



Nº9 Abril 2012

## ALGUNOS MÉTODOS PARA MODELAR TENDENCIAS Y SU APLICACIÓN A LAS SERIES DE EMPLEO SECTORIAL EN PUERTO RICO

*(Different Techniques of Modelling Trend and its applications to Puerto Rico Sectoral Employment Series)*

**Wilfredo Toledo Ph.D.**

[WilfredoToledo@AOL.com](mailto:WilfredoToledo@AOL.com)

**Resumen:** En este artículo se exponen los principales métodos que se utilizan para modelar las tendencias de las series cronológicas. Se consideran métodos para extraer las tendencias deterministas y estocásticas, así como modelos segmentados con las fechas de los quiebres determinadas endógenamente. Se discute, además, la forma de construir intervalos de confiabilidad para las fechas de los cambios estructurales. Dichas técnicas se aplican a las series de empleo total y sectorial de Puerto Rico. Un hallazgo interesante del trabajo fue la identificación de una reducción en la tasa de crecimiento del empleo de manufactura y el empleo total de la isla a partir del segundo lustro de la década de 1990.

**Abstracts:** This paper considers the principal approaches used for modelling trends in economic time series. Deterministic, stochastic, and segmented trends models with dates of break determined endogenously are considered. Also, a review of Quandt/Andrews/Bain method to identify structural changes and construct confidence intervals for the break dates is presented. These techniques are applied to Puerto Rico sectoral employment series.

**Palabras Clave:** Modelos de tendencia, filtro Hodrick-Prescott, descomposición Beveridge-Nelson, cambios estructurales con fechas endógenas, intervalos de confianza para fechas de cambios estructurales.

**Key Words:** Modelling trends, Hodrick-Prescott Filter, Beveridge-Nelson decomposition, structural break and confidence intervals.

**JEL classification codes** C32, C10, E32.

### 1. Introducción

No existe un consenso entre los economistas ni los estadísticos sobre lo que constituye la tendencia general de una serie cronológica y por

consiguiente tampoco existe una sola manera de modelarla. En este artículo se discuten los principales métodos que se utilizan para extraer la tendencia secular de las series de tiempo. Se examinan tanto técnicas deterministas como estocásticas. Además, se discuten los nuevos métodos para identificar cambios estructurales que han afectado la trayectoria de las variables cuando se desconoce la fecha exacta de los mismos y la forma de construir intervalos de confianza para esas fechas.

Las metodologías reseñadas en este trabajo se aplicaron a las series de empleo de los principales sectores económicos de Puerto Rico, a saber: la administración pública, la manufactura, y los servicios. Además, se analiza el empleo total en la isla y en uno de los métodos se añade el sector de comercio. Ese análisis empírico es útil para documentar los cambios sectoriales que están experimentando las economías con costos de mano de obra relativamente altos, que han visto mermado el sector de la manufactura e incrementado los sectores de los servicios y comercio. Para el caso de Puerto Rico parte de esos movimientos intersectoriales del insumo laboral se puede explicar por la proliferación de tratados comerciales, como el NAFTA, y reducción de incentivos contributivos que Estados Unidos le otorga a las empresas de ese país que se ubican en la isla.

## **2. El concepto de tendencia**

Desde el punto de vista estadístico una serie cronológica ( $Y_t$ ) puede ser descompuesta en los siguientes componentes; (i) un elemento estacional (s); (ii) un elemento irregular o aleatorio ( $\epsilon$ ); (iii) un elemento cíclico(c) y; (iv) una tendencia o movimiento de largo plazo ( $f(t)$ ). La serie se puede representar como:

$$Y_t = s + \epsilon + c + f(t)$$

Algunas variables tienen la característica de que las observaciones a través del tiempo manifiestan patrones de comportamiento que se repiten cada (k) períodos. A esas series se les denomina series periódicas. En las variables económicas esos patrones ocurren dentro de un año por lo que se le conocen como series estacionales.

Existen distintas técnicas para remover el elemento periódico de las series de tiempo, los más utilizados son: (i) diferenciación discreta estacional; (ii) uso de variables categóricas asociadas a cada período; y, (iii) el filtro X12 (antes X11) que es utilizado por la Oficina de Análisis Económico del Departamento de Comercio de los Estados Unidos y está incorporado en algunos programas estadísticos computadorizados. Los componentes aleatorios y cíclicos podrían separarse o extraerse de la serie mediante la estimación de algún modelo estadístico.

Por otra parte, para modelar la tendencia de una variable es necesario contar con alguna definición conceptual. Como se mencionó, ese componente de las series no ha sido definido concluyentemente, pero en términos generales se concibe como la parte de una serie que cambia lentamente a través del tiempo. La suavidad es una característica importante en ese concepto. Sin embargo, en el caso de los procesos de caminatas aleatorias la tendencia de la variable es la última observación, por lo que los cambios en ésta no son necesariamente graduales. En ese caso el componente secular podría ser una línea que pase por la última observación de la serie.

Otra forma de definir ese concepto es como aquel componente de la serie que cuando se predice ofrece una indicación de los movimientos de la variable en el largo plazo. Si la tendencia es la parte de la serie que se puede extrapolar, entonces hay que tener una forma de realizar dicha extrapolación.

### **3. Tres métodos para extraer la tendencia general de las series**

La tendencia de largo plazo contiene los cambios que han afectado permanentemente las series cronológicas. Existen varios métodos para identificar ese factor, algunos de los cuales son: (i) polinomios de tendencia (iii) el filtro de Hodrick-Prescott (HP); y la descomposición de Beveridge-Nelson. Esta sección se dedica a la discusión de esas tres técnicas y los resultados de su aplicación a las series de empleo de Puerto Rico para el período de 1970 a 2008.

#### **A. Polinomios de tendencia**

Si la trayectoria del indicador  $Y_t$  es lineal se puede representar como:

$$Y_t = \beta_0 + \beta_1 t \quad t=1,2,\dots,T$$

(1)

Donde: los  $\beta$ 's son parámetros,  $t$  es la variable de tendencia y  $T$  es el número de observaciones.

Por otro lado, cuando el componente secular de la variable no es lineal se puede modelar como:

$$Y_t = \beta_0 + \sum_{j=1}^M \beta_j t^j$$

(2)

El modelo (2) es un polinomio de orden  $M$  en  $t$ , lo que implica que la pendiente de la serie no es fija a través del tiempo. El problema empírico es determinar los valores de los  $\beta$ 's en los dos modelos y el orden del polinomio para (2), deliberación que se puede hacer a base de pruebas estadísticas estándares usadas en la técnica de regresión lineal. Lo que supone que las variables dependientes son aleatorias. En esta investigación se utilizaron las pruebas-t para los parámetros estimados y el coeficiente de determinación para escoger el orden de los polinomios de tendencia para cada serie analizada. Además, se ajustó un modelo lineal a cada una de ellas para tener un punto de comparación.

