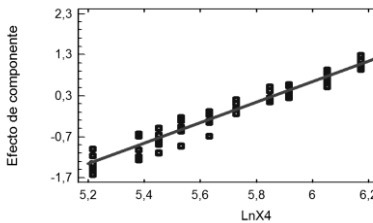


Figura 36

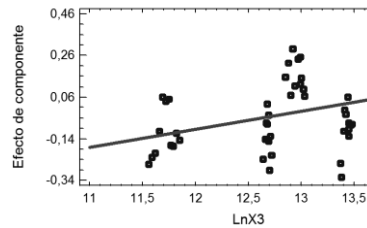


Figura 37

El gráfico de probabilidad normal muestra que los residuos caen aproximadamente sobre la línea recta del diagrama, señalando normalidad.

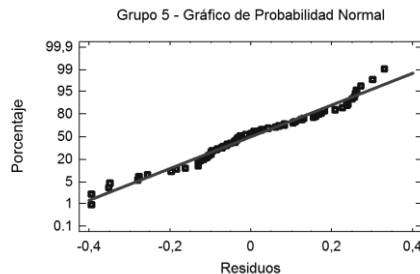


Figura 38

6. *Análisis del Modelo no Linealizado*

La concordancia entre los valores observados y los calculados a

partir del modelo no linealizado se representa en la figura 39 y se sustenta en una prueba de hipótesis para la pendiente de la recta de regresión que contiene al origen ajustada a dichos valores, la que conduce a no rechazar la hipótesis de que la pendiente es igual a uno (*Valor P*=0.74, R^2_{Adj} =0.98).

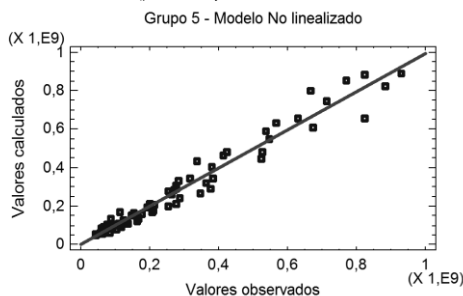


Figura 39

GRUPO VI

CATAMARCA – JUJUY – SANTIAGO DEL ESTERO (2002–2012)

1. *Estadística Descriptiva*

Tabla 18. Estadística descriptiva para la recaudación por el impuesto a los ingresos brutos.

Descripción	Provincias		
	Catamarca	Jujuy	Santiago del Estero
Cantidad de datos	11	11	11
Promedio	1.69514E8	1.87522E8	2.13941E8
Varianza	2.12817E16	2.44148E16	2.83987E16
Mínimo	2.54184E7	3.48092E7	4.67568E7
Máximo	4.53553E8	5.12035E8	5.60384E8
Coefic. estand. de asimetría	1.47374	1.53231	1.48545
Coefic. estand. de curtosis	0.0317353	0.265677	0.208406
Coefficiente de variación	86.0593%	83.3247%	78.7691%

2. Diagrama de Dispersión

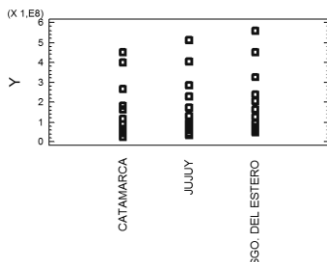


Figura 40

3. Correlaciones

Tabla 19

Gr up o VI	X ₁	X ₃	X ₄	X ₆	Cata marca	X ₁	X ₃	Ju ju y	X ₁	X ₃	S G E	X ₁	X ₃
Y	0.927	0.203	0.976	0.420	Y	0.936	0.808	Y	0.932	0.890	Y	0.939	0.939
	0.000	0.243	0.000	0.015		0.000	0.003		0.000	0.000		0.000	0.000
X ₁		0.101	0.978	0.037	X ₁		0.949	X ₁		0.987	X ₁		0.995
		0.578	0.000	0.032			0.000			0.000			0.000

