

Dossier: «Jornada laboral, productividad e inteligencia artificial» coordinado por Pau Cortadas Guasch

Estimación de la productividad a partir de una aproximación a la cuantificación del flujo de conocimiento en la industria española

Ángel Díaz-Chao

Universidad Rey Juan Carlos

RESUMEN Este estudio propone una metodología para la cuantificación del flujo de conocimiento en empresas manufactureras españolas, utilizando datos de la *Encuesta Sobre Estrategias Empresariales (ESEE)*. Se desarrolla un índice compuesto basado en un modelo de ecuaciones estructurales (SEM) con variables latentes, eliminando la arbitrariedad en la ponderación de las dimensiones. El análisis incluye siete dimensiones clave relacionadas con innovación, I+D, organización del trabajo e internet. Los resultados muestran que las dimensiones asociadas a la I+D tienen un mayor impacto en el flujo de conocimiento y, en última instancia, en la productividad empresarial. La combinación de la inversión en investigación y desarrollo (I+D) y la adopción de nuevos métodos organizativos en el entorno laboral presenta una relación positiva y estadísticamente significativa con el aumento de la productividad. Estos hallazgos respaldan la necesidad de que las empresas adopten estrategias que integren tanto el desarrollo tecnológico como la innovación organizativa. En este contexto, el índice de flujo de conocimiento refuerza la relevancia de ambas dimensiones. Este enfoque ofrece una herramienta estadísticamente válida para evaluar la competitividad empresarial y subraya la importancia de integrar tecnología e innovación organizativa en las estrategias corporativas.

PALABRAS CLAVE flujo de conocimiento; innovación; investigación y desarrollo (I+D); productividad empresarial; encuesta sobre estrategias empresariales (ESEE); ecuaciones estructurales (SEM); organización del trabajo

Estimating productivity from an approximation to quantifying the knowledge flow in the Spanish industry

ABSTRACT This study introduces a methodology to quantify knowledge flow in Spanish manufacturing firms, utilizing data from the Enterprise Strategy Survey (ESEE). A composite index is created through a structural equation model (SEM) featuring latent variables, which removes any subjectivity in weighting the dimensions. The analysis encompasses seven critical dimensions concerning innovation, R&D, work organization, and the Internet. Results indicate that dimensions linked to R&D significantly influence knowledge flow and, consequently, business productivity. The combination of R&D investment and the implementation of new organizational approaches in the workplace shows a positive and statistically significant correlation with enhanced productivity. These findings emphasize companies' need to embrace strategies that merge technological advancement with organizational innovation. In this regard, the knowledge flow index highlights the importance of incorporating both elements. This method provides a statistically sound tool for evaluating business competitiveness, emphasizing the necessity of integrating technology and organizational innovation within corporations' strategies.

KEYWORDS knowledge flow; innovation; research and development (R&D); business productivity; enterprise strategy survey (ESEE); structural equations (SEM); work organization

Introducción

Los flujos de conocimiento han sido analizados en muy diversas ocasiones en la literatura científica, tanto a nivel internacional como a nivel nacional para el caso español. Esta línea de investigación se ha incrementado en las últimas décadas desde la irrupción de las TIC hasta el desarrollo de la inteligencia artificial, muy vinculada en sus desarrollos con el estudio de la productividad en las organizaciones. Sin embargo, pocas veces se ha implementado un estudio de su cuantificación con índices. En este artículo, se aborda una propuesta metodológica para la cuantificación del flujo de conocimiento en las empresas manufactureras españolas a partir de la información contenida en la *Encuesta Sobre Estrategias Empresariales (ESEE)*. El índice propuesto permite demostrar la relación entre los avances tecnológicos y los cambios en la forma de trabajar y los aumentos de productividad.

En este sentido, en este artículo se aborda, en primer lugar, la problemática de la construcción de índices que, habitualmente, tiene su punto débil en la arbitrariedad de la fijación de los pesos que componen los índices. De esta forma, y siguiendo a Díaz-Chao y Montes (2014), donde se aborda la construcción de índice verde, desarrollaremos un modelo basado en ecuaciones estructurales con variables latentes (SEM) para estimar un índice de flujo de conocimiento. La principal bondad de esta metodología radica en que los pesos de las variables consideradas en el índice son estimados por el propio modelo econométrico, eliminando la posible arbitrariedad en su fijación.

