Análisis de curvas de aprendizaje en el sector de la construcción española

JIMÉNEZ TORRES, Fernando #
fjimenez@posta.unizar.es

FAULÍN, Javier *
javier.faulin@unavarra.es

Departamento de Métodos Estadísticos
Universidad de Zaragoza

* Departamento de Estadística e Investigación Operativa
Universidad Pública de Navarra

Palabras Clave: curvas de aprendizaje, economías de escala, regresión múltiple.

Resumen

En este trabajo se presenta un estudio comparativo realizado en el sector de la construcción en España. Se propone un modelo de regresión múltiple para confirmar la existencia de efectos producidos por la experiencia acumulada en el transcurso del tiempo, con objeto de comprobar si realmente hay implícito algún efecto de aprendizaje en las diferentes comunidades autónomas españolas referente al citado sector.

Asimismo, se comprueba que el empleo de curvas de aprendizaje puede ayudar a corroborar el efecto de variables incorrectamente omitidas en un análisis de regresión y cómo implementar caminos alternativos de tests de hipótesis.

1. INTRODUCCIÓN

La existencia de economías de escala y de curvas de aprendizaje tiene importantes implicaciones para la estructura de mercado y beneficio económico, ya que tales fenómenos pueden crear barreras de entrada al presentar reducciones de costes unitarios en la producción. En bien de los intereses de la empresa, y con objeto de atraer la atención de los accionistas, uno de los aspectos que envuelven los logros productivos se centra en la reducción de costes de producción. Así, mientras las economías de escala consiguen reducir costes al aumentar el nivel de producción, las curvas de aprendizaje obtienen el mismo objetivo debido a la eficiencia productiva derivada de la experiencia acumulada por la empresa.

Según esta línea de pensamiento, las economías de escala pueden ser aprovechadas de forma que el coste en media o por unidad disminuya, al mismo tiempo que el nivel de producción aumenta por periodo. Si tales economías de escala son factibles, puede entonces ser factible para la maximización del beneficio o para la minimización del coste acelerar los planes de inversión o reducir precios, y por consiguiente alcanzar niveles más altos de producción con unidades de coste más bajo. Las economías de escala también tienen importantes implicaciones para la estructura competitiva de una industria, aunque casi siempre es necesario un gran gasto fijo de capital inicial que a menudo se requiere antes de que cualquier producción tenga lugar.

Otro importante determinante que afecta a los costes de producción es la llamada curva de aprendizaje. Existen muchos casos prácticos en los que las curvas de aprendizaje han sido muy importantes para la formulación de políticas y de estrategias (Gruber, 1992). Una manera natural para percatarse del efecto del aprendizaje consiste en observar cómo se reducen con el tiempo los costes unitarios de producción al mantener ésta en un nivel constante, o bien, comprobando que el tiempo necesario para la producción de una unidad disminuye conforme aumenta el número de unidades producidas.

El trabajo queda establecido como sigue. En la sección 2 se introduce la metodología propuesta junto con una breve descripción de la terminología a utilizar; en la sección 3 se lleva a cabo la aplicación de la metodología presentada a datos reales procedentes del sector de la construcción en España y en el periodo 1989-1999; por último en la sección 4, se comentan las aportaciones generales obtenidas y las conclusiones del trabajo.

2. METODOLOGÍA

La relación entre factores productivos (inputs) y producción (output) puede resumirse en la conocida función de producción de Cobb-Douglas, cuya formulación general para tres factores productivos x_1, x_2 y x_3 es:

$$P = E\, x_1^{\lambda_1} x_2^{\lambda_2} x_3^{\lambda_3} \qquad (1)$$

donde P es el nivel de producción; E indica el estado del conocimiento tecnológico de la empresa; y λ_1, λ_2, λ_2 son los parámetros del modelo a estimar.

