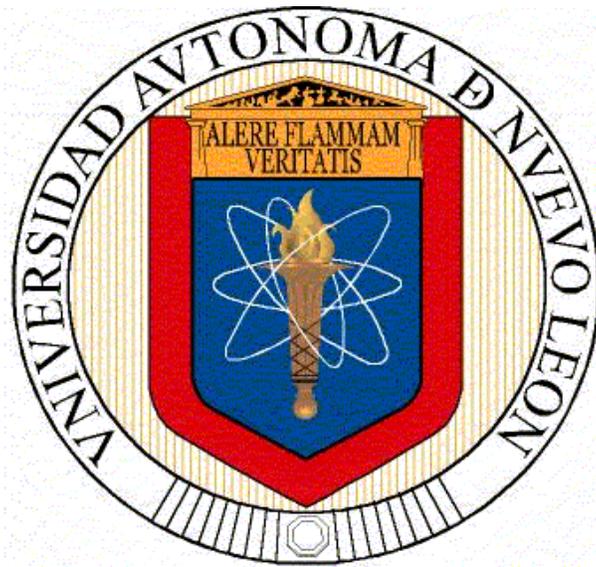


**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN  
FACULTAD DE ECONOMÍA**



**Un Modelo para la Obtención del PIB Trimestral de las Entidades  
Federativas de la República Mexicana: Aplicación del Método de  
Chow-Lin.**

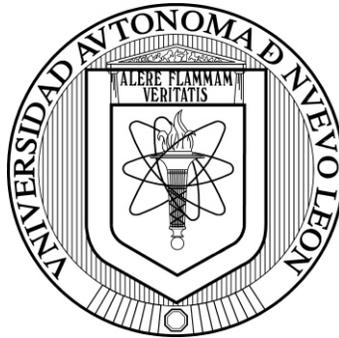
**Por**

**Luis Alberto Cepeda Villasana**

**Tesis presentada como requisito parcial para obtener el grado de  
Maestría en Economía con Orientación en Economía Industrial**

Agosto 2015

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN  
FACULTAD DE ECONOMÍA  
DIVISION DE ESTUDIOS DE POSGRADO**



**Un Modelo para la Obtención del PIB Trimestral de las Entidades  
Federativas de la República Mexicana:  
Aplicación del Método de Chow-Lin.**

**Por**

**LUIS ALBERTO CEPEDA VILLASANA**

**Tesis presentada como requisito parcial para  
obtener el grado de Maestría en Economía con  
Orientación en Economía Industrial**

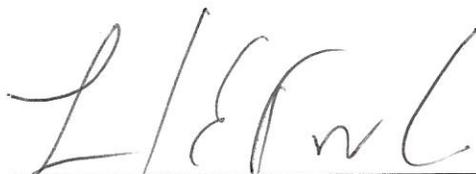
AGOSTO 2015

**“UN MODELO PARA ESTIMAR EL PIB TRIMESTRAL DE LAS  
ENTIDADES FEDERATIVAS DE LA REPÚBLICA MEXICANA:  
APLICACIÓN DEL MODELO DE CHOW-LIN”**

*Luis Alberto Cepeda Villasana*

**Aprobación de Tesis:**

**Asesor de la Tesis**



**DR. LEONARDO E. TORRE CEPEDA**



**DR. ERNESTO AGUAYO TÉLLEZ**



**M.C. PONCIANO MURRILLO DE LA TORRE**

**DR. PEDRO ANTONIO VILLEZCA BECERRA**  
Director de la División de Estudios de Posgrado  
De la Facultad de Economía, UANL  
Junio, 2017



FACULTAD DE ECONOMÍA  
DIV. ESTUDIOS DE POSGRADO

## **Índice**

### **Índice de tablas**

### **Índice de gráficos**

### **Introducción**

### **Capítulo 1**

<b>1. Antecedentes</b>	<b>4</b>
<b>1.1 Métodos triviales</b>	<b>5</b>
<b>1.2 Métodos que no utilizan indicadores</b>	<b>6</b>
<b>1.3 Métodos que utilizan indicadores relacionados</b>	<b>6</b>

### **Capítulo 2**

<b>2. Metodología</b>	<b>14</b>
<b>2.1 Estimación del PIB trimestral por entidad</b>	<b>14</b>
<b>2.2 Modelo de estimación</b>	<b>15</b>
<b>2.3 Descripción de la variable dependiente</b>	<b>17</b>
<b>2.4 Variables independientes</b>	<b>21</b>
<b>2.4.1 Total de trabajadores Asegurados en el IMSS</b>	<b>21</b>
<b>2.4.2 Construcción de la serie TAIMSS 1993-1996</b>	<b>24</b>
<b>2.4.3 Consumo de energía eléctrica por entidad federativa</b>	<b>30</b>
<b>2.4.4 Captación y empleados bancarios</b>	<b>31</b>

### **Capítulo 3**

<b>3. Estimación del Modelo</b>	<b>33</b>
<b>3.1 Aspectos de las series de tiempo</b>	<b>34</b>

<b>3.2 Análisis de correlación</b>	<b>35</b>
<b>Capítulo 4</b>	
<b>4. Resultados</b>	<b>48</b>
<b>4.1 Estimación del modelo General. Variable. Variable rateimss.</b>	<b>49</b>
<b>4.2 Estimación de Efectos Fijos.</b>	<b>50</b>
<b>4.3 Modelo general combinación de variables</b>	<b>52</b>
<b>4.4 Estimación de Efectos Fijo.</b>	<b>55</b>
<b>Conclusiones y recomendaciones</b>	<b>61</b>
<b>Referencias bibliográficas</b>	<b>66</b>
<b>Anexos</b>	<b>67</b>

## Índice de tablas

<b>2.1 Diferencia de la tasa de crecimiento para las entidades federativas base 1993 y 2008.</b>	<b>20</b>
<b>2.2. Estimaciones del Total de trabajadores de asegurados 1997.</b>	<b>23</b>
<b>2.4 Resultados de la regresión T_A vs. Tiempo.</b>	<b>27</b>
<b>2.5 Resultados de la regresión T_A vs. Tiempo. Ecuación logarítmica.</b>	<b>29</b>
<b>3.1 RatePIB y RateIMSS. Coeficientes de correlación por periodos.</b>	<b>36</b>
<b>3.2. Coeficiente de correlación RatePIB vs RateKwh entidad federativa. 2003-2012</b>	<b>38</b>
<b>3.3 RatePIB y RateCapt por entidad federativa. 2001-2010.</b>	<b>41</b>
<b>3.4 Correlación de la tasa de crecimiento del PIB vs. Tasa de crecimiento de las variables explicativas: rateimss, ratebank, ratecapt y ratekwh.</b>	<b>42</b>
<b>3.5. Correlación entre RateIMSS vs Variables explicativas.</b>	<b>44</b>
<b>3.6. Correlación entre RateCapt vs Variables explicativas.</b>	<b>46</b>
<b>3.7 RatePIB y RateCapt por entidad federativa. 2001-2010.</b>	<b>47</b>
<b>4.1 Resultados de estimación modelo general</b>	<b>49</b>
<b>4.2 Resultados de estimación modelo efectos fijos por región</b>	<b>50</b>

<b>Índice de gráficas</b>	
<b>1.1 Índice General de la Actividad Económica. Enero2003/Diciembre 2014.</b>	<b>5</b>
<b>2.1 Sistema de Cuentas Nacionales. Base 2008,2003 y 1998 IGAE metodología de cálculo 1994.</b>	<b>18</b>
<b>2.2 Comparación entre las diferencias de los IGAE de distintas bases. Enero/1994-Enero/2008.</b>	<b>19</b>
<b>2.3 Porcentaje de participación del empleo de los meses de Julio- Diciembre respecto al total anual. Promedio 1998-2014</b>	<b>22</b>
<b>4.1 4.1 Comparación de la tasa de crecimiento observada para los estados (ratePIB) vs los estimados por el modelo general (ratePIBestimado)</b>	<b>49</b>
<b>Gráfico 4.2 Comparación de la tasa de crecimiento observada para los estados (ratePIB) vs los estimados por el modelo de efectos fijos (ratePIBFE)</b>	<b>50</b>

## Introducción

La ciencia económica ha tratado de encontrar herramientas que permitan pronosticar variables relevantes tales como producción, inversión, consumo, etc. con la mayor precisión posible; es decir, se ha buscado anticipar el comportamiento de aquellas variables que tienen relevancia para la toma de decisiones de los agentes económicos (empresas, gobierno y hogares). Pensar en realizar un pronóstico del Producto Interno Bruto en nuestro país implica un alto grado de dificultad, pues depende de variables macroeconómicas internas y externas, y más complicado aún si se intenta estimar un modelo del PIB a nivel de entidades federativas.

En particular, en el ámbito estatal es preciso tener información de manera desagregada de las principales variables económicas mencionadas anteriormente. Pero en la mayoría de los casos, esta no existe o bien tiene un grado de rezago en su publicación y difusión.

Con relación al Producto Interno Bruto por Entidad Federativa, esta variable es publicada hoy día por el Instituto Nacional de Estadística, Geografía (INEGI) con frecuencia anual y 18 meses de retraso, por lo cual pierde interés y relevancia para muchas de las decisiones económicas que deben tomarse.

Obtener la información correspondiente para el cálculo de la producción estatal o regional de manera trimestral es algo relativamente complicado y costoso si no se utilizan métodos de estimación con las actuales fuentes de información disponible<sup>1</sup>.

Considerando las circunstancias mencionadas anteriormente, es preciso encontrar metodologías que permitan hacer estimaciones, con menor rezago utilizando información de variables regionales disponibles en alta frecuencia, es decir de manera trimestral o mensual regularmente.

En 1989 el INEGI realizó un trabajo en este sentido, el cual buscaba crear un modelo de desagregación geográfica para estimar el PIB anual por entidad federativa. Para ello partieron de la metodología de Chow-Lin (1971) utilizando el PIB estatal anual publicado quinquenalmente como

---

<sup>1</sup> Según la metodología del cálculo del Indicador Trimestral de la Actividad Económica Estatal (ITAE) del INEGI: *es un indicador de coyuntura cuyo propósito es facilitar el seguimiento de la economía de los 32 estados del país, ofreciendo un panorama general de la situación y evolución macroeconómica. El indicador satisface la necesidad de disponer de datos que permitan examinar el desempeño trimestral de las entidades federativas.* El uso de este indicador es que presenta un rezago de al menos un trimestre, lo cual en series de tiempo es muy pequeño. [http://www.inegi.org.mx/est/contenidos/proyectos/cn/itae/doc/SCNM\\_Metodologia\\_19.pdf](http://www.inegi.org.mx/est/contenidos/proyectos/cn/itae/doc/SCNM_Metodologia_19.pdf)

variable dependiente y el crédito como variable independiente; y en las últimas dos décadas se han realizado esfuerzos por publicar la información del Producto Interno Bruto por Entidad Federativa (PIBE), mismos que se pueden encontrar en la página web del INEGI a partir de 2003 y hasta 2013, considerando la información con la base de cálculo 2008; y la información correspondiente a 1993-2006 con la base de cálculo de 1993<sup>2</sup>.

El presente documento pretende construir un modelo que pueda estimar el PIB trimestral de las entidades federativas del país a través de variables de publicación trimestral, por ejemplo la captación bancaria de las instituciones financieras, misma que es una variable de publicación mensual dentro de sus características se puede transformar en trimestral a través de un promedio del trimestre o un total a través de la suma de los meses que componen el trimestre, otro ejemplo de información disponible mensual que se podría trimestralizar es el consumo de kilowatts/hora (kwh) aplicando la misma aritmética aplicable a la captación, el total de trabajadores registrados por el IMSS (permanentes y eventuales), y al total de trabajadores registrados en las instituciones bancarias; estas tienen la característica principal de estar disponibles para las 32 entidades federativas.

Adicionalmente la presente investigación pretende incrementar el número de observaciones considerando la información en datos panel y obtener mejores estimadores. A diferencia del trabajo presentado por el INEGI, se tomará el PIB anual de cada entidad durante el periodo 1993-2013 y con la ayuda de las series auxiliares disponibles serie auxiliar se estimará el PIB trimestral.

Al momento de desagregar la información nacional en estados de la república, por lo general se encuentran los problemas de heteroscedasticidad en las estimaciones, es por ello que se acude a la regionalización de las entidades federativas conforme a la agrupación que el Banco de México hace de las mismas; en donde la región Norte está conformada por Baja California, Sonora, Chihuahua, Coahuila, Nuevo León y Tamaulipas; el grupo Centro Norte lo conforman Aguascalientes, Baja California Sur, Colima, Durango, Jalisco, Michoacán, Nayarit, San Luis Potosí, Sinaloa y Zacatecas; por otro lado, el Centro lo integran Distrito Federal, Estado de México, Guanajuato, Hidalgo Morelos, Puebla, Querétaro y Tlaxcala; y la Sur la componen Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz y Yucatán.

---

<sup>2</sup> Estos datos están disponibles en la sección de “Datos que ya no se actualizan”. Los cambios de base se realizan con el objetivo de generar estimaciones más precisas de la actividad económica. Por ejemplo, actividades que anteriormente no tenían relevancia cambian su peso en los ponderadores del año base. Además productos que anteriormente no tenían presencia en el mercado aparecen, al igual que otros son discontinuados.

Uno de los resultados relevantes es que al desagregar la información por regiones, la región Norte y Centro Norte presentaron un mayor grado de ajuste, al igual que sus coeficientes tuvieron resultados más significativos.

Los resultados permiten observar que en regiones con mayor grado de informalidad es más complicado generar un ajuste en el nivel de producción<sup>3</sup>, ya que al utilizar variables relacionadas con la formalidad; por ejemplo: empleo, energía eléctrica, flujos monetarios, etc., comparación con las entidades federativas que presentan un mayor grado de formalidad.

Este trabajo está organizado como sigue. El capítulo 1 se presenta la revisión bibliográfica relacionada con los modelos de desagregación temporal, mismo que desde la década de los 70's han sido utilizados, al grado de conseguir estimaciones mensuales para algunas regiones de Europa. El capítulo 2 describe la metodología y el comportamiento de las variables utilizadas en la estimación; aquí se presentan ejercicios de correlación tanto para la tasa de crecimiento anual de las variables así como las variables sin alguna modificación. En el capítulo 3 se presenta el modelo que se utilizará ya que al utilizar una serie de tiempo se utilizan tasas de crecimiento para verificar las relaciones de las variables a través del tiempo, recordemos que la información utilizada comprende un modelo de datos panel y la escasez de información limita las estimaciones a nivel temporal y en corte transversal la diferencia en la dinámica de crecimiento de los estados obliga a generar separación entre ellos, es por eso que se justifica la agrupación de las entidades federativas. En el capítulo 4 se presentan algunos resultados relevantes de las estimaciones de los modelos planteados al igual que sus transformaciones, cabe mencionar que se presentan algunos cálculos de los resultados trimestrales, para las regiones correspondientes. Las conclusiones, recomendaciones y limitaciones de las estimaciones se presentan el capítulo 5.

---

<sup>3</sup> Según las *El empleo informal en México: situación actual, políticas y desafíos* de la serie de Notas sobre Formalización, del Programa de Promoción de la Formalización para América Latina y el Caribe (2014) de la Organización Internacional del Trabajo, en 2013 los 5 estados con mayor nivel de informalidad, medida como el porcentaje de empleo formal respecto al total del empleo en cada entidad federativa, son Oaxaca, Guerrero, Chiapas, Puebla e Hidalgo y los estados con menor grado de informalidad fueron Nuevo León, Chihuahua, Baja California Sur, Coahuila y Baja California.

# CAPÍTULO 1

## 1. Antecedentes

La necesidad de tener una serie de datos que muestre el ciclo del producto a nivel estatal ha llevado a los investigadores del área económica a generar modelos que permitan predecir y calcular el comportamiento de las principales variables económicas (Producto Interno Bruto, inversión, consumo, remuneraciones, empleo, ventas, ingreso per cápita, etc.). Si este pronóstico se realiza para periodos relativamente pequeños, permitiría decisiones mejor fundamentadas para el desarrollo de la política económica de las entidades federativas.

Diversas metodologías se han aplicado para satisfacer este fin. Una de ellas es la generación de modelos econométricos que, utilizando información del pasado, obtienen ecuaciones que ayudan a presentar una descripción de lo sucedido hasta el punto donde finaliza la serie de datos. En estos casos, la estimación o cálculo se realiza al introducir la serie de valores de las variables explicativas dentro de las ecuaciones resultantes. Otra de las metodologías consiste en formar índices, (como los llamados coincidentes y adelantados) que reflejen el comportamiento de las principales variables económicas.

Estos métodos son llamados “métodos de desagregación temporal” y su objetivo es resolver un problema en el cual la información anual, trimestral, etc., debe ser distribuida en una frecuencia preestablecida (semestral, trimestral, mensual, etc.). En todo caso estos resultados deberían contar con una condición de minimización del error de ajuste respecto a la cifra anual, o bien, conservar, en forma total o parcial, la distribución de alguna variable que pueda estar altamente relacionada con los datos anuales a explicar.

En la literatura relacionada a la estimación de series trimestrales, o de series de frecuencia menor, se pueden considerar varios métodos, todos ellos con sus ventajas y desventajas al momento de ser utilizados.

Por lo general se pueden agrupar en cuatro tipos de métodos como mencionan Cortiñas et al (2008), Elizondo (2012), Rodríguez y Rodríguez (2000), etc.

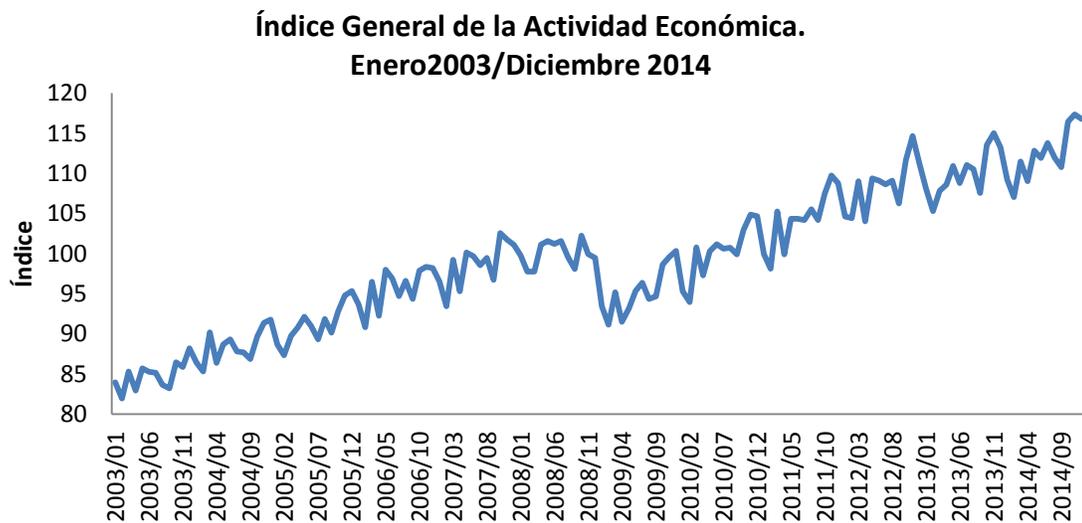
- Métodos triviales.
- Métodos que no utilizan indicadores relacionados.
- Métodos que utilizan indicadores relacionados.
- Métodos de extrapolación.

A continuación se presentan una breve descripción de cada uno de estos métodos:

### 1.1 Métodos triviales:

Dichos métodos son los más simples de utilizar. Ya que consisten en asignar el PIB anual de cada entidad federativa en 4 secciones del mismo valor, lo cual implicaría, económicamente, que en cada uno de los periodos los niveles de producción son los mismos, negando implícitamente las fluctuaciones que existen en las variables económicas. Por ejemplo, los fenómenos económicos relacionados con el incremento en la actividad productiva debido al pago de aguinaldos y beneficios de los trabajadores en el último trimestre del año. El caso contrario se presenta en el primer trimestre de cada año, fenómeno conocido coloquialmente como cuesta de enero, ya que los niveles de crecimiento en el trimestre correspondiente al mes de enero son menores<sup>4</sup>.

