

XXI Concurso Anual de Investigación CIES 2019

Estimación de la relación empírica entre el acceso a infraestructura de riego y la productividad de las unidades agrícolas en el Perú

Cesar Del Pozo •

Con el apoyo de



PERÚ

Ministerio
de Economía y Finanzas

Índice

<i>Introducción</i>	4
1. Marco Teórico	6
1.1 Funciones de Producción y Productividad.....	6
1.2 Rol del Riego en la Productividad Agrícola	10
1.3 Revisión de la Literatura.....	12
2. Metodología	15
2.1 Fuente de Datos y Variables	15
2.2 Métodos de Estimación de la Función de Producción.....	18
2.3 Estrategia de Identificación del Efecto del Riego en la Productividad Total Factorial	26
3. Resultados	29
3.1 Estimación de la Función de Producción.....	29
3.2 Estimación de la Función de Producción Controlando por Calidad de la Tierra	33
3.3 Estimación de la Productividad Total Factorial	40
3.4 Estimación del Efecto del Riego en la Productividad Total Factorial.....	47
3.5 Análisis de Género.....	55
4. Conclusiones	58
5. Recomendaciones de Políticas Públicas	60
<i>Bibliografía</i>	62
<i>Anexos</i>	66

Introducción

De acuerdo con la Política Nacional Agraria¹ y los Lineamientos de la Política y Estrategia Nacional de Riego², definidos por el Ministerio de Agricultura y Riego para el periodo 2015-2025, uno de los problemas más relevantes que afecta a la agricultura peruana es la baja productividad, la cual sería generada por el ineficiente aprovechamiento de los recursos hídricos para uso agrario: excesiva escorrentía, insuficiente infraestructura para riego, inadecuadas prácticas de riego y la escasa implementación de tecnologías de riego más eficientes, entre otros motivos.

De acuerdo con Ludena (2010) y Trindade y Fulginiti (2015), la productividad de la agricultura peruana se habría incrementado a nivel agregado en las últimas décadas; sin embargo, la agricultura peruana comprende muchas realidades que contrastan entre sí. Por un lado, la costa peruana se muestra dotada de sistemas agrícolas dinámicos y modernos, con una mayor inversión, elevada escala de producción y acceso al riego. Por el contrario, la sierra y la selva presentan sistemas agrícolas tradicionales, menor inversión, orientados a la subsistencia y con menor acceso al riego.

De acuerdo con el IV Censo Nacional Agropecuario (CENAGRO, 2012), solamente el 36% de la superficie agrícola a nivel nacional se encuentra bajo algún tipo de riego, sea por gravedad, aspersión, goteo o exudación. La superficie agrícola bajo riego asciende a más de 2.5 millones de hectáreas, de las cuales el 56% están ubicadas en la costa, 40% en la sierra y 4% en la selva. En contraste, la superficie bajo agricultura de secano es de 4.5 millones de hectáreas, de las cuales el 51% se ubican en la sierra, 45% en la selva y solamente 5% en la costa.

¹<https://www.minagri.gob.pe/portal/download/pdf/marcolegal/normaslegales/decretossupremos/2016/ds02-2016-minagri.pdf>

²<https://sinia.minam.gob.pe/normas/aprueban-lineamientos-politica-estrategia-nacional-riego-2015-2025>

Ante ello, desde 2012 el Estado peruano ha ejecutado diversas inversiones focalizadas en el sector agricultura a través de los denominados Programas Presupuestales (PP); en particular, el Ministerio de Agricultura y Riego implementó el PP denominado *Aprovechamiento de los recursos hídricos para uso agrario*. Dicho programa tiene como objetivo mejorar la eficiencia del aprovechamiento de los recursos hídricos, donde su Presupuesto Institucional Modificado se incrementó de S/. 182 millones en 2012 hasta S/. 987 millones en 2019. Dentro de la lógica de intervención de este PP, el incremento en el acceso a la infraestructura de riego puede incrementar la productividad agrícola. Por ello, surge la siguiente pregunta: ¿Cuál es la relación empírica entre el acceso a infraestructura de riego y la productividad de las unidades productivas agrícolas en el Perú?

Para el caso peruano, algunos estudios han estimado la productividad agrícola (Galarza y Díaz, 2014; Banco Mundial, 2017); mientras que otros estudios han analizado el impacto de la infraestructura de riego en el desempeño agropecuario (Del Carpio et al., 2011; MEF, 2019). No obstante, se han hecho relativamente pocos esfuerzos para ir más allá de mediciones agregadas de la productividad agrícola, así como para explorar la relación entre el acceso al riego y la productividad. Lo anterior sugiere la existencia de algunos vacíos de conocimiento en términos de evidencia empírica de la estimación de la productividad a nivel de unidades agropecuarias: la falta de evidencia empírica sobre el efecto del acceso al riego en la productividad en la agricultura peruana.

Así pues, el objetivo general de la presente investigación es estimar la relación empírica entre el acceso a infraestructura de riego y la productividad de las unidades productivas agrícolas en el Perú. De esta manera, los objetivos específicos son: (i) estimar la productividad total factorial de unidades agrícolas en el Perú, con base en metodologías más actuales para la estimación de funciones de producción y datos relativamente más recientes; (ii) estimar la relación empírica entre el acceso a infraestructura de riego y la productividad total factorial.

1. Marco Teórico

1.1 Funciones de Producción y Productividad

a) Funciones de Producción. La producción es el resultado de la transformación de factores productivos tales como el capital, trabajo, materiales o insumos, dada una tecnología. Económicamente, la producción puede ser representada a través de funciones de producción estocásticas (Griliches y Mairesse, 1995):

$$Y_{i,t} = F(K_{i,t}, L_{i,t}, M_{i,t}, \omega_{i,t}, \varepsilon_{i,t}) \quad (1)$$

En la ecuación anterior, $Y_{i,t}$ representa el nivel de producción de la unidad productiva “i” en el periodo “t”, el cual depende del uso de materiales ($M_{i,t}$), trabajo ($L_{i,t}$), capital ($K_{i,t}$), la productividad ($\omega_{i,t}$) y un término de error estocástico ($\varepsilon_{i,t}$). Entre los argumentos del lado derecho de la ecuación anterior, dos de ellos son inobservables: $\omega_{i,t}$, $\varepsilon_{i,t}$. El argumento $\varepsilon_{i,t}$ representa aquellos shocks de producción que no son observables o predecibles por las unidades productivas antes que estas toman sus decisiones de uso de factores de producción en el periodo “t”. En contraste, $\omega_{i,t}$ representa los shocks de productividad que son potencialmente observables o predecibles por las unidades productivas cuando éstas toman decisiones sobre el uso de factores de producción. Intuitivamente, $\omega_{i,t}$ puede resumir la habilidad o experiencia de los conductores de las unidades productivas y/o diversos shocks esperados

Existe un problema de identificación para la estimación empírica, puesto que los materiales ($M_{i,t}$), capital ($K_{i,t}$) y trabajo ($L_{i,t}$) son variables de decisión que dependen a su vez de la productividad, generando un problema de endogeneidad por variable omitida toda vez que la productividad es inobservable.

A pesar de su importancia en el análisis económico, la estimación empírica de la ecuación (1) es afectada por diversos sesgos y desafíos empíricos. Entre los más relevantes se encuentran los siguientes (Aguirregabiria, 2018):

- *Sesgo de selección:* la información necesaria para la estimación de la función de producción ha empleado generalmente muestras del tipo paneles de datos, las cuales combinan información de unidades productivas observadas en varios periodos de tiempo. En paneles de datos, la salida de unidades productivas de la muestra no es exógena y está correlacionada con el tamaño de las mismas, unidades de menor tamaño tienen una mayor probabilidad de salir del mercado que unidades más grandes y consecuentemente no ser observadas en varios periodos de tiempo. Una potencial forma de solucionar el problema es considerar paneles de datos balanceados.
- *Simultaneidad:* determinados factores de producción, tales como el trabajo y capital, pueden estar correlacionados con variables inobservables, tales como la productividad o la calidad de los factores de producción. En consecuencia, se generan problemas de variables omitidas y errores de medición; tales problemas empíricos han sido abordados en la literatura por Olley y Pakes (1996), Levinsohn y Petrin (2003) y Ackerberg, Caves y Fraser (2015). Los errores de medición pueden ser relevantes en el caso de la estimación de la función de producción agrícola; en particular, el error de medición generado por la calidad de la tierra (Restuccia y Santaeulalia-Llopis, 2017; Gollin y Udry, 2019).
- *Dependencia condicional:* este problema empírico se genera porque las variables de decisión (trabajo, insumos e inversión) dependen de las mismas variables de estado (capital y productividad), lo cual implica la presencia de colinealidad entre los factores de producción. Al igual 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. Recientemente, el problema de dependencia condicional en la estimación de función de producción ha sido abordado en la literatura por Ackerberg, Caves y Fraser (2015).

b) Productividad. La productividad es una medición que determina cómo se emplean los factores de producción combinados para generar unidades de producción. La productividad puede ser considerada como el indicador de desempeño económico más intuitivo a distintos niveles (Jorgeson y Griliches, 1967). A nivel de unidades productivas y en un contexto de múltiples factores y productos, la productividad resume todas aquellas externalidades positivas que contribuyen al incremento de la producción.

Sin embargo, a pesar de su relevancia como indicador de desempeño económico, la productividad es una variable inobservable para los investigadores que buscan estimarla. Según la literatura consultada, podemos constatar que se han desarrollado distintas metodologías de estimación de la eficiencia o productividad de las firmas. tales metodologías pueden clasificarse en metodologías econométricas y metodologías basadas en programación matemática, en ambas metodologías se estima la eficiencia alrededor de las fronteras de producción.

En las metodologías econométricas se especifican formas funcionales de costos, beneficios o producción, desarrollando metodologías que tratan de corregir problemas de endogeneidad y, más recientemente, proponiendo metodologías de función de control que estiman la productividad a nivel de unidades productivas como un residuo inobservable resultante de un función de producción que relaciona combinaciones de insumos con resultados de producción. Por su parte, las metodologías basadas en programación matemática se basan en métodos no paramétricos. La literatura abocada al estudio de estos métodos ha desarrollado aproximaciones específicas, tales como *Data Envelopment Analysis* (DEA) o *Free Disposal Hull* (FDH)³.

En esta investigación, se considera relevante emplear la aproximación econométrica con el fin estimar la productividad a nivel de unidades productivas agrícolas para el caso peruano; en particular, aquellas aproximaciones basadas en métodos de función de control. En tal

³ Para una revisión exhaustiva de las metodologías disponibles en la literatura para la estimación de la productividad y eficiencia ver Fried et al., (2008).

contexto, la productividad es aproximada por la Productividad Total de Factores (PTF), la cual es calculada como residuo resultante de descontar de la producción de la contribución de los factores de producción. En este documento, se considera que la PTF resume todas aquellas externalidades positivas inobservables que contribuyen al incremento de la producción (Abramovitz, 1956; Jorgenson y Griliches, 1967).

Generalmente, las estimaciones de la PTF han empleado la función de producción Cobb-Douglas, donde se estiman los parámetros de interés a través de Mínimos Cuadrados Ordinarios (OLS, por sus siglas en inglés) y calculado la PTF como el residuo resultante entre el producto observado y la participación ponderada de los factores productivos (enfoque primal). Sin embargo, al existir problemas de simultaneidad y dependencia condicional, las estimaciones por OLS de los parámetros que asocian el capital y el trabajo en la función de producción resultan inconsistentes.

La inconsistencia en el estimador de OLS implica que, por un lado, la participación del capital (β_k) sufriría un sesgo hacia abajo o sesgo de atenuación ($\hat{\beta}_{k,OLS} < \beta_k$), puesto que el término de error estocástico estaría correlacionado negativamente con las variables omitidas o errores de medición (problema de simultaneidad) o exista colinealidad negativa entre los factores de producción (problema de dependencia condicional). Por otro lado, la participación del trabajo (β_l) sufriría un sesgo hacia arriba ($\hat{\beta}_{l,OLS} > \beta_l$) debido a que el término de error estocástico estaría correlacionado positivamente con las variables omitidas o errores de medición (problema de simultaneidad) o exista colinealidad positiva entre los factores de producción (problema de dependencia condicional).

Ante las implicancias de estimaciones inconsistentes de la participación del trabajo y el capital en la estimación paramétrica de la función de producción, desarrollos recientes (Olley y Pakes, 1996; Levinsohn y Petrin, 2003; y Ackerberg, Cravés y Fraser, 2015), proponen aproximaciones empíricas que podrían ofrecer estimaciones consistentes de los parámetros.

Con tales aproximaciones, se podría ofrecer estimaciones más precisas de la PTF a nivel de unidades productivas agrícolas para el caso peruano, estimaciones de la productividad actualizadas en función a la disponibilidad de datos más recientes y estimaciones menos sesgadas de la relación empírica entre el acceso a infraestructura de riego y la productividad.

1.2 Rol del Riego en la Productividad Agrícola

El agua es un bien económico puesto que genera diversos beneficios de uso (consumo, riego, etc.) y de no uso (servicios ambientales) a los seres humanos. Los diversos beneficios del agua se generan en espacios territoriales específicos (cuencas hidrográficas), en los cuales compiten diversos usos del agua. El agua puede ser considerada como un bien mixto, tanto público como privado, en función de sus beneficios de uso y las características de exclusión⁴ y rivalidad⁵ (Zegarra, 2014).

El caso del uso de agua para riego en la agricultura puede ser considerada como un bien mixto, puesto que la irrigación tiene una alta rivalidad, toda vez que el consumo de agua para fines de riego reduce la disponibilidad total del agua; en el caso peruano, por ejemplo, la agricultura consumiría el 85% de la disponibilidad del recurso hídrico. En contraste, la exclusión es relativamente baja toda vez que no es posible excluir totalmente su uso, donde los mercados no quedan adecuadamente organizados (no hay definición de tarifas, no son óptimas o los usuarios no pagan las tarifas establecidas en los sistemas de riego), como sucede en el caso del agua potable donde no pagar genera el corte del suministro (Zegarra, 2014).

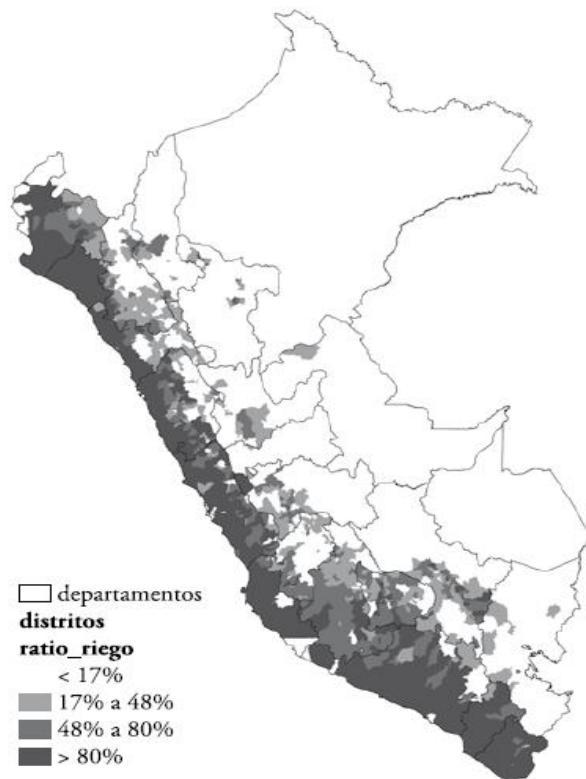
El acceso al riego en la producción agrícola genera mayor estabilidad y control sobre el proceso productivo, incentiva mayores niveles de inversión y estimula la productividad de la tierra (Zegarra, 2014). La generación de infraestructura de riego requiere inversiones para el almacenamiento, control y distribución del recurso hídrico. En el caso peruano, de acuerdo con

⁴ La característica de exclusión de un bien público puro está referida a que no es posible excluir a las personas del consumo del bien: lluvia, aire, etc.

⁵ La característica de rivalidad de un bien público puro está referida a que el consumo de los individuos no reduce el stock del bien, ni incrementa la escasez de este: alumbrado público, defensa nacional.

el IV Censo Nacional Agropecuario, la superficie bajo riego sería de 2.5 millones de hectáreas, de las cuales 56% se ubican en la costa, 40% en la sierra y solamente 4% en la selva. En el Mapa 1 podemos observar que el acceso al riego se encuentra concentrado en la costa peruana:

Mapa 1
Superficie agrícola bajo riego



Fuente: Zegarra (2014).

El acceso al riego puede afectar positivamente al desempeño agrícola a través de dos canales: (i) el riego puede incrementar la extensión de la tierra cultivable y la intensidad agrícola, entendida como el cambio hacia períodos de siembra y cosecha más permanentes y predecibles a lo largo del año agrícola, lo cual es particularmente relevante en un contexto de cambio climático donde las lluvias son en general menos predecibles. (ii) El riego puede generar mayor estabilidad y confiabilidad en los patrones y rendimientos de los cultivos. Para los agricultores,

contar con una fuente constante de agua les permitiría optar por cultivos de mayor rendimiento y rentabilidad (Del Carpio et al., 2011). A través de ambos canales, el acceso al riego puede conducir a la expansión de la tierra cultivable, la intensidad y diversificación de los cultivos, reduciendo el riesgo e incertidumbre de los agricultores mejorando la productividad agrícola (Smith, 2004).

