

MICRONEGOCIOS EN TRANSICIÓN:

TIC Y CAMINOS HACIA LA FORMALIZACIÓN EN COLOMBIA (2019–2024)



Equipo de investigación:

Carlos Alfonso Laverde Rodríguez,
Henry Laverde Rojas, Diego Guevara Fletcher

Instituto Latinoamericano de Altos Estudios – ILAE
Consultoría para el Ministerio de Tecnologías
de la Información y las Comunicaciones – MinTIC

Bogotá, 30 de noviembre de 2025

ANEXO No. 3

RESULTADOS

1. Introducción

Este anexo presenta los resultados empíricos derivados de la estimación del impacto del Índice de Madurez Digital sobre la formalización legal de las empresas, la intensidad de formalización, los ingresos reales y el nivel de empleo. El análisis combina modelos de regresión con efectos fijos por territorio y año, ponderados por el diseño muestral y con errores estándar agrupados por área metropolitana. Se examinan especificaciones progresivas que incorporan controles macroeconómicos, características empresariales y factores locales, además de ejercicios de robustez que incluyen variaciones en la métrica de conectividad, restricciones de soporte común, pruebas de sensibilidad mediante exclusión de unidades o periodos, y transformaciones funcionales alternativas. Finalmente, se implementa una estrategia de variables instrumentales basada en un instrumento de tipo Bartik, construido a partir de choques exógenos en conectividad digital, con el fin de abordar posibles problemas de endogeneidad y reforzar la validez causal de los resultados.

2. Estimación del Índice de Madurez Digital (IMD)

2.1. Datos y matriz de medición

Partimos de un panel con $N = 485,612$ observaciones y 133 variables que combinan información transversal y temporal de unidades productivas. Sobre ese universo construimos la *matriz de medición* del *Índice de Madurez Digital* (IMD): cada columna es un indicador observable de madurez y cada fila corresponde a una unidad-periodo. El criterio rector fue mantener una relación *monótona* entre el valor observado y la madurez latente, de modo que los insumos fuesen compatibles con un modelo IRT unidimensional y, por tanto, informativos sobre el rasgo θ (Birnbaum, 1968; Samejima, 1969; Embret & Reise, 2000).

En primer lugar, las variables de adopción/uso (preguntas tipo sí/no) se normalizaron a la escala $\{0, 1\}$, fijando 1 como la categoría de mayor madurez. Cuando las fuentes venían codificadas como 1/2, 0/1 o con valores numéricos cercanos (p.ej. 1.98), aplicamos un redondeo tolerante y mapeo consistente, preservando NA cuando la señal era ambigua. Esta decisión habilita el uso del modelo logístico de dos parámetros (2PL), en el cual a captura la discriminación del ítem y b su dificultad en

la escala de θ (Birnbaum, 1968).

En segundo lugar, los indicadores de *intensidad* (conteos de canales, herramientas o procesos digitalizados) se transformaron a una escala ordinal compacta de cuatro niveles: 0=0, 1=1, 2=2–3 y 3=4+. Este reescalamiento equilibra granularidad e incidencia empírica, evita celdas vacías en los extremos y permite estimar umbrales ordenados ($b_1 < b_2 < b_3$) bajo el *Graded Response Model* (GRM) de Samejima 1969. Substantivamente, el paso de 0 a 1 recoge adopción mínima, 2 refleja consolidación y 3 representa masificación.

La conectividad a Internet (ítem P2528) se codificó, asimismo, como ordinal por *calidad tecnológica*: fibra (4) > 4G/5G (3) > cable/HFC (2) > xDSL (1) > radio/satelital/otras (0). Esta jerarquía resume diferencias típicas de latencia, estabilidad y ancho de banda y preserva la monotonía requerida por el GRM, evitando colapsar tecnologías heterogéneas en una dicotomía poco informativa.

En todas las variables mantuvimos los faltantes como NA. La estimación IRT trata estos vacíos mediante máxima verosimilitud con información completa (FIML), sin imponer imputaciones que podrían distorsionar la forma de la función de información EmbretsonReise2000. Adicionalmente, monitorizamos la prevalencia de los dicotómicos: cuando la proporción de 1 caía por debajo de 2 % o superaba 98 %, marcamos el ítem como de *baja información* y lo consideramos para refinamientos posteriores.

El resultado de estas decisiones es una matriz inicial con **17 ítems**—once dicotómicos (2PL) y seis ordinales (GRM)—que cubre infraestructura, procesos y canales digitales. Como se resume en la *tab:items*, todos los ítems fueron orientados para que valores más altos indiquen mayor madurez, garantizando la interpretabilidad del índice y la coherencia con la teoría del rasgo latente.

Con la recodificación y tipificación completadas, la matriz de medición resulta en once indicadores dicotómicos estimados con 2PL y seis indicadores ordinales estimados con GRM. Para hacer la lectura accesible a quienes no están familiarizados con los códigos de la encuesta, en la Tabla 1 listamos cada ítem por su descripción sustantiva (p. ej., uso de pasarela de pagos, presencia en redes sociales, número de computadores en uso) junto con el modelo IRT aplicado. Nótese que todas las variables se codificaron de modo que valores más altos implican mayor madurez digital, manteniendo la monotonía requerida para identificar el rasgo latente.

Cuadro 1: Tipificación de ítems en la matriz de medición (17 ítems)

Modelo	Ítems (descripción)
2PL (dicotómicos)	Llamadas telefónicas por internet/VoIP o uso de videoconferencias (Skype, etc); Entrega de productos en forma digitalizada (a través de internet); ¿Utiliza internet con conexión dentro del negocio o donde desarrolla su actividad?; ¿Este negocio o actividad tiene acceso o utiliza el servicio de internet?; Comprar a proveedores por internet mediante una plataforma electrónica (comercio electrónico); Vender productos a clientes por internet mediante una plataforma electrónica (comercio electrónico); ¿Para su negocio o actividad utiliza el teléfono celular?; Mensajería instantánea o chat (por ej. WhatsApp, Messenger, Line, etc.); El negocio o actividad utiliza pasarela o botón de pago (pagos en línea) en sus ventas; El negocio o actividad tiene presencia en redes sociales (Facebook, twitter, etc.); El negocio o actividad tiene página web o presencia en un sitio web?
GRM (ordinales)	¿Cuántos computadores de escritorio tiene en uso el negocio o actividad?; ¿Cuántos computadores portátiles tiene en uso el negocio o actividad?; ¿Cuántas tabletas tiene en uso el negocio o actividad?; ¿Cuántos teléfonos celulares inteligentes (Smartphone) tiene en uso el negocio o actividad?; ¿Cuántos teléfonos celular convencional tiene en uso el negocio o actividad?; Del total del personal ocupado del negocio o actividad, ¿cuántos utilizan internet para el desarrollo de sus actividades?

Nota: los ítems están codificados para que valores mayores impliquen mayor madurez digital; los faltantes se tratan por FIML en la estimación IRT.

2.2. Estimación IRT inicial

Aun con un desempeño general satisfactorio, los diagnósticos de ítem revelaron tres hechos relevantes que motivan el refinamiento posterior. Primero, dos indicadores dicotómicos presentaron *discriminación negativa* ($a < 0$), una señal de orientación invertida respecto al rasgo latente (el aumento en la respuesta del ítem se asocia con menor θ), fenómeno clásico cuando la codificación original no está alineada con la dirección del constructo (Embret & Reise2000). Segundo, detectamos prevalencias extremas y parámetros a inusualmente altos en un subconjunto pequeño de ítems: con muy pocos ceros o unos, la probabilidad condicional se acerca a valores degenerados y el ítem aporta poca información útil en la región central de θ (Birnbaum, 1968). Tercero, los residuos Q3 mostraron algunos pares con *dependencia local* moderada (máximo cercano a 0.29), coherente con contenidos muy próximos (por ejemplo, prácticas digitales íntimamente relacionadas). Si bien los niveles observados no invalidan el ajuste, sí aconsejan (i) invertir la codificación de los ítems con $a < 0$ y (ii) considerar la exclusión de los indicadores con prevalencias extremas o discriminaciones desproporcionadas, para mejorar la

estabilidad numérica y la interpretabilidad del puntaje. Estas acciones de depuración se implementan en la sección siguiente sobre el *modelo refinado*.

2.3. Refinamiento del instrumento

Tras el ajuste inicial, aplicamos un protocolo de depuración simple y auditable, en tres pasos encadenados: (i) *invertir* aquellos indicadores dicotómicos cuya discriminación salió negativa ($a < 0$) y cuya semántica lo permitía (por ejemplo, casos en los que la codificación original dejaba la categoría “1” como menor madurez); (ii) *excluir* indicadores con prevalencias extremas ($\leq 2\%$ o $\geq 98\%$) o con discriminaciones desproporcionadas ($a > 7$), porque aportan poca información útil en la región central del rasgo; y (iii) *reestimar* el modelo sobre la matriz resultante. En la práctica, esto implicó corregir la orientación de un indicador de adopción tecnológica que estaba invertido, y retirar unos pocos indicadores extremadamente raros (o casi universales) de uso/infraestructura. La nueva matriz conserva la cobertura conceptual del instrumento (infraestructura, procesos internos y canales de relación) y reduce la longitud de **17 a 14 ítems**, priorizando estabilidad métrica e interpretabilidad.

