

ANÁLISIS EMPÍRICO DE LOS PRINCIPALES DETERMINANTES DEL GASTO EN INNOVACIÓN Y DESARROLLO PARA PAÍSES DESARROLLADOS

Salvador Rivas Aceves*

63

Resumen

Existen diversos factores, además de la estructura de mercado, la posición de la industria en el mismo y el valor de mercado que determinan el gasto en I+D. Es importante reconocer que el conocimiento, la productividad, el gasto social, las patentes, entre otras variables aportan un grado de explicación dentro de la actividad innovadora de los países desarrollados. En este artículo se presenta evidencia de dicha existencia a través de la estimación de un modelo de Datos Panel, encontrando que estas variables impactan en la trayectoria de la actividad en I+D.

Palabra clave: innovación tecnológica, investigación y desarrollo, conocimiento, datos de panel.

Clasificación JEL: C23, E22, O32.

Introducción

En la actualidad las empresas buscan incrementar su productividad mediante la inversión en Investigación y Desarrollo (I+D), se basan en modelos que explican el comportamiento de las mismas, incorporando la innovación y el crecimiento, así como fundamentos para el asesoramiento de las decisiones en la política científica y tecnológica. Posteriormente se toman decisiones sobre las inversiones a realizar y/o evaluar los resultados de programas concretos.

El nuevo interés en la inversión en I+D tiene como origen la necesidad de la contabilización de los activos intangibles de la empresa, tales como la reputación, patentes, marca, imagen corporativa, etc. aunado, dentro del marco de las decisiones de política científica y tecnológica, la evaluación del incremento en la actividad innovadora, decidir cuánto invertir y qué instrumentos utilizar, evaluar los resultados de las políticas implementadas e intentar establecer un nivel óptimo de subsidios.

* Lic. en Economía, UAM Azcapotzalco. Estudiante de la Maestría en Ciencias Económicas, UAM.

Para que una empresa incremente su actividad innovadora necesita del conocimiento, que se materializa en un incremento en los activos intangibles de la empresa, ocasionando un mejoramiento en los rubros de consultoría, seguridad, servicios financieros, incorporación al mercado de patentes, etc. de la empresa. Existen modelos que incorporan como una de las variables fundamentales del sistema a la tasa de crecimiento del conocimiento como función directa de los recursos dedicados a las actividades de I+D.

Otra de las razones por las cuales el conocimiento es considerado dentro de los modelos, es por que éste se mejora con base al conocimiento ya existente y con materiales nuevos, por lo que la tasa de crecimiento del conocimiento se puede incrementar destinando más recursos a su producción. La producción de nuevo conocimiento genera una serie de consecuencias de distinta naturaleza tales como: exclusión del uso mediante patentes en donde las empresas innovadoras pueden excluir del uso público cierto conocimiento científico y técnico al registrar sus innovaciones ante una autoridad legal; apropiabilidad que induce a las empresas a apropiarse del nuevo conocimiento mediante el aprendizaje empírico; y por último indivisibilidad ya que el nuevo conocimiento no puede ser deshecho una vez que ya se ha generado.

En consecuencia los modelos teóricos buscan resolver problemas tales como si existe o no una relación entre la estructura de mercado y la actividad innovadora de las empresas, si existe o no heterogeneidad respecto de las decisiones de innovación de las empresas y por ende como se explica, y por último cuales son los determinantes de la inversión en I+D.

Conocer con precisión cuales son estos determinantes es muy importante para los países porque el gasto en I+D genera beneficios que a su vez se trasladan del nivel empresarial al nivel social, convirtiéndose en la actualidad el gasto en I+D en parte fundamental del análisis para lograr establecer el presupuesto de egresos de los países.

En la actualidad existen modelos de crecimiento económico que se encuentran explicados fundamentalmente por la actividad innovadora e implícitamente por la inversión en nuevo conocimiento, capacitación de personal para elevar la productividad, generación de patentes, etc., razón por la cual el gasto en I+D se puede convertir en un criterio importante de política económica.

El presente ensayo aborda toda esta problemática analizando los principales determinantes del gasto en I+D a través de variables medidas de diversas formas, encontrando que es posible medir tanto variables cualitativas como cuantitativas, bajo esquemas y supuestos específicos, para lograr explicar la actividad de las empresas en I+D, y por ende la acción de los gobiernos en el mismo rubro. Este análisis se lleva a cabo por medio de una estimación en datos panel.

Innovación Tecnológica y Gasto en I+D

A través de diversos enfoques se ha logrado determinar cuál es el comportamiento de la empresa frente a la actividad de I+D e identificar la relación entre la innovación y la estructura de mercado en donde se dan las siguientes conclusiones:

- Schumpeter (1942): las empresas grandes y monopolísticas innovan más.
- Arrow (1962): los entrantes tienen más incentivos a innovar si la Innovación es radical.
- Gilbert y Newberry (1982): las empresas establecidas pueden actuar estratégicamente para dificultar la entrada.