1.3 Revisión de la Literatura

En primer término, a nivel internacional la estimación de la productividad ha sido realizada empleando distintas metodologías y en distintos contextos. Por ejemplo, Adamopoulos y Restuccia (2014a) estiman la productividad agrícola y su relación con el tamaño de las unidades agropecuarias (fincas) para un conjunto de países, analizando las determinantes de las diferencias en productividad entre países y encontrando que más de la mitad de la brecha de productividad entre países podría ser explicada por la mala asignación de recursos y errores de medición en la actividad agrícola.

En otro trabajo, Adamopoulos y Restuccia (2014b) estiman también la productividad de unidades productivas agrícolas y su relación con una reforma en el acceso a la tierra tomando el caso de Filipinas, donde la reforma impuso un límite a la tenencia de la tierra y restringió severamente la transferencia de tierras entre productores. En este caso, los resultados pusieron en evidencia que dicha reforma tuvo un impacto negativo en el tamaño y la productividad de las fincas, lo que destaca la relevancia del rol de la eficiencia de los mercados de tierras en la mala asignación de recursos y sus implicancias para la productividad en el sector agrícola.

También destacan investigaciones como las de Newman et al., (2015), quienes analizan el efecto de la titulación de las tierras en la productividad agrícola en Vietnam. Para abordar el problema de endogeneidad entre la titulación y la productividad emplean un modelo de efectos fijos encontrando que el acceso a la titulación incrementa la productividad, por lo cual se constituyéndose en un mecanismo efectivo para incrementar el desempeño agrícola. Por otro

lado, Ali y Deininger (2015) estiman la productividad agrícola y su relación con el tamaño de la unidad agrícola (finca) para el caso de Ruanda, encontrando una relación negativa entre productividad y tamaño de la finca, resultado que estaría explicado por imperfecciones en el mercado de trabajo.

En trabajos más recientes, Restuccia y Santaeulalia-Llopis (2017) estiman la productividad de unidades agrícolas (fincas) controlando a partir de la calidad de la tierra como variable para el caso de Malawi. En dicho estudio determinan que la medición de la productividad está afectada por la calidad de la tierra y otros shocks transitorios (principalmente climáticos). Además encuentran que la productividad no está relaciona con el tamaño de las fincas, lo cual implicaría una mala asignación del dicho factor productivo.

Por su parte, Adamopoulos et al., (2017) analizan las consecuencias de la mala asignación de factores productivos en hogares agrícolas de China. En este caso determinan que la mala asignación de factores productivos entre agricultores y la asignación de trabajadores en el sector agrícola reducen la productividad agregada de la agricultura china. En otro caso, Herrera et al., (2018) estiman la productividad de pequeños productores para el caso de Brasil y encuentran que aquellos productores que pertenecen a cooperativas o asociaciones muestran mayores niveles de productividad, ingresos y diversificación productiva.

En segundo término, la literatura empírica reciente sobre la estimación de la productividad agrícola en el Perú ha sido desarrollada, entre otros, por Galarza y Díaz (2014) y el Banco Mundial (2017). Galarza y Díaz (2014), miden la productividad (PTF) y sus determinantes a nivel de productores agrícolas. Usando la estimación de funciones de producción, los autores han implementado mejoras metodológicas para dar cuenta del problema de endogeneidad en la estimación de la función de producción. No obstante, si bien tal estudio da cuenta del problema de endogeneidad, no aborda el problema de dependencia condicional entre los factores de producción. En particular, encuentran que la productividad está

positivamente correlacionada con la edad, sexo (varón) y el nivel educativo de los productores y negativamente correlacionada con el tamaño de la unidad agropecuaria.

Por otra parte, el estudio del Banco Mundial (2017) analiza las tendencias recientes de productividad en el sector agrícola peruano y pone énfasis en analizar el crecimiento de la productividad entre regiones, además de los tipos y tamaños de unidades agropecuarias. Los autores evalúan la Productividad Total de Factores (PTF) con base en un análisis de índices de productividad de Tronqvist-Theil empleando información de productores agropecuarios entre los años 2007 y 2015 provenientes de la Encuesta Nacional de Hogares (ENAHO). Uno de los resultados muestra una relación positiva entre la productividad y el tamaño de la unidad agrícola, lo cual sugiere que las unidades agropecuarias más grandes usan insumos de manera más eficiente que las pequeñas. Respecto a la expansión de la irrigación, se concluye que esta habría tenido efectos positivos en el crecimiento de la PTF, principalmente en la costa peruana. En el caso de la sierra y selva, por el contrario, se registran bajos niveles de productividad a causa de que la práctica agrícola se ejecuta en parcelas muy pequeñas, se emplea mano de obra familiar y no cuenta con sistemas de riego.

En tercer término y de manera específica, estudios recientes analizaron el impacto de la infraestructura de riego en el desempeño agropecuario. Del Carpio et al. (2011) analizaron el impacto de la infraestructura de riego en el desempeño agropecuario. El estudio concluye que un conjunto de proyectos de irrigación y de rehabilitación de infraestructura de riego (PSI⁶) desarrollados en la costa peruana incrementan la producción agrícola de productores de mayor escala en las zonas beneficiarias de los proyectos PSI.

Por su parte, Zegarra (2014) analiza la relación entre el acceso al riego y los ingresos per cápita de agricultores de pequeña escala; este estudio descubre una relación positiva y

⁶ Programa Subsectorial de Irrigaciones

significativa entre ambas variables y termina concluyendo que el acceso al riego incrementa en mayor proporción los ingresos de agricultores de la costa y la selva.

En una investigación más cercana, el MEF (2019) analizó el impacto del Fondo Mi Riego (Sierra Azul) en 16 indicadores de desempeño agropecuario (entre ellos: rendimientos, valor bruto de la producción, ingreso neto). El Fondo Mi Riego implica inversiones públicas focalizadas cuyo propósito es mejorar el acceso al riego en sectores de la sierra peruana con altos niveles de pobreza. Los resultados recogen impactos positivos en los niveles de producción de determinados cultivos (maíz, cebada, trigo), así como efectos positivos en el rendimiento de otros productos (maíz, quinua, etc.); sin embargo, no encuentran efectos positivos en indicadores de rentabilidad, tales como ingresos brutos.

No obstante, a pesar de todas las investigaciones señaladas, es notorio que se han hecho relativamente pocos esfuerzos para ir más allá de mediciones acumuladas en el ámbito nacional para explorar las diferencias en la productividad agrícola entre regiones, entre tipos de productores y entre tamaños de unidades agropecuarias, así como la relación entre el acceso al riego y la productividad, lo cual sugiere la existencia de algunos vacíos de conocimiento. Por lo tanto, este estudio busca contribuir a subsanar dichas ausencias.

2. Metodología

2.1 Fuente de Datos y Variables

Usualmente, para la estimación empírica de las funciones de producción se han empleado paneles de datos de unidades productivas individuales con información anual y mediciones de producción valorizada, ingresos, ventas o valor agregado. Los factores de producción son aproximados a través de mediciones específicas de capital, trabajo e insumos, en función de la disponibilidad de información (Van Beveren, 2012 y Aguirregabiria, 2018). En el presente estudio se consideró emplear como principal fuente de información la Encuesta Nacional Agropecuaria (ENA) correspondiente a los años 2016, 2017 y 2018.

Dicha base de datos permite disponer de información estadística para caracterizar a pequeñas unidades agrícolas en el Perú. La ENA tiene cobertura a nivel nacional, es una encuesta probabilística y su población de referencia está basada en los resultados del IV Censo Nacional Agropecuario. El propósito de la ENA es generar información para calcular variables agrícolas y estimar efectos de diferentes programas públicos implementados en el sector agricultura, la encuesta tiene representación nacional y departamental de unidades agrícolas, lo que hace posible que la información sobre los productores pueda aglomerarse por regiones y dominios.

La muestra empleada en la presente investigación está conformada por 5803 unidades agropecuarias, con información recogida durante los años 2016, 2017 y 2018; dicha muestra presenta una estructura de tipo panel de datos balanceado. El 59% de las unidades productivas se localizan en la sierra peruana, 19% en la costa y 22% en la selva.

LA ENA incluye información detallada sobre la producción agrícola, donde se resalta información sobre las cantidades, los precios y el valor agregado por cultivo en cada parcela de la unidad agrícola. Respecto de insumos, releva información sobre cantidades y gastos por cultivo (semillas, abono, fertilizantes y pesticidas). También se revela información sobre gastos acumulados a nivel de la unidad agropecuaria: alquiler de tierras, trabajadores permanentes y temporales, gasto en agua para riego, asistencia técnica, compra y alquiler de equipos y maquinaria agrícola, compra de combustible y otros; así como información sobre infraestructura productiva y equipo agrícola. Se contempla también información detallada sobre las características de la unidad agropecuaria y la familia, características como: género, edad, educación, lengua materna, número de parcelas, área sembrada, acceso a infraestructura de riego, tipo de riego, número de cultivos, ganado, entre otros.

En la Tabla 1 se muestran las principales variables consideradas para la estimación de la función de producción con base en la información disponible en la ENA. Para aproximar la

producción se ha considerado el valor monetario total de la producción agrícola ($Y_{i,t}$). El factor trabajo ha sido aproximado a través del número de trabajadores remunerados ($LR_{i,t}$) y el número de trabajadores familiares no remunerados que apoyan en la explotación agrícola ($LNR_{i,t}$). El factor capital ha sido aproximado a través del uso de la tierra, medido como el número de hectáreas de la superficie cosechada ($T_{i,t}$). Los insumos fueron aproximados como el valor monetario total de los gastos en semillas, abono, fertilizante, pesticidas y pago por el acceso al agua ($M_{i,t}$). Con el objetivo de complementar las estimaciones para reducir sesgo por variables omitidas, de acuerdo con lo sugerido por Restuccia y Santaella-Llopis (2017) y Gollin y Udry (2019), se incorporaron indicadores proxy de la calidad de la tierra.

En la Tabla A.1 del anexo se presentan el resto de las variables consideradas en el análisis y variables potencialmente relacionadas con la calidad de la tierra.

Tabla 1

Variables

Variables	Definición
$Y_{i,t}$	Producción: valor monetario total de la producción agrícola
$LR_{i,t}$	Trabajo remunerado: número de trabajadores remunerados eventuales y permanentes.
$LNR_{i,t}$	Trabajo no remunerado: número de trabajadores familiares no remunerados que apoyan en la explotación agrícola.
$T_{i,t}$	Tierra: número de hectáreas de superficie cosechada.
$M_{i,t}$	Insumos: valor monetario total de los gastos en semillas, abono, fertilizante, pesticidas y acceso al agua

Elaboración Propia

Finalmente, con el objetivo de mejorar la comparabilidad en el tiempo, las variables expresadas en valores monetarios fueron deflactadas temporalmente a precios de 2009,

empleando para ello el Índice de Precios del Consumidor de Lima Metropolitana entre 2016 y 2018.

2.2 *Métodos de Estimación de la Función de Producción*

Se asume una función de producción de valor agregado de tipo Leontief de la ecuación (1). La especificación de tipo Leontief corresponde a la situación donde una cantidad específica de insumos es necesaria para producir una unidad de producto. Al aplicar logaritmos es posible plantear la función de producción agregada lineal de la forma “*Value-added production*”⁷:

$$va_{i,t} = \beta_0 + \beta_{lr} lr_{i,t} + \beta_{lnr} lnr_{i,t} + \beta_{\tau} \tau_{i,t} + \omega_{i,t} + \varepsilon_{i,t} \quad (2)$$

En la ecuación anterior, $va_{i,t} = y_{i,t} - m_{i,t}$ es el valor agregado de la producción expresada como la diferencia entre el logaritmo de la producción ($y_{i,t}$) y el logaritmo del gasto en insumos ($m_{i,t}$). Por su parte, $lr_{i,t}$ es el logaritmo del número de trabajadores remunerados, $lnr_{i,t}$ es el logaritmo del número de trabajadores familiares no remunerados, $\tau_{i,t}$ es el logaritmo del número de hectáreas cosechadas, $\omega_{i,t}$ es una variable no observable que representa la productividad, la cual es utilizada por los agricultores para tomar decisiones anticipadamente respecto del uso de factores de producción. Finalmente, $\varepsilon_{i,t}$ es un shock estocástico en la producción, el cual ocurre luego de ser tomadas las decisiones de uso de insumos intermedios y de producción por parte de las unidades productivas agrícolas.

Como se discutió previamente en la sección 2.1, la estimación de la ecuación (2) utilizando Mínimos Cuadrados Ordinarios (OLS) de los parámetros que asocian el capital y el trabajo en la función de producción resultan inconsistentes debido a la presencia de diversos sesgos potenciales. Para abordar tales sesgos en la literatura empírica se han desarrollado

⁷ Se plantea una función de producción agregada “*Value-added production*” explícitamente para que los insumos no entren como regresores en la función de producción por estimar.

principalmente tres métodos complementarios, los cuales fueron propuestos por Olley y Pakes (1996), Levinsohn y Petrin (2003) y Ackerberg et al. (2015).

a) Método Olley y Pakes (1996). Los autores de este método proponen utilizar variables observables para controlar la productividad no observable; es decir, proponen específicamente emplear el nivel de inversión ($I_{i,t}$) para aproximar la productividad en la función de producción y un estimador semi-paramétrico consistente ante sesgos de selección. Los autores asumen que, al inicio de cada periodo, cada unidad productiva decide continuar o no continuar con sus operaciones. Si se mantiene en operación, la unidad productiva decide el nivel de factores de producción y de inversión que empleará. Consecuentemente, la unidad productiva maximiza el valor esperado de flujo de caja e inversiones sujeto a la expectativa sobre la estructura futura de producción.

Se asume que el trabajo ($LR_{i,t}$ y $LNR_{i,t}$) es un factor perfectamente flexible, la inversión ($I_{i,t}$) depende del nivel de productividad y de la tierra: $I_{i,t} = I(\omega_{i,t}, T_{i,t})$, y que la inversión es estrictamente creciente con la productividad ($\omega_{i,t}$). Las unidades productivas que esperan un shock positivo de productividad en el periodo “t” incrementaran su nivel de inversión en dicho periodo para un nivel específico de tierra. La productividad ($\omega_{i,t}$) sigue un proceso de Markov de primer orden:

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

Con ello la función inversa para la productividad puede ser expresada como condición de invertibilidad en inversión:

$$\omega_{i,t} = I^{-1}(I_{i,t}, T_{i,t}) = h(I_{i,t}, T_{i,t}) \quad (3)$$

La ecuación anterior expresa una función de control de la productividad no observable, la cual puede ser introducida en la ecuación de valor agregado. Con ello, la función de producción agregada puede ser expresada de la siguiente manera:

$$va_{i,t} = \beta_0 + \beta_{lr}lr_{i,t} + \beta_{lnr}lnr_{i,t} + \beta_\tau\tau_{i,t} + h(I_{i,t}, T_{i,t}) + \varepsilon_{i,t} \quad (4)$$

Expresando la inversión y la tierra en logaritmos, la ecuación anterior puede ser planteada como:

$$va_{i,t} = \beta_{lr}lr_{i,t} + \beta_{lnr}lnr_{i,t} + \varphi(i_{i,t}, \tau_{i,t}) + \varepsilon_{i,t} \quad (5)$$

En la ecuación (5), el término $\varphi(i_{i,t}, \tau_{i,t}) = \beta_0 + \beta_\tau\tau_{i,t} + h(I_{i,t}, T_{i,t})$, es aproximada a través de un polinomio de segundo orden. La estimación de la ecuación anterior por OLS constituye la primera etapa del método de Olley y Pakes (1996), los coeficientes estimados, relacionados con la participación del trabajo remunerado y no remunerado en el valor agregado ($\hat{\beta}_{lr}$ y $\hat{\beta}_{lnr}$), pueden ser consistentes toda vez que la función $\varphi(i_{i,t}, \tau_{i,t})$ estimada controla por la productividad no observada (función de control), donde el término de error no estaría correlacionado con los factores de producción. La ecuación (5) no identifica el coeficiente relacionado con la tierra (β_τ); para ello, en la segunda etapa del método se propone la estimación de la siguiente ecuación:

$$va_{i,t} - \hat{\beta}_{lr}lr_{i,t} + \hat{\beta}_{lnr}lnr_{i,t} - \beta_\tau\tau_{i,t} = \omega_{i,t} + \varepsilon_{i,t} \quad (6)$$

Dado que $\omega_{i,t} = g(\omega_{i,t-1}) + \xi_{i,t}$ y la función $g(\cdot)$ puede ser expresada como $g(\omega_{i,t-1}) = +\gamma_1\omega_{i,t-1} + \gamma_2\omega_{i,t-1}^2 + \gamma_3\omega_{i,t-1}^3$ donde $\omega_{i,t-1} = \hat{\varphi}(i_{i,t-1}, \tau_{i,t-1}) - \beta_\tau\tau_{i,t-1}$, con ello la ecuación anterior puede ser re-expresada como:

$$\begin{aligned}
va_{i,t} - \hat{\beta}_{lr} lr_{i,t} - \hat{\beta}_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} \\
= \gamma_1(\hat{\phi}(i_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1}) \\
+ \gamma_2(\hat{\phi}(i_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1})^2 \\
+ \gamma_2(\hat{\phi}(i_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1})^3 + \xi_{i,t} + \varepsilon_{i,t}
\end{aligned} \tag{7}$$

De la estimación de la ecuación anterior a través de Mínimos Cuadros No Lineales (NLLS) es posible identificar la participación del capital en el producto ($\hat{\beta}_\tau$).

b) Método Levinsohn y Petrin (2003). En este método, los autores asumen que el trabajo ($LR_{i,t}$ y $LNR_{i,t}$) es un factor perfectamente flexible; sin embargo a diferencia del método Olley y Pakes (1996), en lugar de la inversión se asume que la demanda por insumos ($M_{i,t}$) depende del nivel de productividad y de la tierra: $M_{i,t} = M(\omega_{i,t}, T_{i,t})$. Dicha demanda es estrictamente creciente con la productividad ($\omega_{i,t}$). De acuerdo con los autores, 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 dos argumentos: por un lado, 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. Por otro lado, en las bases de datos disponibles para la estimación de la productividad, la inversión reportada por las unidades productivas puede estar truncada a valores próximos a cero ($I_{i,t} = 0$).