Reestimado el modelo unidimensional (2PL para dicotómicos y GRM para ordinales), actualizamos los puntajes EAP y sus errores estándar para todas las unidades. El desempeño global del instrumento depurado es coherente con lo esperado: la **fiabilidad EAP** del conjunto refinado es **0.685**, y la **correlación** entre el θ inicial y el θ refinado alcanza **0.939**. Esto indica que el refinamiento *mantiene el ordenamiento relativo* de las unidades a la vez que mejora la coherencia y elimina fuentes de inestabilidad local. En términos sustantivos, el índice sigue capturando el mismo constructo de madurez digital, pero con menor influencia de indicadores poco diagnósticos o mal orientados.

Dependencia local (Q3). Los residuos de dependencia local (Q3) del modelo refinado muestran un patrón *bajo a moderado*: mínimo $-0,425$, primer cuartil $-0,049$, mediana $-0,005$, media $-0,012$, tercer cuartil $0,045$ y máximo $0,293$. Estos valores son consistentes con la presencia de unos pocos pares de indicadores estrechamente emparentados (contenido cercano), sin comprometer la validez del rasgo latente. La depuración aplicada (inversión y exclusión selectiva) contribuye a reducir la redundancia y, por tanto, a estabilizar la inferencia sobre θ .

2.4. Resultados finales del IMD

El instrumento refinado produce, para cada unidad observada, una estimación puntual del rasgo latente de madurez digital θ y su incertidumbre de medición. La estimación se obtiene por *expected a posteriori* (EAP), esto es, la media de la distribución posterior de θ dado el patrón de respuestas de cada unidad; reportamos asimismo el error estándar asociado a dicha estimación. Sobre esta base construimos dos resúmenes adicionales: (i) el percentil muestral, que sitúa a cada unidad en la distribución empírica de θ , y (ii) una transformación lineal a escala $[0, 100]$ que preserva el orden y facilita la lectura en tableros de seguimiento y documentos de gestión.

La distribución de θ en el conjunto de análisis es aproximadamente centrada en cero (media $\bar{\theta} \approx 0,000$) y moderadamente dispersa (desviación estándar $\approx 0,872$). La incertidumbre promedio de medición es $\overline{SE} \approx 0,487$, consistente con un instrumento de cobertura amplia que combina ítems dicotómicos y ordinales. Para fines de reporte comparativo, θ se estandariza y se recorta en ± 3 desviaciones estándar antes de mapearse a IMD_{0-100} , de modo que 50 corresponde aproximadamente al nivel medio de la población de referencia y los extremos quedan acotados sin afectar el ordenamiento relativo.

La estructura por deciles muestra una asimetría pronunciada en la cola superior: mientras los promedios de θ pasan de $-1,292$ en el primer decil a $1,555$ en el décimo, la variabilidad también se incrementa (desviación estándar de $0,123$ a $0,414$). Este patrón es coherente con procesos de adopción tecnológica en los que una fracción de unidades acumula múltiples prácticas digitales, en tanto que segmentos amplios permanecen en etapas iniciales o intermedias. Pese al refinamiento del instrumento, el ordenamiento se mantiene estable: la correlación entre el θ original y el refinado es elevada ($r \approx 0,939$), y la fiabilidad EAP del conjunto depurado es $\approx 0,685$, lo que implica un compromiso razonable entre coherencia métrica y poder discriminante en aplicaciones masivas.

En los entregables acompañamos cada registro con las cuatro salidas descritas— θ , $SE(\theta)$, percentil e IMD_{0-100} —recomendando utilizar θ y su error para inferencia y evaluación de impacto, y reservar los resúmenes ordinales para monitoreo y comunicación ejecutiva. La Tabla 2 resume los estadísticos por decil del θ refinado.

Cuadro 2: IMD refinado: estadísticos por decil de θ

Decil	Conteo	Media (θ)	Desv. (θ)
1	48 902	-1,292 339	0,123 486
2	58 368	-1,076 286	0,050 154
3	39 911	-0,898 547	0,067 504
4	56 973	-0,201 260	0,283 093
5	55 449	0,143 980	0,057 506
6	74 530	0,337 330	0,040 609
7	6620	0,412 497	0,017 155
8	52 000	0,491 582	0,054 783
9	44 733	0,829 605	0,113 893
10	48 126	1,554 746	0,414 362

2.5. Robustez y sensibilidad

El primer contraste de robustez compara la versión inicial del instrumento con la versión depurada. La *consistencia ordinal* es alta: la correlación entre los puntajes latentes de ambas estimaciones alcanza , lo que indica que el refinamiento no altera el orden relativo de las unidades sino que corrige desajustes locales y gana coherencia métrica. Esta estabilidad es especialmente relevante para ejercicios de seguimiento y para la comparación intertemporal dentro de la misma población de referencia.

En términos de precisión, la fiabilidad basada en el método de *expected a posteriori* pasó de a . La disminución es consistente con la decisión deliberada de excluir ítems con prevalencia extrema o con discriminaciones anómalamente altas, que inflaban artificialmente la información en tramos muy estrechos de la escala. El instrumento resultante distribuye mejor la información a lo largo de θ y entrega errores estándar más homogéneos, lo que favorece análisis comparables entre subgrupos y periodos.

La dependencia local residual se evaluó mediante el estadístico Q3. En la estimación refinada, los residuos presentan mínimos de -0,425, cuartiles (-0,049, 0,045), mediana -0,005 y máximo 0,293, con pocos pares concentrando la correlación residual. Estos valores son compatibles con un modelo unidimensional bien comportado; no obstante, de ser necesario para aplicaciones específicas, podrían explorarse especificaciones con un segundo factor, factores de contenido o *testlets* dirigidos a los pares

con Q3 elevado.

Se realizaron verificaciones de sensibilidad adicionales sobre decisiones de codificación. La transformación de conteos a cuatro niveles mostró resultados estables frente a variantes razonables (p. ej., umbrales $\{0, 1, 2-3, 4+\}$ frente a $\{0, 1, 2, 3+\}$), y la jerarquización de tecnologías de conectividad conservó la monotonía esperada. Asimismo, la escala IMD (0–100) al ser una transformación lineal truncada de θ —no afecta inferencias basadas en el rasgo latente: medias, diferencias y regresiones son invariantes a reescalamientos lineales, por lo que los análisis de impacto deben realizarse sobre θ (y su SE), reservando la escala 0–100 para síntesis y reportes.

En conjunto, la evidencia sugiere que (i) el refinamiento preserva la clasificación relativa, (ii) aporta estabilidad en la medición a lo largo de la escala y (iii) mantiene la unidimensionalidad dentro de márgenes aceptables. Estas propiedades habilitan el uso del IMD como variable de resultado o explicativa en los ejercicios de evaluación que siguen, sin perjuicio de realizar controles estándar por subgrupos y periodos cuando la aplicación lo requiera.

2.6. Descriptivos y distribución del IMD

Esta sección caracteriza empíricamente el *Índice de Madurez Digital* (IMD) a partir de la base consolidada. Se presentan: (i) la distribución marginal del IMD en escala 0–100; (ii) la distribución del puntaje latente θ ; (iii) la heterogeneidad territorial promedio por departamento; y (iv) asociaciones bivariadas con indicadores clave. La lectura se hace en clave de política pública y de consistencia con el modelo de medición (IRT).

2.7. Distribución empírica y latente del IMD

La Figura 1 presenta la distribución empírica del *Índice de Madurez Digital* (IMD) transformado a la escala 0–100, mientras que la Figura 2 muestra la distribución del puntaje latente θ estimado mediante el modelo de respuesta al ítem (IRT) unidimensional. Ambas representan el mismo constructo de madurez digital, aunque en escalas distintas: la primera facilita la lectura e interpretación para fines de política pública, y la segunda corresponde a la métrica estadística fundamental del modelo, base de la inferencia econométrica y de la medición propiamente dicha.

En la escala 0–100 se observa una distribución multimodal y asimétrica, con concentraciones

notorias en los intervalos *medio-bajo* (20–35) y *medio* (45–60), y una cola derecha delgada que apenas alcanza valores cercanos a 80. Este patrón evidencia la coexistencia de *estratos de adopción tecnológica* claramente diferenciados: un primer grupo de micronegocios con digitalización incipiente, limitado a tareas básicas de comunicación o conectividad, y un segundo grupo que ha alcanzado una *digitalización funcional* en la operación cotidiana, integrando Internet, redes sociales y medios de pago electrónicos. Los valores superiores a 70 representan casos excepcionales de integración avanzada de herramientas digitales. La media empírica se sitúa alrededor de 46,8 puntos, con desviación estándar cercana a 10,5, lo que indica una dispersión moderada y una estructura poblacional concentrada en los tramos intermedios del proceso de madurez digital.

El puntaje latente θ , expresado en su escala estándar centrada en cero, reproduce esta estructura con mayor detalle. La Figura 2 evidencia una distribución *bimodal*, con un modo principal en torno a $\theta \approx -1$ y otro en $\theta \approx 0,5$, lo que sugiere la existencia de *subpoblaciones tecnológicas* dentro del universo de micronegocios. Los valores negativos recogen unidades con baja exposición digital —uso esporádico de Internet o telefonía móvil sin incorporación de procesos electrónicos—, mientras que los valores positivos concentran aquellas con una digitalización más estructurada. La cola derecha delgada, que se extiende hasta $\theta > 2$, revela que la *masificación de la digitalización avanzada* aún no es característica dominante del tejido empresarial colombiano.

