

**INSTITUTO TECNOLÓGICO AUTÓNOMO DE  
MÉXICO**



Análisis de Cointegración entre el PIB, Gasto  
en Salud, Educación y Recaudación Fiscal en  
México

**TESIS**  
QUE PARA OBTENER EL TÍTULO DE  
**LICENCIADO EN MATEMÁTICAS APLICADAS**

**P R E S E N T A**

**OMAR DÍAZ LANDA**

**ASESOR: DR. JUAN JOSÉ FERNÁNDEZ DURÁN**

**MÉXICO, D.F.**

**2020**

Con fundamento en los artículos 21 y 27 de la Ley Federal del Derecho de Autor y como titular de los derechos moral y patrimonial de la obra titulada “Caso Práctico de Modelos de Corrección de Error”, otorgo de manera gratuita y permanente al Instituto Tecnológico Autónomo de México y a la Biblioteca Raúl Baillères Jr., autorización para que fijen la obra en cualquier medio, incluido el electrónico, y la divulguen entre sus usuarios, profesores, estudiantes o terceras personas, sin que pueda percibir por tal divulgación una contraprestación.

Omar Díaz Landa

---

Fecha

---

Firma

# Agradecimientos



# Índice general

<b>Introduction</b>	<b>1</b>
<b>1. Análisis Exploratorio</b>	<b>3</b>
1.1. Introducción a las series . . . . .	3
1.1.1. Producto Interno Bruto . . . . .	3
1.1.2. Recaudación Gubernamental . . . . .	7
1.1.3. Gasto en Salud . . . . .	10
1.1.4. Educación Pública . . . . .	11
1.1.5. Desempleo . . . . .	14
<b>2. Pruebas de Raíces Unitarias</b>	<b>17</b>
2.1. Series de Tiempo . . . . .	17
2.2. Raíz Unitaria: Pruebas . . . . .	23
<b>3. Cointegración y Modelos VEC</b>	<b>35</b>
3.1. El concepto de cointegración: Caso 2 Series. . . . .	35
3.1.1. Prueba de Cointegración . . . . .	37
3.1.2. Modelo de Corrección de Errores . . . . .	40
3.2. El concepto de cointegración: Caso 2+ series (VEC) . . . . .	45
3.2.1. Inferencia Sobre los Coeficientes . . . . .	49
3.2.2. Estimación . . . . .	50
3.2.3. Contrastes sobre el rango de cointegración . . . . .	55
<b>4. Aplicación</b>	<b>59</b>
4.1. Rango de Cointegración . . . . .	59

4.2. Estimación del Modelo . . . . . 63

**Conclusiones 73**

**A. Código 77**

**B. Extracción de las Series 79**

# Introducción

Un gasto público ineficiente y de impacto limitado no puede admitirse en un país que subsiste con tantas carencias sociales, económicas y que enfrenta un escenario de recursos públicos escasos. Es por esta razón que en el presente trabajo se investiga la existencia de una relación estable de largo plazo entre factores que denotan la capacidad productiva de la economía mexicana, tales como el Producto Interno Bruto, la Población Total y la Recaudación Impositiva, con factores que denotan el impacto de las políticas públicas en la sociedad mexicana tales como el Desempleo, Gasto en Salud Pública y el Gasto en Educación Pública. Analizar las relaciones de largo plazo entre los factores anteriormente mencionados busca contribuir con evidencia empírica a la toma de decisiones que posteriormente se verán reflejadas en propuestas de políticas públicas que mejoren la gestión de los recursos y el bienestar de los ciudadanos. Las series se analizan en búsqueda de una relación de cointegración, que en caso de existir, se extenderá dicha relación a un modelo de corrección de error vectorial, mejor conocido como VECM por sus siglas en inglés.

El presente trabajo se estructura como sigue: el capítulo 1 muestra la descripción de las series anteriormente descritas que se emplearán en el análisis, su graficación histórica y una breve interpretación histórica de sus respectivas tendencias. Este análisis es relevante para resaltar la importancia de las series en cuestión y será de utilidad para los siguientes capítulos.

Seguidamente, el capítulo 2 presenta los conceptos básicos de una serie de tiempo, así como la teoría requerida para realizar las pruebas de Dickey-Fuller aumentada, las cuales nos ayudan a identificar la presencia de raíces

unitarias en las series de tiempo, siendo así, el primer paso previo al análisis multivariado.

El capítulo 3 inicia con una breve introducción del concepto de cointegración, en primer lugar, bajo el escenario de cointegración entre dos series de tiempo para que después estos resultados se lleven al caso más complejo, en el cual hay existencia de cointegración entre dos o más series. En segundo lugar, la teoría para la estimación de los estadísticos de la traza y del máximo valor propio que nos permitirán determinar el rango de cointegración, junto con el el proceso de estimación sobre los coeficientes del modelo VECM.

Finalmente, el capítulo 4 presenta el estudio del caso práctico de la estimación del modelo de corrección del error presentado en los capítulos 2 y 3. En primer lugar, determinando el rango de cointegración, seguido de la estimación del modelo VECM, para así terminar con una interpretación de los resultados obtenidos. El trabajo finaliza con una breve sección de conclusiones.



# Capítulo 1

## Análisis Exploratorio

### 1.1. Introducción a las series

Las series con las cuales se tiene intención de trabajar, se obtienen de tres fuentes distintas. La primera de ellas es a través del banco de datos del Banco Mundial, la segunda de ellas es a partir de las publicaciones anuales de la Ley de Ingresos de la Federación y por último del Presupuesto de Egresos de la Federación. Las series tienen una periodicidad anual desde 1991 hasta 2016, ya que se trata el horizonte temporal con mayor cantidad de datos continuos informados para cada una de las series.

El proceso de construcción de cada una de las series se encuentra documentado en el apéndice **Extracción de las series**.

#### 1.1.1. Producto Interno Bruto

El producto interno bruto es una medida macroeconómica que expresa el valor monetario total de la producción de bienes y servicios de demanda final de un país durante un periodo determinado de tiempo, el cual generalmente es medido de manera trimestral o anual. Cuando se habla de bienes y servicios de demanda final significa que no se han considerado aquellos bienes que se han elaborado en el periodo en cuestión y que servirán como materia prima

para la fabricación de otros bienes y servicios, es decir, únicamente se incluyen aquellos producidos en el periodo que, por su propia naturaleza, no se van a integrar a ningún otro proceso de producción, así como aquellos otros bienes que no han logrado integrarse en el proceso productivo al final del ejercicio aunque estaban destinados a ello; a estos últimos productos se les conoce como existencias finales. Esto es con la finalidad de evitar una doble contabilidad. Además es importante recalcar el hecho de que el producto interno bruto no contabiliza los bienes y servicios producidos por el trabajo informal, el cual representa un factor importante en nuestra economía.

Debido a que el PIB es un agregado de numerosos componentes, las unidades de medida son heterogéneas, por lo que para obtener un valor total, es preciso transformarlos a términos homogéneos, lo cual se consigue dando valores monetarios a los distintos bienes y servicios. Sin embargo, esto podría generar distorsiones en las comparaciones intertemporales, ya que si un país decide aumentar en un 10% el nivel general de los precios, el PIB se vería aumentado en la misma cantidad. Para evitar estos inconvenientes, se recurre al PIB en términos reales al que no afectan las modificaciones en los precios pues los productos y servicios se encuentran valorados bajo los precios de un año base usado como referencia. Para obtener el PIB real únicamente se debe dividir el PIB nominal por un índice de precios conocido como el deflactor del PIB. En particular, en este trabajo se empleará el año 2010 como el año base.

Los tres sectores económicos principales que componen al PIB de México son:

- **Sector Primario:** actividad proveniente de la agricultura, ganadería, explotación forestal, caza, pesca y minería.
- **Sector Secundario:** actividad industrial de transformación, también incluye al rubro de la construcción.
- **Sector Terciario:** servicios, incluyendo la producción de energía, comunicaciones y agua.

En México el sector servicios es el mayor componente del PIB, seguido por el sector industrial y después el agrícola.

El producto interno bruto es una medida que resume características importantes de una nación, pues es posible interpretarlo desde diferentes ángulos. Por ejemplo, la competitividad de las empresas, ya que si por un lado el PIB no crece, significa que las empresas mexicanas no están creciendo y que no se está invirtiendo en la creación de nuevas empresas, y por lo tanto, la generación de empleos tampoco crece al ritmo deseado. Por otro lado, un crecimiento del PIB representa una mayor recaudación del gobierno vía impuestos y si el gobierno desea conservar dicha tendencia, deberá fortalecer la inversión directa en empresas, fortalecer las condiciones para que las empresas existentes continúen creciendo y asegurarse de alguna manera del bienestar social.

A continuación se muestran las series de tiempo para el Producto Interno Bruto tanto en su formato nominal (a precios corrientes) como en su formato real (a precios constantes):

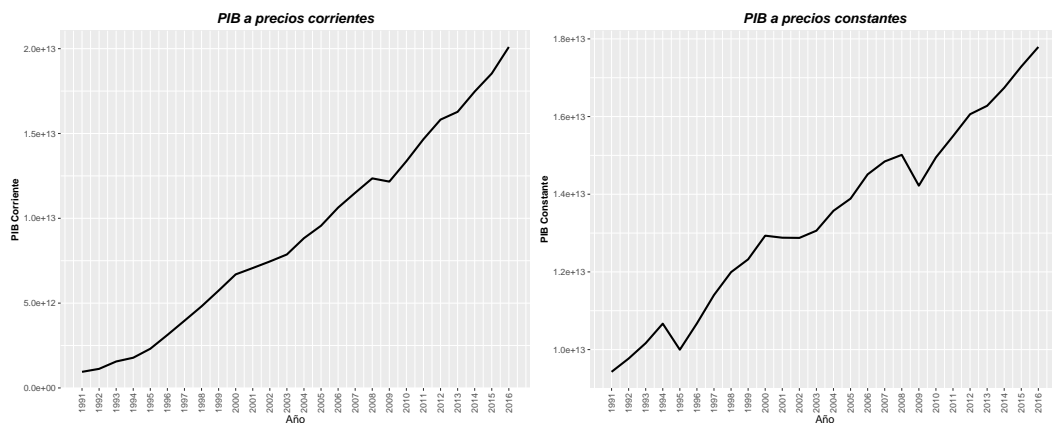


Figura 1.1: PIB a precios corrientes y constantes

Debido a las magnitudes de las series es preferible trabajar con el logaritmo de las mismas, que tal y como se muestra a continuación, se trata de una

transformación monótona por lo que no afecta la tendencia de las series sino únicamente las escalas:

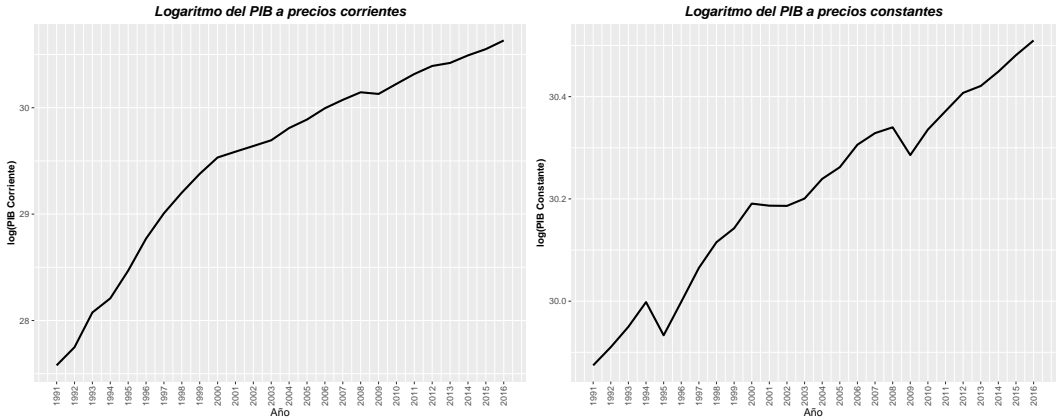


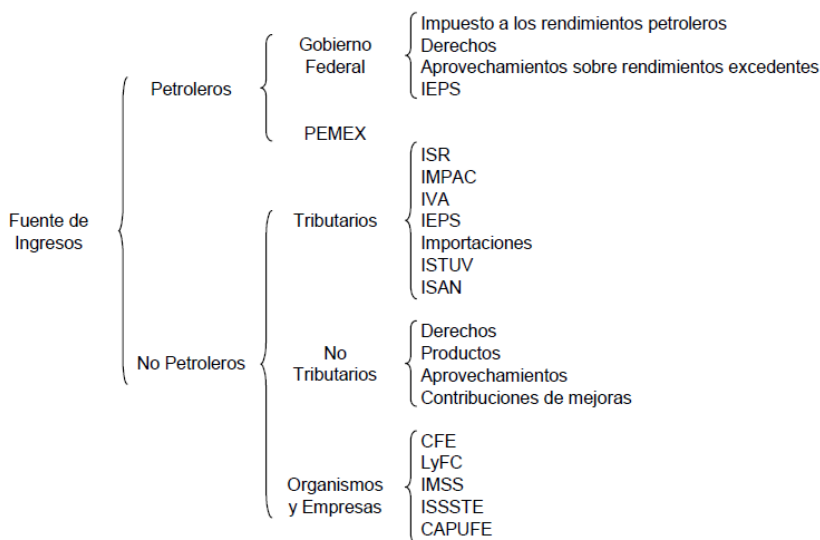
Figura 1.2: Logaritmo Natural del PIB a precios corrientes y constantes

En las gráficas anteriormente mostradas se puede apreciar una tendencia creciente a lo largo del tiempo excepto para los años 1995, 2001 y 2009 en donde se advierte una desaceleración del Producto Interno Bruto justificado por grandes recesiones económicas. En el caso de 1995 la crisis económica conocida como el error de Diciembre de 1994, cuyo costo de la crisis fue trasladado como en la mayoría de los casos a la siguiente administración. En cuanto a los otros decrementos observados, ninguno de los presidentes en turno fue responsable directo de ellos sino que se trató de crisis heredadas. En el 2001 se presenció el ataque terrorista del 11 de Septiembre en Nueva York provocando una caída en la bolsa de valores con un impacto global. Finalmente en el 2009 con el colapso de la burbuja inmobiliaria en los Estados Unidos, se generó una profunda crisis de liquidez, derrumbes bursátiles y en consecuencia una crisis económica a escala internacional. Como es posible observar, el producto interno bruto es sensible a un entorno macroeconómico internacional y es por esta razón que resulta de interés cuantificar el impacto de estos crecimientos o decrementos en aspectos básicos como el bienestar social, así como la productividad del gobierno en términos de recaudación fiscal.

### 1.1.2. Recaudación Gubernamental

La recaudación Gubernamental como su nombre lo indica, es el capital que una nación recibe de acuerdo al sistema tributario que rige en un determinado momento del tiempo. En la actualidad, la debilidad estructural del sistema tributario ha generado que la insuficiencia de los recursos sea uno de los problemas más importantes de la hacienda pública federal, más aún, si se considera que los requerimientos presupuestales de los ejercicios fiscales van en aumento año con año. Esto nos obliga a replantear y generar nuevas políticas de tributación y recaudación que permitan garantizar el bienestar social y una infraestructura sustentable para el desarrollo de la nación.

En nuestro país los grandes rubros de fuentes de ingresos provienen principalmente del petróleo, impuestos no tributarios y organismos y empresas, tal y como se muestra a continuación:



En materia de recaudación, los ingresos tributarios durante 1995 - 2000 crecieron un punto porcentual del PIB como consecuencia de las acciones

enfocadas a la reactivación económica; tres años más tarde parecieron mantenerse constantes debido a las modificaciones fiscales a la Ley del Impuesto Sobre la Renta (ISR) enfocada a disminuir los privilegios fiscales, sin embargo, durante el 2004 - 2006, disminuyeron su participación al registrar 9.3 % del PIB en promedio, derivado de la tendencia a disminuir la tasa impositiva del ISR. El nivel de recaudación fiscal respecto al PIB de México comparado con otros países de la OCDE es muy bajo, incluso comparado con naciones latinoamericanas con economías más pequeñas. De hecho, con relación al producto interno bruto, se ha mantenido históricamente por debajo del 12%. Lo anteriormente mencionado, es un síntoma de que la labor en la política fiscal no ha logrado encontrar un equilibrio entre lo que se recauda y lo que se gasta; es decir, no se ha tenido una programación eficaz de los ingresos y gastos públicos.

Hemos de recordar que la política tributaria es una de las principales herramientas con las que cuenta el estado para cubrir sus funciones relativas al gasto público, tales como son la educación, salud, seguridad, infraestructura, etc. De tal manera que para lograr un sistema tributario eficiente, es de vital importancia establecer estímulos fiscales temporales y efectivos al ahorro, ampliar la base de contribuyentes, fortalecimiento de la misma recaudación, combate a la ilegalidad fiscal, seguridad y certeza jurídica para la autoridad y los contribuyentes. Sin embargo, la historia nos ha demostrado que no se ha logrado llevar a cabo de la manera adecuada ya que en los últimos años el gobierno federal ha iniciado muchos proyectos sociales, que requieren recursos adicionales y ha buscado formas de obtenerlos lo antes posible. Es por esta razón que han sido muy cambiantes las políticas tributarias, tal y como sucedió con la reforma tributaria del 2014, introduciendo algunos aspectos como:

- Incremento de tasas impositivas
- Establecimiento de nuevos gravámenes
- Eliminación de deducciones y regímenes especiales
- Aumento de las facultades de fiscalización

- Restricciones para la aplicación de tratados para evitar la doble imposición

A continuación se muestra la serie histórica de la recaudación impositiva desde 1991 hasta el 2016. Como es posible observar, los efectos de la reforma tributaria se hacen presentes en los años posteriores al 2014. Es relevante destacar que este dato no se encuentra disponible con esta profundidad histórica en ningún sitio oficial y es por esta razón que se ha obtenido la serie histórica del producto interno bruto en el banco de datos del banco mundial y la serie histórica de la recaudación impositiva se ha obtenido a través de las distintas publicaciones anuales de la Ley Federal de Ingresos de la Nación. De esta manera es posible representar a la recaudación impositiva como porcentaje del PIB y con tal profundidad histórica.