Gráfico 1.1



Fuente: INEGI. Índice General de la Actividad Económica

Dentro de los métodos triviales se puede considerar el supuesto de proporciones fijas, es decir si se obtuviera una medida total del PIB a nivel nacional de manera trimestral, con el supuesto mencionado, se podría realizar la distribución de las series trimestrales tomando como base los resultados de la aportación de cada entidad al PIB nacional o al promedio de su aportación durante la serie de tiempo estudiada.

<sup>4</sup> Para comprobar esto se puede realizar la siguiente prueba estadística sencilla, al obtener las diferencias entre los valores del índice para enero en comparación con diciembre, el promedio es de 5.79, lo cual es positivo y significativamente mayor a cero, a pesar de que en 2008 el valor de dicha diferencia es negativo.

## 1.2 Métodos que no utilizan indicadores relacionados

Aquellos métodos que no utilizan indicadores o variables que se vinculen ya sea de manera lógica o de manera económica, según sea el planteamiento del problema, con la serie anual a explicar, generan sus resultados de alta frecuencia utilizando ponderaciones relacionadas con los datos anuales del año en curso, el anterior y el posterior. El objetivo de dichos cálculos es generar una serie de alta frecuencia que se ajuste mediante alguna transformación de la serie de datos disponibles en función de un conjunto de restricciones temporales relativamente coherentes. Una virtud de estos modelos es que se pueden utilizar aun cuando no exista información trimestral que pueda relacionarse con la serie. En todo caso sólo se tomaría la información existente anualmente para determinar el comportamiento de la serie de alta frecuencia a estimar.

Entre estos métodos encuentran el propuesto por Lisman y Sandee (1964), el cual se basa en obtener los valores trimestrales mediante una combinación de ponderadores en función de los valores anuales del actual año, del anterior y del posterior. Mediante el uso de una matriz llamada de paso, la cual está sujeta a ciertas condiciones. Una de las restricciones a las que está sujeto este método es la imposibilidad de generar las estimaciones para el primer y último año de la serie.

Zani (1970) utiliza un método similar con la variante de que los datos trimestrales tienen una relación polinómica de segundo orden de la forma:

$$y_j = a + bt + ct^2 \quad (1.1)$$

Una de las ventajas de este método es que se pueden estimar los trimestres del primer y último año de la serie analizada.

Dentro de los métodos que no utilizan indicadores, Boot, Feibes y Lisman (1967) agregan un criterio que minimiza una función cuadrática, donde dicha función está sujeta a la restricción de que la suma de los trimestres es igual al dato del valor anual observado. Éste es uno de los métodos más utilizados ya que se fundamenta en la minimización de la suma de los cuadrados de las primeras o segundas diferencias entre trimestres consecutivos.

## 1.3 Métodos que utilizan indicadores relacionados.

Este tipo de métodos consideran variables de alta frecuencia que están disponibles y relacionadas en forma lógica o económica con la serie anual a desagregar. Para realizar las estimaciones trimestrales, dichos métodos estiman los resultados en función de la información disponible. En las

décadas de los 70's y 80's fueron muy usados, y aquí podemos mencionar los métodos utilizado por Chow –Lin (1971) y Denton (1971).

Con el objetivo de obtener información mensual a partir de información trimestral, Chow y Lin (1971) desarrollan un modelo que se ha utilizado, con distintas variantes, en la generación de los valores correspondientes a los subperíodos.

Antes de explicar el método de Chow-Lin (1971), es conveniente hacer mención del método de Denton (1971), ya que este último tiene la virtud de que puede ser aplicado utilizando alguna variable como indicador para realizar la desagregación temporal además de que puede omitirse el uso de ésta. Cabe señalar que el procedimiento de Denton (1971) es una aplicación del que utilizan Boot, Feibes y Lisman (1969).

Lo que hace dicho método es realizar una estimación por mínimos cuadrados ordinarios de un modelo lineal de la serie de datos anuales ( $Y$ ) en función de una matriz de indicadores anualizados ( $BZ$ ). Expresando estas ideas en la forma en que son representadas en Cortiñas, Parra y Vicente (2008):

$$Y = \alpha f_n + \beta(BZ) + B\mu \quad (1.2)$$

Se define a  $B$  como una matriz de agregación temporal tal que:

$$B = I_N \otimes f = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & .. & 0 & 0 & 0 & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 1 & 1 & 1 & 1 \end{bmatrix} \quad (1.3)$$

Donde  $I_N$  es la matriz identidad de rango  $1 \times N$ ,  $\otimes$  es producto tensorial, y  $f$ , es una matriz de distribución de los datos de alta frecuencia según el tipo de relación que tiene con el dato anual. Por ejemplo, cuando se habla de datos trimestrales  $f = \{1, 1, 1, 1\}$ ; si se tratara de información trimestral en índices, la matriz  $f = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$ ;  $Z$  es un indicador de referencia (matriz con la información en alta frecuencia), donde  $f$  puede tomar el valor de 1 o  $\frac{1}{4}$ , según sea el caso de información de flujo o índice.

Con las ecuaciones anteriores se obtendrían los datos trimestrales de  $y$ :

$$y = \frac{\alpha_n}{4} + \beta Z \quad (1.4)$$

La ecuación que determina la distribución trimestral una vez que se utilizan las primeras diferencias para minimizar los errores, está dada:

$$X_{fd} = (D'D)^{-1}B'[B'(D'D)^{-1}B]^{-1} \quad (1.5)$$

Y al emplear las segundas diferencias, las soluciones se obtienen por:

$$X_{sd} = (D'D'DD)^{-1}B'[B'(D'D'DD)^{-1}B]^{-1} \quad (1.6)$$

Otro de los métodos más utilizados para la estimación de datos de alta frecuencia con indicadores, es el que aplican Chow y Lin (1971), mismo que ha tenido muchas aplicaciones. Por ejemplo el Instituto Nacional de Estadística de España (1993) genera una metodología para la trimestralización de la economía española en el periodo 1970-1992 basada en esta propuesta. Trujillo, Benítez y López (2000) utilizan el presente modelo para la trimestralización de algunos sectores económicos de Andalucía en España; Rodríguez, Rodríguez y Dávila (2003) lo utilizaron para trimestralizar variables económicas de baja frecuencia en estimaciones de alta frecuencia con variables flujo de alta frecuencia relacionadas para la economía chilena; Hall, Viv y McDermott (2007) lo utilizaron para estimar los datos trimestrales del producto de Nueva Zelanda después de la Segunda Guerra Mundial; en tanto que Basile (2009) utiliza el método para realizar una estimación trimestral de diversas variables de interés para el gobierno italiano.

La aplicación de metodologías que realicen interpolaciones de series de tiempo se transforma en un ejercicio difícil de realizar cuando no se cuenta con series de menor periodicidad que están relacionadas con las variables a interpolar. Tomando en cuenta esto, Álvarez (1981) realiza las estimaciones anuales del PIB estatal en México para el periodo 1950-1978 utilizando como variable auxiliar los depósitos totales (DT)<sup>5</sup>, ya que estos se encontraban desagregados por entidad federativa para el periodo en cuestión.

En el trabajo de Puig y Hernández (1989), se hace una adaptación de la metodología de Chow y Lin (1971), la cual consiste en desagregar la información serial a corte transversal en lugar de series de tiempo de distinta periodicidad. De manera general, la aplicación realizada por Puig y Hernández consiste en hacer estimaciones anuales del PIB estatal, basándose en la información del PIB estatal quinquenal publicado por el INEGI.

---

<sup>5</sup> Definida como el promedio anual de los saldos (cuentas de cheques y depósitos).

Para los países que comprenden la Comunidad Económica Europea se utiliza este método para realizar estimaciones de la producción trimestral para cada país de la región. Por ejemplo, Proietti (2005) presenta una revisión de los métodos que posteriormente son aplicados en la estimación de datos trimestrales. Al considerar el nivel de desagregación regional, algunas regiones españolas como Andalucía utilizan la metodología desarrollada por Chow-Lin para la estimación del PIB trimestral de cada sector de la economía (Arana, Benítez y López, 2000).

En todos los casos donde se utiliza el enfoque de series de tiempo para la predicción y simulación del comportamiento de series, es normal que los errores estén correlacionados. En estos casos los métodos iterativos suelen ser los más utilizados, ya que generan un buen ajuste en el valor de los coeficientes.

En general, la metodología de Chow-Lin (1971) puede reducirse el procedimiento en el siguiente modelo de predicción:

$$y = X\beta + u \quad (1.7)$$

Donde  $y$  es una matriz  $4n \times 1$ ,  $X$  es  $4n \times p$  y debe ser considerada como fija desde el punto de vista del análisis estadístico, y  $u$  es un vector aleatorio con media 0 y una matriz de covarianza  $V$ . las variables a considerarse como explicativas serán aquellas que tengan una alta relación con la dependiente, además de que debe cumplir con la característica de fácil agrupación anual y fácil desagregación temporal.

Utilizando a la matriz  $C$  de rango  $n \times 4n$  como la matriz que convierte los periodos trimestrales en observaciones anuales, misma que permite realizar la interpolación de la información tendrá la estructura descrita a continuación:

$$C_I = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 \end{bmatrix} \quad (1.8);$$

$$C_I = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \dots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 1 & 1 & 1 \end{bmatrix} \quad (1.9)$$

El vector de n observaciones anuales de la variable dependiente será identificado suscribiendo un punto, dicha ecuación de regresión, satisface el siguiente modelo lineal:

$$y_t = Cy_t = CX_t\beta + Cu_t = X_t\beta + u_t \quad (1.10)$$

Con

$$Eu_t u_t' = V_t = CVC' \quad (1.11)$$

Entonces el problema a resolver es encontrar la matriz  $z$  de m observaciones de la variable dependiente, donde  $z$  es idéntica en el caso de la interpolación y distribución. Si quisiéramos solucionar un problema de extrapolación utilizaríamos un conjunto de observaciones que no corresponden al periodo de observación estudiado.

Un estimador linealmente insesgado  $\hat{z}$  de la matriz  $z$  que satisface la siguiente relación con la matriz  $A$  es:

$$\hat{z} = Ay_t = A(X_t\beta + u_t) \quad (1.12)$$

Considerando que “ $X_z$ ” y “ $u_z$ ” participan las variables en el modelo de regresión  $z$ , cabe mencionar que en el caso de la interpolación y distribución,

$$E(\hat{z} - z) = E[A(X_t\beta + u_t) - (X_z\beta + u_z)] = (AX_t - X_z)\beta = 0 \quad (1.13)$$

Las condiciones 1.12 y 1.13 implican

$$AX_t - X_z = 0 \quad (1.14)$$

$$\hat{z} - z = Au_t - u_z \quad (1.15)$$

La matriz de covarianza de  $(\hat{z} - z)$  estará dada por:

$$Cov(\hat{z} - z) = E(Au_t - u_z)(Au_t - u_z)' = AV_t A' - AV_z - V_z A' + V_z \quad (1.16)$$

con  $V_z$  identificamos a  $Eu_t u_z'$  y con  $V_z$  identificamos a  $Eu_z u_z'$ .

Para encontrar el mejor estimador linealmente insesgado de  $\hat{z}$ , minimizamos la traza de la matriz 2.16 con respecto a  $A$ , sujeto a la ecuación matricial  $AX - X_z = 0$  de (1.14).

Usando una matriz “ $M$ ” de rango  $m \times p$  para el multiplicador de Lagrange, la expresión en el lagrangeano se expresa en la siguiente ecuación:

$$L = \frac{1}{2} \text{tr}[AV.A' - AV_z - V_z.A' + V_z] - \text{tr}[M'(AX - X_z)], \quad (1.17)$$

Aplicando las reglas correspondientes a las trazas de matrices:

$$\frac{d \text{tr}AB}{dA} = \frac{d \text{tr}BA}{dA} = B',$$

Se obtiene:

$$(1.11) \quad AV - V_z = MX' \quad (1.18)$$

Al resolver la ecuación 1.26, la solución será

$$A = MX'.V^{-1} + V_z.V^{-1} \quad (1.19)$$

Al ser sustituido el resultado anterior en 2.7 se obtiene la solución para “ $M$ ”:

$$M'.V^{-1}X + V_z.V^{-1}X - X_z = 0 \quad (1.20)$$

Despejando para  $M$  se obtiene:

$$M = X_z(X.V^{-1}X)^{-1} - (V_z.V^{-1})X.(X.V^{-1}X)^{-1} \quad (1.21)$$

Sustituyendo el resultado del despeje de “ $M$ ” en la solución de 1.11 se obtiene la siguiente solución para  $A$ :

$$A = X_z(X.V^{-1}X)^{-1}X.V^{-1} + V_z.V^{-1} [I - X.(X.V^{-1}X)^{-1}X.V^{-1}] \quad (1.22)$$

El estimador que resulta para  $z$  es:

$$\hat{z} = Ay = X_z\hat{B} + (V_z.V^{-1})\hat{u}. \quad (1.23)$$

La solución para el vector de coeficientes:

$$\hat{\beta} = (X'.V.^{-1}X.)^{-1}X'.V.^{-1}y. \quad (1.24)$$

Cabe mencionar que dentro de los resultados presentados por estos autores el inconveniente es que no se puede conocer la matriz de varianza y covarianza, “V.”, misma que relaciona series trimestrales con la estructura de la información anual como lo mencionan Rodríguez, Rodríguez et al (2003). Adicionalmente, dichos autores realizan un ejercicio en donde se muestran los métodos que se han propuesto para la obtención de la matriz de varianza y covarianza que relaciona los datos trimestrales con los anuales en función del tipo de dato disponible. Por un lado, la solución trivial de CL implica sería donde  $V. = I$ , es decir la matriz de varianza y covarianza de relación trimestre-año fuera igual a la matriz identidad, esto implicaría que los errores de estimación anual se distribuyen uniformemente en cada trimestre. Dicho supuesto pareciera poco realista, sobre todo si conocemos que en la economía mexicana el cuarto trimestre tiene una actividad más acelerada que la presentada por el primer trimestre.

Los autores originales de este modelo proponen que la estimación de la matriz de varianza y covarianza se haga por medio de un proceso autoregresivo anual de orden uno (Chow y Lin, 1971), a través del cual se puede obtener el parámetro autoregresivo trimestral  $\emptyset$  a través de la siguiente expresión matemática:

$$\emptyset_a = \frac{\emptyset(\emptyset+1)(\emptyset^2+1)^2}{2(\emptyset^2+\emptyset+2)} \quad (1.26)$$

En el método de Denton (1971) y en Fernández (1981), que son métodos de ajuste, para obtener los valores trimestrales es necesario realizar un ejercicio en dos etapas. En primer lugar, se utilizan los indicadores para encontrar una primera estimación de los datos trimestrales y, posteriormente se aplica algún criterio de optimización que permita ajustar los valores trimestrales con el total de los valores anuales.

Denton (1971) inicia con una estimación de los datos lineales (“Y”) en función del indicador anualizado llamado BZ. Dicho modelo se representa con la siguiente ecuación:

$$Y = \alpha fn + \beta(BZ) + B\mu \quad (1.27)$$

Siendo B la matriz de desagregación temporal, “Z” es el indicador de referencia (es decir, la matriz de valores trimestrales que se relacionan con la serie anual en alguna forma temporal) y  $fn$  es un escalar que toma los valores de “1” cuando “Z” es una matriz de datos índices y de  $1/4$  cuando “Z” tiene valores que son flujo. Por lo tanto, para obtener nuestros datos trimestrales con variables flujo nuestro modelo tendría la siguiente forma:

$$y = \frac{\alpha}{4} + \beta Z \quad (1.28)$$

Por otra parte, el método de Fernández (1981) busca suavizar la serie trimestral estimada, suponiendo que ésta se distribuye como una senda aleatoria, es decir:

$$y = x_j \beta + u_j \quad (1.29)$$

$$u_j = u_{j-1} + \varepsilon_j \quad (1.30)$$

Dicha solución es la misma que presenta Denton (1971), pero aplicada a varias variables explicativas en donde el vector de coeficientes se describe de la siguiente forma:

$$\hat{\beta} = [x' B (B' (D' D) B)^{-1} B' x]^{-1} x' B (B' (D' D)^{-1} B)^{-1} Y \quad (1.31)$$

En donde “D” es la matriz de diferenciación.

## CAPÍTULO 2

### 2. Metodología

#### 2.1 Estimación del PIB trimestral por entidad.

Originalmente tenemos una relación entre las variables “y” y “X” dada por

$$y = X\beta + u \quad (2.1)$$

Donde y se define como la matriz de datos trimestrales del PIB, mismos que aún se encuentran sin estimar ya que el único dato que se conoce es el valor anual de las series. Además, se cuenta con una matriz “C” de agregación.

También se supone que la matriz de errores cuenta con media cero y varianza definida, es decir:

$$E(\varepsilon) = 0 ; E(\varepsilon\varepsilon') = \Omega \quad (2.2)$$

El objetivo principal es encontrar el mejor estimador de y en términos de la matriz de varianza y covarianza definida como  $E(\varepsilon\varepsilon') = \Omega$  y de “C”; por lo tanto, el mejor estimador de “y” está dado por :

$$\hat{y} = X\hat{\beta} + \Omega C'(C\Omega C')^{-1}\hat{u}_a \quad (2.3)$$

En donde el vector de coeficientes estaría determinado por

$$\beta = [X_a(C\Omega C')^{-1}\widehat{X}_a]^{-1}X'_a(C\Omega C')^{-1}y_a \quad (2.4)$$

y el vector de errores

$$\hat{u}_a = y_a - X[X'_a(C\Omega C')^{-1}\widehat{X}_a]^{-1}X'_a(C\Omega C')^{-1}y_a \quad (2.5)$$

Lo que está detrás de este modelo es que la serie trimestral final está constituida por sus dos componentes:

$X\hat{\beta}$ : Los coeficientes estimados de la regresión aplicados a los indicadores trimestrales.

$\Omega C'(C\Omega C')^{-1}\hat{u}_a$ : El vector de errores representa el residual de la regresión anual distribuido en los periodos trimestrales.

Para realizar la estimación trimestral del PIB por entidad es necesaria la desagregación de una serie de tiempo, dicho planteamiento implica tres problemas diferentes: interpolación, extrapolación, y distribución, mismo que se definen de la siguiente manera: la interpolación consiste en obtener estimaciones de subperiodos a partir de los valores de periodos anteriores y posteriores existentes sin imponer restricción alguna a las estimaciones. La extrapolación consiste en obtener valores para los subperiodos con base en periodos anteriores, es decir, toma en cuenta el “historial” de la serie para cada una de las nuevas estimaciones; y la distribución tiene por objeto obtener valores de los subperiodos a partir de un valor total del periodo.

La metodología de desagregación que se utiliza en este documento, se basa en modelos matemáticos y variables auxiliares, además de que no incorpora restricciones

Utilizando la metodología diseñada por Chow y Lin (1971) y su aplicación en distintos trabajos mencionados en los antecedentes, se desarrollará una metodología que permita obtener estimaciones del PIB trimestral para los estados del país.

En este trabajo se realizan estimaciones del PIB estatal trimestral, con base en datos trimestrales observados. También se realiza una interpolación de la serie anual a series trimestrales; y extrapolaciones basadas en la información anual.

## **2.2 Modelo de estimación.**

Para obtener los valores trimestrales del PIB, será necesario generar un modelo anualizado que se define de la siguiente manera:

$$PIB\bar{Est}_{it} = A + \beta \bar{X}_{it} + E_{it} \quad (2.6)$$

Aquí clasificamos a  $\bar{X}_{it}$  como la combinación de las “i” variables que se podrían utilizar como independientes en la estimación de la ecuación de regresión, y  $\beta$  es el vector de coeficientes correspondiente para cada uno de las variables independientes.

