

Estimación del impacto del uso de las TIC y de la inversión en I&D sobre la productividad de las empresas manufactureras en Perú

Juan Manuel García Carpio

Departamento de Países del
Grupo Andino (CAN)

NOTA TÉCNICA N°
IDB-TN-2590

Estimación del impacto del uso de las TIC y de la inversión en I&D sobre la productividad de las empresas manufactureras en Perú

Juan Manuel García Carpio

Diciembre 2022

Catalogación en la fuente proporcionada por la
Biblioteca Felipe Herrera del
Banco Interamericano de Desarrollo

García Carpio, Juan Manuel.

Estimación del impacto del uso de las TIC y de la inversión en I&D sobre la
productividad de las empresas manufactureras en Perú / Juan Manuel García Carpio.
p. cm. — (Nota técnica del BID ; 2590)

Incluye referencias bibliográficas.

1. Research and development projects-Economic aspects-Peru. 2. Technological
innovations-Economic aspects-Peru. 3. Industrial productivity-Effect of technological
innovations on-Peru. 4. Industrial productivity-Information technology-Peru. I. Banco
Interamericano de Desarrollo. Departamento de Países del Grupo Andino. III. Título.
IV. Serie.

IDB-TN-2590

<http://www.iadb.org>

Copyright © 2022 Banco Interamericano de Desarrollo. Esta obra se encuentra sujeta a una licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas (CC-IGO 3.0 BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando el reconocimiento respectivo al BID. No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI (UNCITRAL). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID, no están autorizados por esta licencia CC-IGO y requieren de un acuerdo de licencia adicional.

Note que el enlace URL incluye términos y condiciones adicionales de esta licencia.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.



Estimación del impacto del uso de las TIC y de la inversión en I&D sobre la productividad de las empresas manufactureras en Perú¹

Resumen

El trabajo estima el efecto del uso de las tecnologías de la información y comunicación (TIC) y de la inversión en Investigación y Desarrollo (I&D) sobre la productividad total de factores (PTF) de las empresas de mayor tamaño en la industria manufacturera peruana entre 2016 y 2018, usando datos de panel de la Encuesta Económica Anual (EEA).

Primero, se revisa aspectos conceptuales y metodológicos de la estimación de la PTF a nivel de empresas enfatizando los métodos econométricos modernos que enfrentan los problemas de endogeneidad que afectaban a los métodos tradicionales, y se analiza estimaciones previas de la PTF a nivel empresarial para Perú. También, se revisa la literatura, sobre todo empírica, relativa al impacto del uso de TIC sobre la productividad empresarial, y al efecto de la innovación mediante inversión en I&D, incluyendo la evidencia disponible para Perú. Luego, se estima el logaritmo de la PTF mediante distintos métodos, seleccionándose el método de Levinsohn y Petrin con corrección por *attrition* (salida de empresas del panel). Se explora la relación entre esta PTF estimada y características de las empresas y el uso de las TIC.

Posteriormente, se estima los efectos de la intensidad de la inversión en I&D y del uso de las TIC sobre la PTF utilizando el planteamiento del modelo estándar CDM. Para ello, se estima los determinantes de la inversión en I&D y de su nivel mediante un modelo de corrección del sesgo de selección (Heckman), y se predice la intensidad de esta inversión (valor entre monto de salarios). Respecto al uso de las TIC, se construye índices sintéticos mediante análisis factorial (por la correlación entre las variables disponibles). Además, la estimación incluye *mark-up* estimados para cada empresa.

Los resultados indican que la inversión en I&D tendría un efecto positivo estadísticamente significativo sobre la PTF de las empresas manufactureras. En contraste, no se encuentra un efecto claro del uso de las TIC, aunque sí se halla efectos positivos estadísticamente significativos para variables específicas como el uso de extranet y, en menor medida, el uso de red local. Este tipo de hallazgos mixtos para las TIC podría deberse a que la muestra utilizada proviene de empresas de mayor tamaño. Por ello, resulta importante realizar estudios específicos del uso de distintas TIC enfocados en diferentes segmentos de empresas, a fin de generar evidencia rigurosa de su impacto según características heterogéneas, como el tamaño empresarial.

Clasificación JEL: D24, L2, L6, O3.

Palabras clave: Productividad, productividad total de factores, investigación y desarrollo, tecnologías de información y comunicaciones, manufactura, industria manufacturera, empresas.

¹ El autor agradece la colaboración de Alex Huanca en el desarrollo del análisis cuantitativo del estudio. Asimismo, los valiosos comentarios de Gustavo Crespi, Fernando Vargas, Juan Gallego, Javier Beverinotti, y Kenji Moreno. Cualquier error u omisión es responsabilidad exclusiva del autor.

ÍNDICE

1	Medición de la Productividad a nivel de empresas	3
1.1	Productividad Total de Factores	3
1.2	Métodos de estimación de la PTF	4
1.3	Estudios para Perú	8
2	Impacto del uso de las TIC y de la investigación y desarrollo sobre la productividad ..	11
2.1	Efectos del uso de las TIC sobre la productividad	11
2.2	Uso de las TIC, innovación y productividad	14
2.3	Adopción y uso de tecnologías digitales en Perú	21
2.4	Evidencia del impacto de la innovación en la productividad para Perú	22
3	Metodología de análisis empírico	24
3.1	Pasos de la metodología de estimación	24
3.2	Estimación de la función de producción	24
3.2.1	Limitaciones de los métodos econométricos utilizados	25
3.2.2	Inclusión de capital TIC en la estimación	29
3.3	Relación entre el uso de las TIC y la Productividad	29
4	Información utilizada y estimación de variables	29
4.1	Datos utilizados para estimar la productividad	29
4.2	Estimación de la productividad	31
4.3	Medición del uso de las TIC y de la inversión en I&D	34
4.3.1	Medición del capital TIC	34
4.3.2	Variables sobre uso de las TIC e inversión en I&D	36
4.4	Índice sintético de uso de las TIC	38
4.5	Monto estimado de la inversión en I&D	45
5	Análisis de los determinantes de la PTF	44
5.1	Distribución de la PTF de las empresas de manufactura según variables	45
5.2	Estimación del efecto del uso de las TIC y la inversión en I&D sobre la PTF	52
5.3	Principales resultados	55
6	Conclusiones y recomendaciones	58
	Referencias	61
	ANEXOS	72
	Anexo 1: Métodos de estimación de la Productividad Total de Factores (PTF)	72
	Anexo 2: Características de la Encuesta Económica Anual (EEA)	81
	Anexo 3: Estadística descriptiva - Manufactura Panel 2016-2018	85
	Anexo 4: Modelos Censurados, Truncados, y Sesgo de Selección	86
	Anexo 5: Pruebas de diferencia de medias de la PTF - Manufactura Panel 2016-18	93
	Anexo 6: Estadística descriptiva del uso de las TIC y de inversión en I&D según tamaño de empresa y subsector productivo - Manufactura panel 2016-18	95
	Anexo 7: Pruebas estadísticas para modelos de determinantes de la PTF a nivel de empresas - Manufactura Panel 2016-18	97
	Anexo 8: Determinantes de la PTF a nivel de empresas según tamaño y subsector - Manufactura Panel 2016-18	98

1 Medición de la Productividad a nivel de empresas

1.1 Productividad Total de Factores

En esta sección se realiza la definición del concepto de productividad y cómo se puede medir a nivel de las empresas o unidades productivas de una economía. Para ello, en primer lugar, es necesario indicar que la producción de bienes y servicios es el resultado de la transformación de factores productivos, ya sea insumos, trabajo, capital, entre otros, en productos mediante el uso de una tecnología. Mientras que la productividad es una medida de qué tan eficiente es el uso o combinación de esos factores de producción para generar las unidades de producción (Farrell, 1957).

Por ello, usualmente, la productividad es definida como la ratio entre la medición del volumen de producción y la medición del volumen de factores o insumos utilizados en el proceso productivo (OECD, 2001; Färe *et al.*, 2008). En este contexto, surge el concepto de Productividad Total de factores (PTF) como una medida de la eficiencia global en el uso de los distintos factores productivos. De este modo, la productividad estimada a nivel de empresas la PTF resume las externalidades positivas que contribuyen al incremento de la producción de las empresas (Abramovitz, 1956; Jorgenson y Griliches, 1967)².

Para medir la PTF se han desarrollado distintas enfoques y métodos³ que modelan el proceso productivo mediante una “función de producción” que relaciona los factores de producción o insumos (ambos bajo la denominación de *inputs*)⁴ con el resultado de producción (*outputs*) obtenido mediante la utilización de la tecnología disponible.

En la literatura, para medir la PTF han destacado los métodos paramétricos que asumen una función de producción “estocástica” con una forma funcional determinada (Nadiri, 1970). A partir de distintas técnicas econométricas, se estima la PTF como un residuo no observable resultante entre el nivel de producción obtenido y el nivel predicho dado el uso de los factores de producción a partir de los parámetros obtenidos en la estimación, el cual ha sido denominado “residuo de Solow” (Solow, 1957).

Además, mediante esta estimación es posible determinar la participación de los diversos factores en la producción (específicamente cuando se utiliza una función de producción de tipo Cobb-Douglas), la contribución al crecimiento de la producción, el grado de complementariedad o sustituibilidad entre factores productivos, el rol del cambio

² En general, la PTF puede ser estimada cuando existen múltiples productos y no solo uno.

³ De acuerdo con Tello (2012), existen cuatro métodos para estimar el nivel y variaciones de la PTF: el método contable, el método de los números índices, los métodos no paramétricos y los métodos paramétricos. En cada uno de estos métodos se puede usar, como datos para la estimación, ‘cantidades’ del proceso productivo consistentes con una función de producción o precios consistentes con una función de costos. Al primer caso se le denomina la técnica “primal” y al segundo la “dual”.

Las metodologías econométricas utilizadas para medir la PTF se han basado sobre todo en el enfoque primal.

⁴ En adelante, no se va a diferenciar factores de insumos utilizados en el proceso productivo. En la literatura económica, sin embargo, se suele realizar esa distinción, pues los factores de producción (principalmente capital físico y mano de obra) transforman los insumos o materias primas en un bien o servicio final, a través del proceso productivo (dejan de existir en sí mismos, y se “insumen” en el producto final).

tecnológico, la existencia de economías de escala, así como la evaluación del impacto de las políticas públicas en el desempeño de las empresas (Färe *et al.*, 2008 ;Del Pozo y Guzmán, 2019).

En general, siguiendo a Griliches y Mairesse (1995), conceptualmente la tecnología de producción puede ser representada a través de una función de producción estocástica:

$$Y_{i,t} = f(K_{i,t}, L_{i,t}, M_{i,t}, V_{i,t})$$

Donde $Y_{i,t}$ es el nivel de producción de la empresa "i" en el periodo "t", el cual depende del uso de insumos $M_{i,t}$, el trabajo $L_{i,t}$, el capital $K_{i,t}$, y el shock de productividad $V_{i,t}$.

Si se asume una forma funcional tipo Cobb-Douglas para f, y se toma logaritmos, es posible plantear la ecuación anterior como:

$$y_{i,t} = \beta_0 + \beta_l l_{i,t} + \beta_k k_{i,t} + \beta_m m_{i,t} + v_{i,t}$$

Donde las variables en minúsculas denotan valores en logaritmos, y β_l , β_k y β_m los parámetros correspondientes al trabajo, el capital y los insumos, respectivamente. β_0 representa la eficiencia promedio de las empresas.

El residuo $v_{i,t} = \omega_{i,t} + \varepsilon_{i,t}$ representa al logaritmo de la PTF específica de cada empresa, y está compuesto por dos variables que son no observables ($\omega_{i,t}$, $\varepsilon_{i,t}$) para el investigador.

$\omega_{i,t}$ representa shocks de productividad que son potencialmente observables o predecibles en cada momento t para el conductor de la empresa i cuando toma la decisión del nivel de producción y el uso de los factores de producción. Intuitivamente, $\omega_{i,t}$ puede resumir la habilidad o experiencia de los gerentes o conductores de las empresas y/o diversos shocks esperados, y el nivel tecnológico de la empresa.

Además, $\varepsilon_{i,t}$ son shocks de producción o productividad aleatorios que no son observables o predecibles para las empresas antes que tomen sus decisiones de uso de factores de producción en el periodo "t".

1.2 Métodos de estimación de la PTF

Las metodologías econométricas de estimación de la función de producción han variado a través del tiempo. Inicialmente, el análisis econométrico se ha enfocado en estimar los parámetros estructurales de la función de producción empleando el estimador OLS, empleando la función de producción Cobb-Douglas, y se estiman los parámetros que aproximan la participación del capital y trabajo a través de Mínimos Cuadrados Ordinarios (OLS, por sus siglas en inglés), y se calcula la PTF como el residuo resultante entre el producto observado y la predicción dada la cantidad de factores productivos (enfoque primal) para la forma funcional especificada.

Sin embargo, existen diversos desafíos empíricos importantes para estimar la función de producción. Siguiendo a Aguirregabiria (2018), entre ellos tenemos:

- Errores de medición tanto en los indicadores de la producción (ventas o ingresos) como en los indicadores de factores de producción (valor económico del capital, calidad del factor trabajo, etc.).
- Problemas de especificación en la forma funcional a través de la cual los factores de producción se interrelacionan.
- Endogeneidad por simultaneidad, lo cual implica que determinados factores de producción (trabajo y capital) pueden estar correlacionados con factores de producción y shocks de productividad inobservables (habilidad administrativa, calidad de los factores, etc.). En general, este problema de endogeneidad es causado por un "sesgo de transmisión" de la productividad de una empresa o planta a la elección óptima de insumos o factores de producción de la misma (Gandhi, 2011).
- Multicolinealidad, lo cual implica que los valores de determinados factores de producción, típicamente trabajo y capital, están correlacionados entre sí.
- Endogeneidad por sesgo de selección, en paneles de datos la entrada y la salida de empresas de la muestra no es exógena y está correlacionada con el tamaño, donde empresas menos productivas o de menor tamaño tienen una mayor probabilidad de salir del mercado que empresas más productivas o más grandes⁵.

Dado que problema de endogeneidad genera que las estimaciones por OLS sean inconsistentes, posteriormente se han desarrollado metodologías que tratan de corregir la endogeneidad por la presencia de variables omitidas (Griliches y Mairesse, 1995). Entre las más tradicionales están el método de variables instrumentales (Eslava *et al.*, 2004), el método Generalizado de Momentos (Blundell y Bond, 2000). Sin embargo, esta metodología presenta dificultades para encontrar los instrumentos adecuados (Ackerberg *et al.*, 2007). Si se utilizan "instrumentos débiles", es decir, variables que no están suficientemente correlacionadas con la variable endógena, no se logra corregir la endogeneidad, haciendo que los estimadores del modelo lineal corregido sean también inconsistentes (Staiger y Stock, 1997).

Asimismo, Olley y Pakes (1996) han indicado que existe un problema de sesgo de selección o endogeneidad debido a la entrada y salida de empresas (*endogeneity of attrition*) que implica la correlación entre el capital y el término estocástico no observable de la función de producción ($\varepsilon_{i,t}$). Dicho problema implicaría una subestimación de la participación del capital en el producto.

En las estimaciones de las funciones de producción se han usado paneles con efectos fijos desde trabajos como Isgut *et al.* (1999)⁶, donde se asume que los shocks de

⁵ Este problema fue señalado por primera vez Marschak y Andrews (1944), la usual entrada y salida de las empresas que maximizan las ganancias para estimar las funciones de producción da lugar a un problema de endogeneidad.

⁶ Siendo los antecedentes los estudios de Mundlak (1961) y Mundlak y Hoch (1965).

productividad son invariantes en el tiempo. Bajo dicho supuesto, las estimaciones de los parámetros estructurales pueden ser consistentes, ya que el error idiosincrático fijo de cada empresa considera las posibles variables omitidas en un modelo OLS. Sin embargo, como indica Rovigatti (2020), un inconveniente común de los modelos con efectos fijos es que, al transformarse los datos para restar los efectos fijos, puede eliminarse una parte importante de la información que permite la identificación de los parámetros de la regresión. Ello, a menudo, conduciría a estimaciones del coeficiente de capital y rendimientos a escala que son inverosímilmente bajos, así como a grandes errores estándar (Rovigatti, 2020).

Al respecto, Akerberg *et al.*, (2007) menciona que dicho supuesto es poco plausible, toda vez que la productividad es susceptible a variar en el tiempo, además que la inconsistencia de las estimaciones se incrementa si existen errores de medición en las variables de los factores de producción, y en la práctica se obtienen estimaciones muy bajas del coeficiente del capital que generan retornos a escala decrecientes.

Ante ello, se han propuesto métodos más modernos basados en expresar la productividad no observable como una función de variables observables, mediante el uso de una “función de control” como los métodos propuestos por Olley y Pakes (1996) y el por Levinsohn y Petrin (2003), los cuales utilizan como variables a la inversión y a los insumos, respetivamente, como se explicará posteriormente.

Olley y Pakes (1996) han desarrollado un método semi-paramétrico – denominado método OP- que plantea una relación determinística entre la productividad y la inversión (y la probabilidad de entrada y salida de empresas), asumiendo que dicha relación es la misma para todas las empresas; es decir, que todas las empresas enfrentarían los mismos precios y condiciones de mercado. El método OP impone el supuesto de monotonicidad en la relación entre la inversión y productividad, condición que asegura la invertibilidad de la función de demanda de la inversión.

Sin embargo, de acuerdo con Akerberg, Craves y Fraser (2015), el problema con este método radica en que para muchas empresas la inversión está truncada hacia cero, es decir que la inversión tiene en mismo valor para distintos valores de la productividad, con ello no será posible invertir la función de demanda de la inversión; una posibilidad es excluir todas las empresas para las cuales la inversión sea nula, no obstante, el riesgo potencial es una significativa pérdida de observaciones, lo cual afectaría a las propiedades asintóticas del estimador.

Por su parte, Levinsohn y Petrin (2003) proponen emplear a los insumos en lugar de la inversión en la función de control. Este método, denominado LP, asume que los insumos dependen de la productividad y del capital, y se asume monotonicidad en la relación entre insumos y productividad. Este método estima los parámetros de interés a través del estimador Método Generalizado de Momentos (GMM, por sus siglas en inglés), las condiciones de momentos se plantean en términos de los errores idiosincráticos de la función de producción y el proceso estocástico de la productividad ortogonal a los instrumentos (capital e insumos rezagados un periodo).

Recientemente Akerberg, Craves y Fraser (2007, 2015) han precisado que ambos métodos OL y LP no abordan el potencial problema de la dependencia condicional que implica la existencia de colinealidad entre los factores de producción, puesto que los factores, principalmente, trabajo e insumos intermedios están determinados por las mismas variables de estado (capital y productividad), y la decisión sobre el trabajo y los insumos se toma simultáneamente por parte de las empresas. En la medida que el trabajo sea una variable dinámica o esté asociada a la inversión o a los insumos, el parámetro asociado al factor trabajo no podría ser identificado plausiblemente. Por ello, los métodos OP como LP ofrecerían estimaciones inconsistentes de la participación del trabajo, generando ciertas imprecisiones en la estimación de la PTF a nivel de empresas.

A fin de abordar adecuadamente el problema de dependencia condicional, estos autores proponen el método ACF que asume que la productividad también depende del factor trabajo y que, a través de las condiciones de momentos de los errores ortogonales con el capital y el trabajo, es posible identificar adecuadamente los parámetros de ambos factores. Así, el método AFC podría ofrecer estimaciones consistentes para la función de producción y permitiría un cálculo más preciso de la PTF ante problemas de endogeneidad y dependencia condicional (Del Pozo y Guzmán, 2019).

Por otro lado, Wooldridge (2009) también aborda estos problemas de la estimación mediante los métodos OL y LP proponiendo una mejora a partir de la estimación conjunta de los parámetros del modelo mediante una aproximación polinómica utilizando el método de variables instrumentales. Sin embargo, cabe destacar que, aunque esta estimación simultánea es más eficiente que la estimación en dos etapas, ambas metodologías son consistentes, y que usar este método implica una pérdida importante de grados de libertad en la estimación.

Por último, Mollisi y Rovigatti (2018) han propuesto una modificación al estimador de Wooldridge, basada en la utilización instrumentos dinámico de tipo panel, la cual se desempeñaría mejor que este método para datos simulados con un pequeño número de períodos, permitirá aumentar el tamaño de la muestra en modelos sobreidentificados, y produciría resultados más estables.

El detalle a nivel técnico de los métodos modernos de estimación de la PTF mencionados previamente se indica en el Anexo 1.

En síntesis, siguiendo a Van Beveren (2012), los métodos modernos de estimación de la PTF han buscado enfrentar tres problemas empíricos relevantes en la estimación de la función de producción.

- El sesgo de selección que implica la correlación entre el capital y el término estocástico no observable de la función de producción ($\varepsilon_{i,t}$) dicho problema implicaría una subestimación de la participación del capital en el producto, dicho problema ha sido abordado en la literatura por Olley y Pakes (1996).
- El problema de simultaneidad implica la correlación en los factores de producción y el término de error por la presencia de variables omitidas, en particular la

productividad inobservable, dicho problema empírico ha sido abordado en la literatura por Olley y Pakes (1996) y Levinsohn y Petrin (2003).

- El problema de dependencia condicional implica la presencia de colinealidad entre los factores de producción. De similar modo que el problema de simultaneidad, la dependencia condicional generaría la subestimación de la participación del capital en el producto y la sobreestimación de la participación del trabajo. Este problema ha sido abordado más recientemente por Akerberg, Caves y Fraser (2015).

Cuadro 1: Problemas econométricos en la estimación de la función de producción y métodos asociados

Problemas empíricos	Definición	Dirección del sesgo en la participación de los factores de producción	Métodos
Sesgo de selección (por <i>attrition</i>)	Correlación entre el capital y el término estocástico no observable de la función de producción	Sesgo hacia abajo en la participación del capital (<i>downward bias</i> : $\hat{\beta}_{k,OLS} < \beta_k$)	Olley y Pakes (1996)
Endogeneidad por simultaneidad	Correlación entre los factores de producción y el término estocástico en la función de producción	Sesgo hacia abajo en la participación del capital (<i>downward bias</i> : $\hat{\beta}_{k,OLS} < \beta_k$)	Olley y Pakes (1996) Levinsohn y Petrin (2003) Akerberg, Caves y Fraser (2015)
		Sesgo hacia arriba en la participación del trabajo (<i>upward bias</i> : $\hat{\beta}_{l,OLS} > \beta_l$)	Wooldridge (2009) Mollisi y Rovigatti (2018)
Dependencia condicional	Colinealidad entre los factores de producción	Sesgo hacia abajo en la participación del capital (<i>downward bias</i> : $\hat{\beta}_{k,OLS} < \beta_k$)	Akerberg, Caves y Fraser (2015)
		Sesgo hacia arriba en la participación del trabajo (<i>upward bias</i> : $\hat{\beta}_{l,OLS} > \beta_l$)	Wooldridge (2009) Mollisi y Rovigatti (2018)

Fuente: Adaptado del Del Pozo y Guzmán (2019) en base a Van Beveren (2012).

1.3 Estudios para Perú

Para el caso peruano, diversos estudios han implementado estimaciones de la función de producción y de la PTF para las empresas formales del sector industria

manufacturera⁷ incluyendo las mejoras metodológicas recientes para dar cuenta del problema de endogeneidad en el cálculo de la PTF. Los estudios más relevantes en cuanto a rigurosidad metodológica son los de Tello (2012, 2014), Céspedes *et al.*, (2014), y Del Pozo y Guzmán (2019).

Tello (2012) estima la función para el valor de producción de empresas manufactureras en el periodo 2002-2007 a partir de paneles de datos de la Encuesta Económica Anual (EEA) del Instituto Nacional de Estadística e Informática (INEI), empleando una forma funcional Cobb-Douglas y cinco metodologías alternativas: OLS, paneles de datos con efectos fijos, paneles de datos con efectos aleatorios, OL y LP. Además del trabajo y el capital, se incluye variables de control en la estimación para tamaño de empresa y subsector productivo, se considera dos definiciones de capital (k1: maquinaria y equipo, equipos diversos y de transporte y k2: solo maquinaria y equipo), y se utiliza distintos paneles de datos (panel balanceado 3 años, panel no balanceado 3 y 2 años, panel balanceado 4 años, y panel no balanceado 4, 3 y 2 años). Se encuentra que el método OLS sobreestima la participación de los factores de trabajo y capital, y estima el coeficiente asociado al trabajo entre 0.42 y 0.47 y entre 0.44 y 0.49 con los métodos OL y LP, respectivamente; mientras el capital tendría una participación entre 0,17 y 0.28 con el método OL y entre 0.21 y 0.23 con el método LP.

Tello (2014) realiza estimaciones similares para el período 2000-2010 (con paneles para 2000-2004, 2004-2007, 2007-2009, 2009-2010), encontrando resultados semejantes, con coeficientes ligeramente superiores, para el trabajo de 0.52 y 0.65, usando los métodos OP y LP, respectivamente, y para el capital de 0.24 y 0.33, respectivamente.

Ambos estudios de Tello (2012, 2014) destacan que la PTF habría tenido un bajo crecimiento en los períodos estudiados, y que su aporte habría sido poco significativo para el crecimiento del valor de producción del sector manufactura.

Por su parte, Céspedes *et al.*, (2014) realizan estimaciones para distintos sectores para el periodo 2002-2011 utilizando información de tipo panel de las declaraciones juradas de los estados financieros de empresas registradas por la Superintendencia de Administración Tributaria (SUNAT). Utilizan 8,996 empresas de diversos sectores económicos (Agricultura, Comercio, Construcción, Electricidad, Industria, Intermediación Financiera, Minería, Servicios y Pesca). Asumen una función de producción Cobb-Douglas, y utilizan los métodos OLS, Efectos Fijos, y Arellano-Bond y OP⁸. Para el sector manufactura, los autores encuentran coeficientes de 0.39 y 0.40 para el trabajo usando el método de Arellano-Bond y el método OP, respectivamente.

⁷ El sector manufactura representa cerca al 14% del Producto Bruto Interno (PBI) del Perú y, a fines de la década pasada, incluía casi 200 mil empresas de un total de 2.5 millones de empresas de acuerdo al Directorio Central de Empresas y Establecimientos del Instituto Nacional de Estadística e Informática (INEI).

⁸ En el caso del método de Arellano-Bond, realizan estimaciones en primeras diferencias utilizando los factores de capital y trabajo rezagados (de 1 a 9 años de rezago) como variables instrumentales para los factores capital y trabajo. Según los autores, este método sería consistente incluso en el caso de que la variable PTF de las empresas tenga un componente que varía en el tiempo y que está correlacionado de forma contemporánea con los insumos.

Asimismo, obtienen 0.70 y 0.83 para el capital, de acuerdo con los métodos mencionados.

Aunque estos estudios dan cuenta del problema de endogeneidad a través de los métodos de OL y LP, no abordaban el problema de dependencia condicional entre los factores de producción, lo cual sí fue analizado en el estudio de Del Pozo y Guzmán (2019) mediante el método ACF, debido a que el coeficiente asociado al factor trabajo podría estar incorrectamente identificado. Al estimar la función de producción para el valor agregado de las empresas del sector manufactura para un panel de empresas 2014 a 2015 de la EEA, estos autores encuentran que el coeficiente asociado al trabajo – medido como remuneraciones - sería menor si se hace la estimación utilizando el método ACF respecto a los otros métodos, especialmente con relación a la estimación mediante OLS y panel efectos fijos. Específicamente, el coeficiente del trabajo es 0.49, 0.49, y 0.46 utilizando los métodos OP, LP sin corrección de sesgo de selección, y ACF, respectivamente. Asimismo, en el caso del capital el coeficiente estimado es 0.47, 0.47 y 0.43 al utilizar los tres métodos mencionados.

Por último, PRODUCE (2021) también realiza la estimación de la PTF para el sector manufactura para los años 2015 a 2018 utilizando los datos de la Encuesta Nacional de Empresas (ENE), la cual incluye unidades productivas de menor tamaño (desde 20 UIT de ventas mensuales). En este caso, se utiliza solo el método ACF y se distingue entre trabajo calificado y no calificado, encontrándose un coeficiente para el capital de 0,262 (similar a los estudios de Tello), y una participación muy importante del trabajo calificado.

**Cuadro 2: Estimaciones previas de la función de producción
Sector Industria manufacturera**

Autor / Método	OLS	FE	AB	OP-B	OP	LP	ACF
Tello (2012)							
Trabajo	0,433 - 0,526	0,093 - 0,244			0,422 - 0,466	0,440 - 0,490	
Capital	0,213 - 0,295	0,053 - 0,0963			0,168 - 0,276	0,206 - 0,232	
Céspedes <i>et al.</i> (2013)							
Trabajo	0,353	0,288	0,392		0,400		
Capital	0,650	0,490	0,703		0,830		
Tello (2014)							
Trabajo	0,640	0,24 a 0,44			0,520	0,650	
Capital	0,250	0,25 a 0,12			0,330	0,240	
Del Pozo y Guzmán (2019)							
Trabajo	0,524	0,527		0,495		0,491	0,464
Capital	0,499	0,366		0,470		0,468	0,427
PRODUCE (2021)							
Trabajo calificado							0,589
Trabajo no calificado							0,165
Capital							0,262

Nota: OP-B: Método de Olley-Pakes sin corrección de sesgo de selección (sin estimar probabilidad de que siga activa la empresa).

Fuente: Elaboración propia.

2 Impacto del uso de las TIC y de la investigación y desarrollo sobre la productividad

2.1 Efectos del uso de las TIC sobre la productividad

Las tecnologías de la información y comunicación (TIC) se conciben como el conjunto de servicios, redes, software y dispositivos de hardware que se integran en sistemas de información interconectados y complementarios con la finalidad de gestionar datos, información y procesos de manera efectiva (Concytec, 2016). Actualmente, las TIC se basan en medios digitales, en contraposición a medios analógicos, físicos, o de otra naturaleza. Por ello, actualmente la adopción y uso creciente de las TIC por las empresas forma parte del proceso de adopción digital o “digitalización” empresarial concebido como el uso de tecnologías digitales que permitan generar, procesar y compartir información (Katz y Callorda, 2015).⁹

En general, según Katz y Callorda (2015) la digitalización contribuye a incrementar la producción y el empleo al permitir mejoras de los procesos productivos de las empresas y ampliar sus modelos de negocio para expandirse a nuevos mercados. Asimismo, facilita el desarrollo de marcas y productos mediante el uso de redes sociales permitiendo adaptar mejor los productos a los requerimientos del mercado.

A partir de distintos estudios basados en datos de empresas, se han encontrado evidencias significativas de la relación positiva entre la inversión en las tecnologías de la información (IT) primero, y luego, de forma más general, de la inversión en las TIC, y el crecimiento de la productividad identificando el rol complementario que juegan los cambios organizacionales y según de midan impactos de corto o de largo plazo. En general, los retornos en el corto plazo reflejan los efectos directos de la inversión en TIC, mientras que en el largo plazo se capturan los efectos de las TIC cuando se combinan con el cambio organizacional (Brynjolfsson y Hitt, 1998; Hitt y Brynjolfsson, 2000).

La utilización de las TIC por parte de las empresas promueve cambios en sus procesos que incrementan la productividad y reducen los costos. La evidencia empírica disponible (Billón *et al.*, 2007) indicaba que sus impactos están asociados (sobre todo en el largo plazo) a cambios organizacionales a través de mejoras de coordinación en los flujos de información entre los trabajadores y la dirección (y la creación de equipos de trabajo), la descentralización en la toma de decisiones, la reducción de las jerarquías, promoviendo cambios en la composición del trabajo a favor de los trabajadores más cualificados (*upskilling*) o promoviendo las inversiones en capacitación (Bresnahan y otros, 2002; Brynjolfsson y Hitt 2003; Kohli y Devaraj, 2003; Bertschek y Kaiser, 2004; Bayo-Morions *et al.*, 2008).

⁹ Recientemente, entre las actividades realizada utilizando las TIC de tipo digital, se ha destacado el comercio electrónico (e-commerce) - intercambio de dinero por bienes o servicios entre empresas y consumidores mediante plataformas virtuales, debido a su importante rol en la realización de negocios y en la entrada a nuevos mercados (Kumar & Raheja, 2012).

A su vez, aquellas empresas en las que se han introducido cambios organizativos previos es más probable que registren mayores inversiones en TIC y que ello a su vez afecte a los retornos obtenidos (Doms *et al.*, 1997; Brynjolfsson y Hitt, 2000; Black y Lynch, 2001)¹⁰. Se reconoce la relación complementaria entre la adopción y el uso de las TIC, factores específicos de la empresa y estructura industrial (Bocquet *et al.*, 2007; Fabiani *et al.*, 2005; Giuri *et al.*, 2008. Así, de acuerdo con Hollenstein (2004), la adopción de nuevas prácticas de trabajo favorece la adopción de las TIC, pero, también existiría la relación inversa, lo cual indica que la adopción de las TIC y el cambio organizacional son, en cierta medida, complementos.