4. Matriz de Dispersión

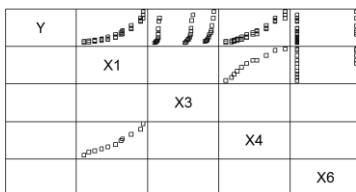


Figura 41

5. *Análisis de la Regresión Múltiple*

La tabla siguiente muestra las estimaciones obtenidas, las estadísticas t y los valores P para pruebas acerca de la significación de los coeficientes de regresión. Todos los valores P son menores que 0.05 señalando regresión significativa.

Tabla 20

Variable Independ.	Variable dependiente $\ln Y$				
	Parámetro	Estimación	Error estándar	Estadística t	Valor P
$\ln x_3$	β_3	0.32379	0.01899	17.0519	0.0000
$\ln x_4$	β_4	2.51037	0.04388	57.2116	0.0000

Las siguientes son las ecuaciones del modelo linealizado ajustado a los datos y el modelo original estimado respectivamente.

$$\ln \hat{y} = 0.32379 \cdot \ln x_3 + 2.51037 \cdot \ln x_4$$

$$\hat{y} = x_3^{0.32379} \cdot x_4^{2.51037}$$

El coeficiente de determinación ajustado por el número de grados de libertad R^2_{Adj} indica que el modelo explica el 99.99% de la variabilidad observada en Y . Una regresión por pasos hacia delante arroja que $\ln x_4$ explica la casi totalidad (el 99.98%) de la variabilidad observada en Y . La figura 42 representa gráficamente la concordancia entre los datos, o valores observados, y los valores calculados a partir del modelo linealizado estimado, o valores predichos.

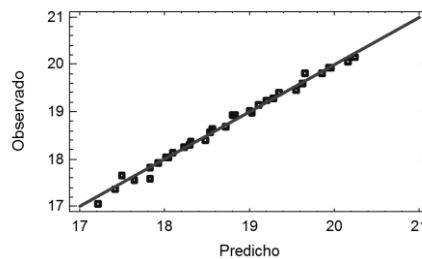


Figura 42

Las gráficas 43 y 44 muestran la parte del modelo ajustado que vincula $\ln Y$ con $\ln x_3$ y con $\ln x_4$ respectivamente, ratificando que la variable $\ln x_4$ es la más importante.

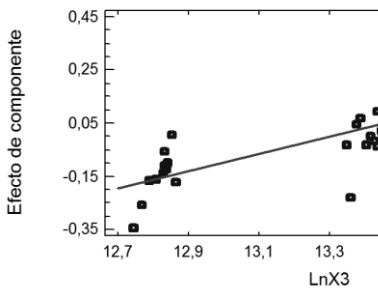


Figura 43

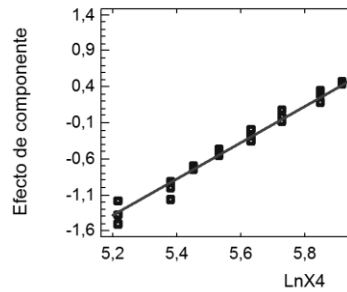


Figura 44

El gráfico de probabilidad normal muestra que los residuos caen aproximadamente sobre la línea recta del diagrama, señalando normalidad.

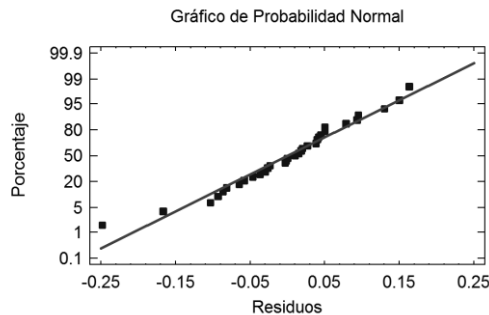


Figura 45

6. *Análisis del Modelo no Linealizado*

La concordancia entre los valores observados y los calculados a partir del modelo no linealizado se representa en la figura 46 y se sustenta en una prueba de hipótesis para la pendiente de la recta de regresión que contiene al origen ajustada a dichos valores, la que conduce a no rechazar la hipótesis de que la pendiente es igual a uno (Valor $P=0.84$, $R^2_{Adj}=0.99$). Tanto para el modelo linealizado como para el modelo sin linealizar, los valores observados son próximos a los estimados por el modelo ajustado, indicando buena capacidad de predicción (Figura 46).