En esta anotación, el intercepto y los coeficientes de las variables explicativas son iguales para cada estado; esto se realiza con el objetivo de tener un mayor número de observaciones que permitan tener más grados de libertad al modelo. Esto tiene validez en el sentido de que son estados que pertenecen a la misma región, y presentan patrones de

crecimiento similar o las relaciones con las variables relevantes de estimación es similar, permitiendo eso que se consideren a las entidades como un grupo similar. Cabe señalar que también entre entidades federativas existe una diferencia en cuanto al comportamiento de las variables económicas, ya que las dinámicas del crecimiento en estos estados podrían ser diferentes.

De hecho, como se ha mencionado anteriormente, en los Reportes sobre las Economías Regionales del Banco de México, se observa una clasificación de cuatro regiones o zonas, mismas que tienen comportamientos similares entre los estados que las integran<sup>6</sup>.

De hecho, regionalizar se consigue a través de la búsqueda de algunas variables de estudio que hagan comparables los elementos que conforman los grupos (Coraggio, 1994), por ejemplo, Isusquiza (2014) utiliza la regionalización para realizar un estudio en el cuál se menciona como es afectado el crecimiento de las entidades federativas y algunas variables sociales una vez que se fomenta el federalismo fiscal, por otro lado Díaz-Bautista (2009) utiliza la información disponible para la región de la frontera norte del país para aplicar modelos de crecimiento en donde introduce variables relacionadas con la corrupción y el buen gobierno, uno de los trabajos más relevantes en el tema de la convergencia económica en las regiones de México (Esquivel, 1999) divide el área nacional en siete regiones distintas considerando sobre todo zonas geográficas, para obtener los niveles de convergencia de las regiones y relacionarlos con los efectos provocados por algunas variables sociales como la educación.

El objetivo de regionalizar es agrupar aquellas entidades federativas que puedan presentar comportamientos similares ante un grupo determinado de variables con el fin de tener grupos homogéneos al momento de hacer las estimaciones correspondientes y reducir, de esta forma, la posibilidad de la enfrentar problemas como la heteroscedasticidad.

---

<sup>6</sup> Recordemos que las zonas están constituidas como siguen: Norte: Baja California, Sonora, Chihuahua, Coahuila, Nuevo León y Tamaulipas; Centro Norte: Aguascalientes, Baja California Sur, Colima, Durango, Jalisco, Michoacán, Nayarit, San Luis Potosí, Sinaloa y Zacatecas; Centro: Distrito Federal, Estado de México, Guanajuato, Hidalgo Morelos, Puebla, Querétaro y Tlaxcala; y Sur: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz y Yucatán.

### 2.3 Descripción de la variable dependiente.

Dado que el presente estudio tiene por objetivo determinar el nivel de producción trimestral por entidad federativa en el periodo 1993-2013, es necesario comentar que en dicho periodo existen tres cambios de base para la determinación del PIB y del Sistema de Cuentas Nacionales de México (SCNM) en su conjunto, dichos cambios se realizan en 1998, 2003 y 2008.

- En primer lugar, el cambio en las clasificaciones de los componentes del PIB, sobre todo que al reclasificar las actividades primarias y secundarias se obtiene como resultado una mayor participación del sector secundario en la actividad económica en general, además de presentar en las principales divisiones industriales mayor actividad excepto por la minería;
- Por otro lado se tienen mejores resultados en la contabilidad nacional, lo cual implica incrementos en el PIB nominal y real.
- Por último con los incrementos de información, a partir del periodo 2008, más actividades y sectores pueden ser vistos en forma más específica.

En el presente caso la pregunta que debería ser respondida es acerca de la compatibilidad de las series con el fin de determinar si es posible considerarlas como una serie única para que puedan ser analizadas. Para ello es conveniente verificar la variabilidad y la dispersión generada en un periodo común al compararse las tasas de crecimiento de alguna de las series relacionadas con el PIB (IGAE, ITAEE, PIB anual, etc.).

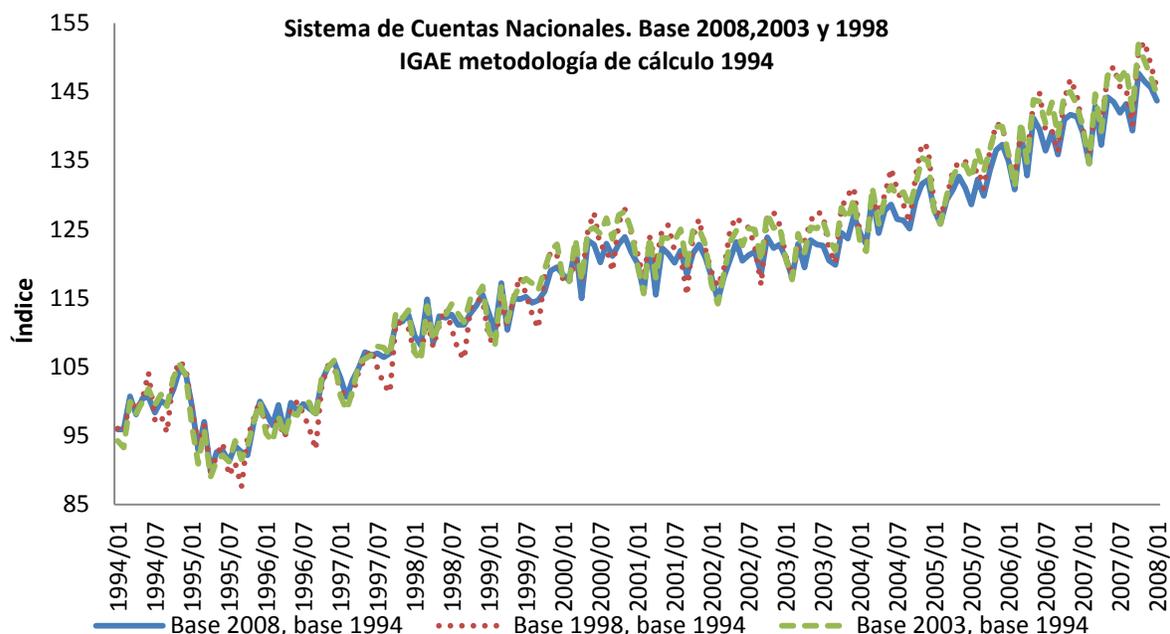
Se agrega un ejercicio para comparar el IGAE mensual a través del tiempo<sup>7</sup>, el objetivo de utilizar esta serie es porque en ellas se manifiestan las tres bases mencionadas anteriormente, para ello se obtuvieron los IGAE con distintas bases, 1998, 2003 y 2008; dado que estos contaban con bases distintas se utilizó el promedio del 1994 como base ya que el periodo traslapado entre las series comprende enero/1994 – enero/2008, es decir 169 observaciones mensuales.

En el gráfico 2.1 se muestra que los IGAE tienen poca variabilidad a través del tiempo lo cual es de esperarse ya que relacionan la actividad económica en el mismo periodo con distintas bases de estimación.

---

<sup>7</sup> Se considera el IGAE mensual debido a las siguientes condiciones, por un lado permite comparar una serie extensa de datos, las series se traslapan en varios periodos del tiempo y al ser extensas se pueden obtener resultados de la variabilidad (varianza) a través del tiempo.

**Gráfico2.1**



Fuente: Cálculos propios con información del INEGI. IGAE transformado en varias bases.

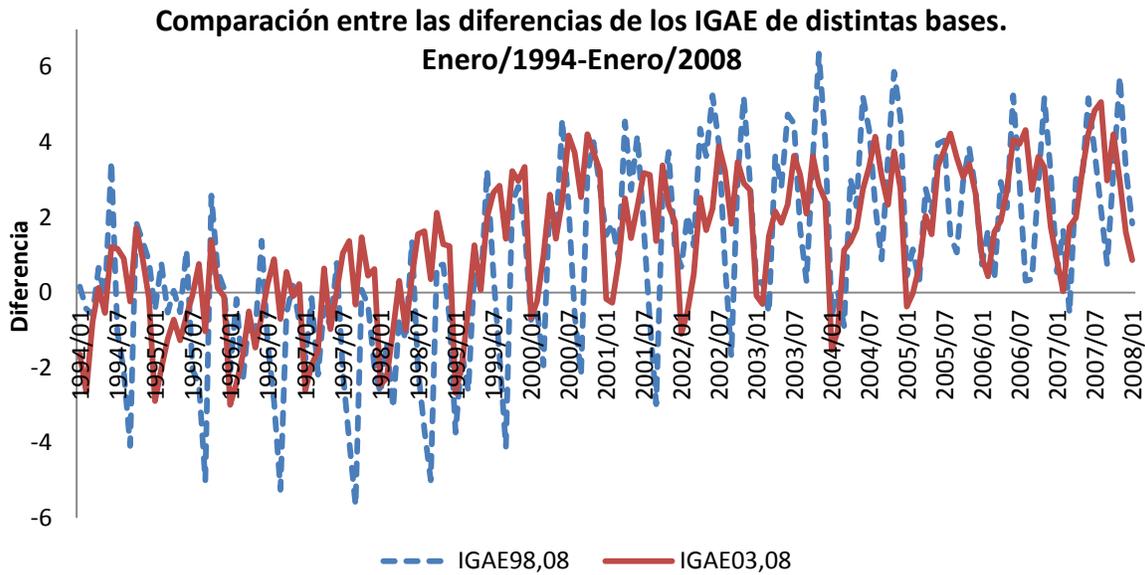
Dado que las series tienen bases distintas a pesar de que indican el movimiento de la actividad económica mensual a nivel nacional, el uso de bases distintas generaría algunas diferencias en los valores de los indicadores; para comparar dichas series y sus magnitudes se crearon dos variables:

$$\text{IGAE98,08} = \text{IGAE base 1998} - \text{IGAE base 2008}. \quad (2.7)$$

$$\text{IGAE03,08} = \text{IGAE base 2003} - \text{IGAE base 2008}. \quad (2.8)$$

La gráfica 2.2 muestra que la variabilidad de las diferencias es mayor cuando las series no son consecutivas, de hecho la varianza de IGAE98,08 es 2.51, mientras que la presentada por IGAE03,08 es de 1.87, esto nos permite observar que las series consecutivas suelen ser más parecidas, este resultado era esperable ya que las condiciones económicas entre los periodos 2008 y 2003 tienden a ser más parecidas, en cambio las condiciones imperantes en 1998 serían diferentes.

**Gráfico 2.2**



Fuente: Cálculos propios con información del INEGI.

Considerando la información de la gráfica 3.2, se realizó un ejercicio para determinar las diferencias de las tasas de crecimiento en los periodos 2004-2006, ya que la información del PIB estatal de los periodos 2003-2006 se traslapan para dos bases de construcción distinta (1993 y 2008). Dicho ejercicio consistió en construir el PIB real considerando los valores nominales de cada serie, posteriormente se deflactan en base 2003, conforme al deflactor de cada uno de los periodos base, es decir la información generada con base en 1993 es deflactada con el deflactor del PIB base 1993; y lo mismo se realiza con la información 2008. Para que ambas series coincidan, ambos deflatores toman el valor de 100 en 2003.

Una vez que se han generado las tasas de crecimiento real de ambas series en condiciones de comparabilidad, se obtuvo la diferencia del valor de la tasa de crecimiento real base 2008 menos el mismo datos para la base 1993 con los datos de los periodos 2004, 2005 y 2006 para cada entidad federativa y el total nacional.

$$\text{Diferencia}_{i,t} = \text{TCPIBE}_{i,t,2008} - \text{TCPIBE}_{i,t,1993} \quad (2.9)$$

Donde,  $i =$  Aguascalientes,...,Zacatecas.       $t = 2004, 2005, 2006$

$\text{TCPIBE}_{i,t, 2008} =$  Tasa de crecimiento del PIB de la entidad  $i$  en el periodo  $t$ , base 2008

$\text{TCPIBE}_{i,t,1993} =$  Tasa de crecimiento del PIB de la entidad  $i$  en el periodo  $t$ , base 1993

Los resultados se presentan en la siguiente tabla 2.1, en ellos podemos observar que en lo referente al total nacional, las diferencias entre tasa de crecimiento con distintas bases es muy pequeña, en cambio, cuando desagregamos la información por regiones estas diferencias se incrementan.

**Tabla 2.1 Diferencia de la tasa de crecimiento para las entidades federativas base 1993 y 2008.**

	Entidad	Diferencia 2004	Diferencia 2005	Diferencia 2006
<b>Región</b>	<b>Total nacional</b>	<b>0.05</b>	<b>0.31</b>	<b>0.16</b>
	<b>Promedio General</b>	<b>-0.24</b>	<b>0.66</b>	<b>0.41</b>
<b>Norte</b>	Baja California	-2.92	-0.26	0.41
	Chihuahua	1.15	-1.24	2.78
	Coahuila de Zaragoza	-2.35	0.58	0.33
	Nuevo León	-0.59	2.43	1.66
	Sonora	-2.65	1.29	-0.78
	Tamaulipas	-3.16	3.23	1.18
	<b>Promedio Región</b>	<b>-1.75</b>	<b>1.00</b>	<b>0.93</b>
<b>Centro Norte</b>	Aguascalientes	2.20	-0.15	-1.65
	Baja California Sur	8.65	3.28	1.26
	Colima	-2.59	-0.56	-1.06
	Durango	-3.48	-0.76	-0.16
	Jalisco	-0.68	0.86	2.68
	Michoacán de Ocampo	-1.92	2.67	-0.04
	Nayarit	4.63	0.45	-4.73
	San Luis Potosí	-1.94	-0.54	-0.28
	Sinaloa	-0.31	0.61	-0.75
	Zacatecas	2.14	3.23	1.51
	<b>Promedio Región</b>	<b>0.67</b>	<b>0.91</b>	<b>-0.32</b>
<b>Centro</b>	Distrito Federal	3.95	0.72	1.05
	Guanajuato	-1.92	2.28	0.42
	Hidalgo	0.37	0.66	-1.01
	México	-0.31	-0.21	-0.77
	Puebla	2.32	-1.13	-0.59
	Querétaro	2.11	1.42	0.88
	Tlaxcala	-2.31	-4.14	2.17
	<b>Promedio Región</b>	<b>4.21</b>	<b>-0.40</b>	<b>2.16</b>
<b>Sur</b>	Campeche	-1.31	-1.12	-1.86
	Chiapas	-7.51	1.23	1.01
	Guerrero	3.34	0.36	0.67
	Morelos	-1.54	-2.51	0.54
	Oaxaca	0.58	-1.52	-1.87
	Quintana Roo	-1.30	3.55	2.66
	Tabasco	0.66	5.37	2.39
	Veracruz de Ignacio de la Llave	0.45	0.79	2.13
	Yucatán	-1.49	0.35	2.89
<b>Promedio Región</b>	<b>-0.90</b>	<b>0.72</b>	<b>0.95</b>	

Fuente: Cálculos propios con información del INEGI.

Si se observan los valores de las diferencia al nivel de las entidades federativas están presentarán valores significativamente diferentes de cero, por ejemplo Baja California Sur en 2004; estos resultados muestran que al momento de utilizar información de los periodos 1993-2003 del PIB por entidades tendremos un sesgo. La advertencia anterior era necesaria ya que se trabajará con la serie correspondiente al PIB de las entidades federativas que será la variable dependiente:

$\overline{PIB}_{Est\ ikt}$  = Valores observados del producto interno bruto para el estado  
Donde  $i = \text{Aguascalientes, ..., Zacatecas}$ , para el año  $t (t= 1993, 1994, \dots, 2012, 2013)$ ,  
y  $k = \text{Norte, Centro Norte, Centro y Sur}$ .

La producción esta medida en pesos reales de 2003 uno de los años donde coinciden las dos series de información, al utilizar la información en precios reales se eliminarían los efectos de la inflación las unidades físicas que se describen a continuación. No se realiza supuesto alguno sobre la productividad de los factores ni del reflejo de esta sobre el producto.

## **2.4 Variables independientes**

### **2.4.1 Total de Trabajadores Asegurados en el IMSS**

En dicha categoría están los trabajadores permanentes y eventuales. Para la construcción de dicha variable se tomó el promedio de cada mes del año y posteriormente se sumaron dichos datos para conformar el total anual.

$TAIMSS_{ikt}$  = Los trabajadores asegurados, permanentes y eventuales, que participaron en la producción durante un año.

$i = \text{Aguascalientes, ..., Zacatecas}$ ;

$t = 1993, 1994, \dots, 2012, 2013$ ), en el grupo  $k$ .

En el caso de la información correspondiente al total de trabajadores del IMSS los datos disponibles anuales correspondieron al periodo 1998-2014, pero solo fueron utilizados hasta 2013 por ser el periodo analizado.

Para extenderlos datos hasta 1997 se utilizaron los datos mensuales disponibles entre julio-diciembre del mismo año y se siguieron los siguientes pasos: primero se determinó el porcentaje de participación de la suma del total de trabajadores asegurados (permanentes y eventuales)

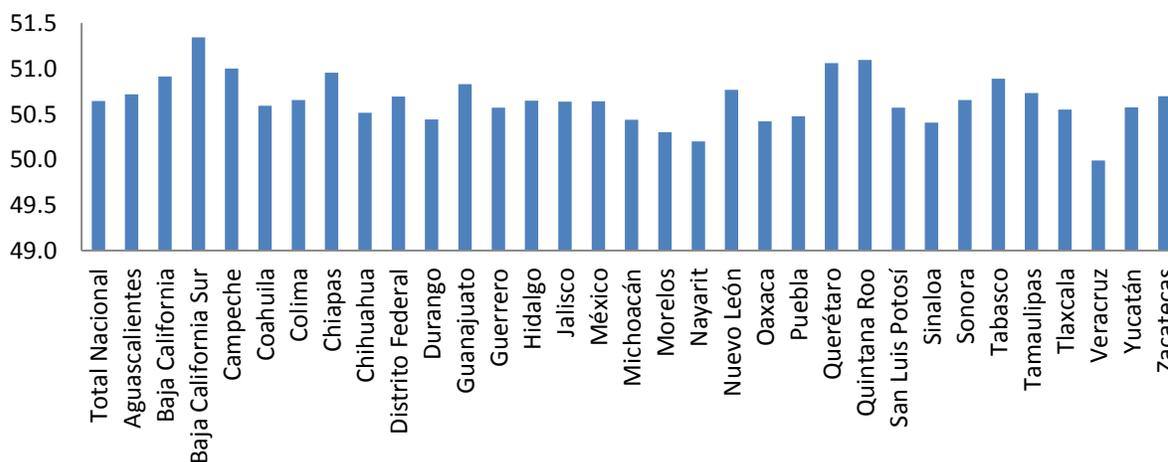
correspondiente al periodo julio-diciembre de cada uno de los años desde 1998 hasta 2013, dicha aplicación se realizó para los 32 estados de la república.

Una vez encontrado el porcentaje de participación, en cada año, de la suma de los trabajadores asegurados del IMSS correspondiente a julio-diciembre se obtuvo un promedio de los 16 años utilizados para tal efecto, mismo que se presenta en la columna 2 de la tabla 2.2 y que están representados en el gráfica para cada estado de la república; el dato mínimo corresponde al estado de Veracruz y el máximo para Baja California Sur. Para efectos estadísticos se calculó la desviación estándar para los estados en las 16 observaciones disponibles al igual que los límites inferiores y superiores (columnas 3, 4 y 5 de la tabla mencionada anteriormente), se puede observar que dichos intervalos son muy pequeños en general, esto permite ver la poca variabilidad en la información correspondiente al total de trabajadores asegurados del IMSS en el segundo semestre de cada año.

Con el porcentaje de los trabajadores asegurados correspondientes al segundo semestre de cada año, se obtuvo el porcentaje del primer semestre a través de una diferencia. Dicho resultado permitió calcular los valores correspondientes al total de trabajadores asegurados del IMSS que participaron en la actividad económica en el primer semestre de 1997, dicha estimación en porcentaje se presenta en la columna 6 y en valor absoluto en la columna 7. Cabe resaltar que la diferencia entre el total de trabajadores de un semestre es porcentualmente muy pequeña, de hecho, asignar el valor del 50% en cada caso tendría resultados similares.