Por otro lado, la evidencia existente ha mostrado la existencia de retornos de las inversiones en TIC superiores a otros tipos de inversión al estimarse funciones de producción que distinguen entre capital TIC y capital no TIC, aunque estos resultados se refieren sobre todo a Estados Unidos, en trabajos tales como Brynjolfsson (1993), Brynjolfsson y Hitt (1995), Lichtenberg (1995), Hitt y Brynjolfsson (1996), Menon *et al.* (2000), Brynjolfsson y Hitt (2003) y Bloom *et al.* (2012). En general, de acuerdo con Bloom *et al.* (2012), al contrario de la "Paradoja de la productividad de Solow"¹¹, los retornos al capital TI hallados en la literatura son mayores de lo que uno esperaría bajo los supuestos estándar de contabilidad de crecimiento.

Por ejemplo, Hitt y Brynjolfsson (1996) obtienen que el producto marginal bruto del capital TI sobre el valor agregado es de un 81% por año en el período 1987-1991, siendo al menos tan grande como el producto marginal de otros tipos de inversión en capital. Además, la contribución del trabajo relacionado a las TI al crecimiento de la producción es tan elevada como la del trabajo no TI. Según Brynjolfsson y Hitt (2003) esto se debería a las inversiones complementarias en "capital organizacional" cuyo efecto se refleja en los coeficientes del capital TI.

A partir de la década de 2000, muchos estudios incorporan variables organizacionales en los modelos, aunque resulta difícil medir los cambios organizativos. Entre las alternativas de estimación destaca el uso de variables ficticias que muestren la presencia o ausencia de una determinada característica organizacional en la empresa (Arvanitis, 2005; Matteucci *et al.*, 2005), como el uso intensivo de equipos de trabajo, la rotación en el empleo y la reducción de niveles jerárquicos, entre otras. Aunque esta opción tiene la limitación de no poder matizar las diferencias de calidad de dichos cambios organizacionales.

Diversos estudios analizan la correlación entre las inversiones en TIC, los cambios en la organización, especialmente en la organización del trabajo y la cualificación de los

¹⁰ Milgrom y Roberts (1990) fueron pioneros en el estudio sobre el papel complementario de las características operativas y organizativas específicas de la empresa como determinantes de la adopción de nuevas TI a nivel de empresas.

¹¹ Referida al hecho de que, a pesar de la creciente incorporación e inversión en TIC en los procesos productivos, en las décadas de 1970 y 1980 en E.E.U.U., no se registró el cambio positivo importante y generalizado esperado en la productividad de las empresas (Brynjolfsson, 1993). La evidencia empírica registrada en muchos países de la OCDE desde los años 90 permite afirmar que la paradoja de la productividad a nivel de empresa no parece ser tal (Billón *et al.*, 2007).

trabajadores, y el incremento de la productividad. Francalanci y Galal (1998) en un estudio aplicado a las compañías de seguros demuestran que la introducción de las TIC viene acompañada de ganancias en productividad en empresas en las que aumenta el peso del componente directivo y se modifica la composición del trabajo a favor de una mano de obra más cualificada y de mayor costo.

Bresnahan *et al.* (2002) hallan que aquellas empresas que combinan un creciente uso de las TIC con cambios en las prácticas organizacionales y la introducción de nuevos productos y servicios incrementan en mayor medida su demanda de mano de obra cualificada. Estos autores encuentran efectos interactivos para las inversiones en TIC y los cambios organizativos por una parte y para las inversiones en TIC y la cualificación y formación por otra. Sin embargo, les resulta difícil determinar el impacto de cada una de las tres variables de manera independiente.

Black y Lynch (2001, 2004) encuentran una correlación entre la introducción de nuevas prácticas en el lugar de trabajo, la creciente difusión del uso de los ordenadores y el crecimiento de la productividad en el sector manufacturero de Estados Unidos en el período 1993-1996. Una mayor participación de los trabajadores en la toma de decisiones sobre mejoras en los procesos de producción asociadas a las TIC, y una mejor la relación entre aquéllos y los directivos de la empresa, conlleva mejores resultados en términos de productividad del trabajo. Asimismo, cuanto más alto es el nivel educativo de los trabajadores y mayor la proporción de éstos y de directivos que utiliza ordenadores, más probable son resultados mejores en la productividad laboral.

Asimismo, al investigar la influencia de las TIC, las nuevas prácticas en la organización del trabajo y la formación sobre la productividad de la empresa, Arvantis (2005) encuentra que las empresas con una mayor participación de trabajadores cualificados obtienen mayor productividad de las TIC, mientras que las nuevas prácticas organizacionales no parecen contribuir de manera significativa a una mayor productividad.

Crespi *et al.* (2006) examina las relaciones entre el crecimiento de la productividad laboral, la inversión en tecnologías de la información (TI) y el cambio organizativo utilizando datos de empresas del Reino Unido. Al incluir el cambio organizativo, el efecto estimado de un mayor capital TI se reduce considerablemente. La inversión en IT y el cambio organizativo interactúan en su efecto positivo sobre el crecimiento de la productividad. Halla que la desaceleración del crecimiento de la PTF estimada para las empresas que no hacen cambio organizativo y/o que se encuentran en las primeras etapas de la inversión en TI, es compatible con la tendencia a nivel agregado de desaceleración del crecimiento de la PTF en el Reino Unido.

Bloom *et. al* (2012) estima la función de producción diferenciando el capital entre capital proveniente de la inversión en TI y capital no TI, e interactúa ambos con el tipo de propiedad de las empresas a fin de comparar el desempeño de las empresas multinacionales de US y las de UK respecto al producto por trabajador. Halla que las empresas multinacionales estadounidenses obtenían una mayor productividad del uso de las TI que las no estadounidenses, particularmente en los mismos sectores

responsables de la aceleración de la productividad de los Estados Unidos. Esta ventaja se debería principalmente a sus mejores prácticas de “gestión de las personas”. Además, como parte de su análisis de robustez, ante el posible error de medición que subestime el capital TI de las multinacionales estadounidenses, utilizan su lugar, la proporción de trabajadores que utilizan computadoras en las plantas, y obtienen resultados similares.

Asimismo, funciones de producción de valor agregado incluyendo variables de adopción de las TIC, Colombo *et al.* (2013) encuentran efectos positivos diferenciados asociados a la adopción de servicios tecnológicos basados en banda ancha por parte de pequeñas y medianas empresas italianas entre 1998 y 2004. Solo encuentran aumentos considerables de la productividad al considerar diferentes niveles de análisis y posibles efectos complementarios. En particular, para aplicaciones avanzadas que son potencialmente pertinentes a sus operaciones, como las referidas a la cadena de suministro y la gestión de clientes, cuya adopción está asociada con la realización de cambios estratégicos y organizativos importantes en la forma de hacer negocios.

Respecto al rol específico del capital humano en la adopción de tecnologías como las TIC, este se ha puesto de relieve en la línea del trabajo pionero de Nelson y Phelps (1966) quienes, al investigar cómo se difunden las innovaciones, observaron que las capacidades individuales son clave (capacidades dinámicas): los trabajadores mejor educados ayudan a que las empresas adopten tecnologías, porque dichos trabajadores tienen mayores habilidades para comprender y usar nueva información. En particular, las tecnologías informáticas serían mejor utilizadas por los trabajadores más calificados, los cuales, a su vez, van adquiriendo mayor capacidad de aprendizaje y se convierten en más productivos a medida que se incrementa su experiencia en el uso de las TIC (Entorf y Kramarz, 1998).

Por otro lado, también se ha estudiado la relación específica entre la adopción de las TIC y la mayor innovación en las empresas como un mecanismo para lograr el incremento de su productividad. Por ejemplo, Reeson y Rudd (2016) encuentran que la inversión y la innovación en TIC en las empresas australianas está correlacionada con un aumento de las tasas de innovación y, a su vez, el aumento de su inversión en TIC conlleva una mayor probabilidad de reportar aumentos de la productividad. Este efecto fue más claro en la industria manufacturera, de la construcción y del transporte, y está mediado por innovaciones más amplias en la práctica empresarial (en aspectos de gestión y de operaciones). Por último, cabe indicar que existen distintos estudios que analizan los factores que llevan a las empresas a adoptar las TIC. Entre ellos, destaca el estudio de Martin & Nguyen-Thi (2015) que incluye como determinantes, entre otros, el tamaño de la empresa y si la empresa pertenece a un grupo económico.

2.2 Uso de las TIC, innovación y productividad

La adopción de las TIC por parte de las empresas forma parte del proceso de innovación empresarial, definida como una aplicación de ideas innovadoras y tecnologías que tienden al desarrollo con la finalidad de mejorar los bienes y servicios o hacer su producción más eficiente. De esta forma, las empresas que invierten tiempo en la

realización de actividades innovadoras resultan ser tecnológicamente más avanzadas y a su vez más productivas (Griffith *et al.*, 2006; Hall, Mairesse & Mohnen, 2010).

En general, existen muchos estudios empíricos que han generado evidencias sobre la relación entre innovación y productividad a nivel de las empresas. El modelo de capital de conocimiento desarrollado por Griliches (1979) ha sido una piedra angular de la literatura acerca del impacto de la inversión en I&D sobre la productividad. Desde su implementación por Pakes y Griliches (1984) la estrategia empírica subyacente ha cambiado poco a lo largo de los años (Mairesse & Sassenou, 1991; Griliches, 1995; y Griliches, 2000). La forma más utilizada para estimar el modelo de capital-conocimiento agrega el logaritmo del capital-conocimiento $c_{i,t}$ como un insumo adicional a una función de producción de Cobb-Douglas

$$y_{i,t} = \beta_0 + \beta_l l_{i,t} + \beta_k k_{i,t} + \beta_m m_{i,t} + \beta_c c_{i,t} + \omega_{i,t} + \varepsilon_{i,t}$$

donde β_c es la elasticidad de la producción con respecto al capital-conocimiento.

Se supone que el capital-conocimiento se acumula a partir de gasto de inversión en I&D de la empresa i , $r_{i,t}$, y se deprecia de un período a otro, según la siguiente ley del movimiento: $c_{i,t} = (1 - \delta)c_{i,t-1} + r_{i,t}$ que sirve para construir $c_{i,t}$, el stock de capital de conocimiento de empresa i en el período t , a partir de los gastos observados de I&D. En general, surgen problemas de determinar una tasa apropiada de depreciación, y datos disponibles de gastos en I&D para períodos a menudo no es muy largos, haciendo poco confiables las estimaciones. Al respecto, estudios como Hall & Mairesse (1995) y Doraszelski y Jaumandreu (2013) asumen que la tasa de depreciación es de 0.15 por período y estiman el capital inicial a partir de la fecha de inicio de la empresa extrapolando los datos de sus gastos promedio de I&D durante el tiempo observado.

Asimismo, posteriormente se ha desarrollado una metodología econométrica específica de estimación planteada por Crépon, Duguet, y Mairesse (1998), basada en Pakes and Griliches (1984), conocida como *Modelo CDM*. El CDM tiene como propósito estimar de forma robusta la relación entre la generación de conocimiento, innovación y desempeño de las empresas. La premisa del modelo recae en que un cambio en el gasto o las actividades realizadas en innovación no son suficientes para incrementar la productividad, sino que son los outputs de estas dos variables las que generan un verdadero impacto sobre la producción.

Siguiendo a Arbeláez y Parra (2011) sobre la base de Lööf and Heshmati (2006) y Masso and Vahter (2008), el modelo CDM puede ser especificado por el siguiente sistema de ecuaciones:

$$g_i = \begin{cases} 1 & \text{if } g_i^* = \mathbf{x}_i^o \mathbf{B}^0 + \varepsilon_i^o > 0 \\ 0 & \text{if } g_i^* = \mathbf{x}_i^o \mathbf{B}^0 + \varepsilon_i^o \leq 0 \end{cases} \quad (1)$$

$$i_i^* = \mathbf{x}_i^1 \mathbf{B}^1 + \varepsilon_i^1 > 0 \quad (2)$$

$$t_i = \beta^k i_i^* + \mathbf{x}_i^2 \mathbf{B}^2 + \varepsilon_i^2 > 0 \quad (3)$$

$$q_{i,\tau} = \beta^t t_i + \mathbf{x}_i^3 \mathbf{B}^3 + \varepsilon_i^3 > 0 \quad (4)$$

En el sistema anterior, las \mathbf{x} representan matrices de variables explicativas de acuerdo con la literatura y \mathbf{B} son los vectores de los parámetros correspondientes a estimar. Los ε corresponden a los términos de error normalmente distribuidos en cada una de las ecuaciones con media cero y varianza constante. Se supone que ε_i^0 y ε_i^1 se distribuyen normalmente de forma conjunta (lo cual puede ser testeado).

La ecuación (1) describe la decisión de las empresas de invertir en actividades de innovación. La inclusión de esta ecuación tiene por objeto modelar la decisión de las empresas que realizan inversión en innovación, y poder corregir la estimación de la ecuación (2), pues puede ser necesario un nivel mínimo de gasto o existir factores que diferencian a las empresas que invierten y a las que no hacen y que, a su vez, influyen en el monto invertido (sesgo de selección).

g_i es igual a uno cuando la empresa realiza actividades de innovación, es decir, cuando la variable de decisión de innovación latente, g_i^* , es mayor que un umbral, digamos cero. Para estimar la ecuación (1) se suele utilizar un modelo probit.

La ecuación (2) describe la intensidad de las actividades de innovación de la empresa dado por la variable latente. En el análisis empírico esta variable corresponde a la inversión total en actividades de innovación i_i^* o la intensidad de la inversión calculada como su valor dividido entre el número de trabajadores o entre el valor real de las remuneraciones.

Las ecuaciones (1) y (2) se deben estimar de forma conjunta mediante modelos para datos censurados, truncados, y sesgo de selección, excepto en el caso de que efectivamente todas las empresas puedan realizar alguna inversión, aunque sea mínima, y que esta decisión sea independiente del monto invertido. Los modelos más utilizados son el Tobit, y el modelo de Heckman para sesgo de selección (en dos etapas denominado *heckit* o el modelo simultáneo o en una etapa). En el Anexo 4 se explica en detalle este tipo de modelos.

La ecuación (3) corresponde a la ecuación de producción de las innovaciones también conocida en la literatura como la ecuación del conocimiento (Griliches, 1979; Pakes y Griliches, 1984). Esta producción t_i se ha estimado generalmente como el número de patentes aplicadas (Griliches, 1990; Crépon, Duguet y Mairesse, 1998) y como una medida relativa de las ventas innovadoras (van Leeuwen y Klomp, 2006; Lööf y Heshmati, 2006; Jefferson *et al.*, 2006, entre otros). Asimismo, se han utilizado indicadores de innovación de productos y procesos (Masso y Vahter, 2008) mediante un conjunto de variables dicotómicas recogidas mediante encuestas especializadas que indican si la empresa ha obtenido resultados innovadores (generalmente es posible clasificar los resultados innovadores en cinco tipos: adaptación e innovación de bienes y servicios, nuevos procesos de producción y nuevos procedimientos de gestión y comercialización).

La ecuación (4) completa el sistema y corresponde al efecto de la innovación sobre la productividad o la rentabilidad dependiendo de la información disponible. En la literatura, q_i se ha medido generalmente como valor agregado por empleado, o como ventas por empleado (por ejemplo, Griffith, Huergo y Mairesse, 2006; Masso y Vahter 2008; y Lööf y Heshmati 2006). Sin embargo, algunos autores (Duguet, 2006; Jefferson *et al.* 2006; Arbeláez y Parra, 2011; Bravo-Ortega *et al.*, 2014) han utilizado como medida de productividad a la productividad total de los factores (PTF).

En cuanto a la evidencia empírica, diversos estudios han demostrado que existe una relación positiva entre la realización de actividades de innovación, sobre todo la inversión en Investigación y Desarrollo (I&D) y la productividad, lo que impulsa el crecimiento económico de largo plazo y promueve la creación del empleo (Crespi y Tacsir, 2011; Guloglu y Tekin, 2012).

En general, una mayor inversión en I&D incrementa la productividad de las empresas al contribuir a alcanzar mayores niveles de resultados en innovación, al incrementar su capacidad de absorción de conocimiento y desarrollo de la tecnología (Griffith *et al.* 2004). Además, la inversión en I&D contribuye a la mejora en la identificación, asimilación y uso de tecnologías, lo que mejora la calidad de capital humano dentro de las empresas (Cohen y Levinthal 1989; Griffith *et al.* 2004). De acuerdo con los resultados de los primeros estudios realizados en países desarrollados, cuanto mayor sea la inversión en I&D, es más probable que la empresa introduzca innovaciones en procesos o productos, y que incremente su productividad.

Lööf y Heshmati (2002) muestran que el capital – conocimiento, definido como la ratio de las ventas de productos de innovación entre las ventas totales, es un factor significativo que contribuye a la heterogeneidad entre las empresas del sector manufactura en Suecia, reflejada en la productividad laboral, incluso cuando se controla por el capital humano, el tipo de producción, el tamaño de la empresa y la entrada, fusión, cierre parcial o salida de empresas. Además, el capital-conocimiento aumenta con las actividades de innovación, con el conocimiento interno que posee la empresa para innovar y con la cooperación en innovación con las universidades nacionales.

En la misma línea, Griffith *et al.* (2006), basándose en la Encuesta de Innovación Comunitaria (CIS), encuentra una correlación positiva entre la innovación y la productividad de las empresas manufactureras de distintos países europeos.

Fazlıoğlu *et al.* (2018) analizan de forma diferenciada el impacto de distintos tipos de innovación sobre las empresas manufactureras turcas usando datos para el período 2003-2014 de encuestas de innovación. Hallan heterogeneidad de las empresas en términos de propensión a innovar y de beneficios obtenidos de las actividades de innovación, pues, aunque todos los tipos de actividad de innovación tienen efectos positivos en la productividad de las empresas, hay pruebas sólidas de impactos diferenciados por tipo.

Recientemente, Mohnen (2019) realiza una revisión sistemática de la evidencia sobre la relación entre inversión en I&D y productividad para países europeos, y estima que existe una contribución por parte de la inversión I&D al crecimiento de la PTF en el rango

del 10 al 15%. Las elasticidades estimadas son generalmente más altas en los sectores de alta tecnología, es decir, I&D+i, que en los de baja tecnología (por ejemplo, Ortega-Argilés et. al, 2015).

Sin embargo, a partir de los resultados de distintos estudios, se encuentra que el efecto de la innovación sobre la productividad sería menor en los países en desarrollo. Acemoglu *et al.* (2006) sugieren que la baja inversión en I&D, la escasa innovación y el reducido impacto de la innovación sobre la productividad en países en desarrollo se deberían a que los incentivos para invertir en innovación son insuficientes considerando que se encuentran lejos de la frontera tecnológica.

En todo caso, muchas empresas que invierten en I&D hacen transferencia tecnológica (es decir, adquisición de maquinaria y equipo o compras de tecnología no incorporada) y, por ende, realizan pocos cambios significativos en sus productos, principalmente mediante la imitación (*learning by copying*) (Anlló y Suárez, 2009). Incluso, la inversión en I&D puede ser difícil (ya sea por costos financieros o por falta de mano de obra calificada), por lo que su efecto sobre la productividad puede darse después de un horizonte de tiempo más largo de lo esperado (Nolazco, 2020).

Respecto a los países latinoamericanos, diversos estudios que utilizan el marco del modelo CDM para estimar los determinantes de la innovación y su impacto en la productividad laboral, muestran una fuerte evidencia de una relación positiva entre actividades de innovación y productividad usando como variable dependiente la productividad laboral. Así, Crespi y Zuñiga (2010) encuentran bastante heterogeneidad en el valor de las elasticidades estimadas: Argentina (0.24), Chile (0.60), Colombia (1.92), Panamá (1.65) y Uruguay (0.80). Asimismo, Arbeláez y Parra (2011) y Cassoni y Ramada (2010) demuestran que las actividades de innovación (ya sea en productos, procesos, marketing y gestión) tienen un impacto directo y positivo sobre la productividad de las empresas colombianas y uruguayas, respectivamente.

Álvarez *et al.* (2011) encuentran evidencia de que un mayor gasto en I&D aumenta la probabilidad de innovar en procesos (no en productos), y que este tipo de innovación repercute en mayor medida sobre la productividad laboral de las empresas del sector manufacturero chileno. Sin embargo, para el caso de las micro, pequeñas y medianas empresas, Restrepo-Morales *et al.* (2019), al aplicar una encuesta específica sobre el desempeño empresarial, encuentran que la innovación en productos sería más importante que la innovación en procesos para mejorar su desempeño en Colombia. Su desempeño depende sobre todo de los esfuerzos internos de este tipo de empresas y, además, los resultados sugieren que los imitadores obtienen un rendimiento casi tan alto como los innovadores.

Por otro lado, Bravo-Ortega *et al.* (2014) analizan el efecto de la I&D utilizando la PTF estimada mediante el método de Levinson y Petrin (LP) para medir la productividad, y encuentran una elasticidad en el rango de 0.30 a 0.36 para la inversión en I&D para la industria manufacturera chilena.

Finalmente, según el estudio de Crespi, Tacsir, & Vargas (2016) para América Latina y el Caribe (ALC) es más probable que las empresas introduzcan innovación de productos

o procesos si invierten más en innovación. Los resultados de la innovación de las empresas están muy condicionados por la Investigación y Desarrollo (I&D), haciendo que en el caso de América Latina el incremento de 10% en I&D genere en promedio un aumento en la propensión a innovar del 1.7%. Mohan, Strobl, & Watson (2018) encuentran resultados similares para empresas de países del Caribe.

En cuanto al análisis conjunto del efecto del I&D y de la TIC sobre la productividad, cabe mencionar el trabajo de Hall et al (2013) que investiga los efectos de la inversión en I&D y en TIC a nivel de las empresas manufactureras italianas a fin de evaluar su importancia relativa y en qué medida son complementos o sustitutos. Utilizando una versión del modelo CDM ha modificado para incluir la inversión en TIC y la I&D como los dos insumos principales para la innovación y la productividad, encuentran que la I&D y las TIC están fuertemente asociadas con la innovación y la productividad, siendo la I&D más importante para la innovación y la inversión en TIC más importante para la productividad. Para la empresa mediana, las tasas de retorno de ambas inversiones son altas sugiriendo que la inversión es insuficiente en ambas actividades.

Dificultades metodológicas

En cuanto a la estimación del efecto de la I&D sobre la PTF, el modelo CDM permite corregir, a partir de las distintas etapas de la estimación, tanto la endogeneidad de la decisión de realizar actividades de innovación como invertir en I&D, como el sesgo de selección o truncamiento en el nivel observado del gasto en innovación o de inversión en I&D.

Recientemente, se han utilizado otras metodologías como la estimación mediante el método econométrico denominado *endogenous switching regression model*, lo cual permite controlar los problemas de endogeneidad y sesgo de selección de la variable de inversión en I&D (Fazlıoğlu, 2018).

Un problema que se ha discutido en la literatura es la incorporación de cambios de calidad en los deflatores de precios (Hall *et al.*, 2010). Los productos nuevos o mejorados se añaden a los índices de precios con un retraso sustancial, por lo cual la contribución de la I&D al valor de los bienes intensivos en I&D, ya sean productos o insumos, resulta subestimada.¹²

Al respecto, al existir datos del panel, las diferencias de calidad pueden ser capturadas mediante variables dummy de tiempo y de sectores, incluso en ausencia de buenos datos de precios, dejando solo las diferencias de calidad entre empresas. En este caso, las estimaciones de I&D serán sesgadas, pero solo en la medida en que los precios

¹² La diferencia que puede hacer el uso de precios hedónicos en la estimación de las tasas de rendimiento de la I&D se observa en el estudio de Griliches (1994) que regresiona las tasas de crecimiento de la PTF respecto a las intensidades de I&D para todas las industrias, obteniendo una tasa estimada de retorno a la I&D del 35,7% para el período 1973-89. Cuando la industria informática, que es la única que tiene cambios de calidad incorporados en su índice de precios de producción, se excluye de la regresión, la tasa de rendimiento cae al 13,4%. Además, cuando el crecimiento de la PTF en semiconductores y productos farmacéuticos también se corrige por el cambio de calidad de los productos, las estimaciones aumentan nuevamente al 34,8%, incluso sin incluir la industria informática.

sectoriales o los variables dummies no captan completamente las diferencias de calidad y estas últimas están correlacionadas con las variables explicativas.

Por último, en la literatura se ha destacado la necesidad de modelar la productividad como una variable endógena a la inversión en I&D (enfoque de “productividad endógena”). Específicamente destaca el estudio de Doraszelski and Jaumandreu (2013) que plantea un método para evaluar el impacto sobre la productividad de las empresas de la inversión en actividades de innovación haciendo endógena la productividad al modelar esta inversión de forma dinámica en una extensión del modelo de capital-conocimiento de Griliches (1979). A partir de datos de la encuesta de estrategias empresariales en España, encuentran que el vínculo entre la I&D y la productividad está sujeto a un alto grado de incertidumbre, no linealidad y heterogeneidad entre las empresas. Destacan que no tomar en cuenta la endogeneidad de la productividad, podría llevar a subestimar de forma importante los diferenciales esperados en la PTF entre las empresas que invierten en I&D y aquellas que no lo hacen.

A partir de este trabajo, distintos estudios comparan los enfoques tradicionales de medición de la PTF como el método OLS y la estimación ajustada resaltando la importancia de tratar los insumos y la productividad como endógenos. Por ejemplo, Constanza Demmel *et al.* (2013) estudian el vínculo innovación-productividad a nivel de empresas para cuatro países de América Latina (Argentina, México, Colombia y Perú) para el sector manufacturero distinguiendo entre innovaciones de procesos y de productos, usando el panel de datos de los *Enterprise Surveys* del Banco Mundial para los años 2006 a 2010. Su estrategia de estimación sigue dos pasos: en primer lugar, estimar la medida de PTF siguiendo el procedimiento de estimación de Wooldridge (2009), el cual permite comparar los resultados considerando un proceso a lo Markov exógeno o uno endógeno para la dinámica de la productividad (en este último caso las innovaciones previas pueden afectarla), y, posteriormente, utilizar la PTF estimada como variable dependiente de las actividades de innovación. Sus resultados para Argentina y México confirman que las empresas más productivas se autoseleccionan en actividades de innovación bajo el enfoque de productividad endógena, siendo este resultado impulsado por las innovaciones de productos. Además, para estos países, encuentran retornos a la innovación en términos de productividad para ambos tipos de innovación, tanto bajo un proceso de productividad a lo Markov exógeno como para uno endógeno.

Recientemente, Guillard, Jaumandreu y Olivari (2018) discuten formas de aplicar el modelo de productividad endógena cuando no hay índices de precios de producción a nivel de empresa disponibles, una limitación de muchas bases de datos. Al incluir la demanda de la empresa en la estimación obtienen que el “compuesto” de productividad, elasticidad de la demanda y heterogeneidad de la demanda que representa la “productividad de ingresos”, no se comporta como la productividad y, en particular, tampoco es mayor para las empresas que realizan I&D ni su distribución muestra dominio estocástico. Su persistencia y rendimientos también dan resultados engañosos, y sugieren cautela en la interpretación de los resultados obtenidos. Por ello, resaltan la importancia de producir bases de datos más completas, especialmente si se quieren extraer implicaciones de política.

2.3 Adopción y uso de tecnologías digitales en Perú

El INEI publica informes anuales donde analiza los resultados obtenidos a partir de la información módulo de Tecnologías de Información y Comunicación de la Encuesta Económica Anual (EEA), el cual incluye también información sobre la inversión en actividades de inversión en ciencia y tecnología (I&D), y es aplicado a las grandes, medianas y pequeñas empresas de las distintas ramas de actividad¹³.

Así, de acuerdo con los datos de la EEA 2018, de las 82 mil 249 empresas que desarrollaron actividades económicas el año 2017, el 94,2% utilizó computadoras, el 92,6% el servicio de internet, el 18,9% hicieron uso de intranet, el 15,4% de PDA (Asistente Digital Personal) / Tablet y el 8,3% de extranet; mientras, que el 93,3% utilizó telefonía móvil y el 89,7% telefonía fija (INEI, 2020). En cuanto a los principales usos de internet por las empresas, el 59,2% lo utilizó para comunicarse vía email/chat, el 49,1% para la búsqueda de productos y/o servicios, el 41,3% para realizar operaciones de banca electrónica, el 35,5% para otras búsquedas de información, el 27,5% para ofrecer servicios a sus clientes y el 25,9% para buscar información de organismos gubernamentales. El menor uso de internet fue para distribuir productos en línea, el cual solo llegó a 2,3% (INEI, 2020).

Por su parte PRODUCE ha realizado análisis específicos de un conjunto de indicadores de adopción y uso de las TIC por parte de las micro, pequeñas y medianas empresas (Mipymes) a partir de los datos recogidos en la Encuesta Nacional de Empresas (ENE), la cual abarca a las empresas que venden más de 13 UIT al año (a excepción del sector agropecuario).

En este caso, respecto al servicio de internet para el año 2017, PRODUCE (2020a) indica que el 91,0% de las Mipyme peruanas contó con el mismo, pero su calidad requiere una mejora significativa, pues el 41,2% de las Mipyme que emplearon el servicio reportaron dificultades en su uso referidas, sobre todo, a las interrupciones en el servicio y la velocidad de conexión. Asimismo, el uso principal del servicio de internet por parte de las Mipyme se vinculaba a la coordinación y gestión, mediante e-mail o chat (88,2%), y a la búsqueda de información (51,2%), especialmente de productos o servicios (49,9%). El uso de internet para la promoción de sus productos o servicios era de 21,3%, y para la venta era de 14,2%, mientras que el uso con motivos de capacitación solo fue de 5,3%.

Asimismo, PRODUCE (2020a) ha elaborado un Índice de adopción de las TIC utilizando los datos de la ENE; siguiendo el estudio de Alderete (2012). Este índice sintético incorpora distintas dimensiones de la disponibilidad, utilización y capacidad del uso de las TIC a nivel de empresas, a partir de tres indicadores compuestos alternativos: uso, preparación e impacto de las TIC mediante la metodología Análisis de Componentes Principales (ACP). Para el año 2017 (usando la ENE 2018) los valores promedio del índice de digitalización estimados a nivel sectorial, indicaban que las empresas de los

¹³ El detalle de las variables recogidas se indica en la sección 4.3.2 sobre los datos utilizados en el estudio.

sectores de servicios y comercio tienen mayor índice de digitalización, lo cual podría deberse a que buscan proporcionar valor añadido a los clientes mediante la integración de un mismo sistema de publicidad y marketing de ventas, y la atención personalizada (INEI, 2015), y que la brecha existente entre sectores es relativamente pequeña.

2.4 Evidencia del impacto de la innovación en la productividad para Perú

En el caso del Perú, un primer estudio sobre el efecto de las TIC en el caso de las empresas de menor tamaño es el de Huaroto (2012), el cual mide el efecto de la adopción de Internet sobre la productividad laboral de las microempresas de la Encuesta Nacional de Hogares (ENAHOG) del período 2007- 2010, medida como el valor agregado por hora trabajada, construyendo un índice de Adopción con escala del 1 al 100 (“Índice de Lefebvre y Lefebvre” – ILL) que considera las distintas aplicaciones que realiza la empresa del Internet. Se estima un modelo de Primeras Diferencias, y se obtiene como resultado que un incremento en un punto del índice de adopción del empresario tiene como resultado un incremento significativo y robusto en la productividad de su microempresa similar al 1.5% de la productividad promedio de la muestra.

Asimismo, utilizando los datos de tipo panel de la Encuesta de Micro y Pequeña Empresa (EMYPE) para el sector manufactura en el período 2011-2013, Viollaz (2018) estima una forma reducida de la productividad laboral como función de la adopción de Internet y otros factores explicativos. Para ello instrumentaliza la adopción de Internet utilizando una medida de la disponibilidad de oportunidades financieras para las empresas. Sus resultados indican que la adopción de Internet aumenta la productividad laboral de las empresas, al mismo tiempo que se produce una reasignación de trabajadores administrativos temporales y trabajadores no remunerados hacia trabajadores permanentes de producción, aumenta la capacitación laboral, y hay una mayor aplicación de nuevas prácticas organizativas (de gestión, innovación y uso de las TIC).

En cuanto a los estudios realizados para empresas de mayor tamaño utilizando la EEA, la investigación de Tello (2017) analiza la relación entre las decisiones de inversión, la intensidad de la inversión, los resultados de la innovación y la productividad del trabajo para empresas de servicios y manufactureras del Perú en 2004, sobre la base de un modelo CDM. Muestra que el tamaño de la empresa juega un rol clave en la decisión de invertir en innovación. Asimismo, encuentra que la intensidad de la inversión en actividades de ciencia, tecnología e innovación (CTI) influye positivamente en la productividad laboral de las empresas.

Tello (2020) analiza y estima el impacto de la inversión en I&D (interna y externa) sobre el proceso de innovación y productividad laboral, tomando en cuenta el uso e importancia de las TIC, para las empresas formales manufactureras del período 2009-2014 a partir de los datos de la Encuesta Nacional de Innovación en la Industria Manufacturera (ENIIM) de 2012 y de 2015 (para los períodos 2009 a 2011 y de 2012 a 2014, respectivamente). Luego de estimar un modelo tipo CDFN encuentra que, en promedio, las empresas manufactureras en el Perú no están aprovechando las ventajas de las actividades de I&D (internas y externas) y de innovación, y continúan dependiendo del capital físico y humano para incrementar la productividad laboral.