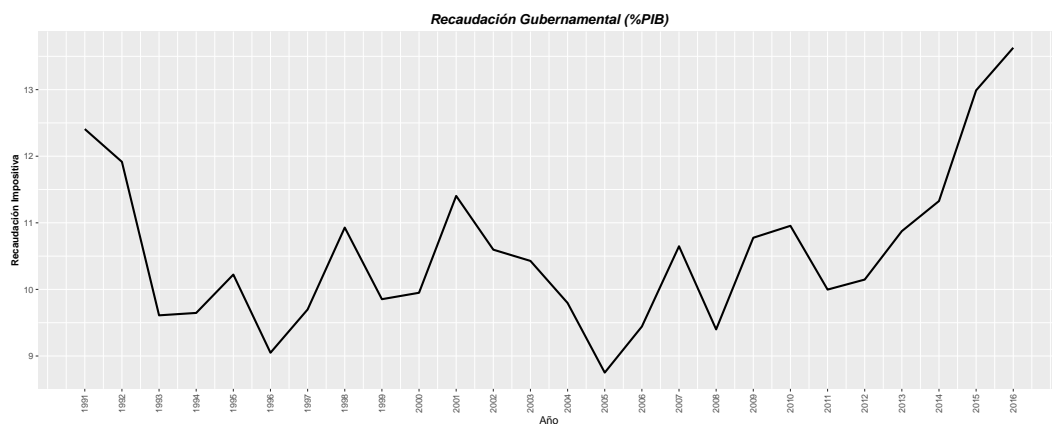


Figura 1.3: Recaudación Impositiva

Por lo tanto, sería de esperarse que cuando el gobierno incrementa las tasas impositivas y aplica una serie de reformas que incrementen la recaudación nacional, esto se vea reflejado en servicios de calidad en términos de salud, seguridad, infraestructura e inversión para la creación de nuevas empresas que incentiven el crecimiento económico de la nación.

### 1.1.3. Gasto en Salud

El Gasto en salud, así como el gasto en educación, infraestructura, gasto federalizado y gasto social son los resultados más tangibles de la administración del gasto público, por lo que sería de esperarse que un gasto público de calidad promueva el crecimiento económico, la equidad de oportunidades y por lo tanto, el desarrollo del país. Durante las últimas décadas, la política gubernamental ha procurado que el financiamiento público para la atención del sector salud del país mantenga una tendencia creciente tal y como puede observarse en la siguiente gráfica:



Figura 1.4: Gasto en Salud Pública ( %PIB)

Dicho incremento significativo se debe principalmente por el origen del Sistema de Protección Social en Salud, mejor conocido como Seguro Popular, el cual pretende representar una ayuda para que las personas en situación de pobreza no incurran en un gasto mayor, en términos de salud, que ponga en riesgo a familias completas. Esta iniciativa busca cerrar la brecha que existe entre la población que cuenta con seguridad social y la que carece de la misma. Sin embargo, lo que el paso de los años nos ha demostrado es que el constante incremento en el gasto en salud no se está viendo reflejado en proporciones similares en resultados que mejoren las condiciones de salud de la población mexicana. De hecho, México ofrece servicios de salud por debajo



de los estándares de la OCDE y tiene un coeficiente de efectividad por debajo de la media de los países del continente americano e incluso, por debajo de países similares como Brasil, Chile o Colombia [18].

Debido a que el Seguro Popular no ha cambiado significativamente el panorama general del sistema de salud y que la manera en que estos servicios se financian es a través de la recaudación de los ingresos presupuestarios del Gobierno, resulta de interés analizar el uso ineficiente del financiamiento público, el cual provoca la pérdida de recursos que no llegan a la atención médica, a partir de la relación que guarda con otro tipo de variables como la misma recaudación, el producto interno bruto, entre otras. Una de las principales razones por la que no se observan mejoras en estos servicios es porque no se cuantifica el impacto de las decisiones de un gobierno.

### 1.1.4. Educación Pública

Al igual que el gasto en salud, el gasto en educación pública representa uno de los factores centrales para la promoción del desarrollo de un país y en México, como se puede apreciar en la siguiente gráfica, se ha manifestado con un gasto creciente durante los últimos 20 años.

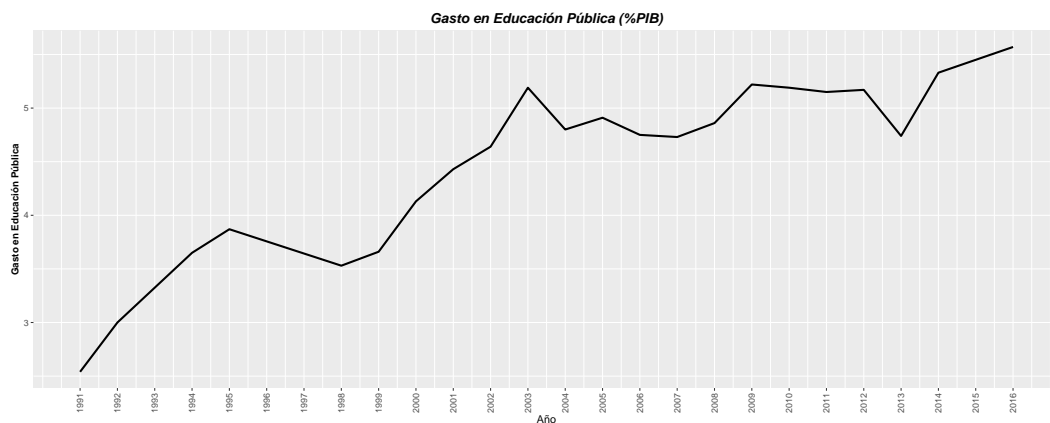


Figura 1.5: Gasto en Educación Pública (%PIB)

Si bien la educación representa el mayor gasto del gobierno federal, dicho gasto no necesariamente implica una mejora en la calidad educativa, ni en el aprovechamiento escolar en la misma proporción que la inversión. Esto ha quedado manifestado en múltiples estudios como el publicado por México Evalúa en el 2011 y el publicado por el Centro de Investigación Económica y Presupuestaria en el 2016 titulado "Gasto Público para una Educación de Calidad", en los cuales se exhibe que México es el país de la OCDE que destina mayores recursos a la educación y que en proporción del PIB, el gasto representa un porcentaje apenas por encima del promedio de la OCDE, además de ser uno de los países que menos recursos destina a cada alumno. Esto significa, que en términos absolutos el esfuerzo en el gasto público educativo podría ser considerado como elevado, sin embargo, en términos relativos se requiere de una mayor inversión y una mejor planificación, ya que aproximadamente el 98 % del financiamiento educativo actual se dirige a gasto corriente, es decir, sueldos y salarios.

Actualmente, ni el aparato burocrático ni las escuelas están sujetas a un proceso de evaluación, medición ni rendición de cuentas sobre las funciones que ejercen, por lo que es menester desarrollar un sistema de información que permita una correcta medición de la eficiencia de los servicios en donde se destina la mayor parte de los recursos, ya que esto permitiría en primer lugar, un esquema transparente en el cual es posible identificar cada uno de los recursos; en segundo lugar, que dado el esquema y los indicadores de eficiencia se justifiquen los recursos destinados y en último, que permita una mejor planificación y toma de decisiones en el sector educativo, incentivando la competencia por los recursos limitados existentes. No obstante, mientras se desarrolla un sistema de información que haga esto posible, continuaremos con un 7 % de la población de 15 años y más que son analfabetas de acuerdo con el último censo de Población (2010), con un México que se ubica en el último lugar de los países de la OCDE en el número de personas que terminaron sus estudios universitarios durante el 2016 y que en el mismo año 2016, la tasa de empleo fue del 65 % para los jóvenes de 25 a 64 años con una educación secundaria, mientras que la tasa de empleo aumenta hasta el 80 % para aquellos que cuentan con licenciatura o equivalente y sube hasta el 85 %

para maestría, doctorado o títulos equivalentes.

Por lo tanto, el panorama actual alrededor de la educación en México se sitúa en una inversión elevada en la educación, con pocos rendimientos en términos de eficiencia, ya que la mayor parte del presupuesto queda destinada al gasto corriente, mismo que no se evalúa con calidad y que, en consecuencia, los que reciben el impacto total de esta serie de desarticulaciones son los alumnos mismos, los cuales carecen de oportunidades para recibir una mejor educación y que al mismo tiempo viven en un México en donde los estudios son aún un factor determinante para la obtención de un empleo y para la distribución de los salarios. Por esta razón, es relevante estudiar cómo es que se relaciona el crecimiento del país en términos de Producto Interno Bruto, con el gasto público en educación con la tasa de desempleo. En efecto, sería de esperarse que estos factores se encuentren relacionados entre sí, sin embargo, dado el contexto del país que hemos visto podría no necesariamente ser explícita la relación, ya que a pesar de presentar un crecimiento a nivel nación que a su vez se refleje en un aumento de inversión a la educación, si las condiciones de evaluación de eficiencia no mejoran, los resultados seguirían siendo pobres.

### 1.1.5. Desempleo

Analizar la tasa de desempleo en México en conjunto con las series descritas anteriormente resulta ser interesante, ya que se buscará determinar una perspectiva del impacto que el comportamiento de la economía mexicana en términos de recaudación, crecimiento del PIB, educación y salud tienen sobre ella. A continuación se muestra la serie histórica de la tasa de desempleo:

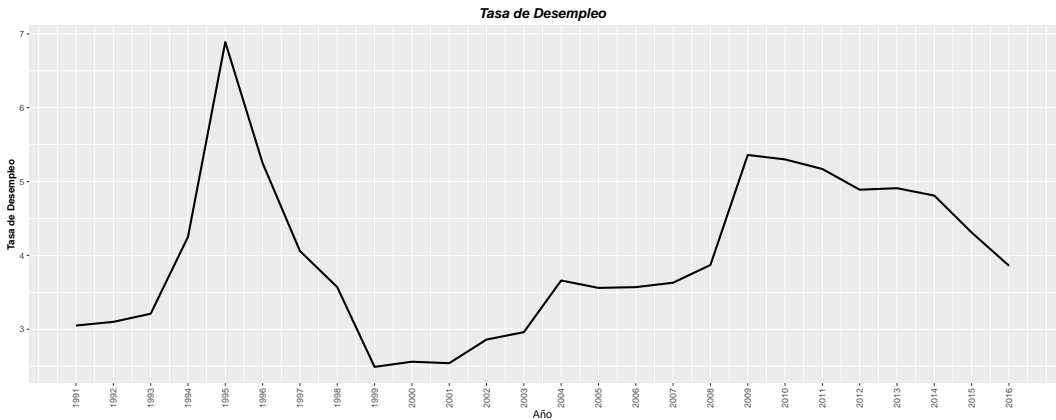


Figura 1.6: Tasa de Desempleo

Para analizar dicha gráfica es pertinente definir cómo está construida la tasa de desempleo. Para que una persona sea considerada desempleada se deben cumplir dos características, en primer lugar, que la persona, en la semana de referencia, no trabajó y en segundo lugar, que la persona en la semana de referencia intentó conseguir un empleo; de tal forma que la tasa de desempleo se calcula dividiendo el número de desempleados entre la Población Económicamente Activa (PEA), la cual está constituida por la suma de las personas ocupadas y las personas desempleadas. Esta manera de calcular el desempleo ha sido muy controversial, ya que tiende a subestimar el valor real, pues existen una infinidad de problemas laborales además del desempleo como el subempleo (personas que sí trabajan y tiene la necesidad de trabajar más horas) y el desempleo encubierto (personas que no trabajan ni buscan

empleo) que no forman parte de la PEA, pero que son personas que están disponibles para trabajar. Si bien hay varias propuestas para definir la tasa de desempleo, como la propuesta por Loría, E. y Ramos, M. (2007)  $Desempleo = ((PEA - PO)/PEA)$  en el presente documento se utilizará la tasa de desempleo oficial obtenida a través del Banco Mundial.

En la figura 1.6 se muestra la evolución histórica de la tasa de desempleo, en donde se puede observar la presencia de las crisis de 1994 y la crisis de liquidez mundial del 2008 que hemos tratado en secciones anteriores, de hecho la serie muestra una tendencia a la alza desde 1999 hasta el punto de quiebre alcanzado en el 2009, seguido de un ligero decremento en la tasa de los últimos años aunque con dificultad para alcanzar los niveles observados durante el periodo 2000-2008. Dicha situación ha llevado a cuestionar las políticas y los planteamientos de política laboral en el país, ya que de acuerdo con Arthur Okun (Okun, 1962, p.2) la desocupación tiene enormes costos sociales y económicos intertemporales, en virtud de que provoca significativos efectos depresivos de largo alcance que se auto-reproducen, constituyendo así un círculo vicioso dinámico. Por esta razón y por una escasa literatura económica que aborde la relación temporal del desempleo con otros factores macroeconómicos, se pretende comprender el fenómeno del desempleo en México a través de la magnitud en que un shock de comportamiento de la misma variable junto con los otros factores macroeconómicos descritos anteriormente, afectan la dinámica del desempleo en el futuro.

Debido al cambio estructural que presenta la serie de tiempo de la tasa de desempleo en el horizonte temporal compartido con el resto de las series y con la finalidad de evitar que dicho cambio estructural impacte en el proceso de estimación del modelo VECM, se ha decidido descartar dicha serie del análisis y continuar con el resto de ellas.



# Capítulo 2

## Pruebas de Raíces Unitarias

### 2.1. Series de Tiempo

Una serie de tiempo, como su nombre lo indica, consiste en la medición u observación sistemática de un fenómeno de interés a lo largo del tiempo, con la particularidad de que dicha información debe ser de naturaleza numérica y registrada en intervalos de tiempo fijos. Esto permite observar la evolución histórica de la información y a partir de ello, llevar a cabo la planeación, la toma de decisiones y en la mayoría de los casos, realizar un pronóstico sobre cuál será el siguiente valor a observar. Es por esta razón que en la actualidad el análisis de series de tiempo ha cobrado interés en áreas de investigación como la economía, las finanzas, la biología e incluso en la astronomía.

Debido a que las series de tiempo suponen un ordenamiento temporal de la información de manera implícita, no es posible utilizar las técnicas estadísticas usuales para su modelación pues en primer lugar, éstas suponen que no hay presencia de información temporal y en segundo lugar, no es de nuestro interés encontrar un conjunto de variables que nos ayuden a hacer inferencia sobre la variable de interés, sino que únicamente se pretende observar la evolución histórica de la serie para poder generar información de valor, hacer análisis y posteriormente realizar pronósticos para la toma de decisiones. Por lo tanto, al tener un problema que no se puede resolver con

metodologías usuales, se abre una nueva rama de investigación conocida como el análisis de series de tiempo, la cual a lo largo de la historia ha encontrado dos corrientes principales para su desarrollo. El primero de los métodos es conocido como **descomposición de series**, el cual supone que una serie de tiempo esta compuesta por elementos como tendencia o ciclo, estacionalidad e irregularidad de tal manera que busca identificar y estimar cada uno de estos componentes por separado. Realizar el análisis de series de tiempo bajo este enfoque tiene sus desventajas pues la modelación e interpretación es de mayor dificultad. A causa de esto surge la segunda línea de trabajo, la cual emplea métodos estadísticos bajo el **enfoque de procesos estocásticos**. Este método es el más común en la actualidad debido a que ha demostrado mayor eficacia en la creación de modelos, pues a partir de una sola familia de ellos se pueden representar un gran número de fenómenos de distinta naturaleza. Además, dicho enfoque presenta flexibilidad para extenderse a modelos para varias series simultáneas.

A lo largo de este capítulo se mostrarán los conceptos básicos que darán apertura al tema principal del presente trabajo.

Para poder analizar series de tiempo desde un punto de vista de procesos estocásticos, es necesario definir en primera instancia lo que se entiende por uno. Se dice que es un **proceso estocástico** cuando se tiene un conjunto índice de números reales  $T$  y una familia de variables aleatorias  $Z$  definidas en un espacio de probabilidad, de tal manera que cada uno de los elementos del conjunto índice tenga asociada una y solo una variable aleatoria, es decir,  $\{Z_t : t \in T\}$ . Por lo tanto, si ahora recordamos la definición anteriormente enunciada de una serie de tiempo es posible equiparar sus elementos con los de un proceso estocástico, pues una serie de tiempo es la medición sistemática en un periodo de tiempo fijo (conjunto índice) de un fenómeno de interés (variable aleatoria) a lo largo del tiempo. Por consiguiente, una serie de tiempo puede ser considerada como un proceso estocástico cuyo conjunto índice es el tiempo.

Es importante aclarar que cuando se tiene una serie de tiempo observada,



es producto de varias realizaciones de una variable aleatoria a lo largo de todo el conjunto índice; es por esta razón que una serie observada representa únicamente una ejecución de todas las posibles realizaciones que el proceso estocástico pudo haber generado. Ante la presencia de este efecto probabilístico, es natural pensar que a partir de la función de densidad conjunta de las variables que conforman el proceso estocástico, es posible describir su comportamiento en su totalidad. Sin embargo, en el caso de series de tiempo obtener dicha función es muy complejo y a diferencia de la mayoría de los desarrollos estadísticos, no es posible suponer que se puede obtener a partir del producto de funciones de densidad marginales, ya que las variables en este caso son altamente dependientes, en realidad, se trata de la misma variable pero medida en diferentes momentos de tiempo. Por consiguiente, el análisis de series de tiempo propone conocer los primeros y segundos momentos, es decir, esperanzas, varianzas y covarianzas para poder caracterizar la serie, ya que estos estadísticos descriptivos resumen en buena medida a la función de densidad conjunta.

Se debe tomar en consideración que cuando se desea analizar la evolución de algún fenómeno o realizar un pronóstico sobre los posibles valores futuros a observar, la serie debe cumplir con la característica de no mostrar tendencias, ciclos o cambios en la varianza pues estos efectos podrían alterar directamente los resultados. Aunque en la practica la mayoría de las series presentan ya sea tendencia o cambios en la varianza, existen técnicas para hacer que la tendencia desaparezca y que la varianza permanezca constante a lo largo del tiempo con la finalidad de hacer posible un pronóstico de mejor calidad. En términos matemáticos, cuando una serie cumple con estas condiciones, se conoce como estacionariedad estricta y estacionariedad débil. Un **proceso estocástico estrictamente estacionario** es un proceso estocástico cuya distribución de densidad conjunta es invariable ante desplazamientos en el tiempo, de manera que si se calculan medias, varianzas y covarianzas, se obtendrán exactamente los mismos resultados sin importar la longitud y momento de tiempo en que se midan, es decir, las medias y varianzas son constantes en el tiempo y las covarianzas solo dependen de la distancia entre los dos periodos de tiempos en que se calcula. Sin embargo,

como es complicado conocer la función de densidad conjunta, se utilizan los primeros y segundos momentos para determinar estacionariedad, a esto se le conoce como **estacionariedad de segundo orden** o **estacionariedad débil**. En resumen, si  $Z_t$  es un proceso estacionario tiene las siguientes características:

$$\begin{aligned}
 \text{Media :} & \quad E(Z_t) = \mu \\
 \text{Varianza :} & \quad Var(Z_t) = E[(Z_t - \mu)^2] = \gamma_0 \\
 \text{Covarianza} & \quad E[(Z_t - \mu)(Z_{t+k} - \mu)] = \gamma_k
 \end{aligned} \tag{2.1}$$

El concepto de estacionariedad es de suma importancia, ya que los modelos econométricos requieren de series estacionarias para poder analizar el comportamiento de la serie en los periodos observados, generalizarlo a otros periodos de tiempo y ajustar modelos para el pronóstico.