A modo de ejemplo, se estiman los efectos de la variable creada como índice sobre la productividad horaria por trabajador medida en logaritmos para el año 2018.

1. Metodología

1.1. El problema de la construcción de índices

La problemática de la construcción de índices es algo que persiste desde hace muchas décadas en el ámbito de la investigación científica. Si bien es cierto que existe consenso respecto a la importancia y necesidad del uso de índices para la cuantificación en una sola variable o matriz de datos de fenómenos sociales o económicos (OCDE, 2008; Eurostat, 1996), no hay tal unanimidad respecto a que la forma de construcción se refiera desde un punto de vista técnico (Melser y Hill, 2007).

En nuestro caso, dado que nuestra variable objeto de medida tiene numerosos factores, el índice que la mida debe ser necesariamente complejo, entendiendo complejo, desde un punto de vista estadístico, como aquella expresión matemática compuesta por más de un factor o variable.

A partir de aquí, surgen dos problemas clásicos en la construcción del índice:

- Las variables a utilizar, denominadas *dimensiones*.
- Segundo, el peso que se asigna a cada una de ellas y que origina la contribución de cada variable al cálculo de índice y, por tanto, a su variación, si fuera el caso.

Como es natural, desde la óptica matemática del propio índice, la ponderación de las variables o factores elegidos se ve afectada significativamente por la propia elección de los factores. La elección del peso de los factores que componen el índice, en muchas ocasiones, no responden a razones científicas, sino que son asignados de forma arbitraria y más bien como un axioma que, por su propio concepto, es irrefutable y que suele estar más vinculado a intuiciones que razones científicas o matemáticas. En definitiva, no responden a una razón matemática sustentada en una teoría contrastada científicamente o por técnicas estadísticas o econométricas. Naturalmente, el desarrollo metodológico no suele ser explícito en los distintos informes que lo explican. Todo ello conduce a afirmar que muchos de estos índices presentan problemas de tipo analítico, metodológico y cuantitativo.

En resumen, la construcción de un índice presenta dos problemas: la elección de variables y la metodología empleada para su propia construcción.

En lo que respecta al primer problema, y como se expone más adelante, en nuestro caso utilizaremos la *Encuesta Sobre Estrategias Empresariales (ESEE)* y será la propia disponibilidad de datos la que condicione la propia elección de variables para el caso analizado.

En el caso del segundo problema, la metodología, en este artículo se propone la consideración del propio índice como una variable latente, esto es, no observable, cuya composición debe venir determinada con los propios factores o dimensiones seleccionados. Para ello, se desarrollará un modelo general de ecuaciones estructurales (SEM) con variables latentes, donde consideraremos el propio índice como variable latente, tal y como se expone a continuación.

Esta metodología solventa el problema de la arbitrariedad de los pesos, siendo el propio modelo econométrico el que estima los valores de las ponderaciones que aplicar.

1.2. Metodología: el modelo general de sistemas de ecuaciones estructurales

Tal y como se analiza en Díaz-Chao y Montes (2014), el modelo general de los sistemas de ecuaciones estructurales es un modelo matemático formal, un grupo de ecuaciones lineales que, como caso particular, se compone de diferentes tipos de modelos como modelos de regresión, sistemas de ecuaciones simultáneas, análisis factorial y análisis de trayectorias. Las variables en el sistema de ecuaciones pueden ser tanto variables directamente observadas y medibles como variables latentes (teóricas), que representan conceptos que no son directamente observables, aquellos que solo pueden ser tratados y medidos a través de variables observadas y medibles. Las variables latentes deben ser continuas; sin embargo, las variables dependientes, observadas, pueden ser continuas, censuradas, binarias, ordenadas, categóricas (ordinales), o combinaciones de todos estos tipos.