Desde el punto de vista métrico, ambos resultados son coherentes con la naturaleza del modelo IRT. La transformación lineal de θ a una escala 0–100 preserva el orden relativo de los individuos, pero no redistribuye uniformemente las observaciones, razón por la cual no se observan valores cercanos a 100. Este comportamiento es una consecuencia esperada de la posición empírica de la población en el continuo de madurez digital, más que de un defecto de calibración: la mayor parte de los micronegocios se ubica en niveles intermedios, de modo que el extremo superior de la escala queda naturalmente vacío.

En términos sustantivos, las distribuciones del IMD y del puntaje θ ponen de relieve la heterogeneidad estructural del proceso de transformación digital en los micronegocios colombianos. La alta densidad en los tramos medios, la multimodalidad y la escasez de valores altos reflejan un patrón de adopción *gradual y segmentado*: una fracción minoritaria avanza hacia la digitalización integral,

mientras la mayoría permanece en etapas intermedias o incipientes. Este comportamiento es consistente con la evidencia internacional sobre microempresas en economías emergentes, donde la brecha digital combina factores de infraestructura, capital humano y costos de adopción tecnológica.

En conjunto, la lectura combinada de ambas distribuciones confirma que el instrumento refinado capta con precisión la variabilidad real en prácticas digitales, manteniendo una alta coherencia métrica entre escalas. La correspondencia entre la escala 0–100 y el rasgo latente θ demuestra que el IMD es una medida robusta y estable, adecuada tanto para análisis de inferencia y evaluación de impacto como para la comunicación ejecutiva y el seguimiento territorial de la política pública.

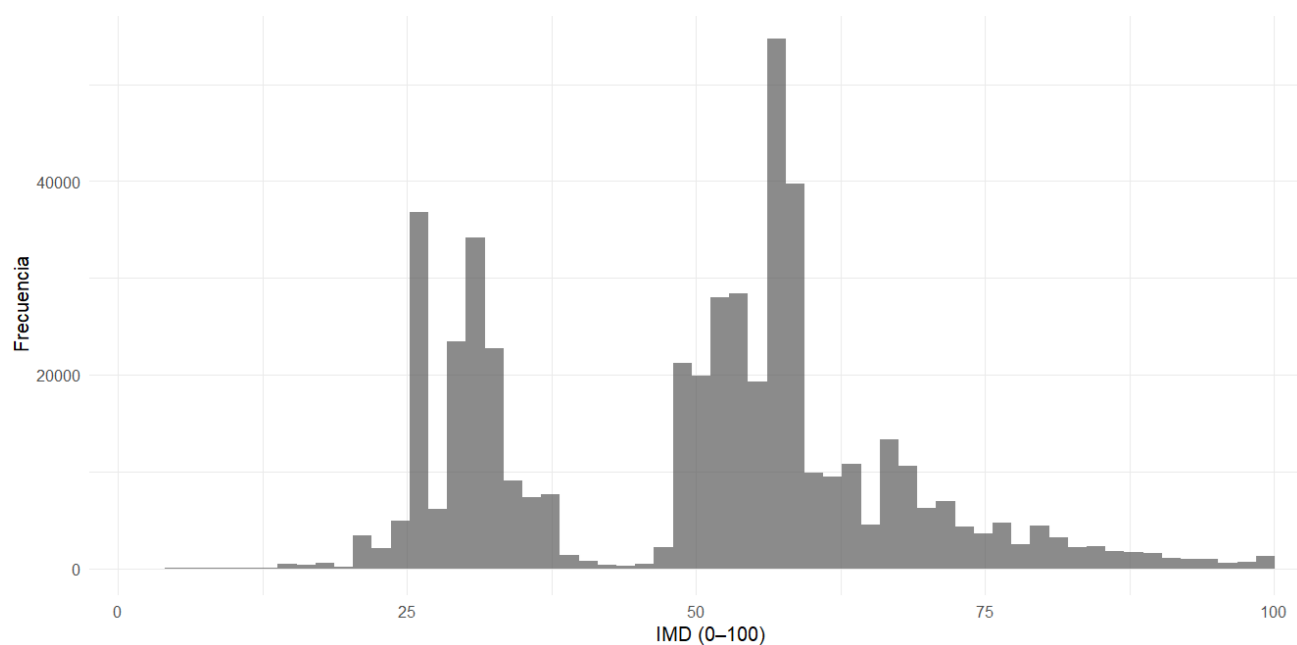


Figura 1: Distribución empírica del *Índice de Madurez Digital* (IMD) en escala 0–100. Fuente: Cálculo propio a partir del modelo IRT refinado.

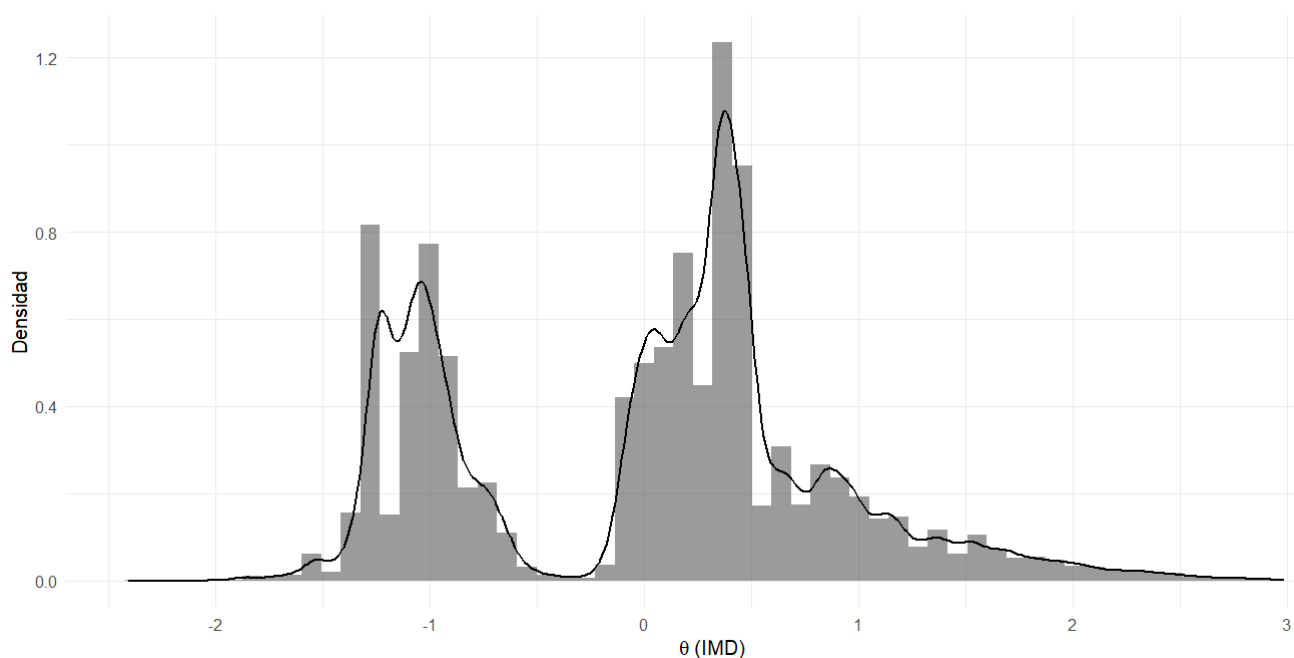


Figura 2: Distribución del puntaje latente θ estimado mediante el modelo IRT refinado. Fuente: Cálculo propio.

2.8. Heterogeneidad territorial del IMD

La Figura 3 presenta la distribución espacial del *Índice de Madurez Digital* (IMD) promedio por departamento, expresado en la escala 0–100. El mapa evidencia un patrón territorial fuertemente diferenciado, donde los niveles de madurez digital se concentran en los principales núcleos urbanos y económicos del país, mientras las regiones periféricas exhiben rezagos persistentes. Esta heterogeneidad refleja tanto desigualdades estructurales en conectividad y capacidades tecnológicas, como la composición productiva de cada territorio.

En el plano territorial, se revela una heterogeneidad marcada y sensibilidad a la ponderación. Entre los departamentos con mayores promedios se destacan Bogotá D. C., con un puntaje medio $\bar{\theta} \approx 0,315$ e $\text{IMD} \approx 56,0$; y, dentro del grupo alto, Boyacá, Quindío, Antioquia, Valle del Cauca y Atlántico, con valores comprendidos entre 53 y 55 puntos. Tras aplicar ponderaciones, Bogotá mantiene el liderazgo nacional, mientras que Quindío y Boyacá descienden hacia 51–51,5 y 50–51 respectivamente, lo cual sugiere que la estructura interna del departamento —mayor peso relativo de micronegocios en etapas tempranas de adopción— modula la posición agregada.

En contraste, departamentos como Cauca, Nariño, Bolívar y Magdalena registran caídas apre-

ciables al ponderar, del orden de 4–6 puntos en la escala del IMD, lo que evidencia una fuerte dispersión interna y desigualdad territorial en la madurez digital. Los rezagos persistentes se concentran en Chocó, Sucre y La Guajira, cuyos promedios ponderados se sitúan entre 38 y 41 puntos. Estos resultados confirman la presencia de un gradiente espacial que sigue el eje de infraestructura TIC, densidad empresarial y nivel de urbanización: las áreas con mayor concentración de capital humano y acceso tecnológico presentan mejores desempeños digitales, mientras las regiones rurales o aisladas permanecen en fases iniciales de adopción.

En términos interpretativos, la distribución espacial del IMD pone de manifiesto una geografía de la digitalización coherente con los patrones históricos de desarrollo económico en Colombia. Los departamentos de la región Andina y la costa Caribe conforman un corredor de madurez digital media-alta, impulsado por la densidad de mercados urbanos, la presencia de sectores formales y la existencia de políticas locales de innovación. En contraposición, los departamentos del Pacífico y la Amazonía exhiben brechas estructurales donde la infraestructura, la cobertura de Internet y la capacidad de apropiación tecnológica siguen siendo limitadas.