Estos resultados fueron encontrados a través de análisis empíricos en I+D mediante regresiones de gasto en I+D sobre medidas de rentabilidad o concentración a través de industrias, utilizando variables exógenas que sirven como instrumentos para determinar preferencias como determinantes del lado de la demanda, tecnología como determinante del lado de la oferta, personal científico contratado, gasto directo en I+D, gasto de capacitación de personal, patentes ponderadas por citas, nuevos productos, crecimiento de la productividad, valor de mercado o de rentabilidad de las empresas, ventas de los nuevos productos y condiciones de apropiabilidad. Además, se tomó en cuenta que existen diferencias institucionales, factores de velocidad de las variables y otros factores que pueden estar determinados fuera del sistema.

Es importante señalar que estos análisis se realizaron a nivel empresarial, pero existen varios análisis que se han hecho a nivel agregado determinando cuales son los factores que afectan a la actividad innovadora, por ejemplo de un país en su conjunto. Uno de los principales modelos que permite visualizar determinantes a nivel agregado es el Modelo de Productividad con I+D de Griliches¹, en donde se supone una función de producción Cobb-Douglas (aproximación de primer orden en logaritmo a la función de producción) utilizando como unidad de observación la línea de producto, empresa, industria o país de la forma:

$$Y = AL^\alpha C^\beta K^\gamma e^u \quad (1)$$

En dicho modelo se concluye que un aumento en el capital de conocimiento se traduce en un aumento en el nivel de producto, que la medida del crecimiento de la productividad depende de la estructura de mercado para los bienes de consumo y que es sensible a la deflación. Estos resultados muestran la razón por la cuál el conocimiento se convierte en una variable fundamental en el análisis de la inversión en I+D.

¹ Griliches, Z. (1979). Issues in assesing the contribution of research and development to productivity growth. *Bell Journal of Economics*, 10(1), 92-116.

Pero además nos aporta otras conclusiones, debido a que la actividad innovadora de un país no nada más puede estar en función del conocimiento o de la inflación o de la estructura de mercado, en realidad existen otros factores igualmente importantes que afectan el gasto en I+D de un país de entre los cuales destacan las patentes, la productividad o la tecnología, etc. Y es precisamente ésta la motivación del presente artículo, es decir, mostrar que hay otros determinantes del gasto en I+D.

Es importante señalar que los análisis de regresión deben ser construidos de tal forma que se eviten problemas de endogeneidad como la causalidad, es decir, se requiere que los regresores sean independientes o no correlacionados con el error estocástico de tal forma que:

$$Y = \beta X + e \quad (2)$$

66

Además se requiere que $E(e/X) = 0$ para $E(\beta^{MCO}) = \beta$. Si se llega a tener este problema puede ser debido, entre otras razones, por:

- Variables omitidas: una variable afecta a X e Y, si la omitimos confundimos el efecto de X y el de la variable omitida.
- Simultaneidad: las variables regresoras miden lo mismo. En este caso no se puede saber quién innova más si las empresas grandes o las chicas, o si las industrias concentradas innovan más o menos y como las empresas que innovan se hacen más grandes, entonces un shock positivo sobre la innovación (“e” más grande) incrementará X. A nivel agregado no se sabe cuales son los principales determinantes del gasto en I+D
- Auto-selección: se traduce en un experimento no aleatorio, y se da porque si las empresas que son complementarias en I+D incrementan su I+D después de la fusión, tienen más incentivos a fusionarse, es decir, se auto seleccionan. Por lo cual ya no es posible medir si las fusiones de empresas incrementan o reducen la inversión en I+D. La solución consiste en escoger empresas con la misma propensión a ser adquiridas, pero con algunas que no son adquiridas.

La forma de encontrar relaciones causales es mediante experimentos naturales en donde se tenga que el instrumento solo mueve a X pero no a “e”. Lo anterior puede encontrar otro tipo de problemas debido a las diferencias institucionales entre países, factores que se mueven más despacio o antes que la variable dependiente, factores que se determinan fuera del sistema que se está analizando. La solución sería ignorar el problema de endogeneidad e interpretar la regresión como la esperanza condicionada de Y dado X, más que una relación causal entre Y y X.

² Gujarati, D. D., Ecometría McGraw-Hill. Cap 16.

Para realizar la estimación se obtuvieron datos que miden costos compartidos en I+D, la contribución de los activos al crecimiento de los servicios de capital, el crecimiento del producto interno bruto, el total de horas trabajadas, la productividad medida a través del cociente del producto interno bruto entre el total de horas trabajadas, el crecimiento de la productividad multifactor, el crecimiento total de los factores, crecimiento de las patentes, patentes industriales, número de patentes registradas en la United States Patent and Trademark Office (USPTO), patentes en información, computadoras y telecomunicaciones; patentes del tratado de cooperación, patentes registradas ante la European Patent Office (EPO), patentes en biotecnología en la EPO, el gasto en I+D, la inversión en conocimiento, el gasto social, inversión en información, computadoras y telecomunicaciones; asistencia oficial en desarrollo y exportaciones de alta tecnología.

