

MODELADO DEL COMPORTAMIENTO ASIMETRICO DE LOS COSTOS EN EMPRESAS ARGENTINAS: INCORPORANDO CASOS ATIPICOS CON GAMLSS

María Inés Stimolo- José Vargas

Universidad Nacional de Córdoba - Facultad de Ciencias Economicas

Wolkshop Instituto de Estadística y Demografía - Octubre 2017 -

La función de costos de las empresas mide el mínimo costo de producción a un nivel dado de actividad para determinado nivel de precio de los factores de producción.

Proporcionalidad simétrica

Establece una proporcionalidad simétrica entre el cambio de los costos y el nivel de producción. Los costos aumentan y disminuyen en la misma proporción ante un aumento o disminución del 1 por ciento en el nivel de producción.

El supuesto de proporcionalidad no se cumple en la práctica

Comportamiento pegadizo

Los costos aumentan una determinada proporción ante un aumento del 1 por ciento en la actividad pero disminuyen en una proporción menor ante una disminución del 1 por ciento en el nivel de actividad.

Este problema fue analizado por distintos autores.

- Malcom 1991
- Noreen 1994,1997
- Mark and Roush 1994

Primer trabajo empirico:

Anderson M.C.; Banker R.y Janakiraman S.N. (2003)

Función de costos de la empresa

A partir de la función de producción de Cobb Douglas, en el corto corto plazo podemos clasificar los factores de producción como fijos (x) y variables (z).

La función de costos a minimizar para un nivel de x

$$\begin{aligned} \text{mín } C(q/x) &= \text{mín } (p_x x + p_z z) \\ \text{donde } q &= z^\alpha x^\beta \end{aligned}$$

α y β representan la proporción de cada factor en la producción total
 $p_x p_z$ representan los precios de los factores de producción fijos y variables, respectivamente.

$p_x x$ representa los costos fijos y $p_z z$ los costos variables.

$$C(q/p_x, p_z, x) = p_x x + p_z \left(q x^{-\beta} \right)^{\frac{1}{\alpha}}$$

La cantidad de producto z necesario para un volumen q de producción para un nivel dado de costo fijo x es igual a $z^*(q/x) = \left(q x^{-\beta} \right)^{\frac{1}{\alpha}}$

Funcion de costos de la empresa

Asumiendo que los precios de los factores son constantes la función de costos a corto plazo depende del nivel de producción

$$C(q) = Kq^{\frac{1}{\alpha}} + F \text{ donde } F = p_x x \text{ y } K = p_z x^{\frac{-\beta}{\alpha}}$$

El cambio de los costos entre dos periodos -1 y t puede ser expresado como:

$$\frac{C_t(q_t)}{C_{t-1}(q_{t-1})} = \frac{K_t}{K_{t-1}} \left(\frac{q_t}{q_{t-1}} \right)^{\frac{1}{\alpha}} + 1$$

y tomando logaritmo en ambos miembros

$$\ln \frac{C_t(q_t)}{C_{t-1}(q_{t-1})} = \ln \frac{K_t}{K_{t-1}} + \frac{1}{\alpha} \ln \left(\frac{q_t}{q_{t-1}} \right)$$

$$\ln \frac{C_t(q_t)}{C_{t-1}(q_{t-1})} = \beta_0 + \beta_1 \ln \left(\frac{q_t}{q_{t-1}} \right) \text{ donde } \beta_0 = \ln \frac{K_t}{K_{t-1}} \text{ y } \beta_1 = \frac{1}{\alpha}$$

Costos para la toma de decisiones

El modelo de costos dado es la base del modelo de Costo Volumen Utilidad (CVP) en contabilidad de gestión y contabilidad estratégica de Costos. La toma de decisiones se basa en tres variables fundamentales

- costos fijos
- unidades producidas o vendidas
- margen de contribución unitaria

Todos los métodos de identificación, medición, rastreo, asignación y reporte de costos están basados en la distinción de costos fijos y variables y la relación proporcionalmente simétrica con el volumen de operaciones

Comportamiento asimétrico

Este problema de proporcionalidad asimétrica entre costos y cambios en el nivel de ventas, empezó a explicitarse en algunos trabajos originando el estudio del mismo y dando inicio al desarrollo de la investigación empírica de este problema. .

