

### 3. DATOS Y METODOLOGÍA

#### 3.1 Metodología.

El enfoque que se siguió para efectos de éste trabajo ha sido utilizado por Ellison y Glaeser (1999); Midelfart-Knarvik, Overman, Redding, y Venables (2000); Midelfart-Knarvik, Overman, y Venables(2001), y Sanguinetti Pablo y Volpe Martincus (2004).

Básicamente, el modelo trata de explicar a través de interacciones de características estatales e industriales, los patrones de localización. La idea se centra en que las industrias que utilizan intensivamente un factor se localicen en los estados que tienen cierta abundancia en dichos factores (Midelfart-Knarvik, Overman, Redding, y Venables, 2000). Dichas interacciones se presentan en la siguiente Tabla.

**Tabla I.**

<b>Variables Estatales, Industriales e Interacciones</b>		
Características Estatales	Abundante en Agricultura Abundante en Recursos Minerales Abundante en Mano de Obra Potencial de Mercado Base Industrial Distancia a las Principales Ciudades Capitales, Puertos Y Ciudades Fronterizas Promoción Industrial Kilómetros pavimentados de Carreteras y Autopistas.	
Características Industriales		Intensa en insumos agrícolas Intensa en recursos energéticos Intensa en mano de obra Economías de Escala Intensa en insumos intermedios Demanda Intermedia Intensa en insumos importados (Trade Cost) Uso de transporte intensivo

Interacciones	<ul style="list-style-type: none"> <li>-Abundante en Agricultura</li> <li>-Abundante en Recursos Minerales</li> <li>-Abundante en Mano de Obra</li> <li>-Potencial de Mercado</li> <li>-Base Industrial</li> <li>-Base Industrial</li> <li>-Distancia a las Ciudades Principales Capitales, Puertos Y Ciudades Fronterizas</li> <li>-Kilómetros pavimentados de Carreteras y Autopistas.</li> </ul>	<ul style="list-style-type: none"> <li>* Intensa en insumos agrícolas</li> <li>* Intensa en recursos energéticos</li> <li>* Intensa en mano de obra</li> <li>* Economías de Escala</li> <li>* Intensa en insumos intermedios</li> <li>* Demanda Intermedia</li> <li>* Costo de Comercio</li>   <li>*Uso de transporte intensivo</li> </ul>
---------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Siguiendo a Sanguinetti, las primeras tres interacciones representan el enfoque de ventajas comparativas. El segundo grupo de interacciones se combinan para tomar el enfoque de la *Nueva economía Geográfica*.

El modelo econométrico utilizado fue Mínimos Cuadrados Ordinarios. El uso de este modelo difiere al utilizado por Sanguinetti debido a la discrepancia que existe en la variable dependiente. Sanguinetti utiliza un modelo Tobit, ya que él se encuentra con gran cantidad de ceros en la variable dependiente.

Por otro lado, siguiendo a E&G, se normalizaron las variables de interacción. Esto con el fin de facilitar la interpretación de resultados. Un valor más grande de  $\beta$  indica que la variación en la proporción en que las industrias utilizan un recurso, o la proporción con que un estado cuenta con éste, tiene un efecto mayor en la concentración del empleo en determinada industria (Ellison & Glaeser, 1999).

El modelo general es el siguiente:

$$\ln S_{ki} = \beta_0 + \beta_1(Interacciones) + \beta_2(var\ nivel) + \beta_3(dummy\_años) + \beta_4(dummy\_PI) + \mu$$

donde  $\beta_0$ , es la ordenada al origen y  $\mu$  es el coeficiente de error<sup>1</sup>.

El modelo propuesto sigue la especificación del E&G, con la extensión de las variables de interacción que se tomaron de Venables et al (2001). En el modelo propuesto por E&G, sólo se toman en cuenta las variables de interacción. El modelo formal es como se muestra a continuación:

$$\ln S_{ki} = \beta_0 + \beta_1(\text{normInteracciones}) + \mu$$

Se corrieron dos tipos de regresiones. La primera fue con todos los datos sin hacer distinción de años, contando con una base de datos de 2906 observaciones. Para validar la significancia de los coeficientes obtenidos en esta ecuación, se corrió una segunda regresión haciendo distinción entre los años, realizando una prueba de betas que sostiene la siguiente hipótesis:

$$\beta_{194}(\text{Interaccion}) = \beta_{199}(\text{Interacción})$$

Se examinó la diferencia *estadística* de los coeficientes de regresiones para distintos años estimadas en conjunto. Esto muestra un tratamiento estadístico que no se encuentra en otros trabajos de este tipo y que se considera importante para efectos de esta tesis.