La demanda de insumos ($M_{i,t}$) depende del nivel de productividad y de la tierra: $M_{i,t} = M(\omega_{i,t}, T_{i,t})$. La función inversa para la productividad puede ser expresada como (condición de invertibilidad en insumos):

$$\omega_{i,t} = M^{-1}(M_{i,t}, T_{i,t}) = f(M_{i,t}, T_{i,t}) \tag{8}$$

Donde la productividad ($\omega_{i,j,t}$) sigue un proceso de Markov de primer orden: $\omega_{i,t} = \mathbb{E}[\omega_{i,t}|\Omega_{i,t}] + \xi_{i,t} = \mathbb{E}[\omega_{i,t}|\omega_{i,t-1}] + \xi_{i,t} = g(\omega_{i,t-1}) + \xi_{i,t}$. Empleando la forma funcional asumida en la ecuación de valor agregado, esta puede ser plantada como:

$$va_{i,t} = \beta_{lr} lr_{i,t} + \beta_{lnr} lnr_{i,t} + \varphi(m_{i,t}, \tau_{i,t}) + \eta_{i,t} \quad (9)$$

En la ecuación anterior, la función $\varphi(m_{i,t}, \tau_{i,t}) = \beta_0 + \beta_\tau \tau_{i,t} + f(M_{i,t}, T_{i,t})$ es aproximada a través de un polinomio de tercer orden, la estimación de la ecuación (9) por OLS constituye la primera etapa del método. Los coeficientes estimados relacionados con la participación del trabajo en el valor agregado ($\hat{\beta}_{lr}$ y $\hat{\beta}_{lnr}$) pueden ser considerados consistentes, toda vez que la función $\varphi(m_{i,t}, \tau_{i,t})$ controla la productividad no observada.

La segunda etapa permite identificar la participación de la tierra (β_τ) en el valor agregado de la producción; para ello, los autores proponen la estimación de la siguiente ecuación:

$$va_{i,t} - \hat{\beta}_{lr} lr_{i,t} - \hat{\beta}_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} = \omega_{i,t} + \varepsilon_{i,t} \quad (10)$$

Dado que $\omega_{i,t} = g(\omega_{i,t-1}) + \xi_{i,t}$ y la función $g(\cdot)$ puede ser expresada como $g(\omega_{i,t-1}) = +\gamma_1 \omega_{i,t-1} + \gamma_2 \omega_{i,t-1}^2 + \gamma_3 \omega_{i,t-1}^3$, donde $\omega_{i,t-1} = \hat{\varphi}(m_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1}$. Con ello, la ecuación anterior puede ser re-expresada como:

$$\begin{aligned} \xi_{i,t} + \varepsilon_{i,t} &= va_{i,t} - \hat{\beta}_{lr} lr_{i,t} - \hat{\beta}_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \gamma_1(\hat{\varphi}(m_{i,t-1}, \tau_{i,t-1}) \\ &\quad - \beta_\tau \tau_{i,t-1}) + \gamma_2(\hat{\varphi}(m_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1})^2 \\ &\quad + \gamma_3(\hat{\varphi}(m_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1})^3 \end{aligned} \quad (11)$$

La ecuación anterior es estimada a través del Método Generalizado de Momentos (GMM, por sus siglas en inglés), donde es posible identificar la participación del capital en el producto ($\hat{\beta}_\tau$) a través de la solución al siguiente problema de optimización:

$$\min_{\beta_k} \sum_t \left\{ va_{i,t} - \hat{\beta}_{lr} lr_{i,t} - \hat{\beta}_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \gamma_1 (\hat{\phi}(m_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1}) \right. \\ \left. + \gamma_2 (\hat{\phi}(m_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1})^2 \right. \\ \left. + \gamma_2 (\hat{\phi}(m_{i,t-1}, \tau_{i,t-1}) - \beta_\tau \tau_{i,t-1})^3 \right\}^2 \quad (12)$$

Este método permite obtener estimaciones de la participación del trabajo y el capital consistentes ante sesgos de selección y simultaneidad.

c) Método Ackerberg, Caves y Frazer (2015). Ackerberg et al. (2007) y Ackerberg et al. (2015) plantean una crítica a los dos métodos anteriores. La crítica se enfoca en que la formulación de la función de control, basada en inversión o insumos, podría generar multicolinealidad entre los factores de producción, problema denominado como dependencia condicional. Como se mencionó en la sección 2.1, la dependencia condicional se genera porque las variables de decisión (trabajo, insumos e inversión) dependen de las mismas variables de estado (capital y productividad), lo cual implica la presencia de una relación directa entre los factores de producción generando sesgos en las estimaciones de la función de producción.

Los autores asumen que las unidades productivas agrícolas acumulan tierra de acuerdo con $\tau_{i,t} = \tau(\tau_{i,t-1}, i_{i,t-1})$, donde las decisiones de inversión son tomadas en el periodo “t-1” y las decisiones sobre la cantidad de trabajo son tomadas en el periodo “t” y “t-1”. La demanda por insumos de la unidad productiva “i” en el periodo “t” puede ser expresada como $m_{i,t} = \tilde{f}_t(\omega_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t})$, la cual depende de la productividad y la tierra, y está condicionada al factor trabajo.

Bajo el supuesto de monotonicidad estricta, la demanda $\tilde{f}_t(\cdot)$ es estrictamente creciente con la productividad $(\omega_{i,t})$, por ello es posible invertir la demanda de insumos de la forma:

$\omega_{i,t} = \tilde{f}_{i,t}^{-1}(m_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t})$, con ello, la función de producción puede ser expresada como:

$$\begin{aligned} va_{i,t} &= \beta_0 + \beta_{lr} lr_{i,t} + \beta_{lnr} lnr_{i,t} + \beta_\tau \tau_{i,t} + \tilde{f}_{i,t}^{-1}(m_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t}) + \varepsilon_{i,t} \quad (13) \\ &= \tilde{\Phi}_t(m_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t}) + \varepsilon_{i,t} \end{aligned}$$

La condición de momentos de la primera etapa puede ser planteada en los siguientes términos: $E[\varepsilon_{i,t} | \Omega_{i,t}] = E[v a_{i,t} - \tilde{\Phi}_t(m_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t}) | \Omega_{i,t}] = 0$. En la primera etapa, el parámetro asociado al trabajo no está conceptualmente identificado, mientras que la función $\tilde{\Phi}_t(m_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t})$ puede ser estimada no paramétricamente ($\widehat{\Phi}_t(m_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t})$) a través de un polinomio de segundo orden en $m_{i,t}, \tau_{i,t}, lr_{i,t}, lnr_{i,t}$.

Con lo anterior se recupera un estimado no paramétrico de $\widehat{\Phi}_t$, donde $\widehat{\Phi}_t = \beta_{lr} lr_{i,t} + \beta_{lnr} lnr_{i,t} + \beta_\tau \tau_{i,t} + \beta_m m_{i,t} + \omega_{i,t}$. Asumiendo que el término de innovación de $\omega_{i,t}$ sigue un proceso de Markov de primer orden: $\omega_{i,t} = \mathbb{E}[\omega_{i,t} | \Omega_{i,t}] + \xi_{i,t} = \mathbb{E}[\omega_{i,t} | \omega_{i,t-1}] + \xi_{i,t} = g(\omega_{i,t-1}) + \xi_{i,t}$, la función de producción de valor agregado puede ser expresada como:

$$va_{i,t} = \beta_0 + \beta_{lr} lr_{i,t} + \beta_{lnr} lnr_{i,t} + \beta_\tau \tau_{i,t} + g(\omega_{i,t-1}) + \xi_{i,t} + \varepsilon_{i,t} \quad (14)$$

Si $\eta_{i,t} = \xi_{i,t} + \varepsilon_{i,t}$, la ecuación (14) puede ser expresada como $\eta_{i,t} = va_{i,t} - \beta_0 - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - g(\omega_{i,t-1})$. La función $g(\cdot)$ es expresada a su vez: $g(\omega_{i,t-1}) = \gamma_1 \omega_{i,t-1} + \gamma_2 \omega_{i,t-1}^2 + \gamma_3 \omega_{i,t-1}^3$, como se definió anteriormente $\omega_{i,t-1} = \widehat{\Phi}_{t-1} - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \beta_m m_{i,t-1}$. Para abordar el problema de colinealidad entre los factores de producción, la segunda etapa del método implica estimar los parámetros de interés empleando la siguiente Condición de Momentos:

$$\begin{aligned}
\mathbb{E}[\eta_{i,t} | \Omega_{i,t-1}] = & E \left[va_{i,t} - \beta_0 - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} \right. \\
& - \gamma_1 \left(\widehat{\Phi}_{t-1} - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \beta_m m_{i,t-1} \right) \\
& - \gamma_2 \left(\widehat{\Phi}_{t-1} - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \beta_m m_{i,t-1} \right)^2 \\
& - \gamma_3 \left(\widehat{\Phi}_{t-1} - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} \right. \\
& \left. \left. - \beta_m m_{i,t-1} \right)^3 \right] | \Omega_{i,t-1}
\end{aligned} \tag{15}$$

De la estimación de la ecuación anterior a través del Método Generalizado de Momentos (GMM, por sus siglas en inglés) es posible estimar conjuntamente todos los parámetros de interés de la función de producción ($\hat{\beta}_{lr}$, $\hat{\beta}_{lnr}$ y $\hat{\beta}_\tau$) a través de la solución al siguiente problema de optimización:

$$\begin{aligned}
\min_{\beta_l, \beta_k} & \sum_t \left\{ va_{i,t} - \beta_0 - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} \right. \\
& - \gamma_1 \left(\widehat{\Phi}_{t-1} - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \beta_m m_{i,t-1} \right) \\
& - \gamma_2 \left(\widehat{\Phi}_{t-1} - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \beta_m m_{i,t-1} \right)^2 \\
& \left. - \gamma_3 \left(\widehat{\Phi}_{t-1} - \beta_{lr} lr_{i,t} - \beta_{lnr} lnr_{i,t} - \beta_\tau \tau_{i,t} - \beta_m m_{i,t-1} \right)^3 \right\} | \Omega_{i,t-1}
\end{aligned} \tag{16}$$

Para la solución del problema de optimización, las condiciones de ortogonalidad son: $\mathbb{E}[\tau_{i,t} \eta_{i,t}] = 0$, $\mathbb{E}[lr_{i,t-1} \eta_{i,t}] = 0$, $[lnr_{i,t-1} \eta_{i,t}] = 0$ y $\mathbb{E}[m_{i,t-1} \eta_{i,t}]$, donde la consistencia asintótica del estimador GMM depende de la ausencia de correlación serial de segundo orden en los errores del modelo.

En el problema de optimización planteado anteriormente, la estimación por GMM de la participación del trabajo ($\hat{\beta}_{lr}$ y $\hat{\beta}_{lnr}$) podría generar una mejor identificación de dicho factor,

reduciendo el sesgo por dependencia condicional respecto de los métodos de Olley y Pakes (1996) y Levinsohn y Petrin (2003). En resumen, este método permitiría obtener estimaciones de la participación del trabajo y el capital consistentes ante sesgos de selección, simultaneidad y dependencia condicional. Ello se debe a que la estimación de los coeficientes de participación del trabajo y capital se realizan simultáneamente dentro del problema de optimización y no en dos etapas como en el caso de los métodos OP y LP.

d) Estimación de la Productividad Total Factorial. La estimación de los coeficientes β_{lr} , β_{lnr} y β_τ a través de los métodos descritos anteriormente permitirían calcular la Productividad Total Factorial ($\omega_{i,t}$) de la forma:

$$PTF_{i,t}: \hat{\omega}_{i,t} = va_{i,t} - \hat{\beta}_{lr}lr_{i,t} - \hat{\beta}_{lnr}lnr_{i,t} - \hat{\beta}_\tau\tau_{i,t} \quad (17)$$

2.3 Estrategia de Identificación del Efecto del Riego en la Productividad Total Factorial

La estrategia de identificación para recuperar efectos causales del riego en la productividad se soporta en un diseño de Regresión Discontinua (RD, por sus siglas en inglés). De acuerdo con Lee y Lemieux (2010), RD es una forma de estimar el efecto de un tratamiento cuando éste es determinado por una variable observable de asignación (*forcing variable* - X), la asignación del tratamiento se realiza de manera total o parcial a partir de determinado umbral (*cutoff*). La variable de asignación o *forcing variable* está asociada con los resultados potenciales, pero dicha asociación es continua a lo largo de una muestra de observaciones, cualquier discontinuidad en la esperanza condicional es interpretada como el efecto causal del tratamiento alrededor de un determinado umbral.

En general, el diseño de RD se soporta sobre los siguientes supuestos:

- Los individuos, familias o agentes económicos no pueden manipular perfectamente la variable de asignación ni el umbral a partir del cual se recibe el tratamiento.

- La variación en el estatus de tratamiento en la proximidad del umbral puede ser considerado como aleatorio. Cuando los agentes tienen un control imperfecto sobre la variable de asignación, los agentes próximos al umbral tienen aproximadamente la misma probabilidad de recibir el tratamiento por encima del umbral y de no recibir el tratamiento por debajo del umbral.
- Cualquier otro factor es continuo respecto de la variable de asignación.

Partiendo del modelo causal de resultados potenciales de Rubin. Para cada agente “i” existe un par de resultados potenciales $Y_i(1)$ si “i” recibe o está expuesto al tratamiento y $Y_i(0)$ si no recibe o no está expuesto al tratamiento. El efecto causal del tratamiento puede ser representado como la diferencia $Y_i(1) - Y_i(0)$. El problema fundamental de la evaluación de impacto es que no se puede observar $Y_i(1)$ y $Y_i(0)$ simultáneamente. En un diseño RD, se asumen dos relaciones de los resultados potenciales condicionadas a la variable de asignación: $E[Y_i(1)|X]$ y $E[Y_i(0)|X]$, donde todos los individuos ubicados a un lado de un umbral c de X están expuestos al tratamiento y los ubicados al otro lado del umbral c no están expuestos al tratamiento. El efecto causal puede ser representado como:

$$E[Y_i(1) - Y_i(0)|X = c] = \lim_{\varepsilon \downarrow 0} E[Y_i(1)|X = c + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[Y_i(0)|X = c + \varepsilon] \quad (18)$$

La ecuación anterior representa el Efecto Causal Promedio (ATE) alrededor del umbral c . La regla de asignación del tratamiento (D_i) es una función determinística de la variable de asignación (X_i), de modo que: $D_i = 1\{X_i \geq c\}$. No obstante, la probabilidad de recibir tratamiento no necesariamente cambia de 0 a 1 en el umbral c . Lo cual implica que: $\lim_{\varepsilon \downarrow 0} Pr(D_i = 1|X = c + \varepsilon) \neq \lim_{\varepsilon \uparrow 0} Pr(D_i = 1|X = c - \varepsilon)$, aspecto definido como

cumplimiento imperfecto. En tales casos, el diseño RD toma el nombre de Fuzzy RD, se pueden generar tres casos:

- (i) Cumplidores: $\lim_{X \downarrow c} D_i(X) = 0, \lim_{X \uparrow c} D_i(X) = 1$
- (ii) Nunca tomadores: $\lim_{X \downarrow c} D_i(X) = 0, \lim_{X \uparrow c} D_i(X) = 0$
- (iii) Siempre tomadores: $\lim_{X \downarrow c} D_i(X) = 1, \lim_{X \uparrow c} D_i(X) = 1$.