Existen diversas pruebas para determinar estacionariedad, pero en la práctica, las más comunes son el **análisis gráfico**, la **función de autocorrelación** y la **varianza muestral**. La primera de ellas es poco confiable, ya que consiste, como su nombre lo indica, en graficar la serie y observar que no muestre tendencias, cambios de nivel o que la variabilidad aumente con el tiempo, por lo tanto, aunque esta prueba depende totalmente de la habilidad de la persona para identificar de manera efectiva estos elementos, sirve para darnos una idea del comportamiento de la serie y se debe considerar como un paso inicial para un análisis más formal como el cálculo de la función de autocorrelación, cuya utilidad se describirá más adelante, y el cálculo de la varianza muestral, las cual nos ayuda a evitar sobrediferenciar una serie, pues sobrediferenciar una serie acarrea problemas en la identificación de algún modelo para representarla, elimina observaciones de manera innecesaria e incrementa la varianza.

Como mencionamos anteriormente, se deben calcular las medias, varianzas y covarianzas para caracterizar, aunque no completamente, a una serie de tiempo, sin embargo, al solo tener una realización con observaciones

finitas del proceso, se debe trabajar con las definiciones muestrales:

$$\begin{aligned}
 \text{Media muestral:} \quad & \bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t \\
 \text{Varianza muestral:} \quad & \hat{\gamma}_0 = \frac{1}{N} \sum_{t=1}^N (Z_t - \bar{Z})^2 \\
 \text{Covarianza muestral:} \quad & \hat{\gamma}_k = \frac{1}{N} \sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}) \quad k = 0, 1, 2, \dots
 \end{aligned} \tag{2.2}$$

A partir de estos términos es posible obtener la función de autocorrelación muestral definida como:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^N (Z_t - \bar{Z})^2} \tag{2.3}$$

Si la función de autocorrelación tiende rápidamente a cero, nos enfrentamos ante un proceso estacionario; si por el contrario, el decaimiento de la función es paulatino, quiere decir que las observaciones del proceso muestran una tendencia que las hace depender de otras observaciones en el tiempo, provocando que el nivel de la serie se vea afectado por dicha tendencia y existan correlaciones distintas a cero en periodos de tiempo distantes. Esto último nos hace pensar que la serie no cumple con alguna de las características de un proceso estacionario de segundo orden.

Cuando una serie es no estacionaria, no es posible hacer uso de los modelos ARMA y cualquier otro tipo de análisis, como por ejemplo una regresión, se debe realizar con particular precaución, ya que se podría estar en presencia de un fenómeno que Yule en 1926<sup>1</sup> identifica como regresión espuria. Sin embargo, existen herramientas que ayudan a convertir una serie no estacionaria en estacionaria, dichas herramientas son las diferencias estacionarias y diferencias estacionales.

---

<sup>1</sup>Yule, G. U., "Why Do We Sometimes Get Nonsense Correlations Between Time Series-? A Study in Sampling and the nature of Time Series", Journal of the Royal Statistical Society, vol. 89, 1926, pp.1-64.

Sea el proceso  $Z_t$  tal que al aplicar el **operador diferencia**  $\nabla$  se tiene que  $\nabla Z_t = Z_t - Z_{t-1}$ . De manera que  $\nabla Z_t$  corresponde al cambio en el valor de la variable  $Z$  del periodo  $t - 1$  al periodo  $t$ , por lo que cuenta con una observación menos que  $Z_t$ . Al aplicar el operador la serie obtenida es estacionaria en cuanto a su nivel, pues elimina la tendencia polinomial adaptiva que se encontraba en el proceso.

Si la tendencia polinomial adaptiva es de orden mayor a uno se debe considerar aplicar el operador tantas veces como sea necesario, es decir, si existe presencia de una tendencia polinomial de orden 2, la serie requiere de tomar dos diferencias para estacionarizarla. Además, es importante mencionar el riesgo que existe con la sobrediferenciación pues conlleva a problemas relacionados con la identificación de un modelo para el proceso, incremento en la varianza y pérdida innecesaria de observaciones, pues se pierde una con cada aplicación del operador.

En ocasiones, nos enfrentamos a series que presentan otro tipo de tendencia que se caracteriza por tener un patrón de cambio que se repite con una intensidad similar, en determinados periodos de tiempo año con año. Este fenómeno es llamado **estacionalidad** y para controlarlo es necesario, en primera instancia, identificar el comportamiento periódico de la serie, ya sea que se trate de meses, trimestres, semestres, años, periodos vacacionales o cualquier otro hecho que provoque que la serie tenga un comportamiento parecido en periodos definidos cada año. Posteriormente, se hace uso del **operador diferencia estacional**  $\nabla_E^k$  donde  $E$  es periodo que contiene la tendencia-ciclo y  $k$  es el número de veces necesarias que se debe aplicar el operador diferencia para hacer la serie estacionaria.

El operador diferencia estacional está definido de la siguiente manera:

$$\nabla_E Z_t = Z_t - Z_{t-E} \quad (2.4)$$

Al igual que en el operador diferencia  $\nabla$  se pierden observaciones cada vez que se aplica, en este caso se trata de  $E$  observaciones.

Cuando a un proceso es necesario aplicarle el operador diferencia  $\nabla$  una vez para hacerlo estacionario, se dice que es un proceso integrado de orden 1 y se denota  $Z_t \sim I(1)$ . De manera similar, si se requieren 2 diferencias para estacionarizar la serie se trata de un proceso integrado de orden dos y si necesitamos  $d$  diferencias es un proceso integrado de orden  $d$ ,  $Z_t \sim I(d)$ , finalmente, si  $z_t$  es estacionario nos podemos referir a este proceso como un proceso estacionario o como un proceso integrado de orden cero. En general, el orden de integración depende del mínimo número de diferencias que se debe aplicar al proceso para convertirlo en estacionario. Este concepto de procesos estocásticos integrados trae consigo las siguientes propiedades:

Sean  $Z_t$ ,  $Y_t$  y  $W_t$  tres procesos estocásticos.

1. Si  $Z_t \sim I(0)$  y  $Y_t \sim I(1)$  entonces  $W_t = (Z_t + Y_t) \sim I(1)$  es decir, una combinación lineal entre un proceso estacionario y uno no estacionario es no estacionario.
2. Si  $Z_t \sim I(d)$  entonces  $W_t = (a + bZ_t) \sim I(d)$ , donde  $a$  y  $b$  son constantes. Una combinación lineal de un proceso no estacionario, es no estacionario. Por lo tanto, la combinación lineal de un proceso estacionario continúa siendo estacionario.
3. Si  $Z_t \sim I(d_1)$  y  $Y_t \sim I(d_2)$  entonces  $W_t = (aZ_t + bY_t) \sim I(d_2)$  ya que  $d_1 < d_2$ .
4. Si  $Z_t \sim I(d)$  y  $Y_t \sim I(d)$  entonces  $W_t = (aZ_t + bY_t) \sim I(d^*)$  donde  $d^*$  generalmente es igual a  $d$ , aunque en ocasiones  $d^* < d$ .

Como es posible observar, trabajar con varias series que no tienen el mismo orden de integración puede resultar complicado. Sin embargo, en el siguiente capítulo se presentarán técnicas apropiadas para reducir estos riesgos.

## 2.2. Raíz Unitaria: Pruebas

Consideremos el siguiente proceso determinista:

$$Z_t = a_0 + a_1 Z_{t-1} \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (2.5)$$

que también puede ser expresado a partir del uso del operador  $L$  que proviene del inglés *Lag* y se define como  $LZ_t = Z_{t-1}$ , por lo tanto  $Z_t = a_0 + a_1 Z_{t-1}$  ahora es

$$(1 - a_1 L)Z_t = a_0, \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (2.6)$$

si se despeja a  $Z_t$  del proceso tendríamos la siguiente operación, que al desarrollarla se sigue que

$$\begin{aligned} (1 - a_1 L)^{-1} a_0 &= \sum_{i=0}^{\infty} (a_1 L)^i a_0 \\ &= a_0 \sum_{i=0}^{\infty} (a_1 L)^i \quad L^i 1 = 1 \\ &= a_0 \sum_{i=0}^{\infty} (a_1)^i \\ &= \frac{a_0}{1 - a_1} \quad |a_1| < 1 \end{aligned} \quad (2.7)$$

En la expresión anterior, el supuesto  $|a_1| < 1$  permite la convergencia de la serie geométrica real, sin embargo,  $a_1$  puede tomar cualquier valor excepto  $a_1 = 1$ . Por lo tanto, la solución general para este proceso sería de la forma

$$Z_t = \frac{a_0}{1 - a_1} + s a_1^t \quad (2.8)$$

donde el término  $s a_1^t$  se agrega para poder brindar soluciones particulares a partir de la constante  $s$ . Ahora bien, si por un lado conservamos el supuesto de la expresión 2.7, es decir, que la raíz de la ecuación  $(1 - a_1 L) = 0$  cumpla que  $|a_1| < 1$  podemos observar que

$$\lim_{t \rightarrow \infty} Z_t = \frac{a_0}{1 - a_1} \quad (2.9)$$

lo cual indica que el proceso determinista alcanzará un punto de equilibrio en el largo plazo. Por otro lado, si  $a_1 = 1$  entonces el proceso diverge, ya que el cociente  $\frac{a_0}{1 - a_1}$  quedaría indeterminado y si  $|a_1| > 1$  el proceso no converge y por lo tanto jamás tenderá a estabilizarse.

Ahora que se entiende la problemática que existe cuando se tienen raíces unitarias en procesos deterministas, es oportuno llevar estos resultados a procesos estocásticos, de manera que si definimos a la constante  $a_1 = 1$ , la constante  $a_0 = (1 - L)Z_0$  y agregamos una variable aleatoria que depende del tiempo  $e_t^2$  con media cero, varianza constante  $\sigma_e^2$  y no correlacionada, tendríamos al siguiente proceso:

$$Z_t = Z_{t-1} + a_0 + e_t \quad (2.10)$$

que a su vez puede ser reexpresada como

$$\begin{aligned} (1 - L)Z_t &= a_0 + e_t \\ (1 - L)Z_t &= (1 - L)Z_0 + e_t \\ (1 - L)\tilde{Z}_t &= e_t \quad \text{con } \tilde{Z}_t = Z_t - Z_0 \end{aligned} \quad (2.11)$$

Este proceso contiene el valor obtenido en el periodo pasado más un error en cada una de las realizaciones, de manera que cada realización de la variable aleatoria  $e_t$  en cada momento  $t$  proporciona una tendencia estocástica que provoca que el proceso tenga fluctuaciones, por lo tanto, no es correcto decir que el proceso alcance un punto de equilibrio, sino que se tendrá que recurrir al concepto de estacionariedad ya conocido. En este particular caso, las fluctuaciones implicarían que si tomamos medias muestrales en diferentes intervalos de tiempo, obtendríamos diferentes resultados, característica que nos indica que la serie es no estacionaria. Para comprobar lo anteriormente mencionado, calculamos los dos primeros momentos del proceso después de haber inferido el comportamiento de la serie con el método iterativo.

$$E(\tilde{Z}_t) = E(e_1 + e_2 + \dots + e_t) = 0 \quad (2.12)$$

$$Var(\tilde{Z}_t) = Var(e_1 + e_2 + \dots + e_t) = t\sigma_e^2$$

---

<sup>2</sup> Mejor conocida como ruido blanco

Por lo que podemos observar que no cumple con los supuestos que requiere la estacionariedad de segundo orden, pues la varianza aumenta conforme se avanza en el tiempo causando que el proceso no regrese a su nivel esperado. Por el contrario, si el valor absoluto de la constante  $a_1$  es menor que la unidad, entonces el proceso es estacionario y se demuestra calculando nuevamente los momentos de primer y segundo orden, para facilitar la operación reexpresamos al proceso usando la serie geométrica de la siguiente manera:

$$\tilde{Z}_t = (1 - a_1 L)^{-1} e_t = e_t + a_1 e_{t-1} + a_1^2 e_{t-2} + a_1^3 e_{t-3} + \dots \quad (2.13)$$

Por lo tanto, la media y la varianza del proceso son

$$E(\tilde{Z}_t) = E(e_t) + a_1 E(e_{t-1}) + a_1^2 E(e_{t-2}) + \dots = 0 \quad (2.14)$$

$$Var(\tilde{Z}_t) = \sigma_e^2 (1 + a_1^2 + a_1^4 + \dots) = \frac{\sigma_e^2}{(1 - a_1^2)}$$

Que demuestran estacionariedad de segundo orden ya que ninguna depende del tiempo. Por lo anterior, es de nuestro interés hacer pruebas sobre el valor de  $a_1$  para determinar si es estadísticamente igual a la unidad, por lo tanto, en la prueba de hipótesis correspondiente, se tendría como hipótesis nula el hecho de que  $a_1 = 1$  que implicaría no estacionariedad en el proceso, mientras que la hipótesis alternativa abarca los posibles valores de  $a_1$  que proporcionan un proceso estable:

$$H_0 : a_1 = 1 \quad H_a : a_1 < 1 \quad (2.15)$$

No obstante es posible reparametrizar al proceso sustrayendo en cada lado de la ecuación el valor  $Z_{t-1}$ , con la finalidad de que la prueba sea similar a la prueba sobre coeficientes individuales de regresión



$$\begin{aligned}
Z_t - Z_{t-1} &= a_1 Z_{t-1} - Z_{t-1} + e_t \\
\nabla Z_t &= (a_1 - 1) Z_{t-1} + e_t \\
\nabla Z_t &= \rho Z_{t-1} + e_t
\end{aligned} \tag{2.16}$$

donde  $\rho = (a_1 - 1)$  de manera que las nuevas hipótesis nula y alternativa son:

$$H_0 : \rho = 0 \quad H_a : \rho < 0 \tag{2.17}$$

Así que  $\rho = 0$  implica que  $a_1 = 1$  tratándose de un proceso no estacionario, mientras que si se rechaza la hipótesis nula tendremos un proceso estable, la parte  $|\rho| > 1$  no se considera pues la serie sería explosiva. Esta última prueba es la llamada **Prueba de Dickey-Fuller** y es complemento formal a los métodos ya mencionados, como la gráfica de los valores de la serie respecto al tiempo, el uso del autocorrelograma y la varianza muestral para determinar estacionariedad, así como el número de diferencias necesarias para convertir la serie en estacionaria.

Para hacer uso de la prueba, es imprescindible conocer el valor del coeficiente  $\rho$  pero como se dijo anteriormente, la prueba es en esencia la misma que para los coeficientes de una regresión, así que se estima este valor por medio de mínimos cuadrados ordinarios. Si  $\rho > 0$  implica que  $a_1 > 1$  por lo que se puede concluir rápidamente que la serie es no estacionaria, no obstante, si  $\rho \leq 0$  se tiene que observar si es estadísticamente diferente de cero, así que se calcula el estadístico de prueba

$$\tau_0 = \frac{\hat{\rho}}{se(\hat{\rho})} \tag{2.18}$$

Si el proceso es estacionario, se puede continuar con el análisis usual de las pruebas sobre coeficientes individuales de una regresión, sin embargo,

cuando el proceso no lo es, se tiene la particularidad de que no se distribuye como una  $t - Student$ , ya que si se aceptara la hipótesis nula,  $Z_t$  sería un proceso no estacionario cuya varianza aumenta conforme se incrementa el tamaño de muestra, provocando que la distribución usual  $t - Student$  se modifique. La contribución de Dickey y Fuller consistió en la creación de los valores críticos del estadístico  $\tau$  para tres diferentes procesos a partir de simulaciones de Monte Carlo. Dichos procesos corresponden en primer lugar, al mostrado en la ecuación 2.16.

En segundo lugar, al proceso

$$\nabla Z_t = \alpha + \rho Z_{t-1} + e_t \quad (2.19)$$

donde cada realización de la variable  $\nabla Z_t$  depende del valor de una constante, más una proporción del valor anterior  $Z_{t-1}$ , más el error  $e_t$ . Estos procesos comúnmente presentan tendencias definidas hacia arriba si el valor de  $\alpha$  es positivo y tendencias definidas decrecientes si es negativo, es decir, tienen una tendencia determinística, ya que se le agrega un valor constante  $\alpha$  en cada instante de tiempo  $t$ . En este caso, si rechazamos la hipótesis nula, es posible hablar de un proceso estacionario pero con media distinta de cero.

Finalmente, tenemos al proceso que incluye constante y una tendencia en el tiempo

$$\nabla Z_t = \alpha + \beta t + \rho Z_{t-1} + e_t \quad (2.20)$$

como es posible observar, el nuevo término proporciona impacto a la tendencia en cada instante de tiempo, si la prueba de hipótesis refleja que el coeficiente  $\rho \neq 0$  entonces se trata de un proceso estacionario alrededor de una tendencia determinística. Para los tres casos se usa la misma prueba de hipótesis y el mismo estadístico de prueba pero diferentes valores críticos que dependen del tamaño de la muestra, tal y como se muestra en la siguiente tabla. El estadístico  $\tau$  debe ser más negativo que el valor crítico correspondiente para rechazar la hipótesis nula ( $\tau \leq \tau^*$ ).

Tabla 2.1: Valores Críticos para la prueba Dickey-Fuller

<i>Modelo</i>	1 %	5 %	10 %
$\nabla Z_t = \rho Z_{t-1} + e_t$	-2.56	-1.94	-1.62
$\nabla Z_t = \alpha + \rho Z_{t-1} + e_t$	-3.43	-2.86	-2.57
$\nabla Z_t = \alpha + \beta t + \rho Z_{t-1} + e_t$	-3.96	-3.41	-3.13
Valores críticos estándar	-2.33	-1.65	-1.28

Adicionalmente, Dickey and Fuller crearon tres  $F$ -estadísticos adicionales llamados  $\phi_1, \phi_2, \phi_3$  para probar pruebas de hipótesis conjuntas:

- **El estadístico  $\phi_1$**  .- Emplea la ecuación numero 2 para probar la hipótesis nula conjunta  $\rho = \alpha = 0$
- **El estadístico  $\phi_2$**  .- Emplea la ecuación numero 3 para probar la hipótesis nula conjunta  $\rho = \alpha = \beta = 0$
- **El estadístico  $\phi_3$**  .- Emplea la ecuación numero 3 para probar la hipótesis nula conjunta  $\rho = \beta = 0$

En donde el término  $\alpha$  corresponde al intercepto y el término  $\beta$  corresponde al coeficiente asociado a la tendencia lineal temporal, tal y como se muestra en la tabla 2.1.

