

4. DATOS Y METODOLOGÍA

Una vez planteada la motivación teórica y una descripción de la experiencia mexicana, pasamos ahora al análisis econométrico del trabajo. El objetivo general del estudio es detallar el vínculo que existe en México entre el crecimiento económico, el sector exportador y su grado de diversificación. La hipótesis que buscamos probar es que la expansión de las exportaciones, y su diversificación, tienen un efecto positivo sobre el crecimiento económico del país, tanto en términos de producto como en productividad.

Antes de presentar los modelos estimados, es importante dedicar algunas líneas para detallar la fuente de los datos utilizados. La sección 4.1 se encarga precisamente de esto. La sección 4.2, por otra parte, está dedicada a exponer de manera intuitiva las bases teóricas del método econométrico elegido para probar nuestra hipótesis.

4.1 Descripción de datos

El modelo desarrollado para la sección empírica del estudio está basado en una función de producción Cobb-Douglass aumentada. Como hemos visto, este tipo de planteamiento es utilizado comúnmente en trabajos que buscan probar la validez de la hipótesis de crecimiento por exportaciones (Awokuse, 2003; Abdulai y Jacques, 2002; Keong, Yusop, Khim-Sen Liew, 2003). Así pues, podemos plantear la función general como

$$(1) \quad Y = F(K, L; X, \text{Gini}, \Omega)$$

Donde Y mide la producción, K es capital, L es trabajo, X es el nivel de exportaciones manufactureras con o sin maquila, Gini es un índice de diversificación de exportaciones, y Ω es el resto de las variables de control utilizadas. Este tipo de

planteamiento resulta coherente con la literatura teórica previamente discutida, dado que permite probar si las exportaciones y el grado de diversificación de las mismas son una fuente de crecimiento económico. Al controlar por capital y trabajo, el planteamiento del modelo supone que el sector externo afecta el crecimiento de la economía por medio de cambios en productividad.

Para poder elaborar índices adecuados sobre el grado de diversificación de las exportaciones en el país resulta de capital importancia trabajar con una base con el mayor grado de desagregación posible. Para el presente trabajo se utilizó la base de datos publicada por el INEGI en su Banco de Información Económica, titulada “Exportaciones según principales productos del Sistema Armonizado de Designación y Codificación de Mercancías”. En ella, se reporta el valor de las exportaciones de México (expresado en miles de dólares) de hasta 560 tipos de productos distintos, con una periodicidad mensual entre 1993 y 2006. La serie del INEGI tiene la ventaja, además, de clasificar el valor de las exportaciones de maquila y no maquila. Hacer esta distinción resulta particularmente interesante para este tipo de estudios, en primer lugar porque las maquiladoras están sujetas a una estructura tributaria distinta al resto del país, y también porque estas empresas no generan eslabonamientos con el resto de las industrias nacionales: importan insumos extranjeros, elaboran el producto y lo vuelven a exportar. Por ello, podemos hablar de las maquiladoras como un sector bastante aislado de la economía mexicana, cuyo comportamiento y efecto será naturalmente distinto al resto.

Por último, utilizando información provista por el Banco de México, fue también posible dividir esta misma base de datos en productos manufacturados y no manufacturados, algo importante para poder evaluar resultados de la liberalización

comercial bajo la luz de la literatura sobre transformación estructural productiva previamente discutida.

Las variables utilizadas para correr el modelo econométrico están expresadas en términos trimestrales entre 1993 y 2006. Con excepción de los índices de concentración, tipo de cambio real y términos de intercambio, todos los datos están reportados en logaritmos naturales. Para medir el producto se utilizó el PIB real de México, a precios de 1993, obtenido del Sistema de Cuentas Nacionales publicado por el INEGI. Debido a que nuestro interés es probar si existe una causalidad positiva de la expansión del sector externo sobre la productividad general del país, el producto será expresado neto de exportaciones manufactureras, o en algunos casos neto de exportaciones manufactureras sin maquila. De esta manera eludimos el problema que supone la incorporación de este sector en la contabilidad del producto nacional¹.