Lo anterior implica que la probabilidad de tratamiento salta menos de 1 en el umbral, y por lo tanto la relación entre Y_i y X_i no puede ser interpretada como un efecto causal. El efecto causal puede ser recuperado dividiendo la discontinuidad en la relación entre Y_i y X_i en el umbral c respecto del cambio en la probabilidad de tratamiento generado en el umbral c ; en otros términos:

$$\tau_{FRD} = \frac{\lim_{\varepsilon \downarrow 0} E[Y|X = c + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[Y|X = c + \varepsilon]}{\lim_{\varepsilon \downarrow 0} E[D|X = c + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[D|X = c + \varepsilon]} \quad (19)$$

De acuerdo con Lee y Lemieux (2010), para que τ_{FRD} sea interpretado como el efecto causal en el umbral c , son necesarios los siguientes supuestos de monotonicidad ($D_i(X)$ es no creciente en $X, X = c$) y excluibilidad (cruzando el umbral, la variable de asignación no tiene impacto en el resultados de interés excepto a través del impacto en la probabilidad de exposición al tratamiento). Cuando tales supuestos se cumplen $\tau_{FRD} = E[Y_i(1) - Y_i(0)|X = c]$ es el Efecto Local Causal Promedio (LATE, por sus siglas en inglés) para el caso de cumplidores. Cumplidores son unidades que reciben el tratamiento cuando cumplen la regla de asignación $X_i \geq c$. LATE es interpretado como un efecto causal ponderado, donde los ponderadores reflejan la probabilidad de proximidad al umbral. De similar modo, en

el contexto de RD τ_{FRD} puede ser interpretado como la “Intención de Tratamiento” o ITT (por sus siglas en inglés).

3. Resultados

3.1 Estimación de la Función de Producción

Los resultados presentados en esta sección se basan en la Encuesta Económica Anual de los años 2016, 2017 y 2018. El análisis se ha restringido a aquellas unidades agrícolas que reportaron información de todas las variables necesarias para la estimación de la función de producción por al menos tres años consecutivos, la muestra contiene información de 5803 unidades agrícolas en una estructura de panel balanceado.

En la Tabla 2 se reportan las estadísticas descriptivas básicas de las variables consideradas para la estimación de la función de producción: producción, trabajo, tierra e insumos⁸. La producción total agrícola fue de S/. 9886, el valor agregado fue de S/. 8168. Además, las unidades agropecuarias cuentan en promedio con ocho hectáreas trabajadas, utilizan en promedio once trabajadores remunerados y dos trabajadores familiares no remunerados. Gastan en promedio S/. 1785 en insumos (semillas, abonos, fertilizantes) y tienen un nivel de inversión promedio de S/. 427.

Tabla 2

Estadísticas descriptivas: variables para la estimación de la función de producción

Variables	Obs.	Media	Desviación estandar	Min.	Max
Producción	17409	9885.68	32797.4	0	1264200
Valor agregado	17409	8168.30	27140.8	0	968840
Tierra	17409	7.68	39.3	0	2000
Trabajo remunerado	17409	11.00	39.2	0	2063
Trabajo familiar no remunerado	17409	1.89	1.6	0	14
Insumos	17409	1784.76	7796.1	0	308530
Inversión	17409	427.13	2984.8	0	104000
(Log) Producción	16690	7.46	1.9	2	13

⁸ En la Tabla A.2 del anexo se presentan las estadísticas descriptivas del resto de las variables consideradas en el análisis y de las variables potencialmente relacionadas con la calidad de la tierra.

(Log) Valor agregado	15611	7.33	1.9	1	13
(Log) Tierra	17202	1.24	1.0	0	4
(Log) Trabajo remunerado	17321	1.44	1.2	0	5
(Log) Trabajo no remunerado	17408	0.90	0.6	0	3
(Log) Insumos	14957	5.55	2.0	0	12
(Log) Inversión	3012	6.26	1.6	1	11

Fuente: Encuesta Nacional Agropecuaria: muestra panel 2016-2017-2018 (INEI)

Elaboración Propia

La participación del trabajo remunerado, trabajo no remunerado y la tierra en la función de producción están representadas empíricamente por los coeficientes β_{lr} , β_{lnr} y β_{τ} , respectivamente, los cuales fueron estimados a través de cinco métodos: Mínimos Cuadrados Ordinarios (OLS, por sus siglas en inglés); Efectos Fijos a nivel de unidad agrícola (FE, por sus siglas en inglés)⁹; Olley y Pakes (OP); Levinhson y Petrin (LP) y Ackerberg, Caves y Fraser (ACF). En la Tabla 3 se reportan los resultados de la estimación de los coeficientes relacionados a los factores de producción con base en los métodos mencionados anteriormente, en general se encuentra que la magnitud de los coeficientes estimados difiere en función al método de estimación empleado, lo cual permite identificar la dirección de los sesgos potenciales en tales coeficientes.

En la columna (1) de la Tabla 3 se muestran los resultados estimados con OLS, donde los coeficientes estimados relacionados con la participación del trabajo remunerado, trabajo familiar no remunerado y tierra en la función de producción son: $\hat{\beta}_{lr,OLS} = 0.589$, $\hat{\beta}_{lnr,OLS} = 0.030$ y $\hat{\beta}_{\tau,OLS} = 0.737$ respectivamente, resultando significativos estadísticamente los coeficientes asociados al trabajo remunerado y la tierra.

En la columna (2) se presentan los coeficientes estimados por FE, los efectos fijos se consideran a nivel de unidades agrícolas, los coeficientes estimados relacionados con la

⁹ Los dos primeros métodos no fueron descritos en la parte metodológica del documento, para mayores detalles sobre ambos ver Wooldridge (2002).

participación del trabajo remunerado, trabajo familiar no remunerado y de la tierra son: $\hat{\beta}_{lr,FE} = 0.210$, $\hat{\beta}_{lnr,FE} = 0.136$ y $\hat{\beta}_{\tau,FE} = 0.616$ respectivamente, tales coeficientes resultaron significativos estadísticamente al 1%. De acuerdo con Van Beveren (2012), el método de FE corregiría el sesgo por simultaneidad ante la presencia de variables omitidas en la estimación por OLS, esto se observa en el coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{lr,FE}$) y también en el coeficiente relacionado con la tierra ($\hat{\beta}_{\tau,OLS} < \hat{\beta}_{\tau,FE}$).

En la columna (3) se muestran los resultados de la estimación de la función de producción por OP, donde los coeficientes estimados relacionados con la participación del trabajo remunerado, trabajo familiar no remunerado y de la tierra en la función de producción son: $\hat{\beta}_{lr,OP} = 0.429$, $\hat{\beta}_{lnr,OP} = 0.252$ y $\hat{\beta}_{\tau,OP} = 1.00$ respectivamente; tales coeficientes resultaron significativos estadísticamente al 1%. Comparando los coeficientes estimados por OP respecto de OLS y FE, se observa que el método OP daría cuenta de los posibles sesgos en la estimación del coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{lr,OP} > \hat{\beta}_{lr,FE}$). Por su parte, el coeficiente relacionado con la tierra resultó mayor respecto del método OLS y FE ($\hat{\beta}_{\tau,OP} > \hat{\beta}_{\tau,OLS} > \hat{\beta}_{\tau,FE}$).

En la columna (4) se muestran los resultados de la estimación de la función de producción a través del método LP, donde los coeficientes estimados relacionados con la participación del trabajo remunerado, trabajo familiar no remunerado y de la tierra en la función de producción son: $\hat{\beta}_{lr,LP} = 0.201$, $\hat{\beta}_{lnr,LP} = -0.027$ y $\hat{\beta}_{\tau,LP} = 0.860$ respectivamente, los coeficientes relacionados con el trabajo remunerado y la tierra resultaron significativos estadísticamente al 1%. Comparando los coeficientes estimados por LP respecto de OLS y FE, se observa que el método LP da cuenta de los posibles sesgos en la estimación del coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{lr,FE} > \hat{\beta}_{lr,LP}$). El coeficiente relacionado con la tierra resultó mayor respecto del método OLS y FE ($\hat{\beta}_{\tau,LP} > \hat{\beta}_{\tau,OLS} > \hat{\beta}_{\tau,FE}$).

Finalmente, en la columna (5) se muestran los resultados de la estimación de la función de producción a través del método ACF, la participación del trabajo remunerado $\hat{\beta}_{lr,ACF} = 0.064$ (significativo estadísticamente al 1%), la participación del trabajo familiar no remunerado es $\hat{\beta}_{lnr,ACF} = 0.233$ (significativo estadísticamente al 5%); mientras que la participación de la tierra es $\hat{\beta}_{\tau,ACF} = 1.01$ (significativo estadísticamente al 1%). Comparando los coeficientes estimados por ACF respecto de OLS y FE, se observa que el método ACF daría cuenta de los posibles sesgos en la estimación del coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{lr,FE} > \hat{\beta}_{lr,ACF}$) y de similar modo con los métodos OP y LP, conceptualmente el método ACF daría cuenta de los posibles sesgos generados endogeneidad y dependencia condicional, en el factor trabajo.

Por su parte, el coeficiente estimado por el método ACF relacionado con la tierra resultó mayor que el estimado por FE, lo cual sugiere que, si bien el método FE daría cuenta del problema de endogenidad en la participación del capital en el producto, no corregiría el problema de dependencia condicional, el cual sería abordado de mejor manera por el método ACF, ofreciendo una aproximación relativamente más precisa que los métodos OP y LP.

Tabla 3

Estimación de la función de producción

	(1)	(2)	(3)	(4)	(5)
	OLS	FE	OP	LP	ACF
Trabajo remunerado: $\hat{\beta}_{lr}$	0.589** *	0.210** *	0.429** *	0.201** *	0.064** (0.012)
Trabajo familiar no remunerado: $\hat{\beta}_{lnr}$	0.030 (0.026)	0.136** *	0.252** *	-0.027 (0.019)	0.233** (0.109)
Tierra: $\hat{\beta}_{\tau}$	0.737** *	0.616** *	1.00*** (0.035)	0.860** *	1.01*** (0.034)
Obs.	15351	15351	15351	13154	8066

Notas:

- (1) OLS: Mínimos Cuadrados Ordinarios.
- (2) FE: Panel de datos con efectos fijos (efectos fijos a nivel de unidades agrícolas).
- (3) OP: Olley y Pakes (1996).
- (4) LP: Levinsohn y Petrin (2003).
- (5) ACF: Ackerberg, Caves y Frazer (2015).

Los símbolos ** y *** representan significancia estadística al 5% y 1%, respectivamente. Errores estándar estimados por cluster a nivel de unidades agropecuarias (200 replicaciones). Elaboración Propia.

3.2 Estimación de la Función de Producción Controlando por Calidad de la Tierra

Como se puede apreciar en la Tabla 3, la tierra es el factor más relevante en la función de producción, y además fue aproximado a través de la cantidad de hectáreas cosechadas por la unidad agrícola; sin embargo, a pesar de incluir en la función de producción una medición razonable del factor tierra, es de esperar que existan diferencias en la calidad de la tierra, lo cual podría afectar los resultados de las estimaciones. En particular, de no considerar explícitamente la calidad de la tierra en la función de producción, este aspecto sería absorbido por el término estocástico de la función de producción y podría generar que el factor tierra esté correlacionado con el término estocástico, incumpliendo así la condición de ortogonalidad entre la tierra y el error estocástico: $\mathbb{E}[\tau_{i,t}\eta_{i,t}] \neq 0$.

De acuerdo con Restuccia y Santaella-Llopis (2017), es posible incluir un indicador de calidad de la tierra en la función de producción. Con ello la ecuación (2) puede ser expresada como:

$$va_{i,t} = \beta_0 + \beta_{lr}lr_{i,t} + \beta_{lnr}lnr_{i,t} + \beta_\tau(q_{i,t}\tau_{i,t}) + \omega_{i,t} + \varepsilon_{i,t} \quad (20)$$

El factor tierra controlando por calidad ($q_{i,t}\tau_{i,t}$), puede ser expresado como: $(q_{i,t}\tau_{i,t}) = \tau_{i,t}^o + q_{i,t} + (q_{i,t} * \tau_{i,t}^o)$, donde $\tau_{i,t}^o$ es la cantidad de tierra empleada y $q_{i,t}$ es un índice de la calidad de la tierra. Por disponibilidad de información, se ha considerado un conjunto de variables demográficas, agrícolas y medioambientales que podrían estar relacionadas con la calidad de la tierra ($q_{i,t}$), las cuales se describen en la Tabla A.1 del Anexo. Se han considerado

las siguientes: demográficas (sexo, edad, tamaño familiar, años de educación, lengua materna); realización de prácticas agrícolas (análisis de suelos, rotación de cultivos, uso de terrazas o zanjas de infiltración, nivelación de terreno, aplicación de control biológico, manejo integrado de plagas, etc.) y medioambientales (altitud promedio a nivel de conglomerado, temperatura promedio a nivel regional, precipitación promedio a nivel regional, humedad promedio a nivel regional).

En la Tabla A.2 del Anexo, se presentan las estadísticas descriptivas de las variables mencionadas. De acuerdo con Mugizi y Matsumoto (2020), con el conjunto de variables relacionadas con la calidad de la tierra es posible construir un índice de calidad de la tierra ($\tilde{q}_{i,t}$). Para ello, se consideró emplear el Análisis de Componentes Principales (PCA, por sus siglas en inglés). El PCA es un método que busca sintetizar la dimensionalidad de un conjunto de variables aleatorias conservando su variabilidad subyacente.

Formalmente, considerando un vector de variables aleatorias (X), con una matriz de varianzas y covarianzas $[V(X)]$. Lo cual implica estimar un conjunto de vectores orto-normales: $a(1), a(2), \dots, a(k)$ que en conjunto resuelven el problema de optimización:

$$\begin{aligned}
 a(1) &= \underset{a:|a|=1}{\operatorname{argmax}} V[a'X] \\
 a(2) &= \underset{\substack{a:|a|=1 \\ a \perp a(1)}}{\operatorname{argmax}} V[a'X] \\
 a(k) &= \underset{\substack{a:|a|=1 \\ a \perp a(1), \dots, a \perp a(k-1)}}{\operatorname{argmax}} V[a'X]
 \end{aligned} \tag{21}$$

El método de PCA permite construir componentes que resumen información contenida en un conjunto mayor de variables que el caso del presente estudio, se asume que dicho conjunto puede determinar la calidad de la tierra de las unidades agrícolas. Los vectores $a(k)$ obtenidos en la ecuación anterior permiten el cálculo de componentes para un vector dado de “p” variables aleatorias ($X=x(1); x(2); \dots x(p)$). El k -ésimo componente es definido por:

$$\tilde{q}_{i,t,k} = q_{i,t,k}X = q_{1,i,t,k}x_1 + q_{2,i,t,k}x_2 + \dots + q_{i,t,p,k}x_p \quad (22)$$

El primer componente (q_1), explica la fracción más importante $[V(X)]$ que los demás componentes siguientes, el segundo componente (q_2) explica la segunda fracción más importante y así sucesivamente. La información disponible se combina en un solo indicador que es una variable aproximada de la calidad de la tierra usada ($\tilde{q}_{i,t,k}$). En la Tabla A.3 del Anexo se presentan los resultados de la estimación del índice de calidad de la tierra ($\tilde{q}_{i,t,k}$). Para esta estimación se analizó la correlación lineal entre las variables consideradas, depuraron todas aquellas variables altamente correlacionadas entre sí para reducir la colinealidad exacta entre variables, así como, las variables que no permitían retener un primer componente del PCA que explicara más del 50% de la varianza.

La calidad predictiva de la varianza fue estimada a través del estadístico Kaiser, Meyer, Olkin. Dicho estadístico reporta un valor entre 0 y 1, cuando el valor estimado está cercano a cero indica que el conjunto de variables consideradas tiene poco en común para garantizar un componente principal; mientras que, un valor cercano a uno representa una mejor calidad explicativa. La calidad predictiva estimada fue de 0.821 (ver Tabla A.3 del Anexo), en general, dicha la calidad predictiva se considera razonable.

La participación del trabajo y la tierra en la función de producción están representadas empíricamente por los coeficientes β_{lr} , β_{lnr} y β_{τ} de la ecuación (20), donde se incluyen indicadores adicionales para controlar por calidad de la tierra. Dichos coeficientes fueron estimados a través de cinco métodos: Mínimos Cuadrados Ordinarios (OLS, por sus siglas en inglés), Efectos Fijos (FE, por sus siglas en inglés), Olley y Pakes (OP), Levinhson y Petrin (LP), y Ackerberg, Caves y Fraser (ACF). En la Tabla 4 se reportan los resultados.

En la columna (1) de la Tabla 4 se muestran los resultados estimados con OLS, donde los coeficientes estimados relacionados con la participación del trabajo remunerado, trabajo

familiar no remunerado y tierra en la función de producción son: $\hat{\beta}_{lr,OLS} = 0.632$, $\hat{\beta}_{lnr,OLS} = -0.009$ y $\hat{\beta}_{\tau,OLS} = 0.631$ respectivamente, resultando significativos estadísticamente los coeficientes asociados al trabajo remunerado y la tierra. Los coeficientes relacionados con el indicador de calidad de la tierra y la interacción entre la calidad y cantidad de tierra usada resultaron: $\beta_{\tau,OLS}^{(q_{i,t})} = -0.504$ y $\beta_{\tau,OLS}^{(q_{i,t}*\tau_{i,t})} = 0.095$, ambos significativos estadísticamente al 1%.