La función de coste dual que se obtiene de la función de producción, tras efectuar la correspondiente transformación logarítmica es:

$$\log C = \log k + \frac{1}{L}\log P + \frac{\lambda_1}{L}\log p_1 + \frac{\lambda_2}{L}\log p_2 + \frac{\lambda_3}{L}\log p_3 + e_t \qquad (2)$$

donde C es el coste total asociado al nivel de producción P; $L = \lambda_1 + \lambda_2 + \lambda_3$ son los retornos de escala[1] de la función de producción; p_1, p_2 y p_3 son los precios unitarios de los inputs; $k = L\left(E\lambda_1^{\lambda_1}\lambda_2^{\lambda_2}\lambda_3^{\lambda_3}\right)^{-1/L}$ y e_t el término aleatorio de error, con las habituales hipótesis para la perturbación: i) Esperanza nula: $E[e_i] = 0$; ii) Homocedasticidad: $Var(e_i) = \sigma^2$; iii) Normalidad: $e_i \sim N(0,\sigma^2)$; iv) Independencia: $E[e_ie_j] = 0$, con $i \neq j$.

La ecuación que normalmente suele emplearse para representar curvas de aprendizaje es:

$$C_t = C_1 N_t^{\alpha_{ca}} \exp(e_t) \qquad (3)$$

---

[1] Medida como L-1. Si L > 1, los retornos de escala son crecientes (economías de escala positivas). Si los retornos de escala son constantes (L = 1), las economías de escala son nulas. Si L < 1, los retornos de escala son decrecientes (economías de escala negativas).

donde C_t indica el coste por unidad durante el periodo de tiempo anterior a t; C_1 es el coste unitario inicial (en el periodo t=0); N_t la producción total acumulada durante el periodo de tiempo anterior a t; α_{ca} la elasticidad del coste por unidad con respecto a la producción acumulada; y e_t la perturbación aleatoria. Una formulación equivalente tomado logaritmos es:

$$\log C_t = \log C_1 + \alpha_{ca} \log N_t + e_t \qquad (4)$$

es decir:

$$\log C_t = \alpha_0 + \alpha_{ca} \log N_t + e_t \qquad (5)$$

Por otra parte, es importante destacar que si los retornos de escala no son constantes ($L \neq 1$), no es posible obtener la curva de aprendizaje en la forma en que ha sido presentada, ya que los costes estarían influidos por el nivel de producción. En este caso la ecuación válida es:

$$\log C_t = \log k' + \frac{\alpha_{ca}}{L} \log N_t + \frac{1-L}{L} \log P_t + e_t \qquad (6)$$

donde $k' = L\left(\lambda_1^{\lambda_1} \lambda_2^{\lambda_2} \lambda_3^{\lambda_3}\right)^{1/L}$. La expresión (6) puede escribirse:

$$\log C_t = \beta_0 + \beta_1 \log N_t + \beta_2 \log P_t + e_t \qquad (7)$$

Si nuestros datos verifican que los retornos de escala no son constantes, siendo por tanto el valor de L diferente de la unidad, y aún así estimamos el modelo con la ecuación (5) en lugar de utilizar (7), se produce un sesgo que podemos calcular mediante la ecuación:

$$\log P_t = \gamma_0 + \gamma_1 \log N_t + e_t \qquad (8)$$

Si denotamos por $\hat{\alpha}'s$, $\hat{\beta}'s$ y $\hat{\gamma}'s$ las estimaciones mínimo cuadráticas de los correspondientes parámetros de las ecuaciones (5), (7) y (8), respectivamente, la diferencia entre utilizar (5) ó (7) para estimar la curva de aprendizaje a nivel del parámetro que más nos interesa (α_{ca}) es:

$$\hat{\alpha}_1 - \hat{\beta}_1 = \hat{\beta}_2 \, \hat{\gamma}_1 \qquad (9)$$

Así, el sesgo de omitir log P_t en (5) es igual a $\hat{\alpha}_1 - \hat{\beta}_1$, y por (9) coincide con $\hat{\beta}_2 \, \hat{\gamma}_1$. Este sesgo sólo será nulo cuando, o bien el nivel de producción actual P_t es incorrelado con la producción acumulada N_t, en cuyo caso $\hat{\gamma}_1 = 0$, o bien, el coste el coste por unidad durante el periodo de tiempo anterior a t no depende de la producción actual ($\hat{\beta}_2 = 0$).