La variable de trabajo se obtuvo a partir de estadísticas trimestrales sobre el total de asegurados en el IMSS. Naturalmente, existe una porción considerable de trabajadores en el país que no se encuentran inscritos en el IMSS. Además, la serie de asegurados del IMSS incluye algunos ramos que no pertenecen estrictamente a trabajadores, como el seguro facultativo y el de continuación voluntaria. No obstante, es una de las pocas series mensuales sobre fuerza laboral en el país que reporta información sobre todo el periodo cubierto por este estudio. Los datos sobre personas económicamente activas u ocupadas a nivel nacional para el periodo abarcado, por ejemplo, se reportan sólo en términos anuales. La Encuesta Nacional de Empleo Urbano (ENEU) presenta datos trimestrales sobre estas variables, pero la composición de las ciudades consideradas como urbanas ha cambiado durante estos años, por lo que

¹ Herzer, Nowak-Lehmann y Siliverstovs (2006), entre otros, recomiendan este método.

utilizar esta serie como proxy de la fuerza laboral mexicana presentaría también problemas.

Para la variable de capital se utilizó la serie de formación bruta de capital fijo publicada en el Sistema de Cuentas Nacionales del INEGI. Si bien hubiera sido más adecuado utilizar el acervo de capital en lugar de su flujo, la estimación de esta variable presentaba varios problemas. En primer lugar, hasta donde tengo conocimiento, no existe una estimación trimestral del acervo de capital para México que cubra este periodo. Se intentó construir esta serie utilizando una fórmula simple del método del inventario perpetuo, pero utilizar una tasa de depreciación fija para toda la serie de inversión resulta poco apropiado. Para realizar una mejor estimación hubiera sido necesario obtener una serie de inversión más desagregada, para poder otorgar tasas diferenciadas de depreciación a los distintos tipos de capital. De cualquier forma, el uso del flujo de inversión como proxy de capital es una práctica bastante común en la literatura empírica referente a este tema.

El nivel de exportaciones manufactureras (con y sin maquila) se calculó a partir de datos del Sistema de Cuentas Nacionales. La serie se obtuvo aplicando las cifras de la proporción de manufacturas con y sin maquila sobre el nivel total de exportaciones.

La variable de diversificación se obtuvo aplicando la fórmula estándar del coeficiente de Gini al valor de los productos exportados en el país para cada trimestre considerado dentro de la muestra. La interpretación del coeficiente es la usual: un valor de 1 implicaría un grado máximo de concentración (México sólo exporta un tipo de producto), mientras que el 0 significaría una diversificación perfecta (el valor de las exportaciones de todos los productos clasificados de México es el mismo).

Existen otras variables que pueden ser relevantes a la hora de plantear este tipo de ecuación, cuya omisión podría generar sesgos en los coeficientes estimados. El tipo

de cambio real es un candidato claro: una depreciación real, por ejemplo, tiene un efecto tanto sobre las exportaciones como sobre el producto por canales distintos (por ejemplo, sobre el salario real y la demanda doméstica). Si controlamos por esta variable, el coeficiente asociado a la relación producto-exportaciones será necesariamente más preciso. Li (1999) y Chee Keong, Yusop y Liew (2003), entre otros, utilizan esta variable en sus modelos. La serie utilizada fue el índice de tipo de cambio real con precios de consumidor y con respecto a 111 países, publicado en la página oficial del Banco de México. En esta serie, un aumento del índice significa una depreciación real del peso. De esta misma fuente, además, se obtuvo el índice de términos de intercambio que se utilizó en varios de los modelos que se presentarán más adelante. Esta es una variable de control utilizada en Li (1999) y en Awokuse (2003).

Por último, se utilizaron variables que controlan por el PIB de Estados Unidos y la importación de bienes de capital. La primera se consideró como una posible proxy de la actividad del sector exportador mexicano, debido a la profunda integración de ambas economías durante el periodo que comprende el estudio. La segunda variable fue utilizada para controlar por el efecto tecnológico positivo que provoca la entrada de este tipo de importaciones al país. Los datos de Estados Unidos fueron obtenidos del US Department of Commerce, expresados en billones de dólares reales del 2000 (ajustados por estacionalidad), y convertidos a logaritmos. La variable de importaciones se obtuvo del Banco de Información Económica del INEGI, en el ramo de importación de bienes para formación bruta de capital fijo. Está expresada en millones de pesos reales de 1993, y también fue transformada a logaritmos.

4.2 Metodología econométrica

Para probar la relación entre crecimiento, exportaciones y su grado de diversificación, se planteará un modelo de vector de corrección de errores (VEC). Este tipo de planteamiento presenta varias ventajas y, como hemos visto, ha sido utilizado ampliamente en la literatura que investiga los vínculos de crecimiento del producto y de las exportaciones. En primer lugar, este tipo de modelo permite estudiar apropiadamente la relación que existe entre múltiples series no estacionarias, un caso muy común cuando se estudian las variables macroeconómicas empleadas para este trabajo. En segundo lugar, resulta atractivo porque no es necesario escribir una forma estructural predeterminada para las variables en cuestión. Esto, además, facilita la estimación de causalidades tipo Granger en el modelo. Por último, los modelos VEC permiten establecer una diferencia entre efectos de corto y largo plazo, lo cual es muy atractivo para este tipo de análisis.

