



FACULTAD DE
CIENCIAS ECONÓMICAS
Y DE ADMINISTRACIÓN

POSGRADOS



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

TICs, Productividad y Exportaciones: Un estudio a nivel de empresas uruguayas.

Daniel Méndez*

Tutora: Dra. Adriana Peluffo[†]. Doctora en Economía Aplicada, Universiteit Antwerpen. Profesora Agregada.

UNIVERSIDAD DE LA REPÚBLICA

FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN

**TESIS PARA OBTENER EL TÍTULO DE
MAGÍSTER EN ECONOMÍA**

Montevideo

URUGUAY

23 de junio de 2025

* Correo electrónico: daniel.mendez@fcea.edu.uy; daniknd14@gmail.com

[†] Correo electrónico: adriana.peluffo@fcea.edu.uy

Página de Aprobación

El tribunal docente integrado por los abajo firmantes aprueba el Trabajo Final:

Título

.....
.....
.....

Autor

.....
.....

Tutora

.....
.....

Posgrado

.....

Puntaje

.....

Tribunal

Profesor (nombre y firma)

Profesor (nombre y firma)

Profesor (nombre y firma)

FECHA

Resumen

Utilizando datos de la encuesta EAAE se obtiene un índice que mide la adopción de Tecnología de la Información y de la Comunicación (TICs) por parte de las firmas uruguayas para el período 2012-2019. A continuación, se obtiene una estimación robusta de la Productividad Total de los Factores (PTF) para estas mismas firmas. Finalmente, se construye y se estima un Modelo de Ecuaciones Simultáneas (SEMs) de tres ecuaciones que da cuenta de las interrelaciones entre las TICs, la PTF y las Exportaciones de las firmas, controlando por diferentes variables sugeridas por la literatura, además de efectos industriales y temporales. El objetivo principal es analizar, empíricamente y en un único modelo, el impacto de la adopción de TICs sobre la PTF y cómo ambas variables impactan sobre las exportaciones de las empresas uruguayas durante el período considerado. Los resultados del SEMs evidencian una relación positiva y significativa entre la adopción de tecnologías digitales, la productividad y el desempeño exportador. Se encuentra que mayores niveles de adopción TICs mejoran la eficiencia productiva, y que ambas dimensiones, tecnológica como productiva, mejoran la competitividad internacional de las firmas.

Palabras clave: TICs, Productividad, Exportaciones, Simultaneidad, Empresas.

JEL: F1, D24, O0, O5, O54

Abstract

Using data from the EAAE survey, an index measuring the adoption of information and communication technology (ICT) by Uruguayan firms from 2012 to 2019 is obtained. Next, a robust estimate of total factor productivity (TFP) is obtained for the same firms. Finally, a three-equation simultaneous equation model (SEMs) is constructed and estimated, accounting for the interrelationships between ICT, TFP, and firms' exports. This model controls for different variables suggested by the literature, as well as industrial and time effects. The main objective is to empirically analyze, in a single model, the impact of ICT adoption on TFP and how these two variables affect the exports of Uruguayan companies during this period. The SEMs results show a positive and significant relationship between digital technology adoption, productivity, and export performance. Higher levels of ICT adoption are found to improve productive efficiency, and both technological and productive dimensions are found to improve the international competitiveness of firms.

Keywords: ICTs, Productivity, Exports, Simultaneity, Firms.

JEL: F1, D24, O0, O5, O54

Índice

Acrónimos	6
1 Introducción	7
2 Marco teórico: TICs, PTF y Exportaciones	9
3 Antecedentes	13
3.1 Internacionales	13
3.2 Regionales y Nacionales	15
4 Datos y Estrategia Empírica	16
4.1 Fuente de información	16
4.2 TICs: una medida resumen	17
4.3 PTF: una estimación robusta	19
4.4 SEMs: metodología econométrica	23
5 Resultados	29
5.1 TICs: un análisis descriptivo	29
5.2 PTF: un análisis descriptivo	34
5.3 EXP: un análisis descriptivo	38
5.4 SEMs: resultados	40
6 Conclusiones	45
7 Anexo	50
7.1 Antecedentes empíricos	50
7.2 TICs: resolución de ACP	51
7.3 TICs: valores faltantes y un algoritmo de imputación	54

7.4	TICs: contribución de variables en ACP	56
7.5	PTF: breve mirada a Olley y Pakes (1992)	57
7.6	PTF: breve mirada a Wooldridge (2009)	61
7.7	PTF: breve mirada a Akerberg et al. (2015)	64
7.8	SEMs: verificación de la condición de rango	67
7.9	SEMs: validez de los instrumentos	69
7.10	SEMs: estimación por GMM-3SLS	79

Acrónimos

Acrónimos	Definiciones
INE	Instituto Nacional de Estadística.
EAAE	Encuesta Anual de Actividad Económica. INE.
CIIU	Clasificación Industrial Internacional Uniforme.
IPI	Índice de Precios Implícito. INE.
PYMEs	Pequeñas y Medianas Empresas.
TICs	Tecnologías de la Información y de la Comunicación.
ACP	Análisis de Componente Principal.
PC1	Primer Componente Principal.
PC2	Segundo Componente Principal.
EXP	Exportaciones.
PTF	Productividad Total de los Factores.
OP	PTF estimada por el método de Olley y Pakes (1992).
LP	PTF estimada por el método de Levinsohn y Petrin (2003).
WRD	PTF estimada por el método de Wooldridge (2009).
ACF	PTF estimada por el método de Akerberg et al. (2015).
SEMs	Modelo de Ecuaciones Simultáneas.
GMM	Método Generalizado de los Momentos.
GMM-S2SLS	Estimador GMM del Sistema de Mínimos Cuadrados en Dos Etapas.
GMM-OWM	Estimador GMM con Matriz de Ponderación Óptima.
GMM-3SLS	Estimador GMM de Mínimos Cuadrados en Tres Etapas.

1 Introducción

El objetivo de este trabajo es analizar el impacto de las Tecnologías de la Información y de la Comunicación (TICs) sobre la productividad y las exportaciones de las empresas uruguayas. Diferentes campos de conocimiento, principalmente en la disciplina económica, han intentado dar respuesta a al menos una de las relaciones que se pretenderán analizar, a saber, la relación entre TICs y productividad, la relación entre TICs y exportaciones o la relación entre productividad y exportaciones, pero son escasos los trabajos que consideran las tres relaciones simultáneamente¹.

Existe una extensa literatura que se ha encargado de analizar el concepto de Productividad Total de los Factores (PTF) y de sus determinantes, entre los cuales se ha identificado a las TICs como uno muy relevante (Syverson, 2011). Por otra parte, la literatura de los denominados “nuevos nuevos” modelos teóricos de Comercio Internacional, vincula la productividad con la probabilidad de exportar de la empresa (Eaton y Kortum, 2002; Melitz, 2003). En ese sentido, estos modelos suponen cierta dirección que va desde la productividad hacia las exportaciones de las empresas. A su vez, la literatura de los negocios internacionales explora el vínculo entre TICs y las exportaciones. Esta literatura, que en general se basa en trabajos de tipo cualitativos o trabajos cuantitativos con pocas observaciones, utiliza marcos teóricos explorando la relación entre ambos conceptos, argumentando que las TICs disminuyen las distancias, reducen los costos de información y facilitan tanto la integración como la inserción internacional (Mol y Koppius, 2002; Alcácer et al., 2016; Prashantham, 2005; Hervé et al., 2020).

Este trabajo busca responder cuál es el impacto de la incorporación de TICs sobre la productividad, y el de la productividad sobre las exportaciones. En lugar de suponer un sistema recursivo, se opta por un sistema de ecuaciones simultáneas que pueda contemplar un conjunto más amplio de relaciones y que permita el análisis simultáneo de los efectos entre TICs y productividad. De esta manera, el objetivo general será analizar el impacto de las TICs sobre la productividad, evitando el sesgo por simultaneidad, al mismo tiempo que se busca responder en qué medida la productividad puede impactar sobre las exportaciones de las empresas. Asimismo, esto nos permitirá evaluar si el impacto de las TICs sobre las exportaciones es capturado totalmente por medio de la productividad o si también existen efectos directos tales como la reducción en los costos del comercio que impactan sobre las exportaciones. En este sentido, en el marco teórico se presentarán los mecanismos que podrían estar afectando la vinculación entre TICs, productividad y exportaciones.

En cuanto a la relevancia del presente trabajo se pretende obtener estimaciones de estas relaciones para un conjunto representativo de empresas uruguayas para el periodo 2012-2019. Siendo Uruguay un país pequeño y abierto, el desarrollo del sector productivo y de servicios adquiere fundamental relevancia para su crecimiento². En particular, es de sumo interés cuantificar el uso e importancia de las TICs y su impacto sobre la productividad y exportaciones. En este sentido, las exportaciones son de suma importancia para pequeños países emergentes dado que permiten alcanzar economías de escala, generar empleo y divisas, promoviendo el crecimiento y posibilitando un mayor bienestar para la población. Este trabajo podría ser de utilidad al brindar

¹ En la sección (4.2) se presenta una definición formal del concepto de TICs.

² En este trabajo se incluyen los sectores pertenecientes a las secciones *B* hasta la *S* de la CIIU revisión 4, con la excepción del sector *F* de la Construcción. No incluye la sección *A* correspondiente a la producción agropecuaria, forestación y pesca, en correspondencia con la fuente de datos que se dispone.

información relevante para el diseño e implementación de políticas industriales con un mayor conocimiento sobre los impactos de las TICs sobre la productividad y exportaciones. En la medida que existan políticas de competencia permitirían que esta mayor productividad se traduzca en menores costos y precios para los consumidores que, a su vez, redundarían en menores costos para la población.³

De esta forma, la relevancia económica reside en la posibilidad de detectar márgenes de mejora a través de políticas orientadas a impulsar la productividad de las empresas, por lo que debería estar acompañado de políticas de competencia adecuadas. En este sentido, estructuras de mercado oligopólicas tienden a fijar precios por encima de los costos productivos (mark-up). Estas conductas anticompetitivas son más frecuentes en países pequeños como Uruguay. En esta misma línea, si la mejora en productividad se tradujera en que las empresas uruguayas sean de menor costo frente a empresas de otros países productoras del mismo bien o servicio, en alusión al modelo de Eaton y Kortum (2002), debería contribuir a que las empresas puedan ser más competitivas y logren acceder a mayores niveles de exportación de sus productos. Por otra parte, los aumentos de productividad podrían estar asociados a mejoras en la calidad de los bienes y servicios, agregando valor a los mismos. El cumplimiento con los estándares internacionales facilitaría acceder a nuevos mercados. Esto último, en línea con efectos positivos de la diversificación de mercados, podría tener efectos positivos sobre el crecimiento económico.

Se destaca que no hay trabajos empíricos para Uruguay que analicen el impacto de las TICs a nivel de firma sobre la productividad total de los factores y/o las exportaciones. Más aún, son inexistentes los trabajos que consideren las relaciones entre las tres variables en forma simultánea y para pequeños países emergentes. Dada esta situación, este trabajo pretende aportar evidencia en un área de la disciplina económica que no ha recibido atención a nivel nacional y que se entiende que puede ser de utilidad para la implementación de futuras políticas públicas.

El artículo está organizado de la siguiente manera. En la sección (2) se presenta el marco teórico, donde se intenta vincular las tres variables relevantes mencionadas y que será el marco de referencia sobre el cual aplicaremos nuestra estrategia empírica. En la sección (3), se hace un breve repaso de antecedentes en la literatura relacionada al uso de TICs por parte de las empresas a nivel internacional, regional y nacional. En la sección (4) se presentan los datos a utilizar, se estima un índice sintético que resume la utilización de las TICs, al tiempo que se realiza una estimación robusta de la productividad total de los factores. Por otra parte, se presenta la metodología econométrica aplicada al conjunto de datos de panel para el periodo 2012-2019. En la sección (5) se presentan, en primer lugar, resultados descriptivos de nuestras variables endógenas TICs, Productividad y Exportaciones y, finalmente, los resultados de estimar un sistema de ecuaciones para las variables de interés. En la sección (6) se discuten los resultados en función de nuestros objetivos e hipótesis y se presentan adicionalmente las limitaciones y agenda de investigación futuras y se concluye.

³ Esta es otra de las predicciones teóricas del modelo desarrollado en Melitz (2003), que surge de resolver el problema de las empresas. Cabe mencionar que demostrar este resultado en la práctica requeriría de datos de precios a nivel de empresas, que escapa al objetivo del presente trabajo.

2 Marco teórico: TICs, PTF y Exportaciones

El siguiente diagrama resume las relaciones entre las variables de interés que se analizarán en este trabajo:

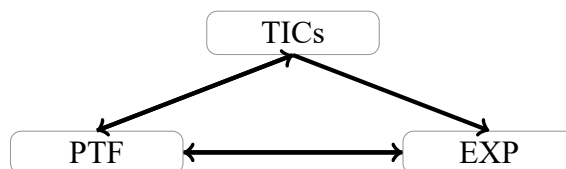


Figura 1: Marco teórico.

Fuente: Elaboración propia.

En cuanto a la relación entre productividad y las exportaciones, si bien existe una serie de trabajos que estudian la relación causal que va desde exportaciones hacia productividad a través del mecanismo de *learning by exporting*, en el presente trabajo no vamos a estudiar esta relación dado que queda por fuera de nuestros objetivos. Sin embargo optamos por incluir en el diagrama la existencia de esta relación en virtud de la literatura que ha estudiado dicha relación. Con esta salvedad, se puede observar que el sentido de las flechas refleja la dirección causal de la relación entre las variables endógenas de nuestro modelo, que se corresponden 1:1 con el modelo empírico que se implementa en la sección (5.4) del presente trabajo.

En las últimas décadas se ha hecho evidente la expansión de las TICs, no sólo a nivel del propio sector de las TICs si no también hacia todos los sectores de la economía. Esto implica que, a nivel general, la economía en su conjunto fue introduciendo y profundizando el uso y la intensidad en las TICs, aunque dicha adopción es heterogénea entre los diferentes sectores. Syverson (2011) identifica a las TICs como una de las denominadas “palancas” de la productividad⁴. El autor realiza una exhaustiva revisión bibliográfica donde identifica hallazgos por parte de numerosos trabajos empíricos que dan cuenta de una relación positiva entre las TICs y el nivel de la productividad total de los factores y, entre las TICs y la varianza de esta productividad. Esto sugiere la existencia de un margen a través del cual pueden existir ganancias de productividad basadas en la adopción de TICs, a su vez, el hecho de que las TICs se relacionen con una mayor variabilidad de productividad nos indica que el proceso de adopción de TICs puede ampliar la brecha entre las empresas cercanas a la frontera tecnológica y las empresas más lejanas a dicha frontera.

De hecho, en Acemoglu et al. (2006) se presenta esta última idea en el contexto de un modelo de crecimiento endógeno que relaciona la distancia a la frontera tecnológica con cambios organizacionales. En particular, relacionando el estado de desarrollo de una economía en función de características como la innovación y la imitación, encuentran que las economías que están en una etapa inicial de desarrollo crecerán más rápidamente si eligen una estrategia de crecimiento basada en la inversión y adopción de tecnología existente, mientras que las economías que están más cerca de la frontera crecerán más rápidamente si adoptan una estrategia basada en la innovación tecnológica⁵. Los autores advierten de la posibilidad de una trampa de no convergencia para las economías que, en sus estados de menor desarrollo lleven la estrategia de adopción

⁴ Ver Syverson (2011), página 335.

⁵ Previo al modelo, utilizando datos a nivel de países para el periodo 1965-1995, los autores estiman un modelo de Variables Instrumentales con Efectos Fijos (VI-EF) donde justifican el supuesto de barreras a la competencia

demasiado lejos, esto es, con muy bajos niveles de competencia, quedando en una situación de crecimiento de equilibrio sin convergencia a la frontera. En este sentido, se podría replicar esta idea junto con el hecho de que las TICs y la varianza de la productividad tengan una relación positiva: empresas muy rezagadas en relación con la frontera tecnológica podrían permanecer en su ciclo de vida sin converger a la misma. Esto destaca la importancia de las TICs en tanto se las puede considerar un factor que promueve la productividad y deja entrever el margen de mejora que subyace a la adopción de tecnología por parte de las empresas que están lejos respecto de la frontera tecnológica.

Por otro lado, en relación con las exportaciones. La literatura de Comercio Internacional, y particularmente los llamados “nuevos nuevos” modelos de Comercio, nos proveen un marco teórico a partir del cual podemos vincular los conceptos de productividad y exportaciones (Eaton y Kortum, 2002; Melitz, 2003).

Melitz (2003) elabora un modelo donde cada firma tiene una productividad distinta y produce un bien distinto, con lo cual se puede indexar las empresas por su nivel de productividad. Las empresas resuelven el problema en competencia monopolística, y dada la existencia de costos fijos de producir, existirá un umbral inferior de productividad a partir del cual las empresas que estén por debajo de dicha cota se verán obligadas a cerrar y retirarse del mercado, mientras que, con la presencia de costos asimétricos de exportar, existirán N umbrales de exportación a cada uno de los países del modelo. Una de las predicciones teóricas del modelo es que, a mayor nivel de productividad de una firma, mayor será la probabilidad de exportar. En relación con el diagrama dibujado al comienzo de esta sección, esto nos marca una relación que va desde la productividad hacia las exportaciones.

Eaton y Kortum (2002), a diferencia del enfoque de Melitz (2003), elaboran un modelo donde cada país puede producir un continuo de bienes, con lo cual puede pasar que más de un país produzcan el mismo bien. Dado que cada país produce los bienes con distinta eficiencia, la optimización buscará que el país J compre determinado bien al productor de mínimo costo para producir ese bien en particular, entre el conjunto de N países. En este modelo, que la firma tenga un nivel de productividad más alto, implica que tendría el mínimo costo, y por lo tanto su probabilidad de exportar aumentaría. Estos modelos pioneros se han extendido teóricamente para incorporar más complejidad, por ejemplo, agregando la matriz de insumo-producto en el análisis y así permite estudiar el comercio intersectorial dentro de cada país del modelo (Caliendo y Parro, 2015).

En cuanto a las exportaciones y su relación con las TICs, los siguientes artículos de la literatura de los negocios internacionales, marketing internacional, la ciencia y tecnología destacan la importancia que fueron tomando las redes informales en la industria, sobre todo en marcos de incertidumbre para la toma de decisiones y cómo esto ayudaba en la inserción internacional de las empresas (Welch, 1996)⁶. Analizando a las empresas transnacionales, se ha argumentado que las TICs permiten una reducción de las distancias y la capacidad de obtener economías de escala

y distancia a la frontera tecnológica. Para países con altas barreras se encuentra una correlación fuerte y negativa entre crecimiento y proximidad a la frontera, implicando que se desempeñan mejor (convergen) lejos de la frontera que cerca de la misma. Para países con bajas barreras la relación también es negativa pero débil, indicando que su desempeño es similar tanto lejos como cerca de la frontera tecnológica.

⁶ Por *inserción internacional* se entienden diferentes formas de internacionalización como exportaciones, importaciones, inversión extranjera directa, franquicias, licencias, entre otras, si bien la más extendida y generalizada son las exportaciones.

como forma de contribuir a las adquisiciones y fusiones de las empresas (Miozzo y Soete, 2001). Otros autores argumentan que las TICs ayudan a superar limitaciones de la contratación global, alterando las limitaciones geográficas y relacionales al permitir una mayor inserción internacional, así como una mejor integración funcional (Mol y Koppius, 2002). Las TICs impactan en la fragmentación de la producción y las estrategias de aprovisionamiento global de las empresas transnacionales, contribuyen a mantener la conectividad y la relación entre distintos campos de conocimiento cuando existe dispersión geográfica (Alcácer et al., 2016).

En cuanto a las pequeñas y medianas empresas (PYMEs), a diferencia de las empresas transnacionales que pueden entrar en los mercados internacionales utilizando muchos de los modos tradicionales, como las exportaciones, franquicias, licencias, contratos de gestión o inversión directa, las PYMEs, debido a las limitaciones de recursos, pueden recurrir a Internet, lo que reduce en parte la dificultad de transferir información a través de las fronteras geográficas (Todd y Javalgi, 2007). Otros trabajos también destacan el rol cada vez más importante de la Internet en la inserción internacional, a través del vínculo entre información y conocimiento (Prashantham, 2005). Las TICs aumentan la propensión a exportar de las PYMEs como estrategia de supervivencia o como estrategia de crecimiento. Sin embargo, no todas las empresas poseen las capacidades o competencias empresariales necesarias para internacionalizarse rápidamente (Y.-Y. Lee et al., 2019). La gestión a distancia permite generar ingresos sin necesidad de grandes inversiones, interactuar directamente con los clientes y las empresas para comprender mejor sus necesidades y personalizar sus servicios y productos. La distancia y la localización se manifiestan principalmente por la desmaterialización de las fronteras entre los países y la aceleración de las operaciones de inserción internacional (Hervé et al., 2020). En general, esta literatura analiza las PYMEs y encuentra que las TICs pueden ayudar en el proceso de inserción internacional. Ello nos invita a pensar en una relación que va desde TICs hacia exportaciones, tal como se muestra en el diagrama, donde los mecanismos que operan apuntan a una reducción de distancias tanto geográficas como relacionales, una reducción de los costos de la información y de obtención de conocimiento.

Retomando la relación entre TICs y productividad. La gran mayoría de la literatura ha analizado la causalidad que va desde TICs a productividad, pero se tienen motivos para creer que las empresas más productivas posiblemente sean capaces de obtener mayores beneficios y esto traducirse en mayor capital financiero disponible. De esta manera, estas empresas pueden estar dispuestas a ampliar su capacidad tecnológica en busca de permanecer cerca o aproximarse a la frontera, con lo cual este argumento iría en el sentido causal desde la productividad hacia las TICs. Si esto ocurre, y considerando lo discutido previamente sobre la causalidad, entonces la productividad y las TICs se determinan simultáneamente. Por lo tanto, estimar por Mínimos Cuadrados Ordinarios (MCO) una especificación para cualquiera de estas dos variables nos arrojaría un estimador inconsistente. En base a lo expuesto, se ha propuesto un marco teórico que relaciona las TICs, la PTF y las exportaciones de las empresas. Esto se ve reflejado en la Figura (1).

Las dos hipótesis principales son las siguientes.

- H1: Un mayor uso de TICs puede tener efectos positivos sobre las exportaciones debido a una reducción de los costos de comercio.
- H2: Un mayor uso de TICs incrementaría la productividad de las empresas, y esto conduciría a un incremento en las exportaciones.

La estrategia empírica utilizada nos permitirá sortear sesgo de simultaneidad entre TICs y productividad, y analizar estas relaciones en forma simultánea.

3 Antecedentes

3.1 Internacionales

Andrews et al. (2018) analizan qué factores afectan las tasas de adopción digital a nivel industrial, en particular estudian el efecto de las *capacidades e incentivos* de las empresas. Del lado de las capacidades se destaca que la baja calidad de los directivos, la falta de competencias de los trabajadores en TICs y la escasa adecuación de los trabajadores a los puestos de trabajo frenan la adopción de la tecnología digital y, por tanto, su tasa de difusión. Por otra parte, las políticas que afectan a los incentivos del mercado son importantes, a mayor competencia, flexibilidad laboral y disponibilidad del capital, mayor adopción. En cuanto al presente trabajo, esto nos puede dar una idea de que variables exógenas utilizar en la especificación para las TICs en la estrategia empírica a implementarse más adelante.

Gal et al. (2019) evalúan cómo la adopción de una serie de tecnologías digitales afecta la PTF de las empresas⁷. Encuentran que la adopción digital en una industria está asociada a aumentos de PTF a nivel de firma y que esta relación es más fuerte para el sector manufacturero respecto del sector servicios, con la única excepción para el caso de la tecnología de Banda Ancha de Alta Velocidad. También encuentran que esta relación es más fuerte para las empresas que tienen mayor intensidad en tareas rutinarias. Por otra parte, esta relación es más débil en las empresas con escasez de personal calificado, lo que apunta a la existencia de complementariedad entre el capital humano calificado y las TICs. Este último aspecto también fue identificado en el trabajo hecho por Syverson (2011). Gal et al. (2019) trabajan con TICs a nivel industrial y con PTF a nivel de empresas y encuentran que las TICs aumentan la PTF.

DeStefano et al. (2023) analizan la existencia de un efecto causal entre la adopción de tecnología de Banda Ancha ADSL y la productividad laboral. Los autores explotan la discontinuidad espacial y temporal en la implementación de banda ancha ADSL en el Reino Unido para capturar efectos causales sobre las empresas. Entre sus resultados no encuentran efectos significativos sobre el proxy utilizado de la productividad. Cabe mencionar respecto a este punto que los autores utilizaron como medida al cociente de las ventas sobre el empleo, y estas ventas contienen valor producido por otras empresas, es decir, consumo intermedio (bienes y servicios producidos por otras empresas) con lo cual no es la medida más adecuada para ver lo que efectivamente produjo la empresa en relación con sus trabajadores. En el presente documento trabajaremos con la productividad total de los factores (PTF).

La innovación es otra de las “*palancas*” que impulsan la productividad, reconocidas por Syverson (2011). Al respecto, Bartelsman et al. (2017) ponen el foco en el proceso de innovación a nivel de empresa y cómo este proceso impacta a nivel agregado en la productividad del trabajo. Usan un marco teórico y una metodología econométrica propuesta originalmente por Crepon et al. (1998) para estudiar la relación entre innovación y productividad. El resultado principal es que el efecto de la innovación sobre la productividad media no es significativa mientras que el efecto de esta sobre la productividad agregada (ponderada por el empleo) es significativo. Esto indica que las empresas más productivas que utilizan intensivamente la Banda Ancha son más grandes. Con respecto al presente trabajo, cabe destacar que se utilizará una medida para las

⁷ Banda Ancha de Alta Velocidad, Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), Cloud Computing (CC).

TICs en lugar de analizar la innovación y, por otra parte, no se utilizará la productividad laboral sino la productividad total de los factores que es una medida multifactorial de la productividad. Otra diferencia, es que el método econométrico utilizado por ellos se basa en un sistema de tipo recursivo.

En cuanto a la literatura sobre TICs y la inserción internacional, Kyläheiko et al. (2011) en un estudio para empresas finlandesas encuentran una correlación positiva entre la capacidad tecnológica, usando una medida compuesta subjetiva, y una medida para la inserción internacional dada por la intensidad exportadora. Gbadegeshin et al. (2019) en un estudio para PYMEs nigerianas encuentran que estas utilizaban herramientas TICs para su inserción internacional, como sitios web y plataformas de medios sociales. Dutot et al. (2014) analizando PYMEs canadienses de fabricación y servicios que se han internacionalizado encuentran que a mayor alineamiento entre los requisitos de información y capacidades de Tecnologías de la Información (TI) mayor rendimiento obtenido de la inserción internacional. Tomski (2015) analiza PYMEs polacas con el objetivo de explorar la influencia de la utilización de las redes por parte del empresario en la fase del proceso de inserción internacional. Este autor encuentra una relación positiva entre tener contactos que tienen relación con el extranjero y el grado de inserción internacional en la empresa. Por otra parte, Lecerf y Omrani (2020) examinan el impacto de las TICs a través de las funciones de planificación de recursos empresariales (ERP) y gestión electrónica de las relaciones con los clientes (e-CRM), en la inserción internacional de las PYMEs alemanas. Los resultados muestran que las TICs no influyen directamente en la inserción internacional.

Centrado en las empresas del tipo Born Global⁸, Vasilchenko y Morrish (2011) investigan el papel de las redes en la inserción internacional de las empresas con altos niveles tecnológicos. Los autores realizan un estudio de casos para cuatro empresas del sector TICs neozelandesas. Encuentran que las redes sociales desarrolladas a partir de conexiones personales permiten explorar oportunidades de inserción internacional, a su vez, estas redes sociales pueden convertirse en redes empresariales formales y estas redes empresariales son un enfoque eficaz para superar las limitaciones de recursos y facilitar la explotación de oportunidades. Por su parte, el objetivo en Sekliuckiene y Maciulskaitė (2013) es analizar qué decisiones estratégicas toman las empresas Born Global lituanas que desean internacionalizar sus operaciones. Estos autores concluyen que la inserción internacional de estas empresas está determinada por factores externos (desarrollo de las TICs, asociación y creación de redes) e internos (conocimientos y habilidades de los directivos, innovación, servicio único) que se convierten en catalizadores para la inserción internacional.

Estos trabajos, en general enmarcados en la literatura de los Negocios Internacionales, utilizan relativamente pocas observaciones y las medidas que utilizan suelen ser tomadas de encuestas que ellos mismos diseñan con métodos empíricos que dan cuenta de correlaciones entre las variables, pero no causalidades. También se destaca el hecho de que la mayoría de estos trabajos se hacen para PYMEs. En general, encuentran una correlación positiva entre TICs y alguna forma de inserción internacional, como podría ser el caso de las exportaciones, que es la forma de inserción internacional considerada en este trabajo.

⁸ En comparación con las organizaciones tradicionales, las “Born Global” se internacionalizan rápidamente. Normalmente, entran en el mercado extranjero en los tres primeros años de su establecimiento nacional. Ver en Sekliuckiene y Maciulskaitė (2013), páginas 19-20, para un mayor detalle de este concepto.

3.2 Regionales y Nacionales

Molina et al. (2013) analizan PYMEs argentinas de la industria manufacturera. Estos autores construyen dos indicadores que recogen, por un lado, el nivel de informatización por áreas a nivel de la firma, y por otro, un patrón de incorporación de TICs. Encuentran que los indicadores de TICs no son significativos para explicar la productividad laboral. Charlo (2011) tiene como objetivo analizar los efectos de las TICs sobre la productividad laboral de las empresas manufactureras en Uruguay. La autora encuentra que las TICs, medidas operativamente como *inversión en capital TICs*, se correlaciona positivamente con la productividad del trabajo, sin embargo, cabe mencionar que la correlación encontrada no es de una magnitud considerable.

Plottier et al. (2013) analiza cómo se procesa la adopción de TICs por parte de las empresas en el Uruguay. Esta investigadora lleva adelante un trabajo cualitativo para el cual realizó entrevistas a un grupo reducido de empresas de cuatro sectores diferentes: biotecnología, industria farmacéutica, sector forestal-maderero e industria naval. Entre los resultados identifica:

1. brecha digital entre pequeñas y grandes empresas
2. desconocimiento de soluciones tecnológicas específicas para el sector
3. dificultades en la coordinación entre la oferta local de TICs y la demanda
4. la falta de capacitación en el personal y los empresarios
5. espacios para mejorar la relación con organismos estatales a través del uso de TICs.

De forma similar a Andrews et al. (2018) realiza un trabajo en el cual sugiere variables exógenas para la especificación de las TICs en nuestro modelo empírico. En el Anexo (7.1) se presenta en formato de tabla un resumen de los antecedentes empíricos.