Modelo ABJ

$$\ln \left[\frac{C_{i,t}}{C_{i,t-1}} \right] = \beta_0 + \beta_1 \ln \left[\frac{V_{i,t}}{V_{i,t-1}} \right] + \beta_2 * dec_{i,t} * \ln \left[\frac{V_{i,t}}{V_{i,t-1}} \right] + \varepsilon_{i,t}$$

$$i = 1, \dots, n \text{ y } t = 1, \dots, n_i$$

$C_{i,t}$ es el costo de la empresa i en el año t

$V_{i,t}$ es el ingreso o nivel de actividad de la empresa i en el año t

dec asume el valor 1 cuando el ingreso o nivel de actividad disminuye en la empresa i en el año t

β_1 es el coeficiente de cambio de los costos ante un aumento del 1 por ciento de los ingresos. Se interpreta como el ratio entre los costos variables y los costos totales $vq/CT(q)$, una mayor pendiente corresponde una estructura de costos a corto plazo menos rígida.

$\beta_1 + \beta_2$ es el coeficiente de cambio de los costos ante una disminución del 1 de los ingresos.

Interpretación de los tipos de asimetría

Relación entre coeficientes	Grado de asimetría $GA = \frac{\beta_1 + \beta_2}{\beta_1}$	Comportamiento costos
$\beta_2 < 0$ $ \beta_2 > \beta_1 $	$GA \leq 0$	Costos aumentan
$ \beta_2 < \beta_1 $	$0 < GA < 1$	Costos pegadizos
$\beta_2 = 0$	$GA = 1$	Proporcionalidad simétrica
$\beta_2 > 0$	$GA > 1$	Costos antipegadizos

Criterios para eliminar casos atípicos

Los autores utilizaron distintos criterios para eliminar datos atípicos:

- Anderson et al. 2003 elimina las observaciones del 5 % de los extremos superior o inferior para las variables del modelo (modela el 90 % central de los datos)

Criterios para eliminar casos atípicos

Los autores utilizaron distintos criterios para eliminar datos atípicos:

- Anderson et al. 2003 elimina las observaciones del 5 % de los extremos superior o inferior para las variables del modelo (modela el 90 % central de los datos)
- Calleja et al. 2006 elimina las ventas y costos cuyas variaciones fueron mayores al 50 %

Criterios para eliminar casos atípicos

Los autores utilizaron distintos criterios para eliminar datos atípicos:

- Anderson et al. 2003 elimina las observaciones del 5 % de los extremos superior o inferior para las variables del modelo (modela el 90 % central de los datos)
- Calleja et al. 2006 elimina las ventas y costos cuyas variaciones fueron mayores al 50 %
- Banker y Byzalov 2014 eliminan las observaciones donde el ratio costos sobre ventas sea mayor a 10 o menor a 0,1

Criterios para eliminar casos atípicos

Los autores utilizaron distintos criterios para eliminar datos atípicos:

- Anderson et al. 2003 elimina las observaciones del 5 % de los extremos superior o inferior para las variables del modelo (modela el 90 % central de los datos)
- Calleja et al. 2006 elimina las ventas y costos cuyas variaciones fueron mayores al 50 %
- Banker y Byzalov 2014 eliminan las observaciones donde el ratio costos sobre ventas sea mayor a 10 o menor a 0,1
- Stimolo y Diaz 2016 se utilizan los residuos estudentizados, y algunas medidas para identificar casos atípicos e influyentes

Resumen de las medidas para detectar atípicos e influyentes

Medida	Expresión analítica	Limites a examinar
$DFITS_i$	$rs_i \sqrt{\frac{h_i}{1-h_i}}$	$2\sqrt{(p+1)/n}$
Distancia de Cook D_i	$\frac{1}{(p+1)} \frac{S_{(i)}^2}{S^2} DFITS_i^2$	$\frac{4}{n}$
Distancia de Welsch W_i	$DFITS_i \sqrt{\frac{n-1}{1-h_i}}$	$3\sqrt{p+1}$
$COVRATIO_i$	$\frac{1}{1-h_i} \left(\frac{n-(p+1)-r_i^2}{n-(p+1)-1} \right)^{(p+1)}$	$ COVRATIO_i - 1 \geq \frac{3(p+1)}{n}$
$DFBETA_i$	$\frac{rS_i u_i}{\sqrt{U^2(1-h_i)}}$ u_i residuos de la regresión de X_i sobre otras X_j $U^2 = \sum_i u_i^2$	$ DFBETA_i > \frac{2}{\sqrt{n}}$

Base de Datos

Empresas argentinas listadas en el mercado de valores (excluye sector financiero) durante el período 2004-2012.