En la columna (2) se muestran los coeficientes estimados por FE (efectos fijos a nivel de unidad agrícola) relacionados con la participación del trabajo remunerado, trabajo familiar no remunerado y de la tierra corregida por calidad son: $\hat{\beta}_{lr,FE} = 0.205$, $\hat{\beta}_{lnr,FE} = 0.141$ y $\hat{\beta}_{\tau,FE} = 0.636$ respectivamente; tales coeficientes resultaron significativos estadísticamente al 1%. Los coeficientes relacionados con el indicador de calidad de la tierra y la interacción entre la calidad y cantidad de tierra usada resultaron: $\beta_{\tau,FE}^{(q_{i,t})} = 0.138$ y $\beta_{\tau,OLS}^{(q_{i,t}*\tau_{i,t})} = 0.066$, ambos significativos estadísticamente al 1%. Cómo se mencionó anteriormente, el método de FE corregiría el sesgo por simultaneidad en la estimación de la función de producción, esto se observa en el coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{lr,FE}$), y en menor medida en el coeficiente relacionado con la tierra luego de controlar por calidad de dicho factor. Respecto del coeficiente relacionado con la calidad de la tierra, las discrepancias entre OLS y FE sugerirían la existencia de variables no observables que determinan la calidad de tierra que generan un sesgo de atenuación en el coeficiente estimado por OLS.

En la columna (3) se muestran los resultados de la estimación de la función de producción por el método OP, los coeficientes estimados relacionados con la participación del trabajo remunerado, trabajo familiar no remunerado y de la tierra (corregida por calidad) en la función de producción son: $\hat{\beta}_{lr,OP} = 0.435$, $\hat{\beta}_{lnr,OP} = -182$ y $\hat{\beta}_{\tau,OP} = 0.940$ respectivamente, los coeficientes relacionados con el trabajo remunerado y la tierra resultaron significativos estadísticamente al 1%. Los coeficientes relacionados con el indicador de calidad de la tierra y

la interacción entre la calidad y cantidad de tierra usada resultaron: $\beta_{\tau,OP}^{(q_{i,t})} = -0.020$ y $\beta_{\tau,OLS}^{(q_{i,t}*\tau_{i,t})} = -0.013$, ambos significativos estadísticamente al 1%. Comparando los coeficientes estimados por OP respecto de OLS y FE, se observa que el método OP da cuenta de los posibles sesgos en la estimación del coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{lr,OP} > \hat{\beta}_{l,FE}$). Por su parte, el coeficiente relacionado con la tierra resultó mayor respecto del método OLS y FE ($\hat{\beta}_{\tau,OP} > \hat{\beta}_{\tau,FE} > \hat{\beta}_{\tau,OLS}$), resultado congruente con los potenciales sesgos generados por la estimación de OLS.

En la columna (4) se muestran los resultados de la estimación de la función de producción a través del método LP, los coeficientes estimados relacionados con la participación del trabajo remunerado, trabajo familiar no remunerado y de la tierra (controlando por calidad) en la función de producción son: $\hat{\beta}_{lr,LP} = 0.233$, $\hat{\beta}_{lnr,LP} = -0.012$ y $\hat{\beta}_{\tau,LP} = 0.828$ respectivamente, los coeficientes relacionados con el trabajo remunerado y la tierra resultaron significativos estadísticamente al 1%. Los coeficientes relacionados con el indicador de calidad de la tierra y la interacción entre la calidad y cantidad de tierra usada resultaron: $\beta_{\tau,LP}^{(q_{i,t})} = -0.177$ y $\beta_{\tau,OLS}^{(q_{i,t}*\tau_{i,t})} = -0.148$, ambos significativos estadísticamente al 1%. Comparando los coeficientes estimados por LP respecto de OLS, se observa que el método LP da cuenta de los posibles sesgos en la estimación del coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{l,LP}$). El coeficiente relacionado con la tierra resultó mayor respecto del método OLS y FE ($\hat{\beta}_{\tau,LP} > \hat{\beta}_{\tau,FE} > \hat{\beta}_{\tau,OLS}$).

Finalmente, en la columna (5) se muestran los resultados de la estimación de la función de producción a través del método ACF. De acuerdo con dicha aproximación, la participación del trabajo remunerado en el producto es $\hat{\beta}_{lr,ACF} = 0.098$ (significativo estadísticamente al 1%), la participación del trabajo familiar no remunerado en el producto es $\hat{\beta}_{lnr,ACF} = 0.214$ (no significativo estadísticamente); mientras que la participación de la tierra (corregida por

calidad de la tierra) resultó $\hat{\beta}_{\tau,ACF} = 0.891$ (significativo estadísticamente al 1%). Los coeficientes relacionados con el indicador de calidad de la tierra y la interacción entre la calidad y cantidad de tierra usada resultaron: $\beta_{\tau,ACF}^{(q_{i,t})} = -0.243$ y $\beta_{\tau,OLS}^{(q_{i,t}*\tau_{i,t})} = -0.032$, donde solamente el coeficiente asociado al indicador de calidad de la tierra resultó significativo estadísticamente al 1%.

Comparando los coeficientes estimados por ACF respecto de OLS y FE, se observa que el método ACF da cuenta de los posibles sesgos en la estimación del coeficiente relacionado con el trabajo remunerado ($\hat{\beta}_{lr,OLS} > \hat{\beta}_{lr,FE} > \hat{\beta}_{lr,ACF}$). Esto sugeriría que conceptualmente el método ACF daría cuenta de sesgos generados por simultaneidad por variables omitidas y dependencia condicional entre factores de producción.

Por su parte, el coeficiente estimado relacionado con la tierra, corregida por calidad, resultó mayor que el estimado por FE, lo cual sugiere que si bien el método FE daría cuenta del sesgo por simultaneidad en la participación de la tierra en el producto, no corregiría el problema de dependencia condicional, problema que sería corregido por el método ACF ($\hat{\beta}_{\tau,ACF} > \hat{\beta}_{\tau,FE} > \hat{\beta}_{\tau,OLS}$). En resumen, los sesgos por simultaneidad (variables omitidas y errores de medición) y dependencia condicional (colinealidad entre los factores de producción) en la estimación de función de producción serían abordados de mejor manera por el método ACF, ofreciendo estimaciones relativamente más precisas de la contribución de los factores productivos en el valor agregado de la producción agrícola.

Tabla 4

Estimación de la función de producción, controlando por calidad de la tierra

	(1)	(2)	(3)	(4)	(5)
	OLS	FE	OP	LP	ACF
Trabajo remunerado: $\hat{\beta}_{lr}$	0.632** *	0.205** *	0.435** *	0.233** *	0.098** *
	(0.012)	(0.013)	(0.014)	(0.011)	(0.024)
	-0.009	0.141** *	-0.182	-0.012	0.214

Trabajo familiar no remunerado:	(0.025)	(0.030)	(0.035)	(0.022)	(0.113)
$\hat{\beta}_{lnr}$	0.631**	0.636**	0.940**	0.828**	0.891**
Tierra: $\hat{\beta}_\tau$	*	*	*	*	*
	(0.025)	(0.033)	(0.032)	(0.035)	(0.033)
Calidad de la tierra: $\hat{\beta}_\tau^{(q_{i,t})}$	-	0.138**	-	-	-
	0.504**	*	0.020**	0.177**	0.243**
	*	*	*	*	*
	(0.025)	(0.037)	(0.035)	(0.022)	(0.029)
Tierra usada*calidad de la tierra:	0.095**	0.066**	-	-	-
$\hat{\beta}_\tau^{(q_{i,t}*\tau_{i,t})}$	*	*	0.013**	0.148**	-0.032
	*	*	*	*	
	(0.016)	(0.021)	(0.022)	(0.015)	(0.018)
Obs.	14979	14979	14979	13154	8066

Notas:

(1) OLS: Mínimos Cuadrados Ordinarios.

(2) FE: Panel de datos con efectos fijos (efectos fijos a nivel de unidades agrícolas).

(3) OP: Olley y Pakes (1996).

(4) LP: Levinsohn y Petrin (2003).

(5) ACF: Ackerberg, Caves y Frazer (2015).

Calidad de la tierra: Índice construido por Componentes Principales con base en variables demográficas, prácticas agrícolas y medioambientales.

Los símbolos ** y *** representan significancia estadística al 5% y 1%, respectivamente.

Errores estándar estimados por cluster a nivel de unidades agropecuarias (200 replicaciones).

Elaboración Propia.

En la Tabla 5, se resumen los resultados obtenidos para la estimación de la función de producción considerando también la calidad de la tierra. El coeficiente relacionado con el factor tierra, estimado tanto por OLS como por ACF, resultó sobreestimado en las regresiones que no controlan por calidad de la tierra; esto último sugeriría que no incorporar la calidad de la tierra podría generar sesgo por variable omitida, la cual se correlacionaría positivamente con el error estocástico de la función de producción.

Considerando lo anterior, se encuentra que empleando el método ACF para la estimación de la función de producción e incorporaron indicadores que controlan por calidad de la tierra. Así, la participación del trabajo remunerado en la producción agrícola es de 0.10 y la participación de la tierra es 0.90. Tales estimaciones serían consistentes ante sesgos por simultaneidad y dependencia condicional. Adicionalmente, se encuentra evidencia que no es

posible rechazar la hipótesis nula de presencia de rendimientos constantes a escala en la función de producción agrícola para el caso peruano ($p\text{-value}=0.706$).

Tabla 5

Estimación de la función de producción, controlando por calidad de la tierra

	Sin controlar por calidad de la tierra		Controlando por calidad de la tierra	
	OLS	ACF	OLS	ACF
Trabajo remunerado	0.589** *	0.064** (0.012)	0.632*** (0.012)	0.098*** (0.024)
Trabajo familiar no remunerado	0.030 (0.026)	0.233** (0.109)	-0.009 (0.025)	0.214 (0.113)
Tierra	0.737** *	1.01*** (0.039)	0.631*** (0.025)	0.891*** (0.033)
Obs.	15351	8066	14979	8066

Notas:

(1) OLS: Mínimos Cuadrados Ordinarios.

(2) ACF: Ackerberg, Caves y Frazer (2015).

Calidad de la tierra: Índice construido por Componentes Principales con base en variables demográficas, prácticas agrícolas y medioambientales.

Los símbolos ** y *** representan significancia estadística al 5% y 1%, respectivamente.

Errores estándar estimados por cluster a nivel de unidades agropecuarias (200 replicaciones).

Elaboración Propia.

3.3 Estimación de la Productividad Total Factorial

De acuerdo con la ecuación (17), la Productividad Total Factorial (PTF) puede ser aproximada a través de residuo de Solow. Para determinar la participación del trabajo de la tierra se ha empleado los coeficientes estimados por el método ACF controlando por la calidad de la tierra. La estimación de la PTF para la muestra de unidades agrícolas se reporta en la Tabla

6, en dicha table se presentan diversos estadísticos, entre ellos: media, desviación estándar, así como los valores en el percentil 25 (P25) y el percentil 75 (P75) y la correspondiente ratio P75/P25. El valor promedio estimado de la productividad resultó 5.8 expresado en logaritmos, lo cual sería equivalente a aproximadamente S/. 330, lo cual sugeriría que la productividad promedio representaría el 22% del valor agregado de la producción de unidades agropecuarias (equivalente a S/ 1525 en la muestra empleada).

Tabla 6
Estimación de la Productividad Total

Factorial

Media	5.80
Desviación estándar	1.64
P25	4.76
P75	6.92
Ratio P75/P25	1.45
Obs.	15331

Notas:

Estimaciones por ACF: Ackerberg, Caves y Frazer (2015).

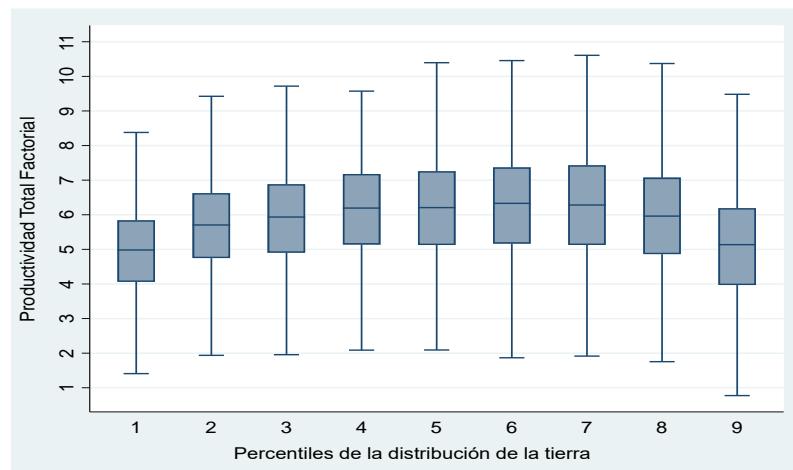
Elaboración Propia.

En los Gráficos 1 al 8 se muestra la relación entre la PTF estimada y diversas características de las unidades agrícolas, entre ellas: tierra utilizada; gasto en insumos; dominios geográficos (costa, sierra y selva); altitud, años de educación; edad de los productores y acceso al riego. Los datos se presentan en gráficos de caja con la finalidad de proporcionar una visión general de la simetría de la distribución de la PTF en función de las características de las unidades agrícolas, permite observar la dispersión de la información con la mediana (centro de la caja), los percentiles 25 y 75 que representan los límites inferior y superior de la caja, respectivamente, así como los valores máximo y mínimo de la distribución de datos.

En el Gráfico 1 se muestra la estimación de la Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por percentiles de la distribución de tierra. En primer término, se observa que la PTF se incrementa desde los percentiles más bajos hasta el sexto percentil; luego de dicho percentil la PTF se reduce considerablemente. En los percentiles más altos, los cuales corresponden a unidades agropecuarias con mayor cantidad de tierra utilizada (unidades con más de siete hectáreas) la productividad es menor, lo que sugiere la potencial presencia de mala asignación de dicho factor. En segundo lugar, se observa que la PTF es relativamente más dispersa en los percentiles más altos de la distribución de la tierra.

Gráfico 1

Productividad Total Factorial por percentiles de la distribución de tierra



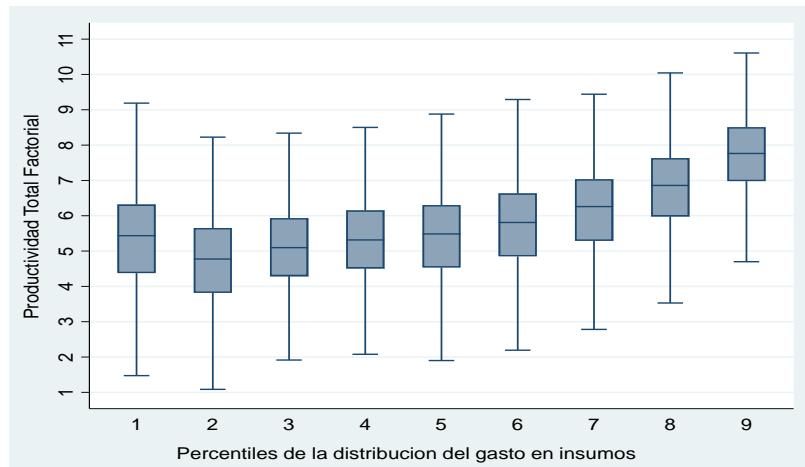
PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).

Elaboración Propia.

En el Gráfico 2 se muestra la estimación de la Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por percentiles de la distribución del gasto en insumos. Se observa que la PTF se incrementa de manera sostenida desde los percentiles más bajos hasta el percentil más alto del gasto en insumos. Adicionalmente, se observa que la PTF es relativamente más dispersa en los percentiles más bajos del gasto en insumos.

Gráfico 2

Productividad Total Factorial por percentiles de la distribución del gasto en insumos

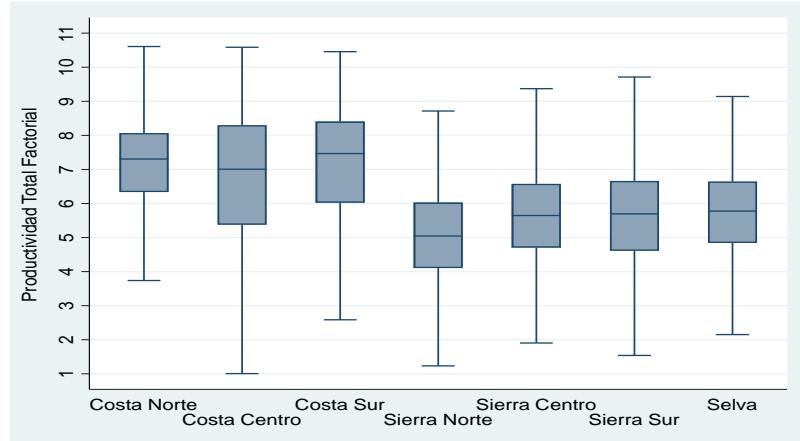


PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).

Elaboración Propia.

En el Gráfico 3 se muestra la estimación de la Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por dominio geográfico: costa (norte, centro y sur), sierra (norte, centro y sur) y selva (único dominio). En primer término, el mayor nivel de la PTF (mediana) corresponde a la costa norte, seguido por el resto de dominios de la costa peruana, sur y centro; en contraste, el menor nivel de productividad corresponde a los dominios de la sierra: sierra norte, centro y sur. En segundo término, la mayor dispersión de la PTF corresponde a los dominios geográficos de costa centro y costa sur. En general, se observa una marcada diferencia en la productividad de las unidades agrícolas de la costa y el resto de dominios geográficos.