La Gráfica 1 muestra la tendencia lineal y la ajustada por un polinomio de cuarto grado para el empleo en manufactura. Como se observa, el empleo en ese sector estuvo por encima de su tendencia lineal desde 1987 hasta 2001, mientras que desde 2002 hasta 2008 asumió valores por debajo de la misma. Por otro lado, al considerar la tendencia generada por el polinomio de cuarto grado, que no toma como constante la tasa de crecimiento de largo plazo, se observa que el empleo en ese sector no alcanzó su nivel de largo plazo durante el período del 2002 al 2005. No obstante, superó el mismo en los años de 2006 y 2007.

La Tabla 1 contiene los resultados más relevantes de la estimación de los dos modelos. Un resultado sobresaliente es que aunque la tendencia lineal del empleo en manufactura es positiva, su tasa de crecimiento sufrió

varios cambios en el período bajo análisis y en la década de los años dos mil decreció. Ese último comportamiento se refleja en el parámetro estimado para el término cúbico de la variable de tendencia que resultó ser negativo.

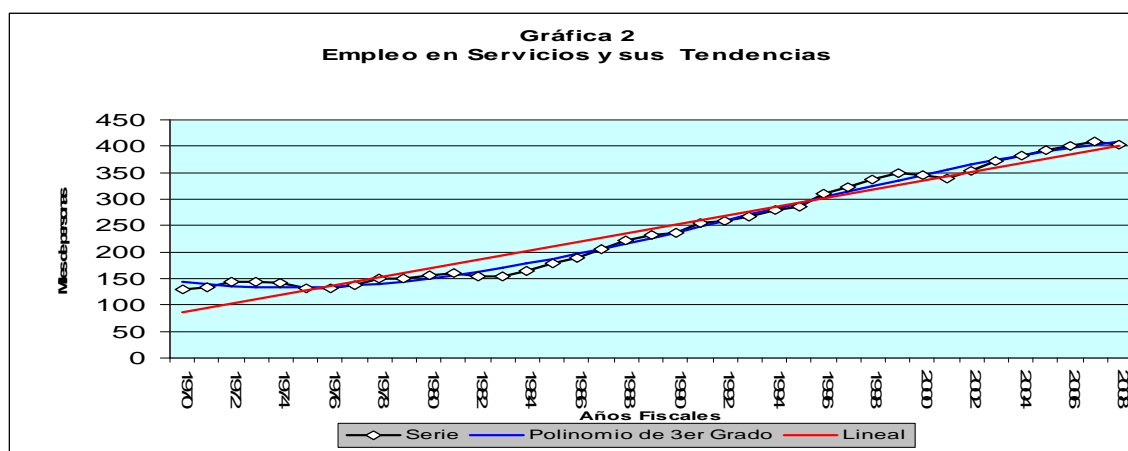


**Tabla 1**  
**Resultados de la estimación de los modelos de tendencia**

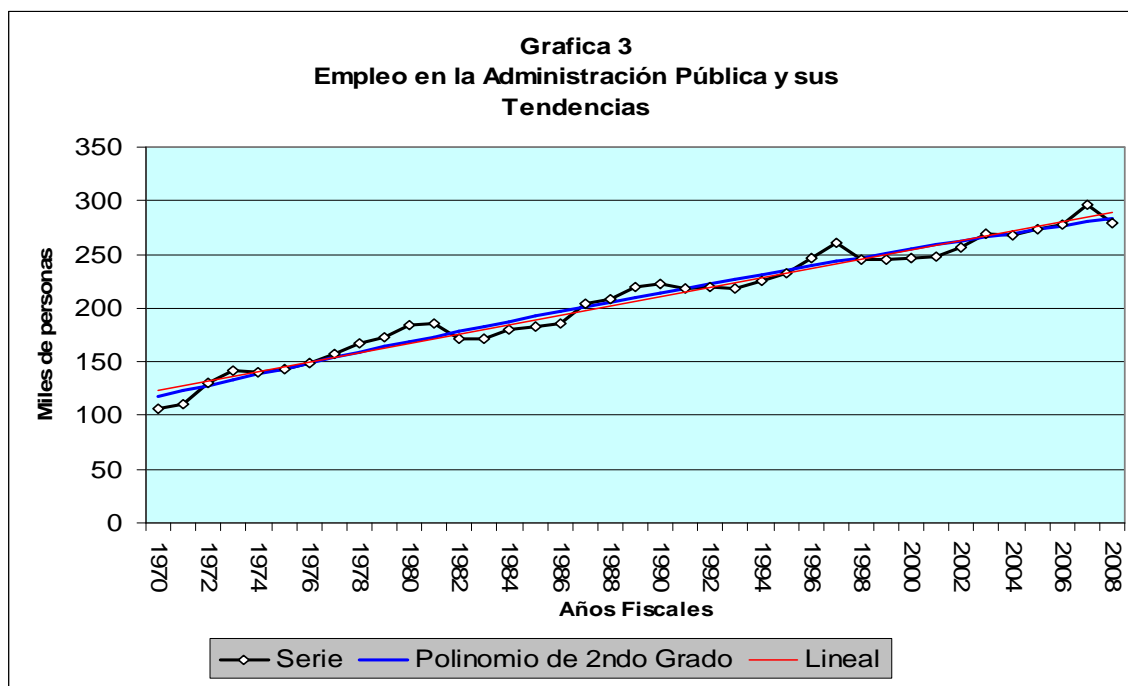
Variable dependiente: Empleo en Manufactura  
Período de estimación: 1970 a 2008

A. Tendencia lineal			
Variable	Coficiente	Desviación estándar	Valor-P
C	141.127	4.317	0.000
t	0.333	0.195	0.096
R <sup>2</sup> =.0731			
B. Polinomio de tendencia			
C	141.017	4.931	0.000
t	-4.389	1.848	0.023
t <sup>2</sup>	0.611	0.202	0.005
t <sup>3</sup>	-0.021	0.008	0.013
t <sup>4</sup>	0.0002	0.0001	0.063
R <sup>2</sup> =.771			

Examinemos ahora, el sector de los servicios. La Gráfica 2 ilustra la estimación de dos modelos de polinomios de tendencia: uno lineal y otro cúbico. Como es evidente en esa gráfica la tendencia general del empleo en servicios de 1970 a 2007 fue positiva. Esa serie estuvo por encima de su crecimiento lineal en la década de los años 2000. Además, al usar el polinomio de tercer grado se encuentra que en esa última década los valores la serie son exactamente iguales a la tendencia.



Los resultados de análisis del empleo público están contenidos en la Gráfica 3 y la Tabla 2. El empleo gubernamental de 1970 hasta 2007 exhibió una tendencia positiva, aunque en el último año su valor no alcanzó la tendencia. Los modelos de tendencia lineal y de orden dos arrojan resultados muy parecidos. Todos los coeficientes estimados en los modelos fueron positivos.