4 Datos y Estrategia Empírica

4.1 Fuente de información

El presente trabajo se basa en datos de la Encuesta Anual de Actividad Económica (EAAE) provista por el Instituto Nacional de Estadística (INE). Su diseño muestral combina aleatorización, estratificación y conglomerados. Las empresas se estratifican según su tamaño y su clase de actividad económica principal. Las grandes empresas son seleccionadas de forma forzosa, mientras que las pequeñas y medianas se eligen aleatoriamente. La EAAE sigue un esquema de panel rotativo incorporando los nacimientos de nuevas empresas. Las empresas seleccionadas aleatoriamente permanecen en la muestra durante al menos tres años, tras lo cual son reemplazadas por otras del mismo estrato ⁹.

Son elegibles empresas que tienen personal ocupado dependiente promedio en el año mayor o igual a 10 y/o presenta un promedio mensual de ventas mayor a 128.3 millones de pesos uruguayos correspondientes al año 2016, cifra que se ajusta año a año. La clase de actividad económica principal de la empresa, siguiendo la Clasificación Industrial Internacional Uniforme (CIIU) revisión 4, se presenta en la siguiente tabla:

CIIU revisión 4 - SEC

- B. Explotación de minas y canteras
- C. Industrias Manufactureras
- D. Suministro de electricidad, gas, vapor y aire acondicionado
- E. Suministro de agua; alcantarillado, gestión de desechos y actividades de saneamiento
- G. Comercio al por mayor y al por menor; reparación de vehículos de motor y de las motocicletas
- H. Transporte y almacenamiento
- I. Alojamiento y servicio de comida
- J. Información y comunicación
- K. Actividades financieras y de seguros
- L. Actividades inmobiliarias
- M. Actividades profesionales, científicas y técnicas
- N. Actividades administrativas y servicios de apoyo
- O. Administración pública y defensa; planes de seguridad social de afiliación obligatoria
- P. Enseñanza
- Q. Servicios sociales y relacionados con la Salud Humana
- R. Artes, entretenimiento y recreación
- S. Otras actividades de servicios

Cuadro 2: Secciones de Actividad

Fuente: Elaboración propia en base al Diseño muestral de la Encuesta Anual de Actividad Económica (EAAE).

Cada clase de actividad económica es estratificada. Estrato 1: comprende a las empresas con personal ocupado entre 10 y 19; Estrato 2: comprende a las empresas con personal ocupado entre 20 y 49; Estrato 3: comprende a las empresas con personal ocupado mayor o igual a 50 y/o promedio mensual de ventas por un monto mayor a 128,3 millones de pesos uruguayos correspondiente al año 2016. El estrato de tamaño 3 es de inclusión forzosa a priori, es decir,

⁹ <https://www.gub.uy/instituto-nacional-estadistica/Encuesta-Anual-Actividad-Económica>

todas las empresas dentro del estrato son encuestadas. Para el presente trabajo, se cuenta con datos del período 2012-2019, y se elabora un panel no balanceado. Todas las variables de gastos e ingresos han sido deflactadas por el Índice de Precios Implícito (IPI) construido por sector industrial, con año base 2012¹⁰.

4.2 TICs: una medida resumen

Existen distintas definiciones sobre las Tecnologías de la Información y la Comunicación (TICs). Si nos basamos en la definición adoptada en OECD (2009), se entiende por TICs a los ordenadores y equipos periféricos, equipos de comunicación, equipos electrónicos de consumo, componentes y bienes TICs diversos, servicios de fabricación de equipos TICs, servicios de licencias y software empresarial y de productividad, consultoría y servicios de tecnologías de la información, servicios de telecomunicaciones, servicios de leasing o alquiler de equipos y de otros servicios TICs.

Para capturar y lograr medir el concepto de TICs se parte del hecho de que en la encuesta EAAE se cuenta con un módulo exclusivo de preguntas asociadas a las TICs. La mayoría de estas preguntas se corresponden a respuestas binarias, y solamente unas pocas se pueden tomar como variables continuas¹¹. Además de estas variables, en la EAAE se encuentran variables de *gasto en servicios informáticos*, *licencia de software* e *inversión en activos intangibles* como el software, que también serán consideradas en el análisis. En total unas 19 variables que podrían estar capturando alguna forma de medir el concepto de TICs, es decir, las podemos interpretar a cada una como que nos proveen una medición distinta de la dimensión TICs. Bajo ese enfoque es que se va a aplicar el método no supervisado de Análisis de Componente Principal (ACP) para obtener una medida que resuma estas 19 variables obtenidas de la EAAE, agrupadas en una matriz $\mathbb{X}$.

El motivo por el que se opta por utilizar esta técnica es que se cuenta con muchas variables que a priori nos dan información sobre un mismo concepto, y lo que se pretende es reducir la dimensionalidad para utilizar una medida resumen como una única variable dependiente, con la particularidad de que esta única dimensión es la que representa la mayor variabilidad del conjunto de todas las variables relacionadas al concepto común, en este caso, las TICs. El método de ACP permite lograr este objetivo y en el Anexo (7.2) se presentan los detalles técnicos para la obtención de estos componentes. Otro motivo reside en el hecho de que las variables usualmente utilizadas en diversos trabajos, como el *gasto en servicios informáticos* o la *inversión en activos intangibles*, tienen la característica de presentar una densidad en cero elevada para nuestros datos provenientes de la EAAE. Esto implicaría un tratamiento particular de esta variable dependiente, pero a su vez esto último podría llevarnos a utilizar una estrategia empírica diferente a la de SEMs, la cual es presentada en el cuarto literal de la presente sección. Al aplicar ACP se logra evitar este problema.

¹⁰ Es el Índice de Precios Implícitos del Producto Interno Bruto por Industrias. Serie anual. Disponible en: www.bcu.gub.uy.

¹¹ Las variables del módulo de TICs fueron procesadas, a través del período de nuestro panel, de forma tal de utilizarlas en el Análisis de Componente Principal. La mayoría de ellas se obtiene directamente como vienen originalmente en la EAAE mientras que otras se construyeron agrupando diferentes binarias originales y formando una variable categórica, como los tramos de banda ancha.

Antes de efectuar la estimación por ACP se debieron estandarizar las variables que conforman la matriz $\mathbb{X}$, es decir, restar por su media y dividir por su desvío estándar. Esto nos permite evitar las diferencias de escala en nuestras x_s y que el proceso no asigne mayores pesos a algunas variables en relación con otras de forma espuria, es decir, producto de que estén medidas en distintas unidades. La siguiente es una tabla que describe las variables TICs que conforman el Análisis de Componente Principal, con el nombre que figuran en la base de datos, su etiqueta y una columna auxiliar que intenta clasificar cada variable procesada a los efectos de ganar en interpretación y de poder conjeturar cuál es el efecto esperado del Primer Componente principal (o del Segundo Componente principal) sobre la productividad y las exportaciones. Volveremos a este punto en la sección (5) al presentar los resultados de ACP e indagar sobre los coeficientes de los autovectores tanto para el Primer como Segundo Componente principal.

Nombre	Definición	Clasificación auxiliar
<i>ctic_1</i>	=2 si la empresa utiliza computadoras	Acceso
<i>ctic_2</i>	Cantidad de personas ocupadas que utiliza computador	Uso interno
<i>ctic_3</i>	=2 si la empresa utiliza Internet	Acceso
<i>ctic_4_baf</i>	Velocidad de Banda Ancha Fija, por tramos.	Acceso
<i>ctic_4_bam</i>	Velocidad de Banda Ancha Movil, por tramos.	Acceso
<i>ctic_5</i>	Cantidad de personas ocupadas que utiliza un computador conectado a Internet	Uso interno
<i>ctic_6_1</i>	=2 si la empresa recibió órdenes de compra a través de Internet	Comercio electrónico
<i>ctic_6_2</i>	Porcentaje declarado de las ventas realizadas a través de internet	Comercio electrónico
<i>ctic_7</i>	=2 si la empresa ordena productos a través de Internet	Comercio electrónico
<i>ctic_88_2</i>	=2 si usa Internet para Banca electrónica y/o otros serv. financieros y/o Transacciones con Org.Gubernamentales	Comercio electrónico
<i>ctic_88_3</i>	=2 si usa Internet para Servicio al cliente y/o Distribuir productos en línea y/o Proporcionar inf. sobre bienes y servicios de la firma	Comercio electrónico
<i>ctic_88_5</i>	=2 si usa Internet para Búsquedas de información diversas	Información
<i>ctic_88_6</i>	=2 si utiliza servicios de almacenamiento y computación en la nube	Información
<i>ctic_88_7</i>	=2 si utiliza internet para Capacitación del personal	Capacitación
<i>ctic_10</i>	=2 si dispone un sitio Web propio	Acceso
<i>ctic_13_1</i>	Cantidad de personas ocupadas que utiliza dispositivos portátiles de alta capacidad	Uso interno
<i>lstot_def</i>	Costo de Licencias de software total. Deflactada	Adopción (licencias)
<i>fb5_9_def</i>	Activos intangibles. FBKF (inversion). Deflactada	Adopción (fbkf)
<i>gal6_def</i>	Servicios informáticos. Deflactada	Mantenimiento (serv. Informáticos)

Cuadro 3: Descripción de variables TICs

Fuente: Elaboración propia en base a datos del INE.

Una vez resuelto el problema de escala, se debe procurar que la matriz $\mathbb{X}$ esté completa. Dado que es muy frecuente que determinadas variables de la base de datos de encuestas tengan valores faltantes, como sucede en este caso, debemos optar por alguna estrategia de imputación de estos valores. En el Anexo (7.3) se presenta una tabla con el porcentaje de valores faltantes de las diferentes columnas de $\mathbb{X}$. Para imputar estos valores faltantes se optó por un algoritmo iterativo, llamado “Matrix Completion”. La idea básica del algoritmo es poder explotar las correlaciones entre las diferentes x_s y de esta forma “predecir” cuales son los valores faltantes. En el mismo Anexo (7.3) se presenta el detalle de cómo opera este algoritmo. Con la matriz completa, esto es, valores originales e imputaciones para los casos en que las observaciones originales tenían valores faltantes, ahora podemos realizar nuestro análisis ACP. En la sección (5) serán presentados los resultados¹².

¹² Se realiza una aproximación utilizando solamente el Primer Componente principal, esto es, usando $M = 1$.

4.3 PTF: una estimación robusta

La productividad se puede entender como cuánto obtenemos de producción para un conjunto dado de insumos, mientras que su dual sería minimizar los insumos para un valor fijo de producción. Se expresa como una relación entre la producción y los insumos. Los niveles de productividad de un solo factor se ven afectados por la intensidad de uso de los insumos excluidos. Por ello, se utiliza la Productividad Total de los Factores (PTF), que es independiente de la intensidad de uso de los factores de producción observables. La variación del precio de los factores que impulsa las diferencias en la intensidad de los factores no afecta a la PTF, ya que induce cambios a lo largo de las isocuantas en lugar de desplazamientos en las isocuantas (Syverson, 2011).

La producción es resultado de una función de insumos observables y un factor neutral (según Hicks).

$$Y_{it} = W_{it}F(K_{it}, L_{it})$$

El término W_{it} es la PTF y captura las variaciones en la producción que no se explican por los cambios en los insumos observables que actúan a través de $F(\cdot)$. Si partimos de que $F(\cdot)$ es una función Cobb-Douglas, tomamos logaritmos y luego definimos a $\ln(W_{it}) = \beta_0 + \omega_{it} = PTF_{it}$, podemos estimar la siguiente función de producción:

$$y_{it} = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it}$$

Sin embargo, una estimación de los residuos resultante de aplicar MCO a dicha función de producción nos arrojaría valores alejados a la verdadera PTF en tanto se calcula mediante la base de coeficientes sesgados e inconsistentes, que suelen ser los asociados a los factores de producción del capital físico y del trabajo, además por construcción el residuo resultante tendrá media cero y no se podrá identificar a la productividad idiosincrática de cada empresa, ω_{it} , que es inobservable por el econometrista, de *todo* el término de error. Estos sesgos se derivan básicamente de dos fuentes de endogeneidad, una de ellas es causada por la simultaneidad que puede existir entre el momento de elección de un insumo y el nivel de productividad en ese momento¹³. Si existe correlación, cuánto más alta sea, mayor será el sesgo positivo asociado al coeficiente del insumo, que se entiende la empresa puede elegir una vez que observa su nivel de productividad. Si bien la empresa conoce su nivel de productividad, el econometrista no observa este dato, donde en la especificación será un factor inobservable, configurando así una primera fuente de sesgo, idea muy bien desarrollada en el trabajo seminal de Olley y Pakes (1992), OP de ahora en adelante.

Adicionalmente, otra fuente de endogeneidad es la inducida por el hecho de observar a las empresas que continúan operando en el mercado. Por un lado, es importante trabajar con un panel no balanceado puesto que restringir la muestra a un panel balanceado que solo contemple las empresas que se observan para todo el periodo induciría un sesgo de selección en la muestra. Esto queda patente en las estimaciones empíricas de OP mostrando como el hecho de trabajar con el panel balanceado induce sesgo en los coeficientes estimados. Sin embargo, aun cuando se trabaje con un panel no balanceado, todavía puede existir sesgo. La empresa que dispone de

¹³ El caso del *trabajo* en Olley y Pakes (1992); el caso del *trabajo* y de los *insumos intermedios* como los materiales o electricidad en el caso de Levinsohn y Petrin (2003).

grandes niveles de stock de capital esperará retornos positivos en los periodos subsiguientes, por lo que su umbral mínimo de productividad que le permite seguir operando será más bajo, y es este comportamiento de cierre de la empresa el que induce a una relación negativa entre el stock de capital y la creencia de la empresa acerca de su productividad corriente: podría continuar operando aun a bajas realizaciones de productividad.

En el trabajo pionero de Olley y Pakes (1992) comienzan por desarrollar un modelo teórico donde su solución contempla una *regla de cierre* para dar cuenta de este comportamiento y, por otro lado, una *demanda de inversión* que servirá como proxy de la productividad inobservable. En la práctica se trata de, por una parte, *invertir* la productividad inobservable como función de variables observables, mientras que la regla de cierre se traduce en estimar una *probabilidad de supervivencia* que dependerá de la mismas variables observables que la función inversa. En el caso de OP tratan la siguiente función de producción con argumentos como el capital físico, la edad y el trabajo.

$$y_{it} = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it} \quad (1)$$

Donde y_{it} representa el logaritmo del *valor agregado de producción*, a_{it} la *edad* de la empresa, k_{it} el logaritmo del *capital físico*, l_{it} el logaritmo de la cantidad de *trabajadores*, ω_{it} es la productividad de la firma no observable por el econometrista y η_{it} representa errores de medida y/o shocks de productividad.

El capital y la edad son consideradas variables de estado en su modelo teórico, mientras que el trabajo es el insumo variable que se ajusta con la productividad actual. No debe confundirse la connotación “variable” con dinámico. Esto es, el trabajo en este caso es una variable estática, puesto que no sigue ninguna dinámica y la empresa la elige en el periodo corriente t , una vez conocida o una vez formada su creencia del valor que toma la productividad en el período corriente. Esto, junto con la utilización de una demanda de inversión como proxy de la productividad son los rasgos del proceso generador de datos que subyace a la propuesta de OP, el cual constituye el punto de partida para los trabajos posteriores de Levinsohn y Petrin (2003), Wooldridge (2009) y Akerberg et al. (2015). En estos desarrollos posteriores se proponen formas levemente diferentes para obtener estimaciones consistentes de estos mismos coeficientes.

En el Anexo (7.5) se presenta un detalle de la parte metodológica de la propuesta de Olley y Pakes (1992) para interiorizar ambos sesgos. En resumidas cuentas, lo que proponen es un algoritmo semiparamétrico de estimación en dos etapas y, de esta forma, obtienen estimaciones consistentes de los coeficientes de la función de producción: capital físico, trabajo y edad. Finalmente, utilizando estas estimaciones pueden obtener una medida de la PTF:

$$PTF_{it} = \hat{\beta}_0 + \hat{\omega}_{it} = y_{it} - \hat{\beta}_a a_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_l l_{it} - \hat{\eta}_{it} \quad (2)$$

Para lograr la identificación, OP se basa en algunos supuestos restrictivos. El supuesto más restrictivo tiene que ver con que la productividad inobservable es un escalar. Esta es una restricción importante puesto que se asume que la función inversa de la inversión de capital sólo varía a través del tiempo, pero no entre las firmas. Esto implica que no es compatible con shocks heterogéneos (específicos de las empresas) asociados a la inversión, puesto que sólo varía en el tiempo. Cuando se dice que no son contemplados, se hace referencia a que son shocks que impactarían en la demanda de inversión pero que no serán tomados en cuenta por esta estimación de la función desconocida, la consecuencia es que las estimaciones son consistentes para un

PGD donde no existan shocks asociados al capital heterogéneo por empresa, en el caso que esto no se pueda sostener como supuesto, cabría esperar obtener una PTF levemente alejada de la verdadera PTF.

Tampoco serán contemplados shocks heterogéneos asociados a la mano de obra puesto que esta última está correlacionada con el nivel de productividad actual y esta última es parte de la función inversa desconocida que sólo varía a través del tiempo. Por el mismo motivo, si aumentáramos OP para agregar a la función de producción los materiales o el consumo de energía, si este insumo intermedio también se elige por parte de la empresa una vez que conoce su nivel de productividad, entonces aduce el mismo problema que la mano de obra. Tampoco serán contemplados shocks heterogéneos asociados al producto. Se podría decir que este es el costo que se debe pagar para poder obtener un proxy de la productividad y así, utilizarlo en la primera etapa del algoritmo de estimación. Otro punto por considerar, relacionado al plano empírico, es que, si la demanda de inversión es utilizada como proxy de la productividad, una manera de lograr invertir esa función de inversión debe ser utilizando los valores estrictamente positivos de la inversión. Esto puede suponer pérdida de datos en la práctica puesto que muchas encuestas poseen alta frecuencia de valores en cero para la inversión.

En tanto que Levinsohn y Petrin (2003) (LP de ahora en adelante) se diferencia del enfoque de OP en el siguiente aspecto clave. Para obtener el proxy de la productividad lo hace mediante la demanda de insumos intermedios (en lugar de la inversión), donde esta dependerá del capital físico y de la productividad inobservable. Luego invierte dicha productividad inobservable como función de las variables observables (capital, edad y los insumos intermedios) a los efectos de lograr su estimación en la primera etapa de formas no paramétricas, similar al enfoque de OP. Si bien el enfoque de LP no puede admitir shocks heterogéneos de las empresas relacionados con los bienes intermedios ni tampoco los relacionados a la mano de obra en tanto siguen correlacionados con la productividad actual, ahora si será posible admitir shocks heterogéneos relacionados con la inversión (como costos de ajuste del capital específico para cada empresa) puesto que una vez que se controla por el capital físico y la productividad, no hay razón para esperar que este tipo de shocks afecte a los insumos intermedios. Esta es la diferencia conceptual con respecto a OP. También podría ser menos probable encontrar muchos valores para nivel cero utilizando insumos intermedios, pero esto dependerá de cada encuesta particular.

Continuando con el desarrollo de la literatura de la PTF, tal como se menciona en el propio artículo de Wooldridge (2009), es en el trabajo de Akerberg et al. (2006) donde muestran el problema de identificación asociado a la estimación del coeficiente del trabajo en la primera etapa del algoritmo en las estrategias seguidas por OP y LP. Si el trabajo es elegido óptimamente por la empresa, será también una función determinista de la productividad inobservable y de las variables de estado, con lo que el coeficiente de la mano de obra no está identificado en forma no paramétrica. En palabras más simples, si el trabajo se determina en función de la productividad y el capital físico, entonces cuando en la primera etapa del algoritmo de OP (y LP) se estima en forma no paramétrica la función inversa desconocida, que depende justamente del capital y de la productividad, la variación de la mano de obra no permitirá identificar el coeficiente del trabajo puesto que gran parte de esa variación en la mano de obra estará tomada en cuenta en la estimación de la función no paramétrica desconocida. Así, el coeficiente de la mano de obra recogerá parte de la relación entre el capital, la productividad y el trabajo. Sobre este punto Akerberg et al. (2006) explicitan que puede ser más importante cuando se utilizan los insumos intermedios como en LP que utilizando la inversión como proxy de la productividad en OP. En Akerberg et al. (2006) proponen una solución híbrida entre los enfoques de Olley y Pakes

(1992) y Levinsohn y Petrin (2003) usando una estimación en dos etapas que no identifica a ningún parámetro en la primera etapa, pero identifica a todos en la segunda¹⁴.

Por su parte Wooldridge (2009) propone un método de estimación para los coeficientes de la función de producción en una única etapa. Presenta su desarrollo utilizando, al igual que LP, la demanda de insumos intermedios como proxy de la productividad. También siguiendo a Akerberg et al. (2006), su desarrollo tiene en cuenta el potencial problema de identificación asociado a la estimación de la mano de obra en la primera etapa de los desarrollos previos. Entonces la novedad por parte de Wooldridge (2009) es presentar un método de estimación en una única etapa, donde estima dos ecuaciones en simultáneo a través del Método Generalizado de los Momentos (GMM) mediante el uso adecuado de instrumentos. La forma que tiene su método de lograr la identificación del coeficiente de la mano de obra (y de todos los coeficientes) es que los instrumentos le permitirán la identificación de ambas ecuaciones y por lo tanto, todos sus parámetros tendrán una única solución. Tal como se muestra en el Anexo (7.6), Wooldridge (2009) plantea su desarrollo en base a condiciones de momentos basadas en el trabajo de Levinsohn y Petrin (2003). El autor encuentra algunas ventajas de este método respecto a los basados en dos etapas. Errores estándar robustos fácilmente obtenibles (a diferencia de las técnicas en dos etapas que deben utilizar Bootstrap). Además, las estimaciones en dos etapas podrían ser ineficientes por dos motivos: al ignorar la correlación contemporánea en los errores a través de las dos ecuaciones y al no poder incorporar eficientemente la correlación serial o la heterocedasticidad en los errores. GMM utiliza la correlación cruzada de las ecuaciones y la matriz de varianza y covarianza con peso óptimo.

Finalmente, en el trabajo de Akerberg et al. (2015) proponen una estimación semiparamétrica en dos etapas, similar a los enfoques de LP y OP, donde se resuelve el problema de identificación asociado al coeficiente del insumo trabajo en la primera etapa. Esto lo logran planteando una modificación al supuesto de escalar inobservable, en ese sentido, ahora la función de demanda de bienes intermedios será *condicional* al insumo trabajo y ello le va a permitir considerar al trabajo como una variable que puede presentar un comportamiento dinámico. Como consecuencia de ello, supera a la estrategia de LP (y también la de Wooldridge (2009) y de OP) en que este proceso generador de datos será compatible con shocks específicos de la empresa asociados a la mano de obra, y por lo tanto producirá estimaciones consistentes aún en presencia de este tipo de shocks. En este caso, todos los coeficientes se identifican en la segunda etapa del algoritmo. En el Anexo (7.7) se presenta una breve discusión del problema de identificación de la primera etapa y de la propuesta desarrollada por estos autores.

Shocks específicos	k_{it}	l_{it}	m_{it}
OP	×	×	×
LP	✓	×	×
Wooldridge (2009)	✓	×	×
Akerberg et al. (2015)	✓	✓	×

Cuadro 4: Shocks específicos de empresa consistentes con cada enfoque

Fuente: Elaboración propia.

¹⁴ Bajo supuestos acerca de los momentos temporales de las elecciones de insumos, configurando otro PGD.

4.4 SEMs: metodología econométrica

Siguiendo el marco teórico desarrollado en la sección (2) disponemos de tres ecuaciones. La primera ecuación será para explicar la utilización de las TICs por parte de la firma, medida que se obtuvo mediante la técnica no supervisada de Análisis de Componente Principal. La segunda ecuación será para explicar la PTF anteriormente estimada, medida que se obtuvo siguiendo la metodología de Akerberg et al. (2015). Mientras que la tercera ecuación dará cuenta de las exportaciones deflactadas de la firma, en logaritmo natural.

Si creemos que las TICs pueden ser explicadas por la PTF, donde una firma por ser más productiva podría generar mayores ingresos y así disponer de mayor capital financiero y adquirir un mayor nivel de TICs. En dicho caso las TICs, además de determinar a la PTF, son determinadas por la PTF, entonces la relación causal va en ambos sentidos, lo que da lugar al concepto de *simultaneidad*. Por otra parte, tal como se aclaró al comienzo de la sección (2), no vamos a analizar la posible existencia de un mecanismo de *learning by exporting*, por lo cual, el modelo que se corresponde 1:1 con nuestro marco teórico es el siguiente:

$$\begin{aligned} \text{TICs}_{it} &= \gamma_{12}\text{PTF}_{it} + \mathbf{x}_{1it}\beta_1 + u_{1it} \\ \text{PTF}_{it} &= \gamma_{21}\text{TICs}_{it} + \mathbf{x}_{2it}\beta_2 + u_{2it} \\ \ln(\text{EXP})_{it} &= \gamma_{31}\text{TICs}_{it} + \gamma_{32}\text{PTF}_{it} + \mathbf{x}_{3it}\beta_3 + u_{3it} \end{aligned} \quad (3)$$

A diferencia de modelos de múltiples ecuaciones que se estiman en forma recursiva, donde implícitamente se asume que la dirección de la causalidad va en un solo sentido, en el presente modelo se permite la posibilidad de que dos o más variables endógenas se determinen mutuamente una a la otra, lo que da sentido al concepto de simultaneidad. Cuando dos variables se determinan conjuntamente, como en este modelo sucede con las TICs y la Productividad Total de los Factores, lo que sucede es que la PTF en la ecuación de TICs estará relacionada con los errores u_{1it} y por lo tanto será endógena en la ecuación de TICs. Lo mismo sucede con la variable TICs en la ecuación de PTF, esta será endógena en la ecuación de PTF al estar correlacionada con el término de error, u_{2it} .

Debido a la presencia de variables endógenas necesitaremos instrumentar dichas variables en cada ecuación. El método de estimación *Método Generalizado de los Momentos* (GMM) nos permite estimar de forma consistente el Modelo de Ecuaciones Simultáneas (SEMs). El desarrollo y aplicación de la estrategia de estimación se basa en los capítulos 5, 7, 8 y 9 del libro Wooldridge (2010). Respecto a la consistencia, el estimador GMM para un SEMs requerirá el supuesto fuerte de exogeneidad:

$$\mathbb{E}(\mathbf{Z}_{it} \otimes \mathbf{u}_{it}) = 0 \quad (4)$$

Se puede demostrar que el producto de Kronecker entre la matriz de instrumentos $\mathbf{Z}_{it}$ y el vector de los errores del sistema $\mathbf{u}_{it}$, implica que cada submatriz del tipo z_{hit} (submatriz de instrumentos para la ecuación h) no debe estar correlacionada en media con ninguno de los errores del sistema, $u_{hit} \forall h = 1, 2, 3$. De esta forma, a modo de ejemplo, necesitamos que la submatriz de instrumentos para la ecuación de TICs, z_{1it} no se correlacionen en media con u_{1it} , pero tampoco con los errores de la ecuación de PTF, u_{2it} ni con los errores de la ecuación de Exportaciones,

u_{3it} .

El método de GMM, al disponer de más instrumentos no triviales que variables endógenas, se pregunta de qué manera los puede combinar. En particular, la expresión que toma el estimador GMM general es la siguiente:

$$\hat{\beta} = (\mathbf{X}'_{it} \mathbf{Z}_{it} \hat{\mathbf{W}} \mathbf{Z}'_{it} \mathbf{X}_{it})^{-1} (\mathbf{X}'_{it} \mathbf{Z}_{it} \hat{\mathbf{W}} \mathbf{Z}'_{it} \mathbf{Y}_{it}) \quad (5)$$

La matriz de ponderación $\hat{\mathbf{W}}$ es una matriz $\mathbf{M} \times \mathbf{M}$ donde $\mathbf{M}$ se compone de la suma de los instrumentos de la ecuación 1, más los instrumentos de la ecuación 2, más los instrumentos de la ecuación 3. Entonces $\mathbf{M} = M_1 + M_2 + M_3$. Además, debe ser una matriz simétrica y semidefinida positiva.

Según el tipo de matriz de ponderación elegida, el estimador GMM tomará diferentes versiones. Si elegimos $\hat{\mathbf{W}} = \left(\frac{\mathbf{Z}'_{it} \mathbf{Z}_{it}}{N} \right)^{-1}$ entonces se puede demostrar que el estimador GMM converge al estimador del Sistema de Mínimos Cuadrados en Dos Etapas (GMM-S2SLS). Dado que en un sistema de ecuaciones simultaneas (SEMs) el sistema *debe* imponer restricciones sobre los parámetros, una consecuencia es que el estimador GMM-S2SLS no será idéntico a estimar por Mínimos Cuadrados en Dos Etapas ecuación por ecuación. Generalmente el estimador GMM-S2SLS se utiliza como forma auxiliar en una primera etapa de otras versiones del estimador GMM en dos etapas, que veremos a continuación.

Si elegimos a $\hat{\mathbf{W}} = \Lambda^{-1} = [\mathbb{E}(\mathbf{Z}'_{it} \mathbf{u}_{it} \mathbf{u}'_{it} \mathbf{Z}_{it})]^{-1} = [\text{Var}(\mathbf{Z}'_{it} \mathbf{u}_{it})]^{-1}$, se puede demostrar que entonces obtendremos la versión de GMM con la Matriz de Ponderación Óptima, GMM-OWM. Como se mencionó, el procedimiento consiste en que en la primera etapa se utiliza la estimación GMM-S2SLS de forma auxiliar, se obtiene el vector de residuos del sistema, y luego, con estos residuos, se puede obtener una estimación consistente de $\hat{\Lambda}$ y, con ello, una estimación consistente para $\hat{\mathbf{W}} = \Lambda^{-1}$. La expresión que toma el estimador GMM-OWM es:

$$\hat{\beta}_{\text{owm}} = (\mathbf{X}'_{it} \mathbf{Z}_{it} \hat{\Lambda}^{-1} \mathbf{Z}'_{it} \mathbf{X}_{it})^{-1} (\mathbf{X}'_{it} \mathbf{Z}_{it} \hat{\Lambda}^{-1} \mathbf{Z}'_{it} \mathbf{Y}_{it}) \quad (6)$$

Con respecto a la estructura de la matriz de varianzas y covarianzas de los coeficientes, la expresión correspondiente al estimador GMM con Matriz de Ponderación Óptima es la siguiente:

$$\frac{\widehat{\text{Var}}(\hat{\beta}_{\text{owm}})}{N} = [(\mathbf{X}'_{it} \mathbf{Z}_{it}) \hat{\Lambda}^{-1} (\mathbf{Z}'_{it} \mathbf{X}_{it})]^{-1} \quad (7)$$

Bajo el supuesto de exogeneidad presentado (4) y los habituales supuestos asociados la condición de rango, la matriz de varianzas y covarianzas definida en (7) es robusta ante la presencia de heterocedasticidad en el sistema: las firmas entre sí pueden presentar distinta varianzas en la ecuación de TICs, presentar distinta varianza en la ecuación de PTF y presentar distinta varianza en la ecuación de EXP. Dentro de una misma ecuación, o lo que es lo mismo, dentro de los bloques de la diagonal principal de la matriz $\mathbb{E}(\mathbf{u}_{it} \mathbf{u}'_{it})$, sus elementos en la diagonal principal serán las varianzas de cada firma para dicha ecuación. Fuera de esa diagonal principal tendremos las correlaciones entre dos firmas diferentes dentro de la misma ecuación. Estas correlaciones

al interior de una misma ecuación para dos firmas diferentes también son admitidas, es decir, pueden ser distintas de cero.