Información de los informes financieros disponibles en el sitio web de la Bolsa de Comercio de Buenos Aires

Se eliminan las observaciones en las que los costos son mayores que las ventas.

En total la muestra es de 733 observaciones para 155 empresas

Estimaciones con distintos criterios de eliminación

Cuadro: Resultados del modelo ABJ con distintos criterios

Criterio	Total	Extremos 5 %	Variación >50 %	Extremos Costo/Ventas	Residuos atípicos
$\hat{\beta}_0$	-0,002	0,015**	0,017**	-0,002	-0,028*
$\hat{\beta}_1$	0,99**	0,81**	0,77**	1,00**	0,916**
$\hat{\beta}_2$	0,11	0,14	0,29**	0,01	0,08
GA	1	1	1,38	1	1
R2 ajustado	0,7802	0,6675	0,7054	0,7854	0,8144
nro obs	633	563	579	631	598
Atípicos	0	70	54	2	35

Regresión Robusta

Estimación Robusta

El enfoque del modelo central asume que el modelo que se estima es el correcto pero parte de los datos están contaminados.

Se estiman los parámetros del modelo central, de modo que sean mínimamente influenciadas por la contaminación, esto es, que la distribución del error es un elemento de la vecindad de e contaminada de la verdadera (central) distribución del error.

Robustez de la estimación

La robustez en una estimación, está dada por: La **función de influencia** es una medida infinitesimal de la influencia de una observación. El **punto de ruptura** mide la proporción de observaciones contaminadas que un procedimiento de estimación robusta puede manejar antes de ser conducido en más o menos al infinito o al límite del espacio de parámetros.

Regresión Robusta MM-Estimadores

Los MM estimadores controlan robustez y eficiencia bajo el modelo normal

Para el modelo de regresión lineal

$$y_i = \mathbf{x}'_i \beta + e_i$$

donde $i = 1, 2, \dots, n$ y $e_i \sim N(0, \sigma^2)$ en el modelo central.

Los M estimadores en regresión están definidos como

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^n \rho \left(\frac{r_i(\hat{\beta})}{\hat{\sigma}} \right) \text{ donde } r_i = r_i(\hat{\beta}) = y_i - \mathbf{x}'_i \hat{\beta}$$

ρ es una función no decreciente de $|r|$ con $\rho(0) = 0$ y estrictamente creciente para $r > 0$ donde $\rho(r) < \rho(\infty)$. Si $\rho(\infty) = 1$, $\hat{\beta}$ se denomina **M estimador redescendente**

Diferenciando se obtiene la ecuación de estimación

$$\sum_{i=1}^n \psi \left(\frac{r_i(\hat{\beta})}{\hat{\sigma}} \right) \mathbf{x}_i = 0$$