**Gráfica 2.3**

**Porcentaje de participación del empleo de los meses de Julio-Diciembre respecto al total anual. Promedio 1998-2014**



La estimación correspondiente al primer semestre del 1997 permitió que se incrementaran en conjunto 32 observaciones de corte transversal y una observación de series de tiempo.

**Tabla 2.2. Estimaciones del Total de trabajadores de asegurados 1997.**

Entidad	Promedio porcentual (Julio-Diciembre)	Desviación estándar	Límite Inferior	Límite Superior	Porcentaje (Enero-Junio)	IMSS 1997 (Julio-Diciembre)	Estimado (Enero-Junio)	Total 1997
<b>Total Nacional</b>	50.64	0.5590	50.58	50.71	49.36	62,806,937	61,208,770	124,015,707
<b>Aguascalientes</b>	50.72	0.6186	50.64	50.80	49.28	863,891	839,393	1,703,284
<b>Baja California</b>	50.91	1.0943	50.78	51.05	49.09	2,841,040	2,739,237	5,580,277
<b>Baja California Sur</b>	51.34	1.9914	51.10	51.59	48.66	385,677	365,487	751,164
<b>Campeche</b>	51.00	0.7447	50.91	51.09	49.00	390,229	374,928	765,157
<b>Coahuila</b>	50.59	0.5244	50.53	50.66	49.41	2,525,095	2,465,883	4,990,978
<b>Colima</b>	50.66	0.7482	50.56	50.75	49.34	366,152	356,682	722,834
<b>Chiapas</b>	50.95	0.6806	50.87	51.04	49.05	600,891	578,368	1,179,259
<b>Chihuahua</b>	50.52	0.9168	50.40	50.63	49.48	3,522,851	3,450,773	6,973,624
<b>Distrito Federal</b>	50.69	0.5504	50.62	50.76	49.31	12,133,199	11,801,802	23,935,001
<b>Durango</b>	50.44	0.7048	50.36	50.53	49.56	940,175	923,700	1,863,875
<b>Guanajuato</b>	50.83	0.4183	50.78	50.88	49.17	2,405,732	2,327,175	4,732,907
<b>Guerrero</b>	50.57	0.7852	50.48	50.67	49.43	664,518	649,475	1,313,993
<b>Hidalgo</b>	50.65	0.4521	50.59	50.70	49.35	722,906	704,447	1,427,353
<b>Jalisco</b>	50.64	0.4408	50.58	50.69	49.36	5,115,780	4,987,090	10,102,870
<b>México</b>	50.64	0.3702	50.60	50.69	49.36	5,068,439	4,939,988	10,008,427
<b>Michoacán de Ocampo</b>	50.44	0.5367	50.37	50.51	49.56	1,234,324	1,212,829	2,447,153
<b>Morelos</b>	50.30	0.3922	50.25	50.35	49.70	788,042	778,601	1,566,643
<b>Nayarit</b>	50.20	1.1879	50.05	50.35	49.80	435,348	431,880	867,228
<b>Nuevo León</b>	50.77	0.5676	50.70	50.84	49.23	4,666,324	4,525,195	9,191,519
<b>Oaxaca</b>	50.42	0.3983	50.37	50.47	49.58	738,326	726,031	1,464,357
<b>Puebla</b>	50.48	0.4834	50.42	50.54	49.52	1,980,983	1,943,606	3,924,589
<b>Querétaro</b>	51.06	0.7240	50.97	51.15	48.94	1,145,833	1,098,266	2,244,099
<b>Quintana Roo</b>	51.10	1.9132	50.86	51.33	48.90	760,350	727,693	1,488,043
<b>San Luis Potosí</b>	50.57	0.8164	50.47	50.67	49.43	1,157,267	1,131,139	2,288,406
<b>Sinaloa</b>	50.41	0.7759	50.31	50.50	49.59	1,594,377	1,568,615	3,162,992
<b>Sonora</b>	50.66	1.2512	50.50	50.81	49.34	1,981,960	1,930,699	3,912,659
<b>Tabasco</b>	50.89	0.8088	50.79	50.99	49.11	558,743	539,181	1,097,924
<b>Tamaulipas</b>	50.73	0.9273	50.62	50.85	49.27	2,384,654	2,315,715	4,700,369
<b>Tlaxcala</b>	50.55	0.7014	50.46	50.64	49.45	364,842	356,919	721,761
<b>Veracruz</b>	49.99	0.4136	49.94	50.04	50.01	2,898,074	2,899,247	5,797,321
<b>Yucatán</b>	50.57	0.4641	50.52	50.63	49.43	1,094,503	1,069,650	2,164,153
<b>Zacatecas</b>	50.70	0.3679	50.65	50.74	49.30	476,412	463,289	939,701

Fuente: Cálculos propios con información del IMSS

## 2.4.2 Construcción de la serie TAIMSS 1993-1996

Adicionalmente para extender la serie de tiempo en el periodo de 1993-1996, se utilizó la información presentada por los Anuarios Estadísticos de las Entidades Federativas, 1994-1997, publicados por el INEGI; cabe mencionar que la información disponible en cada caso eran los “Asegurados permanentes temporales del IMSS”<sup>8</sup>, lo cual representa un grado de diferencia significativo, en algunos casos, con la información correspondiente a la variable TAIMSS.

El procedimiento que se utilizó en la construcción de la serie restante consistió en lo siguiente:

- Con la información del total de Asegurados del IMSS presentada en los Anuarios Estadísticos de las Entidades (1994-1997) de todos los estados de la república.
- Se obtuvo la relación porcentual que existe a través del tiempo (1998-2013) entre el total de trabajadores asegurados y el total de asegurados del IMSS para cada entidad federativa, como a aparecen en la siguiente ecuación

$$TAIMSSvs Asegurados = T\_A = \frac{TAIMSS}{Asegurados IMSS} \times 100 \quad (2.10)$$

- Una vez que se ha obtenido al nueva variable T\_A, se verifica su trayectoria a través del tiempo mediante una regresión lineal simple con la siguiente forma funcional:

$$T\_A_{it} = \alpha_i + \beta_i time_t + \varepsilon_{it} \quad (2.11)$$

Dónde:

---

<sup>8</sup> Cabe destacar que existen diferencias entre las definiciones del “Total (permanentes y eventuales) de trabajadores asegurados” y el “ Total (permanentes y eventuales) de asegurados del IMSS”. Según el Glosario de Términos de la Dirección de Finanzas del IMSS, el total de “Asegurados o cotizantes”: *Se refiere a las personas que están aseguradas en el IMSS de manera directa como titulares. Incluye todas las modalidades de aseguramiento, tanto las relacionadas con trabajadores, como con no trabajadores en las modalidades 32 (seguro facultativo), 33 (seguro de salud para la familia) y 40 (continuación voluntaria al régimen obligatorio). No considera a pensionados, ni a los beneficiarios de asegurados y pensionados. La descripción de las modalidades de aseguramiento en el IMSS se detalla en el apartado de asegurados por tipo de afiliación.* <http://www.imss.gob.mx/sites/all/statics/pdf/informes/GlosarioCubo.pdf>

$T_{A_i}$ = representa la relación que existe entre el total de trabajadores registrados en el IMSS y el total de asegurados registrados para la entidad  $i$  en el periodo  $t$ .

$\alpha_i$ = Representa el valor del intercepto, es decir el valor del porcentaje en el periodo inicial para la entidad  $i$  una vez que la variable de tiempo toma el valor de cero.

$\beta_i$ = Es la razón de cambio de la variable  $T_{A_i}$  cuando cambia la variable el tiempo.

$time_t$ = el tiempo medido consecutivamente 0, 1, 2, ..., 15, el valor de 0 corresponde a 1997.

$\varepsilon_{it}$  =error de estimación para el estado  $i$  en el periodo  $t$ .

Los resultados por entidad federativa se presentan en la tabla 3.4. En ellos se observa que a excepción de los estados de Baja California Sur, Guerrero y Durango, todos los estados de la república presentan  $R^2$  mayores a 0.5, además los coeficientes relacionados con la variable tiempo son altamente significativos y con signo negativo en todos los casos, lo que implica que la tendencia en los últimos años ha sido en dirección de disminuir el total de trabajadores con respecto al total de asegurados, para todas las entidades federativas (Ruiz, 2009) <sup>9</sup>.

Para generar los cálculos correspondientes de la variable TAIMSS correspondientes a los periodos 1993-1997 se utilizó el siguiente procedimiento una vez que se han obtenido los valores correspondientes a las regresiones lineales.

Dado que la variable de estimación es la razón

$$T_{A_i} = \frac{TAIMSS}{Asegurados IMSS} \times 100 = \alpha_i + \beta_i time_t + \varepsilon_{it} \quad (2.11, \text{ para cada entidad})$$

Si  $time = 0$ , entonces  $T_{A_i} = \alpha_i$ , por tanto si reducimos en un periodo de tiempo  $time = -1$ ,

$$T_{A_{i-1}} = \alpha_i + \beta_i(-1), \quad (\text{aplicación de 2.11})$$

---

<sup>9</sup>A través del tiempo en México, se ha observado una disminución importante del número de trabajadores registrados principalmente por los esquemas de sustitución laboral que las empresas pusieron en práctica para reducir los costos generados por la contratación de personal permanente a través de los impuestos y las cuotas de seguridad social.

por lo tanto, la relación  $T_{Ai}$  se incrementaría.

Como ejemplo se calculan los datos de Aguascalientes en 1993 y 1996:

En 1993  $time = -5$  , además en 1996  $time = -1$ ; el resultado de la regresión de dicha entidad es

$$T_{Ai} = 90.22024 - 0.79033 * time$$

En 1993

$$T_{Ai} = 90.22024 - 0.79033 * (-4) = 94.17189$$

En 1996

$$T_{Ai} = 90.22024 - 0.79033 * (-1) = 91.01057$$

Estos resultados muestran la relación entre TAIMSS y los asegurados totales del IMSS, fórmula anexa  $T_{Ai} = \frac{TAIMSS}{Asegurados IMSS} X 100$  , para obtener TAIMSS es necesario despejar dicha variable.

En 1993;  $T_{Ai} = 94.17189$ ; por lo tanto,

$$TAIMSS = \frac{T_{Ait} X Asegurados IMSS}{100} = 1,409,851.13, \text{ es decir } 1,409,852$$

$$TAIMSS_{1993} \approx 1,695,320$$

Cabe destacar que los tres estados de la república que presentaron una  $R^2$  más baja en la tabla 3.4, hecho que era de esperarse ya que según los datos presentados en Ruíz (2009), muestran que tanto Baja California Sur, Durango y Guerrero, son algunas de las entidades federativas con menor grado de formalidad en los últimos años.

**Tabla 2.4 Resultados de la regresión T\_A vs. Tiempo.**

Entidad	Constante	Prueba t	Coefficiente Tiempo	Prueba t	R2
Aguascalientes	90.22024	136.90	-0.79033	-9.87	0.8822
Baja California	95.0726	101.08	-1.1713	-10.24	0.8898
<i>Baja California Sur</i>	<i>88.51652</i>	<i>106.79</i>	<i>-0.2696</i>	<i>-2.68</i>	<i>0.3551</i>
Campeche	80.04481	144.78	-0.26546	-3.95	0.5454
Coahuila de Zaragoza	95.61958	264.46	-0.73207	-16.66	0.9552
Colima	72.43323	95.80	-0.481	-5.23	0.6781
Chiapas	57.13467	84.76	-0.90632	-11.06	0.9039
Chihuahua	93.2782	158.00	-0.31992	-4.46	0.6045
DF	37.41198	91.67	-0.41329	-8.33	0.8422
<i>Durango</i>	<i>618.2156</i>	<i>46.34</i>	<i>-4.50001</i>	<i>-2.77</i>	<i>0.3719</i>
Guanajuato	394.8614	113.59	-8.00377	-18.94	0.9650
<i>Guerrero</i>	<i>74.6491</i>	<i>24.28</i>	<i>-1.16712</i>	<i>-3.12</i>	<i>0.4286</i>
Hidalgo	255.498	66.54	-5.98163	-12.82	0.9267
Jalisco	11.67669	52.43	-0.27872	-10.30	0.8908
México	11.35212	51.52	-0.17857	-6.67	0.7737
Michoacán de Ocampo	322.4893	160.99	-6.8585	-28.16	0.9839
Morelos	319.929	63.15	-4.93834	-8.02	0.8318
Nayarit	346.5981	37.89	-6.27844	-5.65	0.7103
Nuevo León	105.1593	147.79	-0.73689	-8.52	0.8481
Oaxaca	113.419	37.42	-1.4462	-3.93	0.5424
Puebla	29.74757	102.56	-0.32034	-9.09	0.8639
Querétaro	27.16656	54.15	-0.27620	-4.53	0.6121
Quintana Roo	481.1697	35.52	-9.9274	-6.03	0.7365
San Luis Potosí	49.28961	86.07	-.586276	-8.42	0.8451
Sinaloa	112.2781	64.80	-3.22153	-15.29	0.9473
Sonora	56.33466	38.02	.977152	5.42	0.6936
Tabasco	83.82151	26.42	-0.37854	-0.98	0.0690
Tamaulipas	38.431	73.29	0.446049	7.00	0.7902
Tlaxcala	284.459	37.41	-2.02615	-2.19	0.2699
Veracruz-Llave	53.24407	73.09	-1.06487	-12.02	0.9175
Yucatán	39.14288	108.51	.8579546	19.56	0.9672
Zacatecas	373.9335	53.51	-11.8204	-13.91	0.9371

Fuente: Cálculos y tratamientos propios con Información del IMSS.

Adicionalmente se realizó la misma estimación suponiendo un comportamiento logarítmico del porcentaje de los trabajadores respecto al total de asegurados, es decir el comportamiento de la ecuación es el siguiente para cada entidad federativa:

$$\ln(T\_A)_{it} = \alpha_{it} + \beta_i time_t + e_{it} \quad (2.12)$$

Donde

$\ln(T\_A)_{it}$  = es el logaritmo natural de la proporción de los trabajadores asegurados del IMSS respecto al total de asegurados para las entidades federativas en el periodo analizado.

i=Aguascalientes,....., Zacatecas. t=1997,....,2014

Aplicando el antilogaritmo obtenemos la siguiente ecuación:

$$T\_A_{it} = e^{\alpha_{it} + \beta_i time_t + e_{it}} \quad (2.13)$$

Bajo el supuesto de que el error tiene media cero y varianza constante.

$$T\_A_{it} = e^{\alpha_{it} + \beta_i time_t} \quad (2.14)$$

Los resultados correspondientes a las estimaciones se presentan en la tabla 2.5; en ella se muestra que de las 32 entidades federativas sólo 3 presentan una  $R^2$  menor que 0.5 (Nayarit, Sonora y Querétaro), pero en todos los casos el coeficiente que representa la relación con el tiempo tiene valor negativo, el ser una relación logarítmica, por cada unidad porcentual que se incrementa el tiempo, la variable T\_A disminuye en  $\beta$  porciento, podemos observar también que los dos coeficientes son significativos.

Al igual que en el anterior ejercicio, un valor de t=0 implicaría el valor de 1997,

Como se mencionó anteriormente, esto se puede deber a la flexibilización en el uso del factor trabajo y a las relaciones obrero-patronales (Ruiz, 2009). El cálculo de Aguascalientes en 1993 y 1996 con información de la tabla 3.5 se haría de la siguiente manera:

$$T\_A = T\_A_{it} = e^{4.519905 + (-0.008345) (-1)} = 92.5963756$$

Es decir, en 1996 el total de trabajadores respecto a los asegurados correspondía a 92.596% según lo estimado, en comparación al 89.429 de 1997. Un resultado con la ecuación logarítmica es que en los estados de Jalisco y Baja California, T\_A tiene un valor de 100%, lo cual es casi imposible pero a pesar de ello se consideró dicho porcentaje.

**Tabla 2.5 Resultados de la regresión T\_A vs. Tiempo. Ecuación logarítmica.**

Entidad	Constante	Prueba t	Coefficiente de Tiempo	Prueba t	R2
Aguascalientes	4.519905	622.04	-.008345	-11.44	0.8910
Baja California	4.57261	384.18	-.011468	-9.59	0.8519
Baja California Sur	4.506939	518.26	-.004304	-4.93	0.6028
Campeche	4.390035	678.60	-.003184	-4.90	0.6002
Coahuila de Zaragoza	4.566009	720.96	-.006451	-10.14	0.8654
Colima	4.343873	326.90	-.010266	-7.69	0.7872
Chiapas	4.123322	270.83	-.019647	-12.85	0.9117
Chihuahua	4.553878	796.92	-.004135	-7.21	0.7645
DF	4.406725	472.53	-.011230	-11.99	0.8999
Durango	4.536933	436.30	-.013444	-12.87	0.9120
Guanajuato	4.521876	642.61	-.013790	-19.52	0.9597
Guerrero	4.066993	321.78	-.016535	-13.03	0.9139
Hidalgo	4.401871	212.85	-.033531	-16.15	0.9422
Jalisco	4.587648	313.63	-.017777	-12.10	0.9015
México	4.420989	411.98	-.016524	-15.33	0.9363
Michoacán de Ocampo	4.40795	515.22	-.018604	-21.65	0.9670
Morelos	4.414144	328.86	-.016975	-12.59	0.9084
<i>Nayarit</i>	<i>4.258933</i>	<i>134.70</i>	<i>-.011518</i>	<i>-3.63</i>	<i>0.4513</i>
Nuevo León	4.575821	1495.69	-.001604	-5.22	0.6300
Oaxaca	4.194246	146.99	-.020752	-7.24	0.7663
Puebla	4.420398	367.86	-.015560	-12.89	0.9122
<i>Querétaro</i>	<i>4.439422</i>	<i>238.58</i>	<i>-.005684</i>	<i>-3.04</i>	<i>0.3664</i>
Quintana Roo	4.539041	796.67	-.006517	-11.39	0.8902
San Luis Potosí	4.479788	708.10	-.007881	-12.41	0.9058
Sinaloa	4.500242	329.54	-.022424	-16.35	0.9435
<i>Sonora</i>	<i>4.468528</i>	<i>390.06</i>	<i>-.003767</i>	<i>-3.27</i>	<i>0.4012</i>
Tabasco	3.991815	171.93	-.009593	-4.11	0.5141
Tamaulipas	4.511954	829.55	-.009631	-17.63	0.9511
Tlaxcala	4.349014	435.20	-.031904	-31.79	0.9844
Veracruz-Llave	4.448179	531.38	-.022752	-27.07	0.9786
Yucatán	4.491854	1064.35	-.00554	-13.07	0.9144
Zacatecas	4.281933	544.03	-.016471	-20.84	0.9645

Fuente: Cálculos y tratamientos propios con Información del IMSS.

Para comparar y decidir entre el uso de información de uno o de otro método, se calcularon las tasas de crecimiento de TAIMSS de 1994-1997 y se decidió por aquellas que tienen menor varianza, que

para la mayoría de los casos fueron los resultados de la ecuación logarítmica<sup>10</sup>. En el anexo 2, correspondiente a la información estadística de las variables independientes, se presenta el comportamiento de la tasa de decrecimiento de la variable TAIMSS a través del tiempo para cada entidad federativa, se presentan la trayectoria de ambas

### 2.4.3 Consumo de energía eléctrica por entidad federativa

Otra de las variables que se consideró como explicativa es  $Kwh_{ijt}$ , donde:

$Kwh_{ijt}$  = El volumen de ventas de energía eléctrica medido en miles de kilowatts/hora. Para el estado  $i$  ( $i =$  Aguascalientes, ..., Zacatecas), para el año  $t$  ( $t =$  2003, 2004, ..., 2012, 2013), en el grupo  $j$ .