En tanto que para los bloques de la matriz $\mathbb{E}(\mathbf{u}_{it}\mathbf{u}_{it}')$ que están por fuera de los bloques de la diagonal principal, estarán recogiendo las correlaciones entre los residuos de las diferentes ecuaciones. Entonces esta matriz de varianzas y covarianzas también es robusta ante posibles correlaciones cruzadas de los errores de distintas ecuaciones, tanto para la misma firma como para distintas firmas. Con lo cual, debe quedar claro de la robustez de la matriz de varianzas y covarianzas presentada en (7) admite tanto presencia de heterocedasticidad en el sistema como presencia de correlaciones cruzadas entre las ecuaciones del sistema.

Luego se presentará en el Anexo (7.10) la versión del SEMs utilizando el estimador GMM-3SLS, o estimador GMM de Mínimos Cuadrados en Tres Etapas. Es importante notar que, si bien su nombre indica tres etapas, computacionalmente no conlleva tres etapas y el nombre se debe a razones históricas. Lo que hace GMM-3SLS es imponer estructura sobre (7). En particular asume que el sistema de ecuaciones es homocedástico. Esto implica que condicional a la ecuación de TICs, todas las firmas comparten la misma varianza, y así sucesivamente. Pero la varianza de la ecuación de TICs puede diferir de la varianza de la ecuación de PTF y de la varianza de la ecuación EXP respectivamente. Si además imponemos a la matriz (7) que sea *diagonal*, entonces tampoco admitiremos correlaciones cruzadas entre ecuaciones del sistema.

El siguiente cuadro contiene las definiciones de todas las variables que forman parte del SEMs: las variables endógenas y los instrumentos.

Variable	Definición
<i>TICs</i>	Primer componente principal del análisis ACP.
<i>PTF</i>	Estimada mediante el método de Akerberg, Caves y Frazer (2015).
$\ln(EXP)$	Logaritmo natural de las exportaciones deflactadas.
$\ln(tamaño)$	Logaritmo natural del personal total ocupado.
<i>cal_emp_intsec</i>	Salario deflactado de la firma en relación con el salario medio deflactado de su propio sector, en el mismo año.
$\ln(i+d)$	Logaritmo natural del costo de investigación y desarrollo. Deflactado.
$\ln(apalancamiento)$	Logaritmo natural del cociente entre el pasivo total y el patrimonio.
$\ln(roe)$	Logaritmo natural del cociente entre el resultado del ejercicio y el patrimonio.
$\ln(stock_capital)$	Logaritmo natural del stock de capital deflactado.
<i>cuota_mercado</i>	Cociente entre los ingresos por ventas de la firma y los ingresos totales del sector CIIU revisión 4 correspondiente, en el mismo año.
<i>hhi</i>	Índice de Herfindahl-Hirschman (HHI) por industria y año. Mide la concentración sectorial.
<i>capital_ext</i>	Porcentaje de capital extranjero sobre el capital total.
<i>dummies_industria</i>	Variables binarias por industria (industria TICs como referencia).
<i>dummies_año</i>	Variables binarias por año (año 2012 como referencia).
$\ln(merc_est_adq)$	Logaritmo natural del costo de mercaderías vendidas en el mismo estado en que fueron adquiridas. Deflactado.
$\ln(ingreso_total)$	Logaritmo natural de los ingresos totales sin subsidios. Deflactado.
$\ln(publicidad)$	Logaritmo natural del gasto en publicidad. Deflactado.
$\ln(gasto_fin_ban)$	Logaritmo natural de los gastos por servicios bancarios y financieros. Deflactado.

Cuadro 5: Definición de variables del modelo SEMs

Fuente: Elaboración propia.

La mayoría de las variables utilizadas se toman directamente de la encuesta EAAE de forma original. Mientras que otras, como *cal_emp_intsec*, *cuota_mercado* o el índice *hhi* se construyeron sobre la base de variables existente en la EAAE. Como se ha mostrado en las secciones previas, las variables *TICs* y *PTF* surgen de realizar estimaciones con distintos métodos. A la hora de elegir las variables que forman parte de cada una de nuestras ecuaciones estructurales del modelo se tuvo en cuenta el marco teórico y los antecedentes de la literatura relacionada. Las variables asociadas a la rentabilidad de las firmas (*apalancamiento*, *roe*) se estimaron en su transformación de logaritmo natural, esto provocó la pérdida de aproximadamente un 4 % del total de observaciones de la muestra. Esta pérdida de datos no es aleatoria puesto que se debe a empresas que tienen valores negativos en una o ambas variables de rentabilidad. Sin embargo, los coeficientes estimados del resto de las variables del modelo prácticamente no se alteraron cuando estas variables se estimaron en niveles o en logaritmo natural, por lo tanto, y considerando que el ajuste es mejor cuando usamos la transformación logartimo, se opta por esta ultima opción. Por otro lado, lamentablemente no se pudo utilizar la medida de profesionales y técnicos

como proporción del total de personal ocupado ya que no se ha relevado en el encuesta EAAE para el periodo considerado.

El SEMs presentado en (3) no es todavía la versión a estimar por GMM debido a que aún no se han definido los *instrumentos no triviales*. En concreto, los vectores que contienen los instrumentos triviales (o controles) de cada ecuación son:

$$\begin{aligned} x_1 &= [\text{constante}, \ln(\text{tamaño}), \text{cal_emp_intsec}, \ln(i+d), \\ &\quad \ln(\text{apalancamiento}), \ln(\text{roe}), \text{dummies_industria}, \text{dummies_año}]' \\ x_2 &= [\text{constante}, \ln(\text{tamaño}), \text{cal_emp_intsec}, \ln(i+d), \\ &\quad \text{cuota_mercado}, \text{hhi}, \ln(\text{stock_capital}), \text{dummies_industria}, \text{dummies_año}]' \\ x_3 &= [\text{constante}, \ln(\text{tamaño}), \ln(i+d), \text{capital_ext}, \\ &\quad \text{dummies_industria}, \text{dummies_año}]' \end{aligned}$$

Nos queda ahora presentar el SEMs a ser estimado, donde para ello necesitamos disponer de *instrumentos no triviales*, al menos uno para cada variable endógena presente en cada ecuación:

$$\begin{aligned} \text{TICs}_{it} &= \gamma_{12}\text{PTF}_{it} + z_{1it}\delta_1 + u_{1it} \\ \text{PTF}_{it} &= \gamma_{21}\text{TICs}_{it} + z_{2it}\delta_2 + u_{2it} \\ \ln(\text{EXP})_{it} &= \gamma_{31}\text{TICs}_{it} + \gamma_{32}\text{PTF}_{it} + z_{3it}\delta_3 + u_{3it} \end{aligned} \tag{8}$$

Donde cada vector de instrumentos se conforma de las mismas variables exógenas definidas previamente y de los *instrumentos no triviales*:

$$\begin{aligned} z_1 &= [x'_1, \text{cuota_mercado}, \ln(\text{merc_est_adq}), \ln(\text{ingreso_total})]' \\ z_2 &= [x'_2, \ln(\text{publicidad}), \ln(\text{gasto_fin_ban})]' \\ z_3 &= [x'_3, \text{cuota_mercado}, \ln(\text{ingreso_total}), \ln(\text{gasto_fin_ban})]' \end{aligned}$$

Como vemos, disponemos de más de un *instrumento no trivial* para instrumentar la o las variables endógenas en cada ecuación, esto nos permitirá evaluar la validez de estos instrumentos, como presentamos en el Anexo (7.9) mediante el estadístico J de Hansen.

En cuanto a la *identificación*, vemos que $\ln(\text{apalancamiento})$ y $\ln(\text{roe})$ nos ayudan a lograr la identificación de las ecuaciones de PTF y de $\ln(\text{EXP})$, ya que aparecen en z_1 mientras que no forman parte ni de z_2 ni de z_3 . El índice hhi nos ayuda a lograr la identificación de la ecuación de TICs y de la ecuación de $\ln(\text{EXP})$ al formar parte de z_2 mientras que no forma parte de z_1 ni de z_3 . La variable cal_emp_intsec nos ayuda a identificar la ecuación de $\ln(\text{EXP})$. Por su parte, capital_ext nos ayuda a identificar las ecuaciones de TICs y la de la PTF ya que forma parte de z_3 pero no forma parte ni de z_1 ni de z_2 .¹⁵ Si bien este razonamiento intuitivo puede tener su valor, en el Anexo (7.8) se presenta la verificación formal de la *condición de rango* para lograr la identificación en el SEMs. Algebraicamente:

¹⁵ Como queda claro en el Anexo (7.8), los instrumentos no triviales también nos ayudan a lograr la identificación de las respectivas ecuaciones si no están presente en todas las ecuaciones.

$$\text{Rank}[R_h \mathbf{B}] = \mathbf{G} - 1 \rightarrow 3 - 1 = 2 \quad (9)$$

Donde R_h es una matriz de constantes conocidas de dimensión $(J_h \times (\mathbf{G} + \mathbf{M}))$, con J_h la cantidad de restricciones sobre β_h , sin considerar la restricción de normalización de la variable dependiente. Donde además, $\mathbf{G}$ es la cantidad de ecuaciones del sistema y $\mathbf{M}$ es la cantidad de instrumentos de todo el sistema. En concreto, R_h es una matriz de constantes conocidas dado que depende directamente de las restricciones que pongamos sobre los parámetros para la ecuación h , en nuestro caso, utilizaremos solamente restricciones de exclusión por lo cual, será una matriz de ceros y unos. Por su parte, $\mathbf{B}$ es la matriz de parámetros estructurales del SEMs, de dimensión $(\mathbf{G} + \mathbf{M}) \times \mathbf{G}$.

Respecto a la validez de los instrumentos, se optó por la estrategia de trabajar a nivel de cada ecuación por separado a efectos de evaluar su calidad: la *relevancia* y la *exogeneidad* de los instrumentos. Los motivos residen en que computacionalmente el SEMs es muy exigente y además el comando `gmm` de STATA no muestra en su salida estadísticos para poder evaluar ni la relación entre los instrumentos con la variable endógena ni pruebas de hipótesis para evaluar si los instrumentos utilizados están teniendo correlación con los residuos o podemos estar tranquilos de que no rechazamos que sean exógenos.

Trabajando con el comando `ivreg2` de STATA se pudo evaluar tanto la relación de los instrumentos con la variable endógena como los estadísticos relevantes que evalúan si podemos no rechazar la hipótesis nula de que los instrumentos sean exógenos a nivel de una ecuación. Estas salidas se presentan en el Anexo (7.9) para cada una de las tres ecuaciones. De esta forma, una vez que se obtuvieron resultados satisfactorios ecuación por ecuación, se procede a estimar el SEMs con el comando `gmm`. Se presentan los resultados en la sección (5.4).

Si bien se dispone de datos de panel para todas las variables, no vamos a utilizar un método de estimación habitual para este tipo de estructura de datos como pueden ser modelos de efectos fijos o paneles dinámicos (ambos para una sola ecuación). En este caso, la ventaja de disponer de datos de panel está en la mayor cantidad de datos y en que podemos controlar por efectos agregados temporales en cada ecuación del sistema. En las notaciones se opta por mantener el subíndice temporal para hacer explícito que los datos varían por año, excepto en el caso de las variables binarias que capturan el efecto industria, efectos que también forma parte de cada una de nuestras ecuaciones.

5 Resultados

5.1 TICs: un análisis descriptivo

Los resultados de la estimación ACP para la matriz $\mathbb{X}$ utilizando 32479 observaciones, que agrupa a 19 variables TICs de la encuesta EAAE, son los siguientes.

Componente	Valor propio	Diferencia	Proporción	Acumulado
Comp1	11.9126	10.0217	0.6270	0.6270
Comp2	1.8909	0.9281	0.0995	0.7265
Comp3	0.9628	0.0335	0.0507	0.7772
Comp4	0.9292	0.0246	0.0489	0.8261
Comp5	0.9047	0.4691	0.0476	0.8737
Comp6	0.4356	0.1245	0.0229	0.8966
Comp7	0.3111	0.0367	0.0164	0.9130
Comp8	0.2745	0.0506	0.0144	0.9274
Comp9	0.2238	0.0172	0.0118	0.9392
Comp10	0.2066	0.0223	0.0109	0.9501
Comp11	0.1843	0.0194	0.0097	0.9598
Comp12	0.1649	0.0243	0.0087	0.9685
Comp13	0.1407	0.0421	0.0074	0.9759
Comp14	0.0985	0.0027	0.0052	0.9811
Comp15	0.0959	0.0148	0.0050	0.9861
Comp16	0.0811	0.0021	0.0043	0.9904
Comp17	0.0790	0.0136	0.0042	0.9945
Comp18	0.0654	0.0270	0.0034	0.9980
Comp19	0.0384	—	0.0020	1.0000

Cuadro 6: Valores propios de los 19 componentes principales

Variable	Comp1	Comp2
ctic_1_scaled	0.2346	-0.0544
ctic_2_scaled	0.1008	0.5903
ctic_3_scaled	0.2496	-0.0826
ctic_4_baf_scaled	0.2634	-0.0466
ctic_4_bam_scaled	0.2588	-0.0619
ctic_5_scaled	0.2383	0.3350
ctic_6_1_scaled	0.2760	-0.0786
ctic_6_2_scaled	0.2707	-0.0848
ctic_7_scaled	0.2790	-0.0644
ctic_88_2_scaled	0.2708	-0.0688
ctic_88_3_scaled	0.2778	-0.0652
ctic_88_5_scaled	0.2643	-0.0710
ctic_88_6_scaled	0.2719	-0.0686
ctic_88_7_scaled	0.2738	-0.0379
ctic_10_scaled	0.2509	-0.0487
ctic_13_1_scaled	0.1396	0.5231
lstot_def_scaled	0.0099	0.2512
fb5_9_def_scaled	0.0095	0.2410
gal6_def_scaled	0.0131	0.2977

Cuadro 7: Vectores propios de los 2 primeros componentes principales

El Primer Componente Principal, $PC1$, tiene asociado un autovalor $\lambda_1 = 11.91$ y una varianza explicada en relación con la varianza total de 62.70 % mientras que, el Segundo Componente Principal, $PC2$, tiene un autovalor $\lambda_2 = 1.89$, representado una varianza explicada del 9.95 %. Vemos que la mayoría de los restantes Componentes Principales explican un porcentaje muy bajo de la varianza total. Lo siguiente es analizar los autovectores de los dos primeros componentes principales, analizando sus coeficientes o “loadings”, donde será de ayuda recordar el Cuadro (3), presentado en el apartado segundo de la sección (4). Para facilitar una plausible interpretación de los dos primeros componentes principales será útil apoyarse tanto en la clasificación auxiliar realizada previamente, así como en la *contribución* de cada variable a cada uno de estos dos componentes¹⁶.

Comenzando con el autovector δ_1^* , todas las variables TICs contribuyen en la misma dirección del $PC1$, puesto que todos los coeficientes son positivos. De las 14 variables que contribuyen más que el promedio, hay una fuerte relevancia de variables TICs asociadas al comercio electrónico (*ctic_7*, *ctic_88_3*, *ctic_6_1*, *ctic_88_2*, *ctic_6_2*), contribuyendo un 37.78 %, seguidas de variables TICs de acceso o infraestructura (*ctic_4_baf*, *ctic_4_bam*, *ctic_10*, *ctic_3*, *ctic_1*), contribuyendo un 31.67 %. También son relevantes variables TICs de gestión de la información (*ctic_88_6*, *ctic_88_5*), con un 14.38 %, mientras que la única que recoge un uso de Internet con fines de capacitación del personal (*ctic_88_7*) explica un 7.50 %. Dentro de las tres variables de uso interno solamente una de ellas explica levemente más que el promedio (*ctic_5*). Se puede observar que las variables TICs de compras de licencias de software, inversión en activos intangibles como software y el gasto en servicios informáticos apenas contribuyen en esta dimensión de las TICs recogidas en el Primer Componente Principal, con un 0.04 %.

Mientras que por el lado del autovector δ_2^* se observa que algunas variables TICs contribuyen en la misma dirección del $PC2$ mientras que otras contribuyen en la dirección opuesta. En general las que contribuyen en la misma dirección son las variables de uso interno y las de gasto y/o inversión en software. En ese sentido, podemos observar que hay una fuerte relevancia en las variables de uso interno (*ctic_2*, *ctic_5*, *ctic_13_1*), contribuyendo a esta dimensión en un 73.43 %, seguidos de una gran relevancia de las variables de inversión y gasto en software (*gal6_def*, *ls_tot_def*, *fb5_9_def*), contribuyendo en un 20.98 %. El resto de las variables contribuyen menos que el promedio para esta segunda dimensión del concepto TICs, y lo hacen en dirección contraria, con un total de varianza explicada del 5, 58 %. En resumen, esta segunda dimensión de las TICs está ponderando fuertemente a las variables de uso interno seguidas de las asociadas a la inversión y gasto en software.

El siguiente es un gráfico donde se puede apreciar tanto los *scores* de los primeros dos componentes principales ($PC1$, $PC2$) como los coeficientes de sus autovectores (δ_1^* , δ_2^*) que por construcción son ortogonales. En azul se representa a cada empresa para algún año del período considerado. Valores cercanos a 0 de $PC1$ son empresas que disponen de una representación promedio en esta dimensión de TICs, mientras que las empresas que se ubican a la derecha son aquellas que disponen de una mayor intensidad TICs en este Primer Componente Principal, que como vimos, pondera a variables TICs asociadas a comercio electrónico, acceso o infraestructura, gestión de la información y uso en capacitación del personal. Se puede hacer una interpretación similar para $PC2$, donde los valores que estén por encima de 0 serán de empresas que disponen de una mayor intensidad TICs en esta segunda dimensión. Asimismo, las líneas

¹⁶ La contribución de cada variable surge de elevar al cuadrado cada coeficiente y luego dividir cada término por la suma total de los cuadrados. En el Anexo (7.4) se presentan dos gráficos asociados a esta contribución.

rojas que parten desde el origen son representaciones de las distintas variables tomando el valor del coeficiente correspondiente tanto para al Primer como al Segundo Componente Principal, así, por ejemplo, $ctic_5$ tiene un coeficiente asociado con δ_1^* de 0.2383 y un coeficiente asociado con δ_2^* de 0.3350, por lo cual se representa un vector que va desde el origen a este punto del plano.

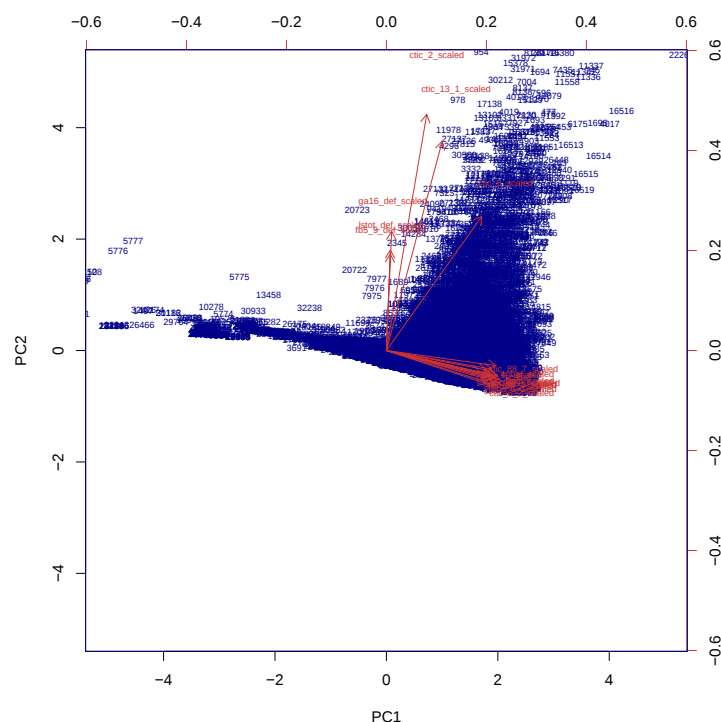


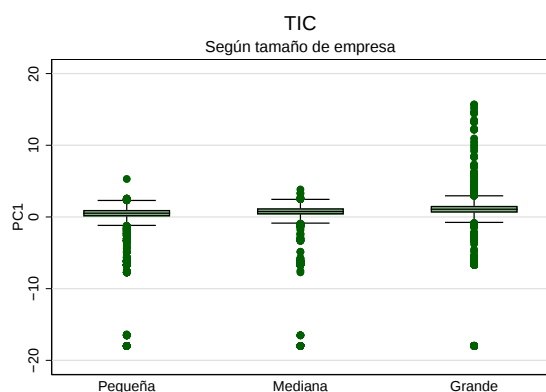
Figura 2: Biplot

Fuente: Elaboración propia.

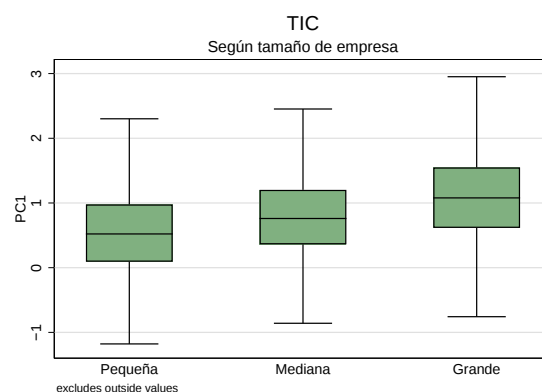
Podemos utilizar el Primer Componente Principal, entendido como un índice de intensidad TICs de las empresas, y analizar si existen diferencias según el tamaño de estas. Para ello utilizamos la cantidad total de personas ocupadas reportadas y se crea una variable categórica para el tamaño de la empresa. Las empresas denominadas *pequeñas* son aquellas cuyo tamaño se ubica entre el mínimo de la encuesta y la mediana, que es de 35 empleados. Las de tamaño *medio* son aquellas que se ubican entre la mediana y el valor dado por el percentil 75, que es de 79 empleados, por lo que las empresas *grandes* son las que sobrepasan dicho umbral.

La siguiente figura (3) presenta cuatro gráficas de la relación entre la intensidad TICs dada por $PC1$ y el tamaño de las empresas. Comenzando por el panel (a) donde se representa un gráfico de caja, se observa que tanto para las pequeñas como para las medianas empresas su distribución es similar. Esto se observa tanto comparando el rango intercuartílico, como el hecho de observar valores atípicos con un patrón similar en la cola izquierda de su distribución. En tanto que, para las grandes empresas, si bien a simple vista el rango intercuartílico parece similar a las PYMEs, se observa una marcada diferencia al encontrar valores atípicos a la derecha de su distribución. El panel (b) vuelve a presentar el gráfico de cajas habiendo eliminado estos valores atípicos. Se observa ahora leves diferencias en la mediana entre las pequeñas y medianas empresas, y una mayor diferencia para las grandes empresas. También una mayor dispersión para las estas últimas. Otra forma de ver la distribución por tamaño de empresa puede ser a través de

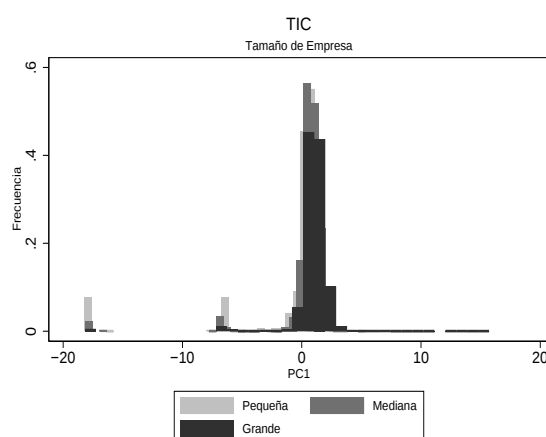
un histograma, como lo muestra el panel (c). En común tienen que están centradas en valores cercanos a 0, lo cual es esperable dado la construcción del índice. Sin embargo, ahora podemos ver mejor sus diferencias, como se observa de una mayor densidad en la cola izquierda para las pequeñas empresas, seguidas de las de tamaño medio, en relación con las grandes empresas. La cola derecha posee densidad para las grandes empresas únicamente. Esto se aprecia aún mejor si comparamos la densidad estimada por Kernel del Primer Componente Principal para las pequeñas empresas por un lado y las grandes empresas por otro. Esto se presenta en el panel (d).



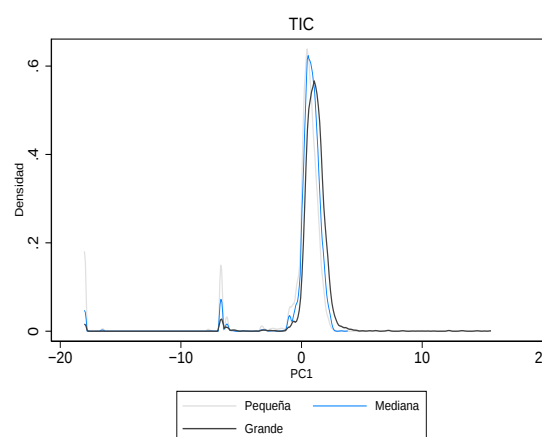
(a) Gráfica 1



(b) Gráfica 2



(c) Gráfica 3



(d) Gráfica 4

Figura 3: TICs (PC1) y tamaño de la empresa

Fuente: Elaboración propia.

En el Cuadro (2) se presentó el listado de industrias o sectores que forman parte de la muestra EAAE. Se presentan dos gráficas con el promedio y la mediana respectivamente, para cada una de estas industrias en la intensidad TICs dada por $PC1$. Se observa que para todos los sectores su valor promedio es inferior a su mediana. La presencia de una mayor densidad en la cola izquierda de la distribución subyace al anterior resultado. En cuanto a los sectores, tanto considerando el promedio como su mediana, se destacan los sectores de *Información y Comunicación*, *Administración Pública y Defensa*, así como de los sectores de *Actividades Inmobiliarias*, *Profesionales y Financieras*, respectivamente. Luego, ya en un nivel medio podemos situar al sector de la *Enseñanza*. Los sectores más bajos corresponden a la *Suministro de Agua, Alojamiento y*

Servicio y el sector de Otras actividades de servicios.

En el caso de la *Industria Manufacturera* incluye una heterogeneidad de sectores importante, desde la farmacéutica, textil, alimentos, madera, papel, combustibles, químicos, plástico, vidrio, etc. Por lo que su nivel, tanto promedio como su mediana, están capturando la información de diversas divisiones de la CIU revisión 4. Mientras en promedio (y en conjunto) es uno de los sectores de menor nivel de TICs, en lo que respecta a la mediana se ubica en una posición más favorable.

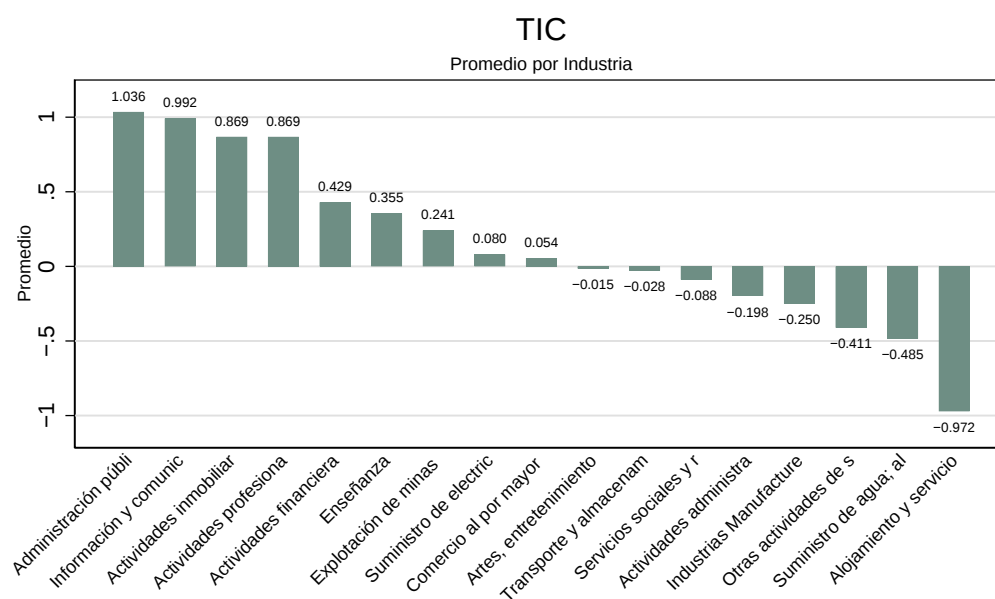


Figura 4: Promedio de TICs (PC1) por Industria

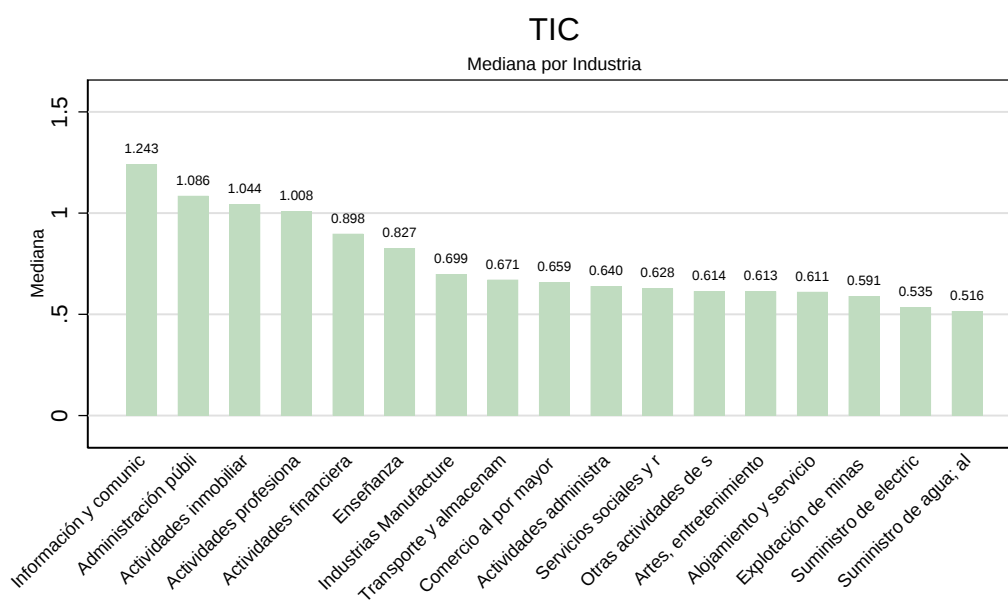


Figura 5: Mediana de TICs (PC1) por Industria

Fuente: Elaboración propia.