Los datos de estas variables, que se proponen como principales determinantes del gasto en I+D, se obtuvieron de las siguientes fuentes:

- Organización para la Cooperación Económica y el Desarrollo (OECD)
- Comisión Europea y agencias estadísticas estatales
- Compustat, Datastream y Worldscope balances
- ISI, Derwent, Lexis-Nexis
- Tres grandes jurisdicciones de patentes:
 - ✓ USPTO: EEUU
 - ✓ EPO: Europa
 - ✓ JPO: Japón

Sin embargo no todas las variables se utilizaron debido a que varias de ellas miden exactamente lo mismo, encontrándonos con problemas de multicolinealidad, por lo tanto para realizar la estimación tomamos en cuenta aquellas que se consideran las más importantes y probablemente las más significativas, siendo éstas las siguientes:

- El Gasto Social (GS)
- Exportaciones de Alta Tecnología (HTEXP)
- Inversión en Conocimiento (ICON)
- Inversión en Información Computadoras y Telecomunicaciones (ICT)
- Número de Patentes en ICT registradas ante la EPO (ICTEPO)
- La Productividad Multifactor (MFP)
- La productividad (PROD)
- La Tendencia de las Patentes (TPF)
- Gasto en I+D (RYD)

Todas las variables están consolidadas y medidas en tasas de crecimiento anuales para el periodo que va desde 1990 hasta 2000, salvo el número de patentes en información, computadoras y telecomunicaciones registradas ante la EPO que está en valores absolutos. Estas variables se obtuvieron para los principales países innovadores como Canadá, Francia, Alemania, Japón, Suecia, Reino Unido y Estados Unidos. Dentro del análisis no se incorporan países como México, Corea o la India debido a que no existen datos sobre estas variables para estos países.

La variable dependiente es el gasto en I+D (RYD) y las variables que son propuestas para explicar a dicha variable son: el gasto social (GS), exportaciones de alta tecnología (HTEXP), inversión en conocimiento (ICON), inversión en información computadoras y telecomunicaciones (ICT), número de patentes en ICT registradas ante la EPO (ICTEPO), la productividad multifactor (MFP), la productividad (PROD) y la tendencia de las patentes (TPF). Se espera que todas las variables expliquen de manera positiva a la variable dependiente, por lo que los coeficientes estimados deberán ser positivos.

El hecho de que se tomen en cuenta estas variables como regresoras significa que la hipótesis de la que se parte es que todos los coeficientes asociados a éstas son significativamente distintos de cero.

Estimación En Datos Panel.

Una vez que ya hemos expuesto cual es la variable dependiente y cuales son las variables regresoras, lo que sigue es mostrar la representación matricial de la estimación en datos panel. Una estimación en datos panel consiste en realizar una regresión de corte transversal y series de tiempo de manera simultánea, esto es:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \dots + \beta_n X_{nit} + u_{it}$$

donde $i = 1, 2, 3 \dots n$

$t = 1, 2, 3 \dots T$

(3)

Donde i significa la i -ésima unidad transversal y t el tiempo. Se supone que hay un máximo de n unidades transversales u observaciones, y un máximo de T periodos.

Para la estimación que se realizó para el presente trabajo tenemos que $i = 1, 2, 3 \dots 7$, debido a que tenemos 7 unidades transversales, a saber, Canadá(1), Francia(2), Alemania(3), Japón(4), Suecia(5), Reino Unido(6) y Estados Unidos(7); y el tiempo que es $t = 1990, 1991, 1992, \dots 2000$, debido al periodo de análisis. La representación matricial del panel estimado, que cabe aclarar es un panel balanceado ya que para cada unidad transversal se tiene el mismo número de observaciones de series de tiempo, quedaría de la siguiente manera:

$$\begin{aligned}
 RYD_{it} = & \beta_1 + \beta_2 GS_{it} + \beta_3 HTEXP_{it} + \beta_4 ICON_{it} + \beta_5 ICT_{it} + \\
 & \beta_6 ICTEPO_{it} + \beta_7 MPF_{it} + \beta_8 PROD_{it} + \beta_9 TPF_{it} + u_{it} \quad (4) \\
 \text{con } i = & 1,2,3...7 \\
 t = & 1990,1991,1992...2000
 \end{aligned}$$

Donde u_{it} , que es el término de error, se distribuye de manera normal es decir tiene media cero y varianza constante (σ^2) y las variables regresoras se suponen no estocásticas. Puntualizando que como tenemos 7 países como unidades transversales y 11 años como variable de tiempo, el total de observaciones que se tienen para el panel es de $i \times t = 77$. Como el panel requiere de la existencia de cada variable económica para cada país, entonces se generaron 63 variables ya que son 7 países y 9 variables económicas.

Dentro de las estimaciones de datos panel existen tres modelos de regresión los cuales son; el modelo de Datos Agrupados (Pooling Data) que consiste en una regresión simple por Mínimos Cuadrados Ordinarios; el Método de Efectos Fijos (MEF) que consiste en hacer suposiciones acerca de la intersección, los coeficientes de las pendientes y el término de error existiendo varias posibilidades:

- La intersección y los coeficientes de las pendientes son constantes respecto al tiempo y el espacio, y el término de error expresa las diferencias en el tiempo y en las unidades.
- Los coeficientes de las pendientes son constantes, pero la intersección varía respecto a las unidades.
- Los coeficientes de las pendientes son constantes pero la intersección varía respecto al tiempo y las unidades.
- Todos los coeficientes varían respecto a las unidades.
- La intersección así como los coeficientes de las pendientes cambian respecto al tiempo y las unidades.

Y por último está el Modelo de Componentes de Error o Efectos Aleatorios (MCE) que consiste en suponer que las unidades de corte transversal se seleccionan en forma aleatoria y que el término de error es compuesto y por lo tanto no cumple con que se distribuye de forma normal.