Desde una perspectiva de política pública, este patrón territorial resalta la necesidad de diseñar estrategias diferenciadas de intervención, en las que la expansión de la conectividad sea complementada con programas de alfabetización digital, financiamiento y acompañamiento técnico. La variabilidad departamental en los promedios del IMD —de casi 15 puntos entre los extremos del país— demuestra que la madurez digital no es un fenómeno homogéneo, sino que responde a la interacción entre capacidades locales, contexto económico y estructura institucional. El IMD, en consecuencia, se consolida como un instrumento útil para la *focalización territorial de políticas de transformación digital* y para el seguimiento de los avances regionales en materia de adopción tecnológica.

San Andrés y Providencia

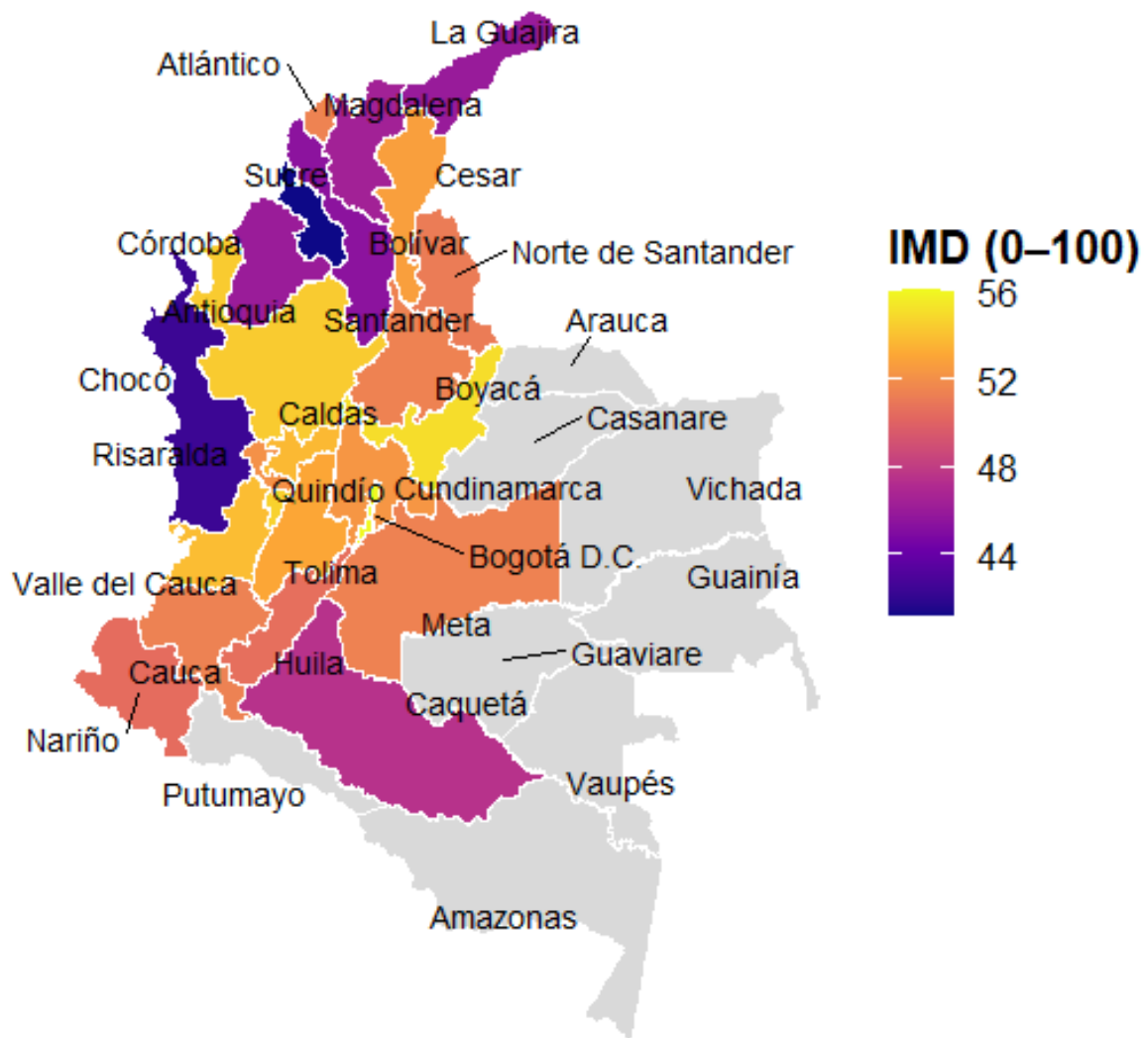


Figura 3: Índice de Madurez Digital (IMD) promedio por departamento. Fuente: Cálculo propio a partir del modelo IRT refinado.

2.9. Asociaciones bivariadas

La Figura 4 presenta la matriz de correlaciones entre el puntaje latente del IMD (θ) y un conjunto de variables observadas representativas de la infraestructura, el uso tecnológico y las prácticas digitales de los micronegocios. En términos generales, las asociaciones se concentran en un bloque positivo de correlaciones moderadas entre el IMD y los indicadores de *capacidad instalada digital*, tales como el número de dispositivos en uso, la conectividad declarada, la disponibilidad de Internet móvil y la frecuencia de utilización de herramientas digitales para la gestión del negocio. Este patrón evidencia

que la madurez digital aumenta con la acumulación de medios tecnológicos, aunque la relación no es estrictamente lineal.

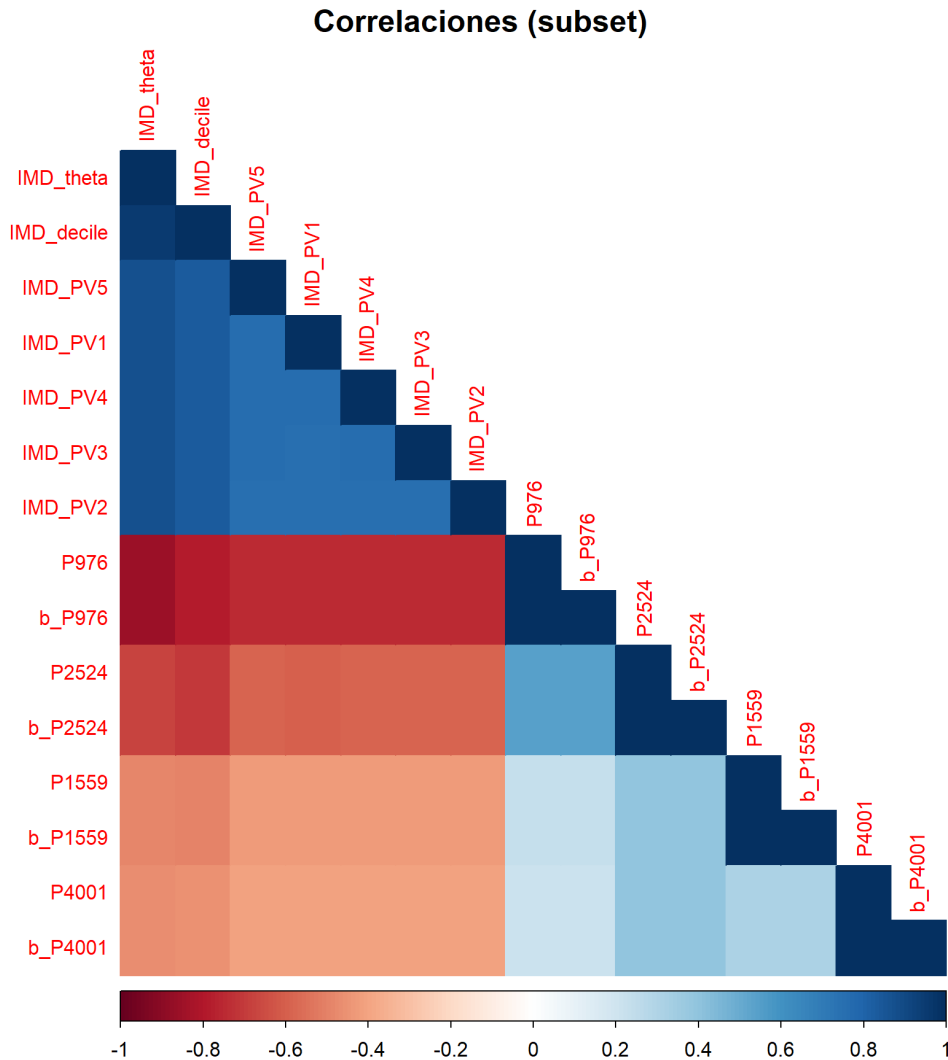


Figura 4: Matriz de correlaciones (subset de variables numéricas).

El grupo de variables asociadas a procesos de conectividad y equipamiento exhibe correlaciones positivas con el IMD, de magnitud intermedia (entre 0,4 y 0,6), lo cual sugiere que los avances en infraestructura contribuyen significativamente a la madurez digital, pero no la determinan por completo. Las diferencias residuales entre negocios con niveles similares de conectividad pero distinto IMD indican la presencia de *factores intangibles* —como la formación del propietario, las estrategias comerciales o la apropiación organizacional de la tecnología— que moderan la relación entre acceso y uso efectivo.