5.2 PTF: un análisis descriptivo

En la presente sección se presentan los resultados de la estimación de la productividad total de los factores a nivel de empresa utilizando cada una de las técnicas presentadas en la sección (4.3). En todos los métodos se utilizó al logaritmo natural del *valor agregado bruto deflactado* como variable dependiente, se especificó al logaritmo natural del *personal total ocupado* como variable libre y al logaritmo natural del *stock de capital deflactado* como variable de estado. Mientras que el logaritmo natural de la *inversión deflactada* se usó como proxy de la productividad en los métodos de OP y ACF, en tanto que el logaritmo natural del *gasto en consumo energético* se utilizó como proxy de la productividad en los métodos de LP y WRD¹⁷.

La siguiente salida muestra las estadísticas descriptivas correspondientes a las estimaciones resultantes de aplicar los métodos de Olley y Pakes (1992), Levinsohn y Petrin (2003), Wooldridge (2009), Akerberg et al. (2015) denotados con OP, LP, WRD y AFC respectivamente. Para cada uno de los métodos se muestra una primera versión sin tendencia y la siguiente incluyéndola, que se entiende puede ajustar mejor los datos. Se presenta, en forma adicional, al logaritmo natural de la productividad laboral definida como el cociente entre el *valor agregado bruto* y el *personal total ocupado* a los efectos de su comparación.

Variable	Obs.	Media	Desv. Est.	Mínimo	Máximo
OP	30338	10.5941	0.8510	2.9940	17.6388
OP_tr	30338	10.2788	0.8509	2.6113	17.3680
LP	30338	10.7782	0.8500	3.2401	17.8156
LP_tr	30338	10.7096	0.8492	3.1388	17.6853
WRD	30338	10.5516	0.8498	3.0000	17.7520
WRD_tr	30338	10.4207	0.8687	2.7243	17.5936
AFC	30338	10.8691	0.8512	3.3457	17.8639
AFC_tr	30338	10.0297	0.8560	2.4381	17.5495
Vab/L	30338	13.0990	0.9874	5.7962	20.4041

Cuadro 8: Estadísticas descriptivas

Como se observa de la salida anterior, el valor medio de la PTF a nivel de empresa estimada por los diferentes métodos econométricos está en un nivel similar, y sucede lo mismo con su desvío estándar, tanto si incluimos una tendencia como si no lo hacemos. En tanto que la productividad laboral muestra tanto un nivel promedio como una variación mayor en comparación de los métodos de estimación presentados.

¹⁷ No se cuenta en nuestra fuente de información con datos para la *edad* de la empresa.

	OP	OP_tr	LP	LP_tr	WRD	WRD_tr	AFC	AFC_tr	Vab/L
OP	1.0000								
OP_tr	0.9963	1.0000							
LP	0.9943	0.9896	1.0000						
LP_tr	0.9932	0.9923	0.9988	1.0000					
WRD	0.9933	0.9903	0.9991	0.9981	1.0000				
WRD_tr	0.9606	0.9753	0.9683	0.9793	0.9686	1.0000			
AFC	0.9921	0.9867	0.9998	0.9985	0.9985	0.9682	1.0000		
AFC_tr	0.9794	0.9832	0.9887	0.9900	0.9938	0.9716	0.9876	1.0000	
Vab/L	0.8376	0.8198	0.8850	0.8830	0.8747	0.8594	0.8950	0.8551	1.0000

Cuadro 9: Matriz de correlaciones

La estimación de la PTF utilizando diferentes técnicas está altamente correlacionada. Esto quiere decir que existe una relación lineal muy alta y en la misma dirección para la PTF a nivel de empresa para cada uno de los métodos. La productividad laboral muestra una correlación más baja en comparación con las técnicas utilizadas, no obstante, se trata de una correlación elevada, resultados que ya fueron encontrados en el trabajo de Adriana Peluffo (2022). Los coeficientes estimados de la función de producción por las diferentes estrategias son los siguientes:

	Capital	Trabajo	Trend	Obs.
OP	0.2687*** (0.0353)	0.5566*** (0.0138)	-	24664
OP_tr	0.2840*** (0.0465)	0.5532*** (0.0138)	0.0274 (0.0263)	24664
LP	0.2343*** (0.0275)	0.6494*** (0.0123)	-	15206
LP_tr	0.2354*** (0.0278)	0.6477*** (0.0156)	0.0178 (0.0248)	15206
WRD	0.2467*** (0.0154)	0.6587*** (0.0099)	-	8772
WRD_tr	0.2346*** (0.0148)	0.6630*** (0.0099)	0.0927 (0.2149)	8772
AFC	0.2505*** (0.0000)	0.6896*** (0.0000)	-	15206
AFC_tr	0.2660*** (0.0000)	0.7048*** (0.0000)	0.0147*** (0.000)	15206

Cuadro 10: Estimaciones de los coeficientes de la función de producción.

*** si son significativas al 1 %. Desvíos estándar entre paréntesis.

Fuente: Elaboración propia en base a datos del INE.

Si creemos que no existieron shocks específicos para cada empresa (que impacten a cada empresa en forma diferente) del capital físico, mano de obra y/o consumo energético, entonces las estimaciones de OP son consistentes. Tienen la ventaja de que las mismas se lograron con una de mayor cantidad de datos que los otros métodos de estimación (24664 observaciones). Por otro lado, si creemos que pudieron haber existido shocks específicos para cada empresa asociados al capital físico (inversión) entonces podríamos pensar que OP es inconsistente, y entonces utilizar en su lugar las estimaciones de LP puesto que serían consistentes en la presencia de shocks asociados a la inversión. En este caso se utilizó el consumo energético como bien intermedio, donde para la estimación de estos coeficientes se usaron 15206 observaciones. Esto se debe a que no se cuenta con el reporte de este dato para algunas empresas a lo largo de nuestro periodo, sin embargo, esta omisión es aleatoria en nuestros datos.

Continuando con el razonamiento, si creemos que la elección del consumo energético está determinada por la productividad de la empresa y el capital físico, entonces podríamos tener un problema de colinealidad en la primera etapa del método de LP. En dicho caso, podríamos optar por la estrategia WRD, donde producto de su estrategia de estimación se pierden más datos, aunque se basa en el consumo energético al igual que LP. Estas estimaciones también son consistentes cuando se tienen shocks específicos de la inversión, sin embargo, y a pesar de resolver el problema de identificación, si tenemos motivos para creer que existieron en el periodo shocks específicos a cada empresa asociado a la mano de obra entonces estas estimaciones (y las anteriores) serán inconsistentes. Las estimaciones que son consistentes con estos shocks específicos de la mano de obra (además de shocks para la inversión) son las estimaciones de AFC. Estas estimaciones también se basan en el consumo energético y se obtienen con 15206 observaciones.

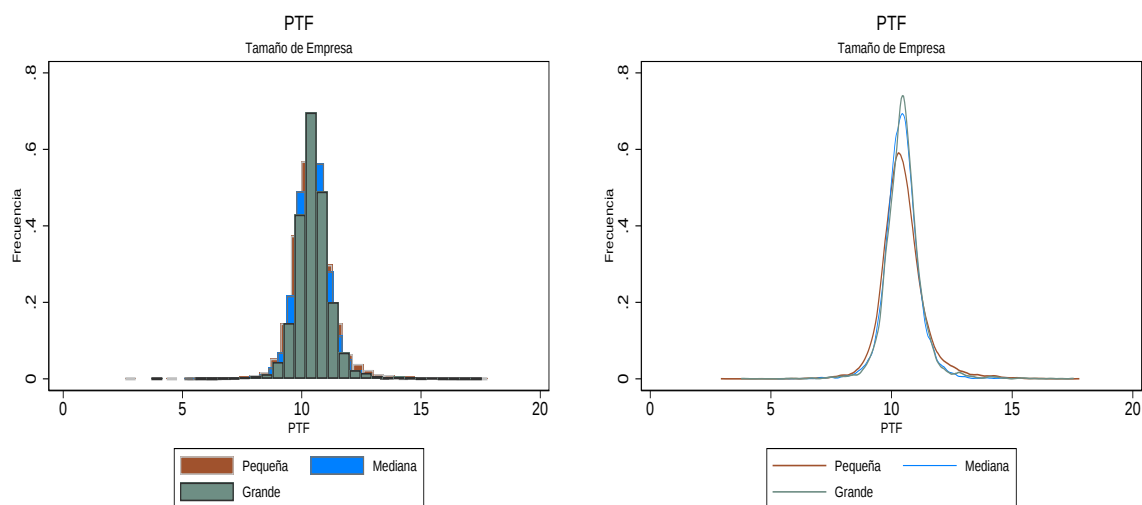
Cuando analizamos las estimaciones puntuales, comenzando con el coeficiente del trabajo, vemos que están en el orden de 0.6404, con un desvío para las mismas de 0.0563. El método de OP obtuvo su estimación con una mayor cantidad de información (utilizando la inversión) y es el que asigna al trabajo el valor más bajo, en el orden de 0.5532. Mientras que, utilizando el consumo energético en lugar de la inversión, los restantes métodos dan valores del coeficiente del trabajo más altos. El método que teóricamente logra obtener estimaciones consistentes bajo una mayor variedad de procesos generadores de datos, AFC, muestra el valor más elevado para el coeficiente del trabajo, con 0.7048. En este sentido, si tenemos motivos para creer que en el periodo existieron shocks específicos a nivel de empresa asociados a la mano de obra, entonces deberíamos quedarnos con este método. Del lado del coeficiente del capital físico las estimaciones puntuales están en el orden de 0.2525 con un desvío promedio para las mismas de 0.0186. En este caso, todos los métodos dan estimaciones puntuales más cercanas entre sí.

Con respecto a la tendencia, el único método que arrojó como resultado una tendencia significativa es el de AFC. Sin embargo, debe mencionarse que se ha estimado la misma especificación en más de una oportunidad y si bien los valores estimados del capital y el trabajo no se mueven, esto no ocurre con la tendencia. En ocasiones da significativa, pero en otras no. Los polinomios de tercer orden que deben ser estimados y el método de bootstrap que subyacen a la estimación podrían estar incluyendo en el coeficiente de la tendencia que, con un valor puntual estimado cercano a cero, es sensible a pequeñas variaciones en la estimación. En resumen, de ahora en adelante **trabajaremos con las estimaciones de la PTF basados en el método AFC que incluye la tendencia, *AFC_tr*.**

A continuación, se presenta la distribución de la mencionada PTF según diferentes tamaños de empresa. Se aprecia una mayor dispersión de la PTF para empresas pequeñas en relación con las medianas o grandes. Por otra parte, las empresas grandes están levemente corridas hacia la derecha en cuanto a su media respecto a las medianas y pequeñas empresas. No obstante, las tres distribuciones presentan un comportamiento similar.

Por otra parte, se presentan el promedio y la mediana para la PTF a nivel de empresa por sectores de la economía uruguaya. En este caso, dada las características de la distribución se espera que el promedio y la mediana estén en niveles similares para cada uno de los sectores. Esto es lo que sucede si comparamos ambas gráficas observando sector por sector. Los sectores más productivos son los de *Actividades Financieras, Información y Comunicación, Administración Pública y Defensa* y el sector de *Comercio al por mayor y al por menor*. Del lado de los sectores menos productivos aparecen los de *Alojamiento y Servicio, Enseñanza, Otras actividades de servicios* y el sector de *Artes y entretenimiento*. Nuevamente debe tenerse presente la heterogeneidad de

divisiones de la CIIU revisión 4 que componen al sector de la *Industria Manufacturera*, el cual queda a mitad de camino entre las más y las menos productivas.



(a) Gráfica 1

(b) Gráfica 2

Figura 6: Distribución de la PTF según tamaño de empresa

Fuente: Elaboración propia.

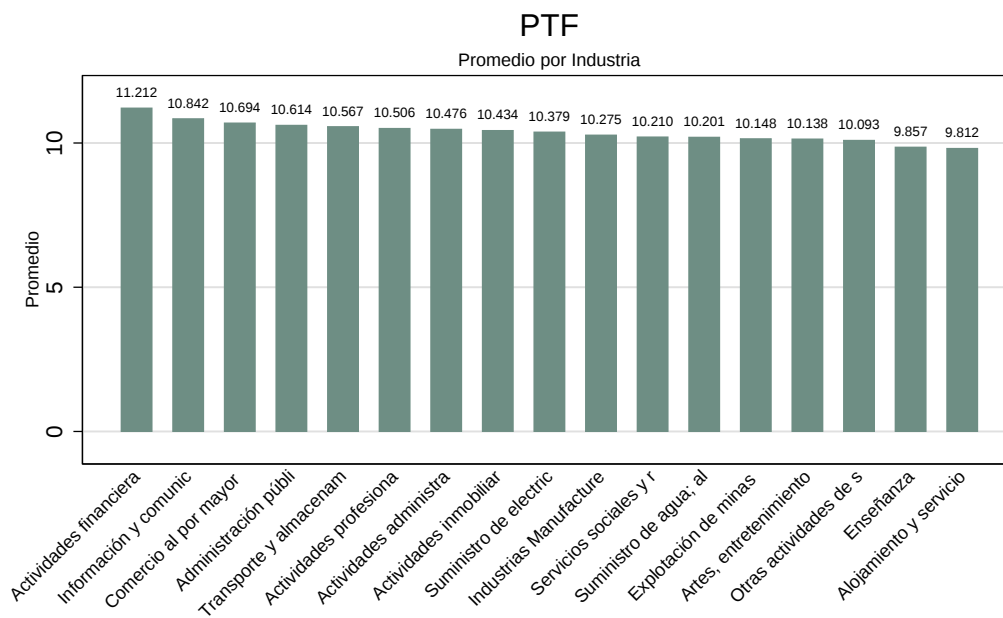


Figura 7: Promedio de la PTF según sector

Fuente: Elaboración propia.

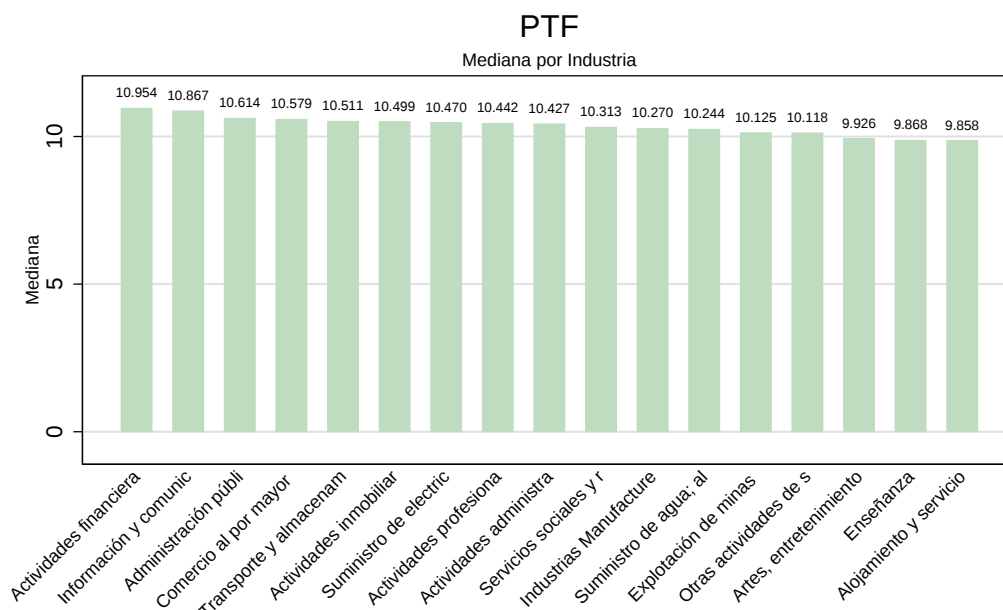


Figura 8: Mediana de la PTF según sector

Fuente: Elaboración propia.

5.3 EXP: un análisis descriptivo

La tercera variable clave del SEMs a estimar son las exportaciones. En particular utilizaremos el logaritmo natural de las *exportaciones deflactadas*. En la siguiente página se presenta un gráfico de la misma en función del tamaño de la empresa.

Las densidades de las exportaciones no parecen tener grandes cambios en función del tamaño de la empresa comprendidas en el periodo 2012-2019. Sí se observa menor densidad para las empresas que no exportan cuando son más grandes en relación con las pequeñas empresas. También observamos valores levemente mayores en la parte positiva de la distribución de las grandes con respecto a las pequeñas empresas.

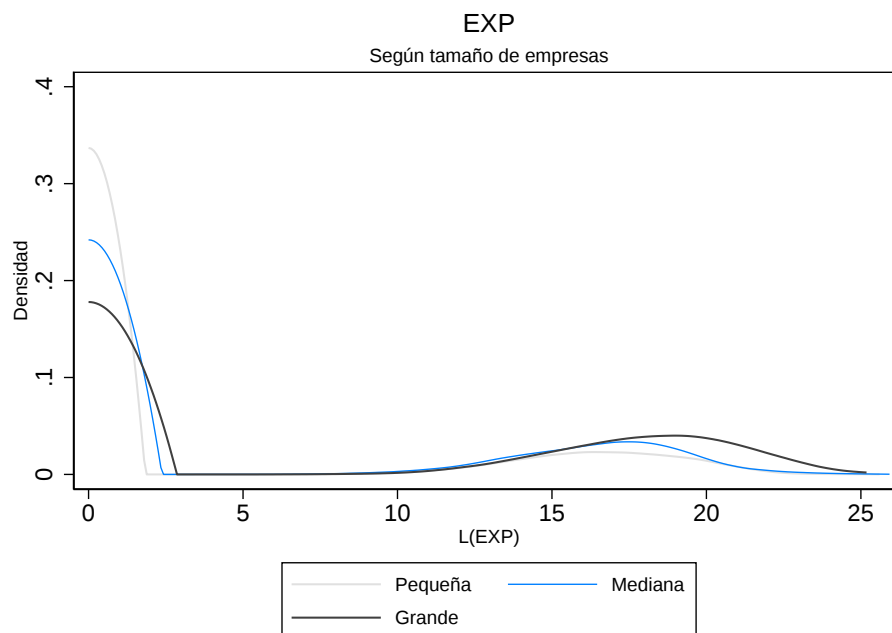


Figura 9: Distribución de las exportaciones según tamaño de empresa

Fuente: Elaboración propia.

A continuación, presentamos el promedio de las exportaciones en función de cada sector industrial.¹⁸

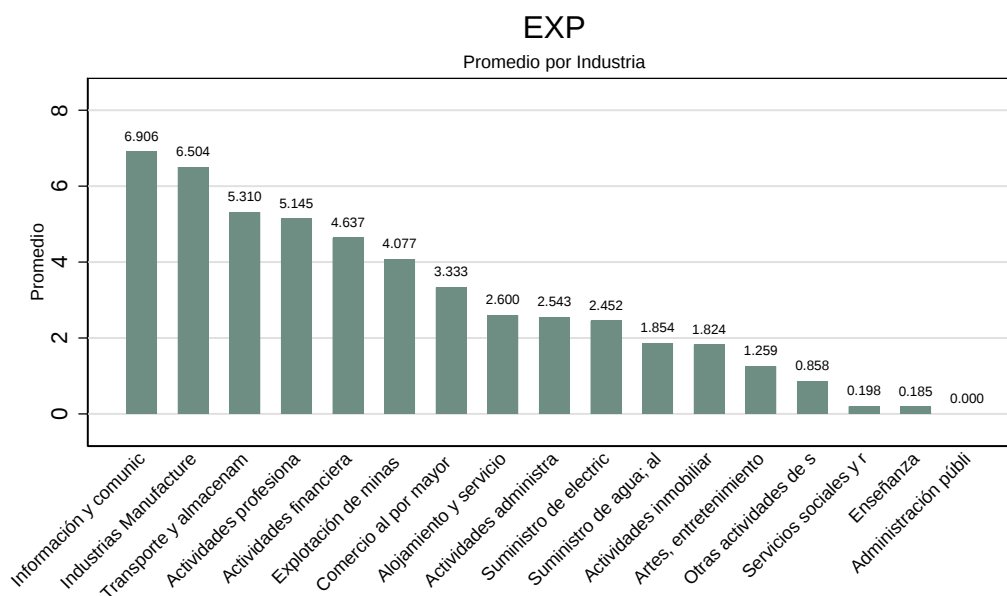


Figura 10: Promedio de las exportaciones según sector

Fuente: Elaboración propia.

¹⁸ No se reporta el gráfico para la mediana por sector industrial debido a que su valor es nulo para cada una de ellas. En la muestra existe elevada cantidad de empresas que no exportan.

Los sectores que en promedio presentaron mayor nivel de exportaciones en el periodo 2012-2019 son los sectores de *Información y Comunicación* y la *Industria Manufacturera* en su conjunto. Los siguen los sectores de *Transporte y almacenamiento*, *Actividades profesionales* y *Actividades financieras* respectivamente. A continuación, pasamos a los resultados del estimar el sistema de ecuaciones simultáneas (SEMs).

5.4 SEMs: resultados

A continuación, se presentan los resultados de estimar (8) por el método de GMM-OWM:

Variable	eq: TICs	eq: PTF	eq: ln(EXP)
PTF	1.2049*** (0.0608)		3.3281*** (0.1976)
TICs		0.2863*** (0.0152)	1.4211*** (0.1050)
ln(EXP)			
ln(tamaño)	0.5875*** (0.0253)	-0.1290*** (0.01265)	0.0834 (0.0787)
cal_emp_intsec	-0.0095 (0.0100)	0.0227*** (0.0044)	
ln(i + d)	0.0081* (0.0042)	0.0220*** (0.0022)	0.0983*** (0.0216)
ln(apalancamiento)	0.1800*** (0.0298)		
ln(roe)	-0.8529*** (0.0829)		
cuota_mercado		0.7667*** (0.0396)	
hhi		-0.2304*** (0.0394)	
ln(stock_capital)		-0.0947*** (0.0060)	
capital_ext			0.0499*** (0.0022)
industria1	-0.4909*** (0.0937)	-0.2690*** (0.0386)	-1.3818*** (0.3019)
industria2	-0.5574*** (0.1259)	0.2123*** (0.0612)	-3.7826*** (0.3602)
industria3	0.5599*** (0.1224)	-0.3502*** (0.0687)	-2.6667*** (0.4977)
industria4	0.4736*** (0.0840)	-0.4775*** (0.0398)	-0.0701 (0.3343)
industria6	-0.4524*** (0.1616)	-0.5204*** (0.0575)	2.5715*** (0.3725)

Cuadro 11 – continuación

Variable	eq: TICs	eq: PTF	eq: ln(EXP)
industria7	0.1690 (0.1141)	-0.5036*** (0.0561)	-1.7652*** (0.3687)
industria8	-0.4687*** (0.0800)	0.0166 (0.0372)	-1.4350*** (0.2553)
industria9	0.6507*** (0.0991)	-0.8837*** (0.0393)	-2.4220*** (0.3056)
industria10	0.3044 (0.2074)	-0.5638*** (0.0828)	1.7923** (0.7580)
industria11	-0.4996*** (0.0849)	-0.2154*** (0.0401)	3.4936*** (0.2805)
industria13	-0.1500 (0.1672)	-0.5050*** (0.0588)	-1.0777*** (0.3559)
industria14	-0.5766*** (0.1064)	-0.2338*** (0.0425)	-2.8015*** (0.2961)
industria15	-0.1994 (0.2658)	-0.4220*** (0.0861)	-1.0272* (0.6186)
industria16	-0.1596 (0.6460)	0.1193 (0.2357)	-2.8185** (1.3882)
industria17	-0.3880*** (.0929)	-0.0166 (0.0418)	1.2377*** (0.3054)
anio2013	0.0872 (0.0734)	0.0716*** (0.0256)	-0.6611*** (0.1745)
anio2014	0.0281 (0.0751)	0.1064*** (0.0259)	-0.6547*** (0.1738)
anio2015	0.3387*** (0.0736)	-0.0525** (0.0256)	-0.8934*** (0.1797)
anio2016	0.4262*** (0.0732)	-0.0961*** (0.0261)	-0.8372*** (0.1843)
anio2017	0.5752*** (0.0709)	-0.1165*** (0.0263)	-1.1747*** (0.1878)
anio2018	0.6927*** (0.0676)	-0.1188*** (0.0265)	-1.7445*** (0.1888)
anio2019	0.8551*** (0.0652)	-0.1549*** (0.0273)	-2.2071*** (0.1954)
_cons	-14.5921*** (0.6780)	12.5539*** (0.1002)	-31.2017*** (2.2826)

Cuadro 11: Resultados de la estimación por GMM-OWM

Notas: Errores estándar entre paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Número de observaciones = 26916

Fuente: Elaboración propia.

Respecto a las variables binarias que capturan el efecto temporal agregado se deja al *año 2012* como referencia. En tanto que para las variables binarias que capturan los efectos agregados correspondientes a distintas industrias se elimina la variable industria correspondiente al sector de Administración Pública y Defensa ya que solamente consistía en una sola empresa para todo el periodo considerado, y se deja como categoría de referencia al sector de *Información y Comunicación*. El siguiente cuadro ayuda a aclarar a que industria están representando cada una de las variables binarias del efecto industria:

Cuadro de correspondencia

Industria10 = B. Explotación de minas y canteras
Industria11 = C. Industrias Manufactureras
Industria16 = D. Suministro de electricidad, gas, vapor y aire acondicionado
Industria15 = E. Suministro de agua; alcantarillado, gestión de desechos y actividades de saneamiento
Industria 8 = G. Comercio al por mayor y al por menor; reparación de vehículos de motor y de las motocicletas
Industria 17 = H. Transporte y almacenamiento
Industria6 = I. Alojamiento y servicio de comida
Industria12 = J. Información y comunicación
Industria2 = K. Actividades financieras y de seguros
industria3 = L. Actividades inmobiliarias
Industria4 = M. Actividades profesionales, científicas y técnicas
Industria1 = N. Actividades administrativas y servicios de apoyo
Industria9 = P. Enseñanza
Industria14 = Q. Servicios sociales y relacionados con la Salud Humana
Industria7 = R. Artes, entretenimiento y recreación
Industria13 = S. Otras actividades de servicios

Cuadro 12: Secciones de Actividad y correspondencia con efecto industria

Fuente: Elaboración propia

La estimación del SEMs permite analizar de forma estructurada las relaciones entre el uso de tecnologías de la información y la comunicación (TICs), la productividad total de los factores (PTF) y el desempeño exportador medido a través de las exportaciones.

En la primera ecuación, donde el índice de adopción de TICs es la variable dependiente, medido como índice, el principal interés recae en el efecto de la Productividad Total de los Factores, que está medido en logaritmo natural. El coeficiente estimado para esta variable es de 1.2049. Al analizar su magnitud, se tiene que ante un aumento en un 1 % para la PTF se espera que aumente el índice de TICs en 0.012 unidades, controlando por el resto de factores. Si bien se encuentra una asociación positiva y significativa entre mayores niveles de productividad y una mayor adopción de tecnologías digitales dada por el uso de TICs, la magnitud de la dirección que va desde la PTF hacia las TICs parece tener un impacto reducido, al tener presente que el aumento de 0.012 unidades representa apenas 0.004 desvíos estándar del índice de TICs.

En la segunda ecuación se modela directamente a la PTF. El coeficiente correspondiente a las TICs es de 0.2863. Esto se traduce en que ante un aumento en 1 unidad para el índice TICs — el cual representa 0.3328 desvíos estándar del índice— se asocia con un aumento en la PTF del 33.149 %, lo que respalda la hipótesis de que la incorporación de tecnologías contribuye a mejorar la eficiencia productiva, incluso luego de controlar por otros factores relevantes. Con lo cual, la dirección del impacto que va desde TICs hacia la productividad es relevante.

Finalmente, en la tercera ecuación, donde la variable dependiente es el logaritmo natural del valor de las exportaciones, los efectos de interés son tanto el de las TICs como el de la PTF.

El coeficiente asociado a las TICs es de 1.4211, mientras que el de la PTF asciende a 3.3281. Si analizamos las magnitudes de estos efectos tenemos que ante un aumento en 1 unidad del índice de TICs —0.3328 desvíos estándar del índice— se espera que incremente el logaritmo natural de las exportaciones en 1.4211, lo que se traduce en un aumento esperado del nivel de las exportaciones en 314.167 %, controlando por los demás factores. Por lo tanto, las TICs tienen efectos directos sobre las exportaciones y como vimos en la sección (2), pueden estar asociados a mecanismos de reducción de los costos del comercio, en el acceso a conocimiento e información. En tanto que ante un incremento en un 1 % de la PTF se asocia con un aumento en las exportaciones de un 3.328 %, *ceteris paribus*. Todo este análisis sugiere que ambas dimensiones —tecnológica y productiva— tienen un rol significativo en el desempeño exportador. Estos resultados son consistentes con la hipótesis de que la digitalización y la productividad actúan como canales complementarios para mejorar la competitividad internacional de las firmas.

De esta manera, se pudo constatar que las TICs impactan tanto directamente en las exportaciones (tercera ecuación) como indirectamente a través de la PTF (segunda y tercera ecuación), respaldando las hipótesis iniciales.

Pasando al análisis de los determinantes de cada una de estas ecuaciones estructurales, tenemos que los resultados obtenidos para el índice de adopción de TICs revelan efectos heterogéneos según las características de las firmas. En primer lugar, se observa que un incremento del 1 % en el tamaño de la firma, medido por el número de personas ocupadas, se asocia con un aumento estimado de 0.0059 unidades en el índice de TICs, lo que sugiere una débil pero positiva relación entre el tamaño y la incorporación de tecnologías digitales. Por otro lado, la variable que relativiza el salario medio de la firma en relación con el salario medio de su sector (*cal_emp_intsec*) no presenta efectos significativos. El gasto en investigación y desarrollo (I+D), por su parte, presenta un coeficiente positivo pero de magnitud muy reducida: un aumento del 1 % en I+D se asocia con un incremento de apenas 0.00008 unidades en el índice. Además, este efecto alcanza significancia estadística al 10 %, pero no al 5 %. En cuanto a la estructura financiera, se encuentra que un aumento del 1 % en el apalancamiento se relaciona con una mejora de 0.0018 unidades en el índice de TICs, si bien la asociación positiva entre estas variables podría indicar que empresas más proctivas hacia la adopción de tecnologías se estén financiando a través de deuda, lo cierto es que la magnitud de la relación es débil. Finalmente, el coeficiente asociado a la rentabilidad sobre el ejercicio (ROE) indica que una mayor rentabilidad se vincula con una caída en la adopción de TICs de 0.0085 unidades, sugiriendo que las firmas más rentables podrían no estar destinando esos excedentes a procesos de transformación digital.