El modelo general se compone de dos submodelos: el modelo estructural, que relaciona las variables latentes entre sí, y el modelo de medida, que relaciona cada variable latente con las variables correspondientes que la miden (denominadas *indicadores*). Normalmente, se asume que existe una estructura causal entre las variables latentes.

Los modelos de ecuaciones estructurales tienen ciertas características distintivas y deseadas, como:

- a) permiten la inclusión explícita del error de medida en el proceso de valoración en todas las variables que se consideren adecuadas;
- b) la valoración simultánea de los parámetros de una serie de relaciones de dependencia, donde una variable puede actuar como dependiente en unas ecuaciones e independiente en otras;
- c) pueden agrupar causales recíprocas y modelos recursivos y no recursivos; y
- d) aunque es una técnica confirmatoria, nuevos desarrollos permiten también su uso como exploratoria.

Una vez descrita la metodología, en el punto 2 se expone la ESEE como base de datos que utilizar, así como las variables y dimensiones que conformarán la estimación del índice como previo a las estimaciones.

2. Los datos: la *Encuesta Sobre Estrategias Empresariales*

La *Encuesta Sobre Estrategias Empresariales (ESEE)*, elaborada por la Fundación SEPI, ofrece datos de la industria manufacturera española sobre un amplio abanico de aspectos relacionados con la estrategia empresarial de dicho sector. Principalmente, cabe destacar, además de los datos relativos a las cuentas y beneficios empresariales, los datos que recaba sobre innovación, producción, empleo y, en las últimas oleadas, inteligencia artificial y sostenibilidad. Al mismo tiempo, uno de los grandes valores añadidos de la ESEE es la disponibilidad de datos de panel, concretamente desde el año 1990 hasta el año 2022, último ejercicio disponible hasta la fecha. Esta característica facilita la investigación con microdatos de empresas desde una óptica dinámica, lo cual permite análisis más consistentes en términos de las relaciones causales que la teoría pueda apuntar en términos intertemporales.

La ESEE tiene dos tipos de variables disponibles: variables de carácter anual, que son recogidas en cada oleada; y variables que son más estables en el tiempo, por su propia naturaleza, y que son recabadas cada cuatro años. A este respecto, la ESEE recoge con un cuestionario completo datos cada cuatro años, mientras que los tres años restantes, recoge datos en cuestionarios reducidos. Para el objeto de análisis en este artículo, utilizaremos el año 2018 por ser año completo en la cumplimentación del cuestionario, de modo que el posible uso de variables cuatrianuales haga más robustas las estimaciones llevadas a cabo.

Para la construcción del índice se han seleccionado quince variables agrupadas en siete dimensiones como se puede ver en la siguiente tabla.

Tabla 1. Variables y dimensiones

Dimensión	Variable
Dimensión 1. Innovación (INNOV)	• Innovaciones de producto, proceso, comercialización y organización del trabajo

Dimensión	Variable
Dimensión 2. Nuevas formas de organización del trabajo (NFOT)	<ul style="list-style-type: none"> Innovaciones de organización del trabajo, innovaciones en la gestión de las relaciones externas, innovaciones de proceso por nuevas técnicas
Dimensión 3. Uso de internet (USOINT)	<ul style="list-style-type: none"> Ventas a consumidores finales por internet Ventas a empresas por internet Incidencia de internet sobre las ventas
Dimensión 4. Tecnologías de internet (TECINT)	<ul style="list-style-type: none"> Dominio propio en internet. Compras a proveedores por internet Página web en servidores de la empresa
Dimensión 5. Tecnologías de I+D (TECID)	<ul style="list-style-type: none"> Dirección o comité de tecnología. Utilización de asesores para informarse sobre internet Evaluación de perspectivas sobre cambios tecnológicos
Dimensión 6. Recursos de I+D (RECID)	<ul style="list-style-type: none"> Reclutamiento de personal con experiencia en I+D Acuerdos de cooperación tecnológica. Aplicación de incentivos fiscales a la I+D
Dimensión 7. Actividades de I+D (AID)	<ul style="list-style-type: none"> Realización o contratación de actividades en I+D

Fuente: elaboración propia

La construcción de cada una de las dimensiones se realiza como suma lineal de cada una de las variables contempladas en cada dimensión, de manera que la variable resultante es una variable discreta que toma valores desde 0, indicando este valor que no tiene actividad en el indicador correspondiente. Lógicamente, a mayor valor de la dimensión, mayor intensidad en la dimensión analizada.

