



Universidad de San Andrés

Departamento de Economía

Tesis presentada para optar el grado de Magister en Economía

Título:

Impacto del acceso a agua y saneamiento en la oferta laboral. Evidencia para Perú a partir del Programa Nacional de Saneamiento Rural

Autor: Yohnny Campana

Pasaporte: 4623166

Director de Tesis: Álvaro Monge Zegarra, Phd.

Lima, Octubre 2018

Impacto del acceso a agua y saneamiento en la oferta laboral. Evidencia para Perú a partir del Programa Nacional de Saneamiento Rural

Yohnny Campana

Resumen

Las intervenciones que dotan de acceso a agua y saneamiento en el hogar tienen importante potencial para generar impactos oferta laboral debido a que liberan recursos que pueden ser utilizados por las personas para ofrecer mano de obra. Este estudio busca generar evidencia en esta dirección a partir de la experiencia del Programa Nacional de Saneamiento Rural (PNSR) en Perú utilizando un estimador *fuzzy* de regresiones discontinuas. El PNSR fue un programa creado en 2012 que seleccionó como localidades beneficiarias a aquellas que según información del Censo de Población y Vivienda de 2007 tuvieran entre 200 y 2000 habitantes. Los resultados del estudio muestran que este programa habría expandido el acceso a agua domiciliaria con conexión a red pública en alrededor de 35 puntos porcentuales, y el acceso a sistemas sanitarios en la vivienda en cerca de 17 puntos porcentuales. Además, habría impactado incrementando las horas de trabajo que realizan las personas en actividades laborales en un rango de 33 a 40 minutos por día. No se encuentra evidencia de impactos en términos de la participación en el mercado laboral o la probabilidad de estar ocupado. Así, tal parece que el PNSR tendría impactos principalmente en el margen intensivo de la oferta laboral, pero no en el margen extensivo.

Abstract

Interventions that provide access to water and sanitation in the household have potential impacts in the labor supply because they release resources that people can to offer work. This study pursue the interest to generate evidence in this direction based on the experience of the National Rural Sanitation Program (PNSR) in Peru by using a fuzzy regression discontinuity design. The PNSR was a program created in 2012 and assigned among rural localities with 200 to 2000 inhabitants according to Population and Housing Census of year 2007. The results show that this program expanded access to residential piped water by around 35 percentage points, and access to sanitation services inside the home by about 17 percentage points. In addition, there is some evidence that it increased the number of hours of productive work in a range of 33 to 40 minutes per day. There is no evidence of impacts in terms of participation in the labor market or the probability of having a job. Thus, it seems that the PNSR had impacts mainly on the intensive margin of the labor supply, but not on the extensive margin.

Keywords: labor supply; piped water and sanitation; rural development

JEL classification: J22; L95; O18

1 Introducción

La falta de acceso a fuentes seguras de agua para consumo y saneamiento es uno de los mayores problemas que enfrentan los países en desarrollo en la actualidad. Según un reporte reciente de la Organización Mundial para la Salud, en 2015 cerca de 2,100 millones de personas en el mundo carecía de acceso a agua potable, y alrededor de 4,500 millones no disponía de sistemas de saneamiento adecuados (OMS, 2017). La reducción de esta brecha es parte de la agenda global de los países y las organizaciones multilaterales, y por ello es uno de los indicadores de medición de las Metas de Desarrollo del Milenio (ver Naciones Unidas 2015).

Perú es un país que no está lejos de esta problemática y debido a ello el Estado ha hecho esfuerzos importantes por ampliar la cobertura entre los hogares. Actualmente, según información de la Encuesta Nacional de Hogares (ENAHOG) que elabora el Instituto Nacional de Estadística e Informática (INEI), a nivel nacional cerca del 89% de la población cuenta con acceso a agua de red en la vivienda (desde niveles iniciales de 70% en 2004), y 74% a formas seguras de saneamiento (desde niveles iniciales de 54% en el 2004). Si bien las brechas aún son importantes, el avance observado ha sido posible gracias a iniciativas desde diversas instancias del Estado y Organizaciones No Gubernamentales que han buscado llevar el agua para consumo doméstico y la construcción de sistemas de saneamiento en las viviendas de los hogares.

Desde el Estado, los proyectos de agua y saneamiento son diseñados siguiendo un modelo causal que en última instancia tiene el objetivo de mejorar indicadores de salud y nutrición infantil (ver, por ejemplo, MEF 2015; Moreano 2010; Smith y Hahaad 2000). Es decir, más allá del cierre de brechas de acceso, la efectividad de las intervenciones de este tipo se evalúa en términos de la reducción de indicadores de prevalencia de enfermedades diarreicas agudas, tasas de desnutrición, prevalencia de enfermedades relacionadas con la falta de acceso a agua, etc. La literatura internacional y nacional se ha encargado de evaluar extendidamente esta relación causal y los resultados sobre su validez son mayormente concluyentes. En el Perú, por ejemplo, la evaluación realizada de Calzada e Iranzo (2016) encuentra que la expansión de la cobertura de agua de red habría reducido la prevalencia de EDAs y la proporción de niños con bajo peso al nacer en tasas de alrededor de 10 puntos porcentuales en cada caso. Esto sería consecuencia del aumento de la disponibilidad de agua durante las 24 horas del día y la mayor cloración adecuada de la misma.