Los determinantes de la Productividad Total de los Factores muestran un panorama diferente al observado para el índice de TICs. En este caso, el tamaño de la firma exhibe una relación negativa: un aumento del 1 % para el tamaño se asocia con una reducción del 0.1290 % en la PTF, manteniendo constantes las demás variables. En contraste, mayores diferencias salariales (*cal_emp_intsec*) se asocian positivamente con la productividad: un aumento de una unidad en esta variable se traduce en un incremento del 2.296 % en la PTF. Pero ha de tenerse en cuenta que un incremento en una unidad para esta medida representaría una mejora sustancial en la relación del salario medio de la firma respecto del salario medio de su sector. De esta forma, ante un incremento moderado de 0.10 unidades, se lo asocia con un incremento de 0.227 % en la PTF. Asimismo, el gasto en I+D presenta una relación positiva pero muy reducida: un aumento del 1 % se vincula con una mejora del 0.022 %. Al considerar la cuota de mercado, se tiene que ante un aumento de 1 unidad en la proporción de cuota —lo cual implicaría una expansión sustancial— se asocia con un incremento del 115.27 % en la PTF. Pero incluso aumentos moderados en la

cuota son considerables, por ejemplo, ante un incremento de 0.10 unidades se asocia con un incremento de 7.97 % en la PTF. Esto implica que las firmas con mayor cuota de mercado son más eficientes. Por el contrario, la concentración de mercado (medida mediante el índice HHI) muestra un efecto negativo significativo: un aumento de una unidad en este índice se traduce en una reducción del 20.58 % en la productividad. Ante un incremento más moderado, por ejemplo, de 0.10 unidades, se asocia con una reducción del 2.27 % en la PTF. Esto implica que cuando los mercados se concentran y existe menos competencia entre las firmas, la eficiencia productiva es menor. Por último, el stock de capital exhibe un coeficiente negativo: un aumento del 1 % en esta variable se relaciona con una leve caída del 0.0947 % en la PTF, lo cual podría estar reflejando dinámicas de sobreacumulación de capital o ineficiencias en su uso.

En relación con el desempeño exportador, se observa que el tamaño de la firma no tiene un efecto significativo en el nivel de exportaciones. El gasto en I+D muestra nuevamente un efecto positivo, esta vez más marcado: un incremento del 1 % en esta variable se asocia con un aumento del 0.0983 % en las exportaciones, resultado consistente con la idea de que las actividades de innovación refuerzan la competitividad internacional. Por último, para la participación de capital extranjero dentro de la firma, se tiene que ante un aumento de una unidad (es decir, pasar del 0 % al 100 % de capital extranjero) se asocia con un incremento del 5.12 % en el nivel de exportaciones. De manera más moderada, un aumento de 10 puntos porcentuales en esta participación (por ejemplo, pasar del 20 % al 30 %) se vincula con un aumento cercano al 0.50 % en las exportaciones, lo que sugiere que incluso cambios parciales en la estructura de propiedad pueden tener implicancias relevantes en el acceso a mercados internacionales.

Los coeficientes de los variables binarias de año capturan el efecto agregado que cada año tiene sobre la variable dependiente, comparado con 2012, manteniendo fijas todas las demás covariables. En la ecuación de TICs, los coeficientes para 2015–2019 son positivos y crecientes, lo que indica que, *ceteris paribus*, el nivel del índice TICs en dichos años está por encima de su valor en 2012. En contraste, en las ecuaciones de PTF y $\ln(\text{EXP})$ los coeficientes de 2015–2019 son negativos, lo que significa que, *ceteris paribus*, en esos años la PTF y las exportaciones están por debajo de sus valores en 2012. Esta interpretación se refiere estrictamente a los efectos temporales agregados, no al comportamiento bruto de las series a lo largo del tiempo. Estos efectos están capturando todo lo que varía en el tiempo e impacta en forma común (transversal) en todas las firmas de la muestra.

Utilizando el cuadro (12) como ayuda, podemos decir que, si una empresa pase de estar en algunas de las industrias 1, 2, 6, 8, 11, 14 y 17 a formar parte de la industria de Información y Comunicación implicaría un aumento en el nivel de las TICs, manteniendo todo lo demás constante. En tanto que para las industrias 7, 10, 13, 15 y 16 no hay cambios significativos respecto a la industria de referencia. Por otro lado, para las empresas que están en las industrias 3, 4 y 9 pasar de estas industrias a formar parte de la industria de Información y Comunicación implica un empeoramiento en su nivel de TICs, manteniendo todo lo demás constante. En cuanto a la productividad, las firmas que pertenecen a las industrias 8, 16 y 17 no presentan cambios con respecto a la industria de Información y Comunicación, mientras que para las firmas que pertenecen a la industria 2 pasar de esta industria a formar parte de la industria de Información y Comunicación implica un empeoramiento en su nivel de PTF, manteniendo todo lo demás constante. Por el contrario, para las firmas del subconjunto restante de industrias, pasar de formar parte de alguna de ellas a formar parte del sector de Información y Comunicación implicaría un aumento en su nivel de productividad, manteniendo todo lo demás constante. Se puede hacer una interpretación similar con la ecuación de las exportaciones.

6 Conclusiones

Dentro de los principales hallazgos del presente trabajo destacan las relaciones causales¹⁹ positivas entre las TICs y la productividad. Si bien los coeficientes en ambas direcciones son positivos y significativos, la dirección que va desde la PTF hacia las TICs es débil mientras que el efecto que va en la dirección desde TICs hacia PTF es fuerte. Por lo que, a priori, estimar estas tres ecuaciones en forma secuencial podría arrojar estimaciones similares a las encontradas en el presente trabajo, donde se utilizó una estrategia de ecuaciones en simultáneo, aunque ello sin considerar que el método de estimación —GMM— incorpora las posibles correlaciones cruzadas entre las ecuaciones del sistema.

Dentro de los principales resultados, destaca el hecho de que tanto las TICs como la PTF tienen un impacto directo y positivo sobre las exportaciones de las firmas. Lo que implica que tanto la dimensión tecnológica como la productiva actúan como canales que favorecen la competitividad internacional de las empresas. Más aún, las TICs no sólo impactan de forma directa en las exportaciones, sino que al impactar en la PTF, también logran impactar sobre las exportaciones vía productividad, esto es, en forma indirecta. Entonces adquiere especial relevancia como recomendación de política el impulso a la adopción de TICs a nivel de firmas y por sobre todo, con particular énfasis sobre aquellos sectores o industrias que presentan menores niveles de adopción de TICs.²⁰

Pasando a los determinantes del sistema, cuanto mayor es el tamaño de las firmas estas se asocian con una mayor adopción TICs mientras que, por otro lado, se asocian a una menor eficiencia productiva. Además, no parecen tener una asociación con respecto al desempeño exportador. Respecto al gasto en I+D se encuentra que sus efectos sobre la adopción TICs y/o la eficiencia productiva, si bien es positiva en ambas ecuaciones, su efecto es muy cercano a cero. Sin embargo, se encuentra que un mayor gasto en investigación y desarrollo tiene efectos positivos en el desempeño exportador, por lo que ayudar a las empresas a que sean capaces de invertir en I+D puede tener efectos positivos en la economía a través de un incremento en la capacidad exportadora.

De los determinantes de las TICs se desprende que empresas de menor tamaño poseen un menor nivel de adopción de TICs y, por otra parte, empresas con mayor apalancamiento financiero disponen de mayor nivel de adopción de TICs, por lo que surge como otra recomendación de política la ayuda a las PYMEs en el acceso del financiamiento para la adopción de TICs.

Asimismo, de los determinantes de la PTF, los resultados sugieren otra recomendación de política en el sentido de fomentar la competencia dentro de cada industria. En este sentido, las firmas que pertenecen a industrias con menor concentración de mercado son firmas con mayor eficiencia productiva, seguramente porque factores como la presión competitiva impulsan a las firmas a innovar, tanto en procesos, productos, como a nivel de organización.

Por otro lado, aprovechar las redes, contactos, y conocimiento que puede extraerse de la presencia de empresas transnacionales puede ayudar a las firmas en su experiencia de internacionalización y mejora de la capacidad competitiva, lo cual puede incrementar sus exportaciones. Por

¹⁹ Si entendemos que cada ecuación es autónoma. Ver discusión en Wooldridge, 2010.

²⁰ En la figura (4) se presentó el promedio de TICs por industria, mientras que las industrias se presentaron en el cuadro (2).

lo tanto, otra recomendación va en la dirección de una política que permita aprovechar derrames de las empresas transnacionales fomentando sinergias, vínculos, redes y colaboraciones entre estas y las empresas nacionales, donde potenciales sinergias pueden mejorar la capacidad de internacionalizarse de las firmas nacionales.

En cuanto al capital humano, si bien no contamos con una medida de la calidad del empleo que mida las habilidades en TICs, que los salarios más elevados en relación con los salarios promedios de su correspondiente sector impliquen un aumento en la eficiencia productiva, nos da indicios de que un mayor fomento de capacitación del capital humano puede reflejarse en un incremento de la productividad, por lo que esta es otra recomendación de política que se desprende de los resultados obtenidos de la estimación del SEMs.

Algunas de las limitaciones que se enfrentan en el presente trabajo están muy relacionadas a los datos. Se mencionó que no se cuenta con una medida concreta para medir las habilidades en TICs, anteriormente la encuesta con la que se trabaja en este documento relevaba la proporción de *profesionales y técnicos* en relación de dependencia con respecto al personal total ocupado, pero dejó de relevarse en el periodo considerado. La literatura sugiere que esta medida podría utilizarse como una medida de la calidad del empleo y posiblemente se hubiese incluido en el SEMs tanto en la ecuación de las TICs como en la de la PTF. Por otro lado, otra variable que hubiese sido positivo contar con ella es la *edad* de la empresa. En la literatura se suele utilizar a la hora de estimar la Productividad Total de los Factores, pero no se pudo contar con ella para el periodo que se abarca en este trabajo. En tanto que, con respecto al índice creado de TICs, si bien sería pertinente que en el futuro, para periodos posteriores la encuesta pueda relevar acerca del uso de la Inteligencia Artificial, el uso de robots en los procesos de producción, y otras medidas como la Computación en Nube, lo cual puede enriquecer el análisis o reflejar un concepto más amplio de adopción digital, lo cierto es que se tienen motivos para creer que el índice de TICs creado refleja, en el *score* final del Primer Componente principal, el grado de adopción tecnológica de las firmas uruguayas en el período considerado y su heterogeneidad. Con lo cual, si bien en este trabajo cuando hablamos del índice de TICs nos referimos a la adopción de tecnologías de la información y de la comunicación, es plausible que el *score* resultante refleje también el grado de adopción de otras formas de tecnología digital como las mencionadas.

Por último, en cuanto a la agenda de investigación, debe mencionarse que, en esta oportunidad, por razones de espacio, no se incluyó en el análisis la hipótesis de *learning by exporting* o aprendizaje a través de las exportaciones, que sugiere que hay un canal que va desde las exportaciones hacia la PTF. Con lo cual, una posible extensión natural de este trabajo es rehacerlo incorporando dicha relación en el modelo de ecuaciones simultáneas. Así como también la posibilidad de trabajar con un período aún más reciente conforme el tiempo transcurra y como se mencionó previamente, quizás con la posibilidad de medir un concepto de adopción digital más amplio que el del presente documento.

Referencias

- Acemoglu, D., Aghion, P., & Zilibotti, F. (2006). Distance to Frontier, Selection, and Economic Growth. *Journal of the European Economic Association*, 4(1), 37-74. <https://doi.org/10.1162/jeea.2006.4.1.37>
- Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, 83(6), 2411-2451. <https://doi.org/10.3982/ECTA13408>
- Adriana Peluffo. (2022). El retrato del éxito: empresas uruguayas en el comercio internacional. *Revista de la CEPAL*, 2022(138), 193-220. <https://doi.org/10.18356/16820908-2022-138-9>
- Alcácer, J., Cantwell, J., & Piscitello, L. (2016). Internationalization in the information age: A new era for places, firms, and international business networks? *Journal of International Business Studies*, 47(5), 499-512. <https://doi.org/10.1057/jibs.2016.22>
- Andrews, D., Nicoletti, G., & Timiliotis, C. (2018, julio). *Digital technology diffusion: A matter of capabilities, incentives or both?* (OECD Economics Department Working Papers N.º 1476) (Series: OECD Economics Department Working Papers Volume: 1476). OCDE. <https://doi.org/10.1787/7c542c16-en>
- Bartelsman, E., Van Leeuwen, G., & Polder, M. (2017). CDM using a cross-country micro moments database. *Economics of Innovation and New Technology*, 26(1-2), 168-182. <https://doi.org/10.1080/10438599.2016.1202517>
- Caliendo, L., & Parro, F. (2015). Estimates of the Trade and Welfare Effects of NAFTA. *The Review of Economic Studies*, 82, 1-44.
- Charlo, G. (2011). IV. Impact of ICT and innovation on industrial productivity in Uruguay. *CEPAL*.
- Crepon, B., Duguet, E., & Mairessec, J. (1998). Research, Innovation And Productivity: An Econometric Analysis At The Firm Level. *Economics of Innovation and New Technology*, 7(2), 115-158. <https://doi.org/10.1080/10438599800000031>
- DeStefano, T., Kneller, R., & Timmis, J. (2023). The (fuzzy) digital divide: the effect of universal broadband on firm performance. *Journal of Economic Geography*, 23(1), 139-177. <https://doi.org/10.1093/jeg/lbac006>
- Dutot, V., Bergeron, F., & Raymond, L. (2014). Information management for the internationalization of SMEs: An exploratory study based on a strategic alignment perspective. *International Journal of Information Management*, 34(5), 672-681. <https://doi.org/10.1016/j.ijinfomgt.2014.06.006>
- Eaton, J., & Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, 70(5), 1741-1779. <https://doi.org/10.1111/1468-0262.00352>
- Gal, P., Nicoletti, G., Renault, T., Sorbe, S., & Timiliotis, C. (2019, febrero). *Digitalisation and productivity: In search of the holy grail – Firm-level empirical evidence from EU countries* (OECD Economics Department Working Papers N.º 1533) (Series: OECD Economics Department Working Papers Volume: 1533). OCDE. <https://doi.org/10.1787/5080f4b6-en>
- Gbadegeshin, S. A., Oyelere, S. S., Olaleye, S. A., Sanusi, I. T., Ukpabi, D. C., Olawumi, O., & Adegbite, A. (2019). Application of information and communication technology for internationalization of Nigerian small- and medium-sized enterprises. *The Electronic Journal of Information Systems in Developing Countries*, 85(1), e12059. <https://doi.org/10.1002/isd2.12059>

- Hervé, A., Schmitt, C., & Baldegger, R. (2020). Internationalization and Digitalization: Applying digital technologies to the internationalization process of small and medium-sized enterprises. *Technology Innovation Management Review*, 10(7), 28-40. <https://doi.org/10.22215/timreview/1373>
- Kyläheiko, K., Jantunen, A., Puumalainen, K., Saarenketo, S., & Tuppur, A. (2011). Innovation and internationalization as growth strategies: The role of technological capabilities and appropriability. *International Business Review*, 20(5), 508-520. <https://doi.org/10.1016/j.ibusrev.2010.09.004>
- Lecerf, M., & Omrani, N. (2020). SME Internationalization: the Impact of Information Technology and Innovation. *Journal of the Knowledge Economy*, 11(2), 805-824. <https://doi.org/10.1007/s13132-018-0576-3>
- Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2022). Valid t-Ratio Inference for IV. *American Economic Review*, 112(10), 3260-90. <https://doi.org/10.1257/aer.20211063>
- Lee, Y.-Y., Falahat, M., & Sia, B.-K. (2019). Impact of Digitalization on the Speed of Internationalization. *International Business Research*, 12(4), 1. <https://doi.org/10.5539/ibr.v12n4p1>
- Levinsohn, J., & Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies*, 70(2), 317-341. <https://doi.org/10.1111/1467-937X.00246>
- Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6), 1695-1725. <https://doi.org/10.1111/1468-0262.00467>
- Miozzo, M., & Soete, L. (2001). Internationalization of Services: A Technological Perspective. *Elsevier Science Inc.*
- Mol, M. J., & Koppius, O. R. (2002). Copyright Idea Group Inc. Information Technology and the Internationalization of the Firm. *Idea Group Publishing*.
- Molina, M., Rotondo, S., & Yoguel, G. (2013). IV. El impacto de las TIC en la productividad del trabajo: algunos indicios para las PyME del sector manufacturero argentino. *CEPAL*.
- OECD. (2009). Information Economy Product Definitions Based on the Central Product Classification (version 2) [Publisher: OECD]. *OECD Digital Economy Papers*, (158). <https://doi.org/10.1787/222222056845>
Series: OECD Digital Economy Papers Volume: 158.
- Olley, G. S., & Pakes, A. (1992, enero). *The Dynamics of Productivity in the Telecommunications Equipment Industry* (inf. téc. N.º w3977). National Bureau of Economic Research. Cambridge, MA. <https://doi.org/10.3386/w3977>
- Plottier, C., Rovira, S., & Stumpo, G. (2013). Una iniciativa sectorial para la difusión de las TIC en las empresas. *CEPAL*.
- Prashantham, S. (2005). Toward a Knowledge-Based Conceptualization of Internationalization. *Journal of International Entrepreneurship*, 3(1), 37-52. <https://doi.org/10.1007/s10843-005-0304-5>
- Sekliuckiene, J., & Maciulskaitė, S. (2013). Internationalization Decision of a Born Global: the Case of Information Technology Companies. *Social Sciences*, 80(2), 17-26. <https://doi.org/10.5755/j01.ss.80.2.4644>
- Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49(2), 326-365. <https://doi.org/10.1257/jel.49.2.326>
- Todd, P. R., & Javalgi, R. (G. (2007). Internationalization of SMEs in India: Fostering entrepreneurship by leveraging information technology (R. V. Eunni, Ed.). *International Journal of Emerging Markets*, 2(2), 166-180. <https://doi.org/10.1108/17468800710739234>

- Tomski, P. (2015). Information Technology and Firm Internationalization. *Applied Mechanics and Materials*, 795, 227-234. <https://doi.org/10.4028/www.scientific.net/AMM.795.227>
- Vasilchenko, E., & Morrish, S. (2011). The Role of Entrepreneurial Networks in the Exploration and Exploitation of Internationalization Opportunities by Information and Communication Technology Firms. *Journal of International Marketing*.
- Welch, L. S. (1996). Information behaviour and internationalization. *International Journal Technology Management*.
- Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3), 112-114. <https://doi.org/10.1016/j.econlet.2009.04.026>
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed) [OCLC: ocn627701062]. MIT Press.

7 Anexo

7.1 Antecedentes empíricos

Artículo	Período	Nivel	Método	Variable explicada	Predictor de interés	Interpretación	Efecto
DeStefano et al., 2023	2000-2004	Empresa	Efectos Fijos	Ventas/L	Banda Ancha	Correlación	+
DeStefano et al., 2023	2000-2004	Empresa	Modelo de Regresión Discontinua	Ventas/L	Banda Ancha* Urbanos	Causalidad	ns
DeStefano et al., 2023	2000-2004	Empresa	Modelo de Regresión Discontinua	Ventas/L	Banda Ancha* Rurales	Causalidad	ns
Bartelsman et al., 2017	2001-2010	Empresa	Inspirados en Crepon et al., 1998. (Mínimos cuadrados asintóticos (ASL))	VAB/L	Banda Ancha	Causalidad	ns
Bartelsman et al., 2017	2001-2010	Empresa	Inspirados en Crepon et al., 1998. (Mínimos cuadrados asintóticos (ASL))	VAB/L	Banda Ancha	Causalidad	+
Andrews et al., 2018	2010-2016	Industria y Países	Diferencia en Diferencia	Adopción Tecnológica (CC, Front office, Back office)	Capacidades (calidad de directivos, habilidad de trabajadores en TICs)	Causalidad	+
Andrews et al., 2018	2010-2016	Industria y Países	Diferencia en Diferencia	Adopción Tecnológica (CC, Front office, Back office)	Incentivos (competencia, disponibilidad de capital)	Causalidad	+
Gal et al., 2019	2010-2015	Empresa y Países-PTF Industrias y Países-Adopción digital	Corrección de error	PTF	Adopción Tecnológica (CC, ERP, etc.)		+

Artículo	Datos	Países	Nivel	Método	Variable explicada	Predictor de interés	Interpretación	Efecto
Kyläheiko et al., 2011	Transversales	Finlandia	Empresas	MCO	Intensidad exportadora	Medida subjetiva de capacidad tecnológica	Correlación	+
Gbadegeshin et al., 2019	Transversales	Nigeria	PYMES	MCParciales	Inserción Internacional	Herramientas TICs (Sitios web y plataformas de medios sociales)	Correlación	+
Dutot et al., 2014	Transversales	Canadá	Propietarios de PYMES	MCParciales	Inserción Internacional	Complementariedad entre requisitos de información y capacidad de TICs	Correlación	+
Tomski, 2015	Transversales	Polonia	Propietarios de PYMES	MCO	Inserción Internacional	Tener contactos que tengan relación con el extranjero	Correlación	+
Lecerf y Omrani, 2020	Transversales	Alemania	PYMES	Probit	Inserción Internacional	TICs (ERP, e-CRM)	Correlación	ns
Sekliuckiene y Maciulskaitė, 2013	Transversales	Lituania	Directivos de empresas Born Global		Inserción Internacional	Factores externos (desarrollo de las TICs, creación de redes) e internos (habilidades de los directivos, innovación, servicio único)	Correlación	+

Artículo	Datos	Período	Países	Nivel	Método	Variable explicada	Predictor de interés	Interpretación	Efecto
Molina et al., 2013	Transversales	2007	Argentina	PYMES	Logit	VAB/L	Nivel de informatización; Patrón de incorporación de TICs	Correlación	ns
Charlo, 2011	Transversales	2001-2007	Uruguay	Empresas manufactureras	MCPonderados	VAB/L	Inversión en capital TICs	Correlación	+

7.2 TICs: resolución de ACP

$$\mathbb{X} = (x_1, x_2, \dots, x_K) \quad (10)$$

Donde $\mathbb{X}$ es una matriz de $NT \times K$. Definamos a T como un factor de la siguiente manera:

$$T = \mathbb{X}\delta_1 \quad (11)$$

Donde los elementos que conforman a δ_1 se los conoce como “loadings” y lo que indica es cuanto peso relativo tendrá cada una de las variables x_s en el factor T . Como se observa en (11), T es una combinación lineal, entonces la pregunta que surge es cual es la mejor combinación lineal de las x_s . La respuesta en este contexto es que la mejor será la que pueda representar la mayor variabilidad de los datos, es decir, de todas las x_s . Al igual que en MCO, al minimizar la suma de residuos al cuadrado, se está maximizando la suma de cuadrados explicados, dado que la suma de cuadrados totales esta fija. La varianza de T es:

$$V(T) = V(\mathbb{X}\delta_1) = \delta_1' V(\mathbb{X}) \delta_1 \quad (12)$$

El problema que se debe resolver es el de maximizar esta varianza, $\delta_1' V(\mathbb{X}) \delta_1$, donde se normaliza δ_1 de forma tal que la solución no pase por llevar a δ_1 a infinito. Entonces:

$$\max_{\delta_1} \delta_1' V(\mathbb{X}) \delta_1 \quad \text{st} \quad \delta_1' \delta_1 = 1 \quad (13)$$

$$L(\delta_1, \lambda) = \delta_1' V(\mathbb{X}) \delta_1 + \lambda(1 - \delta_1' \delta_1) \quad (14)$$

$$\frac{\partial L}{\partial \delta_1} = 2V(\mathbb{X})\delta_1 - \lambda 2\delta_1 = 0 \quad (15)$$

Donde se utiliza el hecho de que $\mathbb{X}$ es simétrica. Si operamos, podemos obtener la siguiente expresión:

$$(V(\mathbb{X}) - \lambda I_k) \delta_1 = 0 \quad (16)$$

La expresión $(V(\mathbb{X}) - \lambda I_k) \delta_1 = 0$, implica que δ_1 es un autovector y λ es el autovalor de la matriz $V(\mathbb{X})$, que se corresponde con δ_1 . Para que δ_1 no sea el vector nulo, $(V(\mathbb{X}) - \lambda I_k)$ debe ser una matriz singular, es decir, su determinante debe ser cero. Hemos encontrado la dirección principal (autovector) en las que los datos x_s varían más, y la magnitud de esta variación está dada por su autovalor. La solución a este problema de maximización la llamaremos δ_1^* y definimos al primer Componente Principal como $PC_1 = \mathbb{X}\delta_1^*$

Una forma de ver esto último es reordenando (15) y premultiplicando por δ_1' :

$$V(\mathbb{X})\delta_1 = \lambda \delta_1 \quad (17)$$

$$\delta_1' V(\mathbb{X}) \delta_1 = \delta_1' \lambda \delta_1 \quad (18)$$

$$\delta_1' V(\mathbb{X}) \delta_1 = \lambda \quad (19)$$

Cuando maximizamos $\delta_1' V(\mathbb{X}) \delta_1$ sujeto a la normalización $\delta_1' \delta_1 = 1$ debemos elegir λ igual al máximo autovalor de $V(\mathbb{X})$, dado que este será el que maximiza la varianza del factor T .

Si tenemos K variables conformando la matriz $\mathbb{X}$, podemos estimar K Componentes Principales. El método ACP los estima en forma recursiva, comenzando con la resolución del problema anterior para obtener δ_1^* y λ y luego estableciendo condiciones de ortogonalidad entre δ_1^* y δ_2 ,

maximiza a $\delta_2' V(\mathbb{X}) \delta_2$ sujeto a la normalización correspondiente, para así obtener δ_2^* y su autovalor correspondiente:

$$\max_{\delta_2} \delta_2' V(\mathbb{X}) \delta_2 \quad \text{st} \quad \delta_2' \delta_2 = 1 \quad \text{y} \quad \delta_2 \delta_1^* = 0 \quad (20)$$

$$L(\delta_2, \lambda, \lambda_2) = \delta_2' V(\mathbb{X}) \delta_2 + \lambda(1 - \delta_2' \delta_2) + \lambda_2(-\delta_2 \delta_1^*) \quad (21)$$

$$\frac{\partial L}{\partial \delta_2} = 2V(\mathbb{X})\delta_2 - \lambda 2\delta_2 - \lambda_2 \delta_1^* = 0 \quad (22)$$

$$= \delta_1^{*'} 2V(\mathbb{X})\delta_2 - \delta_1^{*'} \lambda 2\delta_2 - \delta_1^{*'} \lambda_2 \delta_1^* = 0 \quad (23)$$

$$= 2\delta_1^{*'} V(\mathbb{X})\delta_2 - 2\lambda \delta_1^{*'} \delta_2 - \lambda_2 \delta_1^{*'} \delta_1^* = 0 \quad (24)$$

Dado que $\delta_1^{*'} \delta_2 = 0$ y que $\delta_1^{*'} \delta_1^* = 1$, tenemos:

$$2\delta_1^{*'} V(\mathbb{X})\delta_2 - \lambda_2 = 0 \quad (25)$$

$$2\delta_1^{*'} V(\mathbb{X})\delta_2 = \lambda_2 \quad (26)$$

Dado que $2\delta_1^{*'} V(\mathbb{X})\delta_2 = 0$ por la restricción $\delta_2 \delta_1^* = 0$ tenemos que:

$$2\delta_1^{*'} V(\mathbb{X})\delta_2 = 0 \quad \text{y, por lo tanto} \quad \lambda_2 = 0 \quad (27)$$

Ahora que sabemos que $\lambda_2 = 0$ lo podemos sustituir en (22) y obtener:

$$\frac{\partial L}{\partial \delta_2} = 2V(\mathbb{X})\delta_2 - \lambda 2\delta_2 = 0 \quad (28)$$

$$(V(\mathbb{X}) - \lambda I_k) \delta_2 = 0 \quad (29)$$

Donde obtenemos la expresión $(V(\mathbb{X}) - \lambda I_k) \delta_2 = 0$, implica que δ_2 es un autovector y λ es el autovalor de la matriz $V(\mathbb{X})$, que se corresponde con δ_2 . Para que δ_2 no sea el vector nulo, $(V(\mathbb{X}) - \lambda I_k)$ debe ser una matriz singular, es decir, su determinante debe ser cero. Estamos encontrando la segunda dirección principal en las que los datos x_s varían más y que cumple con la condición de ser ortogonal (linealmente independiente) al autovector hallado anteriormente δ_1^* , y la magnitud de esta variación está dada por su autovalor. La solución a este problema de maximización la llamaremos δ_2^* y definimos al segundo componente principal como $PC_2 = \mathbb{X} \delta_2^*$.

De esta manera, recursivamente, el método ACP va encontrando los K componentes principales, donde el primero explicará una mayor variabilidad de los datos, así sucesivamente donde, a su vez, deben cumplir la condición de ortogonalidad a medida que se resuelve el problema de forma recursiva.

Se mencionó que la magnitud de la variación asociada con un autovector está dada por su autovalor. Para ver ello, recordemos:

$$V(PC_j) = V(\mathbb{X} \delta_j) = \delta_j' V(\mathbb{X}) \delta_j \quad (30)$$

Aplicando la descomposición espectral a la matriz de varianzas de $\mathbb{X}$, es decir, $V(\mathbb{X}) = \mathbb{P} \Gamma \mathbb{P}'$, donde Γ es una matriz diagonal tal que $\Gamma = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$, es decir, contiene los autovalores de $V(\mathbb{X})$ en su diagonal principal, ordenados desde el mayor al más pequeño y siendo, p_1 ,

$p_2, \dots, p_k$ los autovectores correspondientes. Por lo tanto, $\mathbb{P}$ es una matriz cuyas columnas son los autovectores p_j correspondientes a los autovalores ordenados.

$$\delta'_j V(\mathbb{X}) \delta_j = \delta'_j \mathbb{P} \Gamma \mathbb{P}' \delta_j \quad (31)$$

Dado que los autovectores ordenados p_j son iguales a los autovectores ordenados de $V(\mathbb{X})$, tenemos que $p_j = \delta_j$ para todo $j=1,2,\dots,K$.

$$\delta'_j \mathbb{P} \Gamma \mathbb{P}' \delta_j = p'_j \mathbb{P} \Gamma \mathbb{P}' p_j \quad (32)$$

Dado que p_j es un vector de $(K \times 1)$, $\mathbb{P}$ es una matriz de $(K \times K)$ al igual que Γ , el producto anterior es un escalar. Con lo cual si tomamos el operador traza:

$$p'_j \mathbb{P} \Gamma \mathbb{P}' p_j = \text{Tr}[p'_j \mathbb{P} \Gamma \mathbb{P}' p_j] \quad (33)$$

$$= \text{Tr}[p'_j \Gamma \mathbb{P}' \mathbb{P} p_j] \quad (34)$$

$$= \text{Tr}[p'_j \Gamma p_j] \quad (35)$$

$$= p'_j \Gamma p_j \quad (36)$$

Donde se utilizan propiedades de la traza y la ortogonalidad de los autovectores. Por último, usando que $p_j = \delta_j$, con $j=1,2,\dots,K$:

$$p'_j \Gamma p_j = \delta'_j [\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)] \delta_j \quad (37)$$

$$= \lambda_j \quad (38)$$

De donde se utiliza la restricción $\delta'_j \delta_j = 1$. Con lo cual se obtiene que la varianza del j -ésimo Componente Principal es igual al j -ésimo autovalor ordenado de $V(\mathbb{X})$:

$$V(PC_j) = \lambda_j \quad (39)$$

Podemos definir la importancia relativa del j -ésimo componente principal como la proporción de su varianza entre la varianza total, esta última dada por la sumatoria de todos los λ_j con $j=1,2,\dots,K$.