Asimismo, en las estimaciones que las empresas consideren el uso de internet como de mediana y alta importancia por las empresas, incidió positivamente sobre los resultados de la innovación. No obstante, este último resultado no incidió de forma sustantiva en la productividad laboral de las empresas.

Asimismo, Lozada (2019) utiliza un panel de datos de la EEA para el periodo 2009-2015 para estimar el efecto de la diversificación productiva (a nivel de industrias usando un índice de entropía) sobre la PTF (medida mediante el método OL) en las empresas manufactureras peruanas. A fin de controlar los problemas de endogeneidad y sesgo de selección utiliza un modelo en tres etapas: (i) decisión de diversificar, (ii) grado de diversificación y (iii) resultados de la diversificación. Los resultados muestran que un aumento el grado de diversificación de las firmas tiene un efecto positivo y estadísticamente significativo en el crecimiento de su valor agregado y en su productividad.

Por otro lado, Nolazco (2020) examina la relación entre el esfuerzo innovador, el nivel de exportaciones y la productividad en las empresas manufactureras peruanas con datos de la Encuesta Nacional de Innovación en la Industria Manufacturera (ENIIM) de 2015 para el periodo 2012 a 2014. Para ello adapta un modelo tipo CDM que considera los siguientes aspectos: 1) autoselección (*self-selection*), 2) exportaciones como determinante de la innovación (*innovating by exporting*), 3) innovación como generador de una mayor oferta de productos exportables (*exporting by innovating*), 4) aprendizaje por exportar (*learning by exporting*) y 5) el impacto directo e indirecto de la innovación sobre la productividad. Al realizar estimaciones Tobit y regresiones por quintiles, concluye que la productividad inicial de la empresa es determinante para exportar (autoselección), y que exportar, a su vez, incentiva a gastar en innovación (exportaciones como determinante de la innovación) para competir en el mercado internacional. Además, la innovación conlleva una mayor oferta de productos exportables (innovación como determinante de las exportaciones) y un aumento en la productividad (*learning by exporting*). Así, encuentra que los efectos directos e indirectos de la innovación tecnológica implican incrementos de 23 % a 44 % y 16 % a 45 % según quintil de la productividad laboral y de la PTF, respectivamente.

Recientemente PRODUCE (2020b) ha presentado los resultados de la estimación del modelo CDM sobre la base de la Encuesta Nacional de Innovación de la Industria Manufacturera y Empresas de Servicios Intensivos en Conocimiento de 2018 (ENIIMSEC 2018), hallando que la innovación tiene un impacto positivo sobre la productividad laboral de las empresas. Para el periodo 2015-2017, la innovación se asocia, en promedio, con un aumento de 34.2% de la productividad laboral para las empresas de las ramas de actividad analizadas en conjunto. Este porcentaje es ligeramente superior (36.7%) si se analiza solo para el sector manufactura.

3 Metodología de análisis empírico

3.1 Pasos de la metodología de estimación

En el presente trabajo se plantea la estimación del efecto de la inversión en I&D y del uso en TIC sobre la PTF de las empresas manufactureras de Perú utilizando la información panel de la EEA 2016-18, mediante los siguientes pasos:

1. Estimación de la PTF (en logaritmo) mediante distintos métodos econométricos y selección de la estimación que resulte más adecuada a los datos.
2. Estimación de la inversión en I&D mediante modelo de corrección de sesgo de selección de Heckman.
3. Estimación de un índice de digitalización mediante análisis factorial para variables de uso de las TIC dada la alta correlación entre estas variables, y construcción de índices sintéticos de uso de las TIC.
4. Estimación de los *mark-up* de cada empresa.
5. Regresión panel de determinantes de la PTF (en logaritmo) respecto al valor predicho de la intensidad de inversión I&D (nivel de inversión en I&D entre valor real de las remuneraciones) y variables sintéticas del uso de las TIC. Como alternativa se realiza también la estimación utilizando las variables de uso de las TIC desagregadas.

Se incluye distintas variables de control como tamaño, participación femenina, edad o antigüedad, rama de actividad, y región. Asimismo, se incluye el *mark-up* estimado entre las variables determinantes de la PTF.

En este caso, respecto a la posible endogeneidad el uso de las TIC, solo se la está corrigiendo mediante el uso del modelo de efectos fijos que captura el efecto de variables no observables que se mantienen fijas en el tiempo.

3.2 Estimación de la función de producción

A partir de la estimación de los coeficientes β_l y β_k mediante los métodos descritos previamente es posible calcular el logaritmo de la PTF de cada empresa ($\omega_{i,t}$) utilizando la siguiente fórmula:

$$PTF_{i,t}: \hat{v}_{i,t} = \alpha_{i,t} - \hat{\beta}_l l_{i,t} - \hat{\beta}_k k_{i,t}$$

En este trabajo se hace la estimación de la PTF a nivel de empresas (en logaritmo) utilizando el Software STATA 16.0. Las estimaciones se realizan mediante los métodos OLS, panel de datos con efectos aleatorios y con efectos fijos, y para los métodos modernos se utiliza el comando *prodest* desarrollado por Rovigatti y Mollisi (2018), el cual que permite realizar las estimaciones con los métodos siguientes.

1. Mínimos cuadrados ordinarios (OLS)

2. Panel con efectos aleatorios (RE)
3. Panel con efectos fijos (FE)
4. Olley y Pakes (OP)
5. Olley y Pakes con corrección por atrición (OP-A)
6. Olley y Pakes con corrección de Akerberg, Caves y Frazer (OP-ACF)
7. Levinsohn y Petrin (LP)
8. Levinsohn y Petrin con corrección por atrición (LP-A)
9. Levinsohn y Petrin con corrección de Akerberg, Caves y Frazer (LP-ACF)
10. Woolridge (WRDG)
11. Mollisi y Rovigatti (MR)

Para la estimación el comando *prodest* distingue tres grupos de variables:

- Variables de estado: cuyo valor esta dado al momento de la decisión sobre la producción del momento t . Es típicamente el capital que evoluciona de acuerdo con la función de inversión $i(.)$ que se decide en el tiempo $t-1$.
- Variables proxy: aquéllas utilizadas para aproximar el nivel del shock de productividad observado por las empresas para tomar sus decisiones de producción. Son la inversión en el método OP y los insumos en el método LP.
- Variables libres: variables no dinámicas, en el sentido de que su valor en el momento t no afecta los beneficios futuros, y se eligen en el momento t cuando la empresa conoce el shock de productividad. Son la mano de obra y los materiales intermedios o insumos.

3.2.1 Limitaciones de los métodos econométricos utilizados

En los últimos años se ha realizado varias investigaciones que criticado los métodos existentes de medición de la PTF utilizados por la mayoría de los estudios académicos. En principio, se ha estudiado las discrepancias entre la estimación de la productividad utilizando funciones de producción basadas en los ingresos o valores y no en las cantidades específicas producidas (al no contarse con información detallada de precios y cantidades), lo cual pueden distorsionar la estimación de la productividad mediante un residuo o valor no observable respecto al valor de la producción o del valor agregado, a menudo llamado “productividad de ingresos”.

Específicamente, en la estimación se produce un sesgo de estimación al no considerar los precios debido a que la aproximación a las cantidades de la producción y los factores de producción (como capital e insumos intermedios) son obtenidos deflactando los valores con índices de precios agregados (Tello, 2014). Así, en el caso de la producción de un solo producto

$$y_{it} = q_{it} + (p_{it} - IP_{it})$$

Donde p_{it} es el precio del producto para la empresa i en el período t , IP_{it} es el deflactor y q_{it} la cantidad del producto.

Entonces,

$$y_{it} = \beta_0 + \beta_l l_{i,t} + \beta_k k_{i,t} + \beta_m m_{it} + (p_{it} - IP_{it}) + \omega_{i,t} + \varepsilon_{i,t}$$

En las estimaciones de la diferencia de precios no se incluye lo cual origina un sesgo en las estimaciones por omisión de variables. De otro lado, esta diferencia de precios origina correlaciones entre los factores de producción y los “errores” que incluyen esta diferencia de precios. Estos problemas se agravan con los deflatores del capital y del valor de los insumos (Van Beveren, 2012)

Al respecto, se ha enfatizado específicamente que muchos estudios no han tomado en cuenta los efectos de las características de la demanda de los productos y de las imperfecciones en el funcionamiento de los mercados de bienes y de insumos. Los modelos suelen asumir que existe competencia perfecta entre las empresas, pero cuando las empresas tienen algún poder de mercado, por ejemplo, cuando los productos están diferenciados, entonces la demanda específica que enfrenta cada empresa se debe considerar en la especificación de las funciones de demanda de insumos (Mairesse y Jaumandreu, 2005).

Así, si se considera que las empresas que enfrentan una función de demanda con pendiente descendente que depende del precio de la producción p_{jt} y existen factores z_{jt} que influyen en la demanda, la maximización requiere que las empresas establezcan el precio que equipara el costo marginal con el ingreso marginal $P_{jt} \left(1 - \frac{1}{\eta(p_{jt}, z_{jt})}\right)$, donde $\eta(\cdot)$ es el valor absoluto de la elasticidad de la demanda.

Con las empresas minimizando los costos, el costo marginal y la demanda de mano de obra condicional se pueden determinar a partir de la función de costos y combinarse con los ingresos marginales para obtener la función inversa de demanda de mano de obra.

$$h^{IC}(l_{jt}, k_{jt}, w_{jt} - p_{jt}, p_{jt}, z_{jt}) = \lambda_0 + (1 - \beta_l)l_{jt} - \beta_k k_{jt} + (w_{jt} - p_{jt}) - \ln \left(1 - \frac{1}{\eta(p_{jt}, z_{jt})}\right)$$

En este caso, $h^{IC}(\cdot)$ reemplaza a $h(\cdot)$ en la ecuación de estimación, y se espera que tanto p_{jt} como z_{jt} no estén relacionados con la innovación de la productividad $\omega_{i,t}$.

En general, se ha encontrado que las diferencias en resultados haciendo este tipo de corrección en las estimaciones no serían sustantivas (Mairesse y Jaumandreu, 2005), siendo más importantes otras fuentes de errores de especificación de los modelos econométricos.

Respecto a estas dificultades, Forlani *et al.* (2016) desarrollan un instrumental econométrico que permite simultáneamente recuperar la heterogeneidad en la demanda, la PTF y los márgenes entre las empresas, dejando la correlación entre las tres sin restricciones. Para ello explotan sistemáticamente los supuestos que están implícitos en los enfoques anteriores de estimación de la productividad a nivel de empresa. Utilizan datos de producción de las empresas belgas para cuantificar la PTF, la demanda y los márgenes y mostrar cómo se correlacionan entre ellos, a través del tiempo y con las medidas obtenidas mediante otros enfoques. Las tres dimensiones de heterogeneidad permiten obtener información más profunda y nítida sobre el estado y el tamaño de las exportaciones a nivel de empresa.

Asimismo, con el fin tomar en cuenta en la estimación de la PTF a nivel empresas las diferencias en poder de mercado, es posible considerar los trabajos de De Loecker (2011a) y De Loecker y Warzynski (2012) que desarrollan un enfoque micro de estimación de los *mark-up* a partir de la estimación de la función de producción.

Dados los supuestos de *Hick's neutral productivity shocks* y la minimización de costos por parte de las empresas el *mark-up* para la empresa k en la industria s viene dado por:

$$\mu_{kst} = \frac{P_{kst}}{MC_{kst}} = \text{Labour elasticity}_{kst} / \text{Labour share}_{kst}$$

Es decir, el valor del *mark-up* puede calcularse a partir de las elasticidades de demanda del trabajo y la participación del trabajo en el valor agregado, los cuales pueden ser obtenidos de la estimación de la PTF realizada con los métodos estándar mencionados previamente¹⁴.

Sin embargo, de acuerdo con Doraszelsky & Jaumandreu (2019, 2021), esta estimación presenta algunos problemas pues estimar consistentemente las elasticidades de los insumos o factores y la perturbación mediante el procedimiento desarrollado en Olley & Pakes (1996) y Levinsohn & Petrin (2003) generalmente requiere observar y controlar el *mark-up*. Además, Doraszelsky & Jaumandreu (2019) indican que las elasticidades de los insumos o factores son una función de no observables (como costos de ajuste) y plantean ajustes a su estimación. También existen sesgos similares de estimación cuando las empresas producen más de un producto mediante tecnologías distintas para cada producto o si las respectivas demandas son diferentes (produciendo diferentes precios)¹⁵.

Por último, Gandhi *et al.* (2011) resaltan que pueden surgir diferencias en productividad entre las empresas dependiendo de si la productividad se mide utilizando una función de producción bruta o de valor agregado que resta el valor de los insumos “flexibles” de una planta (materiales, energía, etc.) del valor de la producción bruta. Sin embargo, el

¹⁴ A partir de los *mark-up* estimados es posible ajustar el valor estimado de la PTF de cada empresa utilizando la siguiente identidad:

$$TFP_{kst}^{adj} = \log(TFP_{kst}) - \log(\mu_{kst})$$

Sin embargo, dada la fuerte relación entre el *mark-up* y la innovación a través de la I&D, utilizar esta medida de productividad ajustada por *mark-up* conduciría que no se encuentran los efectos de la innovación ampliados observados en la literatura. Dai y Cheng (2018) documentan este aspecto para el caso de las empresas manufactureras chinas. Sus resultados sugieren que la innovación de productos aumenta el margen de beneficio a nivel de empresa y la productividad (TFP basados en datos de ingresos). Sin embargo, dado que el efecto de la innovación de productos en la productividad ajustada es mayormente negativo o no significativo estadísticamente, los autores indican que el impacto positivo de la innovación de productos sobre la productividad (de ingresos) se debería principalmente a los cambios en el margen de beneficio precio-costo en lugar de las mejoras de la productividad física.

¹⁵ De Loecker (2011b) presenta un método simple de abordar el problema usando el número de productos por empresa.

principal obstáculo para utilizar la producción bruta en el trabajo aplicado es la falta de una fuente clara de identificación. Solo en el caso de cumplirse el supuesto de *Hick's neutral productivity shocks* – con funciones de producción tipo Leontief en los insumos y valor agregado- los parámetros de los factores de producción (como el capital y el trabajo), utilizando ambas medidas de la producción, serían equivalentes (Gandhi *et al.*, 2014).

Gandhi destaca que el sesgo de transmisión al estimar la función de producción con el valor de producción es más severo para los insumos flexibles (materiales, energía, etc.) pues en un entorno competitivo estándar, no existen restricciones de exclusión que puedan identificar las elasticidades de los insumos flexibles. Dado que estos mismos insumos se restan de la producción bruta para formar el valor agregado, la función de producción de valor agregado se puede identificar a través de instrumentos cuando los factores, es decir, el capital y el trabajo, son “inflexibles” de alguna manera, ya sea debido a restricciones de tiempo o costos de ajuste (Bond y Söderbom, 2005), Akerberg *et al.* (2006)). Sin embargo, la solución de estimar el valor agregado implica que, si los “insumos flexibles” no entran en la función de producción a la manera de Leontief, entonces la construcción de valor agregado causaría su propia fuente de sesgo para medir la productividad (Basu y Fernald, 1997).

Gandhi *et al.* (2016, 2020) indican que los enfoques de “variable proxy” o funciones de control utilizados para estimar las funciones de producción sufren genéricamente de un problema de identificación fundamental en presencia de insumos “flexibles” que satisfacen los supuestos de “variable proxy” para la invertibilidad como los materiales del consumo intermedio (la fuente de la falta de identificación sería la elasticidad del insumo flexible). Por ello, plantean una estrategia de identificación para las funciones de producción del valor de producción mediante la estimación basada en la condición de primer orden de maximización de ganancias de las empresas, y el uso de fuentes adicionales de variación en la demanda de insumos flexibles (por ejemplo, precios). Al aplicar esta metodología de estimación a datos a nivel de plantas de Chile y Colombia y sus resultados sugieren que el sesgo introducido al usar el valor agregado es al menos tan importante, si no más, que el sesgo de transmisión en sí.

Por último, pueden existir otras fuentes de heterogeneidad de la productividad a nivel de empresas que no se hayan considerado en la estimación de la PTF. Al respecto, actualmente, existe una amplia literatura respecto a varios hechos estilizados sobre la heterogeneidad de la productividad a nivel de empresas. Incluso en industrias definidas de forma específica se ha observado una fuerte dispersión de la productividad no inexplicable (Collard-Wexler, 2010). Además, la productividad está estrechamente relacionada con otras dimensiones de la heterogeneidad a nivel de empresas, como la importación (Kasahara y Rodrigue, 2008) y la exportación (Bernard y Jensen, 2004).

3.2.2 Inclusión de capital TIC en la estimación

Dado que la información disponible solo cuenta con datos de capital TIC relativos al valor del software con el que cuenta la empresa, en el presente trabajo no incluye el capital TIC en la estimación de la función de producción y de la PTF.¹⁶

3.3 Relación entre el uso de las TIC y la Productividad

A partir de la estimación de la PTF para las empresas de la industria manufacturera utilizando los datos de la EEA, es posible realizar un análisis exploratorio respecto a las diferencias en la PTF respecto a variables relativas a características del conductor y de la empresa identificadas en la literatura como determinantes de la productividad.

Entre las variables de la empresa disponibles en la información de la EEA tenemos su tamaño, antigüedad, subsector productivo, localización, diversificación de productos. Respecto al conductor de la empresa, no se tiene mayor información, excepto el género del mismo.

Cabe indicar que el análisis se realiza respecto a un grupo priorizado de las variables TIC que podrían influir positivamente en la PTF de las empresas. Asimismo, la información sobre acceso y uso de las TIC se utilizó para construir índices de uso de las TIC de las empresas de forma similar al índice de digitalización de PRODUCE (2020a) y, posteriormente, se analiza la relación con el nivel de productividad empresarial.

4 Información utilizada y estimación de variables

4.1 Datos utilizados para estimar la productividad

Al ser la productividad – entendida como la PTF - una variable no observable, su medición es difícil y está condicionada a la calidad de los datos.

Para la estimación empírica de las funciones de producción usualmente se han empleado paneles de datos de empresas o plantas con información anual y mediciones de producción valorizada, ingresos, ventas o valor agregado (output), los factores de producción son aproximados a través de mediciones específicas, en función de la disponibilidad de información (Van Beveren, 2012; Aguirregabiria, 2018): trabajo (gasto de personal, número de trabajadores), capital (activos totales, activos fijos), insumos (gasto total en insumos).

En el caso de Perú, la fuente de información más importante para estimar la PTF ha sido la Encuesta Económica Anual (EEA) realizada anualmente por el Instituto Nacional de Estadística e Informática (INEI), la cual permite construir muestras panel de empresas para distintos períodos de tiempo.

¹⁶ Al realizar, un ejercicio incluyendo la ratio del valor de este tipo de capital respecto al valor total del capital, como variable de control, esta resultó con un efecto positivo altamente significativo a nivel estadístico.

La EEA releva información económica y financiera de empresas en 14 actividades económicas: agencias de viajes, agroindustria, centros educativos no estatales, comercio, construcción, establecimiento de hospedaje, hidrocarburos, pesca, manufactura, servicios eléctricos, transporte y comunicaciones, universidades no estatales, restaurantes y servicios. La población está compuesta por más de 75 mil empresas a nivel nacional, de la cual la muestra de empresas relevadas en EEA es probabilística de la población de empresas registradas en el Padrón de Contribuyentes de la Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT)¹⁷.

El conjunto de datos puede ser expresado como (Aguirregabiria, 2018):

$$Data = \{Y_{i,t}, L_{i,t}, K_{i,t}, M_{i,t} : i = 1, 2, \dots, N; t = 1, 2, \dots, T\}_{i,t} \quad (2)$$

Donde $Y_{i,t}$ es el resultado del proceso productivo (output), el proceso productivo se basa en factores de producción: trabajo, capital e insumos ($L_{i,t}, K_{i,t}, M_{i,t}$, respectivamente), N es el número de empresas “ i ” disponibles en la muestra y T es el número de periodos con información disponible de las empresas.

En la Tabla se muestran las principales variables construidas a partir de la información de la EEA de forma similar al trabajo de Del Pozo y Guzmán (2019) tomando el valor monetario reportado en el Estado de Situación Financiera de cada empresa.

El nivel de producción es la suma de la producción del ejercicio correspondiente y el margen comercial. A su vez, la producción es la suma de las ventas netas, prestación de servicios, variación de la producción en almacenes y la producción del activo inmovilizado. El margen comercial es la suma de las ventas netas de mercancías más la variación de existencias de mercancías menos el valor de la compra de mercancías.

En segundo término, el factor trabajo ha sido aproximado a través del valor monetario de los gastos de personal (que incluye a trabajadores de planta, directores y gerentes), se ha considerado dicho indicador para aproximar el precio del trabajo en el proceso productivo de acuerdo con lo sugerido por Aguirregabiria (2018).

En tercer término, el factor capital ha sido aproximado a través del valor monetario del activo total no corriente neto de depreciación, conocido también como activo fijo, que son aquellos activos de la empresa que tengan condición de uso a largo plazo, tales como terrenos, patentes, inversiones y similares (Llamas, 2020).

¹⁷ La EEA incluye información de empresas que registran valor de ventas, en condición de activas y habidas en SUNAT. No incluye información de empresas agrícolas, empresas de educación pública, empresas del gobierno central, organismos extraterritoriales, instituciones financieras y empresas mineras. La muestra es probabilística, estratificada, unietápica e independiente en cada departamento, de estrato forzoso y auto-representada.

Las principales características de la EEA de acuerdo con su ficha técnica se indican en el Anexo 2.

El valor de los insumos se obtiene como la suma de las compras de materias primas, insumos auxiliares, suministros, repuestos, envases y embalajes, más costos vinculados con las compras.

La inversión es el valor de la suma de inversiones inmobiliarias (terrenos y edificaciones) más inversiones en maquinaria y equipo, e intangibles y activos biológicos.

Cuadro 3: Variables utilizadas en la estimación de la PTF

Variables	Definición
$Y_{i,t}$	Producción total: suma de la producción total del ejercicio y el margen comercial.
$L_{i,t}$	Gastos de personal, directores y gerentes.
$K_{i,t}$	Capital: total activo no corriente neto de depreciación.
$M_{i,t}$	Insumos: suma de las compras de materias primas, auxiliares, suministros, repuestos, envases y embalajes más costos vinculados con las compras.
$I_{i,t}$	Inversión: valor de la suma de inversiones inmobiliarias más inversiones en maquinaria, equipo y otros.

Fuente: Basado en Del Pozo y Guzmán (2019) a partir de Encuesta Económica Anual (INEI).

A fin de hacer comparables los valores en el tiempo, las variables expresadas en valores monetarios fueron deflactadas temporalmente utilizando el Índice de precios nacional¹⁸. Asimismo, los paneles se han conformado solo con empresas que presentan información completa de producción e insumos por al menos tres años consecutivos (considerando el año de ejercicio al que corresponde la información).

Tratamiento de observaciones con valores extraños

Dado que en los datos se observa posibles anomalías debido a observaciones que tendrían problemas de medición o constituirían eventos extraños dado el patrón general de la información, se ha procedido a utilizar un método estándar para identificar la existencia de observaciones con valores extraños (*outlier*) en un contexto de análisis multivariado. Específicamente, se ha optado por utilizar el método *Generalized Stahel-Donoho estimator of multivariate location* propuesto por Verardi y Vermandele (2018), el cual se implementa utilizando el comando *sdasym* disponible para Stata 16.0. Dado que las posibles observaciones identificadas con valores extraños en las variables utilizadas para la estimación de la función de producción no muestran valores que no sean verosímiles, no se ha procedido a su eliminación.

4.2 Estimación de la productividad

A continuación, se muestran los resultados de las estimaciones de la función de producción para el valor agregado obtenidos usando el panel de datos no balanceado de los años de ejercicio 2016 a 2018¹⁹.

¹⁸ De acuerdo con la disponibilidad de información.

¹⁹ Los principales estadísticos descriptivos de las variables utilizadas se presentan en el Anexo 3.

Cuadro 4: Estimación de la Productividad Total de Factores del Sector Manufactura: Periodo 2016-2017-2018

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	OLS	FE	RE	OP	OP-A	OP-ACF	LP	LP-A	LP-ACF	WRDG	MR
Trabajo	0.598*** (0.030)	0.456*** (0.159)	0.580*** (0.041)	0.556*** (0.030)	0.556*** (0.030)	0.607*** (0.115)	0.524*** (0.028)	0.524*** (0.028)	0.607*** (0.202)	0.556*** (0.018)	0.528*** (0.031)
Capital	0.230*** (0.022)	0.211 (0.144)	0.225*** (0.031)	0.321*** (0.124)	0.343*** (0.121)	0.239*** (0.063)	0.352*** (0.121)	0.337*** (0.121)	0.239*** (0.088)	0.344*** (0.051)	0.345*** (0.119)
Constant	3.688*** (0.305)	6.160*** (2.038)	4.020*** (0.305)								
Observations	2,240	2,240	2,240	2,240	2,240	2,240	2,240	2,240	2,240	1,221	2,240
R-squared	0.714	0.713	0.714								
RMSE	0.712	0.331	0.438								
Prob > F	0	2.41e-06									
Prob > Wald (CRS)				0.290	0.380	0.330	0.280	0.230	0.580	0.0600	0.260
Prob > Chi2			0								

Notas:

Errores estándar robustos entre paréntesis. *** p<0.01, ** p<0.05, * p<0.1

A: corregido por atrición en los datos, es decir, por salidas de empresas.

ACF: Ackerberg, Caves and Frazer correction

Wald test on Constant returns to scale

Fuente: Elaboración propia en base al Panel EEA 2016-18.

En general, los resultados obtenidos con los distintos métodos no difieren de forma sustancial, excepto en el caso de los métodos OP-ACF y LP-ACF que realizan la corrección propuesta por Akerberg, Caves y Frazer.

Sin embargo, el método OLS y el método panel de efectos aleatorios estarían sobrestimando el coeficiente de la variable trabajo, de acuerdo con lo esperado, y subestimando el valor del coeficiente del capital, mientras el método de efecto fijos lo estaría subestimando al no tomar en cuenta toda la variabilidad observada en los datos. Además, al realizar la corrección por sesgo de selección varía ligeramente el valor estimado del coeficiente del capital, aumentando con el método OP-A y reduciéndose en el caso del método LP-A.

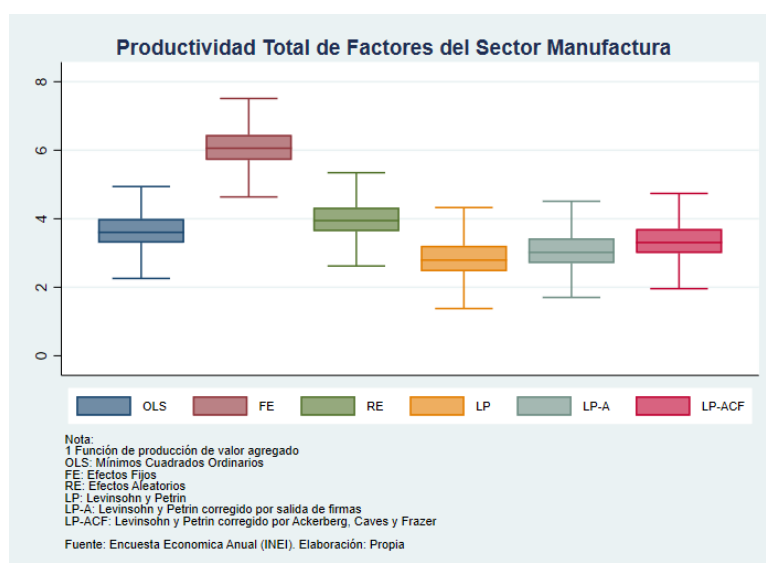
La corrección ACF de la estimación reduce el coeficiente del capital tanto para el método OP como el LP. Mientras tanto, el método WRDG proporciona un coeficiente estimado para el capital significativamente mayor a los métodos OP-ACF y LP-ACF.

Por último, los resultados del método MR son similares a los del método WRDG con coeficientes ligeramente mayores, y proporcionarían resultados más eficientes que los métodos OP y LP. Sin embargo, debe recordarse que estos métodos no corrigen el sesgo de selección por sobrevivencia de las empresas.

En todas las especificaciones se rechaza la hipótesis nula de retornos a escala constantes. Resultan de tipo decreciente al sumar menos de 1 los coeficientes del capital y el trabajo.

En todo caso, dado que en los paneles no balanceados de tres años existen un grupo significativo de empresas que salen en el segundo y en el tercer año, las cuales suelen ser de menor tamaño, y que existen valores de las variables de inversión muy pequeños, para el análisis cuantitativo de la PTF a nivel de empresas se ha optado por utilizar los resultados de la estimación utilizando el método Levinsohn y Petrin con corrección por atrición (LP-A), debido a que un porcentaje significativo de empresas salen del panel el segundo o el tercer año siendo, en promedio, de menor tamaño.

Gráfico 1: Productividad Total de Factores promedio según métodos de estimación Manufactura Panel 2016-2018



Fuente: Elaboración propia en base al panel EEA 2016-18.

Cuadro 5: Productividad Total de Factores promedio según métodos de estimación Manufactura Panel 2016-2018

	mean	sd	p25	p75
OLS	3.677	0.635	3.299	3.994
FE	6.115	0.660	5.715	6.447
RE	4.016	0.635	3.636	4.324
OP	2.848	0.654	2.467	3.182
OP-A	2.502	0.665	2.111	2.852
OP-ACF	3.380	0.638	3.000	3.702
LP	2.853	0.662	2.466	3.210
LP-A	3.087	0.655	2.706	3.430
LP-ACF	3.379	0.638	2.999	3.702
WRDG	2.481	0.665	2.090	2.831
MR	2.884	0.660	2.497	3.240

Nota: El valor estimado de la PTF incluye la constante de los modelos.
 Fuente: Elaboración propia en base al panel EEA 2016-18.

4.3 Medición del uso de las TIC y de la inversión en I&D

4.3.1 Medición del capital TIC

Respecto a la medición del capital TIC, se debe mencionar que los datos de la EEA no permiten distinguir este tipo de capital. En el módulo que recoge el resumen del movimiento del activo fijo e intangibles (capítulos 07M y 07D2 para pequeñas empresas y empresas medianas y grandes, respectivamente) los rubros incluidos no brindan información específica sobre este tipo de capital. Solo se tiene capital en software como parte del activo no intangible para las empresas medianas y grandes. Los rubros específicos son:

Cuadro 6: Información del movimiento del activo fijo e intangibles – EEA

A. Pequeñas Empresas

Código PCGE	Activo inmovilizado e intangibles	Clave
331	Terrenos	01
332	Edificaciones	02
333	Maquinaria, y equipos de explotación	03
334	Unidades de transporte	04
335	Muebles y enseres	05
336	Equipos diversos	06
338	Unidades por recibir	07
337+339	Herramientas y unidades de reemplazo y obras en curso	08
34	Intangibles	09

B. Empresas medianas y grandes

Código PCGE	Activo inmovilizado e intangibles	Clave
311	Inversiones Inmobiliarias – Terrenos	01
312	Inversiones Inmobiliarias – Edificaciones	02
32	Activos adquiridos en arrendamiento financiero (04 al 06)	03
321	Inversiones Inmobiliarias – Terrenos	04
321	Inversiones Inmobiliarias – Edificaciones	05
322	Inmuebles, maquinaria y equipo (07 al 13)	06
322	Terrenos	07
322	Edificaciones	08
322	Maquinarias y equipos de explotación	09
322	Unidades de Transporte	10
322	Muebles y enseres	11
322	Equipos Diversos	12
322	Herramientas y unidades de reemplazo	13
33	Inmuebles, maquinarias y equipo (15 al 23)	14
331	Terrenos	15
332	Edificaciones	16
333	Maquinarias y equipos de explotación	17
334	Unidades de Transporte	18
335	Muebles y enseres	19
336	Equipos Diversos	20
337	Herramientas y unidades de reemplazo	21
338	Unidades por recibir	22
339	Construcciones y obras en curso	23
Sub	Total de activos fijos (01+02+03+14)	24
34	Intangibles (26 al 33)	25
341	Concesiones, licencias y otros derechos	26
342	Patentes y propiedad industrial	27
343	Programas de computadora (software)	28
344	Costos de exploración y desarrollo	29
345	Fórmulas, diseños y productos	30

346	Reservas de recursos extraíbles	31
347	Plusvalía Mercantil	32
348	Otros activos intangibles	33
35	Activos biológicos (35+36)	34
351	Activos biológicos en producción	35
352	Activos biológicos en desarrollo	36
	Sub total de intangibles y activos biológicos (25+34)	37
	TOTAL (24+37)	38

Fuente: ENE.

Asimismo, la Encuesta Nacional de Empresas (ENE) tampoco tiene información separada para el capital en Tecnologías de la información y comunicaciones (TIC).