---

<sup>1</sup> Se corrió una regresión incluyendo la variable dummy de promoción industrial. Sin embargo, se encontró poco relevante para efectos del estudio, razón por la cual se omitió en el capítulo Resultados.

Los resultados de dicha prueba se presentan en la Tabla II del Anexo y se someten a discusión en la interpretación de resultados en el capítulo cinco.

### 3.2 Datos

La fuente principal de información para construir la base de datos fueron los Censos Económicos para los años de 1994 y 1999<sup>2</sup>. Dichos censos fueron realizados por el Instituto Nacional de Estadística, Geografía e Informática (INEGI). Otras fuentes fueron la Secretaría de Economía, los gobiernos de los estados, la Secretaría de Comunicaciones y Transporte, Guía Roji y la Encuesta Nacional de Empleo (1998 y 1995).

Los años analizados en el estudio fueron 1994 y 1999. La muestra consiste de 2,906 observaciones; 1,453 observaciones tanto para el año de 1994 como para 1999.

Se cuenta con datos descriptivos acerca de los diferentes estados e industrias. Dichos datos, siguiendo a Sanguinetti (2004), se usan como variables explicativas en el modelo econométrico. Los datos utilizados fueron de los 31 estados y el Distrito Federal a un nivel de desagregación de 4 dígitos (a nivel de rama, según la CMAP<sup>3</sup>) para los 9 sub sectores de la industria manufacturera. En la Tabla III del Anexo se muestra el listado de industrias manufactureras que se ocuparon para este estudio.

#### 3.2.1 Variable dependiente.

---

<sup>2</sup> Es importante señalar que la información que muestran estos censos es información recolectada en los años de 1993 y 1998.

<sup>3</sup> CMAP Clasificación Mexicana de Actividades y Productos

Como variable dependiente se tomó el logaritmo natural del la participación del empleo de la industria  $k$  en el estado  $i$  como proporción del empleo manufacturero total<sup>4</sup>. Dicha variable es utilizada por Sanguinetti en su artículo para Argentina. El valor del empleo de la industria manufacturera  $k$  en el estado  $i$  esta definido por  $x_{ik}$ , a su vez, . De esta manera nuestra proporción quedaría definida de la siguiente manera ( $S_{kit}$ ):

$$s_{ik}(t) \equiv \frac{x_{ik}(t)}{\sum_{i=1}^N x_{ik}(t)}$$

### 3.2.2 Concentración Industrial.

Para referirnos a la concentración industrial, se hace uso de los índices planteados por Hirschman-Herfindahl, quienes proponen el índice de concentración absoluta (AC), y el índice Florence-Ellison-Glaeser para referirnos a concentración relativa (RC). Entiéndase como concentración absoluta si unas cuantas regiones sin importar su tamaño cuentan con una gran proporción de la actividad de cierta industria<sup>5</sup>. Por otro lado, una industria está relativamente concentrada si su distribución espacial difiere significativamente del total del sector manufacturero<sup>6</sup>. A continuación, se describen los índices formalmente:

---

<sup>4</sup> Sanguinetti (2004).

<sup>5</sup> Las definiciones de los índices referidos en este apartado fueron tomados de Sanguinetti 2004.

<sup>6</sup> (Haaland, Kind, Midelfart-Knarvik, and Torstensson, 1999)

$$AC_k \equiv \sum_{i=1}^N s_{ik}^2$$

$$RC_k \equiv \sum_{i=1}^N (s_{ik} - s_i)^2$$

El índice AC es la suma del cuadrado de la proporción de cada estado en el empleo total manufacturero a nivel nacional. Cuando la concentración está distribuida equitativamente el índice toma valores de  $(1/n)^7$  y 1 cuando sólo una entidad mantiene la totalidad del empleo manufacturero.

El índice RC está construido con base en las diferencias al cuadrado de las proporciones  $S_{ki}$  y  $S_i$ <sup>8</sup>. Al considerar las diferencias al cuadrado se asegura que todas las entidades tengan una ponderación positiva. Este índice es cero cuando el patrón espacial del empleo en alguna industria en específico, concuerda exactamente con el total del sector manufacturero. El índice crece conforme a la divergencia.

### 3.2.3 Variables Independientes.

En este apartado se describen los dos grupos de variables que se utilizaron. El primer grupo está compuesto por variables que describen las características del estado; y el segundo lugar por variables que describen las características de las industrias.

---

<sup>7</sup> Entiéndase por  $n$  el número total de estados.

<sup>8</sup> Proporción del empleo manufacturero del estado  $i$  en empleo manufacturero nacional.