7.3 TICs: valores faltantes y un algoritmo de imputación

Variable	Missing	Total	% Missing
ctic_1	135	32750	0.41
ctic_2	1377	32750	4.20
ctic_3	135	32750	0.41
ctic_4_baf	3497	32750	10.68
ctic_4_bam	3221	32750	9.84
ctic_5	2376	32750	7.25
ctic_6_1	2423	32750	7.40
ctic_6_2	2424	32750	7.40
ctic_7	2423	32750	7.40
ctic_88_2	2423	32750	7.40
ctic_88_3	2425	32750	7.40
ctic_88_5	2420	32750	7.39
ctic_88_6	2421	32750	7.39
ctic_88_7	2426	32750	7.41
ctic_10	2247	32750	6.86
ctic_13_1	15905	32750	48.56
lstot_def	816	32750	2.49
fb5_9_def	223	32750	0.68
gal6_def	816	32750	2.49

Cuadro 13: Porcentaje de valores perdidos por variable

Para imputar estos valores faltantes se optó por un algoritmo iterativo, llamado “Matrix Completion”. La idea básica del algoritmo es poder explotar las correlaciones entre las diferentes x_s y de esta forma “predecir” cuales son los valores faltantes. Por predecir en este caso vamos a entender aproximarnos. Esto último está relacionado con una interpretación alternativa para el método ACP en el cual puede ser visto como que la suma de los productos entre los “score” de los M primeros componentes principales y los “loadings” de estos primeros M componentes principales proveen la mejor aproximación M -dimensional a las observaciones x_{ij} :

$$\min_{\mathbb{A} \in \mathbb{R}^{N \times T \times M}, \mathbb{B} \in \mathbb{R}^{K \times M}} \left\{ \sum_{(i,j) \in O} \left(x_{ij} - \sum_{m=1}^M a_{im} b_{jm} \right)^2 \right\} \quad (40)$$

Donde O es el conjunto de todos los elementos observados en nuestra encuesta. Para cualquier M , las columnas de $\mathbb{A}$ y de $\mathbb{B}$ que resuelven el problema anterior son los “score” y los “loadings” de los primeros M componentes principales.

Se comienza por imputar el valor medio de cada columna de la matriz original, $\bar{x}_j$, a los valores faltantes de las respectivas columnas de $\mathbb{X}$. Los pasos del algoritmo son los siguientes:

1. Se resuelve (40) utilizando M primeros componentes principales y esto dará como resultado M vectores de “scores” y de “loadings” que nos van a permitir aproximarnos a cada x_{ij} . Es decir:

$$\sum_{m=1}^M \hat{a}_{im} \hat{b}_{jm} \rightarrow \hat{x}_{ij} \quad (41)$$

2. Para cada elemento de la matriz $\mathbb{X}$ que era un dato faltante, que después se le imputó la media de su respectiva columna, se lo actualiza por $\hat{x}_{ij}$. En este punto, trabajamos con los datos faltantes dado que es cuando se van a ir reemplazando en cada iteración.
3. En (40) estamos minimizando la suma de residuos al cuadrado, dado que tenemos los valores reales en x_{ij} y los valores predichos por (41). Podemos definir un error de pronóstico, para los valores observados dado por:

$$mss = Err[\hat{\mathbb{X}}] = E(\mathbb{X} - \hat{\mathbb{X}})^2 \quad (42)$$

Esta es la diferencia entre los valores observados originales y los valores observados actualizados tras correr el paso 1 del algoritmo. Esta medida se actualiza tras cada iteración. Se computa este cálculo solo para los datos que se observan desde nuestra encuesta original. Podemos definir otras dos medidas que también aplican solamente para los datos observados en el conjunto dado por O :

La media al cuadrado de los valores observados originales.

$$mss0 = E[\mathbb{X}]^2 \quad (43)$$

Notar que la media de $mss0$ es cero por estar estandarizados previamente. Entonces si escribimos $E[\mathbb{X}]^2 - E[\mathbb{X}] = E[\mathbb{X}]^2 = mss0$.

La media al cuadrado de los valores observados actualizados con la media justo antes de empezar el primer paso del algoritmo:

$$mssold = E[\mathbb{X}_x]^2 \quad (44)$$

Dado que $mssold$ se computa para los datos que pertenecen al conjunto O , su media también es cero. En el inicio del algoritmo tendremos que $mssold$ termina coincidiendo con $mss0$. Entonces se define el error relativo como:

$$err_{rel} = \frac{mssold - mss}{mss0} \quad (45)$$

Es decir, a $mssold$ es la media al cuadrado para las columnas donde se imputó la media en aquellos valores faltantes, esto es justo antes del algoritmo. Luego, empieza y termina la primera iteración del algoritmo, y se calcula mss por primera vez, que sera la media al cuadrado de cada columna una vez que se termina la primera iteración, donde los valores que habian sido imputados por la media de cada columna, luego fueron actualizados por la aproximación M -dimensional. Luego, en la segunda iteración, justo antes de comenzar, $mssold$ se actualiza nuevamente para igualar a mss . Resta definir un umbral y la iteración parará cuando el cómputo del error relativo este por debajo de dicho umbral. Mientras no se cumpla, seguirá iterando.

7.4 TICs: contribución de variables en ACP

Podemos ver cuánto contribuye cada variable TICs a los dos primeros componentes principales.

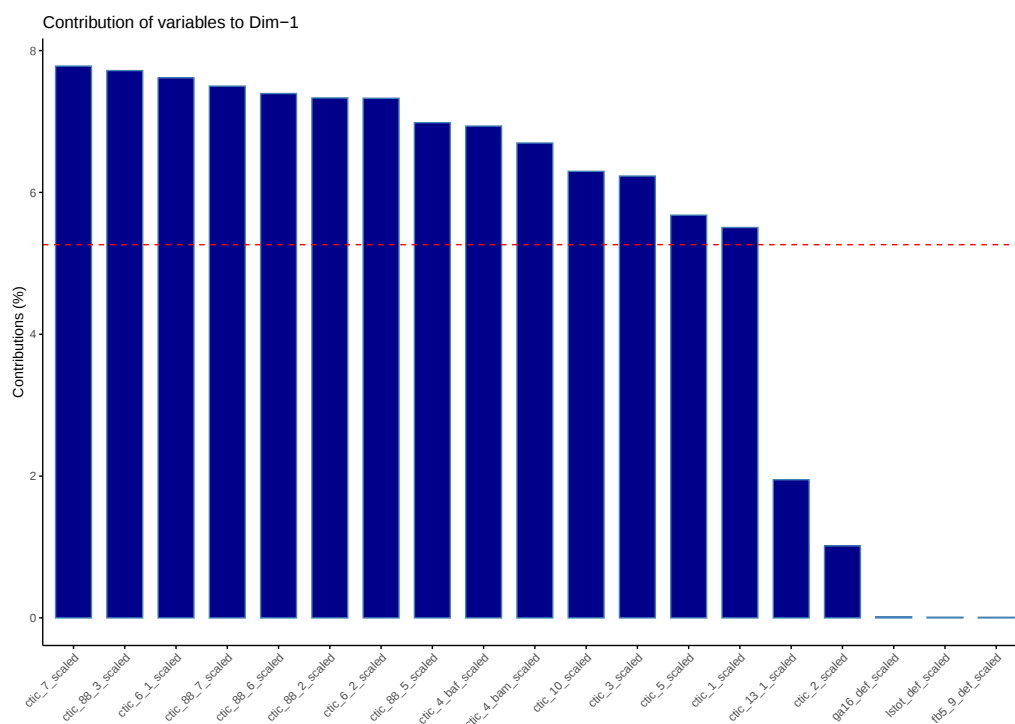


Figura 11: Contribución al primer componente principal

Fuente: Elaboración propia.

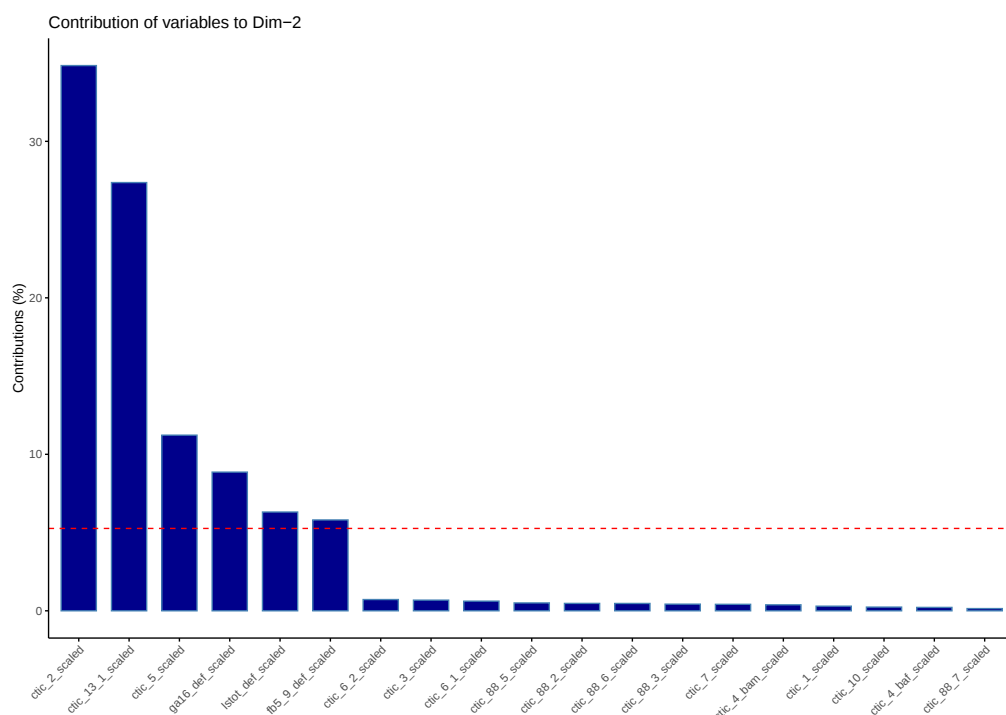


Figura 12: Contribución al segundo componente principal

Fuente: Elaboración propia.

7.5 PTF: breve mirada a Olley y Pakes (1992)

En el presente apartado se pretende ilustrar la parte técnica del algoritmo en dos etapas propuesto en Olley y Pakes (1992) para obtener estimaciones consistentes de los coeficientes de una función de producción. Comenzando con las soluciones del modelo teórico propuesto por los autores, de las cuales se derivan dos relaciones importantes, la regla de cierre de las empresas y la función de demanda de inversión.

Sea χ_{it} el umbral mínimo necesario de productividad para que la firma continúe funcionando, la regla de cierre puede escribirse como:

$$\chi_{it} = \begin{cases} 1 & \text{if } \omega_{it} \geq \underline{\omega}_{it}(a_{it}, k_{it}) \\ 0 & \text{otro caso} \end{cases}$$

La cual depende de las variables de estado (dinámicas). Mientras que la demanda de inversión:

$$i_{it} = i_{it}(\omega_{it}, a_{it}, k_{it}) \quad (46)$$

Depende de las variables de estado y de la productividad. La especificación de partida, que se deriva de una función de producción Cobb-Douglas, es la siguiente:

$$y_{it} = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it} \quad (47)$$

Para valores positivos de la inversión (46), se puede invertir la función de forma tal de expresar la productividad inobservable en función de observables:

$$\omega_{it} = h_t^{-1}(i_{it}, a_{it}, k_{it}) \quad (48)$$

Implícito en la expresión anterior está el hecho de que se asume que la productividad, ω_{it} , es un escalar. En Akerberg et al. (2015) desarrollan este punto en forma sobresaliente. Este supuesto puede ser restrictivo puesto que h_t^{-1} sólo varía en el tiempo, pero no entre las empresas. De esta forma, descarta cualquier heterogeneidad no observada entre empresas. En particular, la presencia de shocks que afecten a la inversión (condicionada por el capital y la productividad) de forma heterogénea entre las empresas no son contempladas por este supuesto. Ejemplos de estos shocks pueden ser la presencia de costos de ajustes del capital no observados que varían entre empresas u perturbaciones no observadas específicas de la empresa en los precios de la inversión. Tampoco serán contemplados por dicha función los shocks que afecten a la mano de obra, en tanto se asume que el trabajo se correlaciona con la productividad, y dado que la productividad es el factor inobservable en la demanda de inversión, solamente serían contemplados casos poco creíbles como que todas las empresas accedan a un mismo y único mercado laboral, en el caso más realista de que las empresas se vean en la búsqueda de trabajadores con diferentes características, estos shocks serán heterogéneos y no serán contemplados por h_t^{-1} . Tampoco admite shocks de otras potenciales variables que no sean de estado y diferentes al trabajo, como por ejemplo, los asociados a los insumos intermedios, puesto que si incluyéramos (en una extensión de OP) a los insumos intermedios, es posible que también se correlacionen con el nivel de la productividad actual y por lo tanto, shocks asociados a estos insumos intermedios (materiales, energía, etc.) tampoco serán contemplados por dicha función, excepto que todas las empresas accedan al mismo mercado de insumos intermedios, situación poco realista. Por lo tanto, estas tres fuentes de shocks heterogéneos (inversión, trabajo e insumos intermedios) no son admitidas por el proceso generador de datos de OP. Cabría agregar una cuarta fuente relacionada a los shocks heterogéneos asociados al producto.

Por otro lado, cabe destacar que en el caso de LP si bien se mantiene el supuesto de que la productividad es un escalar, su PGD es diferente en el sentido de que utilizan como proxy de la productividad a la demanda de insumos intermedios y tratan esta variable igual al trabajo, esto es, como una variable no dinámica que se elige en t una vez que la empresa observa su nivel de productividad. Si bien en el caso de LP tampoco se admitirán shocks heterogéneos del trabajo o de bienes intermedios en tanto relacionados con la productividad, ni tampoco shocks relacionados al producto, la diferencia con OP es que ahora si se admiten shocks heterogéneos relacionados a la inversión (por ejemplo, costos de ajuste del capital) una vez que se condiciona por el capital y la productividad. En otras palabras, los insumos intermedios no son afectados por los shocks relacionados a la inversión, una vez condicionados por el capital físico y la productividad, dado que los insumos intermedios no son dinámicos y sólo afectan a los beneficios actuales. Otra ventaja de LP respecto de OP puede ser el hecho en la práctica de que la inversión tiene una alta densidad en el valor cero, lo que se traduce en una pérdida de datos importante a la hora de trabajar sólo con valores positivos. Retomando el análisis de OP:

$$y_{it} = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + \beta_l l_{it} + h_t^{-1}(i_{it}, a_{it}, k_{it}) + \eta_{it} \quad (49)$$

$$y_{it} = \beta_l l_{it} + \phi_t(i_{it}, a_{it}, k_{it}) + \eta_{it} \quad (50)$$

Donde $\phi_t(i_{it}, a_{it}, k_{it}) = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + h_t^{-1}(i_{it}, a_{it}, k_{it})$. Los autores estiman de forma no paramétrica la función $\phi_t(\cdot)$ y luego la utilizan para estimar β_l mediante (50). Esto permite resolver el problema de simultaneidad al utilizar a la demanda de inversión como proxy de la productividad. Sin embargo, en cuanto los coeficientes para las variables de estado esto implica un problema de identificación al no poder separar su efecto sobre la inversión respecto de sus efectos sobre el valor agregado de producción.

Para resolver el problema de identificación usa el hecho de que el siguiente periodo se observan datos para las empresas que sobreviven, por lo que se necesita una probabilidad de supervivencia.

$$P(\chi_{it+1} = 1 \mid \underline{\omega}_{it+1}(a_{it+1}, k_{it+1}), \zeta_{it}) \quad (51a)$$

$$= P(\omega_{it+1} \geq \underline{\omega}_{it+1}(a_{it+1}, k_{it+1}) \mid \underline{\omega}_{it+1}(a_{it+1}, k_{it+1}), \omega_{it}) \quad (51b)$$

$$\equiv \mathbf{P}_{it} \quad (51c)$$

Como vemos, esta se condiciona por el umbral mínimo de productividad del próximo período y por la productividad del periodo actual.

Las ecuaciones de las variables de estado ($a_{it+1} = a_{it} + 1$; $k_{it+1} = (1 - \delta)k_{it} + i_{it}$) junto con la demanda de inversión, $i_{it} = i_{it}(\omega_{it}, a_{it}, k_{it})$ implican que a_{it+1} y k_{it+1} pueden expresarse como funciones de la terna $(\omega_{it}, a_{it}, k_{it})$. A su vez, dado que $\omega_{it} = h_t^{-1}(i_{it}, a_{it}, k_{it})$ se puede expresar a la probabilidad de supervivencia, $\mathbf{P}_{it}$, como función de la terna (i_{it}, a_{it}, k_{it}) . Esto da pie a que posteriormente procedan a estimar esta probabilidad de supervivencia por aproximaciones polinómicas para la terna (i_{it}, a_{it}, k_{it}) o mediante el uso de Kernel, de forma similar a la estimación anterior para la función $\phi_t(\cdot)$.

Solo observamos los datos del año siguiente para las empresas que sobreviven, entonces:

$$y_{it+1} - \beta_l l_{it+1} = \beta_0 + \beta_a a_{it+1} + \beta_k k_{it+1} + \omega_{it+1} + \eta_{it+1} \quad (52)$$

Podemos sumar y restar el siguiente término, $E[\omega_{it+1}|\omega_t, \chi_{it+1} = 1]$, que surge luego de tomar expectativas de la expresión (52) anterior, por lo tanto:

$$\begin{aligned} y_{it+1} - \beta_l l_{it+1} &= \beta_0 + \beta_a a_{it+1} + \beta_k k_{it+1} + E[\omega_{it+1}|\omega_t, \chi_{it+1} = 1]) \\ &\quad + (\omega_{it+1} - E[\omega_{it+1}|\omega_t, \chi_{it+1} = 1]) + \eta_{it+1} \\ y_{it+1} - \beta_l l_{it+1} &= \beta_0 + \beta_a a_{it+1} + \beta_k k_{it+1} + g(\mathbf{P}_{it}, \omega_{it}) + \varepsilon_{it+1} + \eta_{it+1} \end{aligned} \quad (53)$$

Por un lado, $g(\mathbf{P}_{it}, \omega_{it}) = E[\omega_{it+1}|\omega_{it}, \chi_{it+1} = 1]$ y, por otro, $\varepsilon_{it+1} = \omega_{it+1} - E[\omega_{it+1}|\omega_t, \chi_{it+1} = 1]$. Para esta última expresión observar que su media es cero. Los autores plantean estimar de forma no paramétrica la función $g(\cdot)$ que engloba a la probabilidad de supervivencia y el proxy de productividad.

$$E[y_{it+1} - \beta_l l_{it+1}] = \beta_0 + \beta_a a_{it+1} + \beta_k k_{it+1} + g(\mathbf{P}_{it}, \omega_{it}) \quad (54a)$$

$$E[y_{it+1} - \beta_l l_{it+1}] = \beta_0 + \beta_a a_{it+1} + \beta_k k_{it+1} + g(\mathbf{P}_{it}, h_t^{-1}(i_{it}, a_{it}, k_{it})) \quad (54b)$$

$$E[y_{it+1} - \beta_l l_{it+1}] = \beta_0 + \beta_a a_{it+1} + \beta_k k_{it+1} + g(\mathbf{P}_{it}, \phi_t(\cdot) - \beta_0 - \beta_a a_{it} - \beta_k k_{it}) \quad (54c)$$

Algoritmo semiparamétrico en dos etapas:

1. Estima la función $\phi_t(\cdot)$. Luego se corre la primera etapa para obtener $\hat{\beta}_l$ utilizando (50):
2. Utilizando $\hat{\beta}_l$ para el nuevo vector de variable dependiente y usando $\hat{\phi}_t$ y $\hat{\mathbf{P}}_{it}$ se logra estimar tanto la función $g(\cdot)$ como las estimaciones de interés β_0, β_a y β_k :

$$y_{it+1} - \hat{\beta}_l l_{it+1} = \beta_0 + \beta_a a_{it+1} + \beta_k k_{it+1} + g(\hat{\mathbf{P}}_{it}, \hat{\phi}_t(\cdot) - \beta_0 - \beta_a a_{it} - \beta_k k_{it}) \quad (55)$$

Dado que el efecto de las variables de estado no se puede distinguir de sus efectos sobre la inversión $h_t^{-1}(\cdot)$ de sus efectos sobre la nueva variable dependiente, se siguen dos estrategias alternativas para lograrlo que requieren algoritmos no lineales de convergencia que minimizan la suma de residuos al cuadrado de (55).

Los autores en Olley y Pakes (1992) estiman $\phi_t(\cdot)$ utilizando un polinomio para (i_{it}, a_{it}, k_{it}) de orden cuatro con un conjunto completo de interacciones para poder aproximarse a dicha función. Para la probabilidad de supervivencia, $\mathbf{P}_{it}$, utilizan un polinomio para (i_{it}, a_{it}, k_{it}) de orden cuatro con un conjunto completo de interacciones como regresor en una estimación Probit. Alternativamente, la estiman mediante Kernel y comparan resultados. De sus comparaciones optan por una estimación basado en Kernel para la probabilidad de supervivencia.

Por último, para estimar la función $g(\cdot)$ que depende tanto de $\hat{\mathbf{P}}_{it}$ como de $h_t(\cdot)^{-1}$, la intentan aproximar tanto por un polinomio como por estimación Kernel, donde la diferencia es que, tal como se mencionó, se debe recurrir a técnicas de convergencia no lineales para obtener el vector de coeficientes de capital y edad tal que minimizan la suma de residuos al cuadrado.

Por último, encuentran que incluir una tendencia en la regresión es significativa y optan por mantenerla como parte del modelo al momento de estimar los coeficientes de interés. Luego no se utiliza para la construcción de la medida de la PTF.

Finalmente, utilizando las estimaciones de β_0 , β_l , β_k y β_a y recordando:

$$y_{it} = \beta_0 + \beta_a a_{it} + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it}$$

Podemos obtener una estimación robusta para la PTF:

$$y_{it} - \hat{\beta}_a a_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_l l_{it} - \hat{\eta}_{it} = \hat{\beta}_0 + \hat{\omega}_{it} \quad (56)$$

$$y_{it} - \hat{\beta}_a a_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_l l_{it} - \hat{\eta}_{it} = PTF_{it} \quad (57)$$

Olley y Pakes (1992) nos proponen un algoritmo con el objetivo de estimar en forma consistente los coeficientes de una función de producción a nivel de plantas-firmas y con ello obtener una estimación consistente de la PTF. Los autores utilizan la inversión como proxy a la productividad.

Tal como se detalla en el trabajo de Akerberg et al. (2015), los enfoques de OP y de LP:

- Permiten la endogeneidad de los insumos con respecto al inobservable variable en el tiempo.
- Pueden permitir que subconjuntos de insumos sean dinámicos sin requerir soluciones explícitas derivadas de condiciones de primer orden de estos insumos. Es decir, la función inversa desconocida sería la solución a un problema de programación dinámica, en cambio, en estos desarrollos la tratan como no paramétrica.
- No requieren que el econometrista observe la variación exógena de los precios de los insumos entre las empresas. (En dicho caso podría usarse un enfoque de VI basado en precios de insumos).

Sin embargo, los supuestos que subyacen a los enfoques de OP y LP pueden ser restrictivos. El enfoque de LP tiene una ligera ventaja respecto de OP al permitir shocks heterogéneos relacionados al capital. Pero tanto el enfoque LP como OP pueden adolecer del problema de identificación para el coeficiente de la mano de obra en la primera etapa de su algoritmo semiparamétrico.

7.6 PTF: breve mirada a Wooldridge (2009)

Este trabajo se basa en el proceso generador de datos de LP, sus principales aportes son (a) resolver el problema de la dependencia funcional de las estimaciones de la primera etapa de los enfoques de OP y LP asociados a la mano de obra y (b) aportar una estimación que podría ser más eficiente al considerar correlaciones cruzadas entre las variables, correlaciones en serie y heterocedasticidades de los errores. Para la empresa i en el año t , tenemos la siguiente función de producción:

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + v_{it} + e_{it} \quad (58)$$

El término dependiente y_{it} es el logaritmo del valor agregado de producción. El vector $\mathbf{w}_{it}$ contiene insumos variables (típicamente el trabajo), el vector $\mathbf{x}_{it}$ contiene variables de estado (típicamente el capital físico), todos en logaritmo. La productividad inobservable para el econométrista es representada por v_{it} , mientras que e_{it} son shocks independientes en media de los valores actuales y pasados de los insumos de la función de producción.

Basándose en la idea de Levinsohn y Petrin (2003), la productividad se puede escribir como una función (desconocida) de las variables de estado (capital físico) y el proxy de la productividad, que en este caso serán los insumos intermedios, $\mathbf{m}_{it}$, pero podría ser la demanda de inversión si nos basamos en Olley y Pakes (1992):

$$v_{it} = g(\mathbf{x}_{it}, \mathbf{m}_{it}) \quad (59)$$

Como se mencionó, el shock es independiente en media:

$$E(e_{it} | \mathbf{w}_{it}, \mathbf{x}_{it}, \mathbf{m}_{it}) = 0 \quad (60)$$

El problema de identificación de β surge cuando:

$$E(y_{it} | \mathbf{w}_{it}, \mathbf{x}_{it}, \mathbf{m}_{it}) = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + g(\mathbf{x}_{it}, \mathbf{m}_{it}) \quad (61a)$$

$$= \mathbf{w}_{it}\beta + h(\mathbf{x}_{it}, \mathbf{m}_{it}) \quad (61b)$$

Aparentemente, la ecuación (61b) logra identificar al vector de insumos variables β . Sin embargo, si el insumo trabajo ($l_{it} \subset \mathbf{w}_{it}$) es elegido por la firma en el *mismo momento* que los bienes intermedios $\mathbf{m}_{it}$ (o la inversión en el caso de Olley y Pakes (1992)), entonces será una función determinista de los vectores $\mathbf{x}_{it}$ y $\mathbf{m}_{it}$, esto es, de las variables de estado y del vector de variables proxy de la productividad. Por lo tanto, en este Proceso Generador de Datos (PGD) el coeficiente β no se logra identificar debido a que la estimación de β en (61b) recogerá parte de la relación entre $(\mathbf{x}_{it}; \mathbf{m}_{it})$ y $\mathbf{w}_{it}$.

Llevando la condición (60) más lejos, permite dependencia serial de los shocks, dado que valores pasados de y_{it} y de e_{it} no aparecen en el conjunto condicionante:

$$E(e_{it} | \mathbf{w}_{it}, \mathbf{x}_{it}, \mathbf{m}_{it}, \mathbf{w}_{it-1}, \mathbf{x}_{it-1}, \mathbf{m}_{it-1}, \dots, \mathbf{w}_{i1}, \mathbf{x}_{i1}, \mathbf{m}_{i1}) = 0 \quad (62)$$

Por otro lado, la productividad sigue un proceso de Márkov exógeno:

$$E(v_{it} | v_{it-1}, \dots, v_{i1}) = E(v_{it} | v_{it-1}) \quad (63)$$

Definiendo el shock de productividad como:

$$a_{it} \equiv v_{it} - E(v_{it} | v_{it-1}) \quad (64)$$

Las variables de estado (siguen alguna dinámica temporal) del periodo actual, $\mathbf{x}_{it}$ se asumen incorrelacionadas con estos shocks, a_{it} . Mientras que los insumos variables del periodo actual, $\mathbf{w}_{it}$ sí pueden estar correlacionados con a_{it} debido a que factores como el trabajo si pueden ajustarse (instantáneamente) en el mismo período a un shock de productividad. Como veremos, las variables proxy del período actual, $\mathbf{m}_{it}$, también pueden estar correlacionadas con estos shocks de productividad a_{it} .

Para lograr la identificación de γ también se requiere que a_{it} este incorrelacionada con $\mathbf{x}_{it-1}$ y con $\mathbf{m}_{it-1}$:

$$E(v_{it}|\mathbf{x}_{it}, \mathbf{w}_{it-1}, \mathbf{x}_{it-1}, \mathbf{m}_{it-1}, \dots, \mathbf{w}_{i1}, \mathbf{x}_{i1}, \mathbf{m}_{i1}) = E(v_{it}|v_{it-1}) = f[g(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})] \quad (65)$$

De donde se utiliza el hecho de que $v_{it-1} = g(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})$ y la composición de funciones, donde $f(\cdot)$ es una función desconocida. Observar, por una parte, que se permite que tanto $\mathbf{w}_{it}$ como $\mathbf{m}_{it}$ puedan estar correlacionadas con el shock de productividad, sin embargo, el valor actual de $\mathbf{x}$ y los valores pasados de $(\mathbf{w}, \mathbf{x}, \mathbf{m})$ y todas las funciones de estos valores se suponen incorrelacionadas con el shock de productividad. Esto es lo que nos dice la expresión (65).

Por lo tanto, partiendo de (58):

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + v_{it} + e_{it} \quad (66a)$$

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + a_{it} + E(v_{it}|v_{it-1}) + e_{it} \quad (66b)$$

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + a_{it} + f[g(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})] + e_{it} \quad (66c)$$

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + f[g(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})] + u_{it} \quad (66d)$$

$$\text{donde } u_{it} = a_{it} + e_{it} \quad (66e)$$

Las expresiones (58), (59) y (62) nos dan la primera ecuación:

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + g(\mathbf{x}_{it}, \mathbf{m}_{it}) + e_{it} \quad (67)$$

Mientras que las expresiones (58), (59), (62) junto con las expresiones (63), (64) y (65) nos dan la segunda ecuación:

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + f[g(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})] + u_{it} \quad (68)$$

Con ambas podemos identificar a los vectores de coeficientes β y γ en un único paso.