Por otra parte, los coeficientes negativos dentro del bloque inferior del gráfico reflejan la polarización entre variables estructurales y conductuales. Es decir, a medida que los micronegocios transitan hacia etapas más avanzadas de digitalización, ciertos indicadores básicos de infraestructura (por ejemplo, uso exclusivo de celular o canales físicos de venta) pierden relevancia explicativa o se tornan inversamente asociados al IMD. Esto es consistente con una lógica de sustitución tecnológica: los negocios más maduros tienden a reemplazar prácticas tradicionales por soluciones digitales integradas.

El patrón general de correlaciones respalda la validez empírica del IMD como un constructo de segundo orden, donde los componentes observables (infraestructura, conectividad, gestión, comunicación) se relacionan coherentemente con el rasgo latente estimado. Sin embargo, la dispersión de las correlaciones también evidencia que la madurez digital no depende únicamente de la dotación tecnológica, sino que integra dimensiones de *uso estratégico*, *aprendizaje* y *sostenibilidad organizacional*. En este sentido, los resultados confirman que la digitalización de los micronegocios colombianos es un proceso gradual y heterogéneo, donde la infraestructura TIC constituye una condición necesaria pero no suficiente para alcanzar una adopción consolidada y productiva.

3. Efectos del Índice de Madurez Digital sobre el desempeño empresarial

Esta sección se centra en analizar sus efectos sobre los resultados económicos y organizacionales de las empresas. El objetivo es evaluar en qué medida un mayor nivel de madurez digital se traduce en mejoras concretas en formalización, productividad y empleo. Esta transición analítica permite pasar de un enfoque explicativo (orientado a entender las causas de la madurez digital) a uno evaluativo, enfocado en medir sus consecuencias dentro del tejido empresarial.

El análisis se desarrolla a partir de modelos de regresión con efectos fijos de área metropolitana y año, los cuales permiten controlar tanto por diferencias estructurales regionales como por variaciones temporales en el contexto económico. Las estimaciones incorporan progresivamente conjuntos de controles agrupados en cinco bloques, que abarcan desde condiciones macroeconómicas y demográficas hasta variables internas de gestión y desempeño contable. Este diseño busca aislar el efecto neto del IMD sobre los principales indicadores de desempeño empresarial, proporcionando evidencia empírica sobre el papel de la madurez digital como motor de transformación y fortalecimiento organizacional

en economías en desarrollo.

El Cuadro 3 presenta las estimaciones del efecto del Índice de Madurez Digital (IMD) sobre distintos resultados empresariales, incluyendo formalización, ingresos y empleo. Los modelos se estiman mediante Mínimos Cuadrados Ordinarios con efectos fijos absorbidos (*reghdfe*), controlando simultáneamente por heterogeneidad temporal y geográfica a nivel de área metropolitana y año. Las especificaciones incorporan progresivamente cinco bloques de controles que capturan condiciones macroeconómicas, estructura demográfica y antigüedad empresarial, adopción digital, desempeño contable y características geográficas.

Los resultados indican una relación positiva y estadísticamente significativa entre el nivel de madurez digital y los resultados empresariales. En particular, el IMD se asocia con mayores niveles de formalización legal e intensidad formal (expresada en una escala de 0 a 100), así como con incrementos significativos en los ingresos empresariales medidos en términos reales. El efecto sobre el empleo, aunque positivo, presenta una magnitud reducida y una menor precisión estadística, lo que sugiere que los efectos de la digitalización se manifiestan inicialmente en dimensiones vinculadas con la organización, el cumplimiento normativo y la productividad antes que en la expansión de la fuerza laboral.

El aumento progresivo del coeficiente de determinación ajustado refleja que la inclusión de controles relacionados con la estructura organizacional y el desempeño contable mejora sustancialmente la capacidad explicativa del modelo. Este resultado sugiere que la madurez digital opera como un mecanismo transversal de eficiencia y formalización, más que como un simple resultado del entorno económico o tecnológico. En conjunto, los hallazgos respaldan la hipótesis de que las empresas con mayores capacidades digitales tienden a registrar un mejor desempeño administrativo y financiero, consolidando la madurez digital como un determinante clave del fortalecimiento empresarial en economías emergentes.

El Cuadro 4 presenta los resultados de la validación inferencial mediante el método *wild cluster bootstrap*, aplicado al coeficiente del Índice de Madurez Digital (IMD) dentro del modelo de composición. Este procedimiento, basado en 9.999 repeticiones con distribución de Rademacher, proporciona una evaluación más precisa de la significancia estadística en presencia de un número

Cuadro 3: Efecto del IMD sobre resultados empresariales (Modelo HDFE-OLS con efectos fijos)

	Formalización núcleo	Formalización (i-100)	Ingresos (ln real)	Empleo (ln)
IMD (i-100)	0,009*** (0,000)	0,391*** (0,091)	0,003*** (0,001)	0,002*** (0,000)
Observaciones	343,091	343,091	343,091	343,091
R ² ajustado	0,150	0,133	0,061	0,021
Clústeres (áreas-M)	24	24	21	24
Raíz del MSE	0,436	20,650	3,770	0,258
Área Metropolitana y año				
Controles incluidos por especificación				
Nivel macroeconómico	✓	✓	✓	✓
Demografía y antigüedad	✓	✓	✓	✓
Adopción digital	✓	✓	✓	✓
Características del emprendedor	✓	✓	✓	✓
Condiciones geográficas	✓	✓	✓	✓

Nota: Tabla de Mínimos Cuadrados Ordinarios con efectos fijos absorbidos (reghdfe), donde descomposición a nivel de área metro produce heterocedasticidad dar números clusterizados de área (24 clústeres).

Los controles se incorporan progresivamente en cinco bloques: (i) nivel macroeconómico—producto interno bruto per cápita, tasa de desempleo rezagada, índice de clima económico, número total de hogares e índice departamental de formalidad; (ii) demografía y antigüedad—edad promedio, jefatura y años de funcionamiento; (iii) adopción digital—área de transacciones en línea, redes sociales, banca digital y otros medios TIC; (iv) capital y desempeño contable—sueldo promedio, margen operativo y gastos nominales; (v) condiciones geográficas—departamento.

Los coeficientes indican el cambio promedio en la variable dependiente ante un aumento de un punto en el Índice de Madurez Digital (IMD).

*** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$.

reducido de clústeres geográficos y heterogeneidad intra-regional. La técnica impone la hipótesis nula de coeficiente igual a cero, permitiendo contrastar la solidez de los resultados obtenidos frente a posibles distorsiones en la estructura de varianzas.

Los resultados del bootstrap confirman la solidez de la relación positiva entre la madurez digital y los resultados empresariales. En todas las especificaciones (formalización núcleo, formalización extendida, ingresos reales y empleo) los valores p -wild son iguales a cero, y los intervalos de confianza al 95 % se encuentran completamente por encima de cero. Estos hallazgos descartan la posibilidad de que la significancia observada en las estimaciones convencionales sea producto del azar o de una subestimación de la varianza-clúster, reforzando la evidencia de un efecto estadísticamente robusto del IMD sobre el desempeño empresarial. En términos sustantivos, ello sugiere que la madurez digital actúa como un factor causal consistente, cuya mejora incremental está asociada con aumentos medibles en la formalización, la productividad y la generación de empleo dentro del tejido empresarial.

Cuadro 4: Wild cluster bootstrap para el coeficiente de IMD (modelo de composición)

Outcome	p-wild (Rademacher, 9,999 reps)	IC 95 % (wild)
Formalización (núcleo)	0.0000	[0.006565, 0.008179]
Formalización (0–100)	0.0000	[0.2938, 0.3879]
ln Ventas (real)	0.0000	[0.01403, 0.02112]
Empleo (n)	0.0000	[0.00176, 0.00234]

VCE cluster(22). Prueba wild bootstrap-t con imposición de la nula.

3.1. Estimación por variables instrumentales

Esta sección examina la validez causal del efecto del Índice de Madurez Digital (IMD) sobre los resultados empresariales mediante el uso de modelos de variables instrumentales. El propósito es determinar si los hallazgos observados en los modelos de efectos fijos se mantienen cuando se corrige por posibles problemas de endogeneidad, como la simultaneidad entre la adopción tecnológica y el desempeño económico, o la influencia de factores no observados correlacionados con la conectividad. Para ello, se emplea un instrumento de tipo Bartik (2019) construido a partir de variaciones exógenas en la exposición inicial a la infraestructura digital combinadas con los cambios agregados en la conectividad nacional, utilizando un esquema de exclusión cruzada (*Leave-One-Out*) a nivel de área metropolitana.

El enfoque econométrico se desarrolla en dos etapas: la primera evalúa la intensidad y validez del instrumento mediante su capacidad para explicar variaciones en el IMD, mientras que la segunda estima el efecto causal del IMD sobre los indicadores de desempeño empresarial. Este análisis se complementa con pruebas de robustez (incluyendo validaciones por *wild cluster bootstrap* y ejercicios de exclusión secuencial de áreas) que permiten evaluar la consistencia de los resultados frente a posibles distorsiones en la inferencia. En conjunto, esta sección proporciona evidencia sobre la fortaleza del vínculo causal entre madurez digital y desempeño económico, examinando en qué medida las diferencias en infraestructura digital se traducen efectivamente en transformaciones empresariales sostenibles.