Si ninguna de estas dos condiciones se verifica se producirá un sesgo que será el responsable de que la verdadera curva de aprendizaje sea sobreestimada o infraestimada, dependiendo de cómo sean los retornos de escala (L-1). Si los retornos de escala son decrecientes ($L < 1$) el sesgo será positivo. Si los retornos de escala son constantes ($L = 1$), no existirá sesgo. Si los retornos de escala son crecientes ($L > 1$) el sesgo será negativo.

3. APLICACIÓN DEL MODELO

En este apartado ilustramos una aplicación del concepto de curva de aprendizaje al sector de la construcción en España durante los años 1989-1999. Disponemos del número total de viviendas terminadas (agregando las Viviendas de Protección Oficial (VPO) a las Viviendas Libres) por comunidades autónomas y del número de ocupados en el sector de la construcción, igualmente por comunidades y durante el intervalo temporal especificado. La información considerada, obtenida del Anuario Estadístico de España, ha sido cedida por el Instituto Nacional de Estadística (Dirección General de la Vivienda, la Arquitectura y el Urbanismo, Ministerio de Fomento).

En nuestro caso a analizar, interpretaremos el coste como el número de empleados necesarios para terminar una vivienda (ya sea VPO o Libre). En estas condiciones, el primer paso consiste en estimar el modelo de curva de aprendizaje (7) para contrastar la hipótesis nula de la existencia de retornos de escala constantes ($L = 1$), es decir, que el coeficiente asociado a la producción (β_2) sea nulo.

En la Figura 1 se aprecia claramente que existen cuatro comunidades autónomas, en concreto, Islas Baleares, Islas Canarias, Región de Murcia y la Comunidad Valenciana, donde en principio se produce el efecto contrario al esperado. Dejaremos para otro trabajo el análisis de estas comunidades y nos centraremos en las trece comunidades españolas restantes para ilustrar el cálculo e interpretación de las curvas de aprendizaje que se obtengan. Esta figura muestra la representación gráfica del número de viviendas terminadas acumuladas frente al número de ocupados por vivienda terminada. Puede observarse que, con excepción de las cuatro comunidades mencionadas, en el resto es apreciable el decrecimiento esperado para estudiar el efecto de la curva de aprendizaje.