En donde los tres estadísticos se construyen igual que cualquier prueba  $F$ :

$$\phi_i = \frac{[RSS_{restricted} - RSS_{unrestricted}]r}{RSS_{unrestricted}(T - k)}$$

donde  $RSS$  son las sumas al cuadrado de los residuales del modelo restringido vs el no restringido,  $r$  es el numero de restricciones,  $T$  el número de observaciones y  $k$  el número de parámetros estimados en el modelo sin restricciones. Por lo tanto, la hipótesis nula es que los datos fueron generados por el modelo con restricciones y la alternativa es que los datos fueron generados por el modelo sin restricciones. Si la restricción no se cumple, la suma de los cuadrados de ambos modelos deberían ser muy similares y por lo

tanto  $\phi_i$  será pequeña. Así pues, si el valor de  $\phi$  es grande querrá decir que se rechaza la hipótesis nula.

Tiempo después, los supuestos de la prueba de Dickey-Fuller se vieron relajados permitiendo que existiera correlación entre los errores, la cual puede surgir cuando el modelo no captura la dinámica completa del proceso, es decir, le faltan elementos para explicar a la variable. Lo anterior se corrige agregando un mayor número de diferencias de primer orden, tantas como sean necesarias para asegurar que la correlación en los residuos desaparezca, estas se pueden determinar con el gráfico de las autocorrelaciones considerando todas aquellas que sean grandes y descartando de manera formal las que no son necesarias con la significancia de los coeficientes  $\beta_i$ , obteniendo así el siguiente proceso general:

$$\nabla Z_t = \alpha + \beta_1 t + \rho Z_{t-1} + \sum_{i=1}^m \beta_i \nabla Z_{t-i} + e_t \quad (2.21)$$

Análogamente a la prueba no aumentada, se debe procurar no incluir o excluir el término constante y a la tendencia determinística cuando no es necesario, ya que podría traer problemas de sesgo al estadístico  $\tau$  y pérdida de significancia de la prueba si se eligen erróneamente los valores críticos, es posible hacer uso de los mismos valores mostrados en el cuadro 2.1, ya que la prueba de Dickey-Fuller aumentada sigue la misma distribución asintótica.

Resultados de la prueba de orden de integración para las series del logaritmo del PIB a precios constantes, Recaudación Gubernamental (%PIB), Gasto en Salud (%PIB) y Gasto en Educación (%PIB) se muestran a continuación:

Tabla 2.2: Orden de Integración PIB a precios constantes

$\nabla Z_t = \alpha + \rho Z_{t-1} + \beta_1 t + e_t$		Valores Críticos	¿Rechaza $H_0$ ?
Estadístico	$H_0$		
$\phi_3 = 3.89$	$(\alpha, \rho, \beta_1) = (\alpha, 0, 0)$	1 % = 9.31	No
		5 % = 6.73	
		10 % = 5.61	
$\tau_3 = -2.7442$	$\rho = 0$	1 % = -4.15	No
		5 % = -3.50	
		10 % = -3.18	
$\nabla Z_t = \alpha + \rho Z_{t-1} + e_t$		Valores Críticos	¿Rechaza $H_0$ ?
Estadístico	$H_0$		
$\phi_1 = 8.656$	$(\alpha, \rho) = (0, 0)$	1 % = 7.06	Sí
		5 % = 4.86	
		10 % = 3.94	
$\tau_2 = -0.8775$	$\rho = 0$	1 % = -3.58	No
		5 % = -2.93	
		10 % = -2.60	
$\nabla Z'_t = \rho Z'_{t-1} + e_t$ con $(Z'_t = \nabla Z_t)$		Valores Críticos	¿Rechaza $H_0$ ?
Estadístico	$H_0$		
$\tau_1 = -3.4739$	$\rho = 0$	1 % = -2.66	Sí
		5 % = -1.95	
		10 % = -1.6	

Tabla 2.3: Orden de Integración Recaudación Gubernamental (%PIB)

$\nabla Z_t = \alpha + \rho Z_{t-1} + \beta_1 t + e_t$ lags=3			
Estadístico	$H_0$	Valores Críticos	¿Rechaza $H_0$ ?
$\phi_3 = 1.1716$	$(\alpha, \rho, \beta_1) = (\alpha, 0, 0)$	1 % = 9.31	No
		5 % = 6.73	
		10 % = 5.61	
$\tau_3 = -1.0871$	$\rho = 0$	1 % = -4.15	No
		5 % = -3.50	
		10 % = -3.18	
$\nabla Z_t = \alpha + \rho Z_{t-1} + e_t$ lags=3			
Estadístico	$H_0$	Valores Críticos	¿Rechaza $H_0$ ?
$\phi_1 = 1.0241$	$(\alpha, \rho) = (0, 0)$	1 % = 7.06	No
		5 % = 4.86	
		10 % = 3.94	
$\tau_2 = -0.7408$	$\rho = 0$	1 % = -3.58	No
		5 % = -2.93	
		10 % = -2.60	
$\nabla Z_t = \rho Z_{t-1} + e_t$ lags=3			
Estadístico	$H_0$	Valores Críticos	¿Rechaza $H_0$ ?
$\tau_1 = 1.2038$	$\rho = 0$	1 % = -2.62	No
		5 % = -1.95	
		10 % = -1.61	
$\nabla Z'_t = \rho Z'_{t-1} + e_t$ lags=2 con $(Z'_t = \nabla Z_t)$			
Estadístico	$H_0$	Valores Críticos	¿Rechaza $H_0$ ?
$\tau_1 = -1.9381$	$\rho = 0$	1 % = -2.66	Sí
		5 % = -1.95	
		10 % = -1.6	

Tabla 2.4: Orden de Integración Gasto en Salud (%PIB)

$\nabla Z_t = \alpha + \rho Z_{t-1} + \beta_1 t + e_t$		Valores Críticos	¿Rechaza $H_0$ ?
Estadístico	$H_0$		
$\phi_3 = 2.1475$	$(\alpha, \rho, \beta_1) = (\alpha, 0, 0)$	1 % = 9.31	No
		5 % = 6.73	
		10 % = 5.61	
$\tau_3 = -2.0706$	$\rho = 0$	1 % = -4.15	No
		5 % = -3.50	
		10 % = -3.18	
$\nabla Z_t = \alpha + \rho Z_{t-1} + e_t$		Valores Críticos	¿Rechaza $H_0$ ?
Estadístico	$H_0$		
$\phi_1 = 1.1422$	$(\alpha, \rho) = (0, 0)$	1 % = 7.06	No
		5 % = 4.86	
		10 % = 3.94	
$\tau_2 = -0.977$	$\rho = 0$	1 % = -3.58	No
		5 % = -2.93	
		10 % = -2.60	
$\nabla Z_t = \rho Z_{t-1} + e_t$		Valores Críticos	¿Rechaza $H_0$ ?
Estadístico	$H_0$		
$\tau_1 = 1.1422$	$\rho = 0$	1 % = -2.62	No
		5 % = -1.95	
		10 % = -1.61	
$\nabla Z'_t = \rho Z'_{t-1} + e_t$ con $(Z'_t = \nabla Z_t)$		Valores Críticos	¿Rechaza $H_0$ ?
Estadístico	$H_0$		
$\tau_1 = -4.3671$	$\rho = 0$	1 % = -2.66	Sí
		5 % = -1.95	
		10 % = -1.6	

Tabla 2.5: Orden de Integración Gasto en Educación (%PIB)

$\nabla Z_t = \alpha + \rho Z_{t-1} + \beta_1 t + e_t$ lags=4			
Estadístico	$H_0$	Valores Críticos	¿Rechaza $H_0$ ?
$\phi_3 = 1.236$	$(\alpha, \rho, \beta_1) = (\alpha, 0, 0)$	1 % = 9.31	No
		5 % = 6.73	
		10 % = 5.61	
$\tau_3 = -1.1889$	$\rho = 0$	1 % = -4.15	No
		5 % = -3.50	
		10 % = -3.18	
$\nabla Z_t = \alpha + \rho Z_{t-1} + e_t$ lags=4			
Estadístico	$H_0$	Valores Críticos	¿Rechaza $H_0$ ?
$\phi_1 = 4.7969$	$(\alpha, \rho) = (0, 0)$	1 % = 7.06	Sí
		5 % = 4.86	
		10 % = 3.94	
$\tau_2 = -1.3159$	$\rho = 0$	1 % = -3.58	No
		5 % = -2.93	
		10 % = -2.60	
$\nabla Z'_t = \rho Z'_{t-1} + e_t$ lags=3 con $(Z'_t = \nabla Z_t)$			
Estadístico	$H_0$	Valores Críticos	¿Rechaza $H_0$ ?
$\tau_1 = -3.1472$	$\rho = 0$	1 % = -2.66	Sí
		5 % = -1.95	
		10 % = -1.6	



# Capítulo 3

## Cointegración y Modelos VEC

### 3.1. El concepto de cointegración: Caso 2 Series.

Como extensión de los modelos *ARIMA* se genera la familia de modelos multivariados, que nos permiten representar el comportamiento de varias series de tiempo de manera simultánea, así como capturar la estructura de asociación, tendencia común, que existe entre las variables. El concepto de cointegración hizo acreedor del premio Nobel de economía al británico Clive W. J. Granger en el año 2003; para ilustrar el tema es oportuno recurrir a la metáfora presentada por Michael P. Murray, la cual nos obliga a recordar al proceso no estacionario conocido como caminata aleatoria asociándola con el andar de un cliente ebrio al salir de un bar.

La trayectoria que sigue un borracho está marcada por pasos independientes de un tiempo al siguiente, por lo que la mejor manera de pronosticar donde será el siguiente paso se reduce a conocer en qué punto se encontraba en el momento inmediato anterior, asimismo conforme avance el tiempo, la varianza creciente del caminar de un borracho provoca que sea casi imposible conocer en qué lugar se encontrará en  $k$  momentos futuros. Estas características que tiene su transitar nos hace recapitular conceptos ya antes vistos como caminata aleatoria, raíz unitaria y un proceso estocástico integrado.

Es importante mencionar que nuestro cliente es muy precavido, ya que

acostumbra viajar a todas partes con su fiel compañero canino; mismo que al ver la condición de su dueño piensa: “No puedo dejarlo ir muy lejos, después de todo, mi deber es protegerlo” así que el perro comienza a seguirlo de manera que también adopta una estructura no estacionaria aunque siempre va evaluando hacia qué posición se está moviendo su dueño corrigiendo su camino si se aleja mucho, por lo que se genera una pequeña brecha entre ambos dentro de la cual se van a mantener en el transcurso del tiempo. A la situación anterior, se le conoce formalmente como **modelo de corrección del error**. Ahora bien, si los observamos desde otra perspectiva, por ejemplo, otros dos clientes dentro del bar viendo al borracho y a su perro verán que ambos muestran un comportamiento no estacionario. Después de observarlos detenidamente uno de los usuarios comenta que a pesar de la no estacionariedad del caminar del perro y su dueño, basta con conocer donde se encuentra el borracho, pues el perro no debería estar muy lejos. Él está en lo correcto, ya que la brecha que se ha formado entre el borracho y su fiel acompañante, permite que ocasionalmente se alejen el uno del otro o se acerquen demasiado pero nunca estarán fuera de control. Esto quiere decir que la distancia entre ambos es estacionaria y a esto se le conoce como **cointegración de orden cero**.

Dejando de lado el aspecto metafórico del concepto debemos recordar que algunas series están mejor representadas por sus primeras  $d$  diferencias, lo que nos dice que son integradas de orden  $d$ ,  $I(d)$ , y por lo tanto dichas variables pueden ir tomando valores grandes y pequeños sobre alguna tendencia estocástica sin mostrar evidencia de que en el largo plazo regresarán a su nivel medio. Supongamos dos variables con esta característica, es decir,  $Y_t, Z_t \sim I(d)$  y cualquier combinación lineal de dichas variables

$$Z_t = \beta_0 + \beta_1 Y_t + e_t \quad (3.1)$$

que también puede ser expresada como

$$e_t = Z_t - \beta_0 - \beta_1 Y_t \quad (3.2)$$

Es también  $I(d)$  y conocemos los riesgos que se pueden presentar al intentar hacer regresión con variables no estacionarias. Sin embargo, existe un caso en

el que  $e_t \sim I(d - b)$  con  $d > 0$ , de manera que se tiene ahora una restricción  $\beta_1$  operando sobre las componentes de largo plazo de la serie para cancelarlas. Si en particular proponemos  $d = 1$  y  $b = 1$ , entonces  $Z_t \sim I(1)$  y  $Y_t \sim I(1)$ , por lo que la regresión de una variable no estacionaria sobre otra, no necesariamente será espuria, es decir, que la constante  $\beta_1$  sugiere algún escalamiento entre las variables para que  $e_t$  sea un proceso estacionario  $e_t \sim I(0)$ . En tal caso se dice que  $Y_t$  y  $Z_t$  están cointegradas lo cual implica que comparten la misma tendencia estocástica de manera que no divergirán muy lejos la una de la otra en el largo plazo. Por lo tanto, dos series cointegradas son estacionarias, a pesar de que individualmente no lo sean.

La fuerte contribución que tienen conceptos como raíz unitaria, cointegración, entre otros, nos obliga a averiguar si los residuos de una regresión son estacionarios. Por lo que una prueba de cointegración puede ser pensada también como una pre-prueba para evitar situaciones de regresiones espurias.

Una regresión como la que se muestra en la ecuación 3.1 es conocida como **regresión de cointegración** mientras que el parámetro  $\beta_1$  se le conoce como **parámetro de cointegración**, además el concepto de cointegración se puede extender a una regresión con  $n$  regresores, lo que nos llevaría a tener  $n$  parámetros de cointegración. En este caso, la cointegración que existe entre  $n$  variables integradas, existe si hay al menos una pero no más de  $n - 1$  combinaciones lineales que son estacionarias. Si se tuviera el caso en que las  $n$  combinaciones lineales son estacionarias, entonces se puede concluir que las  $n$  variables son integradas de orden cero desde el momento de hacer la regresión, ya que cualquier combinación lineal de variables estacionarias es estacionaria.

### 3.1.1. Prueba de Cointegración

Como dijimos anteriormente, la prueba de cointegración se basa en demostrar que la ecuación 3.2 es estacionaria, sin embargo no conocemos el valor de los residuales teóricos, así que se demuestra para los residuales estimados por mínimos cuadrados ordinarios. Por lo tanto, como primera

instancia debemos realizar la regresión lineal entre las variables que creemos que pueden estar cointegradas.

$$\hat{Z}_t = \hat{\beta}_0 + \hat{\beta}_1 Y_t \quad (3.3)$$

De manera que observamos si los parámetros estimados son significativos estadísticamente hablando con el uso del estadístico usual  $t$ , si lo son entonces obtenemos los residuales estimados de dicha regresión

$$\hat{e}_t = Z_t - \hat{\beta}_0 - \hat{\beta}_1 Y_t \quad (3.4)$$

El siguiente paso es determinar si efectivamente los residuales estimados son estacionarios usando la prueba de Dickey-Fuller. Por lo tanto, si resultan ser estacionarios entonces hay cointegración entre las variables  $Z_t$  y  $Y_t$ ; si no se rechaza la hipótesis nula entonces los residuales no son estacionarios y cualquier posible relación entre las variables podrá tratarse de una regresión espuria.

La prueba de estacionariedad está basada en la siguiente ecuación característica de la prueba Dickey-Fuller

$$\nabla e_t = \rho \hat{e}_{t-1} + \nu_t \quad (3.5)$$

Como es posible observar, la ecuación no tiene término constante ya que los residuales tienen media cero, además se puede extender la prueba agregando términos  $\nabla \hat{e}_{t-1}$ ,  $\nabla \hat{e}_{t-2}$ , etcétera, para eliminar la posible existencia de autocorrelación en  $\nu_t$  al igual que en la prueba de Dickey-Fuller Aumentada. Sin embargo, puesto que  $\hat{e}_t$  depende de la estimación de  $\beta_1$  los valores críticos de  $DF$  y  $ADF$  no son apropiados para el análisis de estacionariedad.

La prueba de Dickey-Fuller Aumentada para los residuales es de la forma:

$$\nabla e_t = \rho \hat{e}_{t-1} + \sum_{i=1}^m a_i \nabla e_{t-i} \quad (3.6)$$

Los valores críticos apropiados para determinar cointegración son generados para responder la siguiente prueba de hipótesis:

$$\begin{aligned}
H_0 : \beta_1 &= 0 \quad o \quad \rho \neq 0 \\
H_a : \beta_1 &\neq 0 \quad y \quad \rho = 0
\end{aligned}
\tag{3.7}$$

La cual considera en un mismo instante la significancia de la estimación del parámetro  $\beta_1$  de la regresión de cointegración y la prueba de raíz unitaria para la estacionariedad en los residuales. La hipótesis nula nos dice que los residuales no son estacionarios o que la estimación de  $\beta_1$  no es significativa, por lo que podemos concluir que las series no están cointegradas, mientras que la hipótesis alternativa hace referencia de un parámetro  $\beta_1$  distinto de cero y estacionariedad en los residuales y por lo tanto, estamos en presencia de series cointegradas. Existen tres tipos de valores críticos y cada tipo depende de la regresión a partir de la cual provienen los residuales, es decir, si los residuales provienen de una regresión sin término constante (*I*), con término constante (*II*) o con tendencia en el tiempo y término constante (*III*).

Tabla 3.1: **Valores Críticos para la prueba de Cointegración**

Modelo de Regresión	1 %	5 %	10 %
( <i>I</i> ) $Z_t = \beta_1 Y_t + e_t$	-3.39	-2.76	-2.45
( <i>II</i> ) $Z_t = \beta_0 + \beta_1 Z_t + e_t$	-3.96	-3.37	-3.07
( <i>III</i> ) $Z_t = \beta_0 + \delta_t + \beta_1 Z_t + e_t$	-3.98	-3.42	-3.13

De manera similar a la prueba de raíces unitarias, se rechaza la hipótesis nula de no cointegración si nuestro estadístico  $\tau$  es menor o igual a alguno de los valores críticos mencionados en el cuadro 3.1, es decir,  $\tau \leq \tau_c$  lo cual implicaría que el valor del estadístico sea más negativo que el valor crítico deseado. Esto conlleva a concluir que las series  $Z_t$  y  $Y_t$  están cointegradas, ya que la regresión de cointegración estimada es válida (no espuria) y por lo tanto existe una relación entre ambas que será notable en el largo plazo. Consecuentemente no rechazaríamos la hipótesis nula si  $\tau > \tau_c$ . Las pruebas de cointegración buscan únicamente relaciones estables lineales, por lo que debemos tener especial atención cuando se acepta la hipótesis relacionada con

no cointegración, ya que ésta solo sugiere que no hay una relación lineal en el largo plazo entre las variables y no ausencia de una relación a largo plazo estable entre ellas.