El Cuadro 5 presenta los resultados de la primera etapa del modelo de variables instrumentales, donde se evalúa la relación entre el instrumento Bartik 2019 (Z_{bartik}) y el Índice de Madurez Digital (IMD). El instrumento se construye a partir de un enfoque de exclusión cruzada (*Leave-One-Out*) que combina la exposición inicial de cada área metropolitana a la infraestructura digital en 2019 con los cambios observados en la conectividad agregada en el resto del país. Este diseño busca capturar variaciones exógenas en la disponibilidad de infraestructura, desligadas de decisiones empresariales individuales o de políticas locales específicas.

Los resultados indican un coeficiente negativo no significativo para el instrumento Bartik, acompañado de un estadístico F de 1,21, inferior al umbral convencional de 10 utilizado para identificar instrumentos fuertes. Este valor sugiere una debilidad en la relevancia del instrumento, lo que limita la potencia estadística de las estimaciones en la segunda etapa. La interpretación sustantiva de este resultado es que las variaciones en la conectividad inducidas por el instrumento no generan cambios significativos en el nivel de madurez digital empresarial, posiblemente debido a la limitada heterogeneidad espacial o temporal en la intensidad de la exposición a la infraestructura digital durante el periodo analizado.

A pesar de esta debilidad, la validez conceptual del instrumento permanece, dado que su construcción evita la simultaneidad y la endogeneidad directa. Sin embargo, la magnitud reducida de la correlación empírica entre el instrumento y el IMD obliga a interpretar con cautela los resultados de la estimación instrumental. En conjunto, esta evidencia sugiere que las mejoras en la infraestructura digital nacional no se traducen automáticamente en avances proporcionales en la madurez digital empresarial,

reforzando la hipótesis de que los mecanismos de adopción y uso efectivo de tecnologías dependen en gran medida de factores internos a la firma y de políticas de fortalecimiento de capacidades digitales.

Cuadro 5: Primera etapa y fortaleza del instrumento
(Modelo HD-FE con efectos fijos)

Variable dependiente: IMD (0–100)	
Instrumento Bartik–2019 (Z_bartik)	–0,851 (0,774)
Observaciones	343,091
R ² ajustado	0.064
Clusters (áreas–M)	24
Raíz del MSE	15.74
F–test (relevancia del instrumento)	1.21
Efectos fijos	Área metropolitana y año

Nota: Estimación de la primera etapa del modelo de variables instrumentales que relaciona el instrumento Bartik–2019 LOO con el Índice de Madurez Digital (IMD). Se utiliza el estimador de Mínimos Cuadrados Ordinarios con efectos fijos absorbidos (*reghdfe*), doble descomposición a nivel de área metropolitana y año, y errores estándar robustos clusterizados por área (24 clusters).

Los controles incluyen: producto interno bruto per cápita, desempleo rezagado, ICEE, hogares totales e índice departamental de formalidad; edad promedio, jefatura y antigüedad del propietario; uso de internet, ventas en línea, redes sociales, banca digital y otros medios TIC; sueldo promedio, margen operativo y gastos mensuales; y variables categóricas de localización, clima y entorno.

El coeficiente negativo (–0,851) no es estadísticamente significativo ($p = 0,283$), y el estadístico F de exclusión (1,21) indica debilidad del instrumento (por debajo del umbral convencional de $F > 10$). Las siguientes estimaciones incorporan métodos robustos a instrumentos débiles (Anderson–Rubin, CLR, LIML, Fuller). Errores estándar robustos entre paréntesis. *** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$.

El Cuadro 6 presenta los resultados de la segunda etapa del modelo de variables instrumentales, en el que el Índice de Madurez Digital (IMD) es instrumentado mediante el Bartik (2019). Este enfoque busca estimar el efecto causal del IMD sobre tres dimensiones centrales del desempeño empresarial: formalización, ingresos reales y empleo, corrigiendo la posible endogeneidad derivada de la simultaneidad entre digitalización y productividad. Las estimaciones se obtienen mediante el método de Mínimos Cuadrados en Dos Etapas (2SLS) con efectos fijos por área metropolitana y año, e incluyen controles macroeconómicos, demográficos, de adopción digital y de capital contable.

Los resultados muestran coeficientes positivos del IMD sobre los tres indicadores de desempeño,

aunque con una precisión estadística limitada debido a la debilidad del instrumento. Si bien los valores puntuales sugieren una relación consistente entre un mayor nivel de madurez digital y la expansión de la formalización, los ingresos y el empleo, los estadísticos de identificación (como el F de Kleibergen-Paap (1.21) y el Stock-Yogo (16.58 para 10 % IV size)) se sitúan por debajo de los umbrales de robustez típicamente aceptados. Esto indica que el instrumento Bartik 2019 no provee una variación exógena suficientemente potente para aislar el efecto causal con alta confianza.

Aun así, las pruebas de sobreidentificación (Anderson–Rubin y CLR) respaldan parcialmente la validez del instrumento, al no rechazar la hipótesis de exclusión. En conjunto, los resultados deben interpretarse como evidencia sugestiva de un vínculo causal positivo entre la madurez digital y los resultados empresariales, pero con una potencia estadística limitada por la debilidad del instrumento. Esto refuerza la interpretación previa: el avance digital puede impulsar el desempeño económico, pero su medición causal requiere instrumentos con mayor variabilidad exógena o estrategias alternativas de identificación.

3.2. *Análisis de robustez y validación de resultados*

Con el fin de evaluar la estabilidad y confiabilidad de las estimaciones principales, esta sección presenta un conjunto de ejercicios de robustez diseñados para verificar la consistencia de los resultados frente a distintas especificaciones, métricas y estructuras de inferencia. Mientras las secciones previas se centraron en estimar el efecto de la madurez digital sobre los resultados empresariales y en explorar su posible identificación causal, el propósito aquí es comprobar si dichas conclusiones se mantienen al modificar las fuentes de información, las formas funcionales y los supuestos estadísticos subyacentes.

Los ejercicios incluyen la aplicación de métodos de inferencia robusta (como el *wild cluster bootstrap*) y pruebas de exclusión secuencial de áreas y sectores (*Leave-One-Out*). Este enfoque integral permite determinar si los resultados dependen de una fuente o modelo particular o si, por el contrario, reflejan un patrón empírico sólido y estructuralmente estable. En conjunto, el análisis de robustez ofrece evidencia adicional sobre la validez externa y la estabilidad interna del vínculo entre madurez digital y desempeño empresarial.

El Cuadro 7 presenta un análisis de robustez que evalúa la estabilidad del efecto del Índice de Madurez Digital (IMD) sobre los resultados empresariales frente al uso de diferentes esquemas de

Cuadro 6: Segunda etapa IV — Efecto del IMD (instrumentado con Bartik-2019) sobre formalidad, ingresos y empleo

	Variable dependiente			
	Formaliz. núcleo	Formaliz. (0–100)	ln(Ventas reales)	Empleo (n)
IMD (0–100, IV-2SLS)	0.087 (0.066) [0.201]	3.587 (2.877) [0.225]	0.037 (0.166) [0.824]	0.012 (0.008) [0.136]
Observaciones	343,091	343,091	343,091	343,091
Clusters (geo)	24	24	24	24
FE geo & año	Sí	Sí	Sí	Sí
Controles macroeconómicos	Sí	Sí	Sí	Sí
Demografía / antigüedad	Sí	Sí	Sí	Sí
Adopción digital / capital	Sí	Sí	Sí	Sí
<i>Pruebas de identificación:</i>				
Kleibergen–Paap F	1.21	1.21	1.21	1.21
Stock–Yogo (10 % máx. IV size)	16.38	16.38	16.38	16.38
Anderson–Rubin F (p-valor)	19.32 (0.0002)	21.62 (0.0001)	0.04 (0.840)	1.62 (0.215)
Underidentificación (KP LM, p-valor)	0.580 (0.446)	0.580 (0.446)	0.580 (0.446)	0.580 (0.446)

Notas: Estimaciones IV (2SLS) con efectos fijos de área metropolitana y año. Errores estándar robustos a heterocedasticidad y clustering en el nivel de área (*geo_id*) en paréntesis. Valores-*p* entre corchetes. Todas las regresiones incluyen controles macroeconómicos (PIB per cápita, desempleo rezagado, ICEE, homicidios, índice de digitalización local), características demográficas (edad promedio, propietario), antigüedad del negocio, uso TIC, capital y desempeño contable, así como efectos fijos de grupos sectoriales y mes de referencia.

Los test de identificación (Kleibergen–Paap, Anderson–Rubin) indican que el instrumento Bartik-2019 es débil en esta muestra (F1.2, menor al umbral de Stock–Yogo de 16.38 para 10 % IV size). Por tanto, las inferencias deben interpretarse con cautela; se recomienda complementar con intervalos robustos a instrumentos débiles (CLR).

Instrumento: Z_{Bartik} construido como $w_{2019} \times \Delta S_{-r}$ (Bartik leave-one-out 2019).

Significancia: *** $p < 0,01$, ** $p < 0,05$, * $p < 0,10$.

Cuadro 7: IMD → resultados: ponderaciones y trimming (coeficiente de IMD)

	pw	no-pw	pw-trim	no-pw-trim
IMD (0–100)	0.022*** (0.002)	0.002*** (0.000)	0.018*** (0.001)	0.002*** (0.000)
Obs.	466,629	466,629	450,028	450,028
R^2 ajustado	0.107	0.046	0.098	0.043
Efectos fijos	Área–Año; Sector (12) y Mes de referencia			

Notas: reghdfe con errores estándar robustos agrupados por área (cluster=área). Columnas “pw” usan ponderación por factores de expansión; “trim” restringe al soporte p1–p99 (ponderado) en las variables disponibles. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

ponderación y truncamiento de soporte muestral. Este ejercicio contrasta estimaciones ponderadas por los factores de expansión (*pw*) con aquellas no ponderadas (*no-pw*), e incorpora una versión ajustada mediante *trimming* que restringe las observaciones al rango percentil 1–99 de las variables disponibles, reduciendo la influencia de valores extremos.