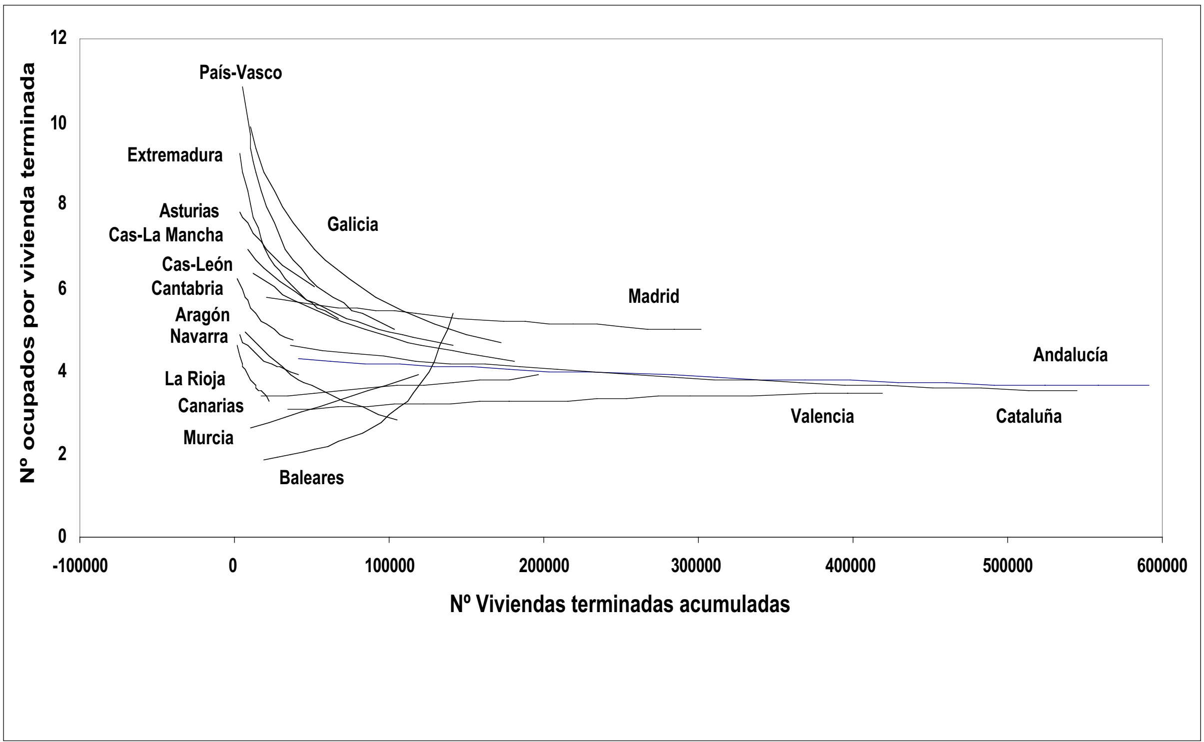

Figura 1: Número de viviendas terminadas acumuladas frente al número de ocupados por vivienda terminada.

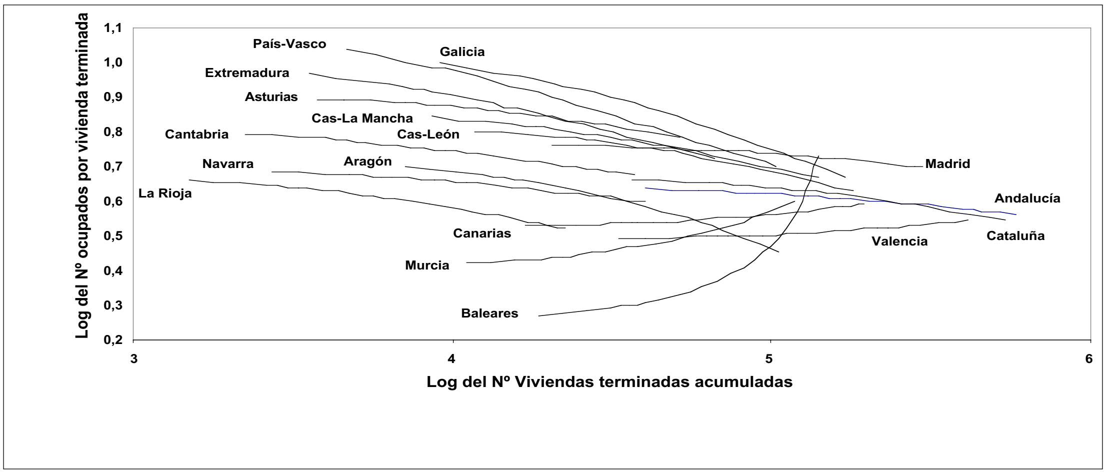

Figura 2: Logaritmo del número de viviendas terminadas acumuladas frente al logaritmo del número de ocupados por vivienda terminada.

En la Figura 2 se representan los logaritmos de las mismas variables que las que aparecen en la Figura 1. En este nuevo gráfico se observa la fuerte relación lineal entre las variables, con una acusada pendiente negativa en algunas comunidades. Así, calcularemos la curva de aprendizaje para todas las comunidades autónomas donde se admite la existencia de retornos de escala constantes ($\beta_2 = 0$). El resultado de este estudio puede verse en la Tabla 1, donde se observa que en las trece comunidades autónomas estudiadas el p-valor es lo suficientemente significativo (mayor de 0,05) como para aceptar la hipótesis nula de que el coeficiente β_2 es nulo en todas ellas. Esto significa que se admite la existencia de retornos de escala constantes en estas comunidades.