### 3.1.2. Modelo de Corrección de Errores

Al igual que en la sección anterior se continuará trabajando con series cuyo orden de integración sea mayor o igual a uno, es decir, que se requiera aplicar una vez o más el operador diferencia ( $\nabla$ ) para convertir las series en estacionarias, ya que en la serie original hay presencia de alguna tendencia polinomial adaptiva; además, es de nuestro interés observar cómo es la tendencia estocástica de los procesos y en caso de que la compartan poder establecer relaciones de equilibrio en el largo plazo. En particular, en este apartado las variables consideradas son integradas de orden uno y cointegradas.

De acuerdo a la metáfora del perro con su dueño ebrio, ambos adoptaron una dinámica no estacionaria, sin embargo, el perro intentaba no alejarse mucho de su amo por lo que corregía su andar cada vez que se alejaba una distancia considerable, manteniéndolos dentro de una brecha a lo largo del tiempo. Esto quiere decir que nuestro principal problema es que dada una variable la cual depende de sus valores pasados y también en los valores actuales y pasados de otra variable exógena poder determinar la relación de equilibrio en el largo plazo entre las variables endógenas y exógenas. Si consideramos únicamente a una variable  $Y_t$  en función de los valores de un conjunto de variables exógenas  $Z_t$  en el mismo momento de tiempo, es decir, sin retrasos, tendríamos un efecto inmediato y completo de  $Z_t$  sobre  $Y_t$ . Sin embargo, si consideramos retrasos en cada una de las variables tendremos que analizar también el efecto en el largo plazo, que estará en función de todos los retrasos.

De manera que estamos enfocados en analizar los efectos de corto y largo plazo de los procesos, los cuales pueden ser encontradas en diferentes formas equivalentes de ecuaciones dinámicas; al referirnos al término “equivalente” lo

hacemos en el sentido de que cualquiera de las formas de la ecuación explica exactamente lo mismo del proceso pero cada una de ellas revelará diferentes tipos de información. Esto se puede observar claramente cuando se tiene un proceso  $ARMA(p, q)$  estacionario e invertible, ya que acepta tanto una representación del tipo modelo lineal general como una representación autorregresiva con una infinidad de retrasos, ambas explican lo mismo del proceso, sin embargo, observamos distintos tipos de información en cada una de ellas pues en la primera podemos identificar cuáles son los retrasos más importantes con la finalidad de tener una mejor interpretación de la serie; en la segunda se encuentra toda la información correspondiente a la inferencia estadística, es decir, la media de la serie, las varianzas y covarianzas, entre otras. Por lo tanto, se pretende usar la representación que sea más conveniente para el análisis, en particular consideraremos un modelo lineal conocido como **Retrasos Distribuidos Autorregresivos** con notación  $ARDL$  por sus siglas en inglés, a partir del cual deduciremos el modelo de corrección del error.

Un modelo  $ARDL$  contiene retrasos tanto de la variable exógena como de la variable endógena, para su representación general consideremos  $p$  retrasos en la variable  $Y_t$  y  $q$  retrasos de  $Z_t$  denotado como  $ARDL(p, q)$  que es de la siguiente forma:

$$Y_t = \delta + \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + \delta_0 Z_t + \delta_1 Z_{t-1} + \cdots + \delta_q Z_{t-q} + e_t \quad (3.8)$$

Donde  $e_t \sim IID(0, \sigma_e^2)$ , la parte autorregresiva del modelo se puede encontrar en la regresión de la variable  $Y_t$  sobre sus propios valores pasados, mientras que el componente de retrasos distribuidos se observa en el efecto de la variable  $Z_t$  junto con sus retrasos. Este tipo de modelos trae consigo una serie de ventajas que son deseables en el análisis, ya que captura la dinámica de las  $Z$ 's y de las  $Y$ 's e incluyendo suficientes retrasos podemos asegurar que se va a eliminar la correlación serial entre los errores.

Para poder derivar de un modelo  $ARDL$  el modelo de corrección del error, es oportuno considerar un  $ARDL(1, 1)$ , el cual considera sólo un retraso en

cada variable:

$$Y_t = \delta + \theta_1 Y_{t-1} + \delta_0 Z_t + \delta_1 Z_{t-1} + e_t \quad (3.9)$$

Si por un momento suponemos que las variables son estacionarias, es decir,  $Y_t \sim I(0)$  y  $Z_t \sim I(0)$ , podemos decir que cualquier combinación lineal de estas dos variables es estacionaria y además, sus dos primeros momentos son invariantes en el tiempo, por lo que sus valores en el largo plazo estarán dados por sus valores esperados  $E(Y_t) = y^*$ ,  $E(Z_t) = z^*$  y  $E(e_t) = 0$  de manera que si calculamos el valor esperado de la ecuación 3.9 y haciendo uso de sus propiedades como operador lineal, se tiene que

$$\begin{aligned} y^* &= \delta + \theta_1 y^* + \delta_0 z^* + \delta_1 z^* \\ y^* - \theta_1 y^* &= \delta + \delta_0 z^* + \delta_1 z^* \\ (1 - \theta_1)y^* &= \delta + (\delta_0 + \delta_1)z^* \end{aligned} \quad (3.10)$$

que también se puede escribir como

$$y^* = \beta_1 + \beta_2 z^* \quad (3.11)$$

con  $\beta_1 = \frac{\delta}{1-\theta_1}$  y  $\beta_2 = \frac{\delta_0 + \delta_1}{1-\theta_1}$  con lo que hemos derivado la relación que mantienen estas dos variables conocido como el **término de error de equilibrio**. Regresando al supuesto de que ambas variables son  $I(1)$  y cointegradas, entonces la ecuación 3.11 será la relación de cointegración entre  $Y_t$  y  $Z_t$ , por lo que se puede asegurar que existe relación en el largo plazo entre ellas, sin embargo, en el corto plazo puede haber desequilibrio que nos gustaría capturar, de manera que usaremos este término para ajustar el comportamiento en el corto plazo de  $Y_t$  a su valor en el largo plazo, corrigiendo el desequilibrio.

Para apreciar esto de manera mas concreta, se deducirá el modelo de corrección del error manipulando el modelo  $ARDL(1,1)$  con una serie de operaciones algebraicas que implicarán la misma relación, ya que una se puede derivar de la otra sin modificar la igualdad. Las primeras



modificaciones consisten en restar en ambos lados de la ecuación  $Y_{t-1}$  y posteriormente sumar y restar en el lado derecho de la ecuación  $\delta_0 Z_{t-1}$ :

$$Y_t - Y_{t-1} = \delta + (\theta_1 - 1)Y_{t-1} + \delta_0(Z_t - Z_{t-1}) + (\delta_0 + \delta_1)Z_{t-1} + e_t \quad (3.12)$$

Finalmente, debemos identificar los casos en donde se puede usar el operador diferencia ( $\nabla$ ) y factorizar  $(\theta_1 - 1)$  en los términos restantes

$$\nabla Y_t = \delta + (\theta_1 - 1)Y_{t-1} + \delta_0 \nabla Z_t + (\delta_0 + \delta_1)Z_{t-1} + e_t \quad (3.13)$$

factorizando tenemos

$$\nabla Y_t = (\theta_1 - 1) \left( \frac{\delta}{\theta_1 - 1} + Y_{t-1} + \frac{\delta_0 + \delta_1}{\theta_1 - 1} Z_{t-1} \right) + \delta_0 \nabla Z_t + e_t \quad (3.14)$$

Sin embargo, al encontrar la relación de cointegración habíamos establecido los valores  $\beta_1 = \frac{\delta}{1-\theta_1}$  y  $\beta_2 = \frac{\delta_0 + \delta_1}{1-\theta_1}$  que vuelven a aparecer en esta ecuación:

$$\nabla Y_t = -\alpha (Y_{t-1} - \beta_1 - \beta_2 Z_{t-1}) + \delta_0 \nabla Z_t + e_t \quad (3.15)$$

con  $\alpha = (1 - \theta_1)$ ,

Para que esta regresión, en términos estadísticos, sea válida y no genere resultados espurios se requiere que todos sus elementos sean estacionarios, sin embargo, por los supuestos establecidos, las variables son  $I(1)$  y es por eso que aparecen con una diferencia, además la relación  $(Y_{t-1} - \beta_1 - \beta_2 Z_{t-1})$  también debe ser estacionaria para que la ecuación esté balanceada, lo cual sucede únicamente cuando las variables están cointegradas, así que acabamos de hallar dentro de un modelo *ARDL* una relación de cointegración. Por lo tanto, la variable  $Y_t$  depende de la diferencia de  $Z_t$  con su periodo anterior ( $\nabla Z_t$ ) y del término de error o desequilibrio pasado, el cual deseamos que sea cero, sin embargo cuando es distinto de cero implica que no hay equilibrio y se debe emplear el mecanismo de corrección del error, ya que este es una corrección para el equilibrio.

Si suponemos  $\nabla Z_t = 0$  y al observar el término de error de equilibrio en el periodo anterior  $u_{t-1} = Y_{t-1} - \beta_1 - \beta_2 Z_{t-1}$ , donde  $u_{t-1}$  es un proceso de ruido blanco, resulta ser positivo,  $u_{t-1} > 0$ , implicaría que no hay equilibrio en la relación, ya que  $Y_{t-1}$  es más grande que su valor de equilibrio en el largo plazo  $\beta_1 + \beta_2 Z_{t-1}$ , consecuentemente el valor de  $\alpha$ , el cual se espera que sea mayor a cero, recibe todo el impacto por tener  $Y_{t-1} \neq \beta_1 + \beta_2 Z_{t-1}$ , dicha discrepancia puede ser producto de malas decisiones pasadas, por lo que el término  $\alpha$  refleja el ajuste en el corto plazo de tales errores para el próximo periodo. En otras palabras, si tenemos que  $u_{t-1} > 0$  y  $\alpha > 0$  tendremos que  $\nabla Y_t < 0$ , esto es  $Y_t < Y_{t-1}$ , por lo que  $Y_t$  empezará a caer en el siguiente periodo para corregir el error de equilibrio.

De manera similar, podemos conocer cuáles son las implicaciones reflejadas en el modelo cuando se tiene  $Y_{t-1} < \beta_0 + \beta_1 Z_{t-1}$ , es decir, nuevamente no hay equilibrio en la relación pero en este caso el valor de la variable  $Y_{t-1}$  resulta ser menor que su valor de equilibrio en el largo plazo, si además conservamos el supuesto de que  $\alpha > 0$  entonces es fácil observar que  $Y_t$  comenzará a aumentar para el siguiente periodo por el hecho de que  $\nabla Y_t$  es positivo.

El supuesto que se establece sobre  $\alpha$  implica que el valor de  $\theta_1$  debe ser menor que la unidad, ya que si este fuera igual que 1 no sería posible determinar una relación de cointegración y en cuyo caso se tendría que la diferencia de  $Y_t$  queda explicada únicamente por la diferencia de una variable exógena  $Z_t$  más el término de ruido blanco, modelo conocido como **Modelo de Vectores Autorregresivos**, *VAR* por sus siglas en inglés.

El trabajo de Granger y Engel ha sido fundamental en el desarrollo de una metodología econométrica, ya que en un inicio se asumía estacionariedad en los datos, sin embargo, el modelo de corrección del error permite la existencia de unión, en una misma ecuación, entre variables que guardan una relación a largo plazo  $(Y_{t-1}, Z_{t-1}) \sim I(1)$  con variables relacionadas en el corto plazo  $(\nabla Y_t, \nabla Z_t) \sim I(0)$  con la única condición de que  $Y_t$  y  $Z_t$  estén cointegradas, de manera que el término del error de equilibrio tenga residuales estacionarios.

Un factor importante que ha contribuido a la rápida adopción de la teoría de cointegración y el modelo de corrección del error en el uso de la econometría

moderna ha sido la sencillez del procedimiento que han propuesto Granger y Engel, que consiste en aplicar un método de estimación que genere estimadores consistentes de los parámetros de la ecuación de equilibrio y posteriormente estimar el *ECM* incluyendo los errores rezagados de la ecuación de equilibrio, ya que los errores estimados nos permitirán conocer si hay una relación de largo plazo entre las variables, es decir, la formulación del mecanismo de corrección del error también puede servir para probar cointegración. Por lo tanto, el *ECM* al ser estimado proporciona el ajuste dinámico en el corto plazo, de manera que se obtiene una mejor aproximación del proceso generador de los datos.

## 3.2. El concepto de cointegración: Caso 2+ series (VEC)

En secciones previas, se ha tratado el concepto de cointegración entre dos series con orden de integración uno, es decir, variables  $I(1)$  tales que existe una combinación lineal entre ellas cuyos residuos son estacionarios. Esto nos permite hacer referencia de la relación a largo plazo existente entre las series, así como proponer un modelo de corrección del error para obtener beneficios de la dinámica de ambas variables. Sin embargo, es tentador extender el análisis a un caso más general que permita rescatar la información de varias series en un mismo instante.

Es lógico pensar que la definición de cointegración resulta invariante respecto a la suma de variables en el sistema, es decir, si se tienen  $k$  series de tiempo cointegradas donde cada una de ellas es integrada de orden  $d$  ( $I(d)$ ), entonces existe al menos una combinación lineal que resulta ser estacionaria. En términos matriciales y de acuerdo a la definición de Engle y Granger (1987) se tiene que los componentes del vector  $\bar{Z}_t$  se dice que están cointegradas de orden  $d, b$ , denotado  $\bar{Z}_t \sim CI(d, b)$ , si todos los componentes de  $\bar{Z}_t$  son  $I(d)$  y existe un vector  $\alpha$  diferente de cero tal que  $\bar{Z}_t = \alpha' \bar{Z}_t \sim I(d - b)$  con  $b > 0$ . El vector  $\alpha$  es llamado vector de cointegración.

La metodología de dos pasos de Granger, presentada anteriormente,

ayuda a identificar si dos series están cointegradas ajustando cada una de las regresiones lineales del sistema y posteriormente observando si los residuos son estacionarios, dicha estrategia se puede extender fácilmente a un caso más general. Sin embargo, la prueba comienza a mostrar debilidades al aumentar el número de variables en el sistema, ya que en primer lugar, ajustar las regresiones lineales del sistema otorgará únicamente uno de los posibles vectores de cointegración; en segundo lugar, las conclusiones pueden ser ambiguas dependiendo de la variable usada para la normalización y finalmente, al tratarse de una estimación en dos pasos, es posible acarrear errores que provoquen graves consecuencias para el análisis.

Existe otra técnica que ayuda a corregir estas deficiencias de la metodología de Granger, como es el caso del **Procedimiento de Johansen** que encuentra su motivación a partir de una generalización multivariada de la prueba de Dickey-Fuller, la cual emprende a partir de un proceso autorregresivo de orden uno de la siguiente forma:

$$\begin{aligned} Z_t &= a_1 Z_{t-1} + e_t \\ Z_t - Z_{t-1} &= (a_1 - 1) Z_{t-1} + e_t \\ \nabla Z_t &= \rho Z_{t-1} + e_t \end{aligned}$$

donde  $\rho = (a_1 - 1)$  y las pruebas de hipótesis correspondientes son  $H_0 : \rho = 0$  y  $H_a : \rho \neq 0$ . Por lo tanto, de manera análoga es posible derivar su versión multivariada

$$\begin{aligned} \bar{Z}_t &= A \bar{Z}_{t-1} + e_t \\ \bar{Z}_t - \bar{Z}_{t-1} &= (A - I_k) \bar{Z}_{t-1} + e_t \\ \nabla \bar{Z}_t &= \Pi \bar{Z}_{t-1} + e_t \end{aligned} \tag{3.16}$$

donde  $\Pi = (A - I_k)$ ,  $\bar{Z}_t = (Z_{1t}, Z_{2t}, \dots, Z_{kt})'$ ,  $e_t$  es un vector de ruido blanco Gaussiano de dimensión  $(k \times 1)$  y las matrices  $I$  y  $A$  son de dimensión  $(k \times k)$ , de manera que la prueba de hipótesis a realizar es respecto al rango de la

### 3.2. EL CONCEPTO DE COINTEGRACIÓN: CASO 2+ SERIES (VEC)47

matriz  $\Pi$ , dicho de otra forma, determinar cuántas relaciones de cointegración se hallan en el sistema.

En consecuencia, si la matriz  $\Pi$  es de rango completo indicaría que todos los vectores son linealmente independientes y por lo tanto vectores de cointegración, por lo que se concluye que todas las variables son estacionarias ya que cualquier combinación lineal de variables estacionarias, es estacionaria. En cambio, si el rango es cero quiere decir que las variables no guardan ningún tipo de asociación a largo plazo y por ende no están cointegradas, finalmente, si  $rank(\Pi) = r$  con  $r < k$  implica que únicamente existen  $r$  combinaciones lineales de los procesos en  $\bar{Z}_t$  que son estacionarias.

Al igual que en la prueba de **Dickey-Fuller Aumentada** es factible aumentar el orden del vector autorregresivo

$$\bar{Z}_t = A_1 \bar{Z}_{t-1} + A_2 \bar{Z}_{t-2} + \cdots + A_p \bar{Z}_{t-p} + e_t \quad (3.17)$$

el cual se puede expresar de manera más compacta y útil al realizar la siguiente recursión:

Restar  $\bar{Z}_{t-1}$ ,

$$\nabla \bar{Z}_t = (A_1 - I_k) \bar{Z}_{t-1} + A_2 \bar{Z}_{t-2} + \cdots + A_p \bar{Z}_{t-p} + e_t$$

sumar y restar del lado derecho de la ecuación  $(A_1 - I_k) \bar{Z}_{t-2}$

$$\nabla \bar{Z}_t = (A_1 - I_k) \bar{Z}_{t-1} + (A_2 + A_1 - I_k) \bar{Z}_{t-2} + \cdots + A_p \bar{Z}_{t-p} + e_t$$

sumar y restar del lado derecho de la ecuación  $(A_2 + A_1 - I_k) \bar{Z}_{t-3}$

$$\nabla \bar{Z}_t = (A_1 - I_k) \bar{Z}_{t-1} + (A_2 + A_1 - I_k) \bar{Z}_{t-2} + (A_3 + A_2 + A_1 - I_k) \bar{Z}_{t-3} \cdots + A_p \bar{Z}_{t-p} + e_t$$

y así sucesivamente para obtener

$$\nabla \bar{Z}_t = \sum_{i=1}^{p-1} \Gamma_i \nabla \bar{Z}_{t-i} + \Pi \bar{Z}_{t-p} + e_t \quad (3.18)$$

con

$$\Gamma_i = - \left( I - \sum_{j=1}^i A_j \right) \quad (3.19)$$

$$\Pi = - \left( I - \sum_{i=1}^p A_i \right) \quad (3.20)$$

A la expresión 3.18, la cual es la versión generalizada multivariada de Dickey-Fuller Aumentado se le conoce como el **Modelo de Vectores de Corrección del error** denotado como *VEC* por sus siglas en inglés.