Donde ψ es proporcional a ρ' y $\psi'(0) = 1$

Regresión Robusta MM-Estimadores

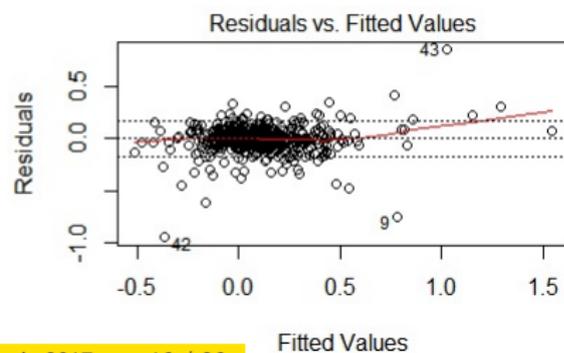
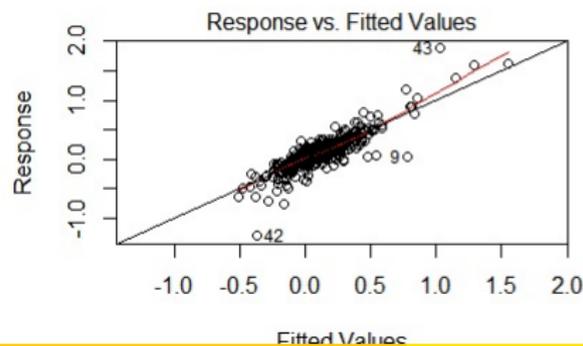
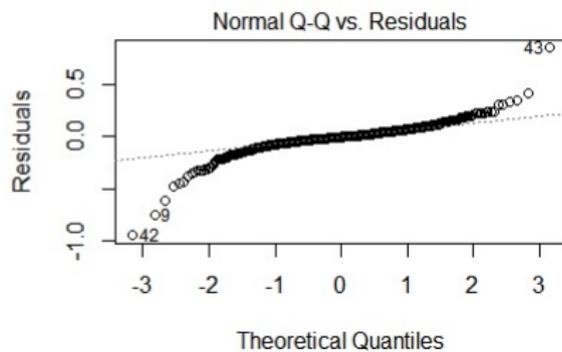
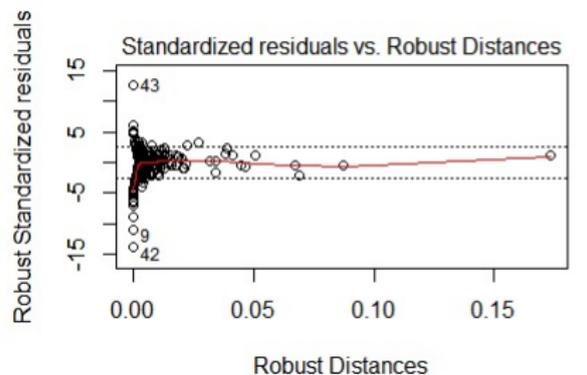
- 1 A partir de los residuos $\mathbf{e} = (e_1, e_2, \dots, e_n)$ se calcula el M estimador $\hat{\sigma}_0$ como solución de $\frac{1}{n} \sum_{i=1}^n \chi \left(\frac{r_i(\hat{\beta})}{\hat{\sigma}} \right) = k$ donde k es una constante de ajuste y $\chi(e_i)$ cumple las mismas propiedades que $\rho(r)$
- 2 S-estimaciones de la regresión son los valores de los parámetros $\hat{\beta}_s$ que minimizan la M-estimación de la escala

$$\hat{\sigma}_s = \hat{\sigma}_s(r_i(\hat{\beta}))$$

y de los residuos asociados, $\hat{\beta}_s = \arg \min_{\beta} \sigma_s(r(\beta))$

- 3 El máximo punto de ruptura de un S estimador $(1 - p/n)/2$ se obtiene cuando $k = (1 - p/n)/2$.
- 4 IRWLS (iteratively reweighted least squares) comienza con un $\hat{\beta}$ inicial a partir de sus residuos se calcula un S-estimador inicial $\hat{\beta}_s$. Para una selección de ρ el MM estimador preserva el punto de ruptura de $\hat{\beta}_s$. Ajustando k , se obtiene la eficiencia asintótica deseada del estimador.

Regresión Robusta. Análisis de residuos



Modelos GAMLSS (*Generalized Additive Model for Location, Scale and Shape*)

- 1 Un modelo GAMLSS, en su versión más simple generaliza un GLM.
- 2 En un GLM, la respuesta tiene distribución en la Flia exponencial y el objetivo es modelar el valor esperado de la respuesta con un predictor lineal a través de una función "link" g

$$Y \sim \text{ExpDist}(\mu, \sigma), g(\mu) = X\beta$$

- 3 En un modelo GAMLSS, la flia exponencial se generaliza a una cantidad mayor de distribuciones de uno, dos, tres o cuatro parámetros (casi 100 distribuciones). Y en vez de modelar un parámetro se modelan hasta cuatro:

$$Y \sim \text{Dist}(\mu, \sigma, \nu, \tau)$$

$$g_1(\mu) = X_1\beta_1, g_2(\sigma) = X_2\beta_2, g_3(\nu) = X_3\beta_3, g_4(\tau) = X_4\beta_4$$

- 4 Los parámetros ν y τ se relacionan con asimetría y kurtosis.