**Tabla 2**  
**Resultados de la estimación de los modelos de tendencia**

Variable dependiente: Empleo en la Administración Pública

Período de estimación: 1970 a 2008

A. Tendencia lineal

Variable	Coefficiente	Desviación estándar	Valor-P
C	123.630	2.782	0.000
t	4.370	0.126	0.000

$R^2=0.970165$

B. Polinomio de tendencia

C	117.265	3.831	0.000
t	5.404	0.466	0.000
$t^2$	-0.027	0.012	0.028

$R^2=0.974$

La estimación de los modelos para el empleo total se resume en la Tabla 3. Esa variable estuvo por encima de su tendencia lineal desde mediados de los años noventa hasta 2006, pero el 2007 asumió un valor por debajo de la misma. El coeficiente estimado para el término de grado tres fue negativo, lo que indica que la tasa de crecimiento de esa serie se ha reducido en los últimos años. Hallazgo que se explica por la recesión que está experimentando la isla desde 2006.

**Tabla 3**  
**Resultados de la estimación de los modelos de tendencia**

Variable dependiente: Empleo Total			
Período de estimación: 1970 a 2008			
A. Tendencia lineal			
Variable	Coefficiente	Desviación estándar	Valor-P
C	604.01	16.248	0.000
t	17.355	0.735	0.000
R <sup>2</sup> =0.937644			
B. Polinomio de tendencia			
C	731.398	17.345	0.000
t	-16.822	4.004	0.000
t <sup>2</sup>	1.948	0.246	0.000
t <sup>3</sup>	-0.030	0.004	0.000
<b>R<sup>2</sup>=0.98</b>			

En esta sección se reseñaron los resultados de la estimación de varios modelos de polinomios de tendencia para algunas series de empleo de Puerto Rico. A base de los hallazgos se puede concluir que la descomposición de una variable en los componentes de largo plazo y cíclico depende del concepto de tendencia que se utilice. Por tanto, determinar si la fluctuación en una variable es un movimiento transitorio o un cambio permanente no es una tarea trivial y no está exenta de la utilización del juicio del investigador. En términos sustantivos vimos, que al igual que otras economías con costos de mano de obra relativamente altos, en Puerto Rico se ha reducido el sector de manufactura y se ha incrementado el de los servicios a juzgar por los datos de empleo examinados. Los datos también documentan la contracción económica que padece ese país desde el 2006.

### B. La tendencia de acuerdo al filtro Hodrick-Prescott

Otra forma de modelar la tendencia de una serie fue propuesta por Hodrick y Prescott (1995), tomando como base el siguiente problema:

$$\text{Min}_{S_t} \sum_1^T (Y_t - S_t)^2 + \lambda \sum_1^T [(S_{t+1} - S_t) - (S_t - S_{t-1})]^2$$

(3)

Donde:  $Y_t$  es la serie bajo análisis;  $S_t$  es la tendencia de la serie en el período  $t$ ;  $T$  es la longitud de la serie; y,  $\lambda$  es un parámetro que controla la suavidad de la tendencia. Hodrick y Prescott recomiendan utilizar  $\lambda$  igual a 100, 1,600 y 14,100 para datos anuales, trimestrales y mensuales, respectivamente.



El primer término de la función objetivo penaliza por valores de la tendencia que se alejen de la serie, mientras que el segundo penaliza por la falta de suavidad en la misma. El resultado constituye una valoración de estos dos criterios para minimizar la penalidad total: el componente de crecimiento secular debe seguir la trayectoria de la variable, pero evitar cambios abruptos. El parámetro  $\lambda$  magnifica la penalidad por falta de suavidad, de forma tal que si el mismo es igual a cero la tendencia será la misma serie y si tiende a infinito la tendencia sería lineal. El filtro Hodrick-Prescott es compatible con la noción de que el ciclo económico son fluctuaciones en la actividad económica alrededor de una tendencia que varía lentamente a través del tiempo.

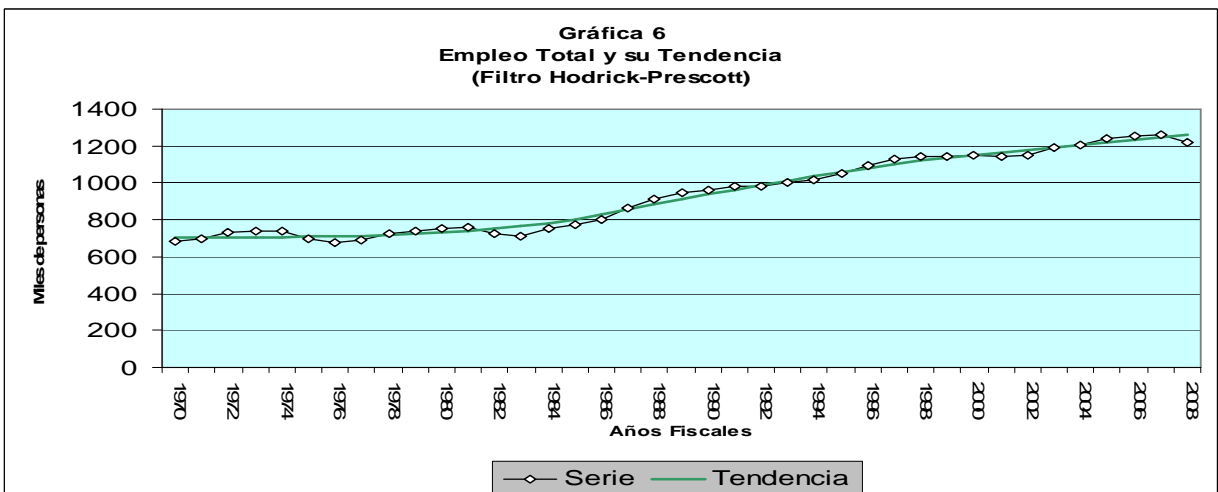
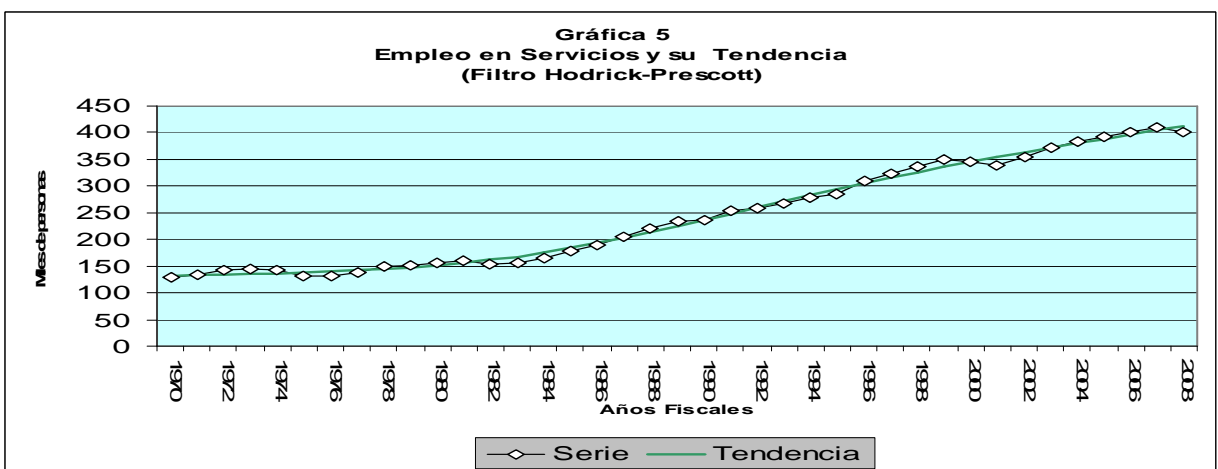
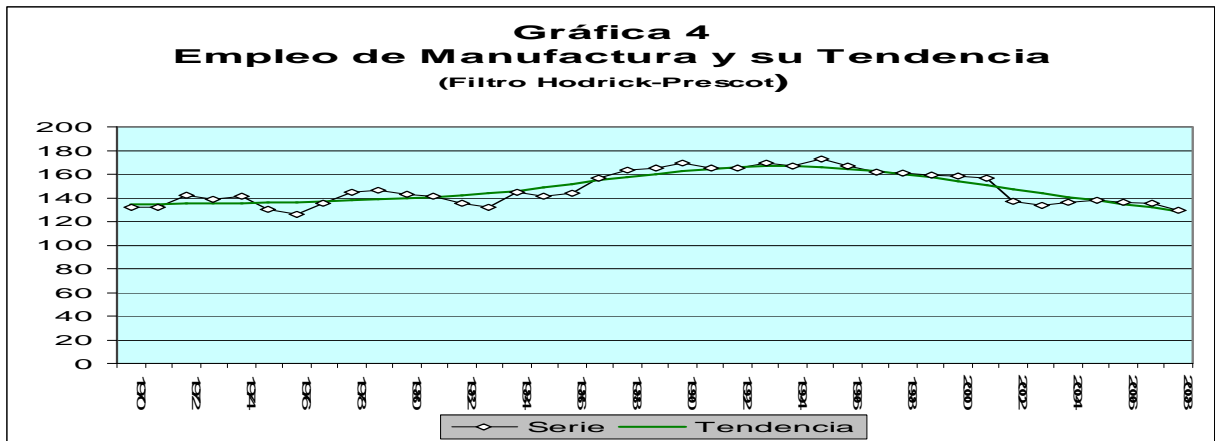
La minimización del problema (5) tiene como resultado el filtro de Hodrick-Prescott que se representa como:

$$\text{Filtro } H - P = [1 + \lambda(1 - L)^2(1 - L^{-1})^2]$$

(4)

Donde L es el operador de rezago. Al aplicar este filtro a una serie ( $Y_t$ ) se obtiene la tendencia de la serie, mientras que la diferencia entre la serie original y la tendencia así obtenida contiene los componentes cíclico y aleatorio, bajo el supuesto de que el elemento estacional ya se ha removido.

Las Gráficas 4 a la 6 presentan las tendencias de acuerdo al filtro H-P para tres de las series de empleo. Ese concepto de tendencia dinámica reproduce mejor el movimiento de las variables a través del tiempo que los polinomios discutidos en la sección anterior. Se observa que para algunas series hay períodos para los cuales no existe diferencia notable entre la tendencia y la serie misma. Por tanto, valores que en los modelos de la sección anterior no estaban muy alejados de la tendencia, parecen formar parte de ésta ahora. Algunos de los movimientos de las variables no son fluctuaciones cíclicas, sino cambios permanentes en sus tendencias. Así que, esta técnica implica que el ciclo económico son cambios en la tendencia secular de algún indicador de la economía y no oscilaciones alrededor de la misma.



### C. La descomposición Beveridge-Nelson (BN)<sup>1</sup>

Beveridge y Nelson (1981) desarrollan un método ingenioso para extraer la tendencia de una serie, partiendo de la descomposición de Wold.

<sup>1</sup> Esta discusión se basa en Mills (2003) páginas 61 a 63.

De acuerdo a esa descomposición, si una serie ( $Y_t$ ) es integrada de grado uno ( $I(1)$ ) su primera diferencia ( $\Delta Y_t$ ) puede representarse como la suma de un elemento determinista y otro estocástico, en símbolos:

$$\Delta Y_t = \mu + \psi(L)\varepsilon_t; \quad \psi_0 = 1$$

(5)

Donde  $\psi(L)$  es un polinomio en el operador de rezago  $L$  ( $LX_t = X_{t-1}$ ) dado por  $(1 + L\psi_1 + \psi_2 L^2 + \dots)$  y  $\varepsilon_t$  es un elemento estocástico con media cero y varianza constante.

Puesto que  $\psi(1) = \sum \psi_i$  es una constante, se puede demostrar que:

$\psi(L) = \psi(1) + \Delta\tilde{\psi}(L)$ , donde  $\Delta\tilde{\psi}(L)$  es un polinomio en el operador de rezagos con coeficientes:  $\psi_j, j=1,2,\dots,\infty$ . Así que,

$$\Delta Y_t = \mu + \psi(1)\varepsilon_t + \Delta\tilde{\psi}(L)\varepsilon_t$$

(6)

La ecuación (6) se puede expresar como la suma de la tendencia ( $TR_t$ ) y el elemento cíclico ( $C_t$ ):

$$\Delta TR_t = \mu + \psi(1)\varepsilon_t$$

(7)

$$C_t = \Delta\tilde{\psi}(L)\varepsilon_t$$

(8)

En esa descomposición la tendencia de la serie es un proceso de caminata aleatoria con un término constante. Las ecuaciones (7) y (8) implican que tanto el elemento cíclico como la tendencia dependen de los elementos aleatorios, por lo que están correlacionados, contrario a las descomposiciones tradicionales que presumen una dicotomía entre estos dos factores.

Una forma de aplicar la descomposición BN es modelando la serie como un proceso ARIMA(p,1,q). En ese caso la ecuación (6) luce como:

$$\Delta Y_t = \mu + \frac{\theta(L)}{\phi(L)}\varepsilon_t$$

(9)

Con:

$$\psi(L) = \frac{\theta(L)}{\phi(L)} \quad \text{y,} \quad \psi(1) = \frac{\theta(1)}{\phi(1)}$$

Donde:  $\psi(1)$  es la medida de persistencia del proceso  $Y_t$ , los  $\theta$ 's son los parámetros asociados al componente MA y los  $\phi$ 's al AR.