Comunidad Autónoma	β_2	**Std. Error**	**p-valor**
Andalucía	-0,027	0,013	0,112
Aragón	-0,035	0,009	0,213
Principado de Asturias	-0,003	0,022	0,103
Cantabria	-0,005	0,025	0,251
Castilla y León	-0,007	0,008	0,192
Castilla La Mancha	-0,002	0,017	0,151
Cataluña	-0,001	0,018	0,223
Extremadura	-0,005	0,003	0,217

Galicia	-0,01	0,004	0,111
Comunidad de Madrid	-0,009	0,012	0,212
Navarra	-0,008	0,004	0,261
País Vasco	-0,002	0,003	0,351
La Rioja	-0,004	0,0017	0,237

Tabla 1: Contraste de retornos de escala constantes ($H_0 \equiv \beta_2 = 0$; obtenido con SPSS).

En la Tabla 2 se muestra un resumen estadístico de las trece regresiones simples efectuadas, para las estimaciones de las curvas de aprendizaje en las comunidades autónomas donde se ha admitido la existencia de retornos de escala constantes. Puede observarse que el modelo es válido para todas estas comunidades, con valores del coeficiente de determinación (R^2) muy cercanos a la unidad, estableciendo que en todos los casos, más del 96% de la variabilidad total queda explicada por el modelo. Por otra parte, la significación general del modelo, con p-valores del F-ratio casi nulos, corrobora la afirmación sugerida por el coeficiente de determinación como medidas de bondad de ajuste. Además, los valores del t-ratio de los coeficientes de cada una de las rectas de regresión efectuadas son también casi nulos, indicando que tales coeficientes son significativos, es decir, no nulos.

Comunid.	α_0	α_{ca}	**Std. error**		R^2	α_0		α_{ca}		**Modelo**	
			α_0	α_{ca}		**t-rat**	**pval**	**t-rat**	**pval**	**F-rat**	**pval**
Andalucía	0,94	-0,06	0,02	0,004	0,96	43,8	0,00	-16,4	0,00	269,9	0,00
Aragón	1,54	-0,21	0,07	0,01	0,95	20,7	0,00	-13,2	0,00	175,3	0,00
Asturias	1,28	-0,10	0,03	0,007	0,95	37,1	0,00	-13,1	0,00	173,7	0,00

Cantabria	1,13	-0,10	0,01	0,004	0,98	64,7	0,00	-23,7	0,00	561,8	0,00
Cast-León	1,45	-0,15	0,04	0,009	0,96	30,3	0,00	-15,7	0,00	247,3	0,00
Cast-Man.	1,44	-0,14	0,03	0,007	0,97	39,0	0,00	-19,0	0,00	364,5	0,00
Cataluña	1,14	-0,10	0,03	0,006	0,96	33,9	0,00	-16,3	0,00	265,8	0,00
Extrem.	1,61	-0,20	0,03	0,007	0,98	49,8	0,00	-25,8	0,00	665,7	0,00
Galicia	2,08	-0,26	0,06	0,012	0,98	34,2	0,00	-21,0	0,00	443,9	0,00
C. Madrid	1,01	-0,0	0,01	0,003	0,96	56,5	0,00	-16,2	0,00	264,8	0,00
Navarra	0,97	-0,08	0,01	0,004	0,96	48,8	0,00	-17,0	0,00	289,0	0,00
País Vasco	1,99	-0,25	0,03	0,006	0,99	65,2	0,00	-38,1	0,00	1455,	0,00
La Rioja	1,07	-0,12	0,00	0,007	0,96	35,4	0,00	-16,1	0,00	267,0	0,00

Tabla 2: Estimaciones de las curvas de aprendizaje (obtenidas con SPSS).