Modelos GAMLSS. Resultados para Costo de Ventas

```
*****
Family: c("TF", "t Family")

Call: gamlss(formula = lnCosto_Ventas_T_1 ~ lnvtas_t_1 +
  DECvtas_t_1, family = TF, data = costos1)

Fitting method: RS()

-----
Mu link function: identity
Mu Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.003550   0.004235   0.838   0.402
lnvtas_t_1   0.954276   0.025809  36.974 <2e-16 ***
DECvtas_t_1 -0.003137   0.063397  -0.049   0.961
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
Sigma link function: log
Sigma Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.88221    0.05798 -49.71 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

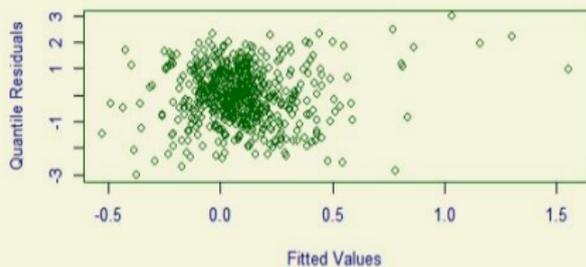
-----
Nu link function: log
Nu Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.7715    0.1121   6.883 1.43e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
No. of observations in the fit: 633
Degrees of Freedom for the fit: 5
Residual Deg. of Freedom: 628
                             at cycle: 9

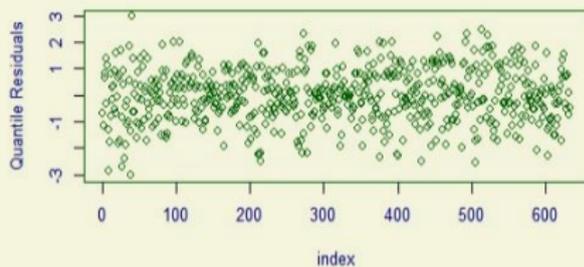
Global Deviance: -1221.239
AIC: -1211.239
SBC: -1188.987
*****
```