### 3.2.3.1 Variables descriptivas estatales:

- ***Distancia entre estados:*** Se utilizó la distancia entre los estados tomando como unidad de medida la distancia en Kilómetros que hay entre las ciudades capitales de cada una de estas<sup>9</sup>.
- ***Distancia a puertos:*** Se tomó en cuenta también la distancia de cada una de las ciudades capitales a los puertos más importantes del país<sup>10</sup>. Esto con el motivo de poder abarcar de la mejor manera los diferentes puntos de exportación e importación con los que podrían contar las industrias a la hora de determinar la localización.
- ***Ciudades Fronterizas:*** con el fin de controlar los diferentes puntos de acceso al mercado nacional, así como también al mercado de Estados Unidos de Norte América, se tomaron en cuenta las distancias de las ciudades capitales a las ciudades fronterizas más importantes.
- ***Infraestructura:*** para esta variable se tomó la longitud de las carreteras pavimentadas de dos, cuatro o más carriles por cada estado de la Republica Mexicana<sup>11</sup> como proporción del área total de cada estado.
- ***Abundancia en recursos energéticos:*** para medir la abundancia en recursos energéticos, especialmente combustibles, se utilizó la participación porcentual de las entidades federativas en las Actividades Económicas de Minería, por Gran División de Actividad Económica. Asignándole una dummy con valor de 1 a

---

<sup>9</sup> Fuente de distancias: Guía Roji y Matriz de distancia proporcionada por Gabriel Lara Ibarra.

<sup>10</sup> Para dicha clasificación se tomaron los puertos administrados por la Administración Portuaria Integral (API). Se tomaron los 16 puertos que son administrados por esta entidad.

<sup>11</sup> Datos tomados de la Secretaria de Comunicaciones y Transporte.

aquellos estados que tuvieran una mayor participación<sup>12</sup> (0 al resto de los estados; min94 y min99). También se tomó directamente la participación del estado en la Actividad Económica (mine94 y mine99)<sup>13</sup>.

- **Producto Interno Bruto por Entidad Federativa:** Se utilizaron los datos del Producto Interno Bruto a precios corrientes por Entidad Federativa y posteriormente fueron deflactados por medio del Índice de Precios Implícitos<sup>14</sup>.
- **Abundancia en Agricultura:** Para efectos de esta variable se tomó la participación porcentual de los estados en la actividad económica agropecuaria. Esta variable nos permite ver la abundancia en agricultura con el que el estado cuenta.
- **Abundancia en Mano de obra:** para medir esta variable, y siguiendo el modelo original de Sanguinetti, se tomo la relación entre el salario promedio de la manufactura en el estado y el promedio Nacional de los salarios manufactureros.
- **Potencial de Mercado:** El potencial de mercado su obtiene a través del índice propuesto por Keeble et al. (1986). Formalmente se define como

$$MP_i \equiv \sum_{j \neq i} \frac{Y_j}{d_{ij}} + \frac{Y_i}{d_{ii}}$$

donde  $i$  es el estado observado,  $j$  es el resto de los estados del país.  $Y_i$  es el PIB del estado  $i$ ,  $d_{ij}$  mide la distancia entre las ciudades más importantes, desde un punto de vista económico, de la provincia  $i$  a la  $j$ . La distancia intra estatal,  $d_{ii}$ , está definida como 1/3 del radio del círculo cuya área es equivalente al área del estado

---

<sup>12</sup> El criterio usado para determinar dicha variable fue tomar los estados cuya participación fuera mayor al 2.5% en la actividad.

<sup>13</sup> Ambas mediciones fueron utilizadas en los ejercicios econométricos. Sin embargo, fue la participación del estado en la actividad la que se muestra en los resultados finales.

<sup>14</sup> Estos datos se obtuvieron del Banco de Información Económica del INEGI.



- en cuestión. Otra opción para medir el acceso o potencial de mercado es el propuesto por Reddings y Venables<sup>15</sup>.
- **Base Industrial:** Para medir dicha variable se utilizó la relación del PIB Estatal Industrial como proporción del PIB Industrial a nivel Nacional.
  - **Población entre 15 y 65 años de edad:** siguiendo el modelo propuesto por Sanguinetti (2004), se utilizó la población entre 15 y 65 años. Para el año de 1999 se utilizó el Censo de Población y Vivienda que lleva a cabo el INEGI. Se tomó población por grupo quinquenal para esta instancia. Como dicho censo sólo se lleva a cabo cada 10 años se tomó el año de 2000 para los datos necesarios para 1999. Para el año de 1994, se hizo lo señalado anteriormente para los datos tomando el censo de 1990 para así poder obtener la información necesaria para 1994.
  - **Proporción del estado en el empleo nacional:** para la proporción del estado en el empleo nacional se tomaron las cifras del Total de la población de 12 años y más económicamente activos de los Censos de Población y Vivienda para los años de 1990 y 2000, ya que eran los años de los Censos más próximos a los años de estudio.