Tabla 2. Indicadores y dimensiones

Variable	Descripción de la variable
IP	Innovaciones de producto
IPR	Innovaciones de proceso
IMO	Innovaciones en métodos organizativos
ICO	Innovaciones de comercialización
IMOPE	Innovaciones en la organización del trabajo
IMOGÉ	Innovaciones en la gestión de relaciones externas
IPRTM	Innovaciones de proceso por nuevas técnicas
WEBB2C	Ventas a consumidores finales por internet
WEBB2B	Ventas a empresas por internet
WEBVEN	Incidencia de internet sobre las ventas
WEBPRO	Dominio propio en internet
WEBCOM	Compras a proveedores por internet
WEBEMP	Página web en servidores de la empresa
DCT	Dirección o comité de tecnología

Variable	Descripción de la variable
UAIT	Utilización de asesores para informarse sobre internet
EPCT	Evaluación de perspectivas sobre cambios tecnológicos
REEID	Reclutamiento de personal con experiencia en I+D
ACT	Acuerdos de cooperación tecnológica
APLIFN	Aplicación de incentivos fiscales a la I+D
AID	Actividades de I+D

Fuente: elaboración propia

Las siete dimensiones analizadas nos permitirán estimar un índice que denominaremos *flujo de conocimiento* (variable IFC) en las empresas manufactureras españolas para el año 2018.

3. 3. Estimación y resultados

La tabla adjunta muestra las estimaciones de los efectos, y por ende pesos, de las siete dimensiones analizadas sobre el índice de flujo de conocimiento. Siendo los estadísticos de bondad de ajuste del modelo superiores a 0,95 en todos los casos (NF, RFI, IPI, TLI y CFI) y siendo el valor del RMSEA inferior a 0,05, podemos afirmar que el modelo es estadísticamente válido.

Respecto a los efectos de cada una de las dimensiones, varias consideraciones.

Primero, todos los coeficientes son positivos y significativos (p-valor), como esperábamos. En este sentido, un indicador contrario a la literatura científica habría invalidado el modelo.

Segundo, atendiendo a los coeficientes estandarizados, las tecnologías de I+D (TECID) suponen la mayor ponderación en el índice, pronosticando una ventaja competitiva en términos de flujo de conocimiento en aquellas empresas que más implementen este tipo de tecnologías. Así, las empresas que cuentan con una dirección o comité de tecnología utilizan asesores para informarse sobre internet y realizan evaluaciones de perspectivas de cambios tecnológicos, tienen mayores valores en el índice de flujo de conocimiento.

Tercero, un segundo grupo de dimensiones con efectos significativos y elevados son las actividades de I+D (AID) y los recursos de I+D (RECID), con efectos estandarizados de 0,697 y 0,635, respectivamente. Efectivamente, las empresas con mayores actividades y recursos de I+D generan ventajas competitivas en términos del índice.

Cuarto, las innovaciones (INNOV) y las nuevas formas de organización del trabajo (NFOT) tienen efectos positivos y significativos con valores estandarizados superiores a 0,5 (concretamente, 0,591 y 0,51, respectivamente). Es decir, el reflejo de la I+D en términos de innovación y la organización del trabajo son aspectos relevantes que completan la explicación del índice.

Quinto, el uso de internet (USOINT) y las tecnologías de internet (TECINT) tienen efectos significativos y positivos, si bien significativamente menores que el resto de las dimensiones, con efectos estandarizados de 0,117 y 0,226, respectivamente.