Modelos GAMLSS. Análisis de los residuos

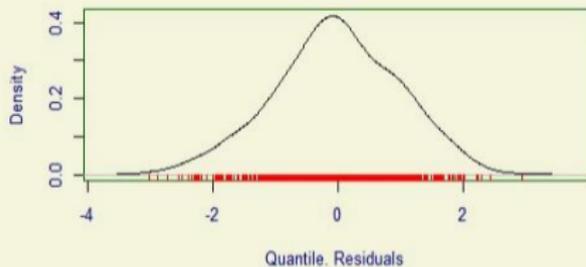
Against Fitted Values



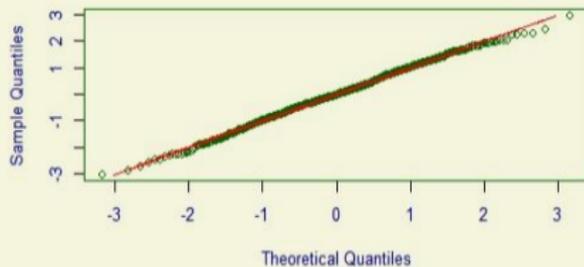
Against index



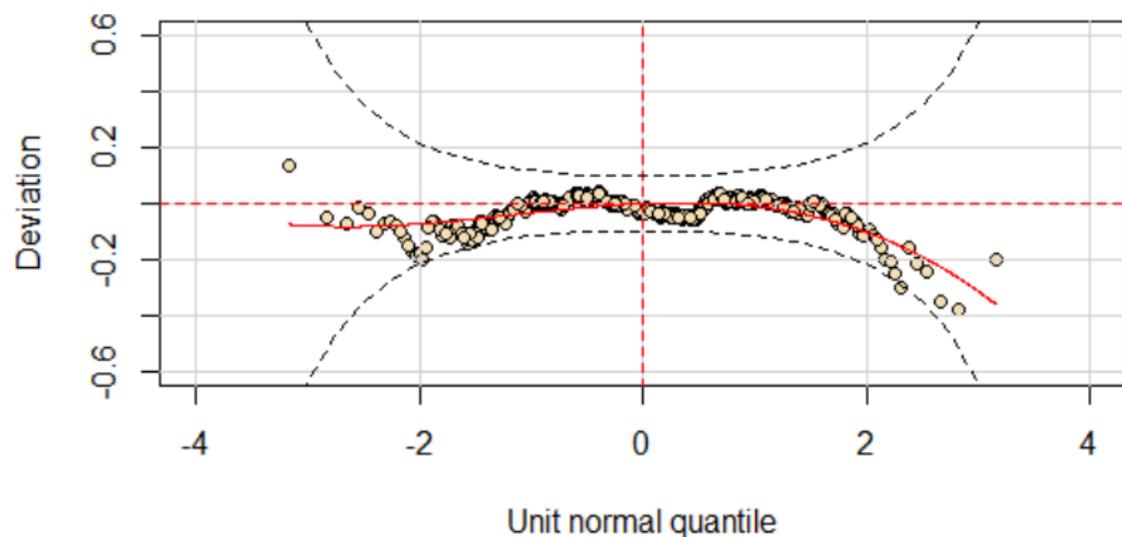
Density Estimate



Normal Q-Q Plot



Modelos GAMLSS



Cantidad de datos eliminados con los distintos criterios

	Extremos 5 %	Variación >50 %	Extremos Costo/Ventas	Residuos regresión	Robusto	GAMLSS
Extremos 5 %	70	37	1	30	7	0
Variacion > 50 %		54	1	17	5	0
Costo/Ventas			2	1	0	0
Residuos regresión				35	4	0
Robusto					16	0
GAMLSS						0

Resultados con distintos criterios de eliminación

Cuadro: Resultados del modelo ABJ con distintos criterios

Modelo ABJ Criterio	Costos de ventas						
	Total	EExt	EVar	ERratio	ERes	Robusto	GLAMMSS
$\hat{\beta}_0$	-0,002	0,015**	0,017**	-0,002	-0,028*	0.005	0.003
$\hat{\beta}_1$	0,99**	0,81**	0,77**	1,00**	0,916**	0.95**	0.95**
$\hat{\beta}_2$	0,11	0,14	0,29**	0,01	0,08	-0.017	-0.003
GA	1	1	1,38	1	1	1	1
R2 ajustado	0.7802	0.6651	0.7054	0.7854	0.8195	0.8592	0.7814
nro obs	633	563	579	631	598	617	633

Resultados con distintos criterios de eliminación

Cuadro: Resultados del modelo ABJ con distintos criterios

Modelo ABJ Criterio	Costos de ventas						
	Total	EExt	EVar	ERatio	ERes	Robusto	GLAMMSS
$\hat{\beta}_0$	-0,002	0,015**	0,017**	-0,002	-0,028*	0.005	0.003
$\hat{\beta}_1$	0,99**	0,81**	0,77**	1,00**	0,916**	0.95**	0.95**
$\hat{\beta}_2$	0,11	0,14	0,29**	0,01	0,08	-0.017	-0.003
GA	1	1	1,38	1	1	1	1
R2 ajustado	0.7802	0.6651	0.7054	0.7854	0.8195	0.8592	0.7814
nro obs	633	563	579	631	598	617	633
Modelo ABJ	Gastos de administración y Comercialización						
$\hat{\beta}_0$	0.05**	0.07**	0.05**	0.05**	0.06**	0.05**	0.05**
$\hat{\beta}_1$	0.44**	0.33**	0.37**	0.49**	0.43**	0,40**	0.41**
$\hat{\beta}_2$	-0.27**	-0.08	-0.14	-0.36**	-0.14	-0,21*	-0.23**
GA	0.39	1	1	0.26	1	0.475	0.439
R2 ajustado	0.2096	0.1372	0.1225	0.2632	0.2372	0.2573	0.2064
nro obs	632	522	564	475	599	612	632