La ecuación (9) puede describirse como:

$$\frac{\phi(L)}{\theta(L)} \Delta Y_t = \mu \frac{\phi(1)}{\theta(1)} + \varepsilon_t, \quad \text{ó}$$

$$\frac{\phi(L)}{\theta(L)} \Delta Y_t \psi(1) = \mu + \psi(1)\varepsilon_t$$

(10)

El lado derecho de (10) es  $\Delta TR_t$ , así que en términos de las  $Y$ 's:

$$TR_t = \frac{\phi(L)}{\theta(L)} \psi(1) Y_t$$

y,

$$C_t = Y_t - TR_t$$

En este trabajo se estimó la tendencia BN para el empleo total. De acuerdo a las pruebas de raíces unitarias ADF no se puede rechazar las hipótesis de que esa variable es  $I(1)$  y su primera diferencia estacionaria<sup>2</sup>. Se ajustó el modelo ARIMA(1,1,0):

$$(1-.48)\Delta L_{\text{Total},t} = 6.58 + \varepsilon_t$$

así que:

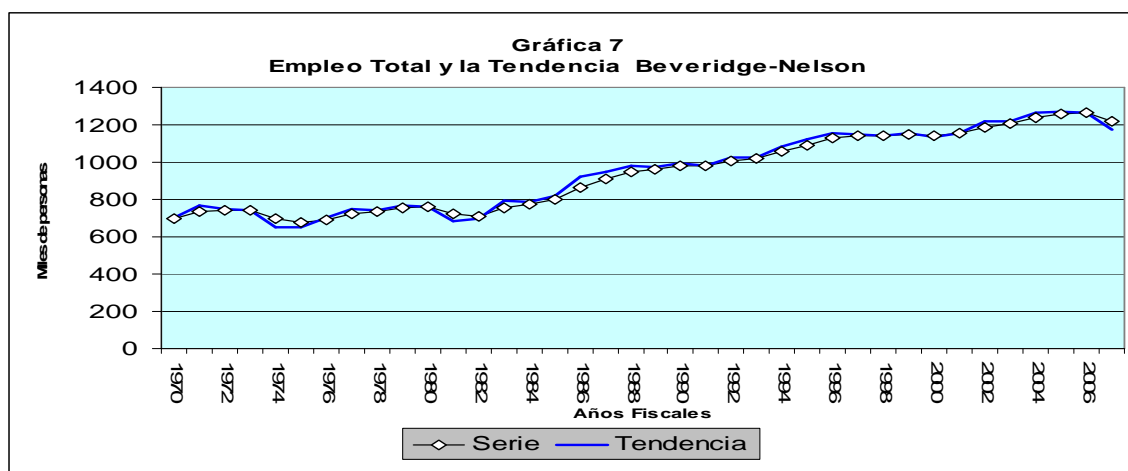
$$\psi(1) = \frac{\theta(1)}{\phi(1)} = \frac{1}{(1-.48)} = 1.92$$

y la tendencia BN es:

$$TR_t = \frac{(1-.48L)}{1} (1.92Y_t) = 1.92Y_t - .92Y_{t-1}$$

En la Gráfica 7 se presenta el empleo total y su tendencia BN. Como vemos, la tendencia es casi igual a la serie, y se asemeja bastante a la que se obtuvo con el filtro H-P, excepto al final de los datos. En términos de las fluctuaciones económicas, este concepto de tendencia implica que el ciclo es un cambio en la tendencia y no movimientos alrededor de la misma, al igual que en el filtro mencionado.

<sup>2</sup> Este asunto se discute en detalle más adelante.



#### 4. Modelos de tendencias segmentadas

Los modelos de tendencias segmentadas parten de una especificación lineal y permiten desplazamientos en los parámetros de la ecuación (interceptos y pendientes) para reflejar cambios estructurales que han afectado la misma. Dichos modelos pueden ser deterministas si se supone que la variable bajo análisis es estacionaria o estocásticos si se asume que la misma es integrada de orden uno. La prueba ADF se utiliza para determinar la naturaleza del proceso que genera la serie (determinista o estocástico), mientras que en ausencia de información exacta sobre la fecha del cambio estructural se pueden usar los criterios estadísticos desarrollados por Andrews (1993) y Quantd (1960) para tales propósitos. Esos dos temas se discuten más adelante.

##### A. Cambios estructurales en series con tendencias deterministas

Un modelo de tendencia segmentada se puede representar como:

$$Y_t = \beta_0 + \beta_1 t + \sum_1^m \beta_{2i} d_{it} + \sum_1^m \beta_{3i} d_{it} t + \varepsilon_t \quad (11)$$

Donde: m indica el número de regímenes, las  $d_{it}$ , para  $i=1,2...m$ , son variables dicótomas que asumen el valor de uno en el régimen i y cero en los otros períodos y  $\varepsilon_t$  es un elemento estocástico.

Para estimar (11) es necesario identificar las fechas de los quiebres en la variable de interés ( $Y_t$ ). La prueba propuesta por Chow (1960) es uno

de los métodos más utilizado para identificar cambios en los parámetros que definen una relación funcional entre variables. Esa prueba se basa en la consideración de un modelo de regresión lineal, que en términos generales se puede representar como:

$$Y_t = X_t \beta + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2) \quad \text{para } t \leq T$$

(12)

donde:  $Y_t$  es la variable dependiente,  $X_t$  es una matriz ( $k \times T$ ) de  $k$  variables independientes que son relevantes para explicar  $Y_t$ ,  $\varepsilon_t$  es un elemento estocástico y  $\beta$  es un vector ( $k \times 1$ ) que contiene los parámetros del modelo.

El modelo (12) implica que la relación entre  $Y$  y las  $X$ 's ha sido constante a través del tiempo. Sin embargo, como el sistema económico está sujeto a impulsos aleatorios continuamente, estos pudieran provocar cambios en la relación entre la variable dependiente y las independientes, por lo que no sería correcto presumir que los  $\beta$ 's son fijos. Por ejemplo, si se supone que la longitud de la serie se puede expresar como,  $T=t_1+t_2$  y que hubo un cambio estructural en  $t_1$  el modelo (12) luciría como:

$$Y_t = X_t \beta_1 + \varepsilon_{t,1} \quad \text{para } t \leq t_1$$

(13)

$$Y_t = X_t \beta_2 + \varepsilon_{t,2} \quad \text{para } t > t_1$$

La Prueba de cambio estructural de Chow consiste en someter a prueba (docimar) la siguiente hipótesis:

$$H_0: \beta_1 = \beta_2$$

$$H_a: \beta_1 \neq \beta_2$$

La hipótesis nula impone la restricción de que los parámetros han sido invariantes en el período bajo análisis y que por lo tanto el modelo (12) es correcto. El estadístico  $F$  nos permite docimar  $H_0$  bajo el supuesto de normalidad de la variable dependiente. Dicho estadístico se calcula como:

$$F_C = \frac{\hat{\varepsilon}'\hat{\varepsilon} - (\hat{\varepsilon}'_1 \hat{\varepsilon}_1 + \hat{\varepsilon}'_2 \hat{\varepsilon}_2) / k}{(\hat{\varepsilon}'_1 \hat{\varepsilon}_1 + \hat{\varepsilon}'_2 \hat{\varepsilon}_2) / T - 2k} \sim F(k, T-2k)$$

El uso de la prueba de Chow requiere que se conozca la fecha del cambio estructural de antemano. Sin embargo, dicho conocimiento en raras ocasiones se tiene por varios motivos. Por ejemplo, cuando el evento exógeno que ocasiona el cambio actúa con rezagos o antes del evento por el efecto anticipación. El investigador tiene dos opciones bajo esa situación: seleccionar una fecha en forma arbitraria; o examinar los datos para determinar el período del cambio en régimen. Como plantea Hansen (2001) en el primer caso la prueba de Chow no sería informativa, mientras que el segundo caso las propiedades estadísticas de la prueba no serían válidas, ya que la fecha de la partición es endógena, y los parámetros estimados estarían sujetos al problema de sesgos asociados a un estimador pre-probado (*pretest estimator*<sup>3</sup>).