4.3.2 Variables sobre uso de las TIC e inversión en I&D

Desde el año 2013 se ha incorporado la EEA el módulo de Tecnologías de Información y Comunicación que recopila información tanto sobre las tecnologías (infraestructura, dispositivos, programas) y los servicios asociados, incluyendo el acceso a internet, el tipo de conexión y los usos del internet por parte de las empresas, así como el comercio electrónico y la capacitación que recibe el personal de las empresas sobre las TIC. Por otro lado, también se obtiene información sobre la inversión que realizan las empresas en ciencia y tecnología, es decir, investigación y desarrollo.

Aunque ha habido ligeras diferencias entre los años, la información contenida en este módulo comprende distintas variables de uso de las TIC, predominando aquellas que reflejan el acceso a las tecnologías o a los servicios asociados por parte de las empresas, y no tanto variables referidas a la intensidad de su uso (que podrían ser relevantes para analizar el impacto de estas tecnologías o servicios sobre la productividad empresarial).

Las variables específicas de uso de las TIC e inversión en I&D cubren los siguientes temas:

1. Infraestructura Informática
 - Uso de computadoras
 - Uso de PDA o Tablet
 - Internet
 - o Acceso a Internet
 - o Usos del internet
 - o Anuncios por internet.
2. Infraestructura de Comunicaciones.
 - Uso de telefonía
 - Red de área local
 - Intranet y Extranet
 - Página o sitio web
 - Uso de software: comprado, libre, desarrollado.
 - Sistemas de gestión
3. Comercio Electrónico

- Compras a través de Internet
- Ventas a través de Internet
- 4. Redes sociales
 - Uso de redes sociales
 - Utilidad de las redes sociales para las empresas
- 5. Capacitación al personal en el uso de las TIC
- 6. Indicadores de Ciencia y Tecnología
 - Empresas que invierten en ciencia y tecnología
 - Inversión en ciencia y tecnología:
 - i. Patentes y propiedad industrial
 - ii. Programas de computadoras y software
 - iii. Costos de exploración y desarrollo
 - iv. Fórmulas, diseños y prototipos
 - v. Reservas de recursos extraíbles
 - vi. Otros tipos de intangibles
 - vii. Capacitación en investigación científica, tecnológica y de gestión
 - viii. Asesoría y consultoría

A partir del análisis de la información, se ha seleccionado un conjunto de variables relevantes, cuyas estadísticas descriptivas para el panel de la EEA 2016-2018 se muestran en el siguiente cuadro.

Cuadro 7: Estadísticas descriptivas de las variables de uso de las TIC y de inversión en I&D – Manufactura panel 2016-18

Variable	Tipo	Observaciones	Media	Desv. Estándar
Porcentaje de trabajadores que usa PC	Porcentaje	2330	51.280	31.197
Porcentaje de trabajadores que usa Internet	Porcentaje	2330	50.505	31.591
Uso de intranet	Dummy	2330	0.490	0.500
Uso de extranet	Dummy	2330	0.219	0.414
Uso de red local	Dummy	2330	0.924	0.266
Uso de banda ancha 1/	Dummy	2330	0.875	0.331
Uso de sistemas de gestión	Dummy	2330	0.960	0.196
Uso de PDA/Tablet	Dummy	2330	0.309	0.462
Uso de página Web	Dummy	2330	0.799	0.401
Pago de anuncios por Internet	Dummy	2330	0.300	0.458
Compra software 2/	Dummy	2330	0.779	0.415
Software libre 2/	Dummy	2330	0.415	0.493
Desarrolla software 2/	Dummy	2330	0.354	0.478
Capacitación en TIC	Dummy	2330	0.375	0.484
Porcentaje de compras por Internet	Porcentaje	2330	3.247	11.045
Porcentaje de ventas por Internet	Porcentaje	2330	1.989	11.047
Inversión en nuevos productos	Dummy	2330	0.274	0.446
Inversión en mejora procesos	Dummy	2330	0.331	0.471

Personal permanente en I&D	Dummy	2330	0.267	0.442
Unidad permanente en I&D	Dummy	2330	0.206	0.405
Inversión en I&D	Dummy	2330	0.462	0.499
Monto de inversión en I&D (en log.)	Monto en S/	1076	11.895	2.479
Intensidad de inversión en I&D (respecto a salarios)	Proporción	1074	0.121	0.279

Notas:

1. Se considera el uso exclusivo de banda ancha (no utilizar también conexiones de menor velocidad).
2. El software comprende software contable, de ventas, de personal, soporte, entre otros.

Fuente: Panel EEA 2016-18.

En el Anexo 6 se muestran los valores promedio de las variables de uso de las TIC e inversión en I&D según tamaño de las empresas y subsector productivo de la manufactura.

El uso de las TIC muestra una tendencia a crecer con el tamaño de las empresas, excepto en el caso del uso de banda ancha. Se identifica un fuerte incremento en el uso de las TIC entre las empresas del segmento corporativo respecto a las demás. En cuanto a los indicadores de inversión en I&D se encuentra un incremento significativo de las actividades orientadas al desarrollo de nuevos productos y a la mejora de procesos, así como de la intensidad de (respecto a salarios) al aumentar el tamaño de las empresas. Sin embargo, respecto a la existencia de personal o unidad permanente en I&D, así como a la inversión en I&D en sí misma, no hay mucha diferencia entre las empresas grandes y las corporativas, aunque su nivel es significativamente mayor al de las empresas medianas.

En cuanto a las diferencias por subsector, en general, el uso de las TIC es mayor en el caso de las manufacturas intensivas en tecnología, las cuales también muestran mayores indicadores de inversión en I&D. Cabe destacar que el porcentaje de trabajadores que usa PC, el uso de intranet, de extranet y de PDA/Table resultan sustantivamente menores para las manufacturas ligeras, mientras que la capacitación en TIC es mayor en este subsector. Respecto a las variables de inversión en I&D, destacan los subsectores de Intensivas en tecnología y de producción de alimentos, bebidas y tabaco, específicamente en cuanto a presentar una mayor realización de actividades de inversión en nuevos productos o en mejora procesos, una mayor proporción que tiene personal o unidad permanente en I&D, así una mayor Intensidad de inversión en I&D (respecto a salarios).

4.4 Estimación del Índice sintético de uso de las TIC

A partir de las variables de uso de las TIC identificadas se ha realizado la construcción de un índice sintético mediante análisis factorial (AF) que capture su información reduciendo la cantidad de variables de análisis, dada la significativa correlación entre estas variables, de forma similar a Colombo *et al.* (2013).

En este caso, una opción, dada la posible endogeneidad del uso de la TIC por las empresas, era instrumentalizar las distintas variables. Sin embargo, ello, no se ha realizado ante la dificultad de encontrar instrumentos adecuados. Aunque en la literatura se ha utilizado como instrumentos para variables relativas al uso (acceso) a servicios como Internet brindados a través de infraestructura de redes, las variables de penetración de estos servicios según zonas o regiones, ya que su despliegue sería

plausiblemente exógeno (Crespi *et al.*, 2006; Akerman *et al.*, 2015), en este caso, la elevada concentración de empresas con domicilio fiscal en Lima Metropolitana en los datos disponibles implica una escasa variación de este tipo de variables. Ello hace que su correlación con las variables de acceso a nivel de empresa sea baja, y que se constituyan en instrumentos débiles, cuya utilización podría ser incluso contraproducente de acuerdo con Bound *et al.* (1995)²⁰.

Para el análisis de las TIC, se ha considerado preferible el AF al análisis de componentes principales (ACP), dado que el AF se concentra en explicar la varianza común de las variables analizadas en vez de la varianza total.

Sobre este método, cabe mencionar que un factor es una variable no observable directamente que es inferida a partir de un conjunto de variables observables, y que representa un elemento o dimensión subyacente de la información de estas variables.

En general, en un conjunto de datos pueden existir “k” variables observables que dependen de “r” factores no observables tal que $k > r$, los cuales son independientes entre sí (es decir, no están correlacionados).

Algebraicamente, el conjunto de factores F se relaciona con las variables observables X a través de la matriz de parámetros A (matriz de cargas factoriales) quedando una variable residual U.

$$X_{n \times k} = F_{n \times r} * A_{r \times k} + U_{n \times k}$$

De esta manera, el análisis de factores busca encontrar un conjunto reducido de factores no observables “r” que reconstruyan a las “k” variables originales. El cálculo de los factores es definido en términos de la predicción de la matriz de varianza-covarianza común de las variables observables. Utilizando los valores propios de la matriz de momentos cruzados de las variables observables, se estima el porcentaje de la varianza explicada por cada factor.

A partir de la matriz F se generan las cargas factoriales que vienen a ser coeficientes de correlación entre las variables y los factores encontrados. Para el cálculo de estas cargas, es posible utilizar distintas rotaciones de las matrices A y F, que permitan facilitar la interpretación de las mismas.

Además, para cada factor estimado es posible obtener un valor propio o autovalor. Como criterio estándar en la literatura se considera que un factor resulta explicativo de los datos, si su valor propio es positivo.

De este modo, al realizar la estimación del AF para las variables TIC de las empresas de manufactura del panel 16-18, se encuentra que utilizar un solo índice sintético para todas las variables implicaría pérdida de información, pues hasta 8 de ellos tienen valor propio positivo.

²⁰ Bound *et al.* (1995) destacan que el uso de instrumentos que explican poco la variación de las variables explicativas endógenas puede conducir a grandes inconsistencias en las estimaciones, incluso si solo existe una relación débil entre los instrumentos y el error en la ecuación estructural (que se busca corregir). También indican que, en muestras finitas, las estimaciones mediante instrumentos están sesgadas en la misma dirección que las de mínimos cuadrados ordinarios (OLS), y que el sesgo de ambas se aproxima cuando el estadístico R² entre los instrumentos y la variable explicativa endógena es cercano a 0.

Cuadro 8: Correlaciones entre variables TIC, Manufactura Panel 2016-18

Variable	Descripción	pc	web	itn	xtn	red	ban	sgs	tab	pag	pin	sfp	sfl	sfd	tic	cmp	vnt	cmp_d	vnp_d
pc	Porcentaje uso de PC	1																	
web	Porcentaje uso de Internet	0,7788*	1																
itn	Uso de intranet	0,1771*	0,1254*	1															
xtn	Uso de extranet	0,1227*	0,0958*	0,4643*	1														
red	Uso de red local	0,1512*	0,1197*	0,1783*	0,1290*	1													
ban	Uso de banda ancha (Modem)	0,0301	0,0181	0,0429*	0,0151	0,0722*	1												
sgs	Uso de sistemas de gestión	0,0990*	0,0911*	0,1164*	0,0922*	0,1807*	0,1020*	1											
tab	Uso de PDA/Table	0,1774*	0,1346*	0,2423*	0,2561*	0,1469*	0,0026	0,1031*	1										
pag	Uso de página Web	0,1538*	0,1243*	0,2006*	0,1419*	0,1377*	0,0726*	0,1601*	0,1782*	1									
pin	Pago por anuncio por Internet	0,0716*	0,0625*	0,1259*	0,1221*	0,0611*	0,0373	0,0711*	0,1040*	0,2395*	1								
sfp	Compra software	0,0915*	0,0600*	0,1423*	0,0669*	0,1665*	0,0839*	0,2193*	0,0825*	0,1408*	0,0632*	1							
sfl	Software libre	0,0662*	0,0387	0,0940*	0,0478*	0,1015*	0,0498*	0,1185*	0,0789*	0,1647*	0,0647*	0,2280*	1						
sfd	Desarrolla software	0,0614*	0,0151	0,1759*	0,1194*	0,1048*	0,03	0,1051*	0,1943*	0,1881*	0,033	0,0909*	0,2846*	1					
tic	Capacitación en TIC	0,1056*	0,0761*	0,2465*	0,2171*	0,0993*	0,0058	0,1308*	0,2416*	0,1855*	0,1261*	0,1599*	0,1239*	0,1511*	1				
cmp	Porcentaje de compras por Internet	0,0535*	0,0627*	0,0063	0,0276	0,0189	0,0045	0,028	0,0391	0,0802*	0,1214*	0,0434*	0,0832*	0,0153	0,0377	1			
vnt	Porcentaje de ventas por Internet	0,0531*	0,0509*	0,0104	0,0408*	0,0051	-0,0191	0,0343	0,0144	0,0753*	0,1193*	0,009	0,0172	0,016	0,0357	0,3791*	1		
cmp_d	Uso de Internet en compras	0,0846*	0,0977*	0,1487*	0,1004*	0,0616*	0,0195	0,0667*	0,1818*	0,1822*	0,1986*	0,1075*	0,1601*	0,1415*	0,1376*	0,4624*	0,1688*	1	
Vnt_d	Uso de Internet en ventas	0,0538*	0,037	0,0539*	0,0236	-0,0104	-0,0193	0,0466*	0,0549*	0,1075*	0,1957*	0,0256	0,0602*	0,0563*	0,0836*	0,2808*	0,5895*	0,2736*	1

* Significativa al 5% de probabilidad de error.

Fuente: Panel EEA 2016-18.

En este caso se ha optado por realizar una rotación de los factores mediante una matriz oblicua (*oblimin*) para facilitar la interpretación de cada factor. Ello, a diferencia de la rotación ortogonal que es más estándar (especialmente la *varimax*), permite que los factores tengan una correlación entre ellos, lo cual resulta natural dado el tipo de dimensiones del uso de la TIC (que, incluso, pueden resultar complementarias para las empresas).

Con ello se identifica 8 factores con valor propio positivo como se muestra en la tabla. Entre ellos destaca el primer factor con un poder explicativo que llega al 65% de la varianza.

Dados estos resultados, se ha decidido seleccionar los 7 primeros factores debido a que el octavo factor no aporta mayor explicación de la varianza conjunta de las variables y, además, como se verá a continuación, no muestra una relación interpretable con variables específicas dentro del conjunto de variables analizadas.

Cuadro 9: Análisis factorial para variables de uso de las TIC

Factor analysis/correlation	Number of obs =	2,330
Method: principal factors	Retained factors =	8
Rotation: (unrotated)	Number of params =	100

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	2.10288	1.04378	0.6509	0.6509
Factor2	1.05909	0.47975	0.3278	0.9787
Factor3	0.57934	0.12969	0.1793	1.1580
Factor4	0.44965	0.26449	0.1392	1.2972
Factor5	0.18515	0.07729	0.0573	1.3545
Factor6	0.10786	0.07692	0.0334	1.3879
Factor7	0.03095	0.02823	0.0096	1.3975
Factor8	0.00271	0.05657	0.0008	1.3983
Factor9	-0.05386	0.03946	-0.0167	1.3816
Factor10	-0.09332	0.04185	-0.0289	1.3527
Factor11	-0.13517	0.01735	-0.0418	1.3109
Factor12	-0.15252	0.01632	-0.0472	1.2637
Factor13	-0.16884	0.03475	-0.0523	1.2114
Factor14	-0.20360	0.02151	-0.0630	1.1484
Factor15	-0.22511	0.02929	-0.0697	1.0787
Factor16	-0.25439	.	-0.0787	1.0000

LR test: independent vs. saturated: chi2(120) = 5264.14 Prob>chi2 = 0.0000

Fuente: Elaboración propia en base al panel EEA 2016-18.

Los resultados de la matriz de cargas factoriales indican el tipo de información que captura cada factor, que usualmente corresponde a variables que se encuentran correlacionadas entre sí. Estos valores se muestran tanto para la rotación *varimax* como la *oblimin*²¹.

De este modo, se identifica una relación entre cada uno de los 7 factores seleccionados y determinadas variables de las 16 variables originales. Esta asociación se resume en lo siguiente:

²¹ La construcción de los factores se ha basado en la rotación ortogonal *varimax* debido a que, posteriormente, se incluyen en la regresión de determinantes de la PTF, es conveniente que los factores calculados sean independientes entre sí, a fin de evitar problemas de multicolinealidad (si hubiera alta correlación entre ellos).

- Factor 1: porcentaje de trabajadores que usan PC y porcentaje de trabajadores que usan Internet.
- Factor 2: uso intranet y extranet.
- Factor 3: compra y venta de productos a través de Internet.
- Factor 4: software libre o desarrollado por las empresas.
- Factor 5: red de área local, software propietario o comprado, y uso de algún sistema de gestión informatizado.
- Factor 6: página web y pago de anuncios por internet.
- Factor 7: uso de Tablet/pda y, en menor medida, a capacitación en TIC a los trabajadores.
- Factor 8: no muestra una relación específica con las variables analizadas²².

Cuadro 10: Matriz de cargas factoriales y unicidad de variables TIC, Manufactura Panel 2016-18

A. Rotación ortogonal – varimax

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Uniqueness
pc	0,834	0,0858	0,0134	0,041	0,0274	0,015	0,0039	0,0022	0,2942
web	0,8303	0,029	0,0239	-0,0093	0,01	0,0092	0,0006	-0,0012	0,3089
itn	0,1338	0,594	-0,0098	0,0857	0,0814	0,0361	-0,021	-0,0024	0,6134
xtn	0,0923	0,5855	0,0314	0,012	0,0109	0,0156	-0,0033	0,0011	0,6472
red	0,1406	0,1966	-0,0095	0,1086	0,2694	0,0306	-0,0072	0,0288	0,8553
ban	0,0232	0,0145	-0,0226	0,0335	0,2006	0,0766	-0,1002	0,0153	0,9412
sgs	0,0943	0,1272	0,0265	0,136	0,3468	0,0786	0,0091	0,0119	0,8291
tab	0,1531	0,3787	0,0134	0,1669	0,0154	0,0821	0,1215	0,0246	0,7828
pag	0,1343	0,2266	0,1111	0,2257	0,1412	0,2934	0,0088	0,0072	0,7612
pin	0,061	0,1704	0,1945	0,0356	0,0586	0,2872	0,0041	-0,0061	0,8422
sfp	0,0707	0,111	0,0246	0,2186	0,3645	0,0376	0,0098	-0,0199	0,7995
sfl	0,0415	0,0636	0,0606	0,4352	0,1715	0,0365	-0,0114	-0,0128	0,7702
sfd	0,0272	0,201	0,0007	0,4261	0,0378	0,0372	0,0215	0,0164	0,7738
tic	0,0785	0,3461	0,042	0,177	0,0994	0,112	0,1287	-0,013	0,8018
cmp	0,0559	0,0055	0,5149	0,0473	0,0235	0,029	0,0006	-0,0014	0,7281
vnt	0,0491	0,0194	0,5125	-0,0006	-0,005	0,0301	0,0034	0,0011	0,7336

B. Rotación oblicua – oblimin

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Uniqueness
pc	0,8303	0,0178	-0,005	0,0214	0,0055	0,0008	0,0042	0,0031	0,2942
web	0,8392	-0,0174	0,0063	-0,0227	-0,0063	-0,0004	-0,004	-0,0029	0,3089
itn	0,0191	0,5912	-0,025	0,0323	0,0295	0,0161	-0,0104	0	0,6134
xtn	-0,0148	0,5952	0,0288	-0,0275	-0,0216	-0,0075	0,0201	0,0106	0,6472
red	0,0526	0,0948	-0,003	-0,0113	0,3234	-0,0168	-0,0029	0,039	0,8553
ban	-0,0171	0,0463	-0,0524	-0,013	0,164	0,0932	-0,161	-0,0004	0,9412
sgs	0,0003	-0,0161	0,0084	-0,0298	0,4091	0,0449	0,0134	0,0146	0,8291
tab	0,0683	0,1298	-0,0057	0,0735	0,0513	0,0673	0,2136	0,0673	0,7828
pag	0,0407	0,0018	-0,0168	0,1092	0,0464	0,3814	0,0161	0,0157	0,7612
pin	-0,0168	0,024	0,0656	-0,0676	-0,0239	0,3919	0,0044	-0,0026	0,8422
sfp	-0,0098	0,0043	0,0114	0,0712	0,403	-0,0088	0,007	-0,0354	0,7995
sfl	0,007	-0,036	0,0399	0,4237	0,0896	0,016	-0,0226	-0,028	0,7702

²² En la rotación *varimax* (aunque en la rotación *oblimin* se captura una ligera relación con tener red de área local y usar tablets).

sfd	-0,0123	0,0539	-0,0107	0,4448	-0,0429	0,0216	0,0462	0,0278	0,7738
tic	-0,0126	0,1044	-0,0021	0,0411	0,1316	0,1091	0,2125	0,0104	0,8018
cmp	0,006	-0,0083	0,5161	0,0286	0,0138	0,0016	-0,0037	-0,0031	0,7281
vnt	-0,0001	0,0134	0,5141	-0,0172	-0,0091	0,0072	0,0026	0,0026	0,7336

Nota: La tonalidad sombreado verde más fuerte indica mayores valores de las cargas factoriales.

Fuente: Elaboración propia en base al panel EEA 2016-18.

En cuanto a la validez estadística del AF realizado, la prueba o test de esfericidad de Bartlett indica que se rechaza la hipótesis nula de que las variables no están correlacionadas con una probabilidad de error de 0²³.

Cuadro 11: Prueba estadística de Bartlett

Determinant of the correlation matrix

Det = 0.106

Bartlett test of sphericity

Chi-square = 5211.544

Degrees of freedom = 105

p-value = 0.000

H0: variables are not intercorrelated

Fuente: Elaboración propia en base al panel EEA 2016-18.

Por último, la medida de ajuste de Kaiser-Meyer-Olkin (KMO) sobre adecuación de los resultados de la muestra de los datos utilizados, resulta en un valor de 0.68, lo cual se considera adecuado.

Cuadro 12: Medida de Kaiser-Meyer-Olkin de ajuste

Variable	kmo
pda_tab_	0.8300
red_loc_	0.8487
int_net_	0.7331
ext_net_	0.7030
uso_pc_	0.5633
uso_int_	0.5443
pag_web_	0.8208
pct_cmp_	0.5557
pct_vnt_	0.5456
ban_anc_	0.7144
pag_int_	0.7582
sft_prp_	0.7435
sft_lib_	0.6907
sft_dev_	0.7186
cap_tic_	0.8399
sis_gst_	0.7977
Overall	0.6743

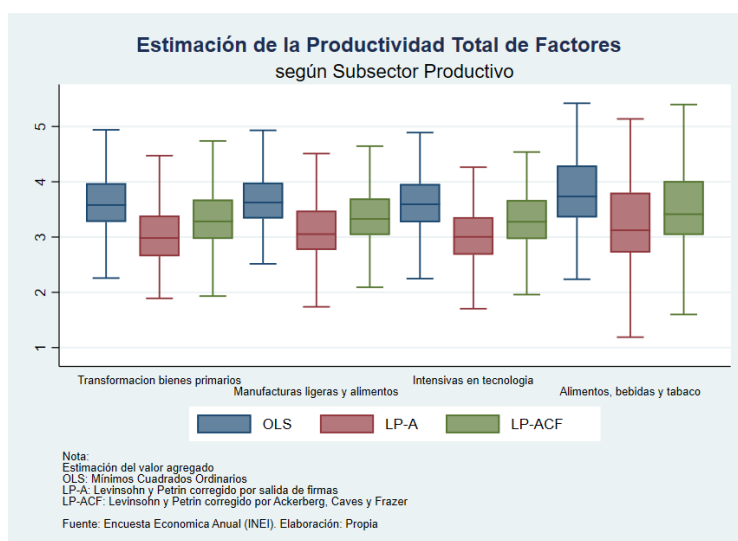
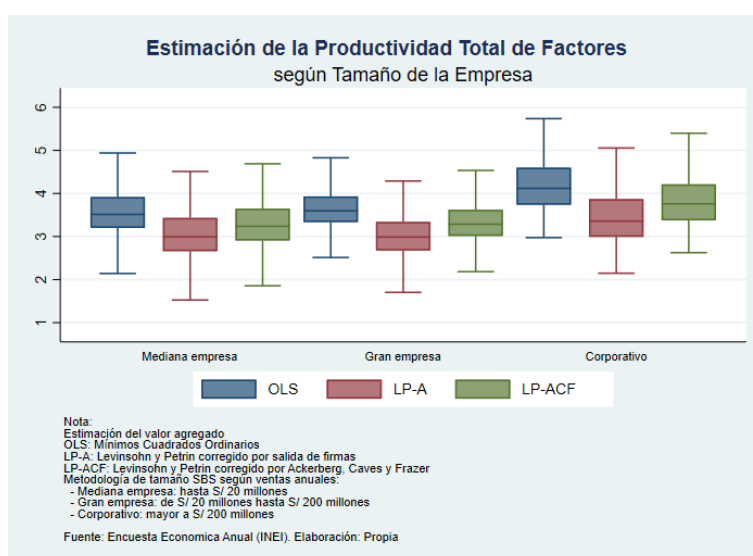
Fuente: Elaboración propia en base al panel EEA 2016-18.

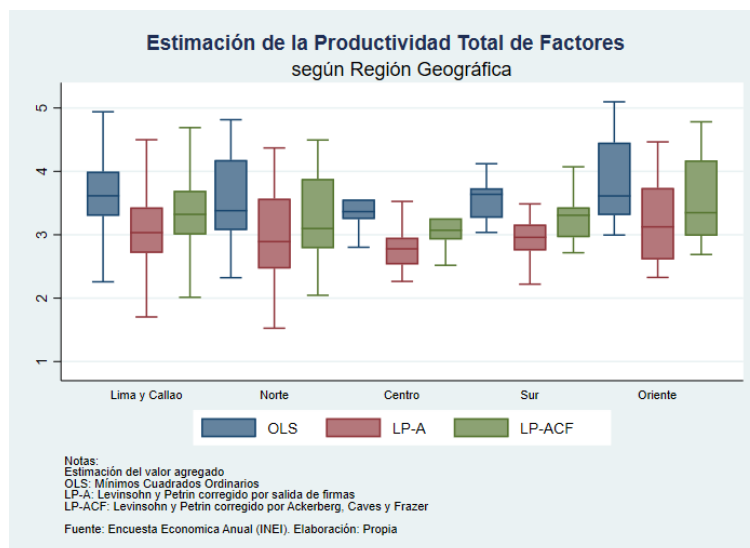
²³ Sin embargo, se encuentran valores elevados del indicador de unicidad (proporción de la varianza de cada variable que no es explicada por los factores comunes) para las variables vinculadas a banda ancha, uso de software, y sistema de gestión, por lo cual el AF podría no ser el método más adecuado para analizar en conjunto la información sobre las TIC (al no ser tan fuerte el grado de correlación entre las variables).

5 Estimación de los determinantes de la PTF

En primer lugar, aunque, como se ha mencionado, se ha optado la estimación de la PTF mediante el método LP-A por utiliza el método de AI analizar los resultados de la PTF estimada con método OLS (utilizado como referencia) y los métodos LP-ACF y LP-A para las principales características de las empresas se observan patrones similares. Al analizar según el tamaño de las empresas se encuentra que, en promedio, la PTF es mayor para las empresas corporativas, sin importar el método utilizado. En el caso del subsector, la PT no se observa diferencias sustantivas en la PTF independientemente del método. A nivel de regiones geográfica, las regiones centro y sur tienen los niveles de PTF más bajos y también presentan menor dispersión en las estimaciones de los distintos métodos.

Gráfico 2: Productividad Total de Factores (PTF) según segmentos de empresas





Fuente: Elaboración propia en base al panel EEA 2016-18.

5.1 Estimación del monto de la inversión en I&D

Dadas la posible endogeneidad y sesgo de selección en la variable relativa al monto registrado de inversión en I&D, en línea como el modelo CDM, se estiman ecuaciones previas para predecir la inversión en I&D respecto a variables identificadas en la literatura.

Específicamente, se realiza la estimación del logaritmo de la inversión en I&D mediante el modelo de Heckman que considera una ecuación lineal de determinantes del nivel de esta variable y un modelo probit de elección binaria para la ecuación de selección que determina si aparece o no un valor de inversión en I&D. Se utiliza la estimación en dos etapas (modelo denominado *heckit*) y la estimación simultánea²⁴.

Siguiendo a Arbeláez (2011), Gallego *et al.* (2011), Martin & Nguyen-Thi (2015), Crespi, Tacsir y Vargas (2016), Ramírez *et al.* (2019), y Tello (2020), entre las variables consideradas en la estimación de ambas ecuaciones se consideran el tamaño de empresa (medido por el monto de remuneraciones), si hay participación extranjera en la propiedad de la empresa, la antigüedad de la empresa, la participación en el mercado (competencia), el tener capital intangible (incluyendo patentes), y el subsector de manufactura²⁵. Como las variables incluidas en las ecuaciones no son las mismas se cumple la restricción de exclusión que hace posible la identificación de los parámetros. Específicamente, el tamaño de empresa en nivel se incluye solo en la primera ecuación²⁶ y las variables *dummy* de subsector y de propiedad extranjera solo en la segunda ecuación.

²⁴ En este caso, si existe sesgo de selección resultarán estadísticamente distintos de 0 el coeficiente estimado para el inverso de la ratio de Mills de la segunda ecuación en el primer modelo, o el estimador del coeficiente de correlación entre los errores de ambas ecuaciones en el segundo modelo.

²⁵ Otras características de las empresas como tener equipamiento o personal permanente para I&D, y el tipo de régimen jurídico (persona natural, sociedad anónima) se excluyeron del análisis por su alta correlación con las demás variables.

²⁶ Se incluye las categorías de tamaño basadas en ventas en la segunda ecuación.

En el caso de la participación en el mercado, dado que no se cuenta con datos del nivel de ventas para actividades económicas desagregadas, se ha utilizado una aproximación a partir del valor agregado. Para ello, se ha dividido el valor agregado de cada empresa entre el dato de la producción total a nivel de Clase CIIU proveniente de los compendios estadísticos del INEI²⁷.

La estimación indica que sí existe sesgo de selección en los datos observados de inversión en I&D, pues el coeficiente de correlación entre los errores de las ecuaciones de selección y del nivel de la inversión en I&D es estadísticamente distinto de 0, al igual que el coeficiente de la ratio de Mills. Ello implica que es necesario hacer la corrección de la estimación de sus determinantes.

Los resultados del modelo de corrección de sesgo de selección de Heckman en una etapa o simultaneo con el modelo de dos etapas (*heckkit*) son similares entre sí para la mayoría de las variables. En este caso, se prefiere utilizar el modelo en una etapa dado que otorga mayor eficiencia en la estimación de los parámetros.

En general, los resultados son consistentes con lo hallado en estudios previos como Tello (2020). En primer lugar, si crece el tamaño de la empresa (medido por el monto de remuneraciones) aumenta la probabilidad de realizar inversión en I&D de forma estadísticamente significativa. Asimismo, pertenecer a segmentos de mayor tamaño incrementa la inversión en I&D, y un incremento en la antigüedad de la empresa implica menor propensión a innovar, aunque se haría con mayor inversión.

Por otro lado, una mayor participación en el mercado incrementa de forma estadísticamente significativa la probabilidad de realizar inversión en I&D, aunque su efecto positivo sobre el nivel de inversión en I&D no resulta estadísticamente significativo.

Además, las empresas que poseen capital intangible previo también invierten un mayor monto en I&D. Sin embargo, al contrario de lo esperado, las empresas con capital extranjero tienden menos a invertir en I&D, lo cual podría deberse a que la inversión en I&D la realizan las empresas “matriz”. Por último, las empresas manufactureras que realizan actividades clasificadas como intensivas en tecnología realizan una mayor inversión en I&D.

Cuadro 13: Estimación de la inversión en I&D - Manufactura panel 2016-18

VARIABLES	Modelo de Heckman simultaneo			Modelo de Heckman en dos Etapas		
	Inversión en I&D	Decisión de inversión	Estadísticos	Inversión en I&D	Decisión de inversión	Estadísticos
Gasto en remuneraciones (en log.)		0.212*** (0.0217)			0.259*** (0.0283)	
Posee capital intangible	0.523*** (0.179)			0.588*** (0.194)		
Participación de mercado	0.705 (0.697)	1.567** (0.774)		1.132 (0.877)	1.223* (0.649)	
Edad	0.00987** (0.00456)	-0.00490*** (0.00143)		0.00951** (0.00442)	-0.00557*** (0.00169)	

²⁷https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1690/COMPENDIO2019.html

Gran empresa (SBS)	0.380*** (0.147)			0.477** (0.215)		
Corporativo (SBS)	1.372*** (0.243)			1.676*** (0.354)		
Manufacturas ligeras	-0.224 (0.138)			-0.220 (0.186)		
Intensivas tecnologías	0.285** (0.145)			0.313** (0.148)		
Alimentos, bebidas y tabaco	0.148 (0.303)			-0.174 (0.330)		
Propiedad extranjera		-0.125* (0.0693)			-0.174 (0.115)	
athrho			-1.931*** (0.139)			
Insigma			1.189*** (0.0420)			
lambda						-2.524*** (0.504)
Constant	13.27*** (0.265)	-3.232*** (0.322)		12.59*** (0.550)	-3.915*** (0.426)	
Observations	2,240	2,240	2,240	2,240	2,240	2,240

Notas:

Errores estándar robustos entre paréntesis. *** p<0.01, ** p<0.05, * p<0.1

IDE: Inversión en I y D en Logaritmo

DID: Decisión de inversión en I y D en Dummy

5.2 Distribución de la PTF de las empresas de manufactura según distintas variables

En esta sección se utilizan los resultados estimados de la PTF a nivel de empresas para el panel 2016-18 utilizando el método de Levinsohn y Petrin que corrige por salida de empresas a fin de realizar un análisis de tipo exploratorio de si existen patrones de heterogeneidad de la PTF según distintas variables relativas a característica de las empresas como su tamaño, subsector, antigüedad, acceso al crédito, entre otros.