Cada uno de estos modelos de regresión fue estimado arrojando como resultado una mejor estimación con el MEF, resultados que son respaldados por los cuadros de regresión que se presentan a continuación:

En la tabla 1 podemos observar una estimación panel de datos agrupados simple con mínimos cuadrados ordinarios, que nos da una primera impresión acerca del grado de significancia de las variables explicativas observando la probabilidad asociada a los coeficientes y de un posible problema de autocorrelación debido al valor que toma el estadístico Durbin-

Watson, podemos apreciar que la R^2 es aceptable ya que alcanza un valor de 0.568966 que nos dice que la variable dependiente se encuentra explicada en un 56% por las variables independientes.

Tabla 1
Panel de Datos Agrupados

Method: Pooled Least Squares
Date: 03/21/06 Time: 14:10
Sample: 1990 2000
Included observations: 11
Cross-sections included: 7
Total pool (balanced) observations: 77

<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-Statistic</i>	<i>Prob.</i>
C	-0.962313	0.383026	-2.512399	0.0144
GS?	0.056305	0.008339	6.752098	0.0000
HTEXP?	0.034189	0.008070	4.236433	0.0001
ICON?	0.313394	0.067205	4.663259	0.0000
ICT?	-0.025781	0.010365	-2.487185	0.0153
ICTEPO?	5.49E-05	2.27E-05	2.413839	0.0185
MFP?	-0.213328	0.083013	-2.569824	0.0124
PROD?	0.174807	0.067995	2.570870	0.0123
TPF?	-0.000554	0.005368	-0.103173	0.9181
R-squared	0.568966	Mean dependent var	2.427922	
Adjusted R-squared	0.518256	S.D. dependent var	0.523011	
S.E. of regression	0.363010	Akaike info criterion	0.920696	
Sum squared resid	8.960791	Schwarz criterion	1.194648	
Log likelihood	-26.44681	F-statistic	11.22003	
Durbin-Watson stat	0.336157	Prob(F-statistic)	0.000000	

Fuente: Elaboración propia con datos de la OCDE, EPO, USPTO, JPTO para 1990-2000.

Ahora bien si analizamos la tabla 2 que nos muestra una estimación de datos panel con efectos fijos podemos apreciar que a pesar que la R^2 aumenta a 0.966585, las variables ICTEPO, MFP, PROD y TPF se vuelven no significativas ya que la probabilidad asociada a estas variables es mayor a 0.05, probablemente por un problema de heterocedasticidad, además se confirma la existencia de autocorrelación debido a que el Durbin-Watson es de 1.008307. Es importante señalar que las pruebas de hipótesis individuales y conjuntas, así como la explicación de todos los estadísticos y coeficientes, se harán para el mejor modelo estimado que se presenta en la tabla 5. A continuación se presenta la tabla 2:

Tabla 2
Panel Efectos Fijos

Dependent Variable: RYD?				
Method: Pooled Least Squares				
Date: 03/21/06 Time: 14:16				
Sample: 1990 2000				
Included observations: 11				
Cross-sections included: 7				
Total pool (balanced) observations: 77				
<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-Statistic</i>	<i>Prob.</i>
C	2.434097	0.307263	7.921871	0.0000
GS?	-0.031610	0.010022	-3.154144	0.0025
HTEXP?	-0.025291	0.009443	-2.678398	0.0095
ICON?	0.180597	0.027961	6.458944	0.0000
ICT?	0.028540	0.009666	2.952578	0.0044
ICTEPO?	-1.80E-05	1.41E-05	-1.276751	0.2065
MFP?	-0.021082	0.028745	-0.733419	0.4661
PROD?	0.027149	0.022848	1.188228	0.2393
TPF?	-0.001863	0.001763	-1.056373	0.2949
Fixed Effects (Cross)				
CAN—C	-1.224032			
FRA—C	0.265630			
GER—C	0.079776			
JAP—C	0.560968			
SWE—C	0.821404			
UK—C	-0.294958			
USA—C	-0.208788			
<i>Effects Specification</i>		<i>Cross-section fixed (dummy variables)</i>		
R-squared	0.966585	Mean dependent var	2.427922	
Adjusted R-squared	0.959040	S.D. dependent var	0.523011	
S.E. of regression	0.105851	Akaike info criterion	-1.480637	
Sum squared resid	0.694669	Schwarz criterion	-1.024052	
Log likelihood	72.00454	F-statistic	128.1034	
Durbin-Watson stat	1.008307	Prob(F-statistic)	0.000000	

Fuente: Elaboración propia con datos de la OCDE, EPO, USPTO, JPPTO para 1990-2000.