Quandt (1960) propone calcular el estadístico de Chow presumiendo cambios en cada período de la muestra y seleccionar aquél con el valor más alto como la fecha del cambio estructural, pero no provee criterios estadísticos para esta selección. Valores críticos para el procedimiento propuesto por Quandt fueron desarrollados por Andrews (1993).

Andrews examina un modelo<sup>4</sup> indexado por los parámetros  $\beta_t$  y plantea las siguientes hipótesis sobre la estabilidad de los parámetros:

$$H_0 : \beta_t = \beta_0 \quad \forall t \geq 1 \quad \text{para } \beta_0 \in B \subset \mathbf{R}^p$$

$$H_A(\pi) : \beta_t = \begin{cases} \beta_1(\pi) & \text{para } t=1, \dots, T\pi \\ \beta_2(\pi) & \text{para } t=T\pi+1 \end{cases} \quad \text{para } B_1(\pi), B_2(\pi) \in B \subset \mathbf{R}^p$$

donde:  $\pi \in (0,1)$ .

Obsérvese que  $\pi$  es una fracción que al ser multiplicada por el indicador de tiempo,  $T$ , segmenta la serie. La hipótesis alterna plantea que los parámetros son distintos en esas particiones por lo que dependen de  $\pi$ . Si la fecha del cambio estructural ( $\pi T$ ) es conocida de antemano la prueba estándar de Chow se puede utilizar para someter a prueba las hipótesis. Sin embargo, cuando se desconoce  $\pi$  la distribución de dicha prueba no es la  $F$  o  $J_i$  cuadrada, debido a que aunque el parámetro  $\pi$  aparece bajo la hipótesis

<sup>3</sup> Ver Judge, Griffiths, Hill, Lütkepohl y Lee (1980) capítulo 3 para una discusión de los estimadores pre-probados.

<sup>4</sup> Este modelo es más general que el descrito por (3), ya que permite cambios estructurales parciales.

alterna no aparece bajo la nula<sup>5</sup>. Si esos estadísticos se calculan considerando  $\pi$  como fija no poseerían las propiedades de muestreo regulares.

Andrews deriva la distribución asintótica para:

$\sup_{\pi \in \Pi} W_T(\pi)$ ,  $\sup_{\pi \in \Pi} LM_T(\pi)$  y  $\sup_{\pi \in \Pi} LR_T(\pi)$ ,  $\pi \in [0, 1]$ , donde: sup es el valor más alto del estadístico calculado en el intervalo de tiempo examinado. Dicho autor presenta una Tabla con los valores críticos (asintóticos) para los tres estadísticos.

### B. Cambios estructurales en series con tendencia estocástica

Dickey y Fuller (1979 y 1981) plantean que algunas series económicas pueden manifestar tendencias estocásticas en lugar de deterministas. Considere un proceso AR(1) que genera la serie  $Y_t$ :

$$Y_t = \mu + \rho Y_{t-1} + \varepsilon_t$$

donde:  $\mu$  es una constante,  $\rho$  es un parámetro y  $\varepsilon$  es un elemento aleatorio.

Ese proceso es una caminata aleatoria cuando  $\rho=1$ , en cuyo caso se dice que el proceso es  $I(1)$  y tiene una tendencia estocástica. Por otro lado, es estacionario ( $I(0)$ ) si y sólo si  $|\rho| < 1$ , en cuyo caso la tendencia sería determinista. En este trabajo se utilizó la prueba desarrollada por Dickey y Fuller (ADF) para determinar si la serie es  $I(1)$  ó  $I(0)$ .

Para identificar los cambios estructurales en las tendencias de variables  $I(1)$  es necesario modelar las primeras diferencias de las series que serían estacionarias y por tanto se podría utilizar los estimadores mínimos cuadrados ordinarios en la estimación. En esta investigación se utiliza la siguiente ecuación para esos propósitos:

$$\Delta y_t = \beta_0 + \phi_1 D_t + \beta_1 t + \phi_2 t D_t + \sum_{j=1}^p \rho_j \Delta y_{t-j} + \varepsilon_t$$

(14)

donde;  $\Delta y_t$  es la primera diferencia del logaritmo de las variables de empleo sectorial; las variables dicótomas de cambio estructural se definen como  $D_t=1$  si  $t > T_B$  y cero de otra forma, siendo  $T_B$  la fecha del cambio en régimen, y las letras

---

<sup>5</sup> Esto es, no se está examinando cambios en los parámetros en una partición de tiempo particular, sino buscando (“pescando”) la partición en que hubo el cambio mayor.



griegas representan parámetros.  $T_B$  se identifica en una forma secuencial, como ya se discutió en la sección anterior. Las hipótesis que se someten a prueba son las siguientes:

$$H_0 : \phi_1 = \phi_2 = 0$$

$$H_A : \phi_1 \neq 0 \text{ y/o } \phi_2 \neq 0$$

Rechazo de la hipótesis nula es evidencia de la existencia de cambio estructural en la serie.

### **C. Construcción de intervalos de confianza para las fechas de los quiebres**

Bai (1997a y 1997b) extiende el trabajo de Andrews en dos direcciones: añade un procedimiento secuencial para identificar múltiples cambios estructurales, y presenta la construcción de intervalos de confianza para la fecha del cambio. En el primer caso el procedimiento consiste en identificar primero la fecha que tiene el estadístico de Chow más alto, dividir la serie en dos períodos en base a esa fecha y repetir el procedimiento para identificar otro cambio. El método finaliza cuando se identifican todos los cambios que la longitud de la serie permita.

Los intervalos de confianza se pueden construir considerando la ecuación dada por (12). Si se define  $\hat{\delta}_j$  como el estimado del cambio en los parámetros del modelo asociado al régimen  $j$ , y  $Z_t$  como el subconjunto de  $X_t$  que se afecta por el cambio estructural (se puede definir  $X_t = [W_t, Z_t]$  donde  $W_t$  son las variables independientes cuyos parámetros se mantienen constantes a través del tiempo) el intervalo de confianza está dado por:

$$[\hat{K} - (c/\hat{L}) - 1, \hat{K} + (c/\hat{L}) + 1] \quad (15)$$

donde:  $\hat{K}$  es la fecha del cambio y  $c$  es la  $(1-\alpha/2)$  cuantíla de la distribución de probabilidad derivada por Andrews.

El estimado de  $\hat{L}$  depende de los supuestos sobre estacionariedad de las variables independientes del modelo y la correlación serial del elemento estocástico. Consideremos dos de los posibles casos<sup>6</sup>:

*Caso 1:  $[\varepsilon_t, Z_t]$  es estacionario de orden dos para toda la muestra y  $\varepsilon_t$  no está auto-correlacionado.*

Para este caso:

$$\hat{L} = \hat{\delta}_j' \hat{Q} \hat{\delta} / \hat{\sigma}^2,$$

$\hat{Q} = T^{-1} \sum_{t=1}^T Z_t Z_t'$ , y  $\hat{\sigma}^2$  es el estimado de máxima verosimilitud de la varianza usando los residuos de de la regresión.