Tabla 3. Estimación del flujo de conocimiento

Dimensión	Coficiente	Coficiente estandarizado	Error estándar	Ratio crítico	p-valor
INNOV	1)	0,591			
NFOT	0,639	0,512	0,026	24,498	0,000
USOINT	0,143	0,117	0,04	3,549	0,000
TECINT	0,343	0,226	0,051	6,758	0,000

Dimensión	Coficiente	Coficiente estandarizado	Error estándar	Ratio crítico	p-valor
AID	0,877	0,697	0,05	17,437	0,000
TECID	1,256	0,873	0,072	17,444	0,000
RECID	0,575	0,635	0,035	16,302	0,000

Bondad de ajuste del modelo. RMSEA 0,025; NFI: 0,98; RFI: 0,95; IFI: 0,984; TLI:0,959; CFI 0,984

Fuente: elaboración propia

4. 4. Regresión lineal

Con el propósito de contrastar la utilidad del índice creado, utilizaremos el logaritmo de la productividad horaria por trabajador (variable LPHN) como variable dependiente. De esta forma, el flujo de conocimiento, creado a partir de las estimaciones por SEM explicadas anteriormente, debería tener un efecto positivo y significativo en la explicación de la productividad. Asimismo, completaremos el modelo con dos variables adicionales únicamente a efectos de comprobar un modelo más completo de explicación de la productividad. En tal sentido, el objeto de esta no es una determinación completa de los causantes de la variable dependiente, sino la mera estimación del efecto en la productividad en un modelo algo más completo. Como variables adicionales se han considerado el logaritmo de las exportaciones (LVX) y el logaritmo del gasto total en formación por trabajador (LGFN).

La tabla adjunta recoge las estimaciones por MCO con el modelo completo comentado.

Tabla 4. Regresión lineal (MCO) para la explicación de la variable LPHN

	Coficientes no estandarizados	Error estándar	Coficientes estandarizados	t	Sig.
C	2,93	0,042		68,999	0
LVX	0,022	0,004	0,211	5,83	0
IFC	0,038	0,012	0,119	3,21	0,001
LGFN	0,035	0,01	0,126	3,431	0,001

R2: 13,1; R: 36,2. F: 43,511 (p-valor del modelo 0,000)

Fuente: elaboración propia

Tal y como puede observarse, los efectos de las tres variables independientes contempladas en el análisis de regresión lineal son positivos y significativos al 99 % de confianza, tal y como cabía esperar. Adicionalmente, el índice creado (variable IFC) genera un incremento de la productividad (medida en logaritmos) del 0,119 % por cada punto porcentual de aumento en el propio índice.

Sobre el resto de las variables, cabe indicar que la apertura exterior de las empresas es la que más explica la variabilidad de la productividad horaria por trabajador (coeficiente estandarizado de 0,211), mientras que el gasto en formación tiene un efecto estadísticamente similar al del índice de flujo de conocimiento (0,126 y 0,119, respectivamente).

Desde la perspectiva de la bondad de ajuste del modelo, cabe decir que el modelo es significativo al 99 % de confianza, siendo el estadístico F de 43,511 con un coeficiente de determinación de 13,1.

Conclusions

En este artículo, se ha presentado una propuesta metodológica para la creación de un índice de flujo de conocimiento empresarial en el ámbito de la empresa industrial manufacturera. La base de datos utilizada (la ESEE) proporciona

información muy completa sobre las empresas de los sectores manufactureros en España teniendo además la virtud de ser una encuesta de panel. Si bien el análisis aquí realizado supone un análisis econométrico *cross section*, la ESEE posibilita investigaciones futuras más profundas en el ámbito analizado con análisis dinámico y efectos intertemporales.

El índice flujo de conocimiento calculado se ha demostrado estadísticamente válido y ofrece efectos significativos y positivos de las siete dimensiones consideradas. En este sentido, los efectos cuantitativamente más importantes son los referidos a las dimensiones relacionadas con investigación y desarrollo, concretamente: tecnologías de I+D, recursos de I+D y actividades de I+D.

Un segundo grupo de efectos significativos y elevados, lo constituyen la innovación y los nuevos métodos de organización del trabajo.