La tabla 3, que presenta una estimación datos panel de efectos fijos con secciones cruzadas ponderadas, muestra que existe un problema de autocorrelación que puede ser corregido ya que la Durbin-Watson sube un poco a 0.853514, muestra también que el grado de explicación proporcionado por la R^2 aumenta a 0.991255. Debido a la estimación con secciones cruzadas ponderadas la variable ICTEPO que era no significativa ahora se vuelve significativa arrojando una probabilidad menor al 0.05, el resto de las variables que antes eran no significativas, como MFP, PROD y TPF, lo siguen siendo tal y como se muestra a continuación:

Tabla 3
Panel Efectos Fijos con Secciones Cruzadas Ponderadas

Dependent Variable: RYD?				
Method: Pooled EGLS (Cross-section weights)				
Date: 03/21/06 Time: 14:17				
Sample: 1990 2000				
Included observations: 11				
Cross-sections included: 7				
Total pool (balanced) observations: 77				
Linear estimation after one-step weighting matrix				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.400206	0.273863	8.764258	0.0000
GS?	-0.027665	0.008757	-3.159185	0.0024
HTEXP?	-0.026087	0.008681	-3.004933	0.0038
ICON?	0.166333	0.034999	4.752569	0.0000
ICT?	0.031442	0.008655	3.632766	0.0006
ICTEPO?	-2.08E-05	1.12E-05	-1.858299	0.0679
MFP?	-0.017089	0.023793	-0.718237	0.4753
PROD?	0.018733	0.018915	0.990358	0.3258
TPF?	-0.001443	0.001620	-0.890745	0.3765
Fixed Effects (Cross)				
CAN—C	-1.225562			
FRA—C	0.248019			
GER—C	0.061257			
JAP—C	0.629165			
SWE—C	0.764582			
UK—C	-0.306137			
USA—C	-0.171324			
<i>Effects Specification</i>		<i>Cross-section fixed (dummy variables)</i>		
		<i>Weighted Statistics</i>		
R-squared	0.991255	Mean dependent var	2.866423	
Adjusted R-squared	0.989281	S.D. dependent var	1.008056	
S.E. of regression	0.104367	Sum squared resid	0.675338	
F-statistic	502.0086	Durbin-Watson stat	0.853514	
Prob(F-statistic)	0.000000			
		<i>Unweighted Statistics</i>		
R-squared	0.965902	Mean dependent var	2.427922	
Sum squared resid	0.708862	Durbin-Watson stat	0.891722	

Fuente: Elaboración propia con datos de la OCDE, EPO, USPTO, JPTO para 1990-2000.

Para la estimación de panel con efectos aleatorios, que se presenta en la tabla 4, se necesitó eliminar dos variables explicativas ya que la condición para poder realizar dicha estimación es que el número de secciones cruzadas sea mayor al número de variables regresoras. Las variables que se eliminaron fueron MFP y TPF debido a que son variables no significativas y por que además siguen existiendo variables que miden productividad y patentes en la regresión, como lo hacen éstas, aún y cuando se eliminen las mismas.

Podemos observar en esta estimación que la R^2 disminuye notablemente a 0.518080 y que el estadístico Durbin-Watson también cae de manera drástica observando un valor de 0.198490. Este efecto puede ser causado muy probablemente por la exclusión de dos variables explicativas, por lo que el grado de explicación de la variable dependiente disminuye como lo muestra la R^2 .

Tabla 4
Panel Efectos Aleatorios

Dependent Variable: RYD?				
Method: Pooled EGLS (Cross-section random effects)				
Date: 03/24/06 Time: 18:20				
Sample: 1990 2000				
Included observations: 11				
Cross-sections included: 7				
Total pool (balanced) observations: 77				
Swamy and Arora estimator of component variances				
<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-Statistic</i>	<i>Prob.</i>
C	-0.600966	0.105936	-5.672895	0.0000
GS?	0.053698	0.002403	22.35067	0.0000
HTEXP?	0.032667	0.002321	14.07341	0.0000
ICON?	0.284658	0.019486	14.60803	0.0000
ICT?	-0.027202	0.002999	-9.070097	0.0000
ICTEPO?	5.40E-05	6.59E-06	8.182866	0.0000
PROD?	0.026156	0.011330	2.308596	0.0239
Random Effects (Cross)				
CAN—C	-3.72E-11			
FRA—C	-7.78E-11			
GER—C	8.97E-12			
JAP—C	8.33E-11			
SWE—C	1.15E-10			
UK—C	-7.16E-11			
USA—C	-2.02E-11			
<i>Effects Specification</i>				
Cross-section random S.D. / Rho			4.87E-07	0.0000
Idiosyncratic random S.D. / Rho			0.106476	1.0000
<i>Weighted Statistics</i>				
R-squared	0.518080	Mean dependent var		2.427922
Adjusted R-squared	0.476773	S.D. dependent var		0.523011
S.E. of regression	0.378317	Sum squared resid		10.01867
F-statistic	12.54206	Durbin-Watson stat		0.198490
Prob(F-statistic)	0.000000			
<i>Weighted Statistics</i>				
R-squared	0.518080	Mean dependent var		2.427922
Sum squared resid	10.01867	Durbin-Watson stat		0.198490

Fuente: Elaboración propia con datos de la OCDE, EPO, USPTO, JPTO para 1990-2000.

El que estos estadísticos sean tan bajos en una estimación de efectos aleatorios puede ser muestra clara de la existencia de una correlación existente entre las variables independientes y los errores. Para verificar que existe una correlación entre las variables explicativas o regresoras y el error o residuo debido a la regresión, se contrastaron mediante una gráfica de dispersión al error contra las variables explicativas para cada país. En dichas gráficas se pudo observar claramente que existe un patrón de asociación o correlación de signo positivo o negativo, esto dependiendo de las unidades en que se encuentren medidas las variables y

por ende dependiendo de la relación entre las mismas, entre el error debido a la regresión y las variables explicativas.