*Caso 2: Variables independientes con tendencia y  $\varepsilon_t$  es estacionario y no está auto-correlacionado*

En este caso.

$$\hat{L} = \hat{\delta}_j' g(k/T) g(k/T)' \hat{\delta} / \hat{\sigma}^2,$$

Obsérvese que en términos generales  $Z_t = [g_1(t/T), g_2(t/T), \dots, g_q(t/T)]'$

#### **D. Aplicación de los métodos a las series de empleo: tendencia determinista**

Esta sección presenta los resultados de la identificación de los cambios estructurales que ha experimentado las series de empleo mencionadas desde 1970 hasta el 2008, usando datos anuales. Se utilizan los niveles de las variables bajo el supuesto que la tendencia es determinista. Para este análisis se incluyó el sector de comercio.

La Tabla 4 contiene los resultados de las pruebas para identificar quiebres en las tendencias de las cuatro series de empleo analizadas, mientras que en la Tabla 5 se reportan los intervalos de confiabilidad construidos. Se identificaron cuatro cambios estructurales que afectaron las series, estos ocurrieron en las décadas de mil novecientos setenta, ochenta,

<sup>6</sup> El intervalo pudiera ser asimétrico para otros casos no considerados aquí.

noventa y a principios de los años dos mil. El primer cambio afectó la tendencia del empleo total y no fue estadísticamente significativo de acuerdo a la distribución de probabilidad desarrollada por Andrews, aunque lo es al usar la distribución F que es aplicable si se hubiera conocido de antemano la fecha del quiebre.

<b>Tabla 4</b>							
<b>Resultados de las pruebas de cambios estructurales para las series de empleo:</b>							
<b>los valores-p y los puntos de cambio</b>							
<b>Empleo</b>	<b>Valores-P de Estadístico LR*</b>						
	<b>1975</b>	<b>1985</b>	<b>1987</b>	<b>1995</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>
<b>Total</b>	<b>6.50</b> <b>(.35)</b>	-	<b>58.09**</b> <b>(.000)</b>	-		<b>9.84</b> <b>(.10)</b>	
<b>Manufactura</b>		-	<b>14.62</b> <b>(.013)</b>	<b>56.65**</b> <b>(.000)</b>			<b>21.65</b> <b>(.000)</b>
<b>Comercio</b>	<b>8.28</b> <b>(.185)</b>		<b>92.23**</b> <b>(.000)</b>				
<b>Servicios</b>	<b>5.62</b> <b>(.464)</b>	<b>110.63**</b> <b>(.000)</b>			<b>11.923</b> <b>(.043)</b>		

\*Usando el procedimiento de *refinement* desarrollado por Bai. \*\*1<sup>er</sup> cambio identificado. Se usó  $\pi$  igual a .15.

La serie del empleo de servicios sufrió un desplazamiento en 1985, mientras que en las otras tres series el estimado puntual para la fecha del cambio en esa década ocurrió en 1987. Por otro lado, sólo la tendencia de crecimiento del empleo en manufactura experimentó un cambio en régimen en la década de los mil novecientos noventa el cual acaeció en 1995. Dicho cambio fue el principal en esa serie de empleo, por lo que fue el primero en identificarse y por tanto el que obtuvo el valor más alto para el estadístico LR. Al examinar la Tabla 5 se puede observar que los límites del intervalo de confiabilidad para ese cambio estructural son los años 1994 y 1996.

<b>Tabla 5</b>				
<b>Intervalos de confiabilidad para las fechas de los cambios estructurales</b>				
<b>Fecha del cambio</b>	<b>Intervalos de confianza de 95%</b>			
	<b>Empleo Total</b>	<b>Manufactura</b>	<b>Servicios</b>	<b>Comercio</b>
<b>1985</b>	-	-	<b>[1984 1986]</b>	<b>[1983 1987]</b>
<b>1987</b>	<b>[1986 1988]</b>	<b>[1985 1989]</b>		
<b>1995</b>	-	<b>[1994 1996]*</b>		
<b>2000</b>	-	-	<b>[1998 2002]</b>	<b>[1995 2005]</b>
<b>2001</b>	<b>[1998 2004]</b>	-		

<b>2002</b>	-	<b>[2000 2004]</b>		
*Para cambio en la pendiente.				

El estimado puntual para el último quiebre en la tendencia de crecimiento de las series se ubica entre los años 2000 al 2002. Los intervalos de confianza revelan que no existe diferencia estadísticamente significativa entre las fechas de ese cambio para las cuatro variables consideradas.

Para conocer la dirección de los cambios estructurales identificados se estimaron modelos de tendencias segmentadas para cada una de las cuatro variables de empleo. Los resultados de esa estimación se informan en la Tabla 6. Se observa que las cuatro series manifiestan tendencias de crecimiento positivas.

El cambio estructural que afectó las series de empleo a mediados de los años mil novecientos ochenta aumentó su tendencia de crecimiento. Ese cambio se puede asociar al período de recuperación de la economía luego de la recesión ocurrida a principios de esa década. Por otro lado, el cambio estructural que afectó el empleo en manufactura 1995, el intervalo de confianza de 95% lo ubica entre 1994 y 1996, tuvo el efecto de reducir su tasa de crecimiento. Es sobresaliente el hecho que de las series de empleo consideradas esa fue la única que sufrió un cambio en su tendencia de crecimiento en la década de los 1990, lo que es evidencia de que el cambio en régimen desfavoreció la producción en la manufactura.

<b>Tabla 6</b>				
<b>Resultados de la estimación de los modelos de tendencias segmentadas</b>				
<b>Variables Independientes</b>	<b>Variables de Empleo (dependientes):</b>			
	<b>Coefficientes (Valores-P de t)</b>			
	<b>Total</b>	<b>Manufactura</b>	<b>Servicios</b>	<b>Comercio</b>
<b>C</b>	<b>693.53</b> <b>(0.000)</b>	<b>134.48</b> <b>(0.000)</b>	<b>130.49</b> <b>(0.00)</b>	<b>133.31</b> <b>(0.000)</b>
<b>t</b>	<b>4.32</b> <b>(0.000)</b>	<b>0.47</b> <b>(0.045)</b>	<b>2.17</b> <b>(0.000)</b>	<b>0.66</b> <b>(0.036)</b>
<b>d85</b>	-	-	<b>-126.42</b> <b>(0.000)</b>	<b>-65.39</b> <b>(0.000)</b>
<b>d85*t</b>	-	-	<b>9.55</b>	<b>5.14</b>

			<b>(0.00)</b>	<b>(0.000)</b>
<b>d87</b>	<b>-169.03</b> <b>(0.000)</b>	-	-	-
<b>d87*t</b>	<b>17.06</b> <b>(0.000)</b>	<b>1.02</b> <b>(0.000)</b>	-	-
<b>d95</b>	-	<b>70.02</b> <b>(0.000)</b>		-
<b>d95*t</b>	-	<b>-2.99</b> <b>(0.000)</b>	-	-
<b>d00</b>	-	-	<b>62.95</b> <b>(.038)</b>	<b>37.07</b> <b>(0.120)</b>
<b>d00*t</b>	-	-	<b>-2.58</b> <b>(0.007)</b>	<b>-1.45</b> <b>(0.051)</b>
<b>d01</b>	-	-	-	-
<b>d01*t</b>	<b>-1.59</b> <b>(0.008)</b>	-	-	-
<b>d02</b>	-	-	-	-
<b>d02*t</b>	-	<b>-0.47</b> <b>(0.004)</b>	-	-
<b>R<sup>2</sup></b>	<b>.99</b>	<b>.91</b>	<b>.99</b>	<b>.99</b>