Los impactos potenciales de la provisión del agua, sin embargo, no se limitan al ámbito de la salud. Más bien, por la naturaleza de este tipo de intervenciones, que tienen la virtud de liberar recursos de los miembros del hogar y hacer más eficiente la ejecución de actividades domésticas, tienen impactos potenciales sobre otras dimensiones de bienestar. A modo de ejemplo, Kremer et al. (2011) estiman que un hogar rural en Kenia Occidental realiza alrededor de 7 viajes de acarreo de agua por día, con un promedio por viaje de 20 minutos a pie. En Perú la situación es menos extrema, pero igualmente importante. De acuerdo a la Encuesta de Uso del Tiempo realizada por el INEI, en el 2010 una persona mayor de 16 años que residía en una vivienda sin acceso a agua de red destinaba alrededor de 2.12 horas por semana al acarreo de agua. Esto sin considerar las actividades de higiene, aseo y limpieza fuera del hogar (como el lavado de ropa, que se puede realizar en la fuente de agua).

Considerando los recursos que se liberan con la provisión de agua y saneamiento, una dimensión de impacto es la vinculada al ámbito laboral. En efecto, existe diversa literatura que sugiere que el acceso a agua y saneamiento domiciliarios puede, a través de diversos mecanismos, incrementar la oferta laboral de los miembros del hogar. A pesar de ello, no son abundantes los estudios que evalúa el impacto de intervenciones de agua y saneamiento sobre variables de

dimensión social, como las de orden laboral, y los que existen son principalmente de naturaleza cualitativa (Mills y Cumming 2016). Excepciones en este campo son los trabajos de Zhou y Turkey (2017), Devoto et al. (2012), Meek (2012), entre otros, que sí evalúan el impacto del acceso a agua y saneamiento sobre variables laborales. Sin embargo, los estudios en este punto no son del todo concluyentes. Por ejemplo, Meek (2012) encuentra impactos positivos del acceso a agua en el tiempo de trabajo dedicado a actividades agrícolas en el norte de Kirguizistán. Zhou y Turkey (2017) obtiene resultados similares en China, pero esta vez sobre el trabajo en actividades no agrícolas, principalmente entre las mujeres. A diferencia, Devoto et al. (2012) no encuentra impactos significativos en Marruecos y más bien descubre que el tiempo liberado se dirigió hacia actividades de ocio, principalmente entre las mujeres.

Este estudio busca generar evidencia sobre el impacto del acceso a agua y saneamiento sobre variables de participación laboral tomando como estudio de caso el Perú, un país de ingresos medios que tiene una brecha de acceso aún importante, como se vio líneas arriba. Para esto, se utiliza la experiencia del Programa Nacional Saneamiento Rural (PNSR), que es un programa creado en 2012 para operativizar la estrategia que despliega el Estado central para llevar agua y saneamiento a los hogares rurales del Perú. Para lidiar con problemas de endogeneidad asociados a la asignación del PNSR entre los hogares, tomamos ventaja de una regla de asignación que rige su focalización. En efecto, de acuerdo a la normativa del programa, el PNSR se asigna sobre localidades de 200 habitantes o más según información del Censo de Población y Vivienda 2007 que elaboró el INEI. Esto crea un contexto adecuado para implementar un estimador de regresiones discontinuas teniendo como variable de asignación el tamaño de población de las localidades.

Para la ejecución del estudio, se ha construido una base de datos que combina información del Censo de Población y Vivienda elaborado por el INEI en 2007 y la ENAHO del año 2015, también elaborado por el INEI. Se utiliza esta base de datos porque recoge información detallada de la vida laboral de las personas, por un lado, y porque en el año 2015 ha recogido una muestra amplia de localidades rurales dispersas, que son las que tienen menos de 200 habitantes.

Los resultados principales del estudio muestran que el PNSR habría expandido el acceso a agua domiciliaria con conexión a red pública en alrededor de 35 puntos porcentuales, y el acceso a sistemas sanitarios en la vivienda (considerando sistemas con conexión a red pública y letrinas, que son las soluciones tecnológicas ofrecidas por el programa) en cerca de 17 puntos porcentuales. Más importante, el PNSR habría impactado incrementado las horas de trabajo que realizan los hogares en actividades laborales en alrededor de entre 33 y 40 minutos. No se encuentra evidencia de impactos en términos de la participación en el mercado laboral o la probabilidad de estar ocupado. Así