Asimismo, se analiza si la PTF es distinta según las variables relativas al acceso y uso de las TIC en las empresas. Específicamente se analiza las siguientes variables:

- Tamaño de empresa: de acuerdo con la clasificación de la Superintendencia de Banca y Seguros (SBS)
- Región geográfica
- Porcentaje de mujeres en puestos de gerencia (directivos)
- Antigüedad de la empresa
- Porcentaje uso de PC
- Porcentaje uso de Internet
- Uso de Intranet
- Uso de Extranet
- Uso de Red Local
- Uso de banda ancha (Modem)
- Uso de Sistemas de Gestión
- Uso de PDA/Tablet
- Uso de página Web
- Pago de anuncios por Internet
- Compra software
- Software libre
- Desarrolla software

- Capacitación en TIC
- Porcentaje de compras por Internet
- Porcentaje de ventas por Internet
- Si la empresa realizó inversiones en nuevos productos (I&D para desarrollo de ciencia y tecnología)
- Si la empresa realizó inversiones en mejora de procesos (I&D para desarrollo de ciencia y tecnología)

El análisis empírico se basa tanto en el nivel de la PTF (en logaritmos) del año como en su variación porcentual entre los años 2016 y 2018 para cada empresa (aproximada como a diferencia de sus valores en logaritmos en 2018 y en 2016), respecto a las características de la empresa en el año 2016.²⁸

Al analizar la distribución del valor de la PTF en el año 2016 se identifica que las empresas corporativas tienen una productividad bastante mayor a las medianas y grandes empresas (no corporativas). Asimismo, en Lima Metropolitana suele ser mayor el nivel de productividad.

En cuanto a la antigüedad de la empresa, no hay una relación clara probablemente porque la mayoría son empresas ya consolidadas (tomando en cuenta que la EEA no incluye micro y pequeñas empresas). Solo destacan con mayor productividad aquellas que tienen menos de 14 años de antigüedad. Además, el tener más de 50% de participación femenina en cargos directivos también se asocia a una mayor PTF en promedio.

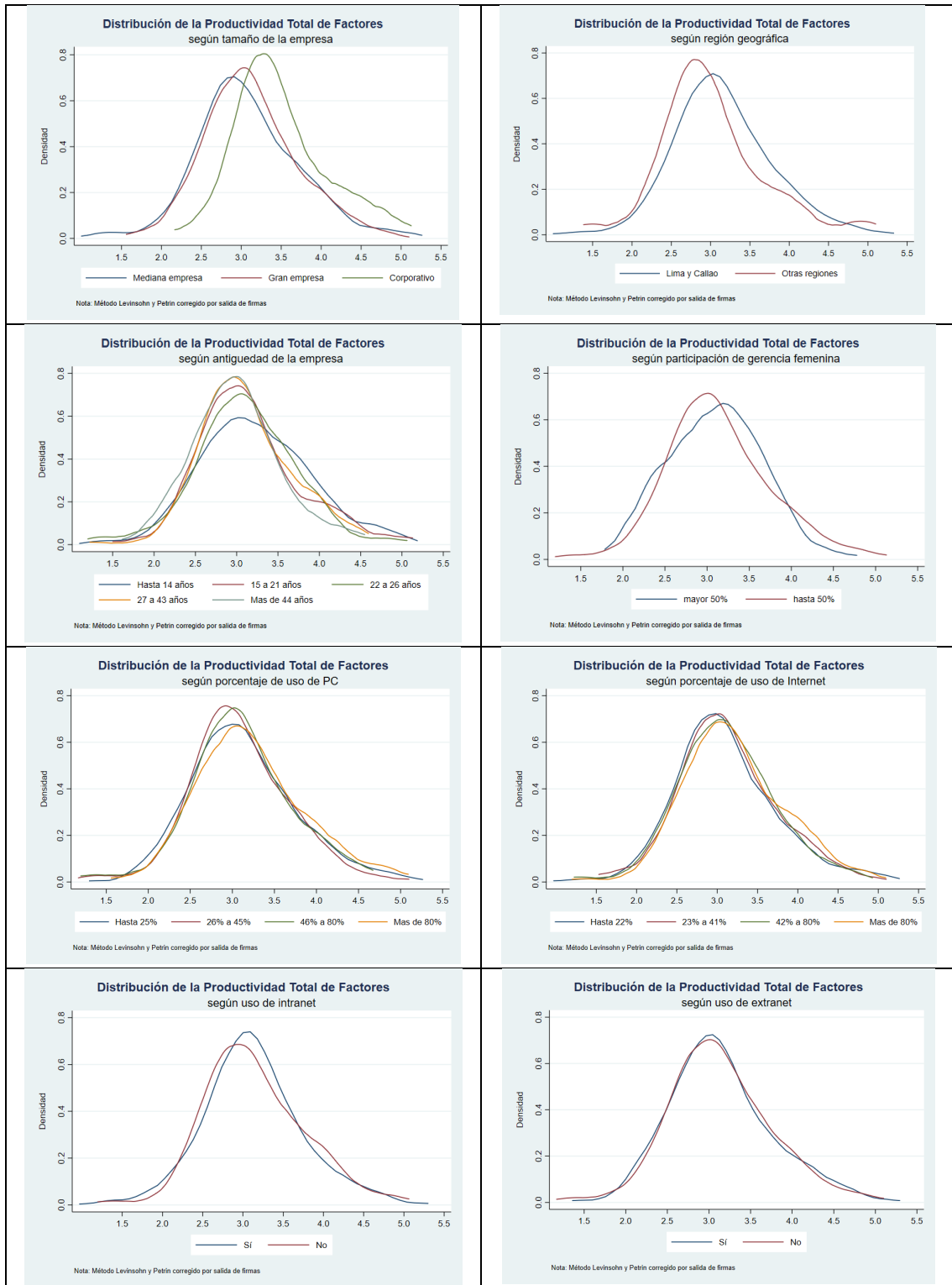
Respecto al uso de las TIC, el porcentaje de trabajadores que usan PC y el uso de Internet no se asocia a cambios significativos en la productividad, excepto ligeramente para las empresas de más de 80% de los trabajadores. Asimismo, el uso de red de área local, de intranet y de algún sistema de gestión informatizado en la empresa se vinculan con una mayor productividad. De igual forma, el uso de extranet y de una conexión de banda ancha no se asocia a un mayor nivel de productividad. Tampoco el uso de Tablet/PDA.

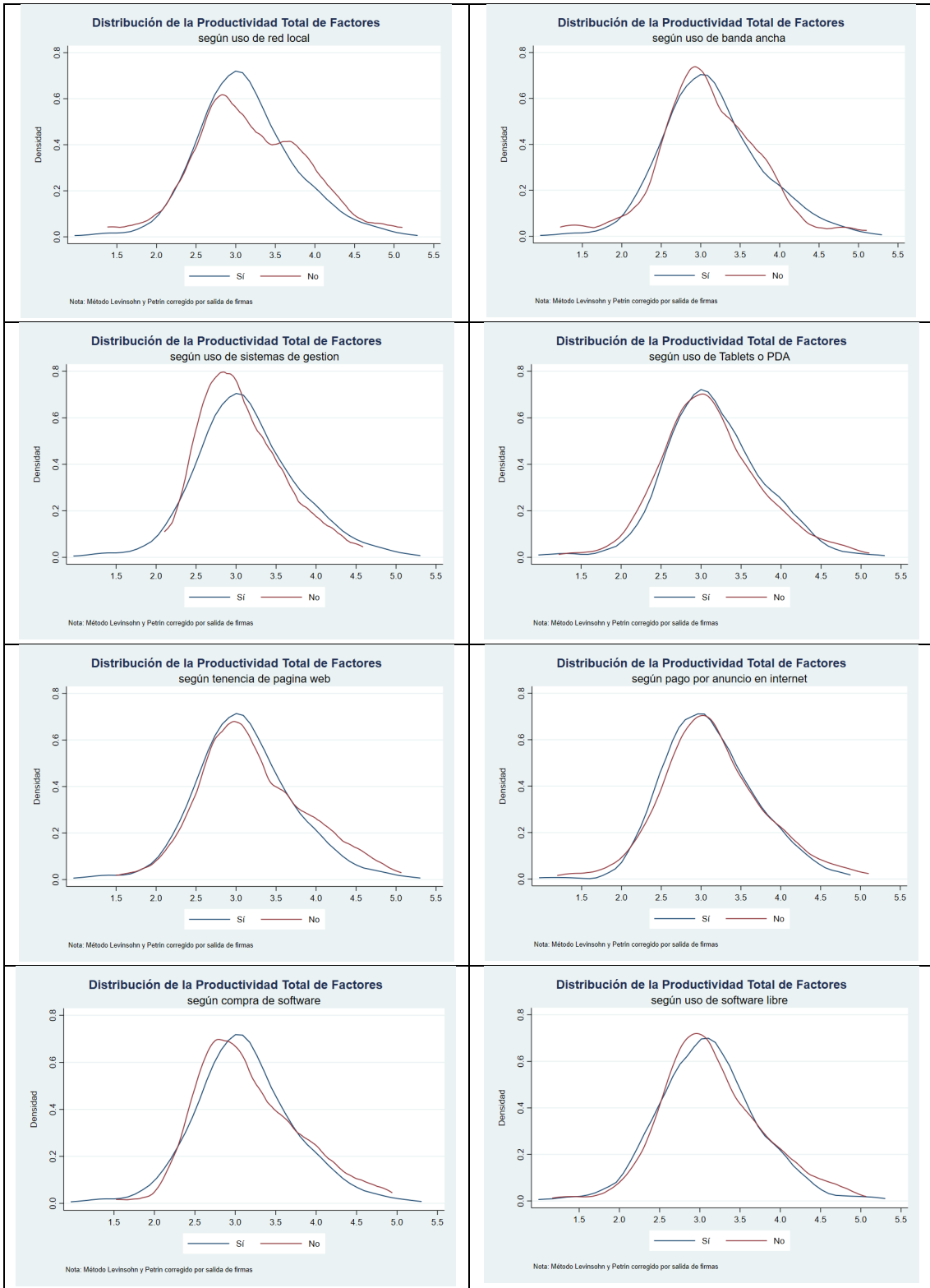
En el caso de uso de página web y uso de anuncios pagados por Internet, donde la relación incluso aparece como negativa. Además, el uso de distintos tipos de software aparece ligeramente asociado a una mayor productividad, pero no la capacitación en TIC al personal. El uso de Internet para la compra de productos se asocia a una menor productividad, y lo contrario sucede con la venta por Internet que se correlaciona positivamente con el nivel de PTF.

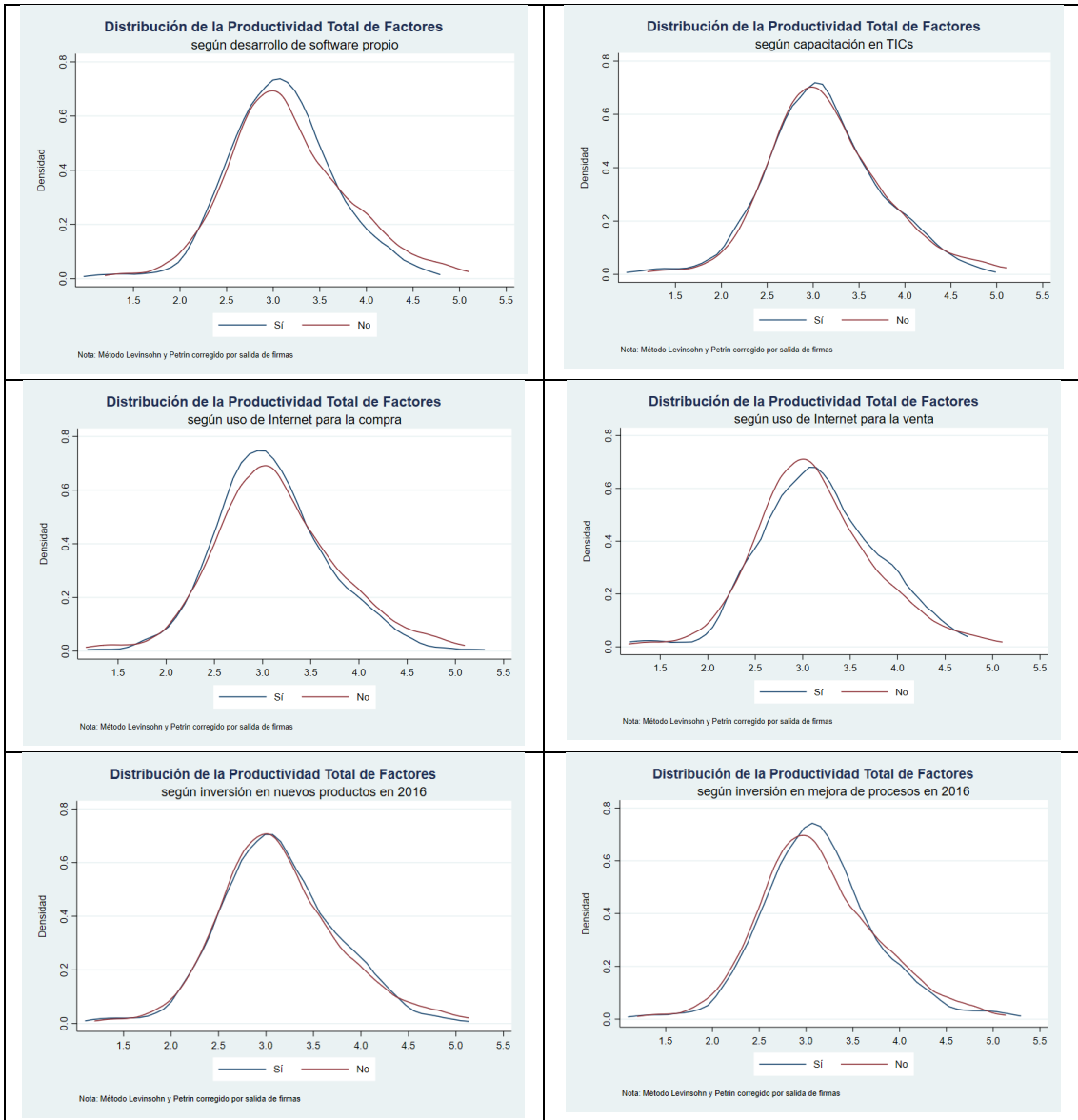
Por último, la inversión en mejora de procesos se asocia a una mayor PTF, debido - posiblemente - a que constituye una práctica más frecuente en las empresas más productivas. No ocurre lo mismo con la inversión en productos nuevos o mejorados significativamente.

²⁸ En el Anexo 5, se incluyen para referencias los resultados de la prueba estadística T-student para la diferencia del nivel de la PTF en el año 2016 según valores de las variables seleccionadas.

Gráfico 3: Distribución del valor de la PTF según características de las empresas del sector manufactura de 2016







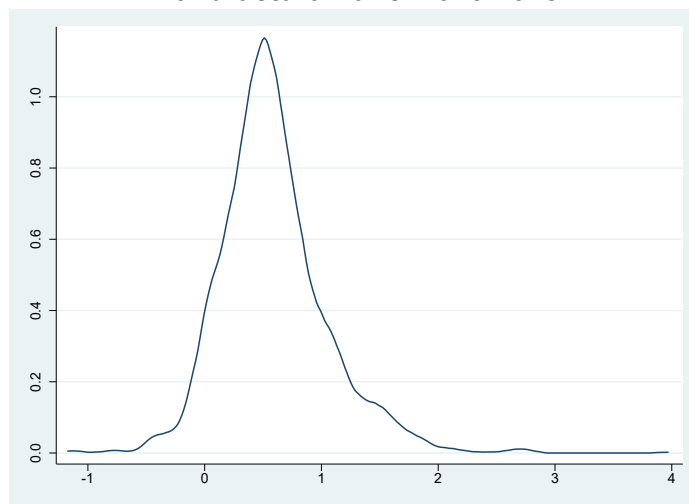
Fuente: Elaboración propia en base al panel EEA 2016-18.

Estimación de los *mark-up*

A fin de considerar el rol de las diferencias en poder de mercado y características de la demanda a cada empresa en la estimación de determinantes de la PTF, se ha realizado la estimación de los *mark-up* por empresa utilizando el comando de Stata *markuppest* desarrollado por Rovigatti (2020) el cual permite estimar los márgenes de ganancia siguiendo a De Loecker y Warzynski (2012) en conjunto con la función de producción.

Los resultados del cálculo de los *mark-up* de las empresas del sector manufactura panel 2016-2018 se muestran en el siguiente gráfico.

**Gráfico 4: Distribución del *mark-up* a nivel de empresas
Manufactura Panel 2016-2018**



Fuente: Elaboración propia en base al panel EEA 2016-18.

Cabe indicar que un potencial problema para realizar la estimación del efecto de la inversión en I&D y de las TIC sobre la PTF de acuerdo con la metodología planteada es la alta correlación entre *mark-up* y la Intensidad de inversión en I&D (coeficiente de Pearson de 0.65).

5.3 Estimación del efecto del uso de las TIC y la inversión en I&D sobre la PTF

En este caso la estimación de determinantes de la PTF se realiza utilizando la PTF estimada mediante el método LP-A sin corregir por *mark-up* con la información de las empresas manufactureras del panel 2016-18.

Se han estimado dos especificaciones del modelo planteado. En la primera de ellas se incluye los 7 factores que sintetizan la información sobre uso de las TIC estimados previamente, y en la segunda se incluyen, en cambio, las distintas variables de uso de las TIC de forma separada lo cual, dada la correlación existente entre estas variables, conlleva un mayor riesgo de que se afecten los test de significancia estadística individual²⁹. En ambas especificaciones se incluye la variable de intensidad de la inversión en I&D estimada previamente a partir del modelo simultáneo de Heckman (de una etapa).

²⁹ Y se tiendan a aceptar con mayor probabilidad la hipótesis nula de que no hay efecto sobre la productividad.

Para la estimación se usan distintos métodos econométricos. Específicamente, el método de mínimos cuadrados ordinario (OLS) para datos ordenados (pool), el método panel con efectos aleatorios (RE), y el método panel con efectos fijos (FE). Dados los resultados del test de Breusch y Pagan y del test de Hausman, resulta preferible la estimación usando el método de panel con efectos fijos para ambas especificaciones³⁰.

A continuación, se muestran los resultados obtenidos para cada especificación.

Cuadro 14: Determinantes de la PTF a nivel de empresas - Manufactura Panel 2016-18

VARIABLES	(1)	(2)	(3)
	Mínimos cuadrados ordinarios (OLS)	Efectos aleatorios (RE)	Efectos fijos (FE)
Ratio I&D predicha / remuneraciones	0.259*** (0.024)	0.235*** (0.025)	0.270*** (0.046)
Logaritmo mark-up	-0.134*** (0.046)	-0.115** (0.055)	-0.084 (0.114)
Antigüedad	-0.009*** (0.002)	-0.011*** (0.003)	-0.015* (0.008)
Antigüedad al cuadrado	0.000* (0.000)	0.000*** (0.000)	0.000** (0.000)
Ratio ejecutivas mujeres	0.179*** (0.057)	0.095 (0.063)	-0.055 (0.100)
Ratio total mujeres	0.154* (0.086)	0.170* (0.101)	0.047 (0.188)
Factor 1 – Uso de PC y de Internet	0.052*** (0.017)	0.034* (0.018)	0.003 (0.026)
Factor 2 – Intranet y extranet	-0.029 (0.022)	-0.012 (0.022)	0.004 (0.030)
Factor 3 – Compra y venta por Internet	0.020 (0.023)	0.003 (0.024)	-0.024 (0.031)
Factor 4 – Software libre o desarrollado	-0.078*** (0.030)	-0.066** (0.031)	-0.053 (0.043)
Factor 5 - Red de área local, software propietario o comprado, y sistema de gestión	0.035 (0.034)	0.053 (0.034)	0.112** (0.046)
Factor 6 - Página web y anuncios por internet	-0.096** (0.043)	-0.115*** (0.043)	-0.190*** (0.060)
Factor 7 - Uso de Tablet/pda	0.085 (0.077)	0.111 (0.074)	0.140 (0.094)
Gran empresa (SBS)	0.207*** (0.034)	0.205*** (0.038)	0.195*** (0.067)
Corporativo (SBS)	0.835*** (0.060)	0.722*** (0.070)	0.357*** (0.137)
Manufacturas ligeras	0.037 (0.038)	0.040 (0.048)	-0.168 (0.144)
Intensivas en tecnología	-0.052 (0.039)	-0.032 (0.051)	0.198 (0.157)
Alimentos, bebidas y tabaco	0.029 (0.062)	-0.021 (0.081)	-0.460 (0.355)
Lima y Callao	-0.014 (0.046)	-0.025 (0.059)	-0.426** (0.214)
Constant	2.923*** (0.071)	2.982*** (0.088)	3.486*** (0.277)

³⁰ Ver Anexo 7.

Observations	2,114	2,114	2,114
R-squared	0.157	0.153	0.0472
RMSE	0.662	0.426	0.423
Prob > F	0	0	0

Notas:

Errores estandar robustos entre paréntesis. *** p<0.01, ** p<0.05, * p<0.1
TFP método Levinsohn y Petrin corregido por atrición en los datos.

B. Segunda Especificación

VARIABLES	(1)	(2)	(3)
	Mínimos cuadrados ordinarios (OLS)	Efectos aleatorios (RE)	Efectos fijos (FE)
Ratio Inversión en I&D predicha / remuneraciones	0.256*** (0.024)	0.233*** (0.025)	0.272*** (0.046)
Logaritmo mark-up	-0.148*** (0.046)	-0.124** (0.055)	-0.094 (0.115)
Antigüedad	-0.009*** (0.002)	-0.012*** (0.003)	-0.015* (0.008)
Antigüedad al cuadrado	0.000* (0.000)	0.000*** (0.000)	0.000** (0.000)
Ratio ejecutivas mujeres	0.183*** (0.057)	0.098 (0.063)	-0.057 (0.100)
Ratio total mujeres	0.151* (0.086)	0.169* (0.101)	0.066 (0.189)
Porcentaje uso de PC	0.033 (0.077)	0.032 (0.072)	-0.048 (0.092)
Porcentaje uso de Internet	0.145* (0.074)	0.087 (0.070)	0.071 (0.090)
Uso de Intranet	-0.067* (0.034)	-0.072** (0.033)	-0.068 (0.041)
Uso de Extranet	0.033 (0.040)	0.062 (0.038)	0.091* (0.047)
Uso de Red Local	0.041 (0.059)	0.079 (0.055)	0.133* (0.071)
Uso de banda ancha	-0.008 (0.045)	-0.021 (0.042)	-0.036 (0.052)
Uso de sistemas de gestión	0.167** (0.084)	0.143* (0.080)	0.163 (0.107)
Uso de PDA/Table	0.015 (0.034)	-0.005 (0.033)	-0.038 (0.041)
Uso de página Web	-0.156*** (0.041)	-0.124*** (0.042)	-0.138** (0.061)
Pago de anuncios por Internet	0.010 (0.033)	-0.029 (0.032)	-0.083** (0.042)
Compra software	-0.046 (0.038)	-0.025 (0.037)	0.025 (0.048)
Software libre	-0.017 (0.032)	-0.023 (0.030)	-0.005 (0.039)
Desarrolla software	-0.028 (0.033)	-0.017 (0.034)	-0.013 (0.047)
Capacitación en TIC	-0.031 (0.032)	0.001 (0.030)	0.024 (0.038)
Porcentaje de compras por Internet	-0.127 (0.143)	-0.168 (0.127)	-0.234 (0.152)
Porcentaje de ventas por Internet	0.234 (0.145)	0.222 (0.156)	0.123 (0.219)
Gran empresa (SBS)	0.213*** (0.034)	0.208*** (0.038)	0.192*** (0.068)
Corporativo (SBS)	0.847*** (0.061)	0.728*** (0.071)	0.354*** (0.137)

Manufacturas ligeras	0.022 (0.038)	0.031 (0.048)	-0.172 (0.145)
Intensivas en tecnología	-0.062 (0.039)	-0.036 (0.051)	0.206 (0.158)
Alimentos, bebidas y tabaco	0.011 (0.063)	-0.027 (0.082)	-0.473 (0.357)
Lima y Callao	-0.169 (0.108)	-0.131 (0.148)	-0.370 (0.262)
Norte	-0.050 (0.136)	-0.053 (0.179)	0.141 (0.455)
Centro	-0.178 (0.143)	-0.080 (0.188)	
Sur	-0.289** (0.129)	-0.233 (0.175)	
Constant	3.018*** (0.156)	3.017*** (0.190)	3.313*** (0.349)
Observations	2,114	2,114	2,114
R-squared	0.167	0.161	0.0504
RMSE	0.660	0.426	0.423
Prob > F	0	0	0

Notas:

Errores estándar robustos entre paréntesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

TFP método Levinsohn y Petrin corregido por atrición en los datos.

Fuente: Elaboración propia en base al panel EEA 2016-18.

Asimismo, se han realizado estimaciones de efectos heterogéneos para las variables de inversión en I&D y de uso de las TIC según tamaño de las empresas o subsector productivo dentro de la manufactura, cuyos resultados utilizando el modelo panel de efectos fijos se muestran para referencia en el Anexo 8³¹.

5.4 Principales resultados

Los resultados de la estimación de efectos fijos para ambas especificaciones del modelo indican que la inversión en I&D tendría un efecto positivo estadísticamente significativo sobre la PTF de las empresas manufactureras de mayor tamaño en Perú. Este efecto sería alrededor de 0.27 por ciento por cada punto de incremento de la ratio de inversión en I&D sobre remuneraciones.

Sin embargo, no se encuentra un efecto claro del uso de las TIC sobre la PTF, aunque sí efectos positivos estadísticamente significativos de variables específicas. En general, ello es consistente con los resultados positivos significativos estadísticamente que se obtienen para los factores que capturan su información en la primera especificación.

La primera especificación permite identificar el efecto positivo estadísticamente significativo del factor asociado al uso de red de área local, software propietario o comprado, y sistemas de gestión informatizados. Mientras que el factor relativo al uso de página web y pago de anuncios por internet muestra un signo negativo estadísticamente significativo.

En cuanto a los resultados para las variables específicas obtenidos con la segunda especificación una mayor intensidad del uso de computadora y del uso de Internet por

³¹ Cabe indicar que el número de observaciones de algunas categorías es reducido, por lo cual debe tomarse sus resultados deben tomarse con mayor cautela, específicamente en el caso del segmento corporativo y la industria de alimentos, bebidas y tabaco.

parte de los trabajadores no tendrían un efecto positivo estadísticamente significativo sobre la PTF de las empresas analizadas.

El uso de extranet como una tecnología que permite una mejor comunicación de la empresa con sus clientes y con sus proveedores, tiene un efecto positivo estadísticamente significativo. Sin embargo, el uso de intranet se asocia negativamente con la PTF.

Asimismo, el efecto positivo de la variable relativa al uso de red local es estadísticamente significativo utilizando 10% de probabilidad de error, mientras que el uso de algún sistema de gestión informatizado tiene un impacto sería positivo que no resulta significativo estadísticamente, aunque el p-value de su coeficiente es 0.13. La significancia estadística de estos resultados estaría afectada, en parte, debido a la alta difusión de ambas tecnologías entre las empresas de la muestra y su correlación con otras variables³². En todo caso, habría indicios de su importancia como un soporte para mejorar la productividad de las empresas. Particularmente, respecto al uso de sistemas informatizados, sería relevante analizar en investigaciones posteriores si alguna forma específica de combinación o integración de estos sistemas resulta más efectiva para incrementar la productividad de las empresas.

Tener una conexión a internet de banda ancha no conlleva diferencias en la productividad de las empresas. El uso de Tablet o PDA se muestra asociado de forma positiva, pero no estadísticamente significativa, a una mayor productividad.

Por otro lado, el uso de página web y el pago de anuncios por Internet se asocia con un menor nivel de productividad. Mientras que el uso de software ya sea comprado, libre o desarrollado por las empresas no se asocia con positivamente con la productividad (su efecto estimado es variable y no es significativo estadísticamente).

Por último, la compra y la venta a través de Internet tampoco tienen efectos positivos estadísticamente significativos sobre la productividad, siendo sus signos contrapuestos. La compra por internet tiene un signo negativo, mientras el porcentaje de ventas por Internet de la empresa muestra un efecto positivo fuerte que pierde su significancia estadística en la especificación del modelo con efectos fijos.

Cabe mencionar que, se realizó un ejercicio de estimación sin considerar la variable de intensidad de la inversión en I&D a fin de determinar cómo podría afectar la estimación del efecto de las TIC. En general, se mantienen los resultados respecto a la significancia estadística de las variables TIC (aumentando ligeramente en algunos casos como el sistema de gestión informatizado llegando el p-value de coeficiente a 0.11).

Asimismo, al no contarse con variables referidas a las capacidades de gestión o sobre la organización de las empresas, se analizó la posibilidad de utilizar como proxy si hay propiedad extranjera de la empresa. Sin embargo, la interacción de esta variable con las variables de las TIC no mostró efectos estadísticamente significativos sobre la PTF, excepto en el caso del efecto positivo del uso de Tablet/PDA que sí fue significativo³³. Ello, en parte, se podría deber a la baja proporción de empresas con propiedad

³² Justamente estas dos variables son las que presentan los mayores valores del factor de inflación de varianza (VIF, por sus siglas en inglés), siendo superiores a 10.

³³ Para el pago de anuncios por internet también se identifica un efecto positivo de la interacción, pero no resulta estadísticamente significativo con menos de 10% de probabilidad de error.

extranjera.

En cuanto a las otras variables incluidas en la estimación, las mismas muestran el efecto esperado sobre la PTF de acuerdo con la literatura empírica previa. Específicamente, el mayor tamaño de la empresa se vincula con una mayor PTF, destacando la fuerte diferencia en PTF a favor del segmento corporativo.

En *mark-up* (en logaritmo) muestra un signo negativo, pero no resulta estadísticamente significativo con menos de 10% de probabilidad de error. Ello se vincula al hecho de haber incluido también variables relativas al tamaño de las empresas con las cuales está correlacionado (su significancia aumenta si no se incluyen)³⁴.

Asimismo, la antigüedad o edad de la empresa tiene un efecto negativo estadísticamente significativo sobre la productividad (Tello, 2020), aunque su significancia estadística disminuya en el modelo de efectos fijos. Este efecto negativo de la mayor edad sería cada vez menor dado que el coeficiente estimado de la edad al cuadrado es positivo y estadísticamente significativo (aunque su valor es muy bajo, de alrededor de 0.0001).

Además, un mayor ratio de mujeres en los trabajadores tendría un impacto positivo sobre la PTF como lo había indicado antes Del Pozo (2019), aunque la significancia estadística se pierde en el modelo de efectos fijos debido sobre todo a su poca variabilidad en el tiempo para cada empresa.

Con relación a las estimaciones realizadas, se debe mencionar como una limitación importante de las mismas es que, dada la información de la EEA, no se ha contado con datos sobre variables que pueden influir en el efecto de la inversión en I&D y el uso de las TIC sobre la productividad de las empresas, de acuerdo con literatura previa. Específicamente, no se ha contado con variables relativas a características de la gestión de las empresas, ni al capital humano de sus gerentes, conductores, y trabajadores.

Asimismo, en este trabajo no se ha llevado a cabo la corrección por posible endogeneidad de la productividad respecto a la inversión en I&D, lo cual podría llevar a subestimar los efectos de la I&D (Doraszelski & Jaumandreau, 2013).

Finalmente, el hecho de que algunos resultados de variables de las TIC no resulten estadísticamente significativos, podría estar vinculado a que la muestra de la EEA se limita a las empresas de mayor tamaño que ya han logrado un mayor uso de las TIC a diferencia de las unidades productivas de menor tamaño. En este sentido, resulta importante realizar investigaciones incluyendo información de las empresas más pequeñas.

Efectos heterogéneos

Respecto a los efectos heterogéneos según tamaño de empresas, en primer lugar, se encuentra que el efecto sobre la PTF de una mayor intensidad de la inversión en I&D sería mayor en las empresas de mayor tamaño, pues el coeficiente estimado crece con

³⁴ En todo caso, su inclusión en la regresión como variable de control sí estaría justificada, sobre todo tomando en cuenta que afecta la significancia de algunas variables de uso de las TIC (p.e. se reduce el p-value del uso de un sistema de gestión informatizado).

el tamaño (aunque en el caso de las grandes empresas solo es estadísticamente significativo con 15% de probabilidad de error)

asimismo, se encuentra que el efecto positivo identificado previamente de determinadas variables de uso de las TIC se da solo para algunos segmentos específicos de tamaño empresarial. Así, el uso de Extranet solo tiene un efecto positivo estadísticamente significativo para las empresas del segmento corporativo, mientras el uso de red local lo tiene solo para las empresas grandes. Mientras tanto, la relación positiva entre el uso de sistemas de gestión y la PTF resulta estadísticamente significativa solo para las empresas medianas. Asimismo, se encuentra un efecto positivo de gran magnitud estantíamente significativo para la compra de software para el segmento corporativo.