Si cada uno de los elementos del vector  $\bar{Z}_t$  contiene una raíz unitaria, es conveniente analizar más a fondo la ecuación 3.18, dado que aquí se encuentra la razón del origen del modelo. Al considerar que  $\bar{Z}_t \sim I(1)$  implica que cada una de las series del vector requieren de una diferencia para ser estacionarias en cuanto a su nivel, por lo tanto,  $\nabla \bar{Z}_t = (\nabla Z_{1t}, \nabla Z_{2t}, \dots, \nabla Z_{kt})'$  es un vector de  $(k \times 1)$  que contiene en cada elemento a una serie estacionaria. Esto significa que la parte izquierda de la expresión del modelo *VEC* es estacionaria y al tratarse de una igualdad, nos obliga a suponer que la parte derecha de la misma también debe de serlo. Es posible observar que el término  $\sum_{i=1}^{p-1} \Gamma_i \nabla \bar{Z}_{t-i}$  es estacionario puesto que tiene a cada una de las series con su primera diferencia, además,  $e_t$  bajo el supuesto de ruido blanco en los errores, es estacionario. Sin embargo, el único término del que no se tiene la certeza de que se trata de un proceso estacionario es  $\Pi \bar{Z}_{t-p}$ . Este es el argumento que llevó a S. Johansen (1991) a considerar que la matriz  $\Pi$  debe ser de tal forma que al ser premultiplicada por  $\bar{Z}_{t-p}$  debe ser estacionaria. Esto quiere decir que se debe analizar el rango de la matriz  $\Pi$  debido a que pueden existir  $r$  relaciones de cointegración  $0 \leq r \leq k$ .

Evidentemente los casos de menor importancia son cuando el rango de la matriz es nulo o es de rango completo, ya que el primero de ellos indicaría que la matriz  $\Pi$  es la matriz nula y no hay ningún tipo de relación de largo plazo entre las variables, haciendo del modelo *VEC* un *VAR*( $p - 1$ ) en primeras diferencias, la cual sería la especificación estable adecuada para  $\nabla \bar{Z}_t$  en este

caso; en el segundo caso, implicaría que el vector  $\bar{Z}_t \sim I(0)$  en cuanto a su nivel y por ende se trata únicamente de un  $VAR(p+1)$ . En consecuencia, el caso para el cual se tiene que el rango de la matriz  $\Pi$  es igual a  $r$  despierta nuestro interés por el hecho de que se desea encontrar una especificación apropiada por el modelo, es decir, proponer un  $VAR(p-1)$  en primeras diferencias no es la mejor selección pues impondría la restricción falsa de  $\Pi = 0$  perdiendo así información valiosa de las relaciones de cointegración presentes en el sistema, de manera que la expresión  $\Pi\bar{Z}_{t-p}$  representa el factor de corrección del error, que ha de ajustar en el corto plazo las desviaciones en las relaciones de largo plazo.

### 3.2.1. Inferencia Sobre los Coeficientes

La inferencia sobre los parámetros del modelo  $VEC$  es, en esencia, la misma que se ha proporcionado para los coeficientes en el modelo de corrección del error. Al observar nuevamente la ecuación 3.15 es posible notar que el vector de cointegración para la relación  $Y_{t-1} - \beta_1 - \beta_2 Z_{t-1}$  es  $(1, -1, -1)$  y  $\alpha$  representa la velocidad de ajuste ante desviaciones en las relaciones de largo plazo. Esta información también se encuentra expresada dentro de la matriz  $\Pi$  del modelo vectorial de corrección del error y para obtenerla de forma explícita es necesario considerar que al emplear una matriz de dimensiones finitas, en particular de  $(k \times k)$ , y con rango  $0 \leq r \leq k$ , admite una factorización de rango. Dicha factorización hace posible expresar a la matriz  $\Pi$  como producto de dos matrices  $\Pi = \alpha\beta'$  donde ambas matrices son de dimensión  $(k \times r)$  con rango  $rank(\alpha) = rank(\beta') = r$ .

Para obtener la factorización solo basta con recapitular que la matriz  $\Pi$  es de rango  $r$  y por lo tanto, tiene  $r$  columnas que son linealmente independientes, equivalentemente, la dimensión del espacio columna es de dimensión  $r$ , así que la matriz  $\beta'$  se construye a partir de colocar la base del espacio columna de  $\Pi$  como vectores columna. En consecuencia, cada vector columna de la matriz  $\Pi$  es una combinación lineal de las columnas de  $\beta'$ , de manera que los coeficientes que se requieren en las combinaciones lineales para generar cada elemento de  $\Pi$  constituyen la matriz  $\alpha$ . A causa de esta factorización se gana interpretación

en el modelo, ya que ahora debe resultar claro que la matriz  $\beta'$  contiene los  $r$  vectores de cointegración existentes en el modelo, mientras que los elementos de  $\alpha$  son los coeficientes del modelo *VEC* en los cuales se podrá observar el ajuste en corto plazo.

Por lo tanto, la ecuación 3.18 también puede ser expresada como:

$$\nabla \bar{Z}_t = \sum_{i=1}^{p-1} \Gamma_i \nabla \bar{Z}_{t-i} + \alpha \beta' \bar{Z}_{t-p} + e_t \quad (3.21)$$

donde por definición se tiene que  $\beta' \bar{Z}_{t-p} \sim I(0)$ .

Finalmente, es necesario enfatizar que las matrices  $\alpha$  y  $\beta$  no son únicas ya que para cualquier matriz no singular  $D$  de dimensión  $(r \times r)$  se tiene que

$$\alpha \beta' = \alpha D D^{-1} \beta' = (\alpha D) (D^{-1} \beta') = (\alpha D) (\beta D^{-1'})' = \alpha^* \beta^{*'} \quad (3.22)$$

de manera que las matrices  $\alpha$  y  $\beta$  no están perfectamente definidas a menos que se impongan restricciones sobre los coeficientes.

### 3.2.2. Estimación

La siguiente pregunta que debe surgir respecto al análisis de un modelo de corrección del error es cómo determinar el rango de la matriz  $\Pi$ . Para ello Johansen (1991) ha desarrollado dos estadísticos de prueba que, basados en Máxima Verosimilitud, permiten hacer una estimación de cuántas posibles relaciones de cointegración se encuentran en el modelo, a través de la técnica de cociente de verosimilitudes.

Sea el modelo *VEC*

$$\nabla \bar{Z}_t = \sum_{i=1}^{p-1} \Gamma_i \nabla \bar{Z}_{t-i} + \Pi \bar{Z}_{t-p} + e_t \quad (3.23)$$



donde  $\Gamma_i = (-I_k + A_1 + A_2 + \cdots + A_i)$  para  $i = 1, \dots, p-1$  y  $\Pi = (-I_k + A_1 + A_2 + \cdots + A_p)$ , suponiendo que  $e_t \sim N(0, \Sigma_e)$  y  $\text{rank}(\Pi) = r$ , dado que no se tienen restricciones sobre  $r$ , el resultado se puede generalizar para cuando  $r = k$ .

Por lo tanto, dado que los errores provienen de una distribución normal multivariada, el logaritmo de la función de verosimilitud de  $T$  observaciones condicionada a los valores iniciales  $(\bar{Z}_{1-p}, \bar{Z}_{2-p}, \dots, \bar{Z}_0)$  correspondiente es:

$$\ln(L(\Pi, \Gamma, \Sigma_e)) = -\frac{Tk}{2} \ln(2\pi) - \frac{T}{2} \ln |\Sigma_e| - \frac{1}{2} \sum_{t=1}^T e'_t \Sigma_e^{-1} e_t \quad (3.24)$$

con  $e_t = \nabla \bar{Z}_t - \sum_{i=1}^{p-1} \Gamma_i \nabla \bar{Z}_{t-i} - \Pi \bar{Z}_{t-p}$  y  $\Gamma = (\Gamma_1, \dots, \Gamma_{p-1})$ .

El esquema a seguir para poder encontrar el estimador de máxima verosimilitud de  $\Pi$  consiste en componer una función de verosimilitud a partir de 3.24 que únicamente dependa de la matriz  $\Pi$ , lo cual será posible si se maximiza respecto a un parámetro considerando el resto como fijos, por ende, si maximizamos la función de verosimilitud respecto a  $\Sigma_e$  sujeta a que  $\Pi$  y  $\Gamma$  están dados, obtendremos que

$$\hat{\Sigma}_e = \frac{1}{T} \sum_{t=1}^T e_t e'_t \quad (3.25)$$

Igualmente, para obtener  $\hat{\Gamma}$  es necesario maximizar 3.24 respecto a  $\Gamma$  sujeto a que  $\Pi$  y  $\Sigma_e$  son dados, sin embargo, la ecuación 3.18 se puede reparametrizar si definimos  $W_{0t} = \nabla \bar{Z}_t$ ,  $W_{pt} = \bar{Z}_{t-p}$  y

$$W_{1t} = \begin{pmatrix} \nabla \bar{Z}_{t-1} \\ \nabla \bar{Z}_{t-2} \\ \vdots \\ \nabla \bar{Z}_{t-p+1} \end{pmatrix} \quad (3.26)$$

por lo que dado un valor de  $\Pi$  se obtiene la nueva expresión  $(W_{0t} - \Pi W_{pt}) = \Gamma W_{1t} + e_t$  o, de manera análoga,  $(W_{0t} - \Pi W_{pt})' = W_{1t}' \Gamma' + e'_t$  para la cual

el estimador de máxima verosimilitud de  $\Gamma'$  dados  $\Pi$  y  $\Sigma_e$  coincide con el estimador de mínimos cuadrados ordinarios ecuación por ecuación, ya que es un sistema de ecuaciones con idénticos regresores en cada ecuación.

Sea,

$$S_{ij} = \frac{1}{T} \sum_{t=1}^T R_{it} R_{jt}' \quad i, j = 0, p \quad (3.27)$$

donde  $R_{0t}$  es el residuo en  $t$  de la regresión de cada uno de los elementos de  $W_{0t} = \nabla \bar{Z}_t$  sobre  $W_{1t} = (\nabla \bar{Z}'_{t-1}, \nabla \bar{Z}'_{t-2}, \dots, \nabla \bar{Z}'_{t-p+1})$  con  $t = 1, \dots, T$  y  $R_{1t}$  es el residuo en  $t$  de la regresión de cada uno de los elementos de  $W_{pt} = \bar{Z}_{t-p}$  sobre  $W_{1t} = (\nabla \bar{Z}'_{t-1}, \nabla \bar{Z}'_{t-2}, \dots, \nabla \bar{Z}'_{t-p+1})$  con  $t = 1, \dots, T$ . Por lo que es posible escribir la función de verosimilitud concentrada solamente en función de  $\alpha$  y  $\beta$ , pues  $\Pi = \alpha\beta'$ .

$$\ln(L_c(\alpha, \beta)) = -\frac{kT}{2}(\ln(2\pi) + 1) - \frac{T}{2} \ln |S_{00} - \alpha\beta' S_{p0} - S_{0p}\beta\alpha' + \alpha\beta' S_{pp}\beta\alpha'| \quad (3.28)$$

a partir de la cual, al maximizar, podemos obtener una estimación de  $\alpha$  considerando a  $\beta$  como valor fijo, de manera que

$$\hat{\alpha} = S_{op}\beta(\beta' S_{pp}\beta)^{-1} \quad (3.29)$$

sustituyendo  $\hat{\alpha}$  en 3.28 es factible concentrar la función de verosimilitud que dependerá exclusivamente de  $\beta$  obteniendo lo siguiente:

$$\ln(L_{cc}(\beta)) = -\frac{kT}{2}(\ln(2\pi)+1) - \frac{T}{2} \ln \left( |\beta' S_{pp}\beta|^{-1} |S_{00}| |\beta' (S_{pp} - S_{p0}S_{00}^{-1}S_{0p}) \beta| \right) \quad (3.30)$$

Maximizar esta última ecuación corresponde a minimizar el siguiente cociente imponiendo la normalización  $\beta' S_{pp}\beta = I$

$$\frac{|\beta' (S_{pp} - S_{p0}S_{00}^{-1}S_{0p}) \beta|}{|\beta' S_{pp}\beta|} \quad (3.31)$$

el cual equivale a resolver el problema de obtención de los  $r$  mayores valores propios de los  $k$  existentes; supongamos que  $\lambda_1 > \lambda_2 > \dots > \lambda_r > \dots > \lambda_k$  para los cuales tenemos que sus correspondientes vectores propios son:  $(v_1, v_2, \dots, v_r, \dots, v_k)$ , de tal suerte que las  $r$  columnas de  $\hat{\beta}_{MV}$  pertenecen a los  $r$  primeros vectores.

Por lo tanto, una vez obtenido  $\hat{\beta}_{MV}$ , el logaritmo de la función de verosimilitud 3.24 evaluada en el máximo considerando  $r$  como el rango de cointegración es:

$$\ln(L_r) = -\frac{Tk}{2} (\ln(2\pi) + 1) - \frac{T}{2} \left( \ln |S_{00}| + \sum_{i=1}^r \ln(1 - \lambda_i) \right) \quad (3.32)$$

con

$$\begin{aligned} \hat{\alpha}_{MV} &= S_{op} \hat{\beta} (\hat{\beta}' S_{pp} \hat{\beta})^{-1} \\ \hat{\Pi}_{MV} &= \hat{\alpha}_{MV} \hat{\beta}'_{MV} \\ \hat{\Gamma}'_{MV} &= \left( \sum_{t=1}^T Z_{1t} Z_{1t}' \right)^{-1} \left( \sum_{t=1}^T Z_{1t} (Z_{0t} - \hat{\Pi}_{MV} Z_{pt})' \right) \\ \hat{\Sigma}_e &= \frac{1}{T} \sum_{t=1}^T e_t e_t' \end{aligned}$$

El resultado 3.32 es utilizado para conseguir el estadístico de prueba denominado **Estadístico Traza**, el cual se obtiene a través de un cociente de verosimilitudes y ayuda a identificar cuántas relaciones de cointegración hay en el modelo de corrección del error.

Para obtener el siguiente estadístico de prueba, es necesario hacer pequeñas modificaciones a la metodología mostrada anteriormente. Evidentemente, deberán otorgar el mismo resultado, ya que parten de la misma función de verosimilitud, pero se ataca el problema desde enfoques diferentes.

La primera de las modificaciones consiste en el procedimiento a seguir para obtener una estimación de  $\hat{\alpha}$ . Para ello, se retoma la reparametrización  $(W_{0t} - \Pi W_{pt}) = \Gamma W_{1t} + e_t$  con  $W_{0t} = \nabla \bar{Z}_t$ ,  $W_{pt} = \bar{Z}_{t-p}$  y  $W_{1t} = (\nabla \bar{Z}_{t-1}, \nabla \bar{Z}_{t-2}, \dots, \nabla \bar{Z}_{t-p+1})'$ . De esta relación se obtienen los residuos  $R_{0t}$  y  $R_{1t}$  de tal suerte que al sustituir en 3.24 se deriva la siguiente función de verosimilitud:

$$\ln(L(\alpha, \beta, \Sigma_e)) = -\frac{Tk}{2} \ln(2\pi) - \frac{t}{2} \ln |\Sigma_e| - \frac{1}{2} \sum_{t=1}^T (R_{0t} - \alpha\beta' R_{pt})' \Sigma_e^{-1} (R_{0t} - \alpha\beta' R_{pt}) \quad (3.33)$$

otra forma de expresar esta última ecuación es

$$R_{0t} = \alpha\beta' R_{1t} + \hat{e}_t \quad (3.34)$$

dicha regresión tiene la misma función de verosimilitud que en 3.33, sin embargo, su estimación no es como en regresión usual, ya que  $\alpha\beta'$  no es de rango completo. Por lo tanto, la estimación de esta regresión debe ser a través de mínimos cuadrados generalizados, otorgando los siguientes estimadores de  $\alpha$  y  $\Sigma_e$  en función de  $\beta$ .

$$\hat{\alpha}(\beta) = S_{p0}\beta(\beta' S_{pp}\beta)^{-1} \quad (3.35)$$

$$\widehat{\Sigma}_e(\beta) = S_{00} - \hat{\alpha}(\beta)(\beta' S_{pp}\beta)\hat{\alpha}(\beta)' \quad (3.36)$$

sustituyendo estos valores en 3.33 encontramos que la función de verosimilitud concentrada únicamente en  $\beta$  es

$$L_{cc}^{-\frac{2}{T}}(\beta) = |S_{00}| \frac{|\beta' (S_{pp} - S_{p0}S_{00}^{-1}S_{0p}) \beta|}{|\beta' S_{pp}\beta|} \quad (3.37)$$

Por lo tanto, nos enfrentamos al mismo problema que en la ecuación 3.31 y el valor que maximiza a la función de verosimilitud en este caso está dado por

$$L_{cc}^{-\frac{2}{T}} = |S_{00}| \prod_{i=0}^r (1 - \hat{\lambda}_i) \quad (3.38)$$

a partir del cual se obtendrá el **estadístico del mayor valor propio** mediante un cociente de verosimilitudes tal y como se muestra en la siguiente sección.

Ambos resultados pueden ser fácilmente reproducidos incluyendo términos determinísticos en el modelo, es importante tener en consideración esta notación, ya que en la práctica la presencia de estos términos resulta ser de uso frecuente.

### 3.2.3. Contrastes sobre el rango de cointegración

El método que brinda resultados razonables para contrastar el rango de  $\Pi$  es el denominado cociente de verosimilitudes, el cual parte de un valor inicial  $r^*$  que se supone, es el número de relaciones de cointegración inherentes al modelo.

Por un lado, para el estadístico traza se define a  $H(r^*)$  como la hipótesis nula de que el rango de  $\Pi$  es menor o igual a  $r^*$ , posteriormente se realiza el siguiente proceso iterativo de pruebas de hipótesis:

$$\begin{aligned}
 H_0^1 : rank(\Pi) \leq r^* \quad vs \quad H_a^1 : rank(\Pi) = 1, 2, \dots, k \\
 H_0^2 : rank(\Pi) \leq r^* + 1 \quad vs \quad H_a^2 : rank(\Pi) = 1, 2, \dots, k \mid rank(\Pi) \not\leq r^* \\
 \vdots \\
 H_0^i : rank(\Pi) \leq r^* + i \quad vs \quad H_a^i : rank(\Pi) = 1, 2, \dots, k \mid rank(\Pi) \not\leq r^* + i - 1
 \end{aligned}$$

La prueba de razón de verosimilitud de la  $i$ -ésima hipótesis nula se basa en el estadístico de prueba llamado por Johansen **estadístico de la traza**, el cual se obtiene al realizar el cociente del logaritmo de la función de verosimilitud bajo la hipótesis nula sobre el logaritmo de la función de verosimilitud bajo la hipótesis alternativa ( $rank(\Pi) = k$ ), de manera que se obtiene

$$\begin{aligned}
\lambda_{traza} &= T \left( \sum_{i=1}^r \ln(1 - \hat{\lambda}_i) - \sum_{i=1}^k \ln(1 - \hat{\lambda}_i) \right) \\
&= -T \sum_{i=r+1}^k \ln(1 - \hat{\lambda}_i)
\end{aligned} \tag{3.39}$$

En consecuencia, dado que se buscan los valores propios que sean significativamente distintos de cero, se tendrá que mientras más alejado esté el valor propio del cero, el valor  $\ln(1 - \hat{\lambda}_i)$  será más pequeño y por lo tanto el valor del estadístico traza será más grande.