La combinación de la inversión en investigación y desarrollo (I+D) y la adopción de nuevos métodos organizativos en el entorno laboral presenta una relación positiva y estadísticamente significativa con el aumento de la productividad empresarial. Estos hallazgos respaldan la necesidad de que las empresas adopten estrategias que integren tanto el desarrollo tecnológico como la innovación organizativa. En este contexto, el índice de flujo de conocimiento refuerza la relevancia de ambas dimensiones.

La convergencia de estos factores resulta esencial para facilitar la transición de las empresas hacia un nuevo paradigma industrial caracterizado por la influencia predominante de la inteligencia artificial (IA), especialmente en su variante generativa. Este enfoque permite no solo responder a las exigencias del entorno económico actual, sino también potenciar la competitividad mediante el aprovechamiento de tecnologías emergentes y la optimización de los métodos laborales.

Por último, las dimensiones relativas a internet (usos de internet y tecnologías de internet) tienen efectos significativos pero moderados, sugiriendo que supone poca ventaja competitiva su desarrollo, pero a su vez necesario, consecuencia lógica de la evolución tecnológica de las últimas décadas en la industria.

La validez del índice estimado se ha contrastado finalmente con la estimación del efecto en la productividad horaria por trabajador, que junto con la apertura en términos de exportaciones y el gasto en formación por trabajador constituyen un modelo válido, desde un punto de vista estadístico, en la explicación de tal variable.

De cara a futuras investigaciones, la construcción del índice sugiere la posibilidad de su ampliación a una estimación de la productividad en un modelo más completo y en términos intertemporales, lo cual es posible con la base de datos utilizada.

Referencias bibliográficas

- DÍAZ-CHAO, Ángel; MONTES, José Luis (2014). «A Measurement for the Evolution of Renewable Energies: The Case of the European Union». *Energy and Environment*, vol. 25, n.º 2, págs. 305-323. DOI: <https://doi.org/10.1260/0958-305X.25.2.305>
- JÖRESKOG, Karl G.; SÖRBORN, Dag (1984). *LISREL-VI user's guide*. Mooresville, IN: Scientific Software.
- MELSER, Daniel; Hill, Robert (eds.) (2007). «Methods for Constructing Spatial Cost of Living Indexes». *Official Statistics Research Series*, vol. 1.
- OCDE (2008). *Handbook of Constructing Composite Indicators. Methodology and User Guide*. OCDE.
- EUROSTAT (1996). *Sources and methods construction price indices*. Eurostat.

Cita recomendada: DÍAZ-CHAO, Ángel. «Estimación de la productividad a partir de una aproximación a la cuantificación del flujo de conocimiento en la industria española». *Oikonomics* [en línea]. Noviembre 2024, n.º 23. ISSN 2330-9546. DOI: <https://doi.org/10.7238/o.n23.2419>



Ángel Díaz-Chao

angel.diaz@urjc.es

Universidad Rey Juan Carlos

Doctor en Economía por la Universidad Complutense de Madrid, máster en Economía Aplicada por la Universitat Autònoma de Barcelona, licenciado en Economía por la Universidad Carlos III y licenciado en Administración y Dirección de Empresas por la universidad Carlos III. Es profesor titular de la Universidad Rey Juan Carlos y ha sido profesor de la Universitat Autònoma de Barcelona y colaborador e investigador en la UOC. Ha sido jefe de investigación de la Fundación SEPI y director general de dicha entidad. Entre sus líneas de investigación destaca el desarrollo de modelos econométricos y su aplicación al estudio de la productividad y la digitación, entre otros, habiendo publicado numerosos artículos científicos en revistas nacionales e internacionales. Actualmente es también miembro del grupo de investigación i2TIC.

Los textos publicados en esta revista están sujetos –si no se indica lo contrario– a una licencia de Reconocimiento 4.0 Internacional de Creative Commons. Puede copiarlos, distribuirlos, comunicarlos públicamente, hacer obras derivadas siempre que reconozca los créditos de las obras (autoría, nombre de la revista, institución editora) de la manera especificada por los autores o por la revista. La licencia completa se puede consultar en https://creativecommons.org/licenses/by/4.0/deed.es_ES.



ODS