Con relación a los posibles efectos negativos del uso de las TIC, el uso de banda ancha solo muestra una asociación negativa con la PTF en el caso de las empresas medianas, mientras la asociación negativa del uso de Intranet solo resulta estadísticamente significativa en el caso del segmento corporativo. En el caso de uso de página Web y de pago de anuncios por Internet, la asociación negativa solo se encuentra para las grandes empresas.

Respecto a las diferencias según subsector, se encuentra que el efecto positivo sobre la PTF de una mayor intensidad de la inversión en I&D es sustantivamente mayor para la industria de alimentos, bebidas y tabaco. No resulta estadísticamente significativo en el caso de las manufacturas ligeras.

Asimismo, en cuanto a los efectos positivos del uso de las TIC, en el caso del uso de red local y sistemas de gestión informatizados, son estadísticamente significativos solo para el subsector de manufacturas ligares. Además, el efecto positivo de la compra software se da en las industrias intensivas en tecnología, y el efecto positivo del porcentaje de ventas por Internet se identifica para la transformación de bienes primarios.

De igual manera, la asociación negativa del pago por anuncio por Internet con la PTF se encuentra solo en los subsectores de manufacturas ligeras y de manufacturas intensivas en tecnología. En el caso del uso de banda ancha, la asociación negativa con al PTF se halla para las manufacturas ligeras.

Por último, en el caso de porcentaje de compras por Internet se encuentra efectos estadísticamente significativos contrapuestos, pues se haya un efecto negativo para la transformación de bienes primaries, y un efecto positivo de mucho mayor magnitud para el subsector alimentos, bebidas y tabaco.

6 Conclusiones y recomendaciones

La literatura económica ha evidenciado el efecto positivo del uso de las tecnologías de la información y comunicación (TIC) y de la inversión en Investigación y Desarrollo (I&D) sobre la productividad total de factores (PTF).

En esa línea, el presente trabajo estima estos efectos para el caso las empresas de mayor tamaño de la industria manufacturera peruana usando la Encuesta Económica Anual (EEA), a partir de la estimación del PTF mediante métodos modernos que corrigen problemas de endogeneidad de los factores o insumos productivos, de los cuales se

selecciona el método de Levinsohn & Petrin (2003) con corrección por *attrition* (salida de las empresas del panel).

A partir del planteamiento del modelo estándar CDM utilizado en estudios previos, se estima los determinantes de la inversión en I&D y de su nivel mediante un modelo de corrección de sesgo de selección de Heckman, a fin de predecir la intensidad de la inversión en I&D a partir de la información disponible. En el caso del uso de las TIC se construye índices sintéticos del a partir de las variables disponibles en la información de la EEA mediante análisis factorial (dada la significativa correlación entre varias de ellas). Además, se considera el rol de la demanda y la estructura de mercado al incluir como determinante de la PTF una estimación de los *mark-up* a nivel de empresas.

Los resultados de la estimación de los efectos sobre la PTF de la intensidad de la inversión en I&D y el uso de las TIC indican que la inversión en I&D tendría un efecto positivo estadísticamente significativo sobre la PTF a nivel de las empresas manufactureras de mayor tamaño en Perú; mientras que, en general, no se encuentra un efecto claro del uso de las TIC sobre la PTF de las empresas.

Por ello, una recomendación de política resultante para este tipo de empresas es priorizar las intervenciones que fomenten la innovación mediante la inversión en I&D en las empresas respecto a la promoción general del acceso y uso de las TIC, dado su mayor impacto sobre la productividad. Dentro de este tipo de intervenciones cabe destacar los fondos públicos concursables de subsidios para el financiamiento de proyectos de I&D y los esquemas de incentivos tributarios para la inversión en proyectos de I&D que hayan sido debidamente validados. El desarrollo de estas intervenciones, que ya existen en el caso de Perú³⁵, deben contar con evaluaciones de desempeño y de impacto rigurosas, cuyos resultados se utilicen para hacer las mejoras en su diseño e implementación que permitan incrementar sus efectos positivos.

Sin embargo, el estudio sí halla efectos positivos estadísticamente significativos para variables específicas de uso de las TIC como el uso de extranet y, en menor medida, el uso de red local, que favorecen la capacidad de las empresas de relacionarse con sus clientes y proveedores, y entre los propios equipos, respectivamente. Además, se encuentra indicios de un efecto positivo del uso de algún sistema de gestión informatizado, siendo factible que este efecto sea más importante si se produce la integración adecuada de distintos sistemas. Al contrario, en el caso de otras variables de las TIC como el uso de página web o de anuncios pagados por Internet, se halla una asociación con un menor nivel de productividad.

Asimismo, existirían efectos heterogéneos de las variables de uso de las TIC según características de las empresas como su tamaño (aunque no se incluya las más pequeñas) o subsector productivo. Sin embargo, en el caso de otras variables como la capacidad de gestión empresarial que podría asociarse a efectos positivos más altos,

³⁵ Específicamente, el Programa ProInnovate (antes Innóvate Perú) cofinancia proyectos de innovación empresarial, desarrollo productivo, emprendimiento e instituciones del ecosistema, a través de concursos nacionales, financiados con recursos de la cooperación internacional (sobre todo del BID), y la Ley N° 30309, Ley que promueve la Investigación Científica, Desarrollo Tecnológico e Innovación Tecnológica, la cual otorga incentivos tributarios a empresas que inviertan en proyectos de I + D + i realizados a través de centros autorizados por el onsejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (Concytec).

no ha podido ser incluida en el presente estudio por restricciones de la información disponible.

En este sentido, es importante que las políticas de promoción del uso de las TIC se concentren en aquellas TIC cuyos servicios tengan un mayor efecto esperado sobre la productividad de las empresas. Para ello es recomendable que se realicen investigaciones que evalúen el impacto del uso de TIC específicas, si fuera posible de tipo experimental, que generen evidencias rigurosas en las cuales se basen las intervenciones públicas diseñadas con este fin, considerando la posibilidad de impactos heterogéneos según características de las empresas.

Entre las limitaciones de las estimaciones realizadas en este estudio cabe indicar que no se ha llevado a cabo la corrección por endogeneidad la productividad respecto a la inversión en I&D, lo cual podría llevar a subestimar los efectos de la I&D (Doraszelski & Jaumandreau, 2013). Este punto debe formar parte de la agenda de investigación pendiente.

Por último, la muestra utilizada en este trabajo - dada la fuente de información disponible - corresponde a las empresas manufactureras de mayor tamaño que ya han logrado un mayor uso de las TIC a diferencia de las unidades productivas más pequeñas, siendo necesario realizar estudios que se enfoquen en las empresas de menor tamaño (micro y pequeñas empresas) o las incluyan como un segmento de análisis. Este tipo de estudios deberían desarrollarse, posteriormente, a medida que se disponga de la información correspondiente.

Referencias

- Abramovitz, M. 1956. Resource and Output Trends in the United States since 1870. *American Economic Review*, 46(2), 5-23.
- Acemoglu D, Aghion P, Zilibotti F. 2006. Distance to frontier, selection, and economic growth. *Journal European Economic Association*. 4, 37–74
- Ackerberg, D. A., L. Benkard, and A. Pakes. 2007. Econometric Tools for Analyzing Market Outcomes. In: *Handbook of Econometrics*. Amsterdam: North-Holland.
- Ackerberg, D. A., K. Caves, and G. Frazer. 2015. Identification properties of recent production function estimators. *Econometrica*, 83 (6), 2411-2451.
- Aguirregabiria, V. 2018. Empirical Industrial Organization, University of Toronto.
- Akerman, A., Gaarder, I., & Mogstad, M. 2015. The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, 130(4), 1781-1824.
- Alderete, M. 2012. *Medición de las tecnologías de información y la comunicación en empresas de servicios de Colombia*. Colombia.
- Álvarez, R.; Bravo-Ortega, C.; Navarro, L. 2011. Innovación, investigación y desarrollo, y productividad en Chile. *Revista Cepal*, 104, 141-166.
- Anlló, G., and Suárez, D. 2009. "Innovación: algo más que I&D. Evidencias ibeoramericanas a partir de las encuestas de innovación. Construyendo las estrategias empresarias competitivas". Cepal-Redes, Buenos Aires.
- Arbeláez, M., and Parra, M. 2011. Innovation, R&D investment and productivity in Colombian Firms. Working Paper Series N° 251. Washington, D. C.: BID.
- Arvantis, S. 2005. Computerization, Workplace Organization, Skilled Labour and Firm Productivity: Evidence for the Swiss Business Sector, *Economics of Innovation and New Technology*, 14(4), 225-249.
- Bayo-Moriones, J.; Billón, M.; Lera, F. 2008. Skills, Technology and Organisational International Innovations in Spanish Firms. *Journal of Manpower*, 29(2), 122-145.
- Benavente, J. 2006. The Role of Research and Innovation in Promoting Productivity in Chile. *Economics of Innovation & New Technology*, 15(4/5), 301-315.
- Bernard. A. and J. Jensen (2004). Exporting and productivity in the US. *Oxford Review of Economic Policy*, 20 (3), 343-357.
- Bertschek, I and U. Kaiser. 2004. Productivity Effects of Organizational Change: Microeconomic Evidence. *Management Science*, L (3), 394-404.

- Billon, M., F. Lera, y S. Ortiz 2007. Evidencias del impacto de las TIC en la productividad de la empresa. ¿Fin de la paradoja de la productividad?. *Cuadernos de Economía*, 30(82).
- Black, S.; Lynch, L. 2001. «How to Compete: the Impact of Workplace Practices and Information Technology on Productivity», *Review of Economics and Statistics*, 83, 435-445.
- Black, S.; Lynch, L. 2004. «What's Driving the New Economy?: The Benefits of Workplace Innovation», *The Economic Journal*, 114(493), 97-116.
- Bloom, N.; Garicano, L.; Sadun, R. and van Reenen, J. 2009. The distinct effects of Information Technology and Communication Technology on firm organization. NBER Working Paper No. 14975. Cambridge, Massachusetts, National Bureau of Economic Research.
- Bloom, N.; Sadun, R. and van Reenen, J. 2012. Americans Do IT Better: US Multinationals and the Productivity Miracle. *American Economic Review*, 102(1), 167-201.
- Blundell, R., and S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87, 115–143.
- Blundell, R. S. Bond 2000. GMM estimation with persistent panel data: an application to production functions. *Econometric Reviews*, 19 (3), 321-340.
- Bound, J., Jaeger, D. and R. Baker. 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak. *Journal of the American Statistical Association*, 90(430), 443-450.
- Bocquet, R.; Brossard, O. ; Sabatier, M. 2007. Complementarities in organizational design and the diffusion of information technologies: An empirical analysis. *Research Policy*, 36, 367-386.
- Bravo-Ortega, C., Benavente, J. M., & Gonzales, Á. 2014. Innovation, exports, and productivity: Learning and self-selection in Chile. *Emerging Markets Finance and Trade*, 50(1), 68-95.
- Bresnahan, T.; Brynjolfsson, E.; Hitt, I. 2002. «Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm Level Comments Firm Level Evidence», *Quarterly Journal of Economics*, 117(1), 339-376.
- Brynjolfsson, E. 1993. The Productivity Paradox of Information Technology, *Communications of ACM*, XXXVI (12), 66-77.
- Brynjolfsson, E.; Hitt, L.M. 2000. Beyond Computation: Information Technology, Organizational Transformation and Business Practices. *Journal of Economic Perspectives*, XIV(4), 23-48.

- Brynjolfsson, E.; Hitt, L.M. 2003. Computing Productivity: Firm-Level Evidence. *Review of Economics & Statistics*, 65(4), 793- 809
- Cameron, A. and P. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Cassoni, A., & Ramada-Sarasola, M. 2010. Innovation, R&D investment and productivity: Uruguayan manufacturing firms. Working Paper Series N° 65. Washington, D.C.: IDB.
- Céspedes, N., M. E. Aquije, A. Sánchez, y R. Vera-Tudela. 2014. Productividad sectorial en el Perú: Un análisis a nivel de empresas. *Revista de Estudios Económicos*, 28(9), 9-26.
- Cohen, W., and D. Levinthal. 1989. Innovation and Learning: Two Faces of R&D. *Economic Journal*, 99, 569–596.
- Colombo, M.G., Croce, A., Grilli, L. 2013. ICT services and small businesses' productivity gains: an analysis of the adoption of broadband internet technology. *Information Economics and Policy* 25(3), 171–189.
- Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (Concytec). 2016. *Programa nacional de ciencia, tecnología e innovación en tecnologías de la información y comunicación*. Lima.
- Constanza Demmel, M. & Máñez, J. & Rochina-Barrachina, M. & Sanchis-Llopis, J. 2013. Innovation and productivity: evidence for 4 Latin American countries manufacturing industry. Working Papers 1307, Department of Applied Economics II, Universidad de Valencia.
- Crépon, B., Duguet, E., Mairesse, J., 1998. Research, innovation and productivity: an econometric analysis at the firm level. *Economics of Innovation and New Technology*. 7 (2), 115–158.
- Crespi, G.; Criscuolo, Ch.; C.; Haskel, J. 2006. Information Technology, Organizational Change and Productivity Growth: Evidence from UK Firms. Working Papers 558, Queen Mary University of London, School of Economics and Finance.
- Crespi, G., & Zuniga, P. 2010. Innovation and productivity: Evidence form six Latin American countries. Inter-American Development Bank. IDB Working Paper Series No. IDB-WP-218.
- Crespi, G., & Tacsir, E. 2011. Effects of Innovation on Employment in Latin America. The Atlanta. Conference on Science and Innovation. Atlanta.
- Crespi, G., Tacsir, E., & Vargas, F. 2016. Innovation dynamics and productivity: Evidence for Latin America. En: *Firm innovation and productivity in Latin America and the Caribbean* (37-71). Palgrave Macmillan, New York.

- Dai, X. y L. Cheng. 2018. The impact of product innovation on firm-level markup and productivity: evidence from China. *Applied Economics*. 50(42), 4570-4581.
- De Loecker, J. 2011a. Recovering markups from production data. *International Journal of Industrial Organization*. 29, 350-355.
- De Loecker, J. 2011b. Product Differentiation, Multi-Product Firms and Estimating the Impact of Trade Liberalization on Productivity. *Econometrica*, 79(5), 1407–1451.
- De Loecker, J. 2013. Detecting Learning by Exporting. *American Economic Journal: Microeconomics*, 5(3), 1-21.
- De Loecker, J., and F. Warzynski. 2012. Markups and firm-level export status. *American Economic Review*. 102(6). 2437-71.
- Del Pozo, C. y E. Guzmán. 2019. *Estimación de la Productividad Total Factorial a nivel de empresas en el Perú: nueva evidencia a través del método Akerberg, Caves y Frazer*. Lima: CIES.
- Doms, M., Dunne, T., Troske, K. 1997. Workers, Wages and Technology. *Quarterly Journal of Economics*, 112 (1), 253-290.
- Doraszelski, U. and J. Jaumandreu. 2013. R&D and Productivity: Estimating Endogenous Productivity. *Review of Economic Studies*, 80, 1338-1383.
- Doraszelski, U. and J. Jaumandreu. 2018. Measuring the Bias of Technological Change. *Journal of Political Economy*, 126(3), 1027 – 1084.
- Doraszelski, U. and J. Jaumandreu. 2019. Using Cost Minimization to Estimate Markups. CEPR Discussion Paper Series.
- Doraszelski, U. and J. Jaumandreu. 2021. “Reexamining the De Loecker & Warzynski (2012) method for estimating markups”. Mimeo.
- Draca, M., R. Sadun, & J. Van Reenen. 2007. Productivity and ICTs: A review of the evidence. In: Robin Mansell, Chrisanthi Avgerou, D. Q. and Silverstone, R. (eds.) *The Oxford Handbook of Information and Communication Technologies*, 100-147. Oxford University Press.
- Duguet, E. 2006. Innovation Height, Spillovers and TFP Growth at the Firm Level: Evidence from French Manufacturing. *Economics of Innovation & New Technology*, 15(4/5), 415-442.
- Entorf, H. and F. Kramarz. 1998. The Impact of New Technologies on Wages: Lessons from Matching Panels on Employees and on their Firms. *Economic Innovation and New Technology*, 5, 165-197.
- Eslava, M., J. Haltiwanger, A. Kugler y M. Kugler 2010. Market Reforms, Factor Reallocation, and Productivity Growth in Latin America. En: Loayza, N. y L. Servén (eds.)

Business Regulation and Economic Performance, 225-263. Washington: The World Bank.

Fabiani, S., Schivardi, F., Trento, S. 2005. ICT adoption in Italian manufacturing: firm-level evidence. *Industrial and Corporate Change*, 14 (2), 225–249.

Färe, R., S. Grosskopf, and D. Margaritis. 2008. Efficiency and Productivity: Malmquist and More. In: H. Fried; C. Knox, and S. Schmidt (eds.). *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, NY.

Farrell, M. J. 1957. The Measurement of Productive Efficiency. *Journal of Royal Statistical Society*. 120, 253-290.

Fazlıoğlu, B.; Dalgıç, B.; Burçin; A. 2019. The effect of innovation on productivity: evidence from Turkish manufacturing firms. *Industry and Innovation*, Taylor & Francis Journals, 26(4), 439-460.

Forlani, E; R. Martin; G. Mion, y M. Muuls. 2016. *Unraveling Firms: Demand, Productivity and Markups Heterogeneity*.

Francalanci, C.; Galal, H. 1998. Information Technology and Worker Composition: Determinants of Productivity in the Life Insurance Industry. *MIS Quarterly*, 22(2), 227-241.

Fried, H. O., C. A. Knox-Lovell, and S. S. Schmidt. 2008. Efficiency and Productivity. In: H. Fried; C. Knox, and S. Schmidt (eds.). *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, NY.

Gallardo, J., A. Arrieta 2000. *Medición y Dinámica de la Producción Industrial*. CIES, Lima-Perú.

Gallego, J., Gutiérrez, L., & Lee, S. 2014. A firm-level analysis of ICT adoption in an emerging economy: evidence from the Colombian manufacturing industries. *Industrial and Corporate Change*, 24(1), 191– 221.

Gandhi, A., S. Navarro, and D. Rivers. 2011. Does Value Added Overstate Productivity Dispersion? Identification and Estimation of the Gross Output Production Function. University of Wisconsin-Madison and University of Western Ontario.

Gandhi, A., S. Navarro, and D. Rivers. 2016. On the Identification of Production Functions: How Heterogeneous is Productivity. Unpublished Report, University of Wisconsin-Madison.

Gandhi, A.; S. Navarro y D. Rivers. 2020. On the Identification of Gross Output Production Functions. *Journal of Political Economy*, 128(8), 2973- 3016.

Giuri, P., Torrisi, S., Zinovyeva, N., 2008. ICT, skills, and organizational change: evidence from Italian manufacturing firms. *Industrial and Corporate Change* 17(1), 29–64.

- Griffith, R, Huergo, E., Mairesse, J., & Peters, B. 2006. Innovation and productivity across four European countries. *Oxford review of economic policy*, 22(4), 483-498.
- Griffith, R.; S. Redding Y J. Van Reenen. 2004. Mapping the Two Faces of R&D: Productivity Growth in a panel of OECD Countries. *Review of Economics and Statistics* 86(4), 883-95.
- Griliches, Z. 1979. Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*, 10, 92-116.
- Griliches, Z. 1994. Productivity, R&D and the data constraint. *American Economic Review*, 84(1), 1-23.
- Griliches, Z., and J. Mairesse. 1995. Production Functions: The Search for Identification. NBER Working Paper 5067. Cambridge, Massachusetts, National Bureau of Economic Research.
- Griliches, Zvi, 1998. *R&D and Productivity. The econometric evidence*. National Bureau of Economic Research Books, University of Chicago Press.
- Griliches, Z. 2000. *R&D, education, and productivity: A retrospective*. Harvard University Press, Cambridge.
- Gu, W., and J. Tang. 2004. Link Between Innovation and Productivity in Canadian Manufacturing Industries. *Economics of Innovation & New Technology*, 13(7), 671–686.
- Guillard, Ch; J Jaumandreu; J. Olivari. 2018. *Exploring Ways to Estimate Endogenous Productivity*. Boston University.
- Guloglu, B. & R. Tekin, 2012. A Panel Causality Analysis of the Relationship among Research and Development, Innovation, and Economic Growth in High-Income OECD Countries. *Eurasian Economic Review*, 2(1), 32-47
- Hall, B. 2011. Innovation and productivity. NBER Working Paper N° 17178. Cambridge, Massachusetts, National Bureau of Economic Research.
- Hall, B.; Lotti, F. and J. Mairesse. 2013. Evidence on the impact of R&D and ICT investment on innovation and productivity in italian firms. NBER Working Paper 18053. Cambridge, Massachusetts, National Bureau of Economic Research.
- Hall, B.H., Mairesse, J., 1995. Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics* 65(1), 263–293.
- Hall, B. H., Mairesse, J., & Mohnen, P. 2010. Measuring the Returns to R&D. UNU-MERIT Working Paper Series; No. 006. Maastricht: UNU-MERIT.
- Hansen, L. 1982. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(3), 1029-1054.

Hitt, L. and E. Brynjolfsson. 1996. Productivity, Business Profitability, and Consumer Surplus: Three Different Measures of Information Technology Value. *MIS Quarterly*, 20(2) 121-142.

Hollenstein, H. 2004. The Decision to Adopt Information and Communication Technologies (ICT) Firm-Level Evidence for Switzerland. En: OECD, *The Economic Impact of ICT. Measurement, evidence and implications* (37-60). Paris: OECD.

Huaroto, C. 2012. Use of the Internet and Productivity of Microbusinesses: Evidence from the Peruvian Case (2007–2010). *Information Technologies and International Development*, 8(4), 113-128.

Huergo, E. and J. Jaumandreu. 2004. Firms' Age, Process Innovation and Productivity Growth. *International Journal of Industrial Organization*, 22(4), 541-559.

Instituto Nacional de Estadística e Informática (INEI). 2015. *Tecnologías de Información y Comunicación en las Empresas*. Lima.

Instituto Nacional de Estadística e Informática (INEI). 2020. *Perú; Tecnologías de Información y Comunicación en las Empresas, 2017. Encuesta Económica Anual 2018*. Lima.

Isgut., A. M.D. Tello, A. Veiderpass. 1999. Microeconomic Adjustment During Structural Reforms: The Nicaraguan Manufacturing Sector 1991-1995. *Canadian Journal of Development Studies*, 20(3), 1-24.

Jefferson, G.H. *et al.* 2006. R&D Performance in Chinese Industry. *Economics of Innovation and New Technology*, 15(4/5), 345-366.

Jorgenson, D., and Z. Griliches. 1967. The explanation of Productivity Change. *Review of Economics Studies*, 34(3), 249-283.

Kasahara, H. and Joel Rodrigue (2008). Does the use of imported intermediates increase productivity? Plant-level evidence. *Journal of Development Economics*, 87 (1), 106-118.

Katz, & Callorda. 2015. *Impacto de arreglos institucionales en la digitalización y el desarrollo económico en América Latina*.

Kohli, R. and S. Devaraj. 2003. Measuring Information Technology Payoff: a Meta-Analysis of Structural Variables in Firm-Level Empirical Research. *Information System Research*, XIV(2), 127-145.

Kumar, V., & Raheja, G. (2012). Business to Business (B2B) and Business to Consumer (B2C) Management. *International Journal of Computers & Technology*, 3(3), pp. 447-451.

Llamas, J. 2020. Capital Fijo. Economipedia. Recuperado el 15 de octubre de 2021, de: <https://economipedia.com/definiciones/capital-fijo.html>

- Lee, Y. A. Stoyanov, and N. Zubanov. 2019. Olley and Pakes-style Production Function Estimators with Firm Fixed Effects. *Oxford Bulletin of Economics and Statistics*, 81(1), 79-97.
- Levinsohn, J., and A. Petrin. 2003. Production Functions using Inputs to Control for Unobservables. *Review of Economic Studies* 70(2), 317-341.
- Lööf, H., & Heshmati, A. 2002. Knowledge capital and performance heterogeneity: A firm-level innovation study. *International Journal of Production Economics*, 76(1), 61-85.
- Lööf H., and A. Heshmati. 2006. On the Relationship between Innovation and Performance: A Sensitivity Analysis. *Economics of Innovation & New Technology* 15(4/5), 317-344.
- Mairesse, J. and M. Sassenou. 1991. R&D and productivity: a survey of econometric studies at the firm level. *STI Review*, 8, 9-46.
- Mairesse, J., Jaumandreu, J., 2005. Panel-data estimates of the production function and the revenue function: what difference does it make? *Scandinavian Journal of Economics* 107(4), 651–672.
- Manjón, M., and J. Mañez. 2016. Production function estimation in Stata using the Akerberg-Caves-Frazer method. *The Stata Journal*, 16(4), 900-916.
- Marschak J. and W.H. Andrews. 1944. Random Simultaneous Equations and the Theory of Production. *Econometrica*, 12(3/4), 143-205.
- Martin, L. and T. U. Nguyen-Thi. 2015. The Relationship between Innovation and Productivity Based on R&D and ICT Use: An Empirical Analysis of Firms in Luxembourg. *Revue économique*, 66, 1105-1130.
- Masso, J., and P. Vahter. 2008. Technological Innovation and Productivity in Late-Transition Estonia: Econometric Evidence from Innovation Surveys. *European Journal of Development Research*, 20(2), 240-261.
- Matteucci, N.; O'Mahony, M.; Robinson, C.; Zwick, T. 2005. Productivity Workplace Performance and ICT: Industry and Firm-Level Evidence for Europe and the U.S. *Scottish Journal of Political Economy*, 52 (3) 359-386.
- Melitz, M. J. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71, 1695–1725.
- Menon, N.; Lee, B. and L Eldenburg. 2000. Productivity of Information System in the Healthcare Industry. *Information Systems Research*, 11(1), 83-92.
- Milgrom, P. and J. Roberts. 1990. *The economics of modern manufacturing*. *American Economic Review*, 80, 511–528.

Ministerio de la Producción (PRODUCE). 2016. *Estudio de la situación actual de la innovación en la industria manufacturera. Análisis de los Resultados de la Encuesta Nacional de Innovación de la Industria Manufacturera 2015*. Lima.

Ministerio de la Producción (PRODUCE). 2020a. *Las Mipyme en cifras 2018*. Lima.

Ministerio de la Producción (PRODUCE). 2020b. *Innovación en la industria manufacturera y en las empresas de servicios intensivos en conocimiento. Análisis de los Resultados de la ENIIMESIC 2018*. Lima.

Ministerio de la Producción (PRODUCE). 2021. *Las Mipyme en cifras 2019*. Lima.

Mohan, P.; Strobl, E.; Watson, P. 2018. In-firm training, innovation and productivity: the case of Caribbean Small Island Developing States. *Entrepreneurship & Regional Development*, 30(9-10), 987-1011

Mohnen, P. 2019. R&D, innovation and productivity. In: *The Palgrave Handbook of Economic Performance Analysis* (97-122).

Mohnen, P., & Hall, B. H. 2013. Innovation and productivity: An update. *Eurasian Business Review*, 3(1), 47-65.

Mollisi, V. y G. Rovigatti. 2018. Theory and practice of total-factor productivity estimation: The control function approach using Stata. *The Stata Journal* 18, 3, 618–662.

Mundlak, Y. 1961. Empirical Production Function Free of Management Bias. *Journal of Farm Economics*, 43, 44-56.

Mundlak, Y, and I. Hoch. 1965. Consequences of Alternative Specifications in Estimation of Cobb-Douglas Production Functions. *Econometrica*, 33, 814-828.

Nelson, R. and Phelps, E. 1966. Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, 56(1/2), 69-75.

Nolazco, J. 2020. Efectos entre las actividades de innovación, exportación y productividad: un análisis de las empresas manufactureras peruanas. *Desarrollo y Sociedad*, 85, 67-109.

Organisation for Economic Co-Operation and Development (OECD). 2001. *Measuring Productivity: Measurement of Aggregate and Industry – Level Productivity Growth*. OECD. París.

Olley, S., and A. Pakes. 1996. The dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*. 64(6), 1263-1297.

Ortega-Argilés, R., M. Piga and M. Vivarelli. 2015. The productivity impact of R&D investment: are high-tech sectors still ahead? *Economics of Innovation and New Technology*, 24(3), 204-222.

Pakes, A., and Z. Griliches. 1984. Patents and R&D at the Firm Level: A First Look." In: Z. Griliches, editor. R&D, Patents, and Productivity. Chicago, United States: University of Chicago Press and National Bureau of Economic Research.

Petrin, A., B. P. Poi, and J. Levinsohn. 2004. Production function estimation in Stata using inputs to control for unobservables. *The Stata Journal* 4, 2, 113-123.

Ramírez, S.; J. Gallego and M. Tamayo. 2019. Human capital, innovation and productivity in Colombian enterprises: a structural approach using instrumental variables, *Economics of Innovation and New Technology*,

Reeson, A. and L. Rudd. 2016. ICT Activity, Innovation and Productivity: An Analysis of Data From Australian Businesses. *Economic Papers*. 35(3), 245–255.

Restrepo-Morales, J.; Loaiza, O. y Vanegas, J. 2019. *Journal of Economics, Finance and Administrative Science*. 24 (47), pp.97-112.

Rovigatti, G. 2020. MARKUPEST: Stata module for markup estimation via Micro and Macro approaches. Statistical Software Components S458820. Boston College Department of Economics.

Saliola, F., and M. Seker. 2011. Total factor productivity across the developing world. *Enterprise Note Series*, 23.

Sargan, J. 1958. The estimation of economic relationships using instrumental variables. *Econometrica*, 26(3), 393-415.

Solow, R. 1957. Technical change and the aggregate production function. *Review of Economics and Statistics*, 39(3), 312-320.

Stock, J. and Staiger, D. 1997. Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3), 557-586.

Tello, M. 2012. Productividad total factorial en el sector manufacturero del Perú 2002-2007. *Economía*, 35(70), 103-141.

Tello, M. 2014. *La productividad del sector manufacturero en el Perú: 2000-2010*. Lima: INEI.

Tello, M. 2017. Innovación y productividad en las empresas de servicios y manufactureras: el caso del Perú. *Revista Cepal*, 121, 73-92.

Tello, M. 2020. Investigación y desarrollo, tecnologías de información y comunicación e impactos sobre el proceso de innovación y la productividad. Documento de Trabajo 487. Lima: PUCP.

Thanh, Q., & Tran, Q. (2020). Firm heterogeneity and total factor productivity: New panel-data evidence from Vietnamese manufacturing firms. *Management Science Letters*, 10(7):1505-1512.

van Beveren, I. 2012. Total Factor Productivity Estimation: A Practical Review. *Journal of Economic Surveys*, 26(1), 98-128.

van Leeuwen, G., and L. Klomp. 2006. On the Contribution of Innovation to Multi-Factor Productivity Growth. *Economics of Innovation & New Technology* 15(4/5): 367-390.

Verardi, V., and C. Vermandele. 2016. Outlier identification for skewed and/or heavytailed unimodal multivariate distributions. *Journal de la Société Française de Statistique*, 157.

Viollaz, M. 2018. ICT Adoption in Micro and Small Firms: Can Internet Access Improve Labour Productivity?. Documento de Trabajo N° 223. CEDLAS.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge MA.

Wooldridge, J. M. 2009. On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3), 112-114.

ANEXOS

Anexo 1: Métodos de estimación de la Productividad Total de Factores (PTF)

En esta sección se resume las principales características de los métodos de estimación paramétricos de la PTF siguiendo a Lee *et al.* (2019), Rovigatti y Mollisi (2018) y Del Pozo y Guzmán (2019).³⁶

En general, se asume una función de producción de sustitución tipo Leontief entre el valor agregado y la demanda intermedia o los insumos³⁷, donde la producción total es proporcional a ambos y el valor agregado se genera de acuerdo a una función Cobb-Douglas (Akerberg *et al.*, 2015) del siguiente modo:

$$Y_{i,t} = \min \left\{ e^{\beta_0} K_{i,t}^{\beta_k} L_{i,t}^{\beta_l} e^{v_{i,t}}, M_{i,t} \right\}$$

En la ecuación anterior, $Y_{i,t}$ es el nivel de producción total de la empresa “i” en el periodo “t”.

La producción es una función del capital (al fin del período) $K_{i,t}$, trabajo $L_{i,t}$ y es proporcional al valor de los insumos $M_{i,t}$. Se asume *Hick's neutral productivity shocks* respecto al término $v_{i,t}$, (Gandhi *et al.*, 2014) que representa el logaritmo de la PTF de cada empresa. $v_{i,t}$, puede ser descompuesto en:

$$v_{i,t} = \omega_{i,t} + \varepsilon_{i,t}$$

Si se cumplen las condiciones de primer orden, se tiene que $Y_{i,t} = M_{i,t} = \min \left\{ \beta_0 K_{i,t}^{\beta_k} L_{i,t}^{\beta_l} e^{v_{i,t}} \right\}$ y tomando logaritmos, se llega a una función de producción agregada “*Value-added production*” que no incluye a los insumos, dada por:

$$va_{i,t} = \beta_0 + \beta_l l_{i,t} + \beta_k k_{i,t} + \omega_{i,t} + \varepsilon_{i,t}$$

Donde $va_{i,t} = y_{i,t} - m_{i,t}$, es el valor agregado de la producción expresada como la diferencia entre el logaritmo de la producción $y_{i,t}$ y el uso de insumos $m_{i,t}$. β_0 representa la eficiencia promedio de las empresas, $l_{i,t}$ es el logaritmo del trabajo, $k_{i,t}$ es el logaritmo del capital, $\omega_{i,t}$ es una variable que representa la productividad no observable por el econométrico, pero posiblemente observable por la empresa i y considerada al tomar sus decisiones *ex - ante* respecto del uso de los factores de producción. En este caso, $\varepsilon_{i,t}$ es un shock estocástico en la producción o la productividad que ocurre luego de ser tomadas las decisiones de uso de insumos intermedios y de producción por parte de las empresas.