Las condiciones de ortogonalidad son diferentes para ambas ecuaciones. Los errores e_{it} de la ecuación (67) vienen dados por (62) mientras que para los errores u_{it} de la ecuación (68) vienen dados por:

$$E(u_{it}|\mathbf{x}_{it}, \mathbf{w}_{it-1}, \mathbf{x}_{it-1}, \mathbf{m}_{it-1}, \dots, \mathbf{w}_{i1}, \mathbf{x}_{i1}, \mathbf{m}_{i1}) = 0 \quad (69)$$

Indica que tanto $\mathbf{w}_{it}$ como $\mathbf{m}_{it}$ pueden estar correlacionadas con u_{it} sólo a través de a_{it} .

En la ecuación (68) se puede utilizar a $\mathbf{w}_{it-1}$ como instrumento de $\mathbf{w}_{it}$, debido a que esta última variable sería endógena en dicha especificación al estar correlacionada con u_{it} a través de a_{it} .

Como se ha mencionado, Olley y Pakes (1992) aproxima la función $g(\cdot)$ mediante una aproximación polinómica de grado cuatro, mientras que en Levinsohn y Petrin (2003) lo hacen con una expresión de grado tres. En este caso, Wooldridge (2009) sigue los pasos de estos últimos.

Con lo cual, si x y m son dos escalares, $g(x, m)$ es lineal en términos de la forma $x^p m^q$ donde p y q son dos enteros no negativos con la restricción $p + q \leq 3$:

$$g(x, m) = \lambda_{00} + \lambda_{01}m^1 + \lambda_{10}x^1 + \lambda_{11}x^1m^1 + \lambda_{02}m^2 + \lambda_{20}x^2 + \lambda_{03}m^3 + \lambda_{21}x^2m^1 + \lambda_{12}x^1m^2 + \lambda_{30}x^3$$

Esto forma un sistema que incluye hasta 10 términos independientes. Para cada par de datos (x_i, m_i, g_i) se reemplaza en el polinomio anterior y se genera un sistema lineal que puede escribirse en formato matricial:

$$g(x_{it}, m_{it}) = c\lambda \quad (70a)$$

$$= \lambda_0 + c(x_{it}, m_{it})\lambda \quad (70b)$$

Donde c es la matriz con los términos monomiales de la aproximación polinómica y λ el vector de coeficientes. Por otra parte, Wooldridge (2009) asume que $f(\cdot)$ puede ser aproximada por un polinomio en v de la siguiente forma:

$$f(v) = \rho_0 + \rho_1 v + \dots + \rho_G v^G$$

Retomando el análisis:

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + g(\mathbf{x}_{it}, \mathbf{m}_{it}) + e_{it} \quad (71a)$$

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + \lambda_0 + c(\mathbf{x}_{it}, \mathbf{m}_{it})\lambda + e_{it} \quad (71b)$$

$$y_{it} = \alpha_0 + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + c(\mathbf{x}_{it}, \mathbf{m}_{it})\lambda + e_{it} \quad (71c)$$

Donde $\alpha_0 = \alpha + \lambda_0$

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + f[g(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})] + u_{it} \quad (72a)$$

$$y_{it} = \alpha + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + \rho_0 + \rho_1(\lambda_0 + c(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})\lambda) + \dots + \rho_G(\lambda_0 + c(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})\lambda)^G + u_{it} \quad (72b)$$

$$y_{it} = \eta_0 + \mathbf{w}_{it}\beta + \mathbf{x}_{it}\gamma + \rho_1(c(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})\lambda) + \dots + \rho_G(c(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})\lambda)^G + u_{it} \quad (72c)$$

Donde $\eta_0 = \alpha + \rho_0 + \sum_{k=1}^G \rho_k \lambda_0$

Los instrumentos para las ecuaciones (71c) y (72c) son, respectivamente:

$$Z_{it1} = (1, \mathbf{w}_{it}, \mathbf{x}_{it}, c(\mathbf{m}_{it})) \quad (73)$$

$$Z_{it2} = (1, \mathbf{x}_{it}, \mathbf{w}_{it-1}, c(\mathbf{x}_{it-1}, \mathbf{m}_{it-1}), q(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})) \quad (74)$$

Donde $c(\mathbf{m}_{it})$ instrumenta a $c(\mathbf{x}_{it}, \mathbf{m}_{it})$. Observar que m podría representar a la demanda de inversión, los insumos intermedios o el consumo en energía eléctrica como una proxy de la productividad.

Mientras que, $c(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})$ instrumenta a $g(\cdot)$ en $t - 1$ y depende tanto de m como de x . Además, $q(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})$ son funciones no lineales de $c(\mathbf{x}_{it-1}, \mathbf{m}_{it-1})$ para instrumentar a $f(g(\cdot))$.

Dado que los instrumentos Z_{it2} son válidos también para la primera ecuación (71c), se pueden redefinir para obtener una matriz de instrumentos tal que:

$$E[Z'_{it} \mathbf{r}_{it}(\theta)] = 0 \quad (75)$$

Siendo $\mathbf{r}_{it}(\theta)$ los residuos (71c) y (72c) respectivamente. Donde θ representa a los coeficientes del sistema formado por ambas especificaciones.

Sin embargo, cabe mencionar que en el trabajo de Akerberg et al. (2015) se analiza en detalle el proceso generador de datos que se corresponde con el propuesto por esta sección.

7.7 PTF: breve mirada a Akerberg et al. (2015)

Estos autores relajan algunos de los supuestos detrás de OP y LP y también logran evitar el problema de la dependencia funcional que tienen OP y LP en su primera etapa. Entre sus aportes está el de señalar claramente cuales son los procesos generadores de datos compatibles con los desarrollos anteriores, además de aportar una estrategia que va a permitir ser compatible con un conjunto más amplio de PGD. No obstante, los supuestos de temporalidad, escalar inobservable y monotonicidad se siguen manteniendo como requisitos para la identificación de los coeficientes.

Problema de dependencia funcional: primera etapa

Akerberg et al. (2015) argumentan que si se considera la condición de momento de la primera etapa de LP:

$$E[\epsilon_{it} | I_{it}] = E[y_{it} - \beta l_{it} - \Phi(k_{it}, m_{it}) | I_{it}] = 0 \quad (76)$$

Entonces esta se corresponde con un modelo parcialmente lineal estudiado por Robinson (1988). La condición de identificación del coeficiente de mano de obra (β_l) es la siguiente:

$$E[\{l_{it} - E[l_{it}|k_{it}, m_{it}, t]\} \{l_{it} - E[l_{it}|k_{it}, m_{it}, t]\}] \text{ es definida positiva.} \quad (77)$$

Que nos dice que, una vez que se controle por el capital físico, los bienes intermedios y el tiempo, debe existir una variación entre el valor actual del trabajo y su valor esperado. Akerberg et al. (2015) muestran como partiendo de una especificación del tipo Cobb-Douglas para la función de producción no se cumple con la ecuación anterior, puesto que una vez que se condiciona por estas variables, la condición de primer orden resultante no contiene el término de la elasticidad de la mano de obra, con lo que esta CPO no puede ser informativa sobre el mismo. En otras palabras, la ecuación anterior sería igual a cero puesto que una vez que se controla por el capital, los insumos intermedios y el tiempo, el trabajo queda completamente determinado, y entonces el valor actual de la mano de obra será igual a su valor esperado.

Dado que podrían utilizarse formas más complicadas que las funciones Cobb-Douglas y entonces derivar CPO que contengan el coeficiente de la mano de obra, el asunto pasa porque los enfoques de LP (y OP) utilizan formas no paramétricas y esto dificulta la identificación. Entonces los autores mencionados lo que hacen es preguntarse cuales son los PGD que implican que la mano de obra esté totalmente determinada por el capital y los insumos intermedios puesto que estos procesos generadores de datos serán aquellos que no permiten identificar a β_l . De forma intuitiva, la contribución del trabajo al valor agregado de producción no se puede separar de la función no paramétrica, en la ecuación de la primera etapa. A la inversa, muestran cuales son los PGD que, en los contextos de los desarrollos de LP (y OP) que permite la identificación de la mano de obra en la primera etapa. Para remitirse a los casos encontrados, que son muy pocos (dos PGD muy específicos para LP, y uno adicional para el caso de OP), se remite al lector a dirigirse a Akerberg et al. (2015, pp. 2424–2426)²¹.

²¹ Tomado de Akerberg et al. (2015, pp. 2426) : (1) error de optimización i.i.d en l_{it} (y no en m_{it} (o i_{it})); (2) shocks i.i.d. en el precio del trabajo o v_{it} después de que se elija m_{it} (o i_{it}) pero antes de que se elija l_{it} ; o (3) (en el contexto OP) el trabajo no es dinámico y se elige en $t - b$ en función de ω_{it-b} , mientras que i_{it} se elige en t .

- Sólo tres (dos) PGD muy específicos permiten identificar al coeficiente de la mano de obra en la primera etapa de las estrategias de OP (LP).
- En la práctica, este problema de dependencia funcional no se observaría (Si los PGD son, de hecho, diferentes). Esto simplemente daría una colinealidad entre la mano de obra y los términos de la función no paramétrica. Con lo cual se observa una estimación numérica para β_l .

Propuesta de los autores |

Mientras que OP y LP invierten las funciones de demanda de inversión y de insumos intermedios respectivamente, y donde ambas son *incondicionales* del insumo trabajo, en Akerberg et al. (2015) se proponen invertir las funciones de demanda de inversión o de insumos intermedios **condicionales** al insumo trabajo. Esto es la diferencia clave que les permite diferenciarse de los desarrollos previos (incluido el de Wooldridge (2009)), en el sentido de lograr romper con la dependencia funcional para el insumo trabajo.

Como veremos, esta ganancia estará dada en que las condiciones de momentos en Akerberg et al. (2015) producen estimaciones consistentes si hay shocks no observados, relacionados serialmente y específicos de la empresa en el precio de la mano de obra (p_l), si la mano de obra (l_{it}) se elige antes que otros insumos variables, o si la mano de obra es dinámica y hay costos de ajuste de la mano de obra no observados y específicos de la empresa.

Considerado siguiente especificación para el logaritmo del valor agregado de la producción:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \epsilon_{it} \quad (78)$$

Donde k_{it} es nuevamente el logaritmo del capital físico, l_{it} es el logaritmo del trabajo y ω_{it} es la productividad no observada por el econometrista y ϵ_{it} es un error de medida. Las empresas acumulan capital de acuerdo con:

$$k_{it} = \kappa(k_{it-1}, i_{it-1}) \quad (79)$$

Donde la inversión i_{it-1} es elegida en el período $t - 1$, al igual que en OP. Ahora se permitirá que la mano de obra pueda ser potencialmente dinámica y que podría ser elegida en el periodo $t-1$ o en el periodo $t-b$ (con $0 < b < 1$) además del periodo t , a diferencia de los enfoques previos de OP y LP donde la mano de obra es estática y elegida en t únicamente. Esta es la diferencia clave de la propuesta de los autores en Akerberg et al. (2015) al permitir que ahora la mano de obra pueda ser una variable de estado. Si se utiliza la demanda de insumos intermedios:

$$m_{it} = f_t(k_{it}, l_{it}, \omega_{it}) \quad (80)$$

Observar que la demanda de insumos intermedios es condicional a la mano de obra. Los insumos intermedios pueden elegirse después o en simultáneo con la mano de obra. El siguiente es el supuesto de monotonidad estricta en el nivel de productividad para la función de demanda de insumos intermedios:

$$f_t(k_{it}, l_{it}, \omega_{it}) \text{ es estrictamente creciente en } \omega_{it}. \quad (81)$$

Utilizando estos dos últimos supuestos:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + f_t^{-1}(k_{it}, l_{it}, m_{it}) + \epsilon_{it} \quad (82)$$

$$y_{it} = \Phi_t(k_{it}, l_{it}, m_{it}) + \epsilon_{it} \quad (83)$$

Donde $\Phi_t(k_{it}, l_{it}, m_{it}) = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + f_t^{-1}(k_{it}, l_{it}, m_{it})$. La especificación anterior representa la primera etapa del método de Akerberg et al. (2015) donde β_0 , β_k , β_l y la función $f_t^{-1}(\cdot)$ quedan sin identificar. Lo que permite es estimar $\hat{\Phi}_t(k_{it}, l_{it}, m_{it})$.

La condición de momento asociada a esta primera etapa es:

$$E[\epsilon_{it}|I_{it}] = E[y_{it} - \Phi_t(k_{it}, l_{it}, m_{it})|I_{it}] = 0 \quad (84)$$

Utilizando los supuestos asociados a los shocks transitorios, $E[\epsilon_{it}|I_{it}] = 0$, donde el conjunto de información I_{it} incluye información sobre los shocks pasados y presentes de productividad, y donde estos últimos evolucionan de acuerdo con un proceso exógeno de Márkov con distribución de ω_{it} conocida para las empresas y creciente en el nivel de productividad, los autores descomponen la productividad en dos componentes:

$$\omega_{it} = E[\omega_{it}|I_{it-1}] + \zeta_{it} = g(\omega_{it-1}) + \zeta_{it} \quad (85)$$

Entonces:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \epsilon_{it} \quad (86)$$

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + g(\omega_{it-1}) + \zeta_{it} + \epsilon_{it} \quad (87)$$

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + g(f_{t-1}^{-1}(k_{it-1}, l_{it-1}, m_{it-1})) + \zeta_{it} + \epsilon_{it} \quad (88)$$

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + g(\Phi_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1}) + \zeta_{it} + \epsilon_{it} \quad (89)$$

Por lo tanto, la segunda condición de momento es:

$$\begin{aligned} E[\zeta_{it} + \epsilon_{it}|I_{it-1}] = \\ E[y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it} - g(\Phi_{t-1}(k_{it-1}, l_{it-1}, m_{it-1}) - \beta_0 - \beta_k k_{it-1} - \beta_l l_{it-1})|I_{it-1}] = 0 \end{aligned} \quad (90)$$

Donde, al igual que en LP y OP, los efectos de las variables de estado (capital y ahora también la mano de obra) no se podrán distinguir de su efecto sobre la función $g(\cdot)$ de su efecto sobre el valor agregado de producción. Por lo que será necesario algoritmos no lineales de convergencia que minimicen la suma de residuos al cuadrado de (90).

Considerando shocks no observados, exógenos y correlacionados en serie a los costos de k_{it} , l_{it} y m_{it} tenemos que el algoritmo OP no puede permitir ninguno de estos shocks, LP puede permitir los relativos a k_{it} aunque no los relativos a l_{it} y m_{it} mientras que los momentos propuestos en Akerberg et al. (2015) pueden permitir los relativos a k_{it} y l_{it} pero no los relativos a m_{it} .

Los autores argumentan que, en el caso de contar con datos de precios para los insumos intermedios, el capital y el trabajo, entonces optar por una estrategia de VI basado en estos precios como instrumentos es una mejor alternativa. En dicho caso se debe fundamentar que los precios de los insumos son exógenos, es decir, que la diferencia entre precios no represente una diferencia en calidad de los insumos y, por tanto, en el fondo, de productividad.

7.8 SEMs: verificación de la condición de rango

Para verificar la identificación de cada ecuación del sistema estructural, se construyeron las matrices de restricciones R_1 , R_2 y R_3 , las cuales imponen restricciones de exclusión sobre el vector de coeficientes estructurales β_h de cada ecuación $h = 1, 2, 3$. También se construyó la matriz $\mathbf{B}$ que contiene todos los parámetros estructurales del sistema:

$$\mathbf{B} = \begin{bmatrix} -1 & \gamma_{21} & \gamma_{31} \\ \gamma_{12} & -1 & \gamma_{32} \\ \gamma_{13} & \gamma_{23} & -1 \\ \delta_{11} & \delta_{21} & \delta_{31} \\ \delta_{12} & \delta_{22} & \delta_{32} \\ \delta_{13} & \delta_{23} & \delta_{33} \\ \delta_{14} & \delta_{24} & \delta_{34} \\ \delta_{15} & \delta_{25} & \delta_{35} \\ \delta_{16} & \delta_{26} & \delta_{36} \\ \delta_{17} & \delta_{27} & \delta_{37} \\ \delta_{18} & \delta_{28} & \delta_{38} \\ \delta_{19} & \delta_{29} & \delta_{39} \\ \delta_{110} & \delta_{210} & \delta_{310} \\ \delta_{111} & \delta_{211} & \delta_{311} \\ \delta_{112} & \delta_{212} & \delta_{312} \\ \delta_{113} & \delta_{213} & \delta_{313} \\ \delta_{114} & \delta_{214} & \delta_{314} \\ \delta_{115} & \delta_{215} & \delta_{315} \\ \delta_{116} & \delta_{216} & \delta_{316} \\ \delta_{117} & \delta_{217} & \delta_{317} \\ \delta_{118} & \delta_{218} & \delta_{318} \\ \delta_{119} & \delta_{219} & \delta_{319} \\ \delta_{120} & \delta_{220} & \delta_{320} \\ \delta_{121} & \delta_{221} & \delta_{321} \\ \delta_{122} & \delta_{222} & \delta_{322} \\ \delta_{123} & \delta_{223} & \delta_{323} \\ \delta_{124} & \delta_{224} & \delta_{324} \\ \delta_{125} & \delta_{225} & \delta_{325} \\ \delta_{126} & \delta_{226} & \delta_{326} \\ \delta_{127} & \delta_{227} & \delta_{327} \\ \delta_{128} & \delta_{228} & \delta_{328} \\ \delta_{129} & \delta_{229} & \delta_{329} \\ \delta_{130} & \delta_{230} & \delta_{330} \\ \delta_{131} & \delta_{231} & \delta_{331} \\ \delta_{132} & \delta_{232} & \delta_{332} \\ \delta_{133} & \delta_{233} & \delta_{333} \\ \delta_{134} & \delta_{234} & \delta_{334} \\ \delta_{135} & \delta_{235} & \delta_{335} \\ \delta_{136} & \delta_{236} & \delta_{336} \end{bmatrix}_{(3+36) \times 3}$$

Los coeficientes de cada columna se corresponden a las variables TICs, PTF, EXP, constante, tamaño, cal_emp_intsec, i+d, apalancamiento, roe, cuota_mercado, hhi, stock_capital, capital_ext, merc_est_adq, ingreso_total, publicidad, gasto_fin_ban, dummies_industria y dummies_año, en ese orden. La matriz $\mathbf{B}$ solo incorpora la restricción de normalización.

La condición de rango requiere que, para un sistema de $\mathbf{G}$ ecuaciones, el rango de la matriz $R_h \mathbf{B}$

sea igual a $\mathbf{G} - 1$ para cada ecuación individual. En este caso, con $\mathbf{G} = 3$, se verifica:

$$\text{rango}(R_1\mathbf{B}) = \text{rango}(R_2\mathbf{B}) = \text{rango}(R_3\mathbf{B}) = 2.$$

A continuación, se presentan las matrices resultantes $R_h\mathbf{B}$ con las restricciones impuestas:

$$R_1\mathbf{B} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & \delta_{28} & 0 \\ 0 & \delta_{29} & 0 \\ 0 & 0 & \delta_{310} \\ 0 & \delta_{213} & 0 \\ 0 & \delta_{214} & \delta_{314} \end{bmatrix}_{6 \times 3} \Rightarrow \text{rango} = 2$$

$$R_2\mathbf{B} = \begin{bmatrix} 0 & 0 & -1 \\ \delta_{15} & 0 & 0 \\ \delta_{16} & 0 & 0 \\ 0 & 0 & \delta_{310} \\ \delta_{111} & 0 & 0 \\ \delta_{112} & 0 & \delta_{312} \end{bmatrix}_{6 \times 3} \Rightarrow \text{rango} = 2$$

$$R_3\mathbf{B} = \begin{bmatrix} \delta_{13} & \delta_{23} & 0 \\ \delta_{15} & 0 & 0 \\ \delta_{16} & 0 & 0 \\ 0 & \delta_{28} & 0 \\ 0 & \delta_{29} & 0 \\ \delta_{111} & 0 & 0 \\ 0 & \delta_{213} & 0 \end{bmatrix}_{7 \times 3} \Rightarrow \text{rango} = 2$$

Dado que se verifica la condición de rango para cada una de las ecuaciones del sistema, podemos esperar que un único vector de parámetros sea el vector solución del método de estimación GMM.

7.9 SEMs: validez de los instrumentos

Como se adelantó, se trabajo a nivel de cada ecuación para verificar la validez de los instrumentos.

ECUACIÓN 1: TICS

Relevancia de los instrumentos. El estadístico F no robusto de la primera etapa es de 395.65, muy por encima del valor de 104.7, como sugiere la literatura reciente sobre variables instrumentales, ver en el trabajo de (D. S. Lee et al., 2022) para referencia. Se reporta el estadístico robusto de Kleibergen-Paap ($F = 1506.41$) junto con los valores críticos de Stock-Yogo (2005) correspondientes al sesgo relativo del estimador IV (maximal IV relative bias). Dado que el valor observado de F supera ampliamente todos los valores críticos relevantes (13.91 para un sesgo del 5 %, 9.08 para un 10 %, 6.46 para un 20 % y 5.39 para un 30 %), se concluye que los instrumentos utilizados son suficientemente fuertes. En consecuencia, el sesgo relativo del estimador IV con respecto al estimador de OLS no supera el 5 %, lo cual valida la relevancia de los instrumentos y respalda la consistencia del estimador IV. La inferencia estadística derivada de este modelo es confiable en términos del sesgo estimado.

Dado que el F de Kleibergen-Paap supera el valor crítico correspondiente al 10 % del tamaño del test t , que es de 22.30, podemos afirmar que la distorsión en el tamaño del test t no supera el 10 %. Es decir, el uso de instrumentos no genera una inflación significativa en la probabilidad de cometer un error de tipo I al aplicar el test t . Por lo tanto, consideramos que los instrumentos son suficientemente fuertes y la inferencia basada en el test t es confiable.

Exogeneidad de los instrumentos. Dado que el número de instrumentos no triviales excede al de variables endógenas y se permite la presencia de heterocedasticidad, se utiliza la prueba de Hansen (estadístico J) para evaluar la validez de los instrumentos. La hipótesis nula de esta prueba establece que los instrumentos son exógenos. En este contexto, un p-valor elevado indica evidencia a favor de la validez de los instrumentos. En particular, el p-valor observado es de 0.2988, por lo que no se rechaza la hipótesis nula. Esto sugiere que los instrumentos utilizados son estadísticamente consistentes con la condición de exogeneidad requerida por el modelo. Se presentan a continuación las salidas del software Stata.

First-stage regressions

Number of obs = 26,928
 F(30, 26897) = 395.65
 Prob > F = 0.0000
 R-squared = 0.5007
 Adj R-squared = 0.5001
 Root MSE = 0.5318

PTF	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
ltamanio	-.3562118	.0076611	-46.50	0.000	-.3712279	-.3411956
cal_emp_intsec	.0078665	.001321	5.95	0.000	.0052772	.0104559
limasd	.0142883	.0015497	9.22	0.000	.0112508	.0173258
lapalancamiento	-.0999003	.0080477	-12.41	0.000	-.1156742	-.0841265
lroe	.5116502	.0215166	23.78	0.000	.4694767	.5538238
industrial	.0254152	.0233057	1.09	0.275	-.0202652	.0710956
industria2	.2141497	.0368289	5.81	0.000	.1419631	.2863363

industria3	-.2793784	.0418552	-6.67	0.000	-.3614169	-.19734
industria4	-.0991896	.0233455	-4.25	0.000	-.1449479	-.0534313
industria6	-.5490576	.0229824	-23.89	0.000	-.5941044	-.5040108
industria7	-.3575853	.0345962	-10.34	0.000	-.4253956	-.289775
industria8	-.3909834	.0224807	-17.39	0.000	-.4350468	-.3469201
industria9	-.3751857	.0256597	-14.62	0.000	-.4254799	-.3248914
industria10	-.5313524	.0504313	-10.54	0.000	-.6302004	-.4325044
industria11	-.4304716	.0193097	-22.29	0.000	-.4683196	-.3926236
industria13	-.3258476	.0282779	-11.52	0.000	-.3812737	-.2704215
industria14	-.1490459	.0222948	-6.69	0.000	-.1927449	-.1053468
industria15	-.2755238	.0370329	-7.44	0.000	-.3481101	-.2029374
industria16	-.4586445	.1507197	-3.04	0.002	-.7540629	-.1632261
industria17	-.2506645	.0211135	-11.86	0.000	-.2920902	-.2092388
anio2013	.08801	.0125948	6.99	0.000	.0633236	.1126964
anio2014	.1161479	.0125486	9.26	0.000	.0915519	.1407438
anio2015	.1020175	.0122125	8.35	0.000	.0780804	.1259546
anio2016	.0973895	.0125992	7.73	0.000	.0726944	.1220846
anio2017	.1209627	.012873	9.40	0.000	.095731	.1461943
anio2018	.1526709	.0125522	12.16	0.000	.1280678	.1772739
anio2019	.1685809	.0128615	13.11	0.000	.1433716	.1937901
cuota_mercado	.1530406	.0252076	6.07	0.000	.1036323	.2024489
merc_est_adq	-.0031635	.0008085	-3.91	0.000	-.0047482	-.0015788
ling_total	.38159	.006031	63.27	0.000	.3697689	.3934111
_cons	5.084512	.087462	58.13	0.000	4.913082	5.255942

Instrumental variables 2SLS regression

Number of obs = 26,928
Wald chi2(28) = 1939.75
Prob > chi2 = 0.0000
R-squared = 0.0612
Root MSE = 2.91

TIC	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
PTF	1.175567	.0622614	18.88	0.000	1.053537	1.297597
ltamano	.5900432	.0253509	23.28	0.000	.5403562	.6397301
cal_emp_intsec	-.011114	.0100843	-1.10	0.270	-.0308788	.0086507
limasd	.0097572	.004268	2.29	0.022	.0013921	.0181223
lapalancamiento	.1781979	.0299546	5.95	0.000	.119488	.2369078
lroe	-.8355721	.0833028	-10.03	0.000	-.9988425	-.6723017
industria1	-.4853224	.0943992	-5.14	0.000	-.6703415	-.3003033
industria2	-.550634	.1277159	-4.31	0.000	-.8009527	-.3003154
industria3	.5600361	.1233217	4.54	0.000	.31833	.8017422
industria4	.4773903	.0851103	5.61	0.000	.3105772	.6442035
industria6	-.4478978	.1619634	-2.77	0.006	-.7653403	-.1304553
industria7	.1725907	.114943	1.50	0.133	-.0526934	.3978748
industria8	-.4540403	.0818089	-5.55	0.000	-.6143828	-.2936978
industria9	.6348269	.1001312	6.34	0.000	.4385733	.8310805
industria10	.311764	.2075084	1.50	0.133	-.0949449	.718473
industria11	-.5158681	.085655	-6.02	0.000	-.6837489	-.3479873
industria13	-.0970196	.1717252	-0.56	0.572	-.4335949	.2395556
industria14	-.5650416	.1075615	-5.25	0.000	-.7758582	-.354225
industria15	-.2057605	.2661656	-0.77	0.439	-.7274355	.3159146
industria16	.0279624	.661606	0.04	0.966	-1.268762	1.324686
industria17	-.3733251	.0943852	-3.96	0.000	-.5583166	-.1883336
anio2013	.0899692	.0735011	1.22	0.221	-.0540903	.2340287
anio2014	.0329725	.0752325	0.44	0.661	-.1144805	.1804255
anio2015	.3435557	.0736603	4.66	0.000	.1991842	.4879273
anio2016	.4311703	.0732567	5.89	0.000	.2875898	.5747508
anio2017	.5812799	.0709144	8.20	0.000	.4422903	.7202695
anio2018	.6923505	.0675464	10.25	0.000	.559962	.824739
anio2019	.8561721	.0652261	13.13	0.000	.7283312	.9840129
_cons	-14.30222	.6909436	-20.70	0.000	-15.65644	-12.94799

Endogenous: PTF

Exogenous: ltamano cal_emp_intsec limasd lapalancamiento lroe industria1
industria2 industria3 industria4 industria6 industria7 industria8
industria9 industria10 industria11 industria13 industria14
industria15 industria16 industria17 anio2013 anio2014 anio2015
anio2016 anio2017 anio2018 anio2019 cuota_mercado merc_est_adq

ling_total

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
 Statistics robust to heteroskedasticity

Total (centered) SS	=	242876.1424	Number of obs =	26928
Total (uncentered) SS	=	243972.9643	F(28, 26899) =	69.20
Residual SS	=	228023.1411	Prob > F	= 0.0000
			Centered R2	= 0.0612
			Uncentered R2	= 0.0654
			Root MSE	= 2.91

	TIC	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
PTF		1.175567	.0622614	18.88	0.000	1.053537	1.297597
ltamano		.5900432	.0253509	23.28	0.000	.5403562	.6397301
cal_emp_intsec		-.011114	.0100843	-1.10	0.270	-.0308788	.0086507
limasd		.0097572	.004268	2.29	0.022	.0013921	.0181223
lapalancamiento		.1781979	.0299546	5.95	0.000	.119488	.2369078
lroe		-.8355721	.0833028	-10.03	0.000	-.9988425	-.6723017
industrial		-.4853224	.0943992	-5.14	0.000	-.6703415	-.3003033
industria2		-.550634	.1277159	-4.31	0.000	-.8009527	-.3003154
industria3		.5600361	.1233217	4.54	0.000	.31833	.8017422
industria4		.4773903	.0851103	5.61	0.000	.3105772	.6442035
industria6		-.4478978	.1619634	-2.77	0.006	-.7653403	-.1304553
industria7		.1725907	.114943	1.50	0.133	-.0526934	.3978748
industria8		-.4540403	.0818089	-5.55	0.000	-.6143828	-.2936978
industria9		.6348269	.1001312	6.34	0.000	.4385733	.8310805
industria10		.311764	.2075084	1.50	0.133	-.0949449	.718473
industria11		-.5158681	.085655	-6.02	0.000	-.6837489	-.3479873
industria13		-.0970196	.1717252	-0.56	0.572	-.4335949	.2395556
industria14		-.5650416	.1075615	-5.25	0.000	-.7758582	-.354225
industria15		-.2057605	.2661656	-0.77	0.439	-.7274355	.3159146
industria16		.0279624	.661606	0.04	0.966	-1.268762	1.324686
industria17		-.3733251	.0943852	-3.96	0.000	-.5583166	-.1883336
anio2013		.0899692	.0735011	1.22	0.221	-.0540903	.2340287
anio2014		.0329725	.0752325	0.44	0.661	-.1144805	.1804255
anio2015		.3435557	.0736603	4.66	0.000	.1991842	.4879273
anio2016		.4311703	.0732567	5.89	0.000	.2875898	.5747508
anio2017		.5812799	.0709144	8.20	0.000	.4422903	.7202695
anio2018		.6923505	.0675464	10.25	0.000	.559962	.824739
anio2019		.8561721	.0652261	13.13	0.000	.7283312	.9840129
_cons		-14.30222	.6909436	-20.70	0.000	-15.65644	-12.94799

Underidentification test (Kleibergen-Paap rk LM statistic): 2136.253
 Chi-sq(3) P-val = 0.0000

Weak identification test (Cragg-Donald Wald F statistic): 5139.826
 (Kleibergen-Paap rk Wald F statistic): 1506.418

Stock-Yogo weak ID test critical values:

5% maximal IV relative bias	13.91
10% maximal IV relative bias	9.08
20% maximal IV relative bias	6.46
30% maximal IV relative bias	5.39
10% maximal IV size	22.30
15% maximal IV size	12.83
20% maximal IV size	9.54
25% maximal IV size	7.80

Source: Stock-Yogo (2005). Reproduced by permission.