Para poder estimar el modelo VEC es necesario primero probar la estacionariedad de cada una de las series, y verificar si existe un vector de cointegración que las vincule. Para ello, el primer paso consiste en desarrollar la prueba Dickey-Fuller aumentada (ADF) para las variables consideradas en el estudio. Esta prueba puede tomar cuatro formas:

$$(2) \Delta y_t = \delta y_{t-1} + e_t$$

$$(3) \Delta y_t = \beta_0 + \delta y_{t-1} + e_t$$

$$(4) \Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + e_t$$

$$(5) \Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + \alpha_i \sum_{i=1}^m \Delta y_{t-i} + e_t$$

La hipótesis nula que se busca probar en cada una de estas ecuaciones es que $\delta=0$. En caso de aceptarse, concluimos que la serie es no estacionaria y tiene un orden de integración positivo. Si se rechaza, entonces concluimos que la serie es $I(0)$. Siguiendo

la ecuación (2), por ejemplo, cuando aceptamos la hipótesis nula estamos concluyendo que la serie es una caminata aleatoria sin deriva; si la rechazamos, entonces es un modelo AR(1) estable. De manera alternativa, el grado de estacionariedad de las series también puede ser verificado con la prueba de Phillips-Perron, cuyos fundamentos intuitivos son básicamente los mismos, pero que utiliza un estadístico t distinto para probar la significancia de δ , el cual se encarga de eliminar el efecto de la correlación serial sobre la distribución asintótica del estadístico (Eviews, 2004).

Una vez que se conoce el grado de integración de cada una de las series, es necesario correr una prueba de cointegración de las variables consideradas. Este es un paso fundamental para poder tener confianza en las inferencias sugeridas por el modelo: como está bien documentado, trabajar con series de tiempo no estacionarias puede arrojar regresiones espurias a menos que las variables estén cointegradas. La cointegración implica que existe una combinación lineal de las variables no estacionarias que es, en efecto, estacionaria. Esta combinación resultante se interpreta como una relación de largo plazo entre las variables, y su naturaleza está descrita en los coeficientes del vector de cointegración.

Existen dos métodos distintos por los cuales se puede investigar la existencia de esta combinación lineal estacionaria de variables: la metodología de Engle-Granger y la de Johansen. Bajo el contexto de Engle-Granger, es fácil entender porqué el vector se entiende como una relación de equilibrio o de largo plazo. Si queremos, por ejemplo, comprobar que las variables integradas de orden 1 y_t y x_t están cointegradas, la metodología sugiere correr la regresión simple

$$(6) y_t = \beta_0 + \beta_1 x_t + e_t$$

y analizar el grado de estacionariedad de la serie de e_t . Si los residuos son estacionarios, entonces concluimos que las variables están cointegradas y que los parámetros de

cointegración son β_0 y β_1 . Si, en cambio, los residuos resultan estar integrados de orden 1, concluimos que las variables no están cointegradas y los coeficientes β_0 y β_1 son poco confiables para estimar la relación entre estas variables, debido a que podríamos estar incurriendo en problemas de regresión espuria.

El grado de estacionariedad de los residuos, como podemos apreciar en la ecuación (6), nos ayuda a descifrar si existe una relación estable entre las dos variables. Si los residuos son estacionarios, entonces e_t tendrá una media de cero y una varianza constante. La relación entre y_t y x_t será de equilibrio en el sentido de que cualquier desviación que experimenten (definida por el término e_t) será necesariamente temporal. Si, en cambio, la serie e_t es no estacionaria, entonces la relación entre las variables será inestable en el tiempo y no se puede hablar de un equilibrio de largo plazo entre ellas. Es importante notar, por último, que para poder hallar cointegración entre variables es necesario que estas sean integradas del mismo orden.

La motivación de plantear un método alternativo al Engle-Granger es que este posee varias debilidades. Entre las citadas por Enders (1995), destaca el hecho de que la prueba de cointegración no es siempre robusta a la variable que se coloca del lado izquierdo de la ecuación: es decir, modificar la variable independiente puede alterar la conclusión encontrada sobre la cointegración de las series. Por esta razón, es importante revisar otras metodologías que no compartan estos mismos problemas. La prueba propuesta por Johansen es la más popular de estas.