³⁶ Respecto a los símbolos se utiliza la notación usada por Del Pozo y Guzmán (2019).

³⁷ La especificación de tipo Leontief de la función de producción corresponde a la situación donde una cantidad específica de cantidad de insumos es necesaria para producir una unidad de producto.

1. Método Olley y Pakes (1996)

Olley y Pakes (1996) proponen utilizar variables observables para controlar por la productividad inobservable, desarrollando un estimador semi-paramétrico consistente ante problemas de endogeneidad empleando el nivel de inversión ($i_{i,t}$) como un proxy de los shocks de productividad no observables. Dado su nivel de cápita al inicio de cada periodo, cada empresa decide si continuar o no continuar con sus operaciones y, de mantenerse en el mercado, el nivel de factores de producción y de inversión. La empresa maximiza el valor esperado del flujo de caja e inversiones sujeto a la expectativa sobre la estructura futura de mercado.

En el método de Olley y Pakes (1996) se asume que el trabajo $l_{i,t}$ es un factor perfectamente flexible. Se asume que la inversión $i_{i,t}$ es una variable ajustable denominada “libre” que depende del nivel de productividad, del capital, y otras variables no observables $z_{i,t}$ mediante una función $i_{i,t} = \phi(\omega_{i,t}, k_{i,t}, z_{i,t})$, siendo $\phi(\cdot)$ estrictamente creciente con la productividad $\omega_{i,t}$.

De este modo, las empresas que esperan un shock positivo de productividad en el periodo “t” incrementarán su nivel de inversión en dicho periodo para un nivel de capital dado, lo cual se conoce como supuesto de *Scalar unobservability* (Lee, 2019).

La función para la productividad puede ser obtenida como la inversa de ϕ (debiendo cumplirse la condición de invertibilidad en inversión):

$$\omega_{i,t} = \phi^{-1}(i_{i,t}, k_{i,t}) = g(i_{i,t}, k_{i,t}, z_{i,t})$$

La función de control g expresada se aproxima con un polinomio de generalmente tercer grado, denotado $\tilde{g}(\cdot)$ que puede ser introducida en la ecuación anterior.

Con ello la función de producción agregada “*Value-added production*” puede ser expresada como:

$$va_{i,t} = \beta_0 + \beta_l l_{i,t} + \beta_k k_{i,t} + g(i_{i,t}, k_{i,t}, z_{i,t}) + \varepsilon_{i,t}$$

Si se define $\varphi(i_{i,t}, k_{i,t}) = \beta_0 + \beta_k k_{i,t} + g(i_{i,t}, k_{i,t}, z_{i,t})$, se obtiene la ecuación

$$va_{i,t} = \beta_l l_{i,t} + \varphi(i_{i,t}, k_{i,t}, z_{i,t}) + \varepsilon_{i,t}$$

La estimación de esta ecuación por OLS mediante un polinomio de tercer orden para $\tilde{g}(\cdot)$ permite estimar el coeficiente relacionado con la participación del trabajo $\hat{\beta}_l$, y constituye la primera etapa del método OP. Esta estimación sería consistente toda vez que la función $\varphi(\cdot)$ estimada controla por la productividad inobservada, y el término de error no estaría correlacionado con los factores de producción.

Como la ecuación previa no identifica el coeficiente relacionado con el capital (β_k), por otro lado, el método OP asume que la productividad $\omega_{i,t}$ sigue un proceso de Markov de primer orden:

$$\omega_{i,t} = \mathbb{E}[\omega_{i,t}|\Omega_{i,t}] + \xi_{i,t} = \mathbb{E}[\omega_{i,t}|\omega_{i,t-1}] + \xi_{i,t}$$

Con lo cual se obtiene

$$va_{i,t} - \hat{\beta}_l l_{i,t} = \beta_k k_{i,t} + \omega_{i,t} + \varepsilon_{i,t} = \beta_k k_{i,t} + \mathbb{E}[\omega_{i,t}|\omega_{i,t-1}] + \xi_{i,t} + \varepsilon_{i,t}$$

Se toma $\mathbb{E}[\omega_{i,t}|\omega_{i,t-1}] + \xi_{i,t} = h(\omega_{i,t-1}) + \xi_{i,t}$, con $\mathbb{E}[\xi_{i,t}|\omega_{i,t-1}] = 0$.

Así, como $h(\omega_{i,t-1}) = h(\hat{\varphi}(i_{i,t-1}, k_{i,t-1}))$, ello permite estimar obtener la siguiente ecuación:

$$va_{i,t} - \hat{\beta}_l l_{i,t} = \beta_k k_{i,t} + \omega_{i,t} + \varepsilon_{i,t} = \beta_k k_{i,t} + h(\hat{\varphi}(i_{i,t-1}, k_{i,t-1})) - \beta_k k_{i,t-1} + \eta_{i,t}$$

Donde $\eta_{i,t} = \xi_{i,t} + \varepsilon_{i,t}$

Esta ecuación no tiene problemas de endogeneidad dado que $\mathbb{E}[\eta_{i,t}|k_{i,t}, \omega_{i,t-1}] = 0$.

Como la función $h(\cdot)$ puede ser expresada como $h(\omega_{i,t-1}) = h(\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1}) = \gamma_1 \omega_{i,t-1} + \gamma_2 \omega_{i,t-1}^2 + \gamma_3 \omega_{i,t-1}^3$ donde $\omega_{i,t-1} = \hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1}$, la ecuación anterior puede ser re-expresada como:

$$va_{i,t} - \hat{\beta}_l l_{i,t} = \beta_k k_{i,t} + \gamma_1 (\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1}) + \gamma_2 (\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1})^2 + \gamma_3 (\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1})^3 + \eta_{i,t}$$

En una segunda etapa del método OP se estima la ecuación anterior a través de Mínimos Cuadros No Lineales (NLLS) a fin de identificar el parámetro $\hat{\beta}_k$ de la participación del capital en el producto. Alternativamente, también puede utilizarse el Método Generalizado de Momentos basándose en $\mathbb{E}[\eta_{i,t}|k_{i,t}, \omega_{i,t-1}] = 0$.

Finalmente, Olley y Pakes (1996) discuten el posible sesgo de selección debido a la no aleatoriedad de las empresas que abandonan la muestra. Más específicamente, las menos empresas productivas podrían verse obligadas a abandonar el mercado precisamente debido a sus bajos niveles de productividad, dejando así solo a las empresas más productivas.

Se supone que una empresa continúa operando siempre que su nivel de productividad exceda un límite inferior, es decir, $\chi_{it} = 1 \iff \omega_{i,t} \geq \check{\omega}_{i,t}$, donde χ_{it} es una variable

binaria de supervivencia de la empresa, y $\tilde{\omega}_{i,t}$ es un umbral de activación de salida específico de la industria³⁸.

Por lo tanto, proponen un tercer paso en la estimación, haciendo que la productividad sea una función de la variable indicadora de supervivencia además de su valor pasado. Así, la corrección de sesgo propuesta por Olley y Pakes (1996) implica estimar la probabilidad condicional de permanecer activo en el mercado, es decir,

$$Pr_{t+1} \equiv \Pr(\chi_{i,t} = 1 | k_{i,t-1})$$

estimada típicamente a través de un modelo de elección discreta sobre un polinomio de la variable de capital (o vector de variables de estado).

La probabilidad estimada se incluye en la ecuación relevante del siguiente modo:

$$va_{i,t} - \hat{\beta}_l l_{i,t} = \beta_k k_{i,t} + h(\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1}, \widehat{Pr}_t) + \eta_{i,t}$$

2. Método Levinsohn y Petrin (2003)

En el método de Levinsohn y Petrin (2003) se asume que el trabajo ($L_{i,t}$) es un factor perfectamente flexible, a diferencia del método anterior, en lugar de la inversión, se utiliza la demanda por insumos $m_{i,t}$ para definir la función de control. Se asume que $m_{i,t}$ depende del nivel de productividad, del capital, y otras variables no observables $z_{i,t}$: $m_{i,t} = M(\omega_{i,t}, k_{i,t}, z_{i,t})$, siendo estrictamente creciente con la productividad ($\omega_{i,t}$).

De acuerdo con Levinsohn y Petrin (2003) la identificación de la función de control a través de la inversión podría generar estimaciones inconsistentes de la participación del trabajo en la producción, debido a que la inversión puede responder solamente a shocks más persistentes en la productividad, en contraste los insumos pueden responder a cualquier tipo de shock de productividad, y a que la inversión reportada por las empresas puede estar truncada a valores próximos a cero, con ello si $I_{i,t} = 0$, con lo cual no se cumpliría la condición de invertibilidad requerida para estimar la función de control.

En el método LP se tiene:

$$\omega_{i,t} = M^{-1}(m_{i,t}, k_{i,t}, z_{i,t}) = f(m_{i,t}, k_{i,t}, z_{i,t})$$

Donde la productividad ($\omega_{i,j,t}$) sigue un proceso de Markov, la función de producción se convierte en:

$$va_{i,t} = \beta_l l_{i,t} + \varphi(m_{i,t}, k_{i,t}, z_{i,t}) + \eta_{i,t}$$

En la ecuación anterior la función $\varphi(m_{i,t}, k_{i,t}) = \beta_0 + \beta_k k_{i,t} + f(m_{i,t}, k_{i,t}, z_{i,t})$ es aproximada a través de un polinomio de tercer orden (Petrin *et al.*, 2004).

³⁸ Véase Melitz (2003).

La estimación de esta ecuación por OLS constituye la primera etapa del método LP.

El estimador del coeficiente relacionado con la participación del trabajo en el valor agregado $\hat{\beta}_l$ sería consistente, pues la función $\varphi(m_{i,t}, k_{i,t}, z_{i,t})$ estimada controla por la productividad no observada.

De forma análoga al método anterior, la segunda etapa permite identificar la participación del capital β_k en el valor agregado de la producción, mediante la estimación de la siguiente ecuación:

$$va_{i,t} - \hat{\beta}_l l_{i,t} = \beta_k k_{i,t} + h(\hat{\varphi}(m_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1}) + \eta_{i,t}$$

Como $\omega_{i,t} = h(\omega_{i,t-1}) + \xi_{i,t}$ y la función $h(\cdot)$ puede ser expresada como $h(\omega_{i,t-1}) = \gamma_1 \omega_{i,t-1} + \gamma_2 \omega_{i,t-1}^2 + \gamma_3 \omega_{i,t-1}^3$ se obtiene:

$$\begin{aligned} \eta_{i,t} = va_{i,t} - \hat{\beta}_l l_{i,t} - \beta_k k_{i,t} - \gamma_1 (\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1}) \\ + \gamma_2 (\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1})^2 \\ + \gamma_3 (\hat{\varphi}(i_{i,t-1}, k_{i,t-1}) - \beta_k k_{i,t-1})^3 \end{aligned}$$

La ecuación anterior es estimada a través del Método Generalizado de Momentos (GMM, por sus siglas en inglés) a partir de la condición $\mathbb{E}[\eta_{i,t} | k_{i,t}, \omega_{i,t-1}] = 0$, con lo cual es posible identificar la participación del capital en el producto ($\hat{\beta}_k$) a través de la solución al siguiente problema de optimización:

$$\begin{aligned} \min_{\hat{\beta}_k} \sum_t \{ va_{i,t} - \hat{\beta}_l l_{i,t} - \hat{\beta}_k k_{i,t} - \gamma_1 (\hat{\varphi}(m_{i,t-1}, k_{i,t-1}) - \hat{\beta}_k k_{i,t-1}) \\ + \gamma_2 (\hat{\varphi}(m_{i,t-1}, k_{i,t-1}) - \hat{\beta}_k k_{i,t-1})^2 \\ + \gamma_3 (\hat{\varphi}(m_{i,t-1}, k_{i,t-1}) - \hat{\beta}_k k_{i,t-1})^3 \}^2 \end{aligned}$$

3. Método Akerberg, Caves y Frazer (2015)

Akerberg *et al.* (2007) y Akerberg *et al.* (2015) critica los dos métodos anteriores por el problema denominado “dependencia condicional” debido a la colinealidad entre los factores de producción, la cual se genera porque las variables de decisión (trabajo, insumos e inversión) dependen de las mismas variables de estado (capital y productividad). Según los autores, el coeficiente de la mano de obra puede no ser identificable en la primera etapa por los métodos OP y LP, debido a que la demanda de trabajo es elegida de manera óptima al observar su productividad, en cuyo caso se convierte en una función $l_{i,t} = f_{1,t}(\omega_{i,t}, k_{i,t}, z_{i,t})$, que depende de las mismas variables que los insumos $m_{i,t} = f_{2,t}(\omega_{i,t}, k_{i,t}, z_{i,t})$, y la inversión $i_{i,t} = f_{3,t}(\omega_{i,t}, k_{i,t}, z_{i,t})$.

Con ello, se tiene que $l_{i,t} = f_{1,t}(f(m_{i,t}, k_{i,t}, z_{i,t}), k_{i,t})$. Es decir, la entrada de trabajo se convierte en una función de las mismas variables que la función de control, lo que impide la identificación del coeficiente β_l en la primera etapa (Lee et al, 2019).

Dado que las empresas acumulan capital de acuerdo con $k_{i,t} = k(k_{i,t-1}, i_{i,t-1})$, donde las decisiones de inversión son tomadas en el periodo “t-1” y las decisiones sobre la cantidad de trabajo son tomadas entre el periodo “t” y “t-1”, con lo cual la demanda por insumos de la empresa “i” en el periodo “t” puede ser expresada como $m_{i,t} = \tilde{f}_t(\omega_{i,t}, k_{i,t}, l_{i,t}, z_{i,t})$, es decir, depende de la productividad, capital, otras variables observables, y está condicionada al valor del factor trabajo.

De acuerdo con el método ACF el parámetro β_l puede ser estimado a partir de la segunda etapa, utilizando la estimación de la función de control de la primera etapa y el supuesto de que $\omega_{i,j,t}$ sigue un proceso de Markov de primer orden, mediante el Método Generalizado de Momentos (GMM, por sus siglas en inglés).

Bajo el supuesto de monotonidad estricta, $\tilde{f}_t(\cdot)$ es estrictamente creciente con la productividad ($\omega_{i,t}$), es posible invertir la demanda de insumos de la forma: $\omega_{i,t} = \tilde{f}_{i,t}^{-1}(m_{i,t}, k_{i,t}, l_{i,t}, z_{i,t})$, con lo cual se define una función $\tilde{\Phi}_t(m_{i,t}, k_{i,t}, l_{i,t}, z_{i,t})$ que cumple:

$$va_{i,t} = \beta_l l_{i,t} + \beta_k k_{i,t} + \tilde{f}_{i,t}^{-1}(m_{i,t}, k_{i,t}, l_{i,t}, z_{i,t}) + \varepsilon_{i,t} = \tilde{\Phi}_t(m_{i,t}, k_{i,t}, l_{i,t}, z_{i,t}) + \varepsilon_{i,t}$$

La condición de momentos de la primera etapa puede ser planteada en los siguientes términos: $E[\varepsilon_{i,t} | \Omega_{i,t}] = E[va_{i,t} - \tilde{\Phi}_t(m_{i,t}, k_{i,t}, l_{i,t}, z_{i,t}) | \Omega_{i,t}] = 0$.

Ahora, en la primera etapa, la función $\tilde{\Phi}_t(m_{i,t}, k_{i,t}, l_{i,t}, z_{i,t})$ se estima $\hat{\Phi}_t(m_{i,t}, k_{i,t}, l_{i,t}, z_{i,t})$ a través de un polinomio de segundo orden en $l_{i,t}$, $k_{i,t}$ y $m_{i,t}$.

Asumiendo que el término de innovación de $\omega_{i,j,t}$ sigue un proceso de Markov de primer orden se tiene que:

$$va_{i,t} = \beta_0 + \beta_l l_{i,t} + \beta_k k_{i,t} + h(\omega_{i,t-1}) + \eta_{i,t}$$

Donde $\eta_{i,t} = \xi_{i,t} + \varepsilon_{i,t}$

La ecuación anterior poder ser expresada como:

$$\eta_{i,t} = va_{i,t} - \beta_0 - \beta_l l_{i,t} - \beta_k k_{i,t} - h(\omega_{i,t-1}).$$

Además, la función $h(\cdot)$ puede ser aproximada como $h(\omega_{i,t-1}) = \gamma_1 \omega_{i,t-1} + \gamma_2 \omega_{i,t-1}^2 + \gamma_3 \omega_{i,t-1}^3$, y como $\omega_{i,t-1} = \hat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1}$,

se obtiene:

$$va_{i,t} = \beta_0 + \beta_l l_{i,t} + \beta_k k_{i,t} + h(\hat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1}) + \eta_{i,t}$$

Con ello, la segunda etapa del método implica estimar de forma conjunta los parámetros de interés usando la siguiente condición de momentos:

$$\begin{aligned} \mathbb{E}[\eta_{i,t}|\Omega_{i,t-1}] = E \left[va_{i,t} - \beta_0 - \beta_l l_{i,t} - \beta_k k_{i,t} \right. \\ \left. - \gamma_1 \left(\widehat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1} \right) \right. \\ \left. - \gamma_2 \left(\widehat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1} \right)^2 \right. \\ \left. - \gamma_3 \left(\widehat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1} \right)^3 \mid \Omega_{i,t-1} \right] = 0 \end{aligned}$$

Mediante la estimación a través del GMM, los estimadores de los parámetros $\hat{\beta}_l$, $\hat{\beta}_k$ y $\hat{\beta}_m$ deben obtenerse sin problemas de identificación minimizando el valor del correspondiente estadístico muestral, mediante la solución al siguiente problema de optimización:

$$\begin{aligned} \min_{\beta_l, \beta_k} \sum_t va_{i,t} - \beta_0 - \beta_l l_{i,t} - \beta_k k_{i,t} - \gamma_1 \left(\widehat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1} \right) \\ - \gamma_2 \left(\widehat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1} \right)^2 \\ - \gamma_3 \left(\widehat{\Phi}_{t-1} - \beta_l l_{i,t-1} - \beta_k k_{i,t-1} - \beta_m m_{i,t-1} \right)^3 \mid \Omega_{i,t-1} \end{aligned}$$

Para la solución del problema de optimización, las condiciones de ortogonalidad son: $\mathbb{E}[k_{i,t}\eta_{i,t}] = 0$, $\mathbb{E}[l_{i,t-1}\eta_{i,t}] = 0$ y $\mathbb{E}[m_{i,t-1}\eta_{i,t}]$.

La consistencia asintótica del estimador GMM depende de la ausencia de correlación serial de segundo orden en los errores del modelo, lo cual se puede verificar empíricamente a través del contraste de Sargan (1958) y Hansen (1982), donde la hipótesis nula de dicho contraste implica que la estructura de instrumentos impuesta en el modelo es válida.

4. Método de Wooldridge (2009)

Wooldridge (2009) propone una mejora a los métodos OP y LP mediante un procedimiento GMM para estimar todos los coeficientes en la función de producción de forma simultánea mediante el método de variables instrumentales. La estimación con este método – denominado WR - es más sencilla de implementar porque es lineal y no se basa en las estimaciones de $\widehat{\Phi}(\cdot)$ de la primera etapa, evitando así el utilizar el método de *bootstrapping* para calcular los errores estándar de forma robusta³⁹.

³⁹ De acuerdo con Woolridge (2009) el método de estimación de ACF que no intenta identificar ningún parámetro de la función de producción en la primera etapa. Su enfoque permite la posibilidad de que la ecuación de la primera etapa de OP o LP contenga realmente información de identificación de parámetros de las variables como el trabajo.

Además, cabe destacar que la estimación simultánea del método de funciones de control es, en general, más eficiente que la estimación en dos etapas, pero que ambas metodologías son consistentes.

De acuerdo con Woolridge (2009), los estimadores de dos pasos son ineficientes por dos razones: (i) ignoran la correlación contemporánea en los errores de las dos ecuaciones; y (ii) no tienen en cuenta eficientemente la autocorrelación o la heteroscedasticidad en los errores. Mediante el método generalizado de momentos (GMM) se utiliza la correlación entre ecuaciones para mejorar la eficiencia, y la matriz de ponderación óptima permite dar cuenta de la autocorrelación y la heteroscedasticidad.

Específicamente, de forma general, las dos ecuaciones a estimar de forma simultánea en el método WR serían:

$$va_{i,t} = \mathbf{w}_{i,t}\beta_w + \mathbf{x}_{i,t}\beta_x + \tilde{g}(\mathbf{x}_{i,t}, \mathbf{p}_{i,t}) + \varepsilon_{i,t}$$

$$va_{i,t} = \mathbf{w}_{i,t}\beta_w + \mathbf{x}_{i,t}\beta_x + \tilde{h}(\tilde{g}(\mathbf{x}_{i,t}, \mathbf{p}_{i,t})) + \eta_{i,t}$$

Donde:

$\mathbf{w}_{i,t}$: variables de insumos incluyendo el trabajo (en logaritmo).

$\mathbf{x}_{i,t}$: variables de estado como el capital (en logaritmo).

$\mathbf{p}_{i,t}$: variables observables denominadas “proxy” utilizadas para estimar la productividad no observable (en logaritmo). Son inversión en OP e insumos en LP.

$\tilde{g}(\cdot)$: función que relaciona la productividad no observable con las variables de estado y las variables proxy.

$\tilde{h}(\cdot)$: función que especifica el proceso de Markov para la productividad no observable

$$\omega_{i,t} = \mathbb{E}[\omega_{i,t} | \omega_{i,t-1}] + \xi_{i,t} = \tilde{h}(\omega_{i,t-1}) + \xi_{i,t}$$

$$\eta_{i,t} = \xi_{i,t} + \varepsilon_{i,t}$$

En la estimación, el enfoque es tratar las formas funcionales desconocidas utilizando polinomios de enésimo grado en $\mathbf{x}_{i,t}$ y $\mathbf{p}_{i,t}$ para una estimación mediante variables instrumentales utilizando el GMM. En general, se utiliza como instrumentos para la estimación los rezagos (usualmente de orden 1) de las variables incluidas en las ecuaciones, es decir, las variables de estado (capital), libres (trabajo o insumos), proxis (inversión o insumos intermedios), y sus rezagos, así como funciones polinómicas de estas variables.

Las condiciones de ortogonalidad para el GMM para la primera ecuación serían:

$$\mathbb{E}[\varepsilon_{i,t} | \mathbf{w}_{i,t}, \mathbf{x}_{i,t}, \mathbf{p}_{i,t}, \mathbf{w}_{i,t-1}, \mathbf{x}_{i,t-1}, \mathbf{p}_{i,t-1}, \dots, \mathbf{w}_{i,1}, \mathbf{x}_{i,1}, \mathbf{p}_{i,1}] = 0$$

Y para la segunda ecuación serían las siguientes:

$$\mathbb{E}[\xi_{i,t} | \mathbf{w}_{i,t}, \mathbf{x}_{i,t}, \mathbf{p}_{i,t}, \mathbf{w}_{i,t-1}, \mathbf{x}_{i,t-1}, \mathbf{p}_{i,t-1}, \dots, \mathbf{w}_{i,1}, \mathbf{x}_{i,1}, \mathbf{p}_{i,1}] = 0$$

Método de Mollisi y Rovigatti (2018)

Mollisi y Rovigatti (2018) proponen una modificación del estimador de Wooldridge basada en una matriz de instrumentos dinámicos de tipo panel a fin de no perder grados de libertad en la estimación de los parámetros al utilizar variables instrumentales. Así, no se pierde información de los primeros años observados, al utilizar como valores de los instrumentos solo los datos rezagados disponibles en cada momento siguiendo la propuesta de instrumentos dinámicos para panel de Blundell and Bond (1998).

Su enfoque permite aumentar las restricciones de momentos sin perder información, que es una característica altamente útil cuando se trata de conjuntos de datos de panel "N grande, T pequeño" frecuentes en la literatura sobre este tema. De acuerdo con sus autores este estimador se desempeña mejor que el de Wooldridge para datos simulados con un pequeño número de períodos, permite aumentar el tamaño de la muestra en modelos sobreidentificados, y produce resultados más estables.

Anexo 2: Características de la Encuesta Económica Anual (EEA)

La EEA es una investigación estadística destinada a recolectar información económico-financiera de los agentes económicos que nos permiten calcular los grandes agregados macroeconómicos, así como obtener directorios actualizados. Es desarrollada por la Dirección Nacional de Censos y Encuestas del INEI en coordinación con la Dirección Nacional de Cuentas Nacionales. Ambas direcciones realizaron actividades de planeamiento (formulación, definición, contenido del formulario, entre otros), ejecución, procesamiento, obtención de resultados y difusión.

Instrumentos de medición

Para el relevamiento de información de las empresas del Sector Manufactura se han diseñado 02 cuestionarios:

- Formato M: para empresas cuyos ingresos netos anuales sean de 150 UIT hasta 1700 UIT, es decir, sean microempresa de acuerdo a la normatividad vigente.
- Formato D2: para empresas cuyos ingresos netos anuales sean más de 1700 UIT. Incluye pequeña, mediana, y gran empresa.

Finalidad

Garantizar la óptima disponibilidad y confiabilidad de estadísticas útiles en la elaboración de agregados macroeconómicos y otros indicadores económicos sectoriales.

Objetivos

Objetivo general: Recoger información económica financiera confiable y oportuna de las principales características de los agregados macroeconómicos, produciendo estadísticas sobre la evolución de la actividad empresarial.

Objetivos Específicos:

- Obtener información estadística básica de las empresas y sus establecimientos que permita caracterizar la estructura económica del país a nivel nacional y por actividad económica.
- Ser fuente permanente de actualización del Directorio Central de Empresas y Establecimientos.
- Organizar e implementar un banco de datos, con información sistematizada y clasificada, creando así un sistema de información permanente para el análisis sectorial.

Actividades investigadas

Respecto al Clasificador Industrial Internacional Uniforme de todas las Actividades Económicas (CIIU Rev. 4), la Encuesta Económica Anual del Sector Manufactura recoge información de las siguientes Divisiones, Grupos y sus respectivas Clases:

División 10 Elaboración de productos alimenticios (excepto las clases 1010 a la 1050, 1061, 1062, 1072 al 1074)

División 11 Elaboración de bebidas (excepto la clase 1103)

División 12 Elaboración de productos de tabaco

División 13 Fabricación de productos textiles

División 14 Fabricación de prendas de vestir

División 15 Fabricación de productos de cuero y productos conexos

División 16 Producción de madera y fabricación de productos de madera y corcho, excepto muebles; fabricación de artículos

División 17 Fabricación de papel y de productos de papel

División 18 Impresión y reproducción de grabaciones

División 19 Fabricación de coque y productos de la refinación del petróleo (excepto la clase 1920)

División 20 Fabricación de sustancias y productos químicos

División 21 Fabricación de productos farmacéuticos, sustancias químicas medicinales y productos botánicos de uso farmacéutico

División 22 Fabricación de productos de caucho y de plástico

División 23 Fabricación de otros productos minerales no metálicos

División 24 Fabricación de metales comunes

División 25 Fabricación de productos elaborados de metal, excepto maquinaria y equipo

División 26 Fabricación de productos de informática, de electrónica y de óptica

División 27 Fabricación de equipo eléctrico

División 28 Fabricación de maquinaria y equipo n.c.p.

División 29 Fabricación de vehículos automotores, remolques y semirremolques

División 30 Fabricación de otro equipo de transporte

División 31 Fabricación de muebles

División 32 Otras industrias manufactureras

División 33 Reparación e instalación de maquinaria y equipo

División 58 Actividades de Edición

Periodo de referencia

Es el ejercicio contable del año previo.

Forma de recojo de la información

La información que se solicita se rige de acuerdo con las normas establecidas en el Plan Contable General Empresarial.

El método de relevamiento de datos es el auto empadronamiento, mediante la remisión vía el portal web del INEI, de los formularios sectoriales, los mismos que deben ser devueltos debidamente diligenciados vía email a la sede que corresponda.

La distribución de los formularios se realiza a través de la página Web del INEI, en el portal de la encuesta, mediante el cual, los informantes, cuenta con el acceso a consultas sobre su participación, y les permite además ingresar la información, utilizando como herramienta el formulario Web.

El Seguimiento a empresas omisas (no entregaron la información debidamente diligenciada en los plazos establecidos) se realiza previa a la evaluación de empresas omisas, esta actividad se efectúa a través de los siguientes medios:

- Visitas a empresas; las empresas omisas podrían recibir más de una visita, en el caso incumplan con el compromiso de entrega de información, en este caso, se lleva a través de un sistema de seguimiento en línea, el control de dichas visitas.

- Llamadas telefónicas; la empresa omisa será contactada por el operador de empresas, las veces que sea necesario a fin de que cumpla con la entrega de información debidamente diligenciada.

- Correos electrónicos; asimismo se utilizará este medio a fin de garantizar la entrega de información. Para todos los casos antes mencionados se llevará el control de estas acciones mediante el Sistema de Información de la EEA.

El procesamiento de datos es centralizado en la ciudad de Lima y está a cargo de la Dirección Ejecutiva de Censos y Encuestas de Empresas y Establecimientos. El procesamiento se realiza de manera simultánea al recojo de la información, debido a que, una vez ingresada la información por el usuario, se podrá verificar la validez de los datos en la base de datos almacenada en el INEI. TIPO DE DATO.

Unidad de análisis

Las empresas y establecimientos dedicadas a la actividad económica con personería jurídica identificadas por el RUC.

Alcance Temático

A nivel de empresa:

- Identificación de la Empresa, Estados Financieros y TIC
- Identificación y ubicación de la empresa
- Estado de situación financiera
- Estado de resultados por naturaleza
- Movimiento de existencias
- Resumen del movimiento del activo inmovilizado e Intangibles
- Resumen del movimiento de la depreciación, amortización y agotamiento acumulado
- Impuesto de la empresa
- Dividendos por pagar y pagados
- Capital social de la empresa
- Gastos de personal en el año
- Personal ocupado en el año
- Datos del (los) establecimiento (s) de la empresa.

A nivel de establecimiento se recoge variables sobre:

- Identificación y ubicación del establecimiento
- Personal ocupado
- Remuneraciones y otros gastos de personal
- Gastos y cargas diversas de gestión
- Existencias
- Impuestos

- Insumos
- Materias primas y auxiliares
- Producción anual
- Producción intermedia
- Movimiento del activo fijo
- Movimiento de la depreciación acumulada
- Ventas netas e ingresos diversos.

Cobertura geográfica

Se recopila información de las empresas manufactureras a nivel nacional en los 24 departamentos y la Provincia Constitucional del Callao.

Diseño muestral

Universo de Estudio: La población de estudio está conformada por todas las empresas comerciales formales que han desarrollado alguna actividad económica durante el año previo y se encuentran ubicadas dentro del territorio nacional

Unidad estadística: La unidad estadística es la empresa.

Marco Muestral: Directorio Central de Empresas y Establecimientos (DCEE),⁴⁰.

Tipo de muestreo: La muestra es probabilística, estratificada, unietápica e independiente en cada departamento. Existen dos estratos:

- Estrato forzoso: conformado por las empresas de mayor tamaño que explican la mayor parte del valor de producción de cada división de la actividad económica. Este estrato es auto representado, cada empresa de este estrato tiene una probabilidad igual a 1.0 de ser incluida en la muestra.
- Estrato no forzoso (muestral): Conformado por las empresas más pequeñas. Se aplica un muestreo aleatorio simple con selección independiente por división de la CIU Revisión 4. Su cálculo es efectuado utilizando el algoritmo de Lavallée-Hidiroglou⁴¹.

Selección muestral: Las empresas fueron seleccionadas mediante un procedimiento sistemático simple al azar con arranque aleatorio.

⁴⁰ Cabe señalar que la Dirección Nacional de Censos y Encuestas cuenta con el DCEE que constituye una base de datos actualizada en forma permanente por dos tipos de fuentes: los directorios de las últimas investigaciones estadísticas del Sistema Estadístico Nacional y el Registro del Padrón de Contribuyentes de la Superintendencia Nacional de Aduanas y de Administración Tributaria (SUNAT). El Directorio (DCEE) es la base para la construcción del marco muestral y la selección para la ejecución de encuestas a empresas del sector manufactura.

⁴¹ Los algoritmos de Lavallée-Hidiroglou están orientados a la definición óptima de estratos de muestreo. Se elige una cantidad de estratos y el coeficiente de variación en cada estrato. El algoritmo construye límites (boundaries), minimizando tamaños muestrales totales. La construcción de los límites se optimiza a través de procesos iterativos. El algoritmo se aplicó sobre las bases correspondientes a las respectivas divisiones de la CIU, buscando la definición de 2 estratos. La tasa de no-respuesta fue supuesta en 15%.