Conclusiones

- Este estudio supone que el comportamiento de los costos no es proporcionalmente simétrico, lo cual es verificado con los modelos propuestos

Conclusiones

- Este estudio supone que el comportamiento de los costos no es proporcionalmente simétrico, lo cual es verificado con los modelos propuestos
- Identificar y medir las asimetrías contribuye a la toma de decisiones, por lo que resulta prioritario la correcta estimación de parámetros por sobre la distribución de errores

Conclusiones

- Este estudio supone que el comportamiento de los costos no es proporcionalmente simétrico, lo cual es verificado con los modelos propuestos
- Identificar y medir las asimetrías contribuye a la toma de decisiones, por lo que resulta prioritario la correcta estimación de parámetros por sobre la distribución de errores
- Este estudio muestra desde un punto de vista metodológico, que la exclusión de outliers puede ser incorrecta si no se tiene en cuenta la distribución de la respuesta.

Conclusiones

- Este estudio supone que el comportamiento de los costos no es proporcionalmente simétrico, lo cual es verificado con los modelos propuestos
- Identificar y medir las asimetrías contribuye a la toma de decisiones, por lo que resulta prioritario la correcta estimación de parámetros por sobre la distribución de errores
- Este estudio muestra desde un punto de vista metodológico, que la exclusión de outliers puede ser incorrecta si no se tiene en cuenta la distribución de la respuesta.
- De los modelos considerados, la regresión robusta y el modelo GAMLSS estiman correctamente los parámetros, no así los otros modelos clásicos con distintos criterios de extracción de outliers.

Conclusiones

- Este estudio supone que el comportamiento de los costos no es proporcionalmente simétrico, lo cual es verificado con los modelos propuestos
- Identificar y medir las asimetrías contribuye a la toma de decisiones, por lo que resulta prioritario la correcta estimación de parámetros por sobre la distribución de errores
- Este estudio muestra desde un punto de vista metodológico, que la exclusión de outliers puede ser incorrecta si no se tiene en cuenta la distribución de la respuesta.
- De los modelos considerados, la regresión robusta y el modelo GAMLSS estiman correctamente los parámetros, no así los otros modelos clásicos con distintos criterios de extracción de outliers.

Cuestiones pendientes

- “En economía, tradicionalmente es prioritario la correcta estimación de parámetros por sobre la distribución de errores” (Margarita Díaz).
- Pero si se considera la verdadera distribución de la respuesta, es posible que no existan outliers, y esto a su vez puede modificar la estimación de los parámetros.
- Dado que GAMLSS ajusta la distribución de la respuesta a una distribución T con muy bajos grados de libertad, existe la cuestión de considerar las posibles razones en este hecho.

Referencias

- Anderson, M., Banker, R., and Janakiraman, S. (2003). Are selling, general and administrative costs sticky? *Journal of Accounting Research*, 41(1):47-63.
- Banker, R. D., Byzalov, D., and Plehn-Dujowich, J. M. (2014b). Demand uncertainty and cost behavior. 26(2):42-79.
- Calleja, K., Stelarios, M., and Thomas, D. (2006). A note on cost stickiness: Some international comparisons. *Management Accounting Research*, 17:127-140.
- Malcom, R. (1991). Overhead control implications of activity costing. *Accounting Horizons*, pages 69-78.
- Mark, Y. and Roush, M. (1994). Flexible budgeting and variance analysis in an activity-based cost environment. *Accounting Horizons*, 8(2):93-103.
- Noreen, E. (1994). Conditions under which activity-based cost systems provide relevant costs. *Journal of Management Accounting Research*, 3:159-168.
- Noreen, E. and Soderstrom, N. (1997). The accuracy of proportional cost models: Evidence from hospital service departments. *Review of Accounting Studies*, 2:89-114.
- Stimolo M.I.y Diaz M. (2017) Comportamiento asimétrico de los costos en empresas argentinas listadas en el mercado de valores de los sectores Construcción, Comercio y Servicios. *Cuad. CIMBAGE 19* : 43-65.
- Stasinopoulos, Rigby, Heller, Voudouris and Bastiani (2017) *Flexible Regression and Smoothing: Using GAMLSS in R*. Chapman and Hall/CRC

Muchas Gracias