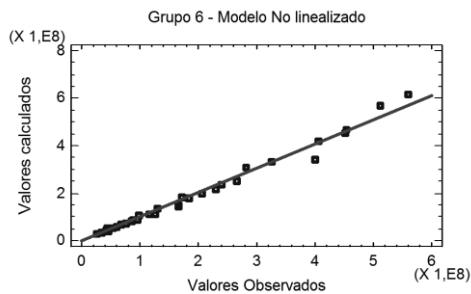


Figura 46

GRUPO VII

FORMOSA - LA RIOJA (2002–2012)

1. *Estadística Descriptiva*

Tabla 21. Estadística descriptiva para la recaudación por el impuesto a los ingresos brutos.

Descripción	Provincias	
	Formosa	La Rioja
Cantidad de datos	11	11
Promedio	1.28719E8	9.54307E7
Varianza	1.30588E16	5.12598E15
Mínimo	1.58726E7	1.22507E7
Máximo	3.70488E8	2.43599E8
Coefic. estand. de asimetría	1.56722	1.29008
Coefic. estand. de curtosis	0.417844	0.253518
Coeficiente de variación	88.7787%	75.024%

2. Diagrama de Dispersión

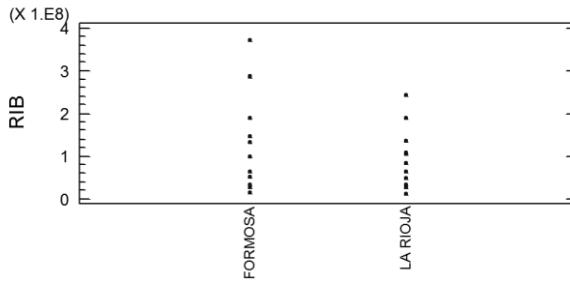


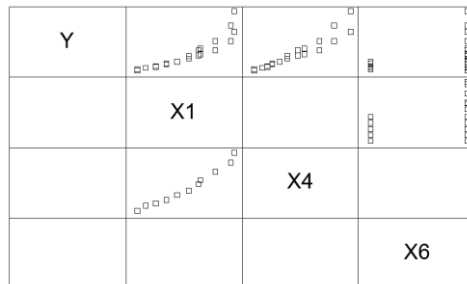
Figura 47

3. Correlaciones

Tabla 22

Grupo VII	X ₁	X ₄	X ₃	X ₆	Formosa	X ₁	X ₃	La Rioja	X ₁	X ₃
Y	0.9042	0.9459	0.3427	0.4333	Y	0.9347	0.9118	Y	0.9566	0.9159
	0.0000	0.0000	0.1184	0.0440		0.0000	0.0001		0.0000	0.0001
X ₁		0.9783	0.1815	0.5159	X ₁		0.9931	X ₁		0.9856
		0.0000	0.4188	0.0140			0.0000			0.0000