Por último, se encontró que hubo una reducción en la tendencia de crecimiento de las cuatro series a principios de los años dos mil. Esa reducción se asocia con la desaceleración económica que ha experimentado la economía de Puerto Rico desde 2001 y que se agudizó luego de 2006. Es menester señalar que las pruebas estadísticas utilizadas arrojan evidencia de que esa contracción económica no ha sido una oscilación de la actividad económica de la isla alrededor de su tendencia de crecimiento secular, sino que ha sido una merma en la senda de expansión de la economía.

#### **E. Aplicación de los métodos a las series de empleo: tendencia estocástica**

En esta sección se discute la identificación de los cambios estructurales ante la hipótesis de que las series de empleo son estacionarias luego de ser diferenciadas. Las pruebas se realizaron usando los logaritmos de las variables. Los resultados de la prueba ADF implican que todas las series tienen una raíz unitaria (ver Tabla 7)<sup>7</sup>.

**Tabla 7**

<sup>7</sup> Es importante señalar que esta prueba tiene un poder bajo, por lo que la probabilidad de no rechazar la nula cuando es falsa es alta.

<b>Resultados de las pruebas de raíces unitarias</b>			
	<b>Valores-P Para la Prueba ADF: H<sub>0</sub>: la variable tiene una raíz unitaria</b>		
<b>Logaritmo del empleo</b>	<b>Niveles</b>	<b>Primeras Diferencias</b>	<b>Orden de integración</b>
<b>Total</b>	<b>0.8073</b>	<b>0.0147</b>	I(1)
<b>Manufactura</b>	<b>0.5956</b>	<b>0.0001</b>	I(1)
<b>Servicios</b>	<b>0.6690</b>	<b>0.0057</b>	I(1)
<b>Comercio</b>	<b>0.9448</b>	<b>0.0000</b>	I(1)
* Con dos rezagos. ** no se rechaza la hipótesis de raíz unitaria a un nivel de 1%. Y <sub>t</sub> es la variable examinada.			

Para identificar los cambios estructurales ocurridos en las tendencias de las series se utilizó la ecuación (4) discutida en la sección II. El resultado de la aplicación de las pruebas de cambios estructurales de Andrews-Quandt implicó el no rechazo de la hipótesis nula de que los parámetros asociados a las tendencias de las tasas de crecimiento de las variables de empleo se mantuvieron constantes para el período examinado, bajo la premisa de que las fechas de los posibles cambios eran desconocidas. Por lo que no se realizaron análisis ulteriores.

## 5. Resumen

En este artículo se discutieron y aplicaron varios métodos para modelar la tendencia de una serie. Vimos que ese no es un concepto que esté muy claramente definido en la literatura y que por tanto pudiera ser ambiguo distinguir entre movimientos cíclicos y cambios en la tendencia de largo plazo. Además, se expusieron las nuevas técnicas econométricas desarrolladas por Andrews (1993) y Bai (1997) para identificar cambios estructurales en los parámetros de los modelos de regresión lineal y construir intervalos de confiabilidad para las fechas de los cambios. Esos procedimientos determinan los momentos de los cambios en regimenes con criterios estadísticos en vez de imponerlos exógenamente. El escrito constituye una introducción a esas últimas técnicas que todavía no han sido incorporadas en los textos de econometría o estadísticas.

Los métodos se aplicaron a algunas series de empleo de Puerto Rico. Se encontró, que al igual que otras economías con costos de mano de obra relativamente altos, en Puerto Rico se ha reducido el sector de manufactura y se ha incrementado el de los servicios. Los análisis también documentan la contracción económica que padece ese país desde el 2006.

## **Bibliografía**

Andrews, D. W.K. (1993), "Test for Parameter Instability and Structural Change with Unknown Change Point" *Econometrica*, 61(4): 821-56.

Bai, J. (1997a), "Estimation of a Change Point in Multiple Regression Models", *Review of Economics and Statistics*, 79(4): 551-63.

Bai, J. (1997b), "Estimating Multiple Breaks Once at a Time". *Econometric Theory*, 13(3): 315-52.

Bai, J., y R. L. Lumsdaine y J. H. Stock. (1998), "Testing for and Dating Common Breaks in Multivariate Time Series." *Review of Economic Studies*, 64(3): 395-432.

Banerjee, A., R. L. Lumsdaine y J. Stock. (1992), "Recursive and Sequential Tests of the Unit-Roots and Trend-Break Hypotheses: Theory and International Evidence", *Journal of Business and Economic Statistics*, 10(3):271-87.

Ben-David, D. y D. H. Papell. (1998), "Slowdowns and Meltdowns: Postwar Growth Evidence from 74 Countries." *Review of Economics and Statistics*, 80(4): 561-71.

Beveridge, S. y C.R. Nelson. (1981), "A New approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle" *Journal of Monetary Economics*, 7:151-174.

Chong, T. T. (1995), "Partial Parameter Consistency in a Misspecified Structural Change Model." *Economics Letters*, 49(4) 351-57.

Chow, G. C. (1960), "Tests of Equality between Sets of Coefficients in Two Linear Regressions." *Econometrica*, 28(3):591-605.

Christiano, L. J. (1992), "Searching for a Break in GNP." *Journal of Business and Economic Statistics*, 10(3) 237-50.

Dickey, D. A. y W. A. Fuller. (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association*, 74:366: 427-31.

Dickey, D. A. y W. A. Fuller. (1981), "Likelihood Ratio Statistics for Autoregressive Time Series with Unit Root", *Econometrica*, 49:1057-1072.

Fernandez, D. G. (1997), "Breaking Trends and the Money-Output Correlation." *Review of Economics and Statistics*, 79(4) 674-79.

Hansen, B. E. (2001), "The New Econometrics of Structural Change: Dating Breaks in US Labor Productivity, *Journal of Economic Perspectives*, 15(4):117-128.

Hodrick, R.J. y E. C. Prescott. (1995), "Postwar Us Business Cycles: an empirical investigation" *Journal of Money Credit and Banking and Banking*, 29: 1-16.

Judge, G.G., W.E. Griffiths, R. Carter Hill, H. Lütkepohl y T.-C Lee. (1985), *The Theory and Practice of Econometrics*, segunda edición, New York:Wiley.

Mills, T. C. (2003), *Modelling Trends and Cycles in Economic Time Series*, Hampshire: Palgrave.

Quandt R. (1960), "Test of the Hypothesis that a Linear Regression Obeys Two Separate Regimes", *Journal of the American Statistical Association*, 55: 324-30.