### 3.2.3.2 Variables descriptivas industriales:

- **Intensidad en agricultura:** los datos para esta variable se obtuvieron de la matriz insumo producto para México. Una limitante en la medición de esta variable

---

<sup>15</sup> Stephen Redding, Anthony J. Venables. Economic geography and international inequality. Journal of International Economics 62 (2004) 53– 82. Dicha medición de Potencial de Mercado no es factible para este estudio debido a la falta de datos de comercio entre las entidades federativas.

- consiste en que la matriz insumo producto utilizada es de 1980. Se tuvo que recurrir a esta matriz ya que era la única fuente disponible de ese tipo de datos.
- ***Unidades económicas:*** Es la unidad estadística de observación sobre la cual se recopilan datos de las industrias; se fundamenta en la necesidad de recolectar información con el mayor nivel de precisión<sup>16</sup>.
  - ***Producción Bruta Total:*** valor de los bienes y servicios producidos o comercializados por la unidad económica como resultado del ejercicio de sus actividades durante el año de referencia. Estos datos fueron tomados de los Censos Económicos de los años en estudio.
  - ***Sueldos y Salarios:*** Se define como los pagos realizados por la unidad económica, para retribuir el trabajo ordinario y extraordinario del personal consignado en nómina, antes de cualquier deducción (Impuesto Sobre la Renta, las aportaciones al IMSS, ISSSTE, INFONAVIT, cuotas sindicales y todas aquellas cuotas a cargo de los trabajadores que fueron retenidas por los empleadores). Incluye aguinaldos, comisiones sobre ventas al personal ocupado remunerado, licencias o permisos temporales, bonificaciones, incentivos y bonos de productividad<sup>17</sup>.
  - ***Economías de escala.*** Apegándonos al modelo original, se tomo la variable *Tamaño del establecimiento promedio* como variable Proxy de las economías de escala. Esta variable nos indica el tamaño promedio de la empresa. Para esto se

---

<sup>16</sup> Glosario Censo Económico de 1999. INEGI. [www.inegi.gob.mx](http://www.inegi.gob.mx)

<sup>17</sup> Glosario Censo Económico de 1999. INEGI. [www.inegi.gob.mx](http://www.inegi.gob.mx)

- tomó al personal ocupado promedio total por industria y se dividió entre el número de unidades económicas por industria<sup>18</sup>.
- ***Intensidad en consumo intermedio***: Describe la característica de aquellas industrias que usan intensivamente como insumos, productos finales de otras industrias maquiladoras. Los datos para esta variable se obtuvieron de la matriz insumo producto para México, con las respectivas limitantes anteriormente mencionadas.
  - ***“Intermediate demand bias”***: Define en qué proporción los productos finales de una industria son utilizados como insumos intermedios en otras industrias. Los datos fueron obtenidos de la matriz insumo producto anteriormente descrita.
  - ***Intensiva en Transporte***: Define a aquellas industrias cuyo uso en transporte es intensivo. Para este motivo se toma la relación Costo Marginal de transporte/Ventas totales. Debido a la dificultad que representa obtener los costos marginales de transporte para cada industria, se tomó la distancia en Km. elevada al cuadrado<sup>19</sup>. Alternativamente, se usa el tabulado de costos de transporte realizado por la CANACAR<sup>20</sup> para 2000. Se utilizan los costos de transportar 42 toneladas en una unidad T3-S2-R4<sup>21</sup>
  - ***Costos de comercio***: Una vez más se tomó la variable directamente del modelo original. Dicha variable usa las tarifas promedio a un nivel de 4 dígitos el ISIC<sup>22</sup>.

---

<sup>18</sup> Los Censos Económicos no cuentan con información a nivel firma, es por eso que se trató de estimar el tamaño del establecimiento por la forma descrita anteriormente.

<sup>19</sup> En un estudio realizado por Combes y Lafourcade, exponen que el uso de la distancia al cuadrado satisface criterios necesarios para utilizarse como proxy.

<sup>20</sup> Cámara Nacional del Auto Transporte de Carga.

<sup>21</sup> T3-S2-R4:

<sup>22</sup> International Standard Industrial Classification. Los datos se obtuvieron de la base de datos de Producción y Comercio del Banco Mundial. Disponible en [www.worldbank.org/trade](http://www.worldbank.org/trade)