4. Matriz de Dispersión



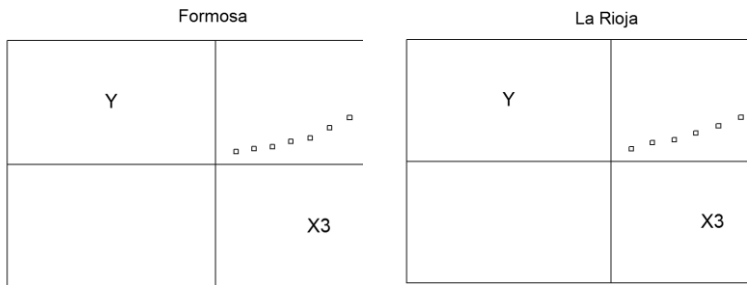


Figura 48

5. *Análisis de la Regresión Múltiple*

La tabla siguiente muestra las estimaciones obtenidas, las estadísticas t y los valores P para pruebas acerca de la significación de los coeficientes de regresión. Todos los valores P son menores que 0.05 señalando regresión significativa.

Tabla 23

Variable dependiente $\ln Y$				
Parámetro	Estimación	Error estándar	Estadística t	Valor P
$\ln \beta_0$	-93.1313	3.2896	-28.3108	0.0000
β_1	3.9886	0.1218	32.7458	0.0000
β_3	0.4121	0.1110	3.7141	0.0015

Para este grupo x_4 no aportó significativamente a la variabilidad observada en $\ln Y$. Durante los años 2002-2012 la provincia de La Rioja tuvo apoyo administrativo, y en apariencia su aporte no es significativo. Las siguientes son las ecuaciones del modelo linealizado ajustado a los datos y el modelo original estimado respectivamente.

$$\ln \hat{y} = -93.313 + 3.9886 \cdot \ln x_1 + 0.4121 \cdot \ln x_3$$

$$\hat{y} = 2.98 \cdot 10^{-41} \cdot x_1^{3.9886} \cdot x_3^{0.4121}$$

El coeficiente de determinación ajustado por el número de grados de libertad R^2_{Adj} indica que el modelo explica el 98.23% de la

variabilidad observada en $\ln Y$. Una regresión por pasos hacia adelante indica que $\ln x_1$ explica el 97.11% de la variabilidad observada en $\ln Y$ y cuando se agrega $\ln x_3$ este porcentaje sube al 98.23%. La figura 49 muestra la concordancia entre los datos y los valores calculados a partir del modelo linealizado estimado.

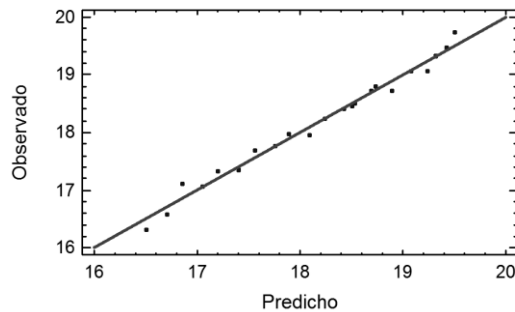


Figura 49

Las figuras 50 y 51, muestran la parte del modelo ajustado que vincula $\ln Y$ con $\ln x_1$ y con $\ln x_3$ respectivamente, confirmando que esta última es la variable más importante para el ajuste.

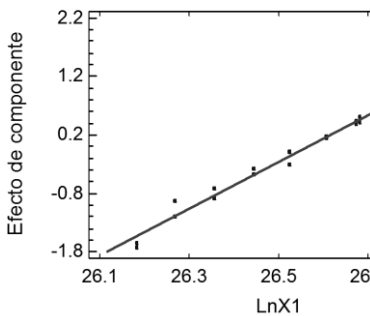


Figura 50

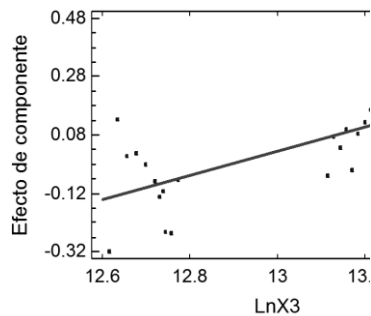


Figura 51

El gráfico de probabilidad normal muestra que los residuos caen aproximadamente sobre la línea recta del diagrama, señalando normalidad.