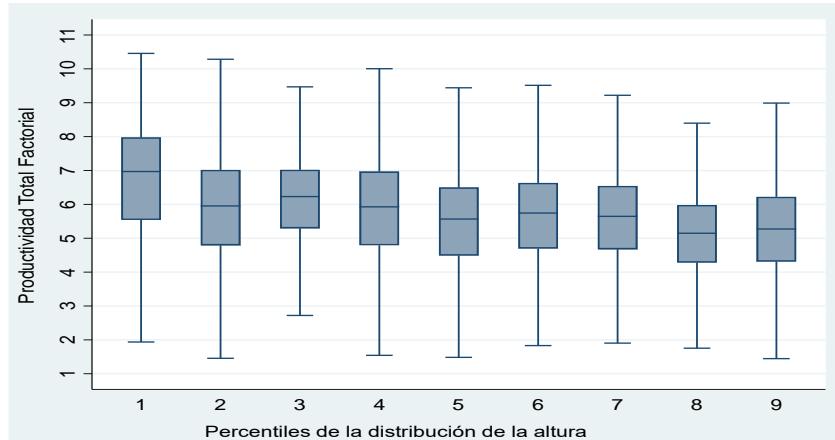
Gráfico 3.
Productividad Total Factorial por dominios geográficos



PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).
Elaboración Propia.

En el Gráfico 4 se muestra la estimación de a Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por percentiles de la distribución de la altura. Se observa que la PTF se reduce de manera relativamente sostenida desde los percentiles más bajos de altitud hasta el percentil más altos. Adicionalmente, se observa que la PTF es relativamente más dispersa en los percentiles más bajos de la distribución de la altura.

Gráfico 4
Productividad Total Factorial por percentiles de la distribución de la altura

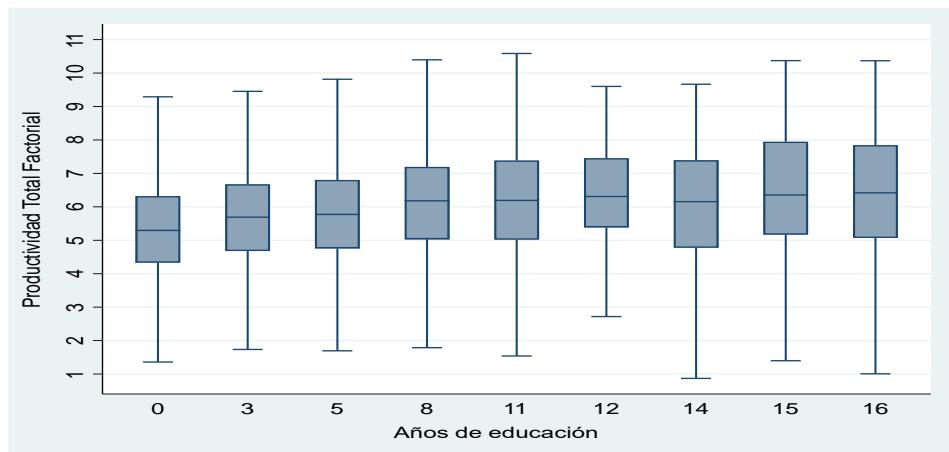


PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).
Elaboración Propia.

En el Gráfico 5 se muestra la estimación de la Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por años de educación de los productores agrícolas. Se observa que la PTF se incrementa con los años de educación de los productores; no obstante, en los niveles más altos de educación se observa una mayor dispersión de la PTF.

Gráfico 5

Productividad Total Factorial por años de educación de los productores

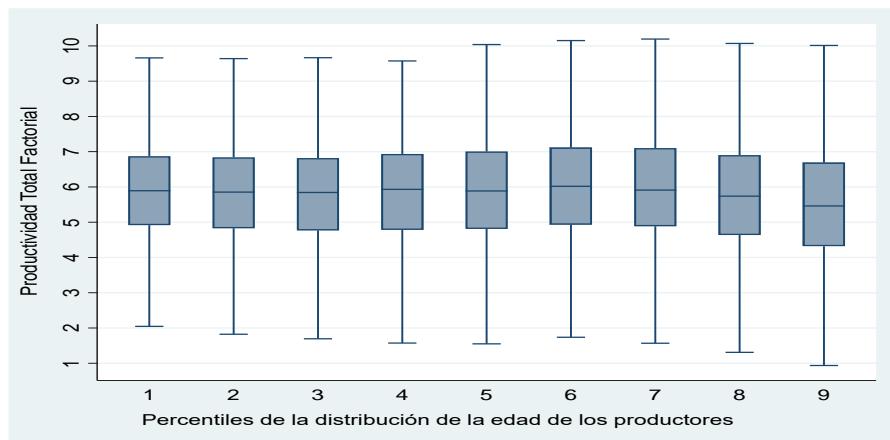


PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).
Elaboración Propia.

En el Gráfico 6 se muestra la estimación de la Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por percentiles de la distribución de la edad de los productores agrícolas. Se observa que la PTF se mantiene relativamente constante a lo largo de los percentiles de la edad. Sin embargo, a partir del octavo percentil la PTF se reduce (percentil que corresponde a productores mayores de 69 años).

Gráfico 6

Productividad Total Factorial por percentiles de la distribución de la edad de los productores agrícolas



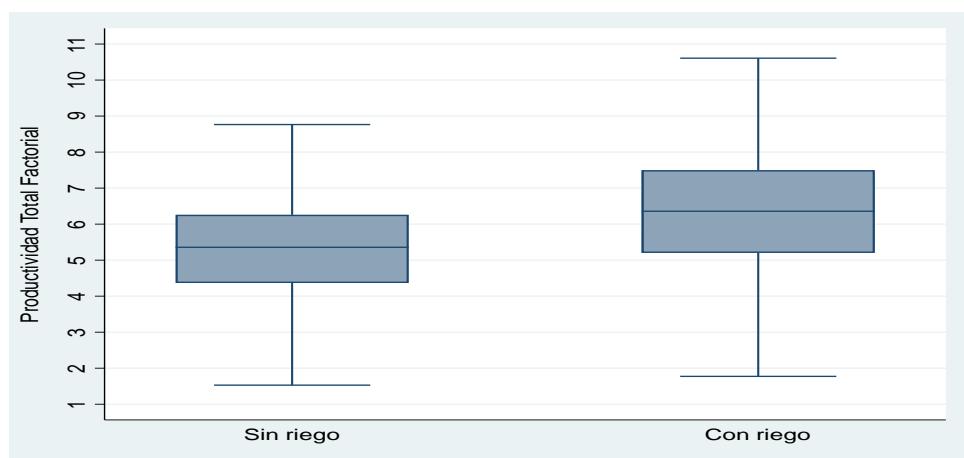
PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).

Elaboración Propia.

En el Gráfico 7 se muestra la estimación de la Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por el acceso al riego. Se observa que la PTF es mayor para aquellas unidades productivas que cuentan con riego; en particular, las unidades productivas con acceso a riego muestran un nivel de productivas mayor en 20% que aquellas unidades sin acceso al riego, siendo dicha diferencia significativa estadísticamente al 1%.

Gráfico 7

Productividad Total Factorial por acceso al riego



PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).
Elaboración propia.

3.4 Estimación del Efecto del Riego en la Productividad Total Factorial

a) Estrategia de identificación. La estrategia de identificación para recuperar efectos causales del riego en la productividad se construye utilizando un diseño de Regresión Discontinua (RD, por sus siglas en inglés). La productividad estimada por el método de ACF (controlando por calidad de la tierra) es comparada entre dos grupos: Por un lado, unidades agropecuarias que cuentan con riego y unidades agropecuarias sin riego. En el diseño del RD se asume que la asignación del riego como tratamiento está determinada, al menos parcialmente, por una variable de asignación a partir de un determinado umbral.

En el caso de la presente investigación se asume que dicha variable de asignación es la altitud. Específicamente, se ha empleado la altitud en metros sobre el nivel del mar partiendo del conglomerado en el cual la unidad agropecuaria se encuentra georeferenciada en la ENA. Se han considerado únicamente unidades agrícolas localizadas en la costa y sierra peruana, ubicados en las regiones de Ancash, Arequipa, Ica, La Libertad, Lambayeque, Lima, Moquegua, Piura y Tacna.

Por las características topográficas de la geografía peruana, el umbral de altitud (cutoff) fue determinado en 2000 metros sobre el nivel del mar para el caso de las unidades agrícolas ubicadas en la costa/sierra norte, y 3000 metros sobre el nivel del mar para el caso de unidades agrícolas ubicadas en la costa/sierra centro y sur del Perú. El umbral de altitud determina que existan unidades agrícolas más expuestas al riego localizadas por debajo del umbral y unidades agrícolas menos expuestas al riego localizadas por encima del umbral de altitud.

En la Tabla 7, se muestran estadísticas básicas de la productividad y la altitud con base en la muestra empleada de la ENA entre los años 2016 y 2018, diferenciando entre los grupos de comparación, unidades agrícolas con y sin riego. En primer término, se observa que la productividad promedio es mayor para unidades agrícolas con riego respecto de aquellas sin riego. En segundo término, la altitud promedio de unidades sin riego es de 3027 metros sobre

el nivel del mar; mientras que, la altitud promedio de unidades con riego es de 1775 metros sobre el nivel del mar.

Tabla 7

Estadísticas de variables en el Diseño de Regresión Contínua

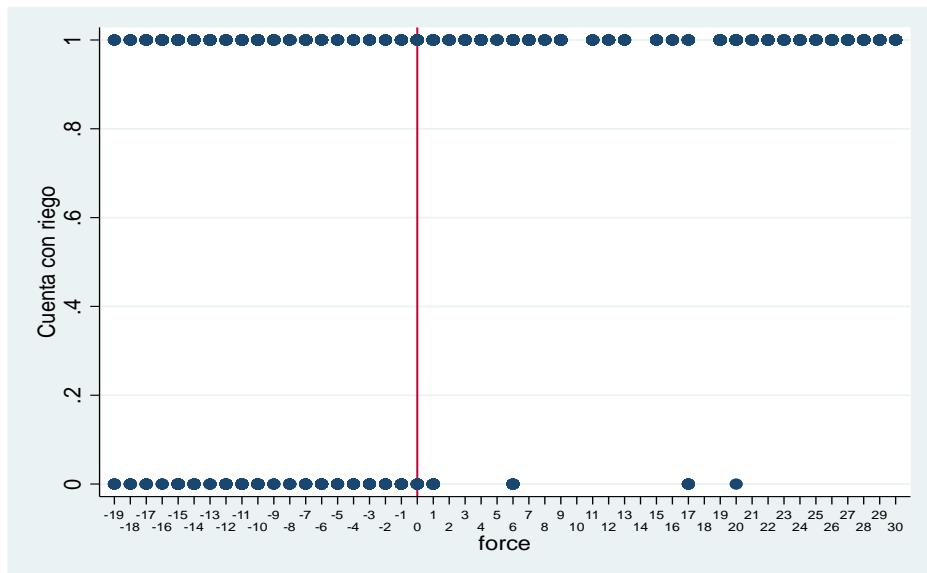
Variables	Desviación			Mínimo	Máximo
	Media	estándar	Máximo		
Muestra completa					
Productividad	6.24	1.77	-0.22	10.6	
Altitud	1941	1601	5	4747	
Sin riego					
Productividad	4.8	1.3	0.4	8.2	
Altitud	3027	813	15	4642	
Con riego					
Productividad	6.5	1.7	-0.2	10.6	
Altitud	1775	1627	5	4747	

Fuente: Muestra ENA 2016-2017-2018

En el Gráfico 8, se muestra la probabilidad de acceso al riego en función de la altura (*forcing variable*) y el umbral de altitud. La variable de asignación de tratamiento ha sido reformulada en función del umbral de altitud (cero es el umbral de altitud), donde valores negativos respecto de cero corresponden a niveles de altitud más altos, cada valor discreto representa incremento de altitud de 100 metros; por su parte, valores positivos respecto de cero corresponde a niveles de altitud más bajos, cada valor discreto representa decrementos de altitud de 100 metros.

Gráfico 8

Probabilidad de acceso al riego, altura y umbral de altitud



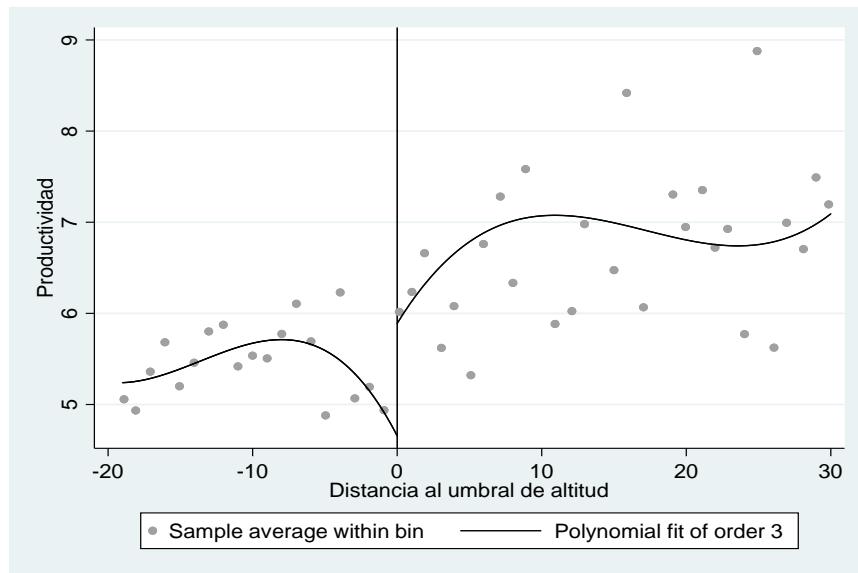
Elaboración Propia.

En el Gráfico 8, se observa también que la probabilidad de acceso al riego no cambia perfectamente a partir del umbral de altitud, lo que fue definido como cumplimiento imperfecto del tratamiento. Ello implica que el diseño de evaluación RD sea de tipo *Fuzzy* (Lee y Lemieux, 2010). De un total de 5511 observaciones, los resultados arrojan que 68% son cumplidores, 30% son siempre tomadores y 2% son nunca tomadores; por tanto, se estimará el efecto LATE para el caso solamente de cumplidores.

Para que el diseño RD (*Fuzzy*) permita recuperar efectos causales, debe ser posible observar alguna discontinuidad de la productividad alrededor del umbral de altitud. En el Gráfico 9 se observa el cambio discreto en la variable de asignación respecto del umbral de altitud (cero), valores hacia la izquierda representan una mayor altitud y valores hacia la derecha representan una menor altitud. Se ha empleado un polinomio de tercer orden para aproximar la media condicional para unidades agropecuarias más y menos expuestas al tratamiento.

Se observa una discontinuidad en torno al umbral de altitud; la localización de las unidades respecto del umbral generaría una discontinuidad en la productividad. Comparados con el grupo de control, las unidades agrícolas ubicadas a altitudes más bajas entorno del umbral muestran un nivel de productividad mayor. Bajo los supuestos descritos en la parte metodológica del documento, dicha discontinuidad puede ser interpretada como el efecto causal (local) del riego en la productividad.

Gráfico 9
Discontinuidad de la productividad entorno al umbral de altitud



Elaboración Propia.

b) Estimación de Efectos Causales por Regresión Discontinua. En la Tabla 8 se explora el efecto del acceso al riego en la productividad de unidades agrícolas, empleando como variable de asignación la altitud a la cual las unidades están localizadas. Los resultados presentados en la tabla se basan en una estimación lineal local (local linear regression, Triangular Kernel), bandas óptimas (optimal bandwith) y errores estándar corregidos por heterocedasticidad empleado cluster a nivel de conglomerado. Se han empleado dos estimaciones alternativas: (i) Por un lado, el modelo RD con inferencia convencional, propuesta en la literatura por Austin (2007), considerando la relación entre la productividad, el acceso al

riego y la altitud (columna 1) e incluyendo en dicha relación covariables (columna 2). (ii) Por otro lado, el modelo RD con inferencia robusta, propuesta en la literatura por Calonico et al., (2017), se ha considerado también dos estimaciones una sin covariables (columna 3) y otra con covariables (columna 4).

Dado el diseño Fuzzy RD en la Tabla 8 se reportan los resultados de la estimación de la ecuación 19, en dos etapas (estimador de Wald), la primera etapa representa la estimación del efecto de la variable de asignación en la probabilidad de acceso al riego ($\lim_{\varepsilon \downarrow 0} E[D|X = c + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[D|X = c + \varepsilon]$); mientras que τ_{FRD} es el efecto causal en la productividad.

El coeficiente estimado para la primera etapa se observa significativa estadísticamente para todas las especificaciones empleadas, lo cual sugiere que entorno al umbral de altitud la probabilidad de acceso al riego se incrementa entre 17 y 38 puntos porcentuales. El coeficiente estimado para la segunda etapa se observa positivo y significativo estadísticamente para todas las especificaciones empleadas, lo cual sugiere que unidades agrícolas con acceso al riego tienen un nivel de productividad mayor que unidades sin acceso al riego. En particular, considerando la estimación de la columna (4) de la Tabla 8, el acceso al riego incrementaría en S/ 26 la productividad de las unidades agrícolas del Perú; es decir, el efecto promedio del acceso al riego en la productividad sería de 17%.

En el Gráfico 10, se presentan los resultados anteriores en dos términos, en el Panel (a), la discontinuidad a lo largo de toda la variable de asignación; mientras que, en el Panel (b) se acotan los resultados en función de la estimación RD con inferencia robusta y covariables en torno al umbral óptimo.