NB: Critical values are for Cragg-Donald F statistic and i.i.d. errors.

Hansen J statistic (overidentification test of all instruments): 2.416
 Chi-sq(2) P-val = 0.2988

Instrumented: PTF
 Included instruments: ltamano cal_emp_intsec limasd lapalancamiento lroe
 industrial industria2 industria3 industria4 industria6

```

industria7 industria8 industria9 industrial0 industrial11
industria13 industria14 industria15 industrial16
industria17 anio2013 anio2014 anio2015 anio2016 anio2017
anio2018 anio2019
Excluded instruments: cuota_mercado merc_est_adq ling_total

```

ECUACIÓN 2: PTF

Relevancia de los instrumentos. El estadístico F no robusto de la primera etapa es de 80.75, algo por debajo del valor sugerido por la literatura reciente de 104.7, en base al trabajo de (D. S. Lee et al., 2022). Sin embargo, el estadístico F de Kleibergen-Paap es de 332.27. La salida de Stata no nos reportó el valor de tabla de Stock-Yogo (2005) para el máximo sesgo relativo con respecto a OLS. Pero es probable que el $F = 332.27$ sea superior a estos valores de tabla.

Dado que el F de Kleibergen-Paap supera el valor crítico correspondiente al 10 % del tamaño del test t , que es de 19.93, podemos afirmar que la distorsión en el tamaño del test t no supera el 10 %. Es decir, el uso de instrumentos no genera una inflación significativa en la probabilidad de cometer un error de tipo I al aplicar el test t . Por lo tanto, la inferencia basada en el test t es confiable.

Exogeneidad de los instrumentos. Dado que el número de instrumentos no triviales excede al de variables endógenas y se permite la presencia de heterocedasticidad, se utiliza la prueba de Hansen (estadístico J) para evaluar la validez de los instrumentos. La hipótesis nula de esta prueba establece que los instrumentos son exógenos. En este contexto, un p-valor elevado indica evidencia a favor de la validez de los instrumentos. En particular, el p-valor observado es de 0.3868, por lo que no se rechaza la hipótesis nula. Esto sugiere que los instrumentos utilizados son estadísticamente consistentes con la condición de exogeneidad requerida por el modelo. Se presentan a continuación las salidas del software Stata.

First-stage regressions

						Number of obs = 29,751
						F(30, 29720) = 80.75
						Prob > F = 0.0000
						R-squared = 0.1352
						Adj R-squared = 0.1343
						Root MSE = 2.9678
	TIC	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]
ltamanio		.2152104	.025328	8.50	0.000	.1655663 .2648544
cal_emp_intsec		.0024093	.0084325	0.29	0.775	-.0141187 .0189373
lstock_capital		.1639833	.0125197	13.10	0.000	.1394442 .1885224
cuota_mercado		.1082522	.1104058	0.98	0.327	-.108148 .3246523
hhi		.5378283	.1055486	5.10	0.000	.3309484 .7447081
limasd		.0168935	.003978	4.25	0.000	.0090965 .0246905
industria1		-.3446763	.0912724	-3.78	0.000	-.5235742 -.1657784
industria2		-.373487	.1148991	-3.25	0.001	-.5986942 -.1482797
industria3		-.0879799	.119813	-0.73	0.463	-.3228187 .1468589
industria4		.3313278	.0778393	4.26	0.000	.1787593 .4838963
industria6		-1.474684	.1409146	-10.47	0.000	-1.750882 -1.198485
industria7		-.4346899	.1288719	-3.37	0.001	-.6872844 -.1820954
industria8		-.7383806	.0770493	-9.58	0.000	-.8894005 -.5873606
industria9		.2419844	.0907226	2.67	0.008	.0641641 .4198047
industria10		-.7504232	.1830945	-4.10	0.000	-1.109297 -.3915499
industria11		-1.224327	.0806607	-15.18	0.000	-1.382425 -1.066228

industrial3	-.8003387	.1581373	-5.06	0.000	-1.110295	-.4903826
industrial4	-.7038944	.1066548	-6.60	0.000	-.9129425	-.4948462
industrial5	-.8129109	.251304	-3.23	0.001	-1.305478	-.320344
industrial6	-1.551054	.6590557	-2.35	0.019	-2.842832	-.2592758
industrial7	-.7664746	.0923416	-8.30	0.000	-.9474682	-.5854811
anio2013	.3254261	.071835	4.53	0.000	.1846263	.4662259
anio2014	.314868	.0731815	4.30	0.000	.171429	.458307
anio2015	.5252798	.0715464	7.34	0.000	.3850458	.6655138
anio2016	.5706941	.0722656	7.90	0.000	.4290504	.7123377
anio2017	.7845652	.0699037	11.22	0.000	.6475509	.9215796
anio2018	.9630868	.0659897	14.59	0.000	.833744	1.09243
anio2019	1.107965	.0639305	17.33	0.000	.9826588	1.233272
lpublicidad	.0327994	.0024488	13.39	0.000	.0279997	.0375991
gasto_fin	.0944809	.0043559	21.69	0.000	.0859431	.1030187
_cons	-4.398449	.1961449	-22.42	0.000	-4.782901	-4.013996

Instrumental variables 2SLS regression

Number of obs = 29,751
Wald chi2(29) = 2879.37
Prob > chi2 = 0.0000
R-squared = .
Root MSE = 1.0785

PTF	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
TIC	.2904042	.0139153	20.87	0.000	.2631307	.3176778
ltamano	-.1274227	.0122697	-10.39	0.000	-.1514708	-.1033746
cal_emp_intsec	.0246728	.0041871	5.89	0.000	.0164663	.0328794
lstock_capital	-.1025302	.0058301	-17.59	0.000	-.113957	-.0911034
cuota_mercado	.7443504	.0532258	13.98	0.000	.6400298	.8486711
hhi	-.2577618	.0400124	-6.44	0.000	-.3361848	-.1793389
limasd	.0229462	.002235	10.27	0.000	.0185657	.0273266
industrial1	-.1870638	.0394552	-4.74	0.000	-.2643945	-.109733
industrial2	.2755706	.0608299	4.53	0.000	.1563462	.394795
industrial3	-.3518496	.0686157	-5.13	0.000	-.4863339	-.2173653
industrial4	-.4625068	.040131	-11.52	0.000	-.5411622	-.3838514
industrial6	-.4644931	.0566797	-8.20	0.000	-.5755833	-.3534029
industrial7	-.3950686	.0605484	-6.52	0.000	-.5137412	-.2763959
industrial8	.0775506	.0375172	2.07	0.039	.0040183	.1510829
industrial9	-.8345384	.0399706	-20.88	0.000	-.9128794	-.7561974
industrial10	-.4698995	.0816512	-5.75	0.000	-.629933	-.3098866
industrial11	-.1607047	.039762	-4.04	0.000	-.2386367	-.0827727
industrial13	-.4690337	.0601598	-7.80	0.000	-.5869447	-.3511227
industrial14	-.2119064	.0435072	-4.87	0.000	-.297179	-.1266339
industrial15	-.3589192	.0852715	-4.21	0.000	-.5260483	-.1917902
industrial16	.1359847	.2232824	0.61	0.543	-.3016408	.5736102
industrial17	.0553612	.0420849	1.32	0.188	-.0271236	.1378461
anio2013	.0633718	.0256278	2.47	0.013	.0131422	.1136014
anio2014	.0886573	.0259044	3.42	0.001	.0378857	.139429
anio2015	-.0672932	.02567	-2.62	0.009	-.1176055	-.0169809
anio2016	-.0980588	.0262888	-3.73	0.000	-.149584	-.0465337
anio2017	-.1308441	.0265078	-4.94	0.000	-.1827984	-.0788897
anio2018	-.1423352	.0264101	-5.39	0.000	-.194098	-.0905725
anio2019	-.1780711	.0272613	-6.53	0.000	-.2315023	-.1246398
_cons	12.59837	.0975258	129.18	0.000	12.40723	12.78952

Endogenous: TIC

Exogenous: ltamano cal_emp_intsec lstock_capital cuota_mercado hhi limasd
industrial2 industrial3 industrial4 industrial6 industrial7
industrial8 industrial9 industrial10 industrial11 industrial13
industrial14 industrial15 industrial16 industrial17 anio2013 anio2014
anio2015 anio2016 anio2017 anio2018 anio2019 lpublicidad gasto_fin

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity

Total (centered) SS	=	18568.33961	Number of obs =	29751
Total (uncentered) SS	=	3254353.271	F(29, 29721) =	99.19
Residual SS	=	34608.22366	Prob > F	= 0.0000
			Centered R2	= -0.8638
			Uncentered R2	= 0.9894
			Root MSE	= 1.079

PTF	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
TIC	.2904042	.0139153	20.87	0.000	.2631307	.3176778
ltamania	-.1274227	.0122697	-10.39	0.000	-.1514708	-.1033746
cal_emp_intsec	.0246728	.0041871	5.89	0.000	.0164663	.0328794
lstock_capital	-.1025302	.0058301	-17.59	0.000	-.113957	-.0911034
cuota_mercado	.7443504	.0532258	13.98	0.000	.6400298	.8486711
hhi	-.2577618	.0400124	-6.44	0.000	-.3361848	-.1793389
limasd	.0229462	.002235	10.27	0.000	.0185657	.0273266
industrial1	-.1870638	.0394552	-4.74	0.000	-.2643945	-.109733
industria2	.2755706	.0608299	4.53	0.000	.1563462	.394795
industria3	-.3518496	.0686157	-5.13	0.000	-.4863339	-.2173653
industria4	-.4625068	.040131	-11.52	0.000	-.5411622	-.3838514
industria6	-.4644931	.0566797	-8.20	0.000	-.5755833	-.3534029
industria7	-.3950686	.0605484	-6.52	0.000	-.5137412	-.2763959
industria8	.0775506	.0375172	2.07	0.039	.0040183	.1510829
industria9	-.8345384	.0399706	-20.88	0.000	-.9128794	-.7561974
industria10	-.4698995	.0816512	-5.75	0.000	-.629933	-.309866
industria11	-.1607047	.039762	-4.04	0.000	-.2386367	-.0827727
industria13	-.4690337	.0601598	-7.80	0.000	-.5869447	-.3511227
industria14	-.2119064	.0435072	-4.87	0.000	-.297179	-.1266339
industria15	-.3589192	.0852715	-4.21	0.000	-.5260483	-.1917902
industria16	.1359847	.2232824	0.61	0.543	-.3016408	.5736102
industria17	.0553612	.0420849	1.32	0.188	-.0271236	.1378461
anio2013	.0633718	.0256278	2.47	0.013	.0131422	.1136014
anio2014	.0886573	.0259044	3.42	0.001	.0378857	.139429
anio2015	-.0672932	.02567	-2.62	0.009	-.1176055	-.0169809
anio2016	-.0980588	.0262888	-3.73	0.000	-.149584	-.0465337
anio2017	-.1308441	.0265078	-4.94	0.000	-.1827984	-.0788897
anio2018	-.1423352	.0264101	-5.39	0.000	-.194098	-.0905725
anio2019	-.1780711	.0272613	-6.53	0.000	-.2315023	-.1246398
_cons	12.59837	.0975258	129.18	0.000	12.40723	12.78952

Underidentification test (Kleibergen-Paap rk LM statistic): 637.097
Chi-sq(2) P-val = 0.0000

Weak identification test (Cragg-Donald Wald F statistic): 458.272
(Kleibergen-Paap rk Wald F statistic): 332.278
Stock-Yogo weak ID test critical values: 10% maximal IV size 19.93
15% maximal IV size 11.59
20% maximal IV size 8.75
25% maximal IV size 7.25

Source: Stock-Yogo (2005). Reproduced by permission.

NB: Critical values are for Cragg-Donald F statistic and i.i.d. errors.

Hansen J statistic (overidentification test of all instruments): 0.749
Chi-sq(1) P-val = 0.3868

Instrumented: TIC
Included instruments: ltamania cal_emp_intsec lstock_capital cuota_mercado hhi
limasd industrial1 industria2 industria3 industria4
industria6 industria7 industria8 industria9 industria10
industria11 industria13 industria14 industria15
industria16 industria17 anio2013 anio2014 anio2015
anio2016 anio2017 anio2018 anio2019
Excluded instruments: lpublicidad gasto_fin

ECUACIÓN 3: EXP

Relevancia de los instrumentos. En esta ecuación tenemos dos primeras etapas. En la primera

etapa para la ecuación de TICs el valor F no robusto es de 87.94 mientras que el valor F no robusto de la primera etapa para la ecuación de PTF es de 322.51. Sin embargo, el estadístico F de Kleibergen-Paap es de 162.12. La salida de Stata no nos reportó el valor de tabla de Stock-Yogo (2005) para el máximo sesgo relativo con respecto a OLS. Pero es probable que el $F = 162.12$ sea superior a estos valores de tabla.

Dado que el F de Kleibergen-Paap supera el valor crítico correspondiente al 10% del tamaño del test t , que es de 13.43, podemos afirmar que la distorsión en el tamaño del test t no supera el 10%. Es decir, el uso de instrumentos no genera una inflación significativa en la probabilidad de cometer un error de tipo I al aplicar el test t . Por lo tanto, la inferencia basada en el test t es confiable.

Exogeneidad de los instrumentos. Dado que el número de instrumentos no triviales excede al de variables endógenas y se permite la presencia de heterocedasticidad, se utiliza la prueba de Hansen (estadístico J) para evaluar la validez de los instrumentos. La hipótesis nula de esta prueba establece que los instrumentos son exógenos. En este contexto, un p-valor elevado indica evidencia a favor de la validez de los instrumentos. En particular, el p-valor observado es de 0.2037, por lo que no se rechaza la hipótesis nula. Esto sugiere que los instrumentos utilizados son estadísticamente consistentes con la condición de exogeneidad requerida por el modelo. Se presentan a continuación las salidas del software Stata.

First-stage regressions

Number of obs = 29,733
F(28, 29704) = 87.94
Prob > F = 0.0000
R-squared = 0.1331
Adj R-squared = 0.1323
Root MSE = 2.9717

TIC	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
ltamano	.169616	.0233133	7.28	0.000	.1239209	.2153112
limasd	.0243188	.0038339	6.34	0.000	.0168042	.0318335
capital_ext	-.0007602	.0004711	-1.61	0.107	-.0016836	.0001632
industrial1	-.4465072	.0905512	-4.93	0.000	-.6239915	-.269023
industrial2	-.3817656	.1166918	-3.27	0.001	-.6104867	-.1530445
industrial3	.2853781	.1120764	2.55	0.011	.0657034	.5050527
industrial4	.3225945	.0753462	4.28	0.000	.1749126	.4702764
industrial6	-1.231496	.1400559	-8.79	0.000	-1.506012	-.9569803
industrial7	-.2762892	.1281966	-2.16	0.031	-.5275603	-.0250182
industrial8	-1.0222	.0806265	-12.68	0.000	-1.180232	-.8641691
industrial9	.4465805	.0913724	4.89	0.000	.2674866	.6256744
industrial10	-.3980676	.1851969	-2.15	0.032	-.7610616	-.0350737
industrial11	-1.116213	.0790825	-14.11	0.000	-1.271218	-.9612073
industrial13	-.5388745	.1580274	-3.41	0.001	-.8486151	-.2291339
industrial14	-.6792907	.1043604	-6.51	0.000	-.8838417	-.4747397
industrial15	-.605041	.2543477	-2.38	0.017	-1.103574	-.1065083
industrial16	-.852329	.5656466	-1.51	0.132	-1.961021	.2563631
industrial17	-.7987736	.0895445	-8.92	0.000	-.9742847	-.6232624
anio2013	.2633893	.0716141	3.68	0.000	.1230224	.4037561
anio2014	.2427642	.0728484	3.33	0.001	.0999781	.3855503
anio2015	.4948947	.0711947	6.95	0.000	.35535	.6344394
anio2016	.5487664	.0719505	7.63	0.000	.4077403	.6897925
anio2017	.765934	.0694281	11.03	0.000	.6298519	.9020162
anio2018	.9282396	.0658299	14.10	0.000	.7992101	1.057269
anio2019	1.072669	.0635964	16.87	0.000	.9480172	1.197321
cuota_mercado	.1649609	.0905816	1.82	0.069	-.012583	.3425047
ling_total	.3271788	.0219168	14.93	0.000	.2842209	.3701367

gasto_fin	.0953026	.0044479	21.43	0.000	.0865846	.1040206
_cons	-7.212172	.3356129	-21.49	0.000	-7.869988	-6.554356

Number of obs = 29,733
 F(28, 29704) = 322.51
 Prob > F = 0.0000
 R-squared = 0.4271
 Adj R-squared = 0.4265
 Root MSE = 0.5983

PTF	Coefficient	Robust std. err.	t	P> t	[95% conf. interval]	
ltamano	-.3598639	.0084504	-42.59	0.000	-.376427	-.3433008
limasd	.0136454	.0016563	8.24	0.000	.0103991	.0168918
capital_ext	.0013311	.0001942	6.85	0.000	.0009503	.0017118
industrial1	.0977592	.0250771	3.90	0.000	.048607	.1469114
industrial2	.2323555	.0408029	5.69	0.000	.15238	.3123311
industrial3	-.253534	.0454112	-5.58	0.000	-.3425421	-.164526
industrial4	-.093526	.0258854	-3.61	0.000	-.1442626	-.0427894
industrial6	-.5450591	.0263364	-20.70	0.000	-.5966796	-.4934385
industrial7	-.327493	.0355726	-9.21	0.000	-.3972168	-.2577691
industrial8	-.4344373	.0222015	-19.57	0.000	-.4779531	-.3909214
industrial9	-.2378025	.0266973	-8.91	0.000	-.2901304	-.1854745
industrial10	-.4851958	.0526291	-9.22	0.000	-.5883511	-.3820404
industrial11	-.4828179	.0211504	-22.83	0.000	-.5242737	-.4413621
industrial13	-.3581491	.0312257	-11.47	0.000	-.4193529	-.2969454
industrial14	-.097523	.0242984	-4.01	0.000	-.145149	-.049897
industrial15	-.2763126	.0437596	-6.31	0.000	-.3620833	-.1905419
industrial16	-.6327091	.1504718	-4.20	0.000	-.9276405	-.3377777
industrial17	-.2194233	.0231853	-9.46	0.000	-.2648675	-.1739792
anio2013	.0693435	.0134597	5.15	0.000	.0429619	.0957251
anio2014	.0917147	.0133295	6.88	0.000	.0655883	.1178411
anio2015	.0617876	.0131185	4.71	0.000	.0360748	.0875004
anio2016	.0605375	.0134872	4.49	0.000	.034102	.0869729
anio2017	.0843433	.0136065	6.20	0.000	.057674	.1110127
anio2018	.1071367	.0134267	7.98	0.000	.0808197	.1334537
anio2019	.1181748	.0137856	8.57	0.000	.0911544	.1451952
cuota_mercado	.1637725	.028763	5.69	0.000	.1073959	.2201492
ling_total	.3819321	.0086114	44.35	0.000	.3650533	.3988109
gasto_fin	-.0044034	.0008707	-5.06	0.000	-.0061099	-.0026968
_cons	5.164363	.1231098	41.95	0.000	4.923062	5.405663

Instrumental variables 2SLS regression

Number of obs = 29,733
 Wald chi2(27) = 6457.79
 Prob > chi2 = 0.0000
 R-squared = .
 Root MSE = 7.2346

LEXP	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
TIC	1.290096	.0863121	14.95	0.000	1.120927	1.459264
PTF	3.023171	.1702943	17.75	0.000	2.6894	3.356942
ltamano	.0056552	.0709095	0.08	0.936	-.1333248	.1446352
limasd	.1031314	.0209197	4.93	0.000	.0621295	.1441334
capital_ext	.0537896	.0021091	25.50	0.000	.0496559	.0579234
industrial1	-.152167	.2823783	-5.39	0.000	-2.075121	-.9682183
industrial2	-3.703707	.3411104	-10.86	0.000	-4.372271	-3.035143
industrial3	-2.957936	.4520057	-6.54	0.000	-3.843851	-2.072021
industrial4	-.4346399	.3108266	-1.40	0.162	-1.043849	.1745691
industrial6	2.183132	.3343751	6.53	0.000	1.527769	2.838496
industrial7	-1.631348	.3454928	-4.72	0.000	-2.308501	-.9541942
industrial8	-1.586013	.2385716	-6.65	0.000	-2.053605	-1.118421
industrial9	-2.397567	.2735176	-8.77	0.000	-2.933652	-1.861482
industrial10	1.263643	.7351585	1.72	0.086	-.1772408	2.704527
industrial11	3.255648	.2602147	12.51	0.000	2.745637	3.76566
industrial13	-1.157867	.3284802	-3.52	0.000	-1.801676	-.5140575

industrial4	-2.531427	.2764862	-9.16	0.000	-3.07333	-1.989524
industrial5	-.9544964	.592132	-1.61	0.107	-2.115054	.206061
industrial6	-3.547042	1.217339	-2.91	0.004	-5.932984	-1.161101
industrial7	1.107166	.2871082	3.86	0.000	.544444	1.669888
anio2013	-.6217697	.162416	-3.83	0.000	-.9400992	-.3034402
anio2014	-.6672717	.1618719	-4.12	0.000	-.9845347	-.3500086
anio2015	-.8710213	.166374	-5.24	0.000	-1.197108	-.5449344
anio2016	-.8383878	.1704094	-4.92	0.000	-1.172384	-.5043915
anio2017	-1.167402	.1743009	-6.70	0.000	-1.509025	-.8257784
anio2018	-1.675357	.1743243	-9.61	0.000	-2.017026	-1.333687
anio2019	-2.035312	.1800757	-11.30	0.000	-2.388253	-1.68237
_cons	-27.42348	1.958257	-14.00	0.000	-31.2616	-23.58537

Endogenous: TIC PTF

Exogenous: ltamano limasd capital_ext industrial industria2 industria3
industria4 industria6 industria7 industria8 industria9
industria10 industria11 industria13 industria14 industria15
industria16 industria17 anio2013 anio2014 anio2015 anio2016
anio2017 anio2018 anio2019 cuota_mercado ling_total gasto_fin

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity

		Number of obs =	29733
		F(27, 29705) =	238.95
		Prob > F =	0.0000
Total (centered) SS	=	Centered R2	= -0.0407
Total (uncentered) SS	=	Uncentered R2	= 0.1812
Residual SS	=	Root MSE	= 7.235

LEXP	Coefficient	Robust std. err.	z	P> z	[95% conf. interval]	
TIC	1.290096	.0863121	14.95	0.000	1.120927	1.459264
PTF	3.023171	.1702943	17.75	0.000	2.6894	3.356942
ltamano	.0056552	.0709095	0.08	0.936	-.1333248	.1446352
limasd	.1031314	.0209197	4.93	0.000	.0621295	.1441334
capital_ext	.0537896	.0021091	25.50	0.000	.0496559	.0579234
industrial	-1.52167	.2823783	-5.39	0.000	-2.075121	-.9682183
industria2	-3.703707	.3411104	-10.86	0.000	-4.372271	-3.035143
industria3	-2.957936	.4520057	-6.54	0.000	-3.843851	-2.072021
industria4	-.4346399	.3108266	-1.40	0.162	-1.043849	.1745691
industria6	2.183132	.3343751	6.53	0.000	1.527769	2.838496
industria7	-1.631348	.3454928	-4.72	0.000	-2.308501	-.9541942
industria8	-1.586013	.2385716	-6.65	0.000	-2.053605	-1.118421
industria9	-2.397567	.2735176	-8.77	0.000	-2.933652	-1.861482
industria10	1.263643	.7351585	1.72	0.086	-.1772408	2.704527
industria11	3.255648	.2602147	12.51	0.000	2.745637	3.76566
industria13	-1.157867	.3284802	-3.52	0.000	-1.801676	-.5140575
industria14	-2.531427	.2764862	-9.16	0.000	-3.07333	-1.989524
industria15	-.9544964	.592132	-1.61	0.107	-2.115054	.206061
industria16	-3.547042	1.217339	-2.91	0.004	-5.932984	-1.161101
industria17	1.107166	.2871082	3.86	0.000	.544444	1.669888
anio2013	-.6217697	.162416	-3.83	0.000	-.9400992	-.3034402
anio2014	-.6672717	.1618719	-4.12	0.000	-.9845347	-.3500086
anio2015	-.8710213	.166374	-5.24	0.000	-1.197108	-.5449344
anio2016	-.8383878	.1704094	-4.92	0.000	-1.172384	-.5043915
anio2017	-1.167402	.1743009	-6.70	0.000	-1.509025	-.8257784
anio2018	-1.675357	.1743243	-9.61	0.000	-2.017026	-1.333687
anio2019	-2.035312	.1800757	-11.30	0.000	-2.388253	-1.68237
_cons	-27.42348	1.958257	-14.00	0.000	-31.2616	-23.58537

Underidentification test (Kleibergen-Paap rk LM statistic): 481.601
Chi-sq(2) P-val = 0.0000

Weak identification test (Cragg-Donald Wald F statistic): 251.919
(Kleibergen-Paap rk Wald F statistic): 162.127

Stock-Yogo weak ID test critical values: 10% maximal IV size	13.43
15% maximal IV size	8.18
20% maximal IV size	6.40
25% maximal IV size	5.45

Source: Stock-Yogo (2005). Reproduced by permission.

NB: Critical values are for Cragg-Donald F statistic and i.i.d. errors.

Hansen J statistic (overidentification test of all instruments):	1.616
Chi-sq(1) P-val =	0.2037

Instrumented: TIC PTF

Included instruments: ltamano limasd capital_ext industrial1 industria2
 industria3 industria4 industria6 industria7 industria8
 industria9 industria10 industria11 industria13 industria14
 industria15 industria16 industria17 anio2013 anio2014
 anio2015 anio2016 anio2017 anio2018 anio2019

Excluded instruments: cuota_mercado ling_total gasto_fin

7.10 SEMs: estimación por GMM-3SLS

Se presentan los resultados de la estimación del SEMs dado en (8) por GMM-Minimos Cuadrados en Tres Etapas.

Variable	eq: TICs	eq: PTF	eq: ln(EXP)
PTF	1.1770*** (0.0440)		3.3694*** (0.1679)
TICs		0.2914*** (0.0130)	1.4269*** (0.1018)
ln(EXP)			
ln(tamano)	0.5891*** (0.0200)	-0.1338*** (0.0089)	0.0852 (0.0759)
cal_emp_intsec	-0.0106 (0.0066)	0.0234*** (0.0023)	
ln(i + d)	0.0096 (0.0080)	0.0216*** (0.0028)	0.0972*** (0.0204)
ln(apalancamiento)	0.1819*** (0.0294)		
ln(roe)	-0.8374*** (0.0575)		
cuota_mercado		0.7673*** (0.0340)	
hhi		-0.2305*** (0.0406)	
ln(stock_capital)		-0.0950*** (0.0050)	
capital_ext			0.0498*** (0.0019)
industria1	-0.4856*** (0.1135)	-0.2691*** (0.0405)	-1.3853*** (0.2934)
industria2	-0.5511*** (0.1503)	0.2153*** (0.0529)	-3.7779*** (0.3865)
industria3	0.5606** (0.2228)	-0.3539*** (0.0786)	-2.6687*** (0.5707)
industria4	0.4769*** (0.1280)	-0.4814*** (0.0450)	-0.0722 (0.3303)
industria6	-0.4463*** (0.1307)	-0.5219*** (0.0478)	2.5829*** (0.3368)
industria7	0.1740 (0.1719)	-0.5087*** (0.0602)	-1.7501*** (0.4388)
industria8	-0.4544*** (0.0938)	0.0144 (0.0339)	-1.4447*** (0.2423)
industria9	0.6370*** (0.1499)	-0.8839*** (0.0509)	-2.3975*** (0.3858)

Cuadro 14 – continuación

Variable	eq: TICs	eq: PTF	eq: ln(EXP)
industria10	0.3129 (0.2788)	-0.5694*** (0.0986)	1.7879** (0.7107)
industria11	-0.5151*** (0.0986)	-0.2095*** (0.0372)	3.5210*** (0.2564)
industria13	-0.0961 (0.1530)	-0.5220*** (0.0541)	-1.1377*** (0.3900)
industria14	-0.5635*** (0.1235)	-0.2343*** (0.0445)	-2.8193*** (0.3207)
industria15	-0.2047 (0.2356)	-0.4194*** (0.0834)	-1.0062* (0.6007)
industria16	0.0280 (0.4098)	0.0806 (0.1481)	-3.1410*** (1.0464)
industria17	-0.3730*** (0.1087)	-0.0200 (0.0396)	1.2254*** (0.2795)
anio2013	0.0904 (0.0663)	0.0710*** (0.0234)	-0.6733*** (0.1690)
anio2014	0.0333 (0.0660)	0.1053*** (0.0232)	-0.6690*** (0.1680)
anio2015	0.3441*** (0.0672)	-0.0559** (0.0243)	-0.9066*** (0.1755)
anio2016	0.4318*** (0.0679)	-0.0996*** (0.0248)	-0.8534*** (0.1797)
anio2017	0.5818*** (0.0688)	-0.1217*** (0.0257)	-1.1932*** (0.1864)
anio2018	0.6927*** (0.0689)	-0.1224*** (0.0266)	-1.7502*** (0.1911)
anio2019	0.8565*** (0.0703)	-0.1597*** (0.0280)	-2.2152*** (0.2016)
_cons	-14.3182*** (0.4902)	12.5780*** (0.0860)	-31.6328*** (1.9717)

Cuadro 14: Resultados de la estimación por GMM-3SLS

Notas: Errores estándar entre paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Número de observaciones = 26916

Fuente: Elaboración propia

Como podemos observar, los resultados obtenidos de asumir que el sistema es homocedástico, aunque sin suponer que la matriz de varianzas y covarianzas de los coeficientes sea diagonal, son muy similares a la estimación GMM-OWM para este caso.