Con la información disponible a nivel entidad federativa anteriormente mencionada se representa la función de producción en la cual la variable energía eléctrica representa una variable instrumental del capital ya que se ha encontrado que en algunos casos se mueven de manera directa (Nicholson, 2001), ya que la energía eléctrica suele ser el factor que permita el movimiento del capital; en mucho caso el objetivo es encontrar la oferta o la función de costos de energía eléctrica y se utiliza como factor explicativo el capital y el trabajo (Maloney, 2001). Pero también nos enfrentamos al inconveniente de que el capital en el muy corto plazo no tiene una gran movilidad como factor productivo, en cambio el consumo de energía eléctrica es afectado por favor diversos factores como precio, ingreso, número de usuarios (Morales y otros 2012) lo que implica que tiene un mayor grado de movilidad, esto genera una distracción para el modelo ya que podría no estimar la relación que tiene originalmente se consideró.

Respecto a dicha variable se puede mencionar que existen dos grandes inconvenientes, el primero es que la información disponible en el INEGI sobre consumo de energía solo está disponible hasta 1995 a nivel nacional. Los datos desagregados para cada entidad federativa ya sea en volumen físico o valor de las ventas de kilowatts/hora, están disponibles en el Banco de Información Económica del INEGI en periodicidad mensual desde 01/2009 hasta 01/2015; además se pudo extender la serie gracias a la publicación “El Sector Energético en México 2009 ” hasta 2003.

Dada la información disponible de esta variable no permitirá extender la serie hacia el pasado ya que en periodos anteriores a 2003 la producción y sus valores se presentan por zonas y no por

---

<sup>10</sup> Solo los estados de Chiapas, Querétaro, Sonora y Tabasco presentaron una varianza menor para el caso de los resultados generados con la ecuación lineal.

entidades federativas. La ventaja para el uso de esta variable es que se puede trimestralizar tan sólo generando una suma de los totales mensuales. Adicionalmente puede ser considerada como uno de los factores que miden la riqueza de los individuos, ya que las ventas por estados de este bien se miden tanto para hogares como para empresa en conjunto.

#### **2.4.4 Captación y empleados bancarios**

Otra variable disponible para realizar el análisis es la captación bancaria, misma que se presenta como sigue:

$Capt_{ikt}$  = Captación bancaria trimestral promedio a precios reales de 2008, misma que se sumó para obtener los valores anuales, para el estado  $i$  ( $i$  = Aguascalientes, ..., Zacatecas), para el año  $t$  ( $t$  = 2003, 2004, ..., 2010), en el grupo  $k$ .

La captación bancaria integral se define, según el Banco de México, como “*la realizada vía instrumentos tradicionales de la banca, aceptaciones bancarias y los valores gubernamentales que la banca maneja por sus clientes o intermediaciones por cuenta de terceros*”<sup>11</sup>. Esa variable indica implícitamente un flujo de ingresos, mismo que se relaciona con que las personas al recibir sus pagos mediante algún instrumento financiero (nóminas, transferencias, cheques, etc.) generarán un movimiento en dicha variable; por otro lado si las personas recibieran sus pagos en efectivo de cuentas bancarias de nómina o en efectivo, en caso de ser en efectivo se considera que eventualmente dicho flujo de efectivo generará un depósito que será registrado dentro de la captación. De hecho González (1980), en su investigación sobre la captación bancaria en México, observa que las variables monetarias son muy sensibles ante cambios de la economía real así como cambios psicológicos de los agentes económicos.

Cabe resaltar que se esperaría que los niveles de dicha variable sean menores en los estados con un PIB menor, el mismo comportamiento sería esperable en los estados de la república con mayor grado de informalidad siempre y cuando la actividad económica manejada por dichos sectores sea relevante (Valero y Flores, 2003).

La Comisión Nacional Bancaria y de Valores (CNBV) presenta la información mensual de la captación bancaria calculando el promedio diario para cada mes; para obtener el valor trimestral solamente se realizó una suma del valor correspondiente a cada mes. Dicha variable tiene algunos

---

<sup>11</sup> Banco de México, Glosario. <http://www.banxico.org.mx/divulgacion/glosario/glosario.html>

inconvenientes, principalmente es la disponibilidad de la información sin que haya cambios en su definición<sup>12</sup>.

Adicionalmente, la CNBV cuenta con información correspondiente al número de empleados de las sucursales bancarias, por entidad federativa, dicha información es generada mensualmente. Esta información se puede encontrar desde enero del 2000 hasta la fecha. Nuevamente, al igual que la captación bancaria, en dicha información pública se encuentre distintos tipos de inconsistencias, por ejemplo, para el estado de Michoacán del 2000 al 2003 mantiene una plantilla de alrededor de 2,300 empleados, y en diciembre del 2000 el total de empleado era superior a los 11,000; dicha situación se repite en algunos otros estados, en tal caso se sustituyó el dato por el promedio del resto de los meses; otra falla observada dentro de la información se mostraba cuando los datos oscilaban en una cifra y repentinamente cambiaban sus valores, esto sucedía conforme al cambio en el orden de los estados, ya que su orden era modificado alfabéticamente pero sus datos no eran modificados.

---

<sup>12</sup> Es decir, durante el periodo de análisis (1993-2013) se presentan al menos 3 cambios de definición en el concepto de captación bancaria; en 1993, 2003 y otro en 2011. Esto complica el uso de la presente variable como parte de las variables explicativas, además debemos recordar que la información presentada por la variable captación bancaria es información precios corrientes, misma que se ha deflactado con el deflactor de PIB correspondiente a cada entidad con base en 2008 para darle a la información la característica de ser real. Fuente: <http://www.cnbv.gob.mx/Paginas/Informacion-Estadistica.aspx>

## CAPÍTULO 3

### 3. Estimación del modelo

Una vez que se ha considerado la información anteriormente expuesta podemos mencionar que nuestro modelo se transforma, para generar las estimaciones anuales, en:

$$PIB_{Est\ tik} = \alpha + \beta k X_{tik} + e_{tik} \quad (3.1)$$

$PIB_{Est\ tik}$  = Producto interno bruto estatal del estado  $i$ ,  
 $t = 1993, \dots, 2013$ ;  $k$  = número de región, 1, 2, 3 o 4.

Para obtener las estimaciones trimestrales se generará la siguiente ecuación:

$$PIB_{Est\ tik} = A + B X_{tik} + e_{tik} \quad (3.2)$$

Donde el subíndice  $j$  representa cada uno de los trimestres que conforman el año y

$A = \alpha / 4$  y  $e = E / 4$ ,

Podemos observar que hay cambios en la ecuación a estimar ya que el valor del intercepto de la ecuación estimada con datos anuales debe ser dividido entre cuatro, puesto que el año está formado por 4 trimestres; en cambio el coeficiente se aplica directamente debido a que la magnitud y la dirección de los movimientos los imprime, implícitamente, las variables explicativas. Las anteriores ecuaciones permitirán estimar el PIB estatal en cada trimestre y para cada estado del país.

Chow y Lin (1971) muestran que la distribución uniforme (dividir entre el número de subperiodos) del error es la más sencilla de aplicar; ya que al incorporar los efectos de la autocorrelación, la distribución de los errores se transforma en un ejercicio matemático de gran complejidad para resolver. Además, al considerar que la autocorrelación es provocada por la falta de variables explicativas, cambiar la distribución del error no generaría un beneficio adicional en la estimación.

Cuando se pretende realizar un análisis de predicción con datos trimestrales de variables anuales, es preciso tener en cuenta el cúmulo de información de que se dispone para concretar las estimaciones anuales objetivo (Quillis, 2001), de esa manera se podrán tener estimaciones más cercanas a las observadas, pero cabe resaltar que los datos mencionados tienen el inconveniente de

### 3.1 Aspecto de las series de tiempo.

Antes de continuar con el análisis de las variables conviene mencionar algunas notas sobre las relaciones de las variables a través del tiempo.

En primer lugar las series que se analizan a través del tiempo están constituidas por tres elementos los cuales muestran el comportamiento de la misma; estos son (Anderson y otros 2011):

- Componente de tendencia: en una serie de tiempo, este componente se refiere al cambio gradual de los valores de la serie en una determinada dirección, cabe resaltar que los cambios son resultado de algunos factores que influyen en el largo plazo. Digamos, la estructura demográfica o las modificaciones en la estructura productiva de una región.
- Componente cíclico: con este componente se puede observar cuando los puntos de observación se están por encima o por debajo de los puntos marcados por la tendencia; Para ciclos más amplios se considera la información correspondiente para periodos de al menos un año.
- Componente estacional: consta del comportamiento regular de una serie a durante un periodo de tiempo, un buen ejemplo es el comportamiento de la actividad económica trimestral en México.
- Componente irregular: coloquial mente lo pudiéramos considerar el factor sorpresa, es al movimiento aleatorio de las series provocado por factores desconocidos y difíciles de predecir.

Cuando analizamos los componentes de dos series distintas se podría encontrar que esta pueden tener la misma tendencia a lo largo de los años, por ejemplo que el ritmo de crecimiento de una ciudad sea el mismo que el crecimiento de una persona; en determinado momento que se compare a el crecimiento de un bebe con el crecimiento de un árbol recién plantado durante un periodo, se podrá observar que ambos están creciendo simultáneamente, y si en determinado momento quisiéramos encontrar una medida de correlación, muy posiblemente se encuentre y esta sea positiva y significativa, pero entre ellos no existe causalidad alguna ya que en ambos se está manifestando el fenómeno del crecimiento, mas no hay una interdependencia de entre ambas variables, el crecimiento del bebe y el crecimiento del árbol recién plantado.

Esta relación aparente se puede presentar cuando dos series distintas tienen la misma tendencia pero en realidad no tienen alguna relación (Cryer y Chan, 2008). Para eliminar los efectos de la tendencia

en estos casos se recurre a la estimación de un modelo en primeras diferencias (Montero, 2013) para eliminar los efectos de la tendencia.

Implícitamente la estimación mediante tasas de crecimiento establece la relación en primeras diferencias con el objetivo de romper las relaciones de tendencia entre la variable dependiente y las variables independientes.

Tomando en cuenta las notas anteriores sobre series de tiempo, una transformación del modelo a estimar sería la siguiente:

$$RatePIBEst_{tik} = a + b_k rateX_{tik} + e_{tik} \quad (3.3)$$

En donde:

$RatePIBEst_{tik}$  = Tasa de crecimiento del Estado i en el periodo t, correspondiente al grupo k.

Para el cálculo trimestral

$$RatePIBEst_{tikj} = \frac{a}{4} + \frac{b_k}{4} rateX_{tikj} + e_{tikj} \quad (3.4)$$

Para tasas de crecimiento muy pequeñas la ecuación (3.4) se cumplirá o al menos se aproximará a las tasas de crecimiento anual.

### 3.2 Análisis de correlación

En primer lugar se realiza un ejercicio de correlación entre las tasas de crecimiento del PIB de los estados con la tasa de crecimiento de los trabajadores asegurados ( $rateIMSS$ ), con la tasa de crecimiento del consumo de Kwh ( $rateKWH$ ) y con la tasa de crecimiento de la captación; lo cual permitirá observar las relaciones que se presentan a través del tiempo eliminando los efectos de las tendencias naturales de crecimiento a través del tiempo.

Dichas correlaciones se estiman para distintos periodos, mismos que están regidos por la disponibilidad de información.

**Tabla 3.1 RatePIB y RateIMSS. Coeficientes de correlación por periodos.**

	Entidades	1994-2003	2004-2013	1994-2013
<b>Norte</b>	<b>Baja California</b>	0.775	0.970	0.836
	<b>Chihuahua</b>	0.551	0.888	0.682
	<b>Coahuila</b>	0.412	0.939	0.726
	<b>Nuevo León</b>	0.808	0.901	0.817
	<b>Sonora</b>	0.732	0.962	0.770
	<b>Tamaulipas</b>	0.519	0.851	0.646
<b>Centro Norte</b>	<b>Aguascalientes</b>	0.735	0.804	0.765
	<b>Baja California Sur</b>	0.128	0.871	0.566
	<b>Colima</b>	0.236	0.832	0.314
	<b>Durango</b>	0.041	0.643	0.190
	<b>Jalisco</b>	0.854	0.904	0.836
	<b>Michoacán</b>	0.250	0.482	0.336
	<b>Nayarit</b>	0.316	0.678	0.388
	<b>San Luis Potosí</b>	0.582	0.730	0.620
	<b>Sinaloa</b>	0.394	0.713	0.552
<b>Zacatecas</b>	-0.253	0.216	-0.102	
<b>Centro</b>	<b>Distrito Federal</b>	0.786	0.684	0.764
	<b>Guanajuato</b>	0.619	0.817	0.647
	<b>Hidalgo</b>	0.771	0.633	0.731
	<b>México</b>	0.805	0.780	0.793
	<b>Puebla</b>	0.789	0.696	0.750
	<b>Querétaro</b>	0.744	0.782	0.739
	<b>Tlaxcala</b>	0.657	0.340	0.540
<b>Sur</b>	<b>Campeche</b>	0.043	0.387	0.287
	<b>Chiapas</b>	0.382	0.684	0.458
	<b>Guerrero</b>	0.186	0.367	0.165
	<b>Morelos</b>	0.713	0.897	0.761
	<b>Oaxaca</b>	0.039	0.099	-0.008
	<b>Quintana Roo</b>	0.685	0.852	0.673
	<b>Tabasco</b>	0.619	-0.110	0.204
	<b>Veracruz</b>	0.471	0.426	0.383
	<b>Yucatán</b>	0.701	0.568	0.630

Fuente: Cálculos propios con información del INEGI e IMSS.

La tabla 3.1 muestra que a excepción de Hidalgo y Puebla, la correlación del periodo 2004-2013 es mayor que en el periodo 1994-2013. Además, Oaxaca, Zacatecas y Tabasco en algunos de los periodos mantienen una correlación negativa de al menos -10%; esto implica

que en la medida que se incrementa el PIB aparece una reducción del número de empleados, de hecho si el factor trabajo fuera limitativo para la producción, al incrementarse la disponibilidad de este factor, había crecimientos directos en la producción<sup>13</sup>.

Cuando se considera el periodo completo para Oaxaca, la correlación es cercana a cero (-0.008), una posible explicación a este fenómeno es que la entidad mencionada es una de las que contó con el mayor número de porcentaje de población participante en el sector informal durante 2013, 81.2% (INEGI, 2014). Al tratar la información por grupos, podemos observar que el grupo con las relaciones más altas para cualquiera de los tres periodos analizados es la región Norte, es decir en tal caso se cumple la teoría de que al aumentar el producto, se ve incrementada la demanda por el factor productivo trabajo.

Por otro lado; la región sur se caracteriza por incluir las entidades con menor relación entre rateIMSS y ratePIB (Campeche, Guerrero, Oaxaca y Tabasco). Con estas relaciones es posible esperar que la región Norte sea la que tenga mejor ajuste al modelar el comportamiento del PIB trimestral.

La región Centro se caracteriza por tener también coeficientes de correlación mayores a 0.5 en todos los casos, en cambio la región Centro Norte tiene dentro de sus elementos a Zacatecas, dicha entidad presenta una correlación menor a -0.2 (dicha entidad presenta una tasa de informalidad del 67% según el reporte del INEGI), y a Durango que tiene una correlación de 0.19 para 1994-2013 (con un sector informal del 58%).

Además de la variable rateIMSS, también se propone utilizar la tasa de crecimiento del volumen de ventas de energía eléctrica en cada entidad (rateKwh). La cual tiene el objetivo de representar el uso intensivo del capital así como un factor de producción que participa en el proceso de generación del PIB (Nicholson, 2001); cabe precisar que existe un factor de imprecisión en el uso de Kwh como medida de intensidad en el uso del capital que se

---

<sup>13</sup> Según la teoría de la producción esto no debería suceder esto, ya que al incrementarse la producción (PIB) se tendría que incrementar la demanda por factores productivos. Sólo en el caso en que el capital resultará más productivo se desplazaría trabajo para integrar más capital en el proceso productivo. (Vilcapona, 1996)

deriva del efecto del uso más eficiente de los energéticos derivado del cambio tecnológico a través del tiempo<sup>14</sup>.

La tabla 3.2 presenta la relación existente entre el ratePIB y rateKwh donde la información anual se obtuvo mediante la suma del volumen de los Kwh vendidos durante cada mes del año y la información trimestral fue el resultado de la suma de los datos del trimestre correspondiente. Los datos para el periodo 01/2003 – 12/2008 se obtuvieron de la publicación “El Sector Energético 2012” del INEGI, en tanto que la información a partir del 2009 se obtuvo del Banco de Información Económica (BIE) del INEGI.

**Tabla 3.2. Coeficiente de correlación RatePIB vs RateKwh entidad federativa. 2003-2012**

<b>Norte</b>	<b>Baja California</b>	0.798	<b>Centro</b>	<b>Distrito Federal</b>	0.166
	<b>Chihuahua</b>	0.385		<b>Guanajuato</b>	-0.098
	<b>Coahuila</b>	0.253		<b>Hidalgo</b>	0.547
	<b>Nuevo León</b>	0.590		<b>México</b>	0.664
	<b>Sonora</b>	0.536		<b>Puebla</b>	0.832
	<b>Tamaulipas</b>	0.277		<b>Querétaro</b>	0.484
				<b>Tlaxcala</b>	0.117
<b>Centro Norte</b>	<b>Aguascalientes</b>	0.388	<b>Sur</b>	<b>Campeche</b>	-0.336
	<b>Baja California Sur</b>	0.529		<b>Chiapas</b>	-0.466
	<b>Colima</b>	0.278		<b>Guerrero</b>	0.204
	<b>Durango</b>	0.334		<b>Morelos</b>	0.834
	<b>Jalisco</b>	0.437		<b>Oaxaca</b>	0.247
	<b>Michoacán</b>	0.923		<b>Quintana Roo</b>	0.527
	<b>Nayarit</b>	0.457		<b>Tabasco</b>	-0.417
	<b>San Luis Potosí</b>	0.906		<b>Veracruz</b>	0.273
	<b>Sinaloa</b>	-0.342		<b>Yucatán</b>	0.070
	<b>Zacatecas</b>	0.309			

Fuente: Cálculos propios con información del INEGI. Banco de Información Económica.

Los datos correspondientes al Distrito Federal presentan algunos problemas, ya que durante el periodo 01/2009- 05/2010, las cifras aparecen en cero o representan un valor muy pequeño en comparación con los meses subsecuentes. Para mitigar los efectos generados

<sup>14</sup> En los estudios realizados por Apergis y Payne (2010), Campo y Sarmiento (2011) y Barreto y Nieto (2012) entre otros, mencionan que los países de América Latina son dependientes del sector energético para realizar su producción. Es decir, la energía eléctrica no sólo reflejaría el uso intensivo del capital sino el consumo de esta como factor de producción independiente.

por esta omisión de datos, se utiliza la tasa de crecimiento promedio anual presentada entre 2008 y 2011 (-0.67 por ciento) que son los periodos más cercanos al punto donde se omiten datos.

Al funcionar Kwh también como variable instrumental del capital de las empresas, y considerando que en el corto plazo el intercambio en el uso de los factores productivos (capital=Kwh y trabajo) es poco flexible, la tasa marginal de sustitución entre ambos factores se mantiene fija. La consecuencia de este razonamiento es simple: podemos usar el factor trabajo como variable independiente para estimar el PIB estatal, o podemos utilizar el factor capital para realizar dichas estimaciones, tal explicación supone que en el periodo analizado la relación capital/trabajo es estable.

Continuando con la interpretación de los resultados presentados en la tabla 4.2, se puede observar que la región Norte es la única donde todas las entidades federativas que la conforman presentan coeficientes de correlación positivos entre ratePIB y rateKwh, de hecho la entidad que presenta una menor correlación es Tamaulipas (0.277). Por otro lado, los estados con correlación negativa son: en la región Centro, Guanajuato (-0.098); Sinaloa (-0.342) en la Centro Norte; y Campeche (-0.336), Chiapas (-0.466) y Tabasco (-0.417) correspondientes a la región Sur. Salvo el caso de Guanajuato, la correlación en el resto de los estados recientemente mencionados, es relativamente alta, dicho signo muestra que al incrementarse la tasa de crecimiento del PIB se redujo la tasa decrecimiento de las ventas de Kwh en dichas entidades<sup>15</sup>.