Por lo tanto es importante señalar que cuando se presenta dicho fenómeno entonces no se aplica el método de efectos aleatorios (MCE) y se recomienda el de efectos fijos (MEF),³ razón por la cual se descarta la estimación por MCE y se realizará un análisis más profundo con el MEF.

En la tabla 5, que se presenta a continuación, se realiza una estimación de Efectos Fijos con Secciones Cruzadas SUR que permite corregir problemas de heterocedasticidad y

Tabla 5
Panel Efectos Fijos con Secciones Cruzadas SUR

Dependent Variable: RYD?				
Method: Pooled EGLS (Cross-section SUR)				
Date: 03/21/06 Time: 14:18				
Sample: 1990 2000				
Included observations: 11				
Cross-sections included: 7				
Total pool (balanced) observations: 77				
Linear estimation after one-step weighting matrix				
<i>Variable</i>	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-Statistic</i>	<i>Prob.</i>
C	2.315193	0.112486	20.58200	0.0000
GS?	-0.026610	0.003256	-8.173475	0.0000
HTEXP?	-0.016110	0.004464	-3.608502	0.0006
ICON?	0.154617	0.018263	8.466158	0.0000
ICT?	0.023293	0.003693	6.306667	0.0000
ICTEPO?	-1.48E-05	5.58E-06	-2.652929	0.0101
MFP?	-0.010036	0.007606	-1.319503	0.1919
PROD?	0.017926	0.007046	2.544185	0.0135
TPF?	-0.001986	0.000629	-3.157728	0.0025
Fixed Effects (Cross)				
CAN—C	-1.084876			
FRA—C	0.210207			
GER—C	0.098755			
JAP—C	0.487397			
SWE—C	0.842199			
UK—C	-0.352379			
USA—C	-0.201303			
Fixed Effects (Cross)		Cross-section fixed (dummy variables)		
		Wighted Statistics		
R-squared	0.999758	8Mean dependent var	36.27143	
Adjusted R-squared	0.999703	S.D. dependent var	60.23322	
S.E. of regression	1.037818	Sum squared resid	66.77809	
F-statistic	18281.43	Durbin-Watson stat	1.660752	
Prob(F-statistic)	0.000000			
<i>Unweighted Statistics</i>				
R-squared	0.965414	2.427922		
Sum squared resid	0.719014	0.718953		

Fuente: Elaboración propia con datos de la OCDE, EPO, USPTO, JPTO para 1990-2000.

³ *Ibidem.*

autocorrelación al estimar un panel de efectos fijos, esto por que E-Views estima un modelo factible por Mínimos Cuadrados Generalizados (GLS) especificando la corrección automática de heterocedasticidad y autocorrelación.

Como se aprecia en el cuadro arriba presentado, la estimación de efectos fijos con secciones cruzadas SUR permite estimar, eliminado de manera simultánea problemas de autocorrelación y heterocedasticidad, un buen modelo explicativo del gasto en I+D.

En primer lugar se observa que, dentro de las estadísticas ponderadas, la R^2 es muy cercana a 1, es decir, tiene un valor de 0.999758 lo cual es muy bueno porque tiene dos interpretaciones importantes; la primera es que la variable dependiente se encuentra explicada por las variables independientes en un 99.97%, o que el grado de explicación del modelo es de esa misma magnitud; la segunda es que el error de la estimación es muy bajo ya que recordando la fórmula de la R^2 tenemos:

$$R^2 = 1 - \left(\frac{SCR}{SCT} \right) \quad (5)$$

Si es cercano a 1 significa que el cociente entre la suma de los residuos al cuadrado (SCR) entre la suma total de los cuadrados (SCT) es muy chico o cercano a cero, lo cual se traduce en un error muy pequeño que significa que la línea de ajuste es muy cercana o idéntica a la original.

Si contrastamos el resultado de la R^2 con el de la R^2 ajustada podemos apreciar que nuestro grado de explicación eleva su validez, ya que la fórmula de la R^2 ajustada es:

$$\tilde{R}^2 = 1 - \left(\frac{(SCR)/(n-k)}{(SCT)/(n-1)} \right) \quad (6)$$

En donde n es el número de observaciones y k el número de parámetros, pero además el término $(n-k)$ agrega un grado de penalización al agregar variables y si analizamos el valor que tiene la R^2 ajustada podemos verificar que está muy cercana a la R^2 , ya que la R^2 ajustada = 0.999703 mientras que la $R^2 = 0.999758$. Esta es la razón por la cual el error estándar de la regresión es tan chico, tan sólo de 1.037818; la media de la variable dependiente a lo largo de todo el periodo y de todas las unidades transversales es de 36.27143, y la desviación estándar de la misma es de 60.23322.

Por otro lado la suma de los residuos al cuadrado se puede verificar que en efecto es muy chica ya que alcanza solo un valor de 66.77809, y el estadístico Durbin-Watson que mide la existencia de autocorrelación o correlación serial es muy cercano a 2 ya que $DW = 1.660752$, recordando la fórmula del estadístico DW esto significa que:

$$DW = \frac{\sum_{t=2}^T \hat{e}_t^2}{\sum_{t=1}^T \hat{e}_t^2} + \frac{\sum_{t=2}^T \hat{e}_{t-1}^2}{\sum_{t=1}^T \hat{e}_{t-1}^2} - 2 \frac{\sum_{t=2}^T \hat{e}_t \hat{e}_{t-1}}{\sum_{t=1}^T \hat{e}_t^2} \quad (7)$$

$$DW \approx 1 + 1 - 2\hat{\rho}$$

$$DW \approx 2(1 - \hat{\rho})$$

Por la fórmula de la DW podemos apreciar que cuando la correlación = 0, entonces el estadístico. $DW \approx 2$. Lo cual significa que en el modelo panel estimado con secciones cruzadas SUR la autocorrelación o correlación serial es positiva pero muy pequeña.