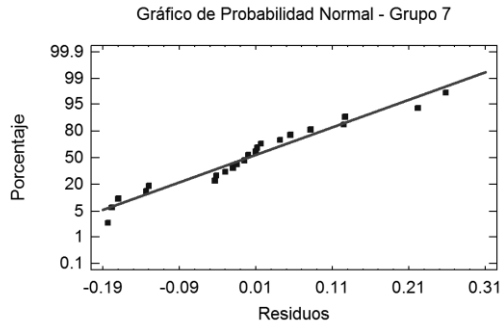


Figura 52

6. *Análisis del Modelo no Linealizado*

La concordancia entre los valores observados y los calculados a partir del modelo no linealizado se representa en la figura 53 y se sustenta en una prueba de hipótesis para la pendiente de la recta de regresión que contiene al origen ajustada a dichos valores, la que conduce a no rechazar la hipótesis de que la pendiente es igual a uno (*Valor P*=0.12, R^2_{Adj} =0.98).

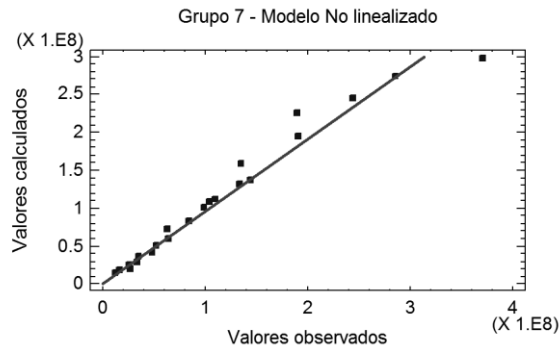


Figura 53

Después de lo expuesto, y a modo de conclusión, podemos afirmar que:

- Se propuso como hipótesis explicativa del comportamiento de la recaudación tributaria provincial, al siguiente modelo:

$$H_0: R_t: f(PBI, PBG, NA, \pi, A, S)$$

- Con salvedades y limitaciones el modelo ha podido ser razonablemente contrastado con datos empíricos mediante el empleo de instrumental econométrico, para las provincias analizadas.
- Se ha podido implementar una metodología, quizás no lo suficientemente difundida en el ámbito de la investigación de la Facultad de Ciencias Económicas y Empresariales de la Universidad Católica de Cuyo, pero imprescindible si se quiere comenzar a transitar un camino que lleve a los estándares internacionales de calidad en la producción académica y científica.
- Se han adquirido nuevas habilidades en el manejo de software diseñados para descubrir y entender relaciones entre variables. Además se pueden ordenar y agrupar y construir y contrastar las hipótesis, imprescindibles para el desarrollo no solo de éste tipo de trabajo de investigación., sino de otros similares pero aplicados a las ciencias sociales, a las ciencias de la alimentación, a las ciencias jurídicas, etc.
- Se ha procedido consistentemente con lo que oportunamente se planteara como una parte fundamental de un “Programa de Investigación para las Ciencias Económicas” Como asimismo con otros trabajos de investigación desarrollados por los autores, presentados en Congresos de importancia nacional y publicados en publicaciones con referato de circulación nacional
- Se ha obtenido valiosa información para toma de decisiones que contribuyan no solo a una mejor y más eficiente administración tributaria sino al diseño de sistemas impositivos más justos y equitativos.
- Se ha hecho una contribución al cumplimiento de la indicación de la Ex Corde Ecclesiae, como asimismo para con la trascendencia que para la Iglesia, y para toda la comunidad tiene la ciencia y el desarrollo: “La Universidad Católica es, por consiguiente, el lugar donde los estudiosos examinan a fondo la realidad con los métodos propios de cada disciplina académica, contribuyendo así al enriquecimiento del saber humano. Cada disciplina se estudia de manera sistemática, estableciendo después un diálogo entre las diversas disciplinas con el fin de enriquecerse mutuamente. Tal investigación, además de ayudar a los hombres y mujeres en la búsqueda constante de la verdad, ofrece un eficaz

testimonio, hoy tan necesario, de la confianza que tiene la Iglesia en el valor intrínseco de la ciencia y de la investigación”