---

<sup>15</sup> Esto parece ser contradictorio, para desmentir dichas relaciones se puede analizar los componentes del PIB de cada una de las entidades federativas, entre ellos los componentes del sector eléctrico. En todo caso, dicho análisis merece una amplia dedicación, mismo que implicaría el desarrollo de otro documento de investigación relacionado con la utilización de factores productivos. A pesar de ellos se puede mencionar que en dichos estados existen actividades relevantes que podrían estar perturbando el comportamiento convencional; por ejemplo: Campeche se relaciona con la actividad petrolera, en Chiapas al igual que las otras entidades mencionadas destacan el “Comercio” y “Servicios inmobiliarios y de alquiler de bienes muebles e intangibles” estas últimas actividades están relacionadas ampliamente con el turismo, por tanto se tendría que analizar la relación de este con el consumo de energía eléctrica, sobre todo por las cuestiones climatológicas, etc. Además se debería realizar un análisis por entidad federativa en el que se relacionen el tipo de cambio, el número de empleados, temperatura, etc. Es decir, variables que afectan el movimiento del consumo de energía eléctrica y el crecimiento del PIB para dar una mejor explicación del fenómeno. Cálculos propios con datos del Sistema de Cuentas Nacionales del INEGI.

Considerando las variables generadas con información del sistema financiero que cumplen con el requisito de ser publicadas trimestralmente, se encuentran el número de sucursales bancaria en las entidades, la captación bancaria promedio mensual y el número promedio de empleados bancarios en el trimestre. La dificultad que se observa con la primera variable mencionada radica en que tiene una menor movilidad, es decir el uso de las sucursales bancarias como indicador de la actividad económica reflejaría movimientos rezagados o tardíos en las regiones; por otro lado, las fluctuaciones económicas podrán observarse con mayor profundidad al revisar la información correspondiente a la captación bancaria.

La tabla 3.3 muestra la relación que existe entre la tasa de crecimiento del PIB de las entidades federativas y la tasa de crecimiento de la captación bancaria, agrupando la información por las regiones sugeridas. En ella se puede observar, nuevamente que la región Norte es la que presenta una mejor correlación respecto a las otras regiones<sup>16</sup>.

Recordemos que la captación bancaria, por definición es el promedio de los saldos monetarios en el total de las cuentas bancarias que conforman el sistema financiero en cada entidad federativa<sup>17</sup>, y para hacerlo comparable con la serie de producción a precios de 2008, se deflactó el valor de cada trimestre, para cada entidad federativa, respecto al Índice Nacional de Precios al Consumidor (INPC) correspondiente del trimestre con base en el promedio del 2008, lo cual permite hacer comparaciones de variables reales en el caso correspondiente.

En el presente documento es conveniente mencionar que la correlación positiva de la mayoría de las Entidades federativas entre su PIB y su captación bancaria tiene ciertas implicaciones; por ejemplo, que al incrementarse la tasa de crecimiento de la economía, la tasa de crecimiento de la captación bancaria se incrementa, mostrando así que es un buen indicador para facilitar el cálculo del crecimiento del PIB. Dicha interpretación tiene su origen en que al incrementarse la actividad económica se debe incrementar el flujo de

---

<sup>16</sup> La captación bancario en primer lugar antes del año 2000 los registros eran llamados saldos relevantes, es decir, solo se presentaba la captación de la o las principales ciudades de la entidad, posteriormente se realizan 3 cambios de definición del captación bancaria, mismos que se sospechan fueron en los periodos 2001-2002, 2008-2009 y por último 2010-2011. De hecho, a partir del segundo trimestre del 2011, la información se presenta por Institución financiera y no por entidad federativa.

<sup>17</sup> Fuente: Comisión Nacional Bancaria y de Valores.

efecto, de tal manera que aquellos que venden sus bienes y servicios en el mercado, depositan sus ingresos en los bancos, y estos últimos les dan un nuevo uso en la economía mediante préstamos que ofrecen al público en general, todo tiene como origen el incremento en la actividad económica real. Conviene agregar que el periodo analizado en la tabla 3.3 corresponde a 2001-2010, ya que es el periodo de coincidencia de disponibilidad entre ambas variables

**Tabla 3.3 RatePIB y RateCapt por entidad federativa. 2001-2010.**

<b>Norte</b>	<b>Baja California</b>	-0.061	<b>Centro</b>	<b>Distrito Federal</b>	0.432
	<b>Chihuahua</b>	0.402		<b>Guanajuato</b>	0.427
	<b>Coahuila</b>	0.070		<b>Hidalgo</b>	0.011
	<b>Nuevo León</b>	0.263		<b>México</b>	0.296
	<b>Sonora</b>	0.413		<b>Puebla</b>	0.232
	<b>Tamaulipas</b>	-0.086		<b>Querétaro</b>	0.341
				<b>Tlaxcala</b>	0.472
<b>Centro Norte</b>	<b>Aguascalientes</b>	0.448	<b>Sur</b>	<b>Campeche</b>	-0.001
	<b>Baja California Sur</b>	0.249		<b>Chiapas</b>	0.278
	<b>Colima</b>	0.296		<b>Guerrero</b>	0.206
	<b>Durango</b>	0.254		<b>Morelos</b>	0.051
	<b>Jalisco</b>	0.355		<b>Oaxaca</b>	0.049
	<b>Michoacán</b>	0.203		<b>Quintana Roo</b>	0.007
	<b>Nayarit</b>	0.135		<b>Tabasco</b>	0.363
	<b>San Luis Potosí</b>	0.264		<b>Veracruz</b>	0.543
	<b>Sinaloa</b>	0.064		<b>Yucatán</b>	0.455
	<b>Zacatecas</b>	0.324			

Fuente: Cálculos propios con información del INEGI y la Comisión Nacional Bancaria y de Valores.

En la anterior tabla también aparecen algunas entidades con correlación negativa, Campeche (-0.001), Tamaulipas (-0.086) y Baja California (-0.061), pero este indicador es pequeño.

Como parte del análisis de regresión, se presentan las correlaciones cruzadas entre las variables independientes y la correlación la variable dependiente. La tabla 3.4 muestra el resumen de la correlación de ratePIB con el resto de las variables explicativas, se observa claramente que para la mayoría de las entidades el mayor grado de relación se muestra

través de la correlación entre ratepib y rateimss; en cambio el menor grado de correlación se obtiene al comparar ratepib con las variable ratebank y ratecapt.

**Tabla 3.4 Correlación de la tasa de crecimiento del PIB vs. Tasa de crecimiento de las variables explicativas: rateimss, ratebank, ratecapt y ratekwh.**

Entidad	Variable	Rateimss (1994-2013)	Ratebank (2001-2013)	Ratecapt (2001-2010)	Ratekwh (2004-2012)
Baja California	ratepib	0.836	0.19	-0.06	-0.09
Chihuahua	ratepib	0.682	-0.19	0.40	0.74
Coahuila	ratepib	0.726	0.34	0.07	-0.31
Nuevo León	ratepib	0.817	0.30	0.26	0.54
Sonora	ratepib	0.770	0.17	0.41	0.37
Tamaulipas	ratepib	0.646	0.09	-0.09	0.56
Aguascalientes	ratepib	0.765	-0.18	0.45	0.02
Baja California Sur	ratepib	0.566	-0.07	0.25	-0.74
Colima	ratepib	0.314	-0.14	0.30	-0.20
Durango	ratepib	0.190	-0.04	0.25	0.32
Jalisco	ratepib	0.836	0.37	0.35	-0.92
Michoacán	ratepib	0.336	0.56	0.20	-0.92
Nayarit	ratepib	0.388	0.00	0.13	-0.74
San Luis Potosí	ratepib	0.620	0.36	0.26	0.33
Sinaloa	ratepib	0.552	-0.49	0.06	-0.68
Zacatecas	ratepib	-0.102	0.15	0.32	0.14
Distrito Federal	ratepib	0.764	0.39	0.43	0.10
Guanajuato	ratepib	0.647	-0.03	0.43	0.17
Hidalgo	ratepib	0.731	0.21	0.01	0.83
México	ratepib	0.793	-0.12	0.30	0.70
Puebla	ratepib	0.750	0.65	0.23	0.50
Querétaro	ratepib	0.739	0.09	0.34	0.04
Tlaxcala	ratepib	0.540	0.01	0.47	-0.11
Campeche	ratepib	0.287	-0.29	0.00	0.60
Chiapas	ratepib	0.458	0.35	0.28	0.66
Guerrero	ratepib	0.165	0.26	0.21	0.40
Morelos	ratepib	0.761	-0.01	0.05	0.57
Oaxaca	ratepib	-0.008	0.24	0.05	-0.48
Quintana Roo	ratepib	0.673	-0.34	0.01	-0.54
Tabasco	ratepib	0.204	0.38	0.36	-0.42
Veracruz de Ignacio	ratepib	0.383	0.05	0.54	-0.43
Yucatán	ratepib	0.630	0.05	0.45	-0.93

Cálculos propios con información de fuentes oficiales en los periodos disponibles.

La tabla anterior muestra el resumen de la relación de la tasa de crecimiento del PIB con respecto a las variables explicativas; adicionalmente se muestra el periodo disponibilidad de la información. Se observa claramente que la mejor relación de presente entre la tasa de crecimiento del PIB y los empleados totales reportados al IMSS. El periodo de coincidencia de información 2004-2006, dicho periodo es relativamente pequeño para atender la serie de tiempo en un modelo de regresión (Contreras, 2011) ya que los supuestos de la regresión lineal permanecen, pero existen algunas dificultades, por un lado el uso de muestras pequeñas incrementa la varianza de las variables independientes del modelo, y por otro lado este aumento genera que la varianza de los residuales de los coeficientes estimados se incrementen, esto provoca que se niegue al significancia del modelo de las variables explicativas.

Debido a que no existen pruebas para determinar los el nivel de error provocado en las muestras pequeñas, la única solución que se propone es disminuir lo más posible el número de variables explicativas para incrementar los grados de libertad del modelo. Esto limita la aplicación del modelo en cuanto al número de variables que se podrían utilizar (Greene, 1999).

La tabla 3.5 nos permite observar las correlaciones cruzadas entre la tasa de crecimiento de los trabajadores totales y el resto de las variables explicativas consideradas para la estimación; es notable que la relación más persistente, sobre todo con signo positivo es la que existe entre *rateimss* con *ratecatp*. Por otro lado, relación entre *rateimss* y *ratekwh* tienen correlación negativa mostrando a través del tiempo un efecto sustitución entre dichas variables para algunas entidades federativas.

Un tema que puede ser objeto de estudio en futuras ocasiones es verificar si en la medida que los estados crecen más, existe una relación positiva o negativa entre las dos variables anteriormente mencionadas (*rateimss* y *ratekwh*).

Obsérvese solamente como ejemplo los estados de Nuevo León, Tamaulipas y algunos estados del centro del país se han caracterizado por los tener las mayores tasas decrecimiento económico en los últimos 10 años, en estos casos existe una demanda mayor de los dos factores productivos (capital y trabajo) representados en las variables analizadas,

cabe mencionar que si las correlaciones entre variables explicativas son altas, entonces se estaría rompiendo uno de los supuestos del análisis de regresión (Greene, 1999). En la Tabla 3.5 se observa que si incluimos simultáneamente rateimss y ratekwh, en la mayoría de los estados estaríamos rompiendo los supuestos básicos de la regresión lineal simple, puesto que para muchos estados la correlación entre ambas series es relativamente alta.

**Tabla 3.5. Correlación entre RateIMSS vs Variables explicativas.**

	Entidad	variable	ratebank	ratecapt	ratekwh
<b>Norte</b>	<b>Baja California</b>	rateimss	0.14	0.03	-0.12
	<b>Chihuahua</b>	rateimss	-0.39	0.34	0.52
	<b>Coahuila de Zaragoza</b>	rateimss	0.20	0.04	-0.27
	<b>Nuevo León</b>	rateimss	0.61	0.37	0.35
	<b>Sonora</b>	rateimss	0.30	0.36	0.54
	<b>Tamaulipas</b>	rateimss	0.32	0.34	0.55
<b>Centro norte</b>	<b>Aguascalientes</b>	rateimss	0.17	0.48	-0.08
	<b>Baja California Sur</b>	rateimss	0.05	0.44	-0.68
	<b>Colima</b>	rateimss	0.09	-0.06	-0.41
	<b>Durango</b>	rateimss	-0.06	0.10	0.60
	<b>Jalisco</b>	rateimss	0.53	0.42	-0.87
	<b>Michoacán de Ocampo</b>	rateimss	-0.04	0.26	-0.41
	<b>Nayarit</b>	rateimss	0.07	0.44	-0.82
	<b>San Luis Potosí</b>	rateimss	0.40	0.42	0.34
	<b>Sinaloa</b>	rateimss	-0.01	0.25	-0.48
<b>Centro</b>	<b>Zacatecas</b>	rateimss	0.04	-0.21	0.20
	<b>Distrito Federal</b>	rateimss	0.54	0.42	-0.25
	<b>Guanajuato</b>	rateimss	-0.25	0.37	0.39
	<b>Hidalgo</b>	rateimss	0.21	0.09	0.81
	<b>México</b>	rateimss	-0.09	0.40	0.66
	<b>Puebla</b>	rateimss	0.73	0.01	0.32
	<b>Querétaro</b>	rateimss	0.12	0.33	0.44
	<b>Tlaxcala</b>	rateimss	0.40	0.01	-0.12
	<b>Sur</b>	<b>Campeche</b>	rateimss	-0.21	0.11
<b>Chiapas</b>		rateimss	0.40	0.53	0.22
<b>Guerrero</b>		rateimss	-0.22	0.41	0.87
<b>Morelos</b>		rateimss	0.42	0.33	0.71
<b>Oaxaca</b>		rateimss	0.31	0.00	-0.56
<b>Quintana Roo</b>		rateimss	-0.42	0.07	-0.71
<b>Tabasco</b>		rateimss	0.62	0.48	-0.14
<b>Veracruz de Ignacio</b>		rateimss	0.34	0.59	0.14
<b>Yucatán</b>		rateimss	0.28	0.13	-0.86

Fuente: Cálculos propios con información de la CNBV en INEGI.

Se observa que existe una amplia correlación entre empleo (rateIMSS) y el consumo de energía eléctrica (rateKwh)<sup>18</sup> para 14 entidades federativas; además de que 15 entidades tienen correlación negativa<sup>19</sup> lo que implica que, para estas entidades, un incremento en la tasa de crecimiento del trabajo, la tasa de crecimiento del factor capital se reduce, es decir existe un efecto sustitución entre uno y otro factor productivo. Teóricamente parece imposible, pero recordemos que en periodos donde el crecimiento del PIB no es alto el factor que es desplazado para reducir los costos de producción es el factor trabajo.

---

<sup>18</sup> Se considera una amplia correlación, para nuestro caso cuando esta es mayor o igual al 50%.

<sup>19</sup> Entre ellos: Baja California Sur, Nayarit, Oaxaca, Quintana Roo y Yucatán, quienes tuvieron las correlaciones significativamente más negativas.

En el periodo en que coinciden las variables presentadas en la tabla 3.6, las correlaciones son relativamente pequeñas, esto permite que en un modelo de regresión puedan participar la combinación de variables ratecapt-ratebank y ratecapt-ratekwh

**Tabla 3.6. Correlación entre RateCapt vs Variables explicativas.**

	Entidad	variable	Ratebank	ratekwh
<b>Norte</b>	<b>Baja California</b>	ratecapt	-0.23	-0.10
	<b>Chihuahua</b>	ratecapt	0.02	0.42
	<b>Coahuila de Zaragoza</b>	ratecapt	-0.28	-0.39
	<b>Nuevo León</b>	ratecapt	0.42	0.33
	<b>Sonora</b>	ratecapt	0.17	0.12
	<b>Tamaulipas</b>	ratecapt	0.36	0.37
<b>Centro norte</b>	<b>Aguascalientes</b>	ratecapt	0.07	0.23
	<b>Baja California Sur</b>	ratecapt	-0.12	-0.33
	<b>Colima</b>	ratecapt	-0.26	0.13
	<b>Durango</b>	ratecapt	-0.37	-0.36
	<b>Jalisco</b>	ratecapt	0.35	-0.17
	<b>Michoacán de Ocampo</b>	ratecapt	0.13	-0.17
	<b>Nayarit</b>	ratecapt	-0.05	-0.18
	<b>San Luis Potosí</b>	ratecapt	0.25	0.31
	<b>Sinaloa</b>	ratecapt	-0.13	-0.15
	<b>Centro</b>	<b>Zacatecas</b>	ratecapt	0.32
<b>Distrito Federal</b>		ratecapt	-0.01	-0.09
<b>Guanajuato</b>		ratecapt	0.28	0.01
<b>Hidalgo</b>		ratecapt	0.07	0.17
<b>México</b>		ratecapt	0.13	0.20
<b>Puebla</b>		ratecapt	0.27	-0.02
<b>Querétaro</b>		ratecapt	0.17	-0.29
<b>Tlaxcala</b>		ratecapt	0.33	0.33
<b>Sur</b>	<b>Campeche</b>	ratecapt	-0.13	0.10
	<b>Chiapas</b>	ratecapt	-0.09	-0.02
	<b>Guerrero</b>	ratecapt	0.05	0.19
	<b>Morelos</b>	ratecapt	0.03	-0.06
	<b>Oaxaca</b>	ratecapt	0.05	-0.28
	<b>Quintana Roo</b>	ratecapt	-0.05	-0.25
	<b>Tabasco</b>	ratecapt	0.19	-0.26
	<b>Veracruz de Ignacio</b>	ratecapt	0.11	-0.36
	<b>Yucatán</b>	ratecapt	0.19	-0.30

Fuente: Cálculos propios con información de la CNBV en INEGI.

En la tabla 3.7 se muestra la correlación existente entre el ratebank y ratekwh, en la mayoría de estos casos la correlación es inferior al 50%, lo cual muestra que estas variables no están correlacionadas; pero usar la variable ratebank arrastrará problemas de estimación

**Tabla 3.7 RatePIB y RateCapt por entidad federativa. 2001-2010.**

<b>Norte</b>	<b>Baja California</b>	-0.21	<b>Centro</b>	<b>Distrito Federal</b>	0.14
	<b>Chihuahua</b>	-0.49		<b>Guanajuato</b>	0.78
	<b>Coahuila</b>	0.01		<b>Hidalgo</b>	0.00
	<b>Nuevo León</b>	0.14		<b>México</b>	-0.30
	<b>Sonora</b>	-0.03		<b>Puebla</b>	0.00
	<b>Tamaulipas</b>	-0.42		<b>Querétaro</b>	0.41
				<b>Tlaxcala</b>	-0.15
<b>Centro Norte</b>	<b>Aguascalientes</b>	-0.36	<b>Sur</b>	<b>Campeche</b>	-0.33
	<b>Baja California Sur</b>	0.39		<b>Chiapas</b>	0.44
	<b>Colima</b>	0.25		<b>Guerrero</b>	-0.37
	<b>Durango</b>	0.59		<b>Morelos</b>	0.51
	<b>Jalisco</b>	-0.47		<b>Oaxaca</b>	-0.49
	<b>Michoacán</b>	-0.54		<b>Quintana Roo</b>	0.22
	<b>Nayarit</b>	0.70		<b>Tabasco</b>	0.05
	<b>San Luis Potosí</b>	-0.54		<b>Veracruz</b>	0.44
	<b>Sinaloa</b>	0.51		<b>Yucatán</b>	-0.15
	<b>Zacatecas</b>	0.63			

Fuente: Cálculos propios con información de la CNBV en INEGI.