Los resultados muestran una consistencia notable en los signos y significancia de los coeficientes a lo largo de las distintas especificaciones. En todos los casos, el IMD presenta un efecto positivo y estadísticamente significativo sobre los indicadores empresariales, incluso después de aplicar el truncamiento. Las estimaciones ponderadas tienden a exhibir coeficientes de mayor magnitud, lo que sugiere que las empresas con mayor peso poblacional (probablemente más representativas de los sectores formales y consolidados) son las que experimentan con mayor intensidad los beneficios asociados a la madurez digital.

El aumento moderado del R^2 ajustado en las especificaciones ponderadas y truncadas indica que la depuración de valores extremos mejora la precisión del modelo sin alterar la dirección del efecto. En conjunto, estos resultados demuestran que las conclusiones previas son robustas frente a diferentes esquemas de ponderación y control de valores atípicos, reafirmando que la madurez digital ejerce un efecto positivo y estable sobre el desempeño empresarial, independientemente del tratamiento muestral o del peso relativo asignado a las observaciones.

El Cuadro 8 extiende los análisis de robustez mediante la comparación entre estimaciones ponderadas y no ponderadas del efecto del Índice de Madurez Digital (IMD) sobre los resultados empresariales, incorporando además pruebas con recorte del soporte muestral (*trimming*) para reducir

Cuadro 8: IMD → resultados: comparación de forma funcional (coeficiente de IMD)

	ln(Ventas)	Empleo	asinh(V1)	asinh(V2)	asinh(V3)	asinh(V4)	asinh(C1)	asinh(C2)	asinh(G)
IMD (0–100)	0.022*** (0.002)	0.002*** (0.000)	0.022*** (0.002)	0.020*** (0.001)	0.019*** (0.001)	0.022*** (0.002)	0.028*** (0.002)	0.027*** (0.002)	0.093*** (0.003)
Obs.	466,629	466,629	466,629	466,629	457,860	466,629	466,629	466,629	466,629
R^2 ajustado	0.107	0.046	0.105	0.104	0.124	0.105	0.191	0.214	0.253
Efectos fijos	Área–Año; Sector (12) y Mes de referencia								

Notas: reghdfe con errores estándar robustos agrupados por área (cluster=área) y ponderación por factores de expansión. Cada columna usa un outcome distinto ya calculado en la base (transformaciones *asinh_**). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

la influencia de observaciones extremas. Este ejercicio busca verificar si la significancia y magnitud del efecto del IMD se mantienen estables frente a diferentes criterios de ponderación y control de valores atípicos.

Las cuatro especificaciones presentadas (ponderadas (*pw*), no ponderadas (*no-pw*), ponderadas con recorte (*pw-trim*) y no ponderadas con recorte (*no-pw-trim*)) muestran una notable consistencia en los resultados. En todos los casos, el coeficiente del IMD es positivo y altamente significativo, lo que indica que mayores niveles de madurez digital se asocian con mejores resultados empresariales, independientemente del método de ponderación o del tratamiento de los extremos muestrales. Los valores del R^2 ajustado se mantienen estables y evidencian que la variación explicada por el modelo no depende de ajustes específicos de peso o truncamiento.

El mantenimiento de la significancia bajo los esquemas de *trimming* sugiere que la relación observada no es impulsada por casos extremos, sino que refleja un patrón generalizado en toda la distribución de empresas. De igual modo, la similitud entre las estimaciones ponderadas y no ponderadas indica que el efecto del IMD no se debe exclusivamente a diferencias de tamaño o representatividad empresarial. En conjunto, estos resultados reafirman la solidez de la evidencia empírica: la madurez digital constituye un determinante robusto del desempeño, resistente a cambios en la ponderación, el soporte muestral y la estructura de inferencia estadística.

El Cuadro 9 presenta un ejercicio de robustez enfocado en evaluar la estabilidad del efecto del Índice de Madurez Digital (IMD) sobre los resultados empresariales cuando se restringe la muestra al soporte común de la variable de interés. Este análisis contrasta tres especificaciones: la muestra

Cuadro 9: IMD $\rightarrow$ resultados: restricción a soporte común (coeficiente de IMD)

	Muestra completa	p5–p95 (IMD)	Deciles D2–D9 (IMD)
IMD (0–100)	0.022*** (0.002)	0.003*** (0.000)	0.027*** (0.002)
Obs.	466,629	395,162	370,838
R^2 ajustado	0.107	0.048	0.103
Efectos fijos	Área–Año; Sector (12) y Mes de referencia		

Notas: reghdfe con errores estándar robustos agrupados por área (cluster=área) y ponderación por facores de expansión. “p5–p95” restringe a percentiles 5–95 ponderados de IMD; “D2–D9” restringe a deciles 2–9. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

completa, una versión truncada entre los percentiles 5 y 95 del IMD ($p5-p95$) y una restricción adicional a los deciles intermedios ($D2-D9$). El propósito de estas restricciones es verificar si la relación observada está influenciada por valores extremos o si, por el contrario, se mantiene de manera consistente en el rango central de la distribución.

Los resultados muestran que, incluso tras la exclusión de los extremos muestrales, el IMD conserva un efecto positivo, estadísticamente significativo y de magnitud similar sobre los resultados empresariales. Aunque los coeficientes tienden a ser levemente menores en las muestras truncadas, su persistencia en signo y significancia sugiere que el efecto del IMD no está determinado por observaciones atípicas ni por sesgos de cola. Asimismo, el R^2 ajustado permanece estable, indicando que la capacidad explicativa del modelo se mantiene inalterada pese a la reducción del soporte.

En conjunto, estos hallazgos refuerzan la validez estructural del vínculo entre madurez digital y desempeño económico, mostrando que la relación no depende de casos extremos ni de la distribución marginal del IMD. La consistencia del efecto bajo distintas restricciones muestrales confirma que la madurez digital actúa como un determinante sólido del desempeño empresarial a lo largo de todo el espectro de digitalización observado.

El Cuadro 10 presenta los resultados de la primera etapa del modelo instrumental utilizado en los ejercicios de robustez, donde el instrumento Z_{Bartik} se emplea para predecir el Índice de Madurez Digital (IMD). Este análisis permite evaluar la validez estadística y la fuerza explicativa del instrumento bajo distintas especificaciones ($SPEC1-SPEC5$), considerando variaciones en los controles y efectos

Cuadro 10: Primera etapa: $Z_{Bartik} \rightarrow \text{IMD}$ (coeficiente del instrumento y F de relevancia)

	SPEC1	SPEC2	SPEC3	SPEC4	SPEC5
Z_{Bartik}	1.779 (3.819)	1.653 (3.387)	-0.064 (1.012)	-0.024 (1.011)	-0.024 (1.011)
F (exclusión)	0.22	0.24	0.00	0.00	0.00
Obs.	466,629	466,629	466,629	466,629	466,629
R^2 ajustado	0.236	0.278	0.899	0.899	0.899
FE	Área–Año (absorbidos)				

Notas: reghdfe con errores estándar robustos agrupados por área (cluster=área) y ponderación por factores de expansión. El estadístico F de exclusión es extremadamente bajo en todas las especificaciones, indicando instrumento débil.

fijos. El objetivo es determinar si el instrumento conserva suficiente poder de predicción sobre la variable endógena, condición esencial para garantizar la validez de las estimaciones de segunda etapa.

Los resultados muestran que el coeficiente asociado a Z_{Bartik} no es estadísticamente significativo en ninguna de las especificaciones, y los valores del estadístico F de exclusión son extremadamente bajos (menores a 1). Esto indica una débil correlación entre el instrumento y el IMD, lo cual sugiere problemas de relevancia en la primera etapa. A pesar de la estabilidad del R^2 ajustado, la falta de significancia del instrumento implica que las estimaciones basadas en este modelo deben interpretarse con cautela, dado el riesgo de sesgo por instrumentos débiles.

En términos metodológicos, estos resultados subrayan la necesidad de considerar enfoques alternativos de identificación o el uso de instrumentos complementarios que fortalezcan la variación exógena del IMD. En conjunto, el ejercicio confirma que el instrumento Z_{Bartik} , en su configuración actual, no ofrece suficiente potencia estadística para inferencias causales sólidas, aunque sirve como prueba de consistencia frente a potenciales fuentes de endogeneidad.

El Cuadro 11 presenta los resultados de la segunda etapa del modelo de variables instrumentales, en el cual el Índice de Madurez Digital (IMD) se instrumenta mediante el Z_{Bartik} para estimar su efecto sobre la formalización, los ingresos y el empleo. Esta estimación se realiza empleando el método de Mínimos Cuadrados en Dos Etapas (2SLS), con errores estándar agrupados por área y ponderación por factores de expansión. El objetivo es verificar si, tras corregir la endogeneidad potencial del IMD, los efectos observados en las secciones previas se mantienen consistentes.

Los resultados muestran que el estimador de 2SLS no identifica coeficientes válidos, reflejando

una debilidad extrema del instrumento. Los valores del estadístico F de la primera etapa (0,15) son considerablemente inferiores al umbral convencional de 10, lo que indica que el instrumento carece de poder explicativo suficiente para generar variación exógena en el IMD. En consecuencia, los coeficientes reportados no son informativos y no pueden ser interpretados como efectos causales robustos.