La pendiente de la recta de regresión es la elasticidad estimada (α_{ca}) del número de ocupados por vivienda terminada con respecto al número de viviendas terminadas acumuladas. Así, al duplicar el número de viviendas terminadas acumuladas, también se duplica la experiencia acumulada, y por tanto, el número de ocupados por vivienda terminada, aproximadamente, debe decrecer en un porcentaje igual a $2^{\alpha_{ac}}$. Esta cantidad recibe el nombre de pendiente de la curva de aprendizaje, y en la Tabla 3 se presentan las trece comunidades autónomas, en orden decreciente, según su pendiente.

Comunidad Autónoma	α_{ca}	$2^{\alpha_{ac}}$ (%)
Comunidad de Madrid	-0,05	96,59
Andalucía	-0,066	95,53
Navarra	-0,08	94,61
Cantabria	-0,1	93,3
Cataluña	-0,103	93,11
Principado de Asturias	-0,105	92,98
La Rioja	-0,125	91,7
Castilla La Mancha	-0,149	90,19
Castilla y León	-0,155	89,81
Extremadura	-0,2	87,06
Aragón	-0,213	86,27
País Vasco	-0,256	83,74
Galicia	-0,268	83,05

Tabla 3. Pendiente de la curva de aprendizaje.

Continuando el análisis de la Tabla 3, por mera inspección de las pendientes ($2^{\alpha_{ac}}$), son la Comunidad de Madrid y Andalucía, con valores prácticamente similares y seguidas muy de cerca por Navarra, las comunidades autónomas en las que puede observarse fielmente el efecto de aprendizaje. Si el nivel de producción acumulada aumenta en una unidad, el coste producido va a aumentar en una cuantía inferior al del resto de las demás comunidades.

Por otro lado, en el estudio realizado destaca Galicia. Es la comunidad donde el aprendizaje está menos marcado, y donde los gastos son mayores con relación al resto, aunque las diferencias no son muy elevadas y teniendo siempre en mente que se han excluido de este estudio las comunidades autónomas de Islas Baleares, Islas Canarias, Región de Murcia y la Comunidad Valenciana.

4. CONCLUSIONES

En este trabajo se ha presentado una metodología para el cálculo de curvas de aprendizaje basada en modelos de regresión simple y múltiple, según se confirme la existencia de retornos de escala constantes. Se define también el sesgo producido al emplear erróneamente una regresión simple para estimar la curva de aprendizaje, cuando en realidad se debiera utilizar un modelo de regresión múltiple.

Para ilustrar la metodología proyectada se ha elegido el sector de la construcción española ya que, tradicionalmente, es uno de los sectores donde mejor puede observarse el fenómeno del aprendizaje. El objetivo de esta aplicación se centra en orientar al posible inversor sobre las comunidades autónomas españolas donde el efecto de aprendizaje está más acentuado, siendo especialmente la Comunidad de Madrid y Andalucía, donde se observa con mayor claridad.

5. REFERENCIAS

Abernathy, W.J. y Wayne, K. (1974): “Limits of the Learning Curve”. *Harvard Business Review*, septiembre-octubre.

Adler, P.S. (1990). “Shared Learning”. *Management Science*, agosto.

Adler, P.S. y Clarck, K.B. (1991). "Behind the Learning Curve: A Sketch of the Learning Process". *Management Science*, marzo.

Bailey, C.D. y McIntire, E.V. (1992): "Some Evidence on the Nature of Relearning Curves". *The Accounting Review*, abril.

Gruber, H. (1992): "The Learning Curve in the Production of Semiconductor Memory Chips". *Applied Economics*, agosto.

Hirschman, W.B. (1964): "Profit from the Learning Curve". *Harvard Business Review*, enero-febrero.

Pattison, D.D. y Teplitz, Ch.J. (1989): "Are Learning Curves Still Relevant?". *Management Accounting*, febrero.