Nivel de inferencia

Se puede estimar resultados para los niveles de inferencia siguiente:

- Nacional
- División de la Clasificación Industrial Internacional Uniforme Revisión 4.

Anexo 3: Estadística descriptiva - Manufactura Panel 2016-2018

	Tipo	Observaciones	Media	Desviación estándar
Value-added output	Ln S/	2240	16.543	1.329
Trabajo	Ln S/	2240	15.310	1.300
Capital	Ln S/	2240	16.129	1.808
Capital no software	Ln S/	2240	16,106	1,822
Capital software	Ln S/	2240	6,941	5,968
Ratio software	Porcentaje	2240	0,020	0,070
Inversión	Ln S/	2240	11.444	5.164
Materiales	Ln S/	2240	16.183	1.477
Mediana empresa (SBS)	Dummy	2240	0.360	0.480
Gran empresa (SBS)	Dummy	2240	0.536	0.499
Corporativo (SBS)	Dummy	2240	0.104	0.305
Transformación de bienes primarios	Dummy	2240	0.352	0.478
Manufacturas ligeras	Dummy	2240	0.338	0.473
Intensivas tecnologías	Dummy	2240	0.241	0.428
Alimentos, bebidas y tabaco	Dummy	2240	0.069	0.253
Lima y Callao	Dummy	2240	0.880	0.325
Norte	Dummy	2240	0.031	0.173
Centro	Dummy	2240	0.023	0.151
Sur	Dummy	2240	0.044	0.205
Oriente	Dummy	2240	0.021	0.142
Ratio total mujeres	Porcentaje	2240	0.248	0.188
Edad	Número	2240	27.195	17.950
Propiedad extranjera	Dummy	2240	0.069	0.253
Capital intangible	Dummy	2240	10.003	5.169
Sociedad anónima	Dummy	2240	0.909	0.287

Fuente: Panel EEA 2016-19.

Anexo 4: Modelos Censurados, Truncados, y Sesgo de Selección⁴²

En muchos casos la variable dependiente de una regresión lineal solo es observada dentro de un intervalo del rango de sus valores. Existen valores de la variable que no son observados y que, por tanto, no conocemos (en el caso de los que no gastan), y que presentan un valor igual a cero en lugar de reflejar la verdadera valoración del bien.

En general, las estimaciones por MCO con variables censuradas no permite obtener estimadores consistentes, debido a que los datos obtenidos no son representativos de la población. Asimismo, tiene que replantearse la inferencia acerca del significado de los estimadores.

Si los datos vienen de una muestra de N observaciones en donde el vector de regresores x_i es completamente observado, pero y_i no siempre se puede observar. Específicamente, algunos y_i toman el valor 0.

Una interpretación de los valores cero es que son observaciones censuradas. Por ejemplo, supongamos que una familia tenga una demanda latente (no observada) por ropa, denotado por y^* , la cual no es posible observar hasta el momento de la compra, la que ocurre luego de que tal variable latente supere un umbral L . Entonces, observamos y^* si y solo si $y^* > L$. En este caso, el gasto cero puede ser interpretado como una variable censurada por la izquierda cuando $y^* \leq L$.

Por otro lado, así como las variables pueden encontrarse censuradas por la izquierda, también pueden serlo por la derecha o por ambos lados.

Si se tiene una variable dependiente censurada pero los regresores completamente observados, entonces el escenario es similar a tener observaciones perdidas en y . Es decir, tenemos una pérdida de información.

En general, en estos casos, se especifica una variable latente no observada:

$$y^* = x'\beta + \varepsilon$$

Normalmente, se asume una distribución normal para el error, es decir, $\varepsilon \sim N(0, \sigma^2)$.

El valor esperado de esta variable latente será:

$$E(y^*) = x'\beta$$

I. Regresiones con datos truncados

En este caso, los valores de la variable observada y son truncados de acuerdo a la siguiente regla:

$$y = \begin{cases} y^* & \text{si } y^* > L \\ - & \text{si } y^* \leq L \end{cases}$$

⁴² Esta sección se basa en Cameron y Trivedi (2005).

Entonces, la probabilidad de que una observación sea truncada es:

$$\Pr(y^* \leq L) = \Pr(x'\beta + \varepsilon \leq L) = \Phi\left\{\frac{L - x'\beta}{\sigma}\right\}$$

Donde $\Phi(\cdot)$ se refiere a la distribución acumulada normal estándar.

En general, se puede asumir que L es igual a 0, aunque en este caso las estimaciones no permiten identificar el valor de la constante β_1 .

De este modo se tiene, la siguiente expresión:

$$\Pr(y^* \leq 0) = \Pr(x'\beta + \varepsilon \leq 0) = \Phi\left\{\frac{-x'\beta}{\sigma}\right\} = 1 - \Phi\left\{\frac{x'\beta}{\sigma}\right\}$$

Así, la esperanza de la variable observada y será:

$$\begin{aligned} E(y | x, y > 0) &= E(y^* | y^* > 0) = E(x'\beta + \varepsilon_i | x'\beta + \varepsilon > 0) \\ &= E(x'\beta | x'\beta + \varepsilon > 0) + E(\varepsilon | x'\beta + \varepsilon > 0) \\ &= x'\beta + E(\varepsilon | \varepsilon > -x'\beta) \end{aligned}$$

En este caso, se utiliza los resultados siguientes para los momentos truncados de una variable distribuida normal estándar $z \sim N(0,1)$.

i) $E[z|z > c] = \phi(c)/[1 - \Phi(c)]$ y $E[z|z > -c] = \phi(c)/\Phi(c)$,

ii) $E[z^2|z > c] = 1 + c\phi(c)/[1 - \Phi(c)]$

iii) $V[z|z > c] = 1 + c\phi(c)/[1 - \Phi(c)] - \phi(c)^2/[1 - \Phi(c)]^2$

De este modo, la media truncada del error tiene la siguiente esperanza:

$$\begin{aligned} E[\varepsilon | \varepsilon > -x'\beta'] &= \sigma E\left[\frac{\varepsilon}{\sigma} \middle| \frac{\varepsilon}{\sigma} > \frac{-x'\beta}{\sigma}\right] \\ &= \sigma \phi\left(\frac{-x'\beta}{\sigma}\right) / \left[1 - \Phi\left(\frac{-x'\beta}{\sigma}\right)\right] \\ &= \sigma \phi\left(\frac{x'\beta}{\sigma}\right) / \left[\Phi\left(\frac{-x'\beta}{\sigma}\right)\right] \\ &= \sigma \lambda\left(\frac{x'\beta}{\sigma}\right), \end{aligned}$$

Donde $\lambda(z) = \frac{\phi(z)}{\Phi(z)}$ se denomina "Inverso de la ratio de Mills".

II. Regresiones con datos censurados (Modelo Tobit)

En este caso, se construye la variable observada, y , de acuerdo a la siguiente regla:

$$y = \begin{cases} y^* & \text{si } y^* > L \\ L & \text{si } y^* \leq L \end{cases}$$

Entonces, la probabilidad de que una observación sea censurada es similar al caso anterior:

$$\Pr(y^* \leq L) = \Pr(x'\beta + \varepsilon \leq L) = \Phi \left\{ \frac{L - x'_i\beta}{\sigma} \right\}$$

Donde $\Phi(\cdot)$ se refiere a la distribución acumulada normal estándar.

En este caso, también se suele asumir que L es igual a 0.

Se define la siguiente variable binaria d :

$$d = \begin{cases} 1, & y > L \\ 0, & y = L \end{cases}$$

La esperanza de la variable y observada será:

$$\begin{aligned} E[y|x] &= E_d [E_{y|d}[y|d]] \\ &= \Pr[d = 0] \times E[y|d = 0] + \Pr[d = 1] \times E[y|d = 1] \\ &= 0 \times \Pr[y^* \leq 0] + \Pr[y^* > 0] \times E[y^*|y^* > 0] \\ &= \Pr[y^* > 0] \times E[y^*|y^* > 0], \\ &= \Pr[\varepsilon > -x'\beta] \{x'\beta + E[\varepsilon|\varepsilon > -x\beta]\} \end{aligned}$$

Momentos condicionados

Dependiendo del tipo de análisis que se quiera realizar, se debe tomar en cuenta la forma de la esperanza condicional de la variable resultado, ya sea esta la variable potencial (que representa al conjunto de posibilidades observadas y no observadas), o de la variable observada ya sea censurada o truncada (que representan el conjunto de los datos medibles).

Dado que $\Pr[\varepsilon > -x'\beta] = \Pr[-\varepsilon < x'\beta] = \Pr[-\varepsilon/\sigma < x'\beta/\sigma] = \Phi(\beta/\sigma)$ se obtiene las siguientes expresiones para las medias condicionadas de la variable de resultado:

Variable latente:

$$E[y^*|x,] = x'\beta$$

Truncada por la izquierda (en 0):

$$E[y|x, y > 0] = x'\beta + \sigma\lambda(x'\beta/\sigma)$$

Censurada por la izquierda (en 0):

$$E[y|x] = \Phi(x'\beta/\sigma)x'\beta + \sigma\phi(x'\beta/\sigma)$$

Estos resultados indican que, aunque la media condicional original es lineal en los regresores, las medias de la variable censurada o de la truncada no lo son, por lo cual al utilizar MCO los estimadores serán inconsistentes.

De forma similar, las varianzas condicionadas vienen dadas por

Variable latente:

$$V[y^*|x,] = \sigma^2$$

Truncada por la izquierda (en 0):

$$V[y|x, y > 0] = \sigma^2[1 - \omega\lambda(\omega) - \lambda(\omega)^2],$$

Censurada por la izquierda (en 0):

$$V[y|x,] = \sigma^2\Phi(\omega)\{\omega^2 + \omega\lambda(\omega) + 1 - \Phi(\omega)[\omega + \lambda(\omega)]\}^2$$

donde $w = x'\beta$.

Tanto en el caso de variables truncadas como de censuradas existe heterocedasticidad.

Además, la variabilidad se reduce en el caso de truncamiento (como era esperado) pues la varianza es menor a σ^2 .

Efectos marginales

Variable latente:

$$\partial E[y^*|x,]/\partial x = \beta$$

Truncada por la izquierda (en 0):

$$\partial E[y|y > 0|x]/\partial x = \{1 - \omega\lambda(\omega) - \lambda(\omega)^2\}\beta$$

Censurada por la izquierda (en 0):

$$\partial E[y|x]/\partial x = \Phi(\omega)\beta$$

III. Modelos de Sesgo de Selección (Selección muestral)

Existen muchos modelos de selección, desde que hay muchas maneras en que una muestra con selección pueda ser generada. Es difícil saber si se está usando una muestra con selección.

La selección puede deberse a autoselección, siendo el resultado determinado en parte por la elección individual de participar o no en la actividad de interés. También puede

ser resultado de la selección muestral cuando aquéllos que participan en la actividad son deliberadamente sobremuestrados – siendo el caso extremo cuando la muestra es de solo participantes.

Modelo de Sesgo de Selección Bivariado

Es conocido como el modelo Tobit tipo 2, y fue planteado por James Heckman (1979). Se especifican dos variables latentes, y no solo una, a fin de capturar los determinantes tanto de la posibilidad de selección muestral como de los resultados observados en la variable dependiente analizando los determinantes en ambos casos (tanto a nivel de variables observables como de errores o variables no observables). Su aplicación más común tiene que ver con la decisión de trabajar de las personas pues su propensión a hacerlo no es observada, y debe considerarse en el análisis de los determinantes de las horas de trabajo o los ingresos laborales de las personas a partir de las observaciones de estas variables.

Ecuación de participación:

$$y_1 = \begin{cases} 1 & \text{si } y_1^* > 0, \\ 0 & \text{si } y_1^* \leq 0 \end{cases}$$

Ecuación de resultados:

$$y_2 = \begin{cases} y_2^* & \text{si } y_1^* > 0 \\ - & \text{si } y_1^* \leq 0 \end{cases}$$

En este caso, las variables latentes se especifican de la siguiente forma:

$$y_1^* = x_1' \beta_1 + \varepsilon_1,$$

$$y_2^* = x_2' \beta_2 + \varepsilon_2,$$

Los problemas para la estimación de β_2 surgen si es que los errores están correlacionados. El modelo Tobit es un caso especial donde $y_1^* = y_2^*$.

Método de estimación

En este caso, se realiza estimaciones mediante métodos de Máxima Verosimilitud, asumiendo que la distribución conjunta de los errores es normal, que son homocedásticos y que están correlacionados de acuerdo a:

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \sim N \left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \right]$$

El valor de la covarianza de los errores ε_1 y ε_2 se suele expresar de la siguiente forma:

$$\sigma_{12} = \rho \sigma_1 \sigma_2$$

Donde ρ es el coeficiente de correlación.

De este modo, la función de Máxima Verosimilitud viene dada por:

$$L = \prod_{i=1}^n \{\Pr [y_{1i}^* \leq 0]\}^{1-y_{1i}} \{f(y_{2i}|y_{1i}^* > 0)\} \Pr [y_{1i}^* > 0]\}^{y_{1i}},$$

En este caso, el valor esperado de la variable observada y_2 condicionada será:

$$\begin{aligned} E[y_2|xy_1^* > 0] &= E[x'_2\beta_2 + \varepsilon_2|x'_1\beta_1 + \varepsilon_1 > 0] \\ &= x'_2\beta_2 + E[\varepsilon_2|\varepsilon_1 > -x'_1\beta_1], \end{aligned}$$

Si los errores son independientes el último termino es $E[\varepsilon_2] = 0$, y la regresión MCO dará un estimador consistente de β_2 .

En el caso general, de correlación de los errores, Heckman utiliza la relación:

$$\varepsilon_2 = \sigma_{12}\varepsilon_1 + \xi$$

y demuestra lo siguiente:

$$\begin{aligned} E[y_2|x, y_1^* > 0] &= x'_2\beta_2 + E[(\sigma_{12}\varepsilon_1 + \xi)|\varepsilon_1 > -x'_1\beta_1] \\ &= x'_2\beta_2 + \sigma_{12}E[\varepsilon_1|\varepsilon_1 > -x'_1\beta_1], \\ &= x'_2\beta_2 + \sigma_{12}\lambda(x'_1\beta_1) \end{aligned}$$

Asimismo, la varianza de la variable observada y_2 condicionada será:

$$V[y_2|x, y_1^* > 0] = \sigma_2^2 - \sigma_{12}^2\lambda(x'_1\beta_1)(x'_1\beta_1 + \lambda(x'_1\beta_1))$$

En algunos modelos el valor de y_2 no está truncado, sino censurado con el valor 0. En este caso. A partir de este análisis se obtiene el valor esperado no condicionado de la variable y_2 :

$$\begin{aligned} E[y_2|x] &= E_{y_1^*}[E[y_2|x, y_1^*]] \\ &= \Pr[y_1^* \leq 0|x] \times 0 + \Pr[y_1^* > 0|x] \times E[y_2^*|x, y_1^* > 0] \\ &= 0 + \Phi(x'_1\beta_1)\{x'_2\beta_2 + \sigma_{12}\lambda(x'_1\beta_1)\} \\ &= \Phi(x'_1\beta_1)x'_2\beta_2 + \sigma_{12}\phi(x'_1\beta_1), \end{aligned}$$

La variable de la variable y_2 censurada será heterocedástica.

Dada la expresión de las esperanzas, se tiene que, si el valor de σ_{12} es 0 (lo cual implica que ρ es 0) no existirá sesgo de selección en variables no observables, por lo cual es posible la estimación de la esperanza de y_2 a partir de los datos observados. Si su σ_{12} es positivo es necesario utilizar técnicas distintas planteadas por Heckman.

Tradicionalmente, la técnica de estimación utilizada para este tipo de modelos ha sido el Método de Heckman en dos etapas (denominado "Heckit") que se basa en la estimación de dos regresiones: una para la participación, a partir de la cual se obtiene estimaciones del ratio de Mills, y otra para los resultados, donde se incluye el valor del

ratio de Mills a fin de corregir el posible sesgo de selección (y analizar examinar si éste existe).

Sin embargo, actualmente, debido al desarrollo de los métodos computacionales se puede estimar de forma simultánea las dos ecuaciones del modelo selección bivariado, mediante el método de Máxima Verosimilitud, obteniendo estimaciones más eficientes de los parámetros al utilizar mejor la información de la muestra observada.

Efectos marginales

En el caso general del modelo de selección bivariado, se puede unir las variables independientes x_1 y x_2 en un solo conjunto x , y reescribir $x_1'\beta_1 = x'\gamma_1$ y $x_2'\beta_2 = x'\gamma_2$ (los vectores de parámetros γ_1 y γ_2 tienen ceros en las posiciones de las variables no incluidas según corresponda). En este caso, los efectos marginales respecto a cambios en las variables x son:

Variable latente:

$$\partial E[y_2^*|x,]/\partial x = \gamma_2$$

Truncada por la izquierda (en el valor 0):

$$\partial E[y_2|x, y_1 = 1]/\partial x = \gamma_2 - \sigma_{12}\lambda(x'\gamma_1)(x'\gamma_1 + \lambda(x'\gamma_1))$$

Censurada por la izquierda (en 0):

$$\partial E[y_2|x, y_1 = 1]/\partial x = \gamma_1\phi(x'\gamma_1)x'\gamma_2 + \Phi(x'\gamma_1)\gamma_2 - \sigma_{12}x'\gamma_1\phi(x'\gamma_1)\gamma_1$$

Anexo 5: Pruebas de diferencia de medias de la PTF - Manufactura Panel 2016-18

Variable (X)	Media de X=0	Media de X=1	Diferencia 3/
Tamaño (Mediana empresa=1) Obs.	3.1785 522	3.0900 409	-0.089 (0.046)
Tamaño (Gran empresa=1) Obs.	3.1630 487	3.1139 444	-0.049 (0.046)
Tamaño (Corporativo=1) Obs.	3.1024 853	3.5461 78	0.444 *** (0.081)
Región (Lima y Callao=1) Obs.	3.0725 108	3.1484 823	0.076 (0.071)
Antigüedad (Hasta 14 años=1) Obs.	3.1090 676	3.2207 255	0.112 ** (0.051)
Antigüedad (15 a 21 años=1) Obs.	3.1302 720	3.1717 211	0.042 (0.054)
Antigüedad (22 a 26 años=1) Obs.	3.1454 762	3.1134 169	-0.032 (0.059)
Antigüedad (27 a 42 años=1) Obs.	3.1411 778	3.1317 153	-0.009 (0.061)
Antigüedad (Mayor a 42 años=1) Obs.	3.1673 788	2.9869 143	-0.18 *** (0.063)
Gerencia (Femenina 50%≤1) Obs.	3.074024 79	3.145663 852	0.072 (0.082)
Uso de PC ≥ 40% (Sí=1) Obs.	3.1312 419	3.1465 512	0.015 (0.046)
Uso de PC ≥ 40% (Sí=1) Obs.	3.1196 423	3.1562 508	0.037 (0.046)
Uso de Intranet (Sí=1) Obs.	3.1600 509	3.1149 422	-0.045 (0.046)
Uso de Extranet (Sí=1) Obs.	3.1412 754	3.1327 177	-0.008 (0.058)
Uso de Red Local (Sí=1) Obs.	3.1623 77	3.1375 854	-0.025 (0.083)
Uso de Banda Ancha (Sí=1) 1/ Obs.	3.1527 127	3.1375 804	-0.015 (0.066)
Uso de Sistemas de Gestión (Sí=1) Obs.	3.1316 40	3.1399 891	0.008 (0.112)
Uso de PDA/Tablet (Sí=1) Obs.	3.1366 673	3.1473 258	0.011 (0.051)
Uso de Página Web (Sí=1) Obs.	3.2913 182	3.1027 749	-0.189 *** (0.057)
Pago de anuncios por Internet (Sí=1) Obs.	3.1537 673	3.1029 258	-0.051 (0.051)
Compra de software (Sí=1) 2/ Obs.	3.1922 204	3.1248 727	-0.067 (0.055)
Uso de software libre (Sí=1) 2/ Obs.	3.1713 204	3.0911 727	-0.08 * (0.055)

Obs.	563	368	(0.046)
Desarrolla software propio (Sí=1) 2/ Obs.	3.1684 617	3.0829 314	-0.086 * (0.048)
Capacitación TIC (Sí=1) Obs.	3.1532 597	3.1153 334	-0.038 (0.047)
Compra por Internet \geq 0% (Sí=1) Obs.	3.1665 617	3.0683 314	-0.098 * (0.051)
Venta por Internet \geq 0% (Sí=1) Obs.	3.1359 860	3.1839 71	0.048 (0.086)
I&D en nuevos productos (Sí=1) Obs.	3.1480 689	3.1155 242	-0.033 (0.052)
I&D en mejora procesos (Sí=1) Obs.	3.1436 635	3.1309 296	-0.013 (0.049)

Notas:

1. Se considera el uso exclusivo de banda ancha (no utilizar también conexiones de menor velocidad).
2. El software comprende software contable, de ventas, de personal, soporte, entre otros.
3. Errores estándar de la diferencia de medias estimada entre paréntesis

Fuente: Panel EEA 2016-18.

Anexo 6: Estadística descriptiva del uso de las TIC y de inversión en I&D según tamaño de empresa y subsector productivo - Manufactura panel 2016-18

A. Según tamaño de empresa (SBS)

Variable	Mediana empresa	Gran empresa	Corporativo
Porcentaje de trabajadores que usa PC	44,367	53,260	65,378
Porcentaje de trabajadores que usa Internet	45,741	52,059	59,286
Uso de intranet	0,320	0,536	0,845
Uso de extranet	0,111	0,232	0,534
Uso de red local	0,869	0,947	0,996
Uso de banda ancha 1/	0,852	0,901	0,824
Uso de sistemas de gestión	0,945	0,967	0,983
Uso de PDA/Table	0,209	0,316	0,626
Uso de página Web	0,703	0,842	0,912
Pago de anuncios por Internet	0,276	0,294	0,412
Compra software 2/	0,714	0,807	0,866
Software libre 2/	0,332	0,459	0,483
Desarrolla software 2/	0,237	0,404	0,508
Capacitación en TIC	0,259	0,394	0,685
Porcentaje de compras por Internet	3,367	3,313	2,492
Porcentaje de ventas por Internet	2,047	1,803	2,778
Inversión en nuevos productos	0,195	0,284	0,500
Inversión en mejora procesos	0,228	0,350	0,597
Personal permanente en I&D	0,165	0,287	0,517
Unidad permanente en I&D	0,103	0,227	0,462
Inversión en I&D	0,352	0,485	0,723
Monto de inversión en I&D (en log.)	10,652	11,880	14,074
Intensidad de inversión en I&D (respecto a salarios)	0,139	0,108	0,135

B. Según subsector productivo

Variable	Transformación de bienes primarios	Manufacturas ligeras	Intensivas en tecnología	Alimentos, bebidas y tabaco
Porcentaje de trabajadores que usa PC	50,993	47,723	56,396	52,915
Porcentaje de trabajadores que usa Internet	49,860	47,039	54,964	55,689
Uso de intranet	0,514	0,435	0,511	0,567
Uso de extranet	0,229	0,183	0,231	0,311
Uso de red local	0,941	0,911	0,921	0,915
Uso de banda ancha 1/	0,880	0,878	0,860	0,890
Uso de sistemas de gestión	0,979	0,937	0,975	0,933

Uso de PDA/Table	0,292	0,279	0,389	0,268
Uso de página Web	0,854	0,711	0,846	0,799
Pago de anuncios por Internet	0,311	0,279	0,265	0,457
Compra software 2/	0,792	0,738	0,828	0,756
Software libre 2/	0,429	0,395	0,423	0,427
Desarrolla software 2/	0,324	0,383	0,396	0,220
Capacitación en TIC	0,381	0,305	0,446	0,451
Porcentaje de compras por Internet	3,805	2,838	3,085	3,062
Porcentaje de ventas por Internet	1,723	1,406	3,672	0,427
Inversión en nuevos productos	0,271	0,229	0,323	0,341
Inversión en mejora procesos	0,351	0,282	0,357	0,390
Personal permanente en I&D	0,262	0,213	0,337	0,311
Unidad permanente en I&D	0,192	0,164	0,272	0,262
Inversión en I&D	0,461	0,421	0,513	0,488
Monto de inversión en I&D (en log.)	12,028	11,364	12,361	11,831
Intensidad de inversión en I&D (respecto a salarios)	0,124	0,093	0,139	0,153

Notas:

1. Se considera el uso exclusivo de banda ancha (no utilizar también conexiones de menor velocidad).
2. El software comprende software contable, de ventas, de personal, soporte, entre otros.

Fuente: Panel EEA 2016-18.

Anexo 7: Pruebas estadísticas para modelos de determinantes de la PTF a nivel de empresas - Manufactura Panel 2016-18

A. Primera Especificación

Breusch and Pagan Lagrangian multiplier test for random effects

$$tfp_lp_a[iruc,t] = Xb + u[iruc] + e[iruc,t]$$

Estimated results:

	Var	SD = sqrt(Var)
tfp_lp_a	.5147917	.7174898
e	.1791433	.4232532
u	.258473	.5084024

Test: Var(u) = 0

chibar2(01) = 530.90
Prob > chibar2 = 0.0000

Test de Hausman

Test of H0: Difference in coefficients not systematic

chi2(19) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 46.76
Prob > chi2 = 0.0002

B. Segunda Especificación

Breusch and Pagan Lagrangian multiplier test for random effects

$$tfp_lp_a[iruc,t] = Xb + u[iruc] + e[iruc,t]$$

Estimated results:

	Var	SD = sqrt(Var)
tfp_lp_a	.5147917	.7174898
e	.1790287	.4231179
u	.2584891	.5084182

Test: Var(u) = 0

chibar2(01) = 524.16
Prob > chibar2 = 0.0000

Test de Hausman

Test of H0: Difference in coefficients not systematic

chi2(28) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 51.65
Prob > chi2 = 0.0029

Fuente: Elaboración propia en base al panel EEA 2016-18.

Anexo 8: Determinantes de la PTF a nivel de empresas según tamaño y subsector - Manufactura Panel 2016-18

A. Según tamaño de empresas (SBS) – Modelo panel de efectos fijos

VARIABLES	Mediana empresa	Gran empresa	Corporativo
Ratio I&D predicha / remuneraciones	0.200*** (0.074)	0.273 (0.178)	0.676* (0.367)
Logaritmo <i>mark-up</i>	-0.046 (0.237)	-0.034 (0.186)	-0.546*** (0.193)
Antigüedad	-0.018 (0.022)	-0.035* (0.019)	-0.010 (0.013)
Antigüedad al cuadrado	0.000 (0.000)	0.000* (0.000)	0.000 (0.000)
Ratio ejecutivas mujeres	0.251 (0.177)	-0.167 (0.146)	-0.199 (0.281)
Ratio total mujeres	-0.242 (0.300)	0.122 (0.287)	-0.080 (0.559)
Porcentaje uso de PC	-0.039 (0.150)	-0.188 (0.143)	-0.163 (0.254)
Porcentaje uso de Internet	-0.172 (0.153)	0.244* (0.138)	0.381 (0.271)
Uso de intranet	-0.076 (0.077)	-0.057 (0.059)	-0.282** (0.140)
Uso de extranet	-0.063 (0.104)	0.087 (0.069)	0.402*** (0.100)
Uso de red local	-0.140 (0.108)	0.313*** (0.117)	0.026 (0.402)
Uso de banda ancha	-0.165* (0.098)	-0.068 (0.083)	0.154 (0.106)
Uso de sistemas de gestión	0.319** (0.150)	0.088 (0.171)	-0.384 (0.426)
Uso de PDA/Table	-0.058 (0.081)	-0.053 (0.061)	-0.086 (0.095)
Uso de página Web	-0.116 (0.101)	-0.161* (0.089)	-0.007 (0.213)
Pago de anuncios por Internet	0.114 (0.074)	-0.167*** (0.063)	-0.097 (0.107)
Compra software	-0.017 (0.079)	0.004 (0.073)	0.727*** (0.175)
Software libre	-0.057 (0.071)	0.013 (0.057)	-0.072 (0.105)
Desarrolla software	0.027 (0.102)	-0.030 (0.066)	0.040 (0.114)
Capacitación en TIC	0.064 (0.075)	0.006 (0.054)	-0.135 (0.096)
Porcentaje de compras por Internet	-0.053 (0.313)	-0.220 (0.199)	-0.492 (0.803)

Porcentaje de ventas por Internet	0.583 (0.507)	0.113 (0.351)	0.575 (0.378)
Manufacturas ligeras	-0.006 (0.205)	-0.221 (0.255)	0.038 (0.331)
Intensivas en tecnología	0.654** (0.255)	0.032 (0.253)	0.065 (0.327)
Lima y Callao	-1.542*** (0.492)	0.010 (0.286)	-1.357*** (0.426)
Alimentos, bebidas y tabaco		-0.826 (0.513)	
Constant	4.521*** (0.619)	3.465*** (0.512)	4.796*** (0.831)
Observations	744	1,144	226
R-squared	0.00493	0.0220	0.151
RMSE	0.387	0.459	0.316
Prob > F	0.000382	0.0548	0.00165

Notas:

Errores estandar robustos entre paréntesis. *** p<0.01, ** p<0.05, * p<0.1

B. Según subsector - - Modelo panel de efectos fijos

VARIABLES	Transformación de bienes primarios	Manufacturas ligeras	Intensivas en tecnología	Alimentos, bebidas y tabaco
Ratio I&D predicha / remuneraciones	0.241** (0.115)	0.237 (0.165)	0.140* (0.083)	0.801** (0.305)
Logaritmo mark-up	-0.181 (0.183)	-0.136 (0.207)	0.476* (0.285)	-0.667 (0.430)
Antigüedad	-0.032* (0.018)	-0.018 (0.015)	-0.013 (0.026)	0.050 (0.052)
Antigüedad al cuadrado	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Ratio ejecutivas mujeres	0.207 (0.166)	-0.249 (0.167)	-0.255 (0.255)	-0.146 (0.326)
Ratio total mujeres	0.204 (0.339)	0.212 (0.263)	-0.576 (0.633)	-1.665* (0.959)
Porcentaje uso de PC	-0.142 (0.155)	-0.130 (0.149)	0.158 (0.201)	-0.114 (0.464)
Porcentaje uso de Internet	0.173 (0.145)	0.248 (0.152)	-0.149 (0.223)	-0.450 (0.396)
Uso de intranet	-0.087 (0.066)	-0.028 (0.077)	-0.149 (0.092)	-0.269 (0.166)
Uso de extranet	0.004 (0.081)	0.135 (0.084)	0.138 (0.106)	0.011 (0.202)
Uso de red local	0.146 (0.112)	0.337** (0.145)	-0.148 (0.141)	0.115 (0.423)
Uso de banda ancha	0.034 (0.083)	-0.344*** (0.093)	0.085 (0.125)	0.443* (0.263)
Uso de sistemas de gestión	-0.082	0.372**	0.170	0.053

	(0.187)	(0.165)	(0.425)	(0.328)
Uso de PDA/Table	0.099	-0.087	-0.219**	0.013
	(0.066)	(0.077)	(0.092)	(0.166)
Uso de página Web	-0.111	-0.117	-0.070	0.249
	(0.106)	(0.105)	(0.132)	(0.267)
Pago de anuncios por Internet	-0.079	-0.175**	-0.154*	0.127
	(0.068)	(0.079)	(0.092)	(0.186)
Compra software	-0.028	0.022	0.234*	-0.009
	(0.070)	(0.094)	(0.120)	(0.310)
Software libre	-0.060	0.008	0.073	-0.125
	(0.066)	(0.065)	(0.087)	(0.259)
Desarrolla software	-0.041	0.067	0.015	-0.023
	(0.074)	(0.088)	(0.099)	(0.274)
Capacitación en TIC	0.096	-0.072	-0.027	0.011
	(0.059)	(0.070)	(0.091)	(0.175)
Porcentaje de compras por Internet	-0.424**	0.086	-0.502	4.468***
	(0.214)	(0.337)	(0.321)	(1.203)
Porcentaje de ventas por Internet	0.533*	-0.235	-0.071	2.971
	(0.310)	(0.860)	(0.372)	(4.451)
Gran empresa (SBS)	0.256**	0.060	0.307**	-0.322
	(0.102)	(0.129)	(0.145)	(0.405)
Corporativo (SBS)	0.489**	0.325	0.352	-0.117
	(0.212)	(0.264)	(0.311)	(0.539)
Lima y Callao	-0.556**			-0.277
	(0.259)			(0.400)
Constant	3.899***	2.995***	2.736***	2.683**
	(0.486)	(0.427)	(0.695)	(1.134)
Observations	762	704	505	143
R-squared	0.0885	0.116	0.00901	0.00291
RMSE	0.404	0.407	0.438	0.420
Prob > F	0.0128	0.00535	1.96e-06	0.0148

Notas:

Errores estándar robustos entre paréntesis. *** p<0.01, ** p<0.05, * p<0.1

Fuente: Elaboración propia en base al panel EEA 2016-18.