## Capítulo 4

### 4. Resultados

Una vez que hemos presentado la teoría detrás del modelo de estimación se presentan las estimaciones de los mismos el primer conjunto de ecuaciones mismos que aplican para una regresión lineal simple en donde la tasa de crecimiento del PIB para cada estado está en función de la tasa de crecimiento del total de trabajadores del IMSS para los años 1998-2013, una vez que se agregan más variables independientes estas, los periodos de estimación se verán modificados. Por ejemplo al considerar el modelo general por ejemplo, al agregar la variable *rateKwh*, el periodo para realizar las estimaciones se reduce a 2001-2011. En cada caso se aplicarán distintos métodos de estimación entre ellos efectos fijos, efectos aleatorios (mínimos cuadrados generalizados y máxima verosimilitud) con el objetivo de precisar los resultados.

$$a) \text{RatePIB}_{i,j,t} = \alpha + \sum \beta X_{i,j,t} + e_{i,j,t}$$

En esta ecuación se considera a cada observación ya sea temporal o estatal como un dato no relacionado entre si, al menos como supuesto. El subíndice *i* considera las regiones económicas del Banco de México; el *j* cuenta a los estados de la república incluyendo el Distrito Federal y el subíndice *t* representa los periodos analizados

$$\begin{aligned} b) \quad & 1.- \text{RatePIB}_{1,j,t} = \alpha + \beta_1 \text{RateIMSS}_{1,j,t} + e_{1,j,t} \\ & 2.- \text{RatePIB}_{2,j,t} = \alpha + \beta_2 \text{RateIMSS}_{2,j,t} + e_{2,j,t} \\ & 3.- \text{RatePIB}_{3,j,t} = \alpha + \beta_3 \text{RateIMSS}_{3,j,t} + e_{3,j,t} \\ & 4.- \text{RatePIB}_{4,j,t} = \alpha + \beta_4 \text{RateIMSS}_{4,j,t} + e_{4,j,t} \end{aligned}$$

Las ecuaciones representan la estimación del modelo aplicado para cada región según la clasificación presentada en los informes regionales del banco de México, en este caso el subíndice *i* toma los valores de 1=Norte; 2=Centro Norte; 3=Centro y 4=Sur.

#### 4.1 Estimación modelo General. Variable *rateimss*.

La estimación correspondiente al modelo general

$$\text{RatePIB}_{i,j,t} = \alpha + \beta \text{RateIMSS}_{i,j,t} + e_{i,j,t} ; \text{ presenta los siguientes resultados:}$$

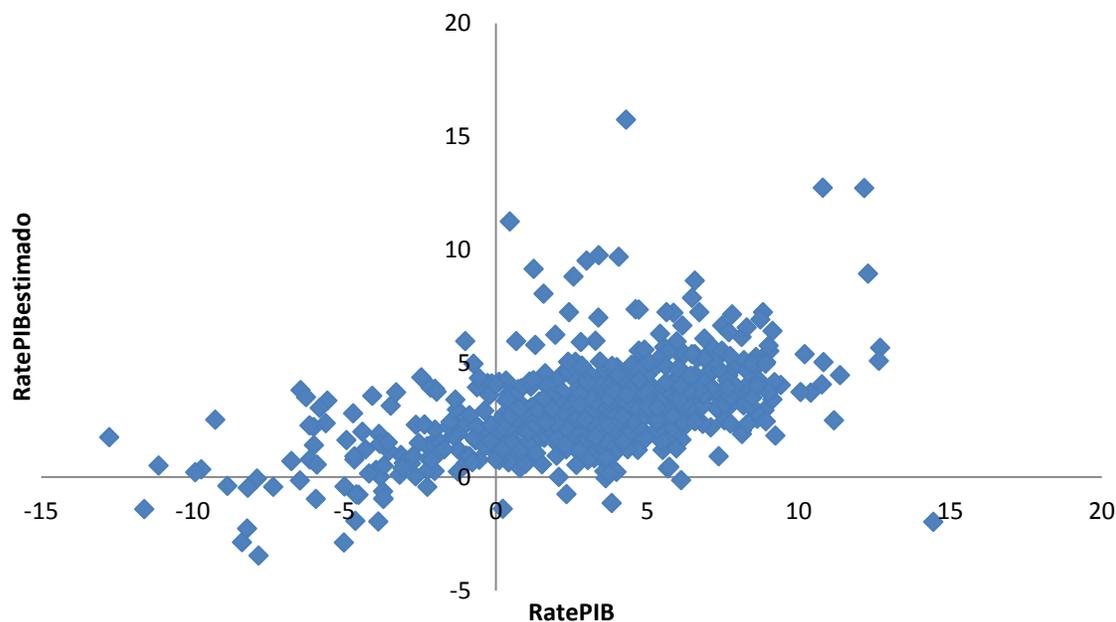
**Tabla 4.1 Resultados de estimación modelo general**

		Coef.	Std. Err.	T	P>t
Modelo general.	Rateimss	0.443525	0.0304763	14.55	0.000
Variable IMSS 1997-2013	_cons	1.324868	0.1769461	7.49	0.000

R2 ajustado=0.248 F(1,638)=211.79

El resultado es aparentemente válido con una R2 ajusta de 24%, se estima un valor de F de 211.79; se interpreta que el modelo es válido estadísticamente. Para probar de un posible heterocedasticidad se realiza la prueba de Breush-Pagan, en ella se obtiene que las varianzas son similares entre los grupos de Banxico, pero que si existe para los estados de manera aislada<sup>20</sup>. Aunado a esto podemos observar en la gráfica4.1 que existe los resultados estimados son mayores que los observados.

**Gráfico 4.1 Comparación de la tasa de crecimiento observada para los estados (ratePIB) vs los estimados por el modelo general (ratePIBestimado)**



De hecho podemos observar que si existe una relación positiva entre ratePIB y ratePIB estimado, pero con algunas fallas, específicamente las estimaciones que se encuentran en el cuadrante IV.

<sup>20</sup> En el “Anexo: Resultados de regresión”, apartado de modelo general se encuentran las pruebas estadísticas correspondientes.

## 4.2 Estimaciones Efectos Fijos

Una de las metodologías para estimación de la regresión lineal en datos panel, es la conocida como efectos fijos. Dicha metodología supone implícitamente que el comportamiento de los integrantes del grupo es similar.

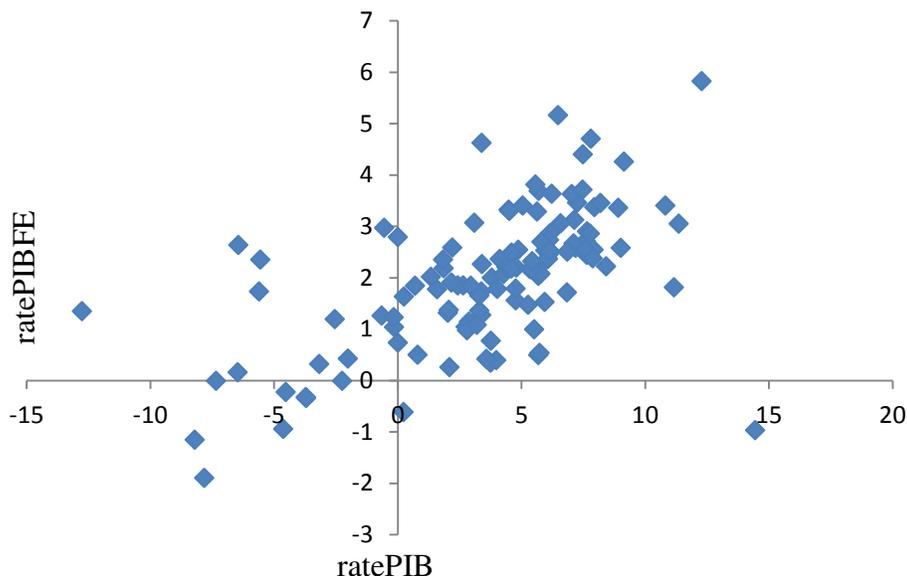
**Tabla 4.2 Resultados de estimación modelo efectos fijos por región**

		Coef.	Std. Err.	T	P>t
Región Norte	Rateimss	0.557412	0.072294	7.71	0.000
R2 =0.339; F(1,120)=59.45	_cons	1.785956	0.421053	4.24	0.000
Región Centro Norte	Rateimss	0.471938	0.066540	7.09	0.000
R2 =0.219; F(1,189)=50.30	_cons	1.309065	0.343942	3.81	0.000
Región Centro	Rateimss	0.626716	0.056472	11.10	0.000
R2 =0.498; F(1,132)=123.16	_cons	1.167627	0.308774	3.78	0.000
Región Sur	Rateimss	0.275521	0.050129	5.50	0.000
R2 =0.127; F(1,170)=30.21	_cons	1.08072	.3169115	3.41	0.001

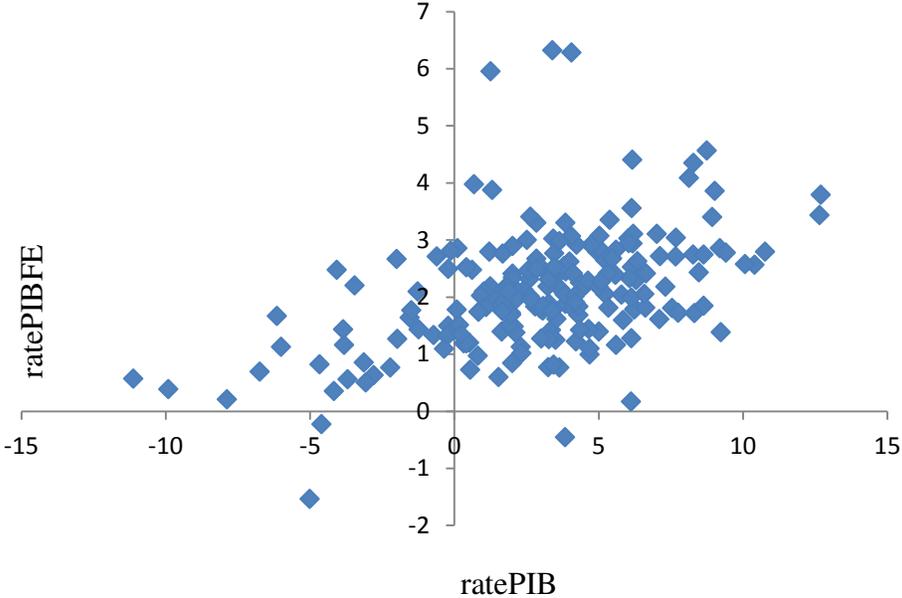
Podemos observar que en todos los casos las ecuaciones son significativas, dado que tenemos valores del estadístico F relativamente altos, considerando el estadístico de la R2, encontramos que el grupo con mejor ajuste en los resultados es la región Centro, seguida de la región Norte, la Centro Norte y al final la región Sur. A continuación se presentan los gráficos correspondientes:

**Gráfico 4.2 Comparación de la tasa de crecimiento observada para los estados (ratePIB) vs los estimados por el modelo de efectos fijos (ratePIBFE)**

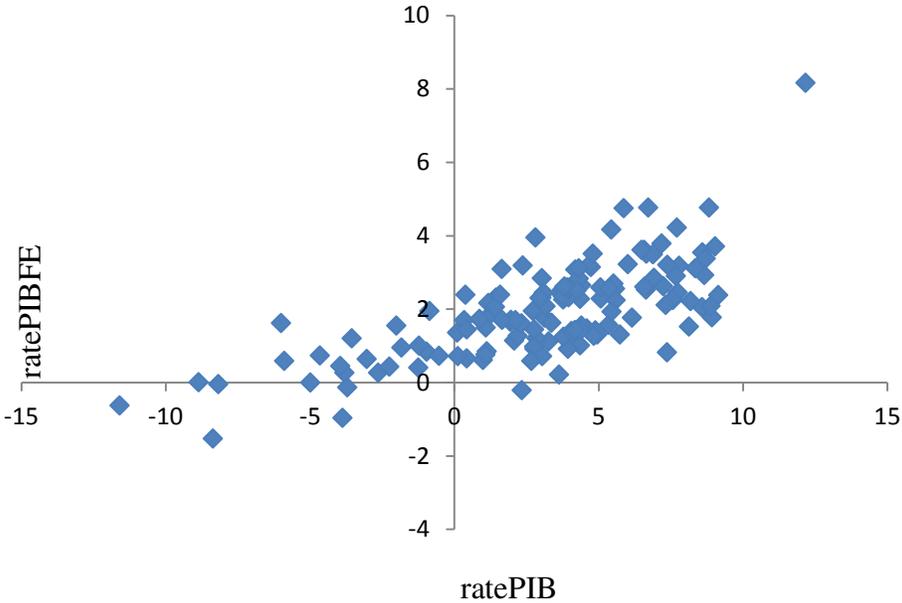
### Región Norte



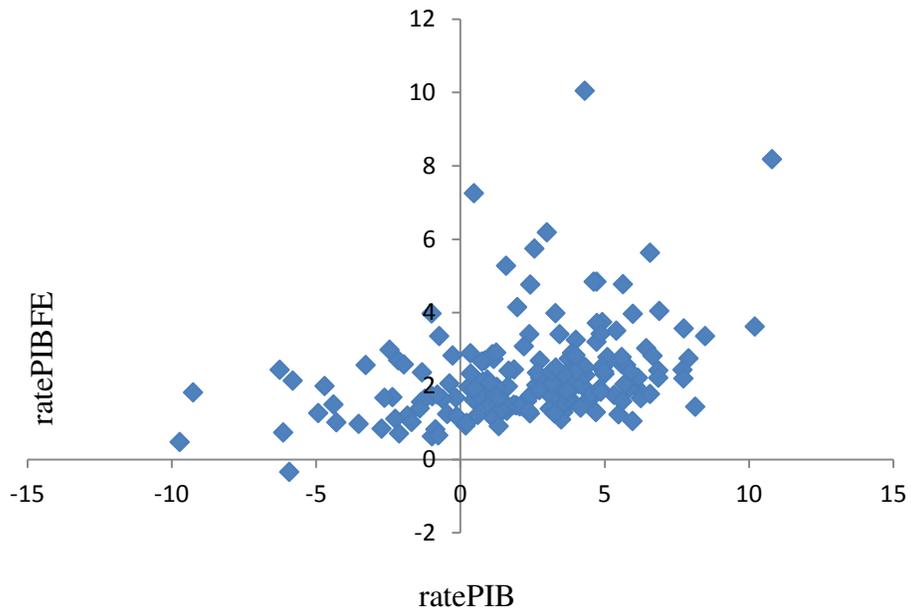
**Región Centro Norte**



**Región Centro**



## Región Sur



### 4.3 Estimación del modelo general. Todas las variables.

## Conclusiones

Como se ha presentado, tenemos grandes dificultades en México para obtener un PIB trimestral para cada entidad federativa, principalmente por la poca disponibilidad de información respecto a la información nacional, pero sobre todo a la periodicidad anual.

Una de las principales dificultades que se enfrentó es que la publicación de la información no es confiable, por wjwmppli

Pero en general se puede hablar de tres grandes conclusiones;

La primera está relacionada con la demanda de trabajo

- Productividad total de factores
- La sustitución de factores, la productividad total de factores.
- El modelo de convergencia
- Las fallas en la estimación.

La convergencia e un estado esta

## Referencia bibliográfica.

Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2011). *Métodos cuantitativos para los negocios*. Cengage Learning. Ed. 11ª. México.

Arana, F., Benítez y López, “Trimestralización de los valores añadidos sectoriales mediante indicadores. Aplicación al caso de Andalucía”. *Revista de Estudios Regionales, Universidades Públicas de Andalucía*, núm. 57, 2000, pp. 59-100, España. 2000.

Apergis, N. y J. Payne, “Renewable and non Renewable Energy Consumption- Growth Nexus: Evidence from a Panel Error Correction Model”. *Energy Economics*, 34, 733-738, 2012.

Cortiñas, P., F. Parra y J. Vicente Virseda, “Trimestralización con la macro solver de EXCEL”, *Econometria WorldPress*. 2008

Elizondo, R., “Estimaciones del PIB Mensual Basadas en el IGAE” *Documentos de Trabajo, Banco de México*, México, 2012.

Barreto, C. y J. Campos. “Relación a largo plazo entre consumo de energía y PIB en América Latina: Una evaluación empírica con datos panel”. *Revista Ecos de Economía*. Vol. 16 Núm. 35. Colombia, 2012.

Basile, R., “Quarterly Estimates for Italy's Government Accrual Data”. *Department of Economics, Università 'La Sapienza' di Roma*, Italia, 2009.

Bassie, V., “Economic forecasting”. *McGraw-Hill*, New York, 1958.

Boot J., G. Feibes, y W. Lisman. “Further methods of derivation of quarterly figures from annual data”. *Applied Statistics*, No. 1, Vol: 16, pp: 65-75. 1967.

Campo, J. y V. Sarmiento. “Relación consumo de energía y PIB: Evidencia desde un panel cointegrado de 10 países de América Latina entre 1971-2007”. *MPRA*, Paper 31772, 2011.

Chans, WS (1993). “Dissagregation of annual time-series data to quaterly figures – A comparative study”. *Journal of Forecasting*, No: 8, Vol: 12, pp: 677-688, 1993.

Chow, G. y Lin, A., "Best linear unbiased distribution and extrapolation of economic time series by related series", *Review of Economic and Statistics*, No. 4, Vol. 53, pp. 372-375, 1971.

Coraggio, José Luis (1994). “Territorios en transición. Crítica a la Planificación regional de América Latina”. *Universidad Autónoma del Estado de México*. 3ra edición.

Cryer, Jonathan, Kung-Sik Chang (2008). “Time Series Analysis with applications in R”. 2d Ed. Spirnger. Pp 501.

Denton , F., “Adjustment of monthly or quarterly series to annuals totals: An approach based on quadratic minimization”. *Journal of the American Statistical Association*, No: 333, Vol: 66, pp: 99-102, 1971.

Díaz-Bautista, Alejandro (2009). “Un modelo empírico de crecimiento económico y la nueva geografía económica urbana en México”. *Revista Paradigma Económico*, Facultad de Economía, Universidad Autónoma del Estado de México (UAEM), Vol. 1 Núm. 1. pp. 9-39.

Esquivel, Gerardo (1999). “Convergencia Regional en México, 19940-1995”. Documento de trabajo, *Centro de Estudios Económicos, El Colegio de México*. Núm. IX. México.

García, G., J. Pérez y J. Rodríguez. “Un Modelo de Predicción de alta frecuencia para la economía española”. Documento de trabajo, *Departamento de Economía Aplicada Universidad Autónoma de Madrid*. España, 2006.

González Méndez, Héctor E. (1980). Comportamiento de la Captación Bancaria en México. *Banco de México*. Documento No. 20. Mayo. México.

Greene, William H. (1999). “Análisis Econométrico”. Editorial Prentice Hall Iberia. 3ra. Edición.

H., Viv y J. McDermott. “A Quarterly Post-World War II Real GDP”. *Series for New Zealand. Motu Economic and Public Policy Research*. Motu Working Paper 07-13, 2007.

Harvey, A. y Pierse, R., “Estimating missing observations in economic time series”. *Journal of the American Statistical Association*, No: 385, Vol: 79, pp: 125-131, 1984

Isusquiza Martínez, Edgar Enrique (2014). *Desigualdad, crecimiento económico y descentralización fiscal: Un análisis empírico para México*. Premio Nacional de Finanzas Públicas 2014. México.

Maloney, Michael, (2001). *Economies and diseconomies: Estimating electricity cost function*. *Review of Industrial Organization*, 19(2), Sept. 2001, 165-180.

Montero. R (2013): Variables no estacionarias y cointegración. Documentos de Trabajo en Economía Aplicada. Universidad de Granada. España

Morales Ramírez, Dionicio; Luyando Cuevas, José Raúl y Flores Curiel, Daniel (2012). “Determinantes del consumo de energía eléctrica residencial de la Zona Metropolitana de Monterrey, Nuevo León, en México”. *Universidad & Empresa*, núm. 22, enero-junio, 2012, pp. 79-98 Universidad del Rosario. Bogota, Colombia

Nicholson, Walter (2001) Microeconomía Intermedia. *McGrawHill*. Ed. 8ª. México.