Por otro lado, para realizar el contraste del mayor valor propio se realiza de manera análoga un proceso iterativo de pruebas de hipótesis y también un cociente de verosimilitudes para obtener el **estadístico del mayor valor propio**. En este caso, la hipótesis nula propone que el rango de cointegración de la matriz  $\Pi$  es exactamente  $r^*$ , de manera que las iteraciones para contrastar su rango se dan de la siguiente manera:

$$\begin{aligned}
H_0^1 : \text{rank}(\Pi) = r^* \quad \text{vs} \quad H_a^1 : \text{rank}(\Pi) = r^* + 1 \\
H_0^2 : \text{rank}(\Pi) = r^* + 1 \quad \text{vs} \quad H_a^2 : \text{rank}(\Pi) = r^* + 2 \mid \text{rank}(\Pi) \neq r^* \\
\vdots \\
H_0^i : \text{rank}(\Pi) \leq r^* + i \quad \text{vs} \quad H_a^i : \text{rank}(\Pi) = 1, 2, \dots, k \mid \text{rank}(\Pi) \neq r^* + i - 1
\end{aligned}$$

Por consiguiente, el contraste de razón de verosimilitudes es

$$\begin{aligned}
\lambda_{max} &= T \left( \sum_{i=1}^r \ln(1 - \hat{\lambda}_i) - \sum_{i=1}^{r+1} \ln(1 - \hat{\lambda}_i) \right) \\
&= -T \ln(1 - \hat{\lambda}_{r+1})
\end{aligned} \tag{3.40}$$

de igual manera, valores propios cercanos a cero implican valores pequeños del estadístico  $\lambda_{max}$ . Los valores críticos para ambos estadísticos

### 3.2. EL CONCEPTO DE COINTEGRACIÓN: CASO 2+ SERIES (VEC)57

fueron desarrollados a través de un estudio de simulación por Johansen y Juselius (1990), la tabla se presenta al final de este documento. La distribución asintótica de estos estadísticos depende de la forma de los términos determinísticos presentes en el modelo y del número de componentes no estacionarios bajo la hipótesis nula, es decir, si la hipótesis nula supone  $r$  relaciones de cointegración, entonces habrá  $(n - r)$  componentes no estacionarias.





# Capítulo 4

## Aplicación

### 4.1. Rango de Cointegración

En esta sección se llevarán a cabo las pruebas de orden de integración sobre un contexto multivariado para analizar la existencia de una relación estable de largo plazo entre factores que denotan la capacidad productiva de la economía mexicana, tales como el Producto Interno Bruto y la Recaudación Impositiva, con factores que reflejan el impacto de las políticas públicas en la sociedad mexicana tales como el Desempleo, Gasto en Salud Pública y el Gasto en Educación Pública. Recordemos que las series tienen una periodicidad anual desde 1991 hasta 2016 y en el caso particular del Producto Interno Bruto, se encuentra expresado a precios constantes y en logaritmo natural. Con base en los resultados mostrados en la sección 2.2 en la cual se exhibe que las series incluidas en este análisis son integradas de orden 1,  $I(1)$ , nos permite realizar la búsqueda de combinaciones lineales estacionarias (rango de cointegración) a través de las pruebas de hipótesis de la traza y del máximo eigenvalor para determinar el rango de la matriz  $\Pi$  del modelo VECM especificado de la siguiente manera:

$$\nabla Z_t = \mu + \Pi Z_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \nabla Z_{t-i} + e_t \quad (4.1)$$

donde el vector  $Z_t$  contiene a los elementos  $(\ln(PIB), Edu, Salud, Rec)'$ , el vector  $\mu$  es un vector de constantes y el proceso del error de 4 dimensiones  $e_t$  se asume que es  $N(0, \sigma)$  *i.i.d* para  $t = 1, \dots, T$ . De tal manera que la prueba de hipótesis  $H_i(r) : \Pi = \alpha\beta'$  evalúa si  $\Pi$  es de rango reducido probando los estadísticos del máximo eigenvalor y de la traza. A continuación, se muestran los resultados de las pruebas:

Máximo Eigenvalor				
Prueba	Estadístico	10 %	5 %	1 %
$r \leq 3$	0.2680195	6.50	8.18	11.65
$r \leq 2$	8.2179845	12.91	14.90	19.19
$r \leq 1$	18.2540706	18.90	21.07	25.75
$r \leq 0$	34.9647411	24.78	27.14	32.14

Tabla 4.1: Rango de Cointegración: estadístico del máximo eigenvalor

Por un lado, considerando el estadístico del máximo eigenvalor, el cual evalúa la prueba de hipótesis  $H_0 : rank(\Pi) = r$  vs  $H_a : rank(\Pi) = r + 1$ , la hipótesis de no cointegración puede ser rechazada a un nivel del 1% de significancia. Sin embargo, la prueba de existencia de 1 relación de cointegración vs 2 relaciones de cointegración no puede ser rechazada a un nivel de 5%. Por lo tanto, de acuerdo con el estadístico del máximo eigenvalor existe únicamente una relación de cointegración entre las series.

Traza				
Prueba	Estadístico	10 %	5 %	1 %
$r \leq 3$	0.2680195	6.50	8.18	11.65
$r \leq 2$	8.4860040	15.66	17.95	23.52
$r \leq 1$	26.7400746	28.71	31.52	37.22
$r \leq 0$	61.7048157	45.23	48.28	55.43

Tabla 4.2: Rango de Cointegración: estadístico de la Traza

Por otro lado, en la tabla anterior se pueden observar los resultados de la prueba de la traza que evalúa la hipótesis  $H_0 : rank(\Pi) \leq r$ , a partir de la cual se pueden reforzar las conclusiones obtenidas en la prueba del máximo eigenvalor. Es decir, la hipótesis de no cointegración puede ser rechazada a un nivel de significancia del 1% y la prueba  $rank(\Pi) \leq 1$  no puede ser rechazada a un nivel del 10% de significancia, por lo cual se acepta la hipótesis nula concluyendo que el espacio de cointegración es de  $r = 1$  y por lo tanto, únicamente existe una relación de cointegración estacionaria entre las series del respectivo análisis. Se ha considerado también, si esta conclusión podría ser errónea debido a la cercanía de los eigenvalores, por lo que se muestran a continuación:

Eigenvalores			
$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$
0.76703434	0.53260767	0.28994725	0.01110535

Tabla 4.3: Eigenvalores de la matriz  $\Pi$ 

En donde se puede observar que  $\lambda_2$  no está lo suficientemente cerca de  $\lambda_1$  como para sesgar las pruebas de hipótesis. Además, Johansen y Juselius [1992] analizan las matrices  $\hat{\alpha}$  y  $\hat{\beta}$ , así como las relaciones de cointegración  $\hat{\beta}'Z_t$ . La matriz de cointegración ha sido normalizada al logaritmo natural del PIB y por lo tanto, la matriz  $\hat{\alpha}$  ha sido ajustada con dicha normalización.

Matriz $\hat{\beta}$				
	ln_pib_cte.l1	edu.l1	salud.l1	rec_imp.l1
ln_pib_cte.l1	1.0000000	1.00000000	1.0000000	1.0000000
edu.l1	-0.2734779	-0.12722367	-0.8329062	-0.4709689
salud.l1	-0.3187240	-0.14485981	1.3555847	-0.7250595
rec_imp.l1	0.2359534	-0.02050015	-0.2056044	-0.4166424

Tabla 4.4: Eigenvectores

En la matriz  $\hat{\beta}$  se pueden observar los vectores de cointegración, de tal manera que la primera columna corresponde al vector de cointegración asociado con el eigenvalor más grande.

Matriz $\hat{\alpha}$				
	ln_pib_cte.l1	edu.l1	salud.l1	rec_imp.l1
ln_pib_cte.l1	0.02575836	-0.1945384	-0.01292212	-0.004630415
edu.l1	0.81182331	1.8215325	0.14885276	0.008440650
salud.l1	-0.10602593	0.7318614	-0.18202612	-0.004282401
rec_imp.l1	-2.42505459	8.1784032	0.50872868	-0.058362273

Tabla 4.5: Velocidades de Ajuste

Al observar la matriz  $\hat{\alpha}$  se puede concluir que las velocidades de ajuste de las relaciones de cointegración parecen ser distintas de cero, esto significa que la relación de cointegración juega un papel importante en la dinámica de corto plazo de las series, aunque aún falta validar que dicha relación de cointegración es estadísticamente significativa en cada serie. Finalmente, podemos analizar la relación de cointegración de manera visual:

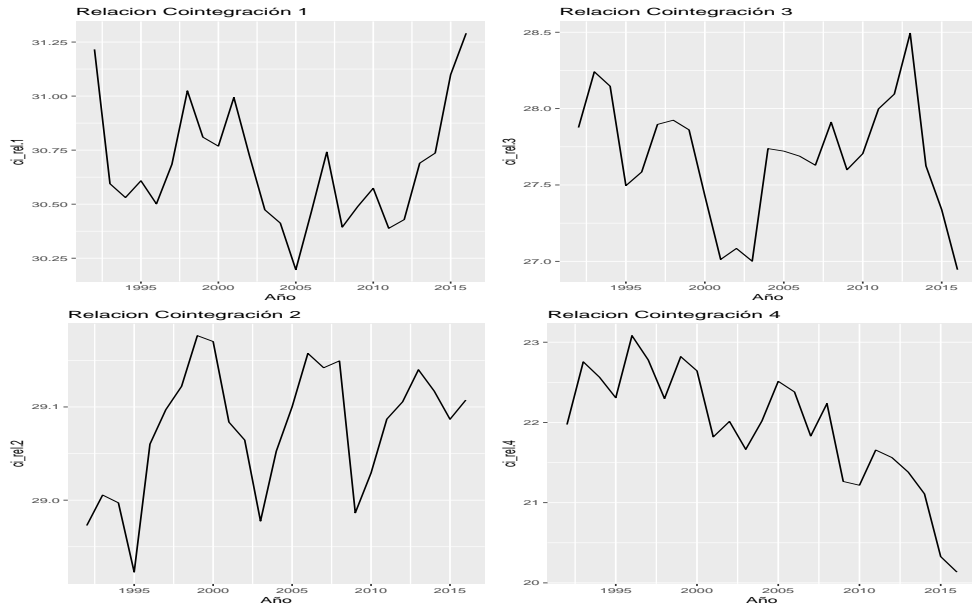


Figura 4.1: Gráfica Relaciones de Cointegración

Debido a que el rango de la matriz  $\Pi$  fue  $r = 1$ , entonces la primera relación de cointegración debería comportarse como un proceso estacionario, lo cual se ratifica en la gráfica anterior. Por lo tanto, basado en los resultados de las pruebas de hipótesis, los elementos de la matriz  $\hat{\alpha}$  y  $\hat{\beta}$  y la forma de las trayectorias de las relaciones de cointegración, se puede concluir que existe únicamente una relación de cointegración entre las series.

## 4.2. Estimación del Modelo

Con los análisis realizados hasta este momento, se cuenta con dos de los tres factores necesarios para realizar la estimación de un modelo VEC. Es decir, se requiere conocer el orden de integración de las series, el número de retrasos a incluir en el modelo y finalmente conocer la existencia de una o más relaciones de cointegración entre las series. Por lo que en primer lugar y con base en lo exhibido en el capítulo 2.2 podemos concluir que a partir del

cálculo de los estadísticos básicos, las gráficas de las series de tiempo, así como las pruebas de raíz unitaria vía ADF hemos logrado concluir que las series son integradas de orden uno  $I(1)$ . En segundo lugar, para la estimación del modelo VEC se considera que un valor apropiado de rezagos es de  $p = 1$  debido a que en la medida en que se agregan más rezagos al modelo, el número de grados de libertad se reduce drásticamente. Esto es causado por la pérdida de datos que ocurre al aplicar el operador diferencia sobre las series y por el incremento en el número de parámetros adicionales a estimar. De manera ilustrativa analizaremos la pérdida de datos cuando se estima un modelo VEC con  $p = 1$  y un modelo VEC con  $p = 2$ , es decir, si consideramos que cada una de las series cuenta con  $n = 26$  datos y que  $p = 1$  o dicho de otra manera, que uno de los componentes que intervienen en el modelo es  $\nabla z_{t-1}$  cuya construcción implica una pérdida de 2 datos del total, uno por el retraso de la serie y otro por el operador diferencia, adicionalmente en un modelo VEC con  $p = 1$  se deben estimar 6 parámetros como veremos a continuación, por lo que los grados de libertad finales son  $gl = 26 - 2 - 6 = 18$ . En el caso del modelo VEC con  $p = 2$  el componente a construir es  $\nabla z_{t-2}$  que representa perder 3 datos y los parámetros a estimar son 10 por lo que los grados de libertad son  $gl = 26 - 3 - 10 = 13$ , lo cual implica una reducción del 28% en los grados de libertad y que para la estimación se utilice únicamente el 57% de la información, lo cual tiene un impacto negativo en el proceso de estimación. Finalmente, con base en lo reportado en el capítulo 4.1 se determinó que únicamente existe una relación de cointegración entre las series de tiempo.

Por lo tanto, ya nos encontramos en posición para realizar la estimación del modelo VEC bajo la restricción de la existencia de una única relación de cointegración. De acuerdo con lo mencionado anteriormente, se agregará un solo retraso de las variables al modelo y los resultados se muestran a continuación:

VEC Model				
	$\nabla \ln\_pib\_cte_t \sim$	$\nabla edu_t \sim$	$\nabla salud_t \sim$	$\nabla rec\_imp_t \sim$
ect1	0.03 (0.03)	0.81 ** (0.22)	-0.11 (0.14)	-2.43 * (0.93)
constant	-0.75 (1.02)	-24.68 ** (6.83)	13.22 (4.28)	73.99 * (28.43)
$\nabla \ln\_pib\_cte_{t-1}$	-0.25 (0.25)	-2.43 (1.66)	1.05 (1.04)	14.00 (6.91)
$\nabla edu_{t-1}$	-0.03 (0.03)	-0.29 (0.19)	0.21 (0.12)	0.99 (0.80)
$\nabla salud_{t-1}$	-0.06 (0.05)	0.08 (0.34)	0.15 (0.21)	-2.00 (1.42)
$\nabla rec\_imp_{t-1}$	-0.00 (0.01)	-0.15 * (0.06)	-0.03 (0.03)	0.35 (0.23)
$R^2$	0.48	0.52	0.30	0.37
Adj. $R^2$	0.31	0.37	0.06	0.17
Num. obs.	24	24	24	24
RMSE	0.03	0.22	0.14	0.93
*** p < 0.001	** p < 0.01	* p < 0.05		

Tabla 4.6: VECM

En la tabla anterior se muestra la ecuación estimada en el modelo para cada una de las series. Como se puede observar, los retrasos de cada una de las series seleccionadas parecen no ser significativas en el modelo VECM para las series del PIB, Gasto en Salud y Recaudación Impositiva a diferencia de la serie del Gasto en Educación, en la cual el retraso de la recaudación impositiva parece ser significativa y con coeficiente negativo. Esto quiere decir que ante una reducción de la recaudación impositiva se observará un impacto positivo en el gasto en educación. Si bien esto va en contra de lo que la lógica nos indica, los datos respaldan el signo del coeficiente en el modelo. A continuación se muestra la relación que guarda la serie  $\nabla edu_t$  con  $\nabla rec\_imp_{t-1}$  en donde se puede observar que en general, cuando ocurre una reducción en la recaudación

impositiva, el gasto en educación pública incrementa.

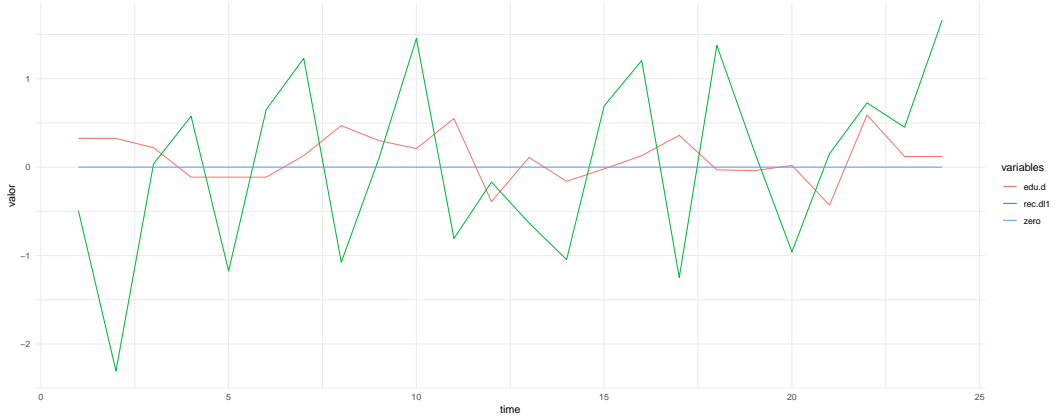


Figura 4.2: Relación  $\nabla edu_t$  con  $\nabla rec\_imp_{t-1}$

Interpretar así este coeficiente resulta ser muy arriesgado y podría llevarnos a conclusiones erróneas tal y como a la que llegamos en el párrafo anterior. Se debe tener presente que el efecto de la variable  $rec\_imp_{t-1}$  se encuentra tanto en el componente  $\nabla rec\_imp_{t-1}$  como en la relación de cointegración  $ect1$ , por lo que si desagregamos ambos componentes y observamos cómo quedaría el impacto de la recaudación gubernamental sobre el gasto en educación pública en la ecuación tendremos lo siguiente:

$$\nabla \hat{edu}_t = -24.68 + 0.07rec\_imp_{t-1} + 0.12rec\_imp_{t-2} + 0.812ln\_pib\_cte_{t-1} - 0.22edu_{t-1} - 0.26salud_{t-1} \quad (4.2)$$

De tal manera que una vez desagregados los componentes  $\nabla rec\_imp_{t-1}$  y  $ect1$  y ajustando el coeficiente de la recaudación impositiva se puede apreciar que el efecto es el esperado. Es decir, que ante un incremento en la recaudación gubernamental, se observará un incremento en el gasto en educación.