Tabla 8*Estimaciones RD del efecto del riego en la productividad agrícola*

	RD		RD			
	Inferencia		Inferencia robusta			
	convencional (a)	(b)	(1)	(2)	(3)	(4)
Primera etapa:			0.253***	0.208***	0.172***	0.378***
$\lim_{\varepsilon \downarrow 0} E[D X = c + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[D X = c + \varepsilon]$			(0.060)	(0.059)	(0.056)	(0.078)
Segunda etapa:			1.39***	1.238***	1.187***	1.242***
$\lim_{\varepsilon \downarrow 0} E[Y X = c + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[Y X = c + \varepsilon]$			(0.185)	(0.181)	(0.152)	(0.182)
Efecto Local Causal Promedio: τ_{FRD}			5.48***	5.95***	6.90*	3.286***
			(1.467)	(1.869)	(3.553)	(1.126)
Optimal bandwith (c)			5.18	5.18	6.31	3.996
Número de observaciones			5511	5511	5511	5511
Media de la variable dependiente entorno a la izquierda del umbral (grupo de control)			5.22	5.22	5.31	5.06
Control por covariables			No	Sí	No	Sí

(a) Austin (2007): rd command (STATA)

(b) Calonico et al., (2017): rdrobust command (STATA)

(c) Optimal bandwith by Mean Squared Error for the RD treatment

Local Linear Regressions, Triangular Kernel

Errores corregidos por cluster a nivel de conglomerado

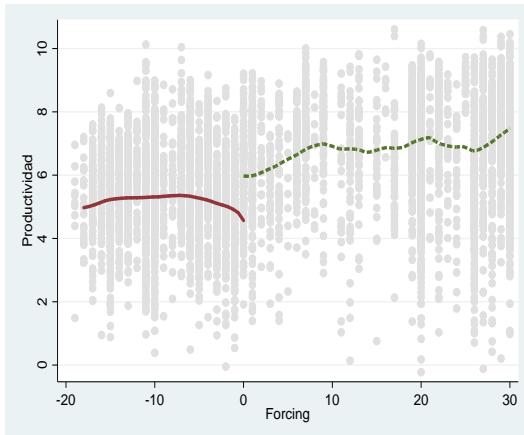
*** 1% de significancia; ** 5% de significancia y * 10% de significancia

Fuente: Muestra ENA 2016-2017-2018

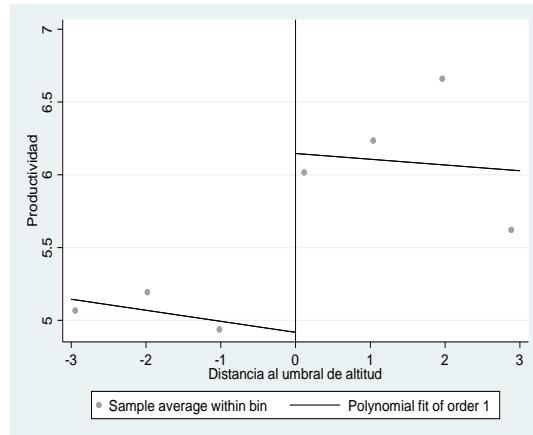
Gráfico 10

Efecto del riego en la productividad

Panel A: Toda la forcing variable



Panel B: Entorno al umbral



Fuente: Muestra ENA 2016-2017-2018

c) Pruebas placebo. Para verificar que los resultados presentados en la tabla anterior representan el efecto causal del acceso al riego en la productividad en lugar de diferencias sistemáticas entre los grupos de tratamiento y de control, en este apartado se presentan pruebas placebo. Como punto de partida se emplea el diseño RD (Fuzzy) para evaluar el efecto del acceso al riego en algunas covariables considerándolas como variables dependientes; se analiza específicamente el efecto del acceso al riego en la edad y el tamaño familiar de las unidades agrícolas, los resultados se presentan en las columnas (1) y (2) de la Tabla 9, los coeficientes estimados asociados a las características predeterminadas no resultaron significativas estadísticamente.

En segundo lugar se realizaron pruebas placebo sobre otros indicadores agrícolas; en particular, se analiza el efecto del acceso al riego en la superficie cosechada y el número de cultivos por parcela, los resultados se presentan en las columnas (3) y (4) de la Tabla 9, los

coeficientes estimados asociados a otros indicadores agrícolas no resultaron significativos estadísticamente.

Tabla 9

Estimaciones RD del efecto del riego en variables pre-determinadas y otros resultados agrícolas

	Edad	Tamaño familiar	Superficie cosechada	Cultivos
	(1)	(2)	(3)	(4)
Efecto Local Causal Promedio: τ_{FRD}	2.55 (5.64)	-0.536 (0.828)	0.07 (0.30)	-1.35 (1.15)
Optimal bandwidth (c)	2.56	2.76	3.29	2.42
Número de observaciones	5511	5511	5511	5511
Control por covariables	No	No	No	No

Optimal bandwidth by Mean Squared Error for the RD treatment

Local Linear Regressions, Triangular Kernel

Errores corregidos por cluster a nivel de conglomerado

*** 1% de significancia; ** 5% de significancia y * 10% de significancia

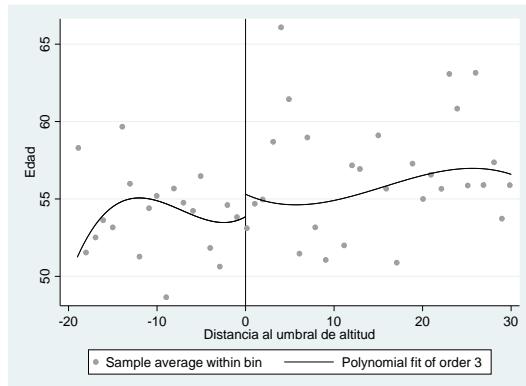
Fuente: Muestra ENA 2016-2017-2018

En el Gráfico 11, se reportan los resultados presentados en la tabla anterior

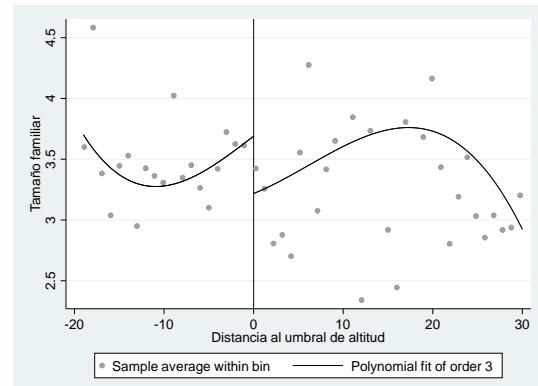
Gráfico 11

Efecto del riego en variables pre-determinadas y otros resultados agrícolas

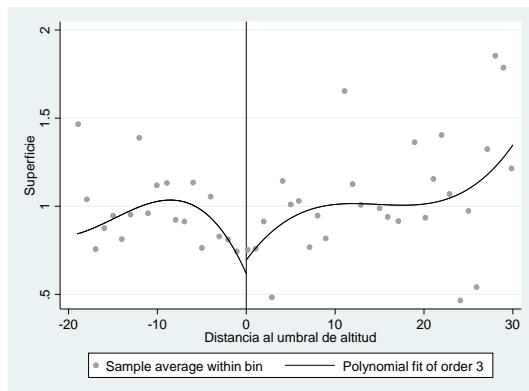
Panel A: Edad



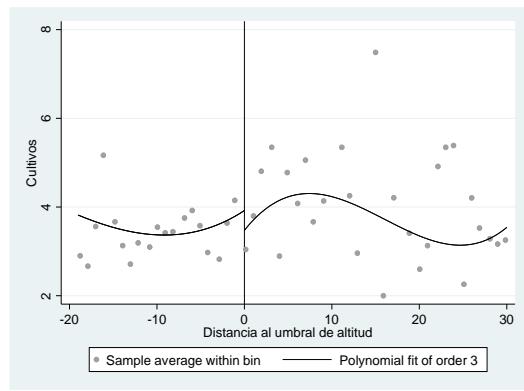
Panel B: Tamaño familiar



Panel C: Superficie cosechada



Panel D: Número de cultivos



Fuente: Muestra ENA 2016-2017-2018

3.5 Análisis de Género

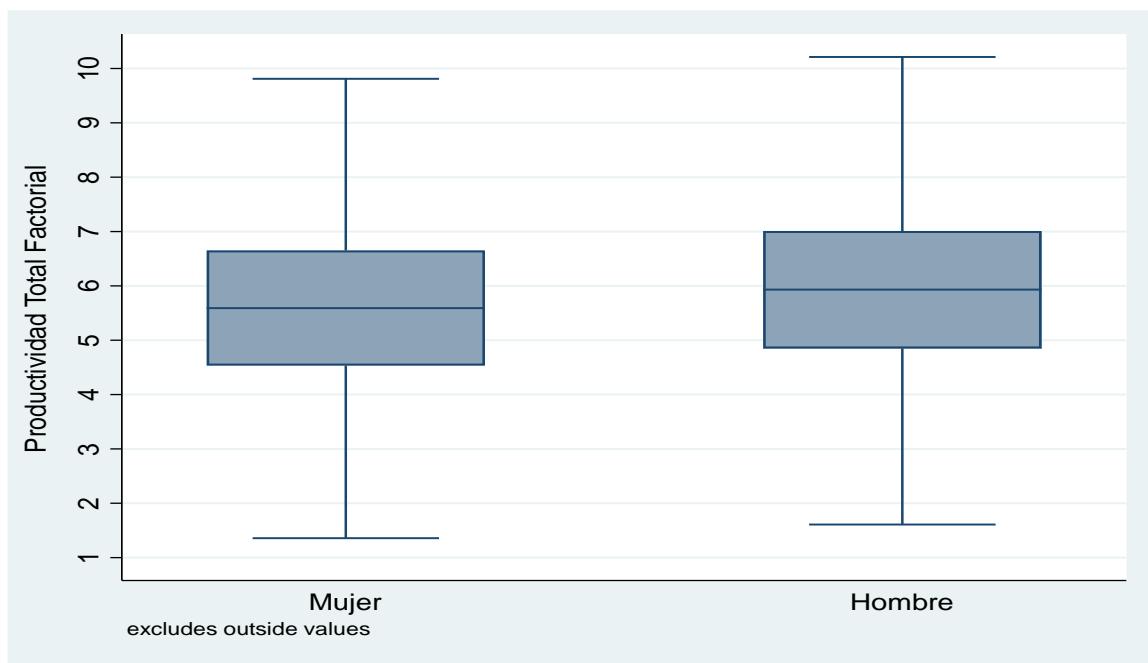
En el Gráfico 12 se muestra la estimación de la Productividad Total Factorial (PTF) de las unidades agrícolas desagregada por sexo de los productores agrícolas. Se observa que la PTF es mayor para aquellas unidades productivas conducidas por hombres; en particular, unidades productivas conducidas por hombres muestran un nivel de productivas mayor en 6% que unidades conducidas por mujeres, siendo dicha diferencia significativa estadísticamente al 1%.

En la Tabla 7 se presentan algunas estadísticas básicas correspondientes a productores agrícolas desagregados por sexo: hombres y mujeres. En primer término, se observan diferencias sistemáticas entre hombres y mujeres en todas las variables comparadas, en términos de variables agrícolas, se observan diferencias a favor de productores hombres en términos de productividad, valor de la producción, valor agregado, tierra, trabajo empleado en la actividad agrícola, gato en insumos y nivel de inversión. No obstante, se observan diferencias a favor de las mujeres en términos de calidad de la tierra y acceso al riego. No se observan diferencias entre hombres y mujeres en términos del número de parcelas explotadas y el número de cultivos. En segundo término, se observan diferencias en términos de algunas características sociales entre hombres y mujeres productores agrícolas. Los productores hombres son dos años

mayores que sus pares mujeres, y acumularon en promedio 1.4 más años de educación. El tamaño familiar es en promedio mayor para el caso de unidades productivas conducidas por hombres respecto de las mujeres; mientras que, las productoras mujeres son predominantemente indígenas.

Gráfico 12

Productividad Total Factorial por sexo de los productores



PTF estimada en base al método ACF (Ackerberg, Caves y Frazer, 2015).

Elaboración Propia.

Tabla 7*Estadísticas descriptivas por sexo de los productores y pruebas de diferencias en medias*

Variables	(1) Mujeres	(2) Hombres	Diferencia (1) - (2) [t-student]
<i>Variables agrícolas</i>			
Productividad Total Factorial	5.53	5.88	-0.34*** [-11.27]
Valor de la producción	5149	11545	-6396*** [-11.30]
Valor agregado	4300	9523	-5223*** [-11.16]
Tierra	5.65	8.38	-2.73*** [-4.01]
Indicador compuesto de calidad de la tierra	0.098	-0.19	0.29*** [15.93]
Trabajo remunerado	7.19	12.34	-5.14*** [-7.58]
Trabajo familiar no remunerado	1.58	1.99	-0.41*** [-14.79]
Insumos	919	2088	-1168*** [-8.67]
Inversión	140	528.00	-387*** [-7.51]
Cuenta con riego (1=Sí, 0=No)	0.576	0.53	0.044*** [5.12]
Numero de parcelas	3.37	3.46	-0.088 [-1.51]
Numero de cultivos	3.64	3.76	-0.106 [-2.35]
<i>Características sociales y demográficas</i>			
Edad del productor	55	53	2*** [9.04]
Años de educación	5	6.40	-1.4*** [-19.52]
Tamaño familiar	3.1	3.70	-0.6*** [-17.34]
Lengua materna indígena (1=Sí, 0=No)	0.46	0.38	0.076*** [8.94]
Obs.	4506	12896	

Los símbolos ** y *** representan significancia estadística al 5% y 1%, respectivamente.

Fuente: Encuesta Nacional Agropecuaria 2016-2017-2018 (INEI)

Elaboración Propia

4. Conclusiones

Los resultados de la estimación de la Productividad Total Factorial (PTF) por diversos métodos complementarios, sugieren que los sesgos por endogeneidad (variables omitidas y errores de medición) y dependencia condicional (colinealidad entre los factores de producción) en la estimación de función de producción serían abordados de mejor manera por el método propuesto por Ackerberg, Caves y Fraser (ACF). Este método ofrece estimaciones relativamente más precisas de la contribución de los factores productivos en la función de producción agrícola.

El coeficiente relacionado con el factor tierra resultó sobreestimado en las regresiones que no controlan por calidad de la tierra, lo cual sugeriría que no incorporar la calidad de la tierra podría generar sesgos en las estimaciones de la función de producción agrícola. Se encuentra que empleando el método ACF para la estimación de la función de producción e incorporando indicadores que controlan por calidad de la tierra, la participación del trabajo remunerado en la producción agrícola es de 10% y la participación de la tierra es de 90%. Adicionalmente, se encuentra evidencia que no es posible rechazar la hipótesis nula de presencia de rendimientos constantes a escala en la función de producción agrícola para el caso peruano.

El valor promedio estimado de la productividad resultó aproximadamente S/. 330, lo cual sugeriría que la productividad promedio representaría el 22% del valor agregado de la producción de unidades agropecuarias.

Respecto del análisis de heterogeneidades de la PTF, se encuentra que:

- En unidades agropecuarias con mayor cantidad de tierra utilizada la productividad es menor, lo cual sugiere la potencial presencia de mala asignación del factor tierra en el proceso productivo.

- Se observa una marcada diferencia en la productividad de las unidades agrícolas de la costa y el resto de las áreas geográficas.
- La PTF se incrementa con los años de educación de los productores.
- La PTF es mayor para aquellas unidades productivas que cuentan con riego.

Respecto del análisis de género, se encuentra que:

- Se observan diferencias sistemáticas entre hombres y mujeres en todas las variables comparadas. En términos de variables agrícolas observamos diferencias a favor de productores hombres en términos de productividad, valor de la producción, valor agregado, tierra, trabajo empleado en la actividad agrícola, gasto en insumos y nivel de inversión
- Las unidades productivas conducidas por hombres muestran un nivel de productivas mayor en 6% que unidades conducidas por mujeres.
- Los productores hombres son dos años mayores que sus pares mujeres, y acumularon en promedio 1.4 más años de educación. El tamaño familiar es en promedio mayor para el caso de unidades productivas conducidas por hombres respecto de las mujeres; mientras que, las productoras mujeres son predominantemente indígenas.

La estrategia de identificación para recuperar efectos causales del riego en la productividad se soporta en un diseño de Regresión Discontinua (RD, por sus siglas en inglés). La productividad estimada por el método de ACF y controlando por calidad de la tierra a nivel de unidades agropecuarias es comparada entre dos grupos: un grupo de unidades agropecuarias que cuentan con riego y otro grupo de unidades agropecuarias sin riego.

Se asume que el acceso al riego está determinado, al menos parcialmente, por una variable de asignación a partir de un determinado umbral. En el presente estudio se utiliza como variable de asignación la altitud en metros sobre el nivel del mar a partir del conglomerado en el cual la unidad agropecuaria se encuentra georeferenciada en la ENA. Se han considerado

únicamente unidades agrícolas localizadas en la costa y sierra peruana ubicados en las regiones de Ancash, Arequipa, Ica, La Libertad, Lambayeque, Lima, Moquegua, Piura y Tacna.