El método de Johansen para probar cointegración parte de un concepto teórico algo distinto. La intuición de esta prueba es similar a una generalización multivariada de las pruebas ADF. Partiendo de un VAR de orden p , tal que

$$(7) \quad y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t$$

donde y_t es un vector k de variables $I(1)$, y ε_t es un vector de innovaciones, encontramos que es posible reescribir la ecuación de manera que tengamos

$$(8) \Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + Bx_t + \varepsilon_t$$

Donde $\Pi = \sum_{i=1}^p A_i - I$, y $\Gamma_i = -\sum_{j=i+1}^p A_j$. Como podemos apreciar, la derivación del

vector Π es similar al del término δ de la prueba ADF. El rango de la matriz Π indica el número de vectores de cointegración independientes. Si es igual a cero, siguiendo la misma lógica que en la prueba ADF, entonces no encontramos una combinación lineal estacionaria de las variables y las series no están cointegradas. Si el rango es 1, en cambio, encontramos que existe una ecuación de cointegración entre las variables. En realidad, si existen n variables en el modelo, pueden existir hasta $n-1$ vectores de cointegración. El método de Johansen está incorporado a Eviews, por lo que no es necesario correr pruebas adicionales para encontrar el rango de la matriz Π . Los dos estadísticos comunes en esta metodología son reportados en el programa: el eigenvalue y el estadístico trace.

Una vez hallado el o los vector de cointegración, es posible finalmente estimar el modelo de corrección de errores. La ecuación de cointegración, como hemos discutido, funciona para entender la relación a largo plazo de las variables: la segunda parte del modelo, en cambio, se encarga de describir las variaciones en corto plazo al calcular la sensibilidad de las variables ante desajustes temporales. Siguiendo a la ecuación (6), y asumiendo que hayamos encontrado que las variables están cointegradas, el modelo de corto plazo se escribe de la siguiente forma:

$$(9) \Delta y_t = \alpha_1 + \gamma_1 \Delta x_t + \delta_1 (y_{t-1} - \beta x_{t-1}) + \phi_1 \Delta y_{t-1} + u_t$$

$$(10) \Delta x_t = \alpha_2 + \gamma_2 \Delta y_t + \delta_2 (y_{t-1} - \beta x_{t-1}) + \phi_2 \Delta x_{t-1} + u_t$$

El coeficiente δ_i es conocido como el término de corrección del error, ya que calcula la velocidad con la que la variable dependiente responde a desequilibrios en el periodo anterior. Este modelo sólo es apropiado si se cumplen el resto de condiciones que hemos comentado: debido a que las variables son I(1), los términos en diferencias que entran en el modelo son estacionarios. Asimismo, los términos en paréntesis son iguales al residuo en la ecuación de cointegración y, como hemos visto, deben ser I(0) si existe una relación de equilibrio entre las variables. Como todos los términos en el modelo son I(0), es posible aplicar a los coeficientes estadísticos t y f usuales con confianza, y la ecuación puede ser estimada mediante MCO.

El modelo expuesto en las líneas anteriores es el más sencillo posible, pero extenderlo no genera problemas, dado que la intuición es la misma. Es importante notar, sin embargo, que no todas las variables del modelo necesitan tener un coeficiente δ significativo. Si sólo uno de estos términos es significativo, entonces las demás variables no responden al desequilibrio en largo plazo y todo el ajuste se lleva a cabo por la variable con el término significativo. Como podemos apreciar, además, para hallar que la variable se ajusta a los shocks temporales debemos de encontrar que δ sea negativo y menor a uno en valor absoluto. Esto es fácil de ver: si el término en paréntesis de (6) es positivo, para que y_t regrese al equilibrio es necesario que δ sea menor que cero. Si fuera positivo, el incremento con respecto al equilibrio de y_t sería seguido por un mayor incremento en el siguiente periodo, lo cual haría que la variable jamás regresara al equilibrio original. Siguiendo una lógica similar, podemos ver que si δ supera a la unidad en valor absoluto, tampoco será posible que la variable se ajuste al equilibrio. Si $\delta=1$, el ajuste será inmediato: si $\delta=0$, la variable no se ajusta. Así, entendemos porqué conforme δ tiende a uno decimos que la velocidad de ajuste es alta, y conforme δ tiende a cero decimos que el ajuste es lento.

Una vez expuesta la intuición de los métodos de Engle-Granger y Johansen para probar cointegración, y ya que ha quedado expresa su relación con la estimación de los modelos de corrección de errores, pasaremos a exponer los resultados de nuestras estimaciones para la economía mexicana.