- Por último, se ha hecho un aporte original, que por pequeño y limitado que sea, contribuirá al progreso y desarrollo de la ciencia económica en general, de la economía política, y de la microeconomía en particular. El aporte de este trabajo podría alentar a los investigadores a encontrar las variables propias de su jurisdicción y aplicar el modelo aquí desarrollado, a los fines de establecer los parámetros de las variables que influyen en la función, determinar cuál de ellos es el más influyente de tal manera que al ajustarlos mediante políticas públicas, puedan conseguir un aumento en la recaudación.

Bibliografía

- ARMSTRONG, J. (2001) “Principles of Forecasting a Handbook for Researchers and Practitioners”. Kluwer Academic Publishers. Norwell.
- GARCÍA, R. (1992) “Inferencia Estadística y Diseño de Experimentos”. Editorial Eudeba. Pg. 257.
- HARSHBARGER, R. (2005) “Matemática Aplicada a la Administración, economía y ciencias sociales”, Editorial McGraw–Hill, México. Págs. 268, 276, 381, 574, 575, 580, 586, 749, 749, 769, 773, 775, 810 y 912 a 958.
- JENKINS, G; CHUN-YAN K.; SHUKLA G. (2000) “Tax analysis and revenue forecasting. Issues and Techniques”. Harvard Institute for International Development. Harvard University.
- LUQUE MARTINEZ, T. (2012) “Técnicas de análisis de datos en investigación de Mercados”. Ediciones Pirámide. Grupo Anaya S.A. Pág. 337.
- PARENT, B. (1995) "La Estimación de los Ingresos Tributarios y el Análisis de las Desviaciones", en Conferencia Técnica del C.I.A.T., San Carlos de Bariloche. Pág. N° 163.
- ROSA DONATI, S. (2007) “Los Programas de Investigación en la Producción Científica Universitaria: Un aporte para el desarrollo de la Teoría y la Política Económica”. Paper.

Facultad de Ciencias Económicas y Empresariales. Universidad Católica de Cuyo.

- ROSA DONATI, S.; GARCÍA, L. y otro (2007) “Interoperabilidad Informática y algo más. Desarrollo del “Sistema de Administración Tributaria Integral”, en el marco de un Proyecto de Soporte Tecnológico y de Gestión, con alto grado de Interoperabilidad Informática, Científica, Académica y Social. El caso de la Dirección General de Ingresos Provinciales de la Provincia de La Rioja y su relación con Programas y Proyectos de Investigación.” Presentado en el Congreso de Informática del Nuevo Cuyo. San Juan.
- ROSA DONATI, S.; GARCÍA, L.; GUTIÉRREZ, M. (2014) “Relevamiento de la situación de recaudación en las provincias. Determinantes de la función de Recaudación Impositiva Provincial”. Revista Vozes dos Vales: Publicações Acadêmicas. Brasil, en <http://www.ufvjm.edu.br/voze>
- WATSON, F. (1995) "La Utilización de modelos de simulación para la proyección de los ingresos tributarios", en Conferencia Técnica del C.I.A.T., San Carlos de Bariloche. Solución de Aspectos Críticos en el Control de los Tributos. Pág. 191.

Webgrafía

- MARTÍN, F. “Los factores determinantes de la recaudación tributaria”, en www.mecon.gov.ar/sip/basehome/factores_recaudacion_tributaria.pdf
- MARTÍN, F. “Elementos para un Mejor Análisis de la Recaudación Tributaria”. Revista de la ASAP Nro. 37. Noviembre de 2001. Buenos Aires, Argentina, en http://www.mecon.gov.ar/hacienda/info_sal.htm