Por las características topográficas de la geografía peruana, el umbral de altitud fue determinado en 2000 metros sobre el nivel del mar para el caso de unidades agrícolas ubicadas en la costa-sierra norte y 3000 metros sobre el nivel del mar para el caso de unidades agrícolas ubicadas en la costa-sierra centro-sur del Perú. El umbral de altitud determina que existan unidades agrícolas más expuestas al riego localizadas por debajo del umbral y unidades agrícolas menos expuestas al riego localizadas por encima del umbral de altitud.

La evidencia encontrada sugiere que unidades agrícolas con acceso al riego tienen un nivel de productividad mayor que unidades sin acceso al riego, en particular el acceso al riego incrementaría en 17% la productividad de las unidades agrícolas del Perú.

5. Recomendaciones de Políticas Públicas

Los resultados obtenidos permiten formular algunos lineamientos para políticas públicas:

- La evidencia encontrada sobre el impacto del riego en la productividad agrícola sugiere la necesidad de que se incrementen y potencien los esfuerzos generados hasta ahora para la promoción del acceso al riego. Siendo el agua un bien económico que suele tener múltiples usos, se debería promover proyectos de inversión pública para la ampliación, mejoramiento, regulación y distribución del recurso hídrico con objetivos de uso multipropósito, entre ellas el agua para riego.
- Resulta necesario promover el Programa Presupuestal (PP 0042) denominado: “Aprovechamiento de los Recursos Hídricos para uso agrario”. Este PP tiene como objetivo mejorar la eficiencia del aprovechamiento de los recursos hídricos y a finales de 2019 contaba con un presupuesto de S/. 987 millones.

- De manera específica, se debería incrementar la instalación y operación de infraestructura de riego en la sierra y selva del Perú. El PP 0042, no cuenta con objetivos relacionados con la promoción del acceso al riego de manera concreta, la intervención se reduce a actividades de extensión, capacitación y asistencia técnica.
- Se sugiere por un lado que el MEF evalúe la efectividad del PP 0042 en su población y ámbito objetivo; por otro lado, se sugiere que el PP 0042 sea más ambicioso en la promoción de la infraestructura de riego. Para ello resulta necesario una mayor articulación entre los niveles de gobierno relacionados, particularmente entre el MINAGRI y los gobiernos regionales, para la formulación de proyectos de inversión en infraestructura de riego.
- De acuerdo con el portal de transparencia económica del MEF a finales de 2019, el gasto (devengado) en la división funcional 025 Riego de la función 10 Agropecuaria fue de S/ 1,967 millones, de los cuales S/ 867 millones corresponden a los gobiernos regionales, S/ 580 millones a los gobiernos locales y S/ 519 millones al Gobierno Nacional.
- Los principales proyectos identificados se concentran en la promoción del riego en la costa peruana, entre ellos: Proyecto Chavimochic, Majes Siguas II, Proyecto Alto Piura. El principal proyecto de gran magnitud para incrementar al riego en la selva peruana es el proyecto Irrigación Amoajo. Las inversiones buscan promover áreas de producción agroexportadora, básicamente en la sierra, sin considerar la producción agrícola para el consumo local, lo cual puede afectar la diversificación agrícola, el consumo de alimentos saludables y la soberanía alimentaria.

Bibliografía

- Abramovitz, M. (1956). Resource and Output Trends in the United States since 1870. *American Economic Review*. 46(2). pp.5-23.
- Ackerberg, D., Benkard, L. y Pakes, A. (2007). Econometric Tools for Analyzing Market Outcomes, in *Handbook of Econometrics*. North-Holland.
- Ackerberg, D., Caves, A. y Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*. 83 (6), pp. 2411-2451.
- Adamopoulos, T., y Restuccia, D. 2014a. The size distributions of farms and international productivity differences. *American Economics Review* 104(6), 1667-1697.
- Adamopoulos, T., y Restuccia, D. 2014b. Land reform and productivity: a quantitative analysis with micro data. University of Toronto, Working Paper.
- Adamopoulos, T., Brandt, L., Leight, J., y Restuccia, D. 2017. Misallocation, selection and productivity: a quantitative analysis with panel data from China. NBER Working Paper Series 23039.
- Aguirregabiria, V. 2018. Empirical Industrial Organization, University of Toronto.
- Ali, D. A, y Deininger, K. 2015. Is there a farm size-productivity relationship in African Agriculture? Evidence from Rwanda. *Land Economics*. 91(2), 317-343.
- Athey, S., y Imbens, G. 2018. Design-based analysis in difference-in-differences settings with staggered adoption. NBER Working Paper 24963.
- Austin, N. 2007. Causal inference with observational data. *The Stata Journal* 7, No. 4, pp. 507-541
- Banco Mundial. 2017. Tomando impulso en la agricultura peruana: oportunidades para aumentar la productividad y mejorar la competitividad del sector.
- Calonico, S., M. D. Cattaneo, and M. H. Farrel. 2017. Rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17 No. 2, pp. 372-404.

- Del Carpio, X., Loayza, N., y Datar, G. 2011. Is irrigation rehabilitation good for poor farmers? An Impact evaluation of a non – experimental irrigation project in Peru. *Journal of Agricultural Economics*, Vol. 62, No. 2, 449-473
- Fried, H. O., Knox-Lovell, C., y Schmidt, S. 2008. Efficiency and Productivity. In H. O. Fried, C. A. Knox Lovell, and S. S. Schmidt (Eds.). 2008. *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press.
- Galarza, F. B., y J. G. Díaz. 2014. Productividad y poder de mercado en mercados agrícolas. CIES, Lima.
- Gandhi, A., S. Navarro, and D. Rivers. 2014. “On the Identification of Production Functions: How Heterogeneous is Productivity”. Unpublished Report, University of Wisconsin-Madison.
- Gollin, D., y Udry, C. 2019. Heterogeneity, measurement error and misallocation: evidence from African Agriculture, NBER Working Paper 25440.
- Griliches, Z., and J. Mairesse. 1995. Production Functions: The Search for Identification. NBER Working Paper 5067.
- Herrera, G. P., R. Lourival, R. B. da Costa, D. R. Fonseca-Medes, T. B. Silva-Moreira, and U. Gomez Pinto de Abreu. 2018. Econometric analysis of income, productivity and diversification among smallholders in Brazil. *Land Use Policy*, <https://doi.org/10.1016/j.landusepol.2018.02.025>
- Jorgenson, D., and Z. Griliches. 1967. “The explanation of Productivity Change”. *Review of Economics Studies* 34(3), pp. 249-283.
- Lee, D. S., and T. Lemieux. 2010. Regression Discontinuity Designs in Economics. *Journal of Economics Literature* 48: 281-355.
- Levinsohn, J., and A. Petrin. 2003. “Production Functions using Inputs to Control for Unobservables”. *Review of Economic Studies* 70(2), pp. 317-341.

- Ludena, C. E. 2010. Agricultural productivity growth, efficiency change and technical progress in Latin America and the Caribbean. Working Papers No. 18. IDB, Washington. DC.
- Newman, C., F. Tarp, and K. van den Broeck. 2015. Property rights and productivity: the case of joint land titling in Vietnam. *Land Economics* 91(1), 91-105.
- Ministerio de Agricultura y Riego. 2016. *Política Nacional Agraria*, Lima.
- Ministerio de Agricultura y Riego. 2015. Lineamientos de Política y Estrategia Nacional de Riego, Resolución Ministerial N° 0507-2015-MINAGRI.
- Ministerio de Economía y Finanzas – MEF. 2019. Evaluación de impacto del Fondo Mi Riego/Sierra Azul, Lima.
- Mugizi, F., y Matsumoto, T. 2020. Population pressure and soil quality in Sub-Saharan Africa: Panel evidence from Kenya. *Land Use Policy*, <https://doi.org/10.1016/j.landusepol.2020.104499>
- Olley, S., y Pakes, A. 1996. The dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica* 64(6), pp. 1263-1297.
- Quisumbing, A. 1996. Male-Female differences in agricultural productivity: Methodological Issues and Empirical Evidence. *World Development*, Vol. 24, No. 10, 1579-1595.
- Restuccia, D., y Santaeulalia-Llopis, R. 2017. Land misallocation and productivity. NBER Working Paper Series 23128.
- Slavchevska, V. 2015. Gender differences in agricultural productivity: the case of Tanzania. *Agricultural Economics* 46 (2015), 335-355
- Smith, L. 2004. Assessment of the contribution of irrigation to poverty reduction and sustainable livelihoods. *International Journal of Water Resources Development*, Vol. 20, No. 2, 243-257

Trindade, F. J., y Fulgintini, L. 2015. Is there a slowdown in agricultural productivity growth in South America? *Agricultural Economics*, Vol., 46, No. 1, pp.: 69-81.

Van Beveren, I. 2012. Total Factor Productivity Estimation: A Practical Review. *Journal of Economic Surveys*, Vol. 26, No.1, pp. 98-128.

Zegarra, E. 2014. *Economía del Agua: Conceptos y Aplicaciones para una mejor gestión*.

GRADE.

Anexos

Tabla A.1. Variables socioeconómicas y variables relacionadas con la calidad de la tierra

Sexo del productor (1=Hombre, 0=Mujer)
Edad del productor
Tamaño familiar
Años de educación
Lengua materna indígena (1=Sí, 0=No)
Numero de parcelas
Numero de cultivos
Altitud promedio a nivel de conglomerado (m.s.n.m)
Temperatura promedio a nivel regional (C°)
Precipitación promedio a nivel regional (milímetros)
Humedad promedio a nivel regional (porcentaje)
Realiza análisis de suelos (1=Sí; 0=No)
Realiza mezcla del suelo con materia orgánica (1=Sí; 0=No)
Realiza rotación de cultivos (1=Sí; 0=No)
Cuenta con terrazas o zanjas de infiltración (1=Sí; 0=No)
Realiza arar o voltear la tierra (1=Sí; 0=No)
Realizar desmenuzado de la tierra (1=Sí; 0=No)
Realiza nivelación de terrero (1=Sí; 0=No)
Realiza surcos en pendiente (1=Sí; 0=No)
Realiza calidad de agua (1=Sí; 0=No)
Realiza medición del tiempo de riego (1=Sí; 0=No)
Realiza medición de la calidad de agua al ingreso a la parcela (1=Sí; 0=No)
Realiza mantenimiento de sistema de riego (1=Sí; 0=No)
Realiza análisis de agua (1=Sí; 0=No)
Usa abonos (1=Sí; 0=No)
Usa fertilizantes (1=Sí; 0=No)
Usa plaguicidas (1=Sí; 0=No)
Aplica control biológico (1=Sí; 0=No)
Aplica manejo integrado de plagas (1=Sí; 0=No)
Usa residuos de cultivos como abonos (1=Sí; 0=No)
Usa residuos de animales como abonos (1=Sí; 0=No)
Existen minas cerca de los predios rurales (1=Sí; 0=No)
Cuenta con producción certificada (1=Sí; 0=No)

Fuente: Encuesta Nacional Agropecuaria 2016-2017-2018 (INEI), INEI.

Elaboración Propia

Tabla A.2

Estadísticas descriptivas: Variables socioeconómicas y variables relacionadas con la calidad de la tierra

Variables	Obs	Media	Desviació n estándar	Min.	Max
Sexo del productor (1=Hombre, 0=Mujer)	17409	0.74	...	0	1
Edad del productor	17409	53.18	14.7	0	1
Tamaño familiar	17409	3.58	2.0	16	98
Años de educación	17409	6.11	4.2	0	1
Lengua materna indígena (1=Sí, 0=No)	17409	0.40	...	0	16
Numero de parcelas	17409	3.44	3.4	1	10
Numero de cultivos	17409	3.73	2.6	1	99
Altitud promedio a nivel de conglomerado (m.s.n.m)	17409	2214.9 3	1569.2	3	4749
Temperatura promedio a nivel regional (C°)	17409	17.95	5.7	5	28
Precipitación promedio a nivel regional (milímetros)	17409	733.42	681.2	0	2995
Humedad promedio a nivel regional (porcentaje)	17409	74.67	11.0	47	93
Realiza análisis de suelos (1=Sí; 0=No)	17409	0.02	...	0	1
Realiza mezcla del suelo con materia orgánica (1=Sí; 0=No)	17409	0.57	...	0	1
Realiza rotación de cultivos (1=Sí; 0=No)	17409	0.52	...	0	1
Cuenta con terrazas o zanjas de infiltración (1=Sí; 0=No)	17409	0.09	...	0	1
Realiza arrar o voltear la tierra (1=Sí; 0=No)	17409	0.71	...	0	1
Realizar desmenuzado de la tierra (1=Sí; 0=No)	17409	0.66	...	0	1
Realiza nivelación de terrero (1=Sí; 0=No)	17409	0.40	...	0	1
Realiza surcos en pendiente (1=Sí; 0=No)	17409	0.29	...	0	1
Realiza calidad de agua (1=Sí; 0=No)	17409	0.15	...	0	1
Realiza medición del tiempo de riego (1=Sí; 0=No)	17409	0.23	...	0	1
Realiza medición de la calidad de agua al ingreso a la parcela (1=Sí; 0=No)	17409	0.07	...	0	1
Realiza mantenimiento de sistema de riego (1=Sí; 0=No)	17409	0.48	...	0	1
Realiza análisis de agua (1=Sí; 0=No)	17409	0.01	...	0	1
Usa abonos (1=Sí; 0=No)	17409	0.58	...	0	1
Usa fertilizantes (1=Sí; 0=No)	17409	0.50	...	0	1
Usa plaguicidas (1=Sí; 0=No)	17409	0.54	...	0	1
Aplica control biológico (1=Sí; 0=No)	17409	0.01	...	0	1

Aplica manejo integrado de plagas (1=Sí; 0=No)	17409	0.08	...	0	1
Usa residuos de cultivos como abonos (1=Sí; 0=No)	17409	0.25	...	0	1
Usa residuos de animales como abonos (1=Sí; 0=No)	17409	0.44	...	0	1
Existen minas cerca de los predios rurales (1=Sí; 0=No)	17409	0.06	...	0	1
Cuenta con producción certificada (1=Sí; 0=No)	17409	0.01	...	0	1

Fuente: Encuesta Nacional Agropecuaria: muestra panel 2016-2017-2018 (INEI)

Elaboración Propia

Tabla A.3

Matriz de coeficientes del Primer Factor Principal de variables potencialmente relacionadas con la calidad de la tierra

Variables	Factor 1	Kaiser-Meyer-Olkin of sample adequacy
Sexo del productor (1=Hombre, 0=Mujer)	-0.1442	0.7232
Edad del productor	0.1476	0.5333
Tamaño familiar	-0.1386	0.605
Años de educación	-0.0681	0.5477
Lengua materna indígena (1=Sí, 0=No)	0.4722	0.799
Numero de parcelas	0.3907	0.8826
Numero de cultivos	0.056	0.5281
Realiza análisis de suelos (1=Sí; 0=No)	-0.0359	0.6691
Realiza mezcla del suelo con materia orgánica (1=Sí; 0=No)	0.68	0.9069
Realiza rotación de cultivos (1=Sí; 0=No)	0.6565	0.9268
Cuenta con terrazas o zanjas de infiltración (1=Sí; 0=No)	0.2132	0.7823
Realiza arar o voltear la tierra (1=Sí; 0=No)	0.7602	0.8584
Realizar desmenuzado de la tierra (1=Sí; 0=No)	0.7663	0.8716
Realiza nivelación de terrero (1=Sí; 0=No)	0.4559	0.8526
Realiza surcos en pendiente (1=Sí; 0=No)	0.3329	0.8611
Realiza calidad de agua (1=Sí; 0=No)	0.0824	0.7551
Realiza medición del tiempo de riego (1=Sí; 0=No)	0.1248	0.7571
Realiza medición de la calidad de agua al ingreso a la parcela (1=Sí; 0=No)	0.048	0.8415
Realiza mantenimiento de sistema de riego (1=Sí; 0=No)	0.2568	0.8161
Realiza análisis de agua (1=Sí; 0=No)	0.0273	0.7381
Usa abonos (1=Sí; 0=No)	0.6535	0.8483
Usa fertilizantes (1=Sí; 0=No)	0.1747	0.6578
Usa plaguicidas (1=Sí; 0=No)	0.2102	0.6846
Aplica control biológico (1=Sí; 0=No)	0.0154	0.6974

Aplica manejo integrado de plagas (1=Sí; 0=No)	0.0524	0.7222
Usa residuos de cultivos como abonos (1=Sí; 0=No)	0.2611	0.873
Usa residuos de animales como abonos (1=Sí; 0=No)	0.636	0.8906
Existen minas cerca de los predios rurales (1=Sí; 0=No)	0.1322	0.8645
Cuenta con producción certificada (1=Sí; 0=No)	-0.1235	0.6597
Altitud promedio a nivel de conglomerado (m.s.n.m)	0.751	0.8656
Temperatura promedio a nivel regional (C°)	-0.5889	0.8596
Precipitación promedio a nivel regional (milímetros)	-0.4161	0.8442
Humedad promedio a nivel regional (porcentaje)	-0.3694	0.7614
Overall		0.8211

Fuente: Encuesta Nacional Agropecuaria: muestra panel 2016-2017-2018 (INEI)

Elaboración Propia

XXI Concurso Anual de Investigación CIES 2019



CIES
consorcio de investigación
económica y social

Construyendo conocimiento para mejores políticas