Organización Internacional del Trabajo (2014). *El empleo informal en México: situación actual, políticas y desafíos*. Notas sobre Formalización. Programa de Promoción de la Formalización para América Latina y el Caribe. OIT.

Proietti, T., *Temporal Disaggregation by State Space Methods: Dynamic Regression Methods Revisited*, European Commission Working Paper Studies, Luxemburgo, 2005.

Qualis, Enrique M. (2001). *Notas sobre la desagregación temporal de series económicas*. Papales de trabajo N.º 1. 01. *Instituto de estudios Fiscales*. Madrid España.

Rodríguez Feijoo, Santiago y Alejandro Rodríguez Caro (2000). *Los métodos de trimestralización sin indicador. Un estudio comparativo mediante simulación*. Documento de trabajo. Universidad de Las Palmas de Gran Canaria, España.

Rodríguez Feijoo, Santiago, Alejandro Rodríguez Caro y Delia Dávila Quintana, (2003). *La trimestralización de variables de flujo. Un estudio de simulación de los métodos con indicador de desagregación temporal*. Documento de trabajo. Universidad de Las Palmas de Gran Canaria, España.

Ruiz Durán, Clemente (2009). *México. Las dimensiones de la flexibilidad laboral*. Serie Macroeconomía del desarrollo. CEPAL. Santiago de Chile.

Trujillo, Aranda, Francisco, Benítez, Márquez, M<sup>a</sup> Dolores, López, Delgado, Pilar (2000). *Trimestralización de los valores añadidos sectoriales mediante indicadores. Aplicación al caso de Andalucía*. Revista Electrónica. Revista de Estudios Regionales

Valero Gil, Jorge y Daniel Flores Curiel (2003). “Tamaño del Sector informal y su potencial recaudatorio”. Universidad Autónoma de Nuevo León, Facultad de Economía, Centro de Investigaciones Económicas. Documento de trabajo realizado para el Servicio de Administración Tributaria.

Wackerly, Denis, William Mendenhall III y Richard L. Sheaffer. (2010). *Estadística Matemática con aplicaciones*. 7a Edición. Cengage Learning. México. D.F.

## Anexo1

Correlaciones de las variables relevantes.

Total Nacional				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.991	0.940	0.154
TAIMSS		1	0.930	0.066
KWH			1	0.165
Captación				1

Chiapas				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.943	0.890	0.024
TAIMSS		1	0.984	0.037
KWH			1	-0.067
Captación				1

Aguascalientes				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.983	0.960	0.030
TAIMSS		1	0.959	0.056
KWH			1	0.230
Captación				1

Chihuahua				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.711	0.836	0.234
TAIMSS		1	0.482	0.401
KWH			1	0.192
Captación				1

Baja California				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.963	0.954	0.281
TAIMSS		1	0.846	0.364
KWH			1	0.254
Captación				1

Coahuila de Zaragoza				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.986	0.834	0.161
TAIMSS		1	0.893	0.105
KWH			1	-0.149
Captación				1

Baja California Sur				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.945	0.958	0.124
TAIMSS		1	0.861	0.114
KWH			1	-0.019
Captación				1

Colima				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.967	0.936	-0.004
TAIMSS		1	0.932	-0.064
KWH			1	0.123
Captación				1

Campeche				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	-0.936	-0.980	0.171
TAIMSS		1	0.981	-0.007
KWH			1	-0.080
Captación				1

Distrito Federal				
	PIBEst	TAIMSS	KWH	Captación
PIBEst	1	0.990	0.854	0.198
TAIMSS		1	0.894	0.100
KWH			1	-0.152
Captación				1

<b>Durango</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.945	0.940	0.107
<b>TAIMSS</b>		1	0.878	-0.020
<b>KWH</b>			1	-0.124
<b>Captación</b>				1

<b>México</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.984	0.126	0.074
<b>TAIMSS</b>		1	0.111	-0.021
<b>KWH</b>			1	0.157
<b>Captación</b>				1

<b>Guanajuato</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.988	0.953	0.170
<b>TAIMSS</b>		1	0.979	0.036
<b>KWH</b>			1	-0.004
<b>Captación</b>				1

<b>Michoacán de Ocampo</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.926	0.328	0.035
<b>TAIMSS</b>		1	-0.015	-0.075
<b>KWH</b>			1	0.111
<b>Captación</b>				1

<b>Guerrero</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.884	0.940	0.065
<b>TAIMSS</b>		1	0.913	0.067
<b>KWH</b>			1	0.094
<b>Captación</b>				1

<b>Morelos</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.981	0.830	0.115
<b>TAIMSS</b>		1	0.745	-0.002
<b>KWH</b>			1	0.122
<b>Captación</b>				1

<b>Hidalgo</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.954	0.272	0.020
<b>TAIMSS</b>		1	0.228	-0.170
<b>KWH</b>			1	0.128
<b>Captación</b>				1

<b>Nayarit</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.985	0.966	0.050
<b>TAIMSS</b>		1	0.982	0.126
<b>KWH</b>			1	-0.027
<b>Captación</b>				1

<b>Jalisco</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.968	0.946	0.122
<b>TAIMSS</b>		1	0.992	0.004
<b>KWH</b>			1	0.056
<b>Captación</b>				1

<b>Nuevo León</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.988	0.963	0.075
<b>TAIMSS</b>		1	0.973	0.009
<b>KWH</b>			1	0.150
<b>Captación</b>				1

<b>Oaxaca</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.964	0.953	0.027
<b>TAIMSS</b>		1	0.970	-0.031
<b>KWH</b>			1	-0.024
<b>Captación</b>				1

<b>Sinaloa</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.945	0.912	0.007
<b>TAIMSS</b>		1	0.983	-0.042
<b>KWH</b>			1	0.037
<b>Captación</b>				1

<b>Puebla</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.955	0.962	0.159
<b>TAIMSS</b>		1	0.907	-0.121
<b>KWH</b>			1	-0.020
<b>Captación</b>				1

<b>Sonora</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.993	0.950	0.189
<b>TAIMSS</b>		1	0.937	0.172
<b>KWH</b>			1	0.133
<b>Captación</b>				1

<b>Querétaro</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.977	0.945	0.072
<b>TAIMSS</b>		1	0.985	0.036
<b>KWH</b>			1	-0.046
<b>Captación</b>				1

<b>Tabasco</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.986	0.977	0.034
<b>TAIMSS</b>		1	0.986	0.011
<b>KWH</b>			1	-0.052
<b>Captación</b>				1

<b>Quintana Roo</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.972	0.962	-0.003
<b>TAIMSS</b>		1	0.948	0.065
<b>KWH</b>			1	-0.096
<b>Captación</b>				1

<b>Tamaulipas</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.906	0.934	0.046
<b>TAIMSS</b>		1	0.768	0.270
<b>KWH</b>			1	0.171
<b>Captación</b>				1

<b>San Luis Potosí</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.995	0.902	0.090
<b>TAIMSS</b>		1	0.889	0.121
<b>KWH</b>			1	0.254
<b>Captación</b>				1

<b>Tlaxcala</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.151	0.456	0.064
<b>TAIMSS</b>		1	0.361	0.269
<b>KWH</b>			1	0.409
<b>Captación</b>				1

<b>Veracruz - Llave</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.965	0.951	0.206
<b>TAIMSS</b>		1	0.954	0.058
<b>KWH</b>			1	0.211
<b>Captación</b>				1

<b>Yucatán</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.975	0.970	0.409
<b>TAIMSS</b>		1	0.941	0.475
<b>KWH</b>			1	0.549
<b>Captación</b>				1

<b>Zacatecas</b>				
	<b>PIBEst</b>	<b>TAIMSS</b>	<b>KWH</b>	<b>Captación</b>
<b>PIBEst</b>	1	0.988	0.937	0.844
<b>TAIMSS</b>		1	0.928	0.919
<b>KWH</b>			1	0.685
<b>Captación</b>				1

## Anexo: Resultados de regresión

### a) Modelo general

. regress ratepib rateimss

```
Source |   SS   df    MS              Number of obs =   640
-----+-----+-----+-----+-----+-----
Model | 2508.58143   1 2508.58143          F( 1, 638) = 211.79
Residual | 7556.78138 638 11.8444849          Prob > F   = 0.0000
-----+-----+-----+-----+-----
Total | 10065.3628 639 15.7517415          R-squared   = 0.2492
                                          Adj R-squared = 0.2481
                                          Root MSE   = 3.4416
```

```
-----+-----
ratepib |   Coef.   Std. Err.   t   P>|t|   [95% Conf. Interval]
-----+-----+-----+-----+-----+-----
rateimss | .4435249   .0304763   14.55 0.000   .383679   .5033709
_cons | 1.324868   .1769461   7.49 0.000   .977401   1.672335
-----+-----
```

### a.1) Prueba de Breush-Pagan para grupos Banxico

. estat hettest banxico, rhs

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: banxico rateimss

chi2(2) = 2.78

Prob > chi2 = 0.2490

### a.2) Prueba de Breush-Pagan para entidades federativas

. estat hettest entidad1, rhs

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance

Variables: entidad1 rateimss

chi2(2) = 14.79

Prob > chi2 = 0.0006

### b) Modelo simple para efectos fijos

. by banxico, sort : xtreg ratepib rateimss, fe

### b.1) Estimaciones Banxico=1, Región Norte

```
Fixed-effects (within) regression      Number of obs   =   120
Group variable: entidad1                Number of groups =    6
```

```
R-sq:  within = 0.3447          Obs per group: min =   20
        between = 0.0502          avg   =   20.0
        overall = 0.3388          max   =   20
```

```
corr(u_i, Xb) = -0.0685          F(1,113)       = 59.45
                                          Prob > F       = 0.0000
```

```
-----+-----
ratepib |   Coef.   Std. Err.   t   P>|t|   [95% Conf. Interval]
-----+-----+-----+-----+-----+-----
rateimss | .557412   .0722947   7.71 0.000   .4141832   .7006409
_cons | 1.785956   .4210531   4.24 0.000   .9517738   2.620138
```

```
-----+-----
sigma_u | .46340524
sigma_e | 3.7883092
rho | .01474282 (fraction of variance due to u_i)
-----
```

F test that all u\_i=0: F(5, 113) = 0.30 Prob > F = 0.9132

-----

**b.2) Estimaciones Banxico=2, Región Centro Norte**

```
Fixed-effects (within) regression      Number of obs   =   200
Group variable: entidad1                Number of groups =   10
```

```
R-sq: within = 0.2102      Obs per group: min =   20
      between = 0.3760      avg =   20.0
      overall = 0.2187     max =   20
```

```
corr(u_i, Xb) = 0.0426      F(1,189) = 50.30
                          Prob > F = 0.0000
```

```
-----+-----
ratepib |   Coef.  Std. Err.   t   P>|t|   [95% Conf. Interval]
-----+-----
rateimss | .4719381 .0665398   7.09 0.000   .340682 .6031942
_cons | 1.309065 .3439423   3.81 0.000   .6306063 1.987524
-----
```

```
-----+-----
sigma_u | .77708647
sigma_e | 3.4144165
rho | .04924631 (fraction of variance due to u_i)
-----
```

F test that all u\_i=0: F(9, 189) = 1.03 Prob > F = 0.4144

-----

**b.3) Estimaciones Banxico=3, Región Centro**

```
Fixed-effects (within) regression      Number of obs   =   140
Group variable: entidad1                Number of groups =    7
```

```
R-sq: within = 0.4827      Obs per group: min =   20
      between = 0.7954      avg =   20.0
      overall = 0.4979     max =   20
```

```
corr(u_i, Xb) = 0.0522      F(1,132) = 123.16
                          Prob > F = 0.0000
```

```
-----+-----
ratepib |   Coef.  Std. Err.   t   P>|t|   [95% Conf. Interval]
-----+-----
rateimss | .6267162 .0564727  11.10 0.000   .5150076 .7384247
_cons | 1.167627 .3087744   3.78 0.000   .5568404 1.778413
-----
```

```
-----+-----
sigma_u | .46044831
sigma_e | 2.9755469
rho | .02338575 (fraction of variance due to u_i)
-----
```

F test that all u\_i=0: F(6, 132) = 0.48 Prob > F = 0.8240



```

-----+-----
sigma_u |    0
sigma_e | 3.7883092
rho |    0 (fraction of variance due to u_i)
-----

```

-> banxico = 2

```

Random-effects GLS regression      Number of obs   =   200
Group variable: entidad1          Number of groups =   10

```

```

R-sq: within = 0.2102              Obs per group: min =   20
      between = 0.3760                avg =   20.0
      overall = 0.2187                max =   20

```

```

                                Wald chi2(1)   =  55.11
corr(u_i, X) = 0 (assumed)       Prob > chi2   =  0.0000

```

```

-----+-----
ratepib |   Coef.  Std. Err.   z  P>|z|  [95% Conf. Interval]
-----+-----
rateimss | .479405  .0645802   7.42  0.000   .3528302   .6059799
_cons | 1.281577  .3495474   3.67  0.000   .5964765   1.966677

```

```

-----+-----
sigma_u | .27613642
sigma_e | 3.4144165
rho | .00649805 (fraction of variance due to u_i)
-----

```

-> banxico = 3

```

Random-effects GLS regression      Number of obs   =   140
Group variable: entidad1          Number of groups =    7

```

```

R-sq: within = 0.4827              Obs per group: min =   20
      between = 0.7954                avg =   20.0
      overall = 0.4979                max =   20

```

```

                                Wald chi2(1)   = 136.86
corr(u_i, X) = 0 (assumed)       Prob > chi2   =  0.0000

```

```

-----+-----
ratepib |   Coef.  Std. Err.   z  P>|z|  [95% Conf. Interval]
-----+-----
rateimss | .6315444  .0539848  11.70  0.000   .5257362   .7373526
_cons | 1.152309  .3018928   3.82  0.000   .56061    1.744008

```

```

-----+-----
sigma_u |    0
sigma_e | 2.9755469
rho |    0 (fraction of variance due to u_i)
-----

```

-----  
-----  
-> banxico = 4

Random-effects GLS regression                    Number of obs    =    180  
Group variable: entidad1                        Number of groups =    9

R-sq: within = 0.1509                            Obs per group: min =    20  
          between = 0.0083                                avg =    20.0  
          overall = 0.1265                                max =    20

Wald chi2(1)    =    29.94  
corr(u\_i, X) = 0 (assumed)                        Prob > chi2    =    0.0000

-----  
-----  
ratepib |    Coef.   Std. Err.    z   P>|z|   [95% Conf. Interval]  
-----+-----  
rateimss | .2702396   .0493876    5.47   0.000   .1734417   .3670374  
\_cons | 1.104108   .5010986    2.20   0.028   .1219723   2.086243  
-----+-----  
sigma\_u | 1.1725915  
sigma\_e | 3.0347313  
rho | .12990338   (fraction of variance due to u\_i)  
-----  
-----

by banxico, sort : xtreg ratepib rateimss, fe

-----  
-----  
-> banxico = 1

Fixed-effects (within) regression                Number of obs    =    120  
Group variable: entidad1                        Number of groups =    6

R-sq: within = 0.3447                            Obs per group: min =    20  
          between = 0.0502                                avg =    20.0  
          overall = 0.3388                                max =    20

F(1,113)        =    59.45  
corr(u\_i, Xb) = -0.0685                            Prob > F        =    0.0000

-----  
-----  
ratepib |    Coef.   Std. Err.    t   P>|t|   [95% Conf. Interval]  
-----+-----  
rateimss | .557412   .0722947    7.71   0.000   .4141832   .7006409  
\_cons | 1.785956   .4210531    4.24   0.000   .9517738   2.620138  
-----+-----  
sigma\_u | .46340524  
sigma\_e | 3.7883092  
rho | .01474282   (fraction of variance due to u\_i)  
-----  
-----

F test that all u\_i=0:    F(5, 113) =    0.30                        Prob > F = 0.9132

-> banxico = 2

Fixed-effects (within) regression      Number of obs    =    200  
Group variable: entidad1                Number of groups =    10

R-sq: within = 0.2102                    Obs per group: min =    20  
      between = 0.3760                        avg =    20.0  
      overall = 0.2187                        max =    20

corr(u\_i, Xb) = 0.0426                    F(1,189)        =    50.30  
    Prob > F        =    0.0000

```
-----  
ratepib |   Coef.  Std. Err.   t  P>|t|  [95% Conf. Interval]  
-----+-----  
rateimss | .4719381 .0665398   7.09 0.000   .340682 .6031942  
_cons | 1.309065 .3439423   3.81 0.000   .6306063 1.987524  
-----+-----  
sigma_u | .77708647  
sigma_e | 3.4144165  
rho | .04924631 (fraction of variance due to u_i)  
-----
```

F test that all u\_i=0: F(9, 189) = 1.03      Prob > F = 0.4144

-> banxico = 3

Fixed-effects (within) regression      Number of obs    =    140  
Group variable: entidad1                Number of groups =    7

R-sq: within = 0.4827                    Obs per group: min =    20  
      between = 0.7954                        avg =    20.0  
      overall = 0.4979                        max =    20

corr(u\_i, Xb) = 0.0522                    F(1,132)        =    123.16  
    Prob > F        =    0.0000

```
-----  
ratepib |   Coef.  Std. Err.   t  P>|t|  [95% Conf. Interval]  
-----+-----  
rateimss | .6267162 .0564727  11.10 0.000   .5150076 .7384247  
_cons | 1.167627 .3087744   3.78 0.000   .5568404 1.778413  
-----+-----  
sigma_u | .46044831  
sigma_e | 2.9755469  
rho | .02338575 (fraction of variance due to u_i)  
-----
```

F test that all u\_i=0: F(6, 132) = 0.48      Prob > F = 0.8240

-> banxico = 4

Fixed-effects (within) regression      Number of obs    =    180  
Group variable: entidad1                Number of groups =    9



```

/sigma_e | 3.700303 .2388535          3.260562  4.19935
rho |      0 (omitted)
-----
Likelihood-ratio test of sigma_u=0: chibar2(01)= 0.00 Prob>=chibar2 = 1.000

```

```

-----
-> banxico = 2

```

```

Fitting constant-only model:
Iteration 0: log likelihood = -556.24845
Iteration 1: log likelihood = -553.36428
Iteration 2: log likelihood = -553.21326
Iteration 3: log likelihood = -553.21265

```

```

Fitting full model:
Iteration 0: log likelihood = -528.54375
Iteration 1: log likelihood = -528.53871

```

```

Random-effects ML regression      Number of obs   =    200
Group variable: entidad1          Number of groups =    10

```

```

Random effects u_i ~ Gaussian      Obs per group: min =    20
                                   avg =    20.0
                                   max =    20

```

```

                                   LR chi2(1)      =    49.35
Log likelihood = -528.53871         Prob > chi2   =    0.0000

```

```

-----
ratepib |   Coef.  Std. Err.   z  P>|z|  [95% Conf. Interval]
-----+-----
rateimss | .4803221 .0641994   7.48 0.000   .3544937 .6061506
_cons | 1.2782 .3371178   3.79 0.000   .6174616 1.938939
-----+-----
/sigma_u |      0 .9587004
/sigma_e | 3.399931 .1699694          3.082597  3.749931
rho |      0 (omitted)

```

```

-----
Likelihood-ratio test of sigma_u=0: chibar2(01)= 0.00 Prob>=chibar2 = 1.000

```

```

-----
-> banxico = 3

```

```

Fitting constant-only model:
Iteration 0: log likelihood = -417.13131
Iteration 1: log likelihood = -400.79314
Iteration 2: log likelihood = -397.21823
Iteration 3: log likelihood = -396.92848
Iteration 4: log likelihood = -396.92554

```

```

Fitting full model:
Iteration 0: log likelihood = -348.70308
Iteration 1: log likelihood = -348.69587
Iteration 2: log likelihood = -348.69587

```



rho | .1031322 .064305 .0246429 .2871613  
-----  
Likelihood-ratio test of sigma\_u=0: chibar2(01)= 8.78 Prob>=chibar2 = 0.002