76

En cuanto al análisis de las variables explicativas podemos apreciar en primer lugar que, salvo la variable MFP para todos los países en su conjunto, todas las variables explicativas son significativas para los países en su conjunto y que por ende explican a la variable dependiente del gasto en I+D. En segundo lugar y debido al grado de significancia de las variables todas las variables, excepto MFP, pasan la prueba de hipótesis que consiste en rechazar la hipótesis nula de que las variables independientes no explican a la variable dependiente, esto es:

$$Sea \quad t = \frac{b_k - \beta_k}{S.E.(b_k)} \quad (8)$$

Se rechaza H_0 si se cumple que $|t| > t_{(T-k)}$ donde $(T-k)$ son los grados de libertad. Como para todas las variables, excepto para MFP, el valor de la $|t|$ es mayor a 2 con una muestra de 77 observaciones por regla se cumple que para cada una de las variables explicativas, excepto para MFP, se rechaza la hipótesis nula H_0 de que las variables independientes no explican a la variable dependiente o alternativamente que el coeficiente asociado a las variables regresoras es igual a cero. Si se llegaran a tener dudas se puede contrastar $|t|$ con $t_{(T-k)}$ para cada regresor, el valor de la t de tablas con $(T-k)$ grados de libertad, en donde tenemos T número de observaciones y k número de parámetros obteniendo $77-15=62$, y con un nivel de confianza del 95%, es decir, con probabilidad del 0.05 distribuido en dos colas por lo que en cada cola tendremos 0.025, es de 2. Por lo tanto todas las variables menos MFP pasan la prueba de hipótesis.

Como se puede ver para la variable MFP no se cumple esta regla ya que el $|t| = 1.3195$, por lo tanto para esta variable regresora no se rechaza H_0 por ende que su probabilidad sea 0.1919 que es mayor a 0.05. Sin embargo como esta variable económica contribuye a la

explicación del modelo no se omitió, el hacerlo implicaría la disminución de la R^2 y la R^2 ajustada, así como probablemente del estadístico DW.

Por otro lado se realizó la Prueba de Hipótesis Conjunta a través de la prueba F siguiente:

$$Sea \quad F = \frac{(R_{NR}^2 - R_R^2) / q}{(1 - R_{NR}^2) / (nT - k)} \quad (9)$$

Donde q es el número de restricciones hechas al modelo, nT es el número de observaciones del panel y k el número de parámetros. Se toma como modelo no restringido (NR) al estimado con efectos fijos y secciones cruzadas SUR, y como modelo restringido (R) al estimado en datos agrupados.

Por lo tanto el número de restricciones es 7 ya que en el modelo de datos agrupados se supone que los interceptos son iguales para cada uno de los países, es decir, se eliminan 7 parámetros en el modelo de datos agrupados. El valor de la prueba F es:

$$F = \frac{(.999758 - .568966) / 7}{(1 - .999758) / 62} = 15766.88548 \quad (10)$$

Como es altamente significativo el valor de la prueba F entonces se rechaza la hipótesis nula (H_0) de que los interceptos son iguales para todos los países en el modelo restringido. En la tabla 5 se puede ver que el valor del estadístico F , que arroja E-Views, es de 18281.43 que es todavía mayor al arriba calculado, lo que confirma que se rechaza la hipótesis nula ya comentada y por lo tanto la prueba de hipótesis conjunta es satisfactoria.

La interpretación de los coeficientes sería la siguiente, dado que todas las variables están medidas en tasas de crecimiento anuales, menos la variable ICTEPO, entonces tenemos que cada vez que aumenta en 1% el gasto social, el gasto en I+D disminuye en 0.02% manteniendo constantes todas las demás variables. Cada vez que aumenta en 1% las exportaciones de alta tecnología el gasto en I+D cae en 0.01%, manteniendo constantes las demás variables; cada vez que aumenta un 1% la inversión en conocimiento el gasto en I+D aumenta en un 15% manteniendo constante las demás variables, podemos apreciar que esta es la variable que tiene una mayor incidencia en el gasto en I+D lo cual era de esperarse.

Cuando aumenta en 1% la inversión en información, computadoras y telecomunicaciones el gasto en I+D disminuye 0.023% manteniendo constantes las demás variables; cada vez que aumenta en una unidad las patentes registradas en la EPO de información, computadoras y telecomunicaciones, el gasto en I+D cae en 0.000015% manteniendo constantes las demás variables. Por otro lado cuando aumenta en 1% la productividad multifactorial el gasto en I+D cae en 0.01% manteniendo constantes las demás variables; cada vez que au-

menta la productividad medida en términos de horas trabajadas entre el PIB en 1%, el gasto en I+D aumenta en 0.017% manteniendo constantes las demás variables; y por último cuando aumenta en 1% la tendencia de las patentes, el gasto en I+D cae 0.001% manteniendo constantes las demás variables.