A continuación, se muestra la matriz de cointegración estimada  $\hat{\beta}$  y la matriz de ajuste  $\hat{\alpha}$ . Debido a que el coeficiente de velocidad de ajuste  $\alpha$  resultó ser significativo en el Modelo VEC para la ecuación del gasto en educación y



recaudación impositiva, se puede decir que tanto el gasto en educación como la recaudación impositiva responden ante desviaciones en el equilibrio de largo plazo y la proporción de desequilibrio que será corregida dentro de un periodo de tiempo será de 0.81 y de -2.43 respectivamente.

	Matriz $\hat{\beta}, \hat{\alpha}$			
	$\ln\_pib\_cte_{t-1}$	$edu_{t-1}$	$salud_{t-1}$	$rec\_imp_{t-1}$
$\hat{\beta}$	1.0000000	-0.2734779	-0.3187240	0.2359534
$\hat{\alpha}$	0.03	0.81 **	-0.11	-2.43 *

Tabla 4.7: VECM: Eigenvector normalizado y velocidades de ajuste

La matriz de cointegración estimada  $\hat{\beta}$  que representa la única relación de cointegración en el modelo también puede expresarse de la siguiente manera:

$$\ln\_PIB = 0.27 * edu + 0.32 * salud - 0.24 * rec\_imp \quad (4.3)$$

o bien

$$(\ln\_PIB + 0.24 * rec\_imp) - (0.27 * edu + 0.32 * salud) = 0 \quad (4.4)$$

Vale la pena detenerse un momento para analizar lo hasta ahora comentado, por un lado, podemos concluir que la dinámica temporal de cada una de las series en general no se ve afectada por los retrasos de las variables en incrementos  $\nabla Z_{t-1}$  indicando que no hay una relación directa entre las variables, excepto para el gasto en educación. Por otro lado, tenemos evidencia de la existencia de una combinación lineal entre las series del producto interno bruto, el gasto en salud pública, el gasto en educación pública y la recaudación impositiva (relación de cointegración) que sí afecta la dinámica temporal de las series, sin embargo, esta relación de cointegración solo resultó ser significativa para el gasto en educación pública y la recaudación fiscal. Estas conclusiones resultan interesantes dado que el producto interno bruto al depender de factores externos y no únicamente

nacionales tiene prácticamente vida propia en el modelo, es decir, su dinámica no se ve impactada ante la ruptura del equilibrio en la relación de cointegración y que los efectos de las políticas públicas se ven reflejados en términos de recaudación fiscal y gasto en educación. Finalmente, cabe resaltar que el gasto en salud al tratarse de un sector muy delicado, tampoco parece reaccionar ante la ruptura en el equilibrio o expresado de otra manera, es un sector tan importante para la población que la inversión destinada para ello no irá en función del crecimiento o recesión del país.

Si analizamos particularmente el efecto que la relación de equilibrio, bajo el enfoque de la ecuación 4.4, tiene sobre las series del gasto en educación y la recaudación impositiva se puede apreciar que el modelo ha capturado correctamente la dinámica de acuerdo con lo que la lógica indica, ya que si la relación de cointegración rompe su equilibrio de manera negativa, es decir, mayor gasto que recaudación entonces la serie del gasto en educación deberá corregir su dinámica y presentar una reducción en el largo plazo, el efecto tenderá a ser con una baja velocidad de reacción comparado con la reacción que tendría la dinámica de la serie de recaudación impositiva, la cual tras una ruptura del equilibrio de manera negativa mostrará una corrección a la alza mucho más veloz, es decir, el gobierno comenzará a recaudar más a través de los impuestos. La respuesta de las variables se puede interpretar de la misma manera cuando la relación de equilibrio se rompe de manera positiva.

Lo anterior puede ser representado con mayor detalle si observamos la serie de tiempo de la relación de cointegración para analizar las medidas correctivas que ha tomado el gobierno que provocan el retorno de la ecuación a su punto de equilibrio, para ello hemos considerado aquellos puntos que se encuentran por encima o por debajo de una banda de  $\pm 1$  desviación estándar que hemos establecido. De tal manera que podemos observar en los periodos posteriores qué cambios en las series lograron el retorno al equilibrio.



Figura 4.3: Relación de cointegración

Corrección Equilibrio					
Año	ect1	ln_pib_cte.l1	rec_imp.l1	edu.l1	salud.l1
1997	30.685	30.065	9.699	3.64	2.11
1998(↑)	31.024	30.115	10.929	3.53	2.21
1999	30.810	30.143	9.853	3.66	2.06
2000	30.768	30.191	9.950	4.13	2.01
2001(↑)	30.994	30.187	11.404	4.43	2.11
2002	30.726	30.186	10.597	4.64	2.17
2004	30.412	30.239	9.796	4.80	2.59
2005(↓)	30.197	30.262	8.751	4.91	2.47
2006	30.464	30.306	9.443	4.75	2.42
2014	30.736	30.449	11.327	5.33	2.91
2015(↑)	31.099	30.481	12.989	5.45	3.00

Tabla 4.8: Análisis Relación de Cointegración Histórico

En la tabla anterior, se puede observar que la relación de cointegración se encuentra por fuera de la banda de equilibrio durante los periodos de

1997-1999, 2000-2002, 2004-2006 y 2014-2015 en donde la relación en algunos casos se rompe por encima de la banda y en otros casos por debajo de la misma. Esto nos permitirá interpretar con más detalle la relación de equilibrio, ya que contextualizaremos cada uno de estos periodos con los hechos históricos que posiblemente hayan generado la ruptura del equilibrio, así como las decisiones que hayan provocado el retorno al equilibrio. Para esto, es importante recordar que por lo establecido en la ecuación 4.4 tanto la recaudación fiscal como el producto interno bruto impactan de manera positiva a la relación de cointegración y que el gasto en salud y el gasto en educación pública impactan de manera negativa.

En el periodo de 1997-1999 la relación de cointegración se encuentra por encima de la banda de equilibrio y en el año posterior ésta regresa dentro de la banda, por lo que podríamos decir que la combinación de factores que producen la salida del equilibrio, está dado por un incremento tanto en el PIB como en la recaudación impositiva, en conjunto con una reducción del gasto en educación pública; mientras que el retorno de la ecuación al equilibrio está dada por una reducción en la recaudación impositiva y un incremento en el gasto en educación pública. Recordemos que en este periodo sucedieron hechos relevantes que de alguna manera soportan lo observado en nuestro modelo; ya que en este periodo por un lado, el sistema tributario mexicano sufrió una de las modificaciones más drásticas de la historia, pues el ISR, el cual representa junto con el IVA aproximadamente el 85 % de los ingresos tributarios y siendo los principales objetos de las diversas reformas tributarias, pasó de un 34 % a un 35 % para personas morales y de un 35 % a un 40 % para personas físicas. Por otro lado, la crisis económica de 1994-1995 representó un incremento de los costos financieros derivados del quebranto del sistema financiero nacional. La necesidad de reconocer los pasivos del rescate bancario hacia 1998, junto con los programas de apoyo a los deudores y ahorradores de la banca, incrementó nuevamente el costo financiero reduciendo así el gasto público destinado al resto de sectores.

En el periodo de 2000-2002 se aprecia también que la relación de cointegración se encuentra por fuera de la banda de equilibrio y que los factores que provocan el retorno al mismo, están dados nuevamente por una

reducción de la recaudación impositiva e incrementos en la inversión por parte del gobierno para el sector salud y el sector educativo. En este caso, México resintió los efectos de la desaceleración económica global, particularmente la de los Estados Unidos provocando una reducción en el PIB, no obstante, la inflación disminuyó y prevaleció un entorno financiero estable. En ese mismo año, el Banco de México publicó que se esperaba una recuperación de la economía mundial en el 2002, aunque resultaba poco probable que la actividad económica de los principales socios comerciales de México regresaran a los niveles observados previos a la recesión, de tal manera que para reactivar la economía resultaba imperativo reforzar los elementos internos que le permitieran a México alcanzar tasas de crecimiento elevadas y sostenibles en un entorno internacional menos propicio que el de años anteriores. Entre dichos elementos, se mencionaron la reforma del sector eléctrico, reforma laboral, el estímulo a la inversión en infraestructura, tecnología, educación y salud, entre otras. Las acciones tomadas por el gobierno en torno a este último elemento se ven reflejadas en el incremento en el gasto en educación y salud, incremento que provoca el retorno de la relación de cointegración a la banda de equilibrio.

En el año 2005 a diferencia de los dos casos anteriores, la relación de cointegración se encuentra por debajo de la banda de equilibrio por lo que este caso tiene un particular interés. En primer lugar, es necesario recordar que durante el 2005 el PIB presentó un incremento, aunque dicho incremento no fue el esperado por los analistas económicos. En segundo lugar, el gasto público presupuestario reflejó un incremento de 4.8 por ciento. Este incremento en el gasto público se aprecia en la alza del gasto en educación, que aunado al decremento en la recaudación gubernamental y que se trata de un año previo a las elecciones del 2006, obliga a la relación de cointegración a salir del equilibrio. Posteriormente, el año 2006 fue un año en donde la economía mundial mantuvo un crecimiento favorable y en particular en la economía mexicana tuvo un crecimiento económico, producto de un incremento del producto interno bruto, un crecimiento de la actividad industrial y de construcción que implicaron un aumento significativo en la creación de empleos en el sector formal de la economía, mismo crecimiento

que se tradujo en una mayor recaudación fiscal, provocando así el retorno de la relación de cointegración a la banda de equilibrio.

Finalmente, en el último punto de la serie observada de la relación de cointegración se observa que se encuentra nuevamente por encima de la banda de equilibrio, sin embargo, con base en la experiencia previa podemos darnos una idea de qué es lo que deberíamos observar para lograr nuevamente el equilibrio, es decir, sería natural pensar que lo próximo a observarse será una reducción de la recaudación gubernamental junto con un incremento en el gasto público. Además, podríamos apoyarnos en el parámetro de velocidad de ajuste que resultó significativo para estimar el tiempo en el que dichas políticas podrían verse materializadas.

# Conclusiones

En secciones anteriores se analizó la existencia de una relación estable de largo plazo, así como la factibilidad de utilizar modelos de corrección de error (VECM) entre factores que denotan la capacidad productiva de la economía mexicana, tales como el Producto Interno Bruto y la Recaudación Impositiva, con otros factores que reflejan el impacto de las políticas públicas en la sociedad como puede ser el Gasto en Salud Pública y el Gasto en Educación Pública. En esta sección, se discuten tanto las implicaciones y el alcance de los resultados obtenidos, como el trabajo futuro derivado del análisis.

El primer resultado de interés que tiene un impacto directo en el análisis y que se encuentra fuera de nuestro alcance, consiste en la dificultad para obtener información histórica de datos públicos en México. Información del gasto en servicios públicos como la educación y la salud, junto con la recaudación impositiva deberían ser datos con suficiente profundidad histórica, con periodicidad mayor a la anual y de fácil acceso. No obstante, encontrar información estructurada resulta muy complicado en fuentes oficiales nacionales, mientras que fuentes internacionales como el banco mundial ofrecen una mejor perspectiva. Sin duda, se trata de un área con una gran oportunidad de mejora.

Las pruebas de estacionariedad de las series en conjunto con las pruebas de cointegración nos permiten cumplir con otro de los propósitos del presente trabajo, el cual consiste en analizar la posibilidad de realizar el ajuste de un modelo VECM con las series anteriormente descritas. Si bien una vez ajustado el modelo, los resultados mostrados por el mismo fueron acordes con

lo indicado por la lógica (mayor recaudación fiscal y mayor crecimiento del PIB implican un mayor gasto en servicios públicos) el modelo conduce a otras dos conclusiones de importancia. En primer lugar, el modelo no solo nos muestra que las relaciones son tal y como se esperan sino que nos presenta la posibilidad de cuantificar y de medir dichas relaciones y efectos, la velocidad de ajuste que tendrán las series en caso de romperse el equilibrio, así como la tendencia que deberá llevar cada una de ellas para regresar al mismo. En segundo lugar, nos demuestra que a pesar de existir una relación de cointegración entre las series, ésta no resulta ser significativa para cada una de ellas, ya que existen series que tienen vida propia y no reaccionan ante la ruptura del equilibrio de la relación de cointegración. En el caso del PIB, esto puede ser causado porque la dinámica de las series a su vez son afectadas por otras series exógenas. En cuanto al gasto en salud, la serie parece no responder al desequilibrio, ya que se trata de un sector muy delicado para la sociedad que no puede estar indexado al crecimiento del país. Además, a pesar de que la economía presente incrementos, estos no se verán reflejados fácilmente en el gasto en salud mientras no exista una mejor distribución de los recursos eficientando los altos costos administrativos, no exista competencia en la provisión de los servicios ni competencia en la organización y administración del cuidado de la salud y no exista un modelo de remuneración de los servicios de salud que reflejen las actividades generadas.

Finalmente, como trabajo futuro, se propone el análisis de la relación de cointegración conforme se publiquen más datos de las series, ya que por ahora la relación de cointegración se mantiene fuera de la banda establecida de  $\pm 1$  desviación estándar, por lo que resulta de interés la evolución de las mismas, así como las decisiones gubernamentales que provoquen el regreso a la banda de equilibrio.



# Bibliografía

- [1] D. N. AGUILERA, *10 puntos para entender el gasto en salud en México: En el marco de la cobertura universal*, México Evalúa, (2011).
- [2] B. DE MÉXICO, *Resumen informe anual 1998, 1999, 2001*, Banco de México, (1998-2001).
- [3] W. ENDERS, *Applied Econometric Time Series*, John Wiley & Sons, Inc., 2014.
- [4] R. F. ENGLE AND C. W. J. GRANGER, *Co-integration and error correction: Representation, estimation and testing*, *Econometrica*, (1987).
- [5] J. C. T. GARCÍA, E. C. R. HERNÁNDEZ, AND H. R. BOLÍVAR, *Análisis de la histéresis del desempleo en México ante shocks macroeconómicos*, Instituto Politécnico Nacional, (2016).
- [6] D. N. GUJARATI, *Basic Econometrics*, Mc. Graw Hill, 2008.
- [7] R. C. HILL, W. E. GRIFFITHS, AND G. C. LIM, *Principles of econometrics*, John Wiley & Sons, Inc., 2011.
- [8] R. HYNDMAN, G. ATHANASOPOULOS, C. BERGMEIR, G. CACERES, L. CHHAY, M. O'HARA-WILD, F. PETROPOULOS, S. RAZBASH, E. WANG, AND F. YASMEEN, *forecast: Forecasting functions for time series and linear models*, 2018. R package version 8.4.

- [9] R. J. HYNDMAN AND Y. KHANDAKAR, *Automatic time series forecasting: the forecast package for R*, Journal of Statistical Software, 26 (2008), pp. 1–22.
- [10] S. JOHANSEN, *Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models*, Econometrica, (1991).
- [11] —, *A small sample correction for the test of cointegration rank in the vector autoregressive model*, Econometrica, (2002).
- [12] E. LORÍA AND M. RAMOS, *La ley de okun: una relectura para méxico, 1970-2004*, Estudios Económicos, (2007).
- [13] H. LÜKTHEPOHL, *New Introduction to Multiple Time Series Analysis*, Springer, 2005.
- [14] T. C. MILLS, *The Econometric Modelling of Financial Time Series*, Cambridge, 1999.
- [15] M. P. MURRAY, *A drunk and her dog: An illustration of cointegration and error correction*, American Statistical Association, (1994).
- [16] B. PFAFF, *Analysis of Integrated and Cointegrated Time Series with R*, Springer, 2008.
- [17] B. PFAFF, *Analysis of Integrated and Cointegrated Time Series with R (urca package)*, Springer, New York, second ed., 2008. ISBN 0-387-27960-1.
- [18] L. SANTIBÁÑEZ, M. CAMPOS, AND B. JARILLO, *10 puntos para entender el gasto educativo en méxico: Consideraciones sobre su eficiencia*, México Evalúa, (2011).

# Apéndice A

## Código

Tanto las bases insumo, como el código que reproducen los resultados mostrados en este desarrollo se encuentran en el siguiente repositorio de Github:

[https://github.com/omisimo/Thesis\\_VEC\\_Models](https://github.com/omisimo/Thesis_VEC_Models)



# Apéndice B

## Extracción de las Series

Las series se han extraído de la siguiente manera:

### 1. PIB a precios constantes y corrientes

- Los datos referentes al PIB a precios corrientes en moneda local, se encuentran disponibles en el catálogo del Banco Mundial con el código **NY.GDP.MKTP.CN**. En el cual se pueden encontrar los detalles de la construcción de la serie, así como sus supuestos y limitaciones
- Los datos referentes al PIB a precios constantes en moneda local, se encuentran disponibles en el catálogo del Banco Mundial con el código **NY.GDP.MKTP.KN**. En el cual se pueden encontrar los detalles de la construcción de la serie, así como sus supuestos y limitaciones

### 2. Desempleo

- La serie se puede consultar en el banco de datos del Banco Mundial con el código **|SL.UEM.TOTL.NE.ZS**. En el cual se pueden encontrar los detalles de la construcción de la serie, así como sus supuestos y limitaciones

### 3. Gasto en Salud (%PIB)

- La serie en un inicio se podía consultar en el banco de datos del Banco Mundial con el código **SH.XPD.PUBL.ZS**. En el cual se podían encontrar los detalles de la construcción de la serie, así como sus supuestos y limitaciones
- Esta serie dejó de estar disponible en el Banco Mundial y por esta razón, la serie se obtuvo a partir de una descarga que el periódico Expansión realizó de la misma

#### 4. Gasto en Educación (%PIB)

- La serie se puede consultar en el banco de datos del Banco Mundial con el código **SE.XPD.TOTL.GD.ZS**. En el cual se pueden encontrar los detalles de la construcción de la serie, así como sus supuestos y limitaciones
- La serie descargada directamente del Banco Mundial presentaba valores faltantes para los años 1993, 1996, 1997, 2015 y 2016.
- Para los años 1993, 1996 y 1997, se ha hecho uso de la función `na.interp` de la librería de R `forecast` de Rob Hyndman, el cual usa una interpolación lineal de series no estacionarias
- Para los años 2015 y 2016 se ha utilizado la función `auto.arima` para ajustar un modelo ARIMA a la serie y posteriormente se ha hecho uso de la función `forecast` para pronosticar los dos periodos correspondientes a 2015 y 2016
- Los resultados del punto anterior, fueron un modelo ARIMA(0,1,0) con  $drift = 0.1224$  a partir del cual se realizó el pronóstico

#### 5. Recaudación Gubernamental (%PIB)

- La serie mostrada en este desarrollo se construyó a partir de los importes mostrados en la Ley de Ingresos de la Federación, en el apartado de recaudación impositiva total. Dicho importe se dividió entre el PIB a precios corrientes para cada año

Debido a que las series descargadas del Banco Mundial, en ocasiones, al momento de actualizarse sufren modificaciones históricas, para este desarrollo

se ha realizado una efecto foto de la descarga. Por esta razón, los datos mostrados en este trabajo podrían diferir a los descargados directamente del Banco Mundial y en fechas posteriores.