Metodológicamente, esta evidencia confirma los hallazgos previos de la primera etapa, donde se advertía la debilidad del instrumento Z_{Bartik} . La ausencia de resultados significativos en esta segunda etapa refuerza la necesidad de emplear pruebas robustas frente a instrumentos débiles —como los estimadores de Anderson–Rubin o CLR— o de recurrir a una forma reducida que capture el efecto total sin depender de una primera etapa ineficiente. En conjunto, el ejercicio evidencia la fragilidad de la estrategia instrumental actual y sugiere que las conclusiones más sólidas deben apoyarse en los modelos de efectos fijos y en los análisis de robustez previos.

El Cuadro 11 presenta los resultados de la estimación en forma reducida, en la cual el instrumento Z_{Bartik} se vincula directamente con los resultados empresariales, omitiendo la etapa intermedia de predicción del Índice de Madurez Digital (IMD). Esta aproximación permite examinar el efecto total del componente exógeno capturado por el instrumento sobre las variables dependientes, evitando los problemas de identificación asociados a una primera etapa débil.

Los resultados muestran que el instrumento Z_{Bartik} exhibe un efecto negativo y significativo sobre las medidas de formalización —tanto en su versión núcleo como en el índice extendido—, mientras que no se observan efectos significativos sobre los ingresos ni el empleo. Este patrón sugiere que los shocks instrumentales asociados a la conectividad podrían estar correlacionados con factores estructurales que, en el corto plazo, reducen los niveles observados de formalización, posiblemente debido a ajustes organizacionales o transitorios en la adopción tecnológica. El hecho de que las variables de desempeño económico no respondan de manera significativa refuerza la idea de que el canal de impacto es institucional y no puramente productivo.

A nivel metodológico, esta estimación sirve como una verificación complementaria a las pruebas instrumentales anteriores. Dado que el Z_{Bartik} presenta una debilidad estadística en la primera etapa, la forma reducida ofrece una alternativa más confiable para evaluar la dirección general del efecto exógeno. En conjunto, los resultados sugieren que los cambios inducidos por el instrumento afectan

Cuadro 11: Forma reducida: $Z_{Bartik} \rightarrow$ resultados (OLS con FE)

	Formalización (núcleo)	Formalización (0–100)	ln(Ventas reales)	Empleo (n)
Z_{Bartik}	-0.074** (0.025)	-2.290** (0.816)	-0.018 (0.501)	0.003 (0.004)
Obs.	466,629	466,629	466,629	466,629
R^2 ajustado	0.325	0.371	0.190	0.904
Efectos fijos	Área–Año (absorbidos) + composición y controles SPEC5			

Notas: reghdfe con errores estándar robustos agrupados por área (cluster=geo_id) y ponderación por peso_exp. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

principalmente los patrones de formalización más que las dimensiones de ingreso o empleo, y que estos efectos probablemente reflejan transiciones estructurales en lugar de impactos permanentes sobre la productividad.

Como último análisis de robustez, el Cuadro 12 presenta los resultados del ejercicio *Leave-One-Year-Out* (LOYO) aplicado a la primera etapa del modelo instrumental, en el cual el instrumento Z_{Bartik} se reestima sucesivamente excluyendo un año completo del periodo analizado (2019–2024). Este procedimiento permite evaluar la estabilidad temporal del instrumento y determinar si su debilidad se concentra en un año específico o si, por el contrario, refleja una falta de poder predictivo generalizada.

Los resultados muestran una marcada inestabilidad en los coeficientes de Z_{Bartik} a lo largo de las exclusiones anuales. Las estimaciones varían tanto en signo como en magnitud, y presentan errores estándar considerablemente elevados, lo que indica una alta imprecisión estadística. En particular, los años intermedios (2021–2023) exhiben coeficientes positivos de mayor magnitud, aunque no significativos, mientras que los extremos (2019 y 2024) muestran valores más bajos y con signo opuesto. Este patrón sugiere que las fluctuaciones anuales en la relación entre el instrumento y el IMD no son sistemáticas, sino probablemente producto de ruido muestral o de variaciones coyunturales en los componentes que integran el instrumento Bartik.

Desde una perspectiva metodológica, la combinación de coeficientes inestables y errores estándar amplios confirma la debilidad estructural del instrumento. La falta de consistencia temporal implica que el Z_{Bartik} no genera una variación exógena sostenida en la madurez digital, reduciendo su utilidad para la identificación causal. En conjunto, este ejercicio refuerza la interpretación de las sec-

Cuadro 12: LOYO (1ra etapa): exclusión secuencial de un año en $Z_{\text{Bartik}} \rightarrow \text{IMD}$

Año excluido	Coef. Z_{Bartik}	EE (cluster)
2019	-1.131	4.068
2020	1.917	3.801
2021	3.965	4.293
2022	2.326	3.961
2023	1.707	3.620
2024	0.471	3.832
FE & controles	Área–Año (absorbidos) + SPEC5 (macro, comp., etc.)	

Notas: en cada fila se elimina por completo el año indicado y se re-estima la primera etapa con ponderación por factores de expansión y errores agrupados por área. La inestabilidad y la baja precisión (EE elevados) son consistentes con debilidad del instrumento.

ciones anteriores: los resultados instrumentales deben considerarse exploratorios y complementarios, y la evidencia más sólida proviene de los modelos de efectos fijos y de los análisis de robustez no instrumentales.

4. Conclusiones e implicaciones

El conjunto de resultados presentados a lo largo de este estudio permite ofrecer una visión integral sobre la relación entre madurez tecnológica y desempeño empresarial. En cuanto al desempeño empresarial, los resultados obtenidos a partir de los modelos con efectos fijos y las pruebas de robustez indican que la madurez digital tiene efectos positivos y consistentes sobre la formalización, los ingresos y el empleo. Este patrón confirma que la digitalización no solo constituye una estrategia tecnológica, sino también un vector de formalización y crecimiento económico (DNP, 2020; OIT, 2021; Solano-Romo, 2022). Las estimaciones ponderadas y truncadas, junto con los ejercicios de soporte común, muestran que estos efectos son generalizables y no dependen de valores extremos o de sectores específicos, lo que refuerza su relevancia estructural (DNP, 2023). Por otro lado, los ejercicios instrumentales basados en el enfoque Bartik revelan limitaciones en la identificación causal, dado el bajo poder predictivo del instrumento. Esta debilidad metodológica, sin embargo, no invalida los hallazgos principales, sino que delimita su interpretación causal estricta. La consistencia de los resultados en los modelos no instrumentales y en las pruebas de robustez sugiere que la relación entre madurez digital y desempeño empresarial es sustantiva y persistente, aunque no pueda atribuirse completamente a

variación exógena (DNP, 2023; OIT, 2021). En conjunto, los resultados apuntan a una conclusión clave: la conectividad digital es un insumo esencial pero insuficiente para impulsar la madurez tecnológica y la productividad empresarial, punto central para fortalecer el ecosistema de los micronegocios. El avance digital requiere complementar la infraestructura con políticas orientadas a la gestión del cambio, la formalización y el desarrollo de capacidades internas (DNP, 2020; OIT, 2021; Solano-Romo, 2022). En ese sentido, instrumentos de financiamiento y acompañamiento que alivien restricciones de capital y permitan inversiones no triviales en activos, por ejemplo, inyecciones de capital y servicios de desarrollo empresarial, resultan coherentes con la evidencia sobre retornos al capital y heterogeneidad de emprendimientos en micronegocios y Mupimes (de Mel, McKenzie & Woodruff, 2008; de Mel, McKenzie & Woodruff, 2019; Schoar, 2015). Las políticas públicas deberían, por tanto, trascender la provisión de conectividad y concentrarse en la creación de entornos organizacionales y financieros que permitan a las empresas transformar dicha conectividad en innovación, eficiencia y crecimiento sostenido (DNP, 2023; OIT, 2021).

Referencias bibliográficas

Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Inferring an Examinee's Ability [Incluye la formulación del modelo logístico de dos parámetros (2PL)]. En F. M. Lord & M. R. Novick (Eds.), *Statistical Theories of Mental Test Scores* (pp. 395-479). Addison- Wesley.

de Mel, S., McKenzie, D., & Woodruff, C. (2019). Labor drops: Experimental evidence on the return to capital for female micro-entrepreneurs. *American Economic Journal: Applied Economics*, 11(2), 124–151.

de Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *Quarterly Journal of Economics*, 123(4), 1329–1372.

Departamento Nacional de Planeación [DNP]. (2020). Documento CONPES 3956: Política de formalización laboral en Colombia.

Departamento Nacional de Planeación [DNP]. (2023). Índice Departamental de Competitividad 2023. Consejo Privado de Competitividad y Universidad del Rosario.

Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3), 109–126.

Organización Internacional del Trabajo [OIT]. (2021). Trabajo, empleo, protección laboral y economía informal. Schneider, F., & Enste, D. H. (2000). Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38(1), 77–114.

Samejima, F. (1969). Estimation of Latent Ability Using a Graded Response Model. *Psychometrika Monograph Supplement*, (17), 1-100.

Schoar, A. (2015). The divide between subsistence and transformational entrepreneurship. In J. Lerner & S. Stern (Eds.), *Innovation Policy and the Economy* (Vol. 10, pp. 57–81). University of Chicago Press.

Solano-Romo, G. (2022). Entendiendo la adopción de herramientas digitales en las MIPYMES latinoamericanas. *Revista de Administración de Empresas*, 62(2), 1–20.