Vale la pena destacar que se esperaba que las variables afectaran de manera positiva todas y cada una de ellas a la variable dependiente, cosa que no sucede ya que como se puede verificar las variables de gasto social, exportaciones de alta tecnología, las patentes registradas en la EPO de información, computadoras y telecomunicaciones, la productividad multifactorial y la tendencia de las patentes explican de manera negativa a la variable dependiente.

Además es importante señalar que la estimación de efectos fijos denotada por la estimación $LS(CX=F,B,WGT=CXSUR) RYD? C GS? HTEXP? ICON? ICT? ICTEPO? MFP? PROD? TPF?$ calcula un intercepto para cada uno de los países siendo éstos los siguientes:

<i>Intercepto</i>	<i>Coficiente</i>
Canadá	-1.084876
Francia	0.210207
Alemania	0.098755
Japón	0.487397
Suecia	0.842199
Reino Unido	-0.352379
Estados Unidos	-0.201303

Conclusiones

A través de la estimación en datos panel realizada podemos concluir que en realidad sí existen otros factores, además de la estructura de mercado, la posición de la industria en el mismo, el valor de mercado, etc., que determinan el gasto en I+D. Es importante reconocer que el conocimiento, la productividad, el gasto social, las patentes, entre otras variables aportan un grado de explicación dentro de la actividad innovadora de los países desarrollados y que, sin duda, la innovación y el desarrollo tecnológico son piezas claves para lograr el desarrollo económico de los países.

Ahora bien, hay que tomar en cuenta que este análisis fue realizado para países integrantes de la OCDE y que además son desarrollados, razón por la cual las variables son tan significativas, debido a que no existen datos disponibles en las variables utilizadas para países subdesarrollados, quedando abierta la posibilidad de un análisis para países como México, Corea, India, Turquía, etc., que logre enriquecer este tipo de análisis. La esencia de los modelos en datos tipo panel es rescatar las divergencias o polaridades que se desprenden al variar las unidades de análisis que en este caso son los países ya sean desarrollados o en vías de desarrollo.

Por otro lado queda abierta también la posibilidad de un desarrollo teórico que logre explicar las relaciones encontradas por este análisis, tomando en cuenta que variables como las exportaciones de alta tecnología, patentes y productividad multifactorial explican de manera inversa al gasto en I+D, cuando se esperaba que fuera de manera directa, por lo que se complica el ejercicio teórico para lograr compatibilizar la teoría económica con la teoría econométrica. Una posible salida sería tomar en cuenta la existencia de costos altos por derecho de autor y piratería, para explicar de manera teórica el carácter negativo de los coeficientes asociados a estas variables.

Bibliografía

- Arrow, K (1962). "The Economic Implications of Learning by Doing", *The Review of Economic Studies*, Vo. 29, No 3, PP 155-173.
- Blundell, R., Griffith, R. & Van Reenen, J. (1999). "Market share, market value, and innovation in a panel of British manufacturing firms". *Review of Economic Studies*, 66, 529-554.
- Griliches, Z. (1979). "Issues in assessing the contribution of research and development to productivity growth". *Bell Journal of Economics*, 10(1), 92-116.
- Griliches, Z. and J. Mairesse (1984). "Productivity and R&D at the Firm Level," in: Z. Griliches (Ed.). *R&D, Patents, and Productivity*, pp. 339-374. Chicago: University of Chicago Press.
- Griliches, Z. (1986). "Productivity, R&D, and Basic Research at the Firm Level in the 1970's," *American Economic Review*. Vol. 76, No. 1, pp. 141-154.
- Griliches, Z. 1990. "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature* XXVIII: 1661-1707.
- Griliches, Z., B.H. Hall, and A. Pakes (1991). "R&D, Patents and Market Value Revisited: Is there a Second (Technological Opportunity) Factor?" *Economics of Innovation and New Technology* 1: 183-202.
- Gujarati, D.D. (2004), *Econometría*. McGraw-Hill. Cap 16.
- Hall, B. (2000). Innovation and market value. In R. Barrell, G. Mason & M. O'Mahoney (Eds.), *Productivity, Innovation and Economic Performance*. Cambridge: Cambridge University Press.
- Hall, B. & Ziedonis, R. (2001). The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979-1995. *Rand Journal of Economics*, 32(1), 101-28.
- Hayashi, Fumio, and Tonru Inoue. 1991. "The Relation between Firm Growth and q with Multiple Capital Goods: Theory and Evidence from Panel Data on Japanese Firms," *Econometrica* 59 (3): 731-753.
- Help E-views 5, Pool Estimate, Cross-sections SUR.
- Lerner, J. (1997). Empirical exploration of a technology race. *Rand Journal of Economics*, 28(2), 228-247.
- Mansfield, E. (1986). "Patents and Innovation: An Empirical Study," *Management Science* 32(2): 173-181.
- Romer, P (1990). "Endogenous Technological Change", *Journal of Political Economy*, Vol. 98, No. 5, pp. 71-103.
- <http://www.european-patent-office.org/index.en.php>
- <http://www.jpo.go.jp/>
- <http://www.oecd.org/topicstatsportal>
- <http://stats.oecd.org/WBOS>
- <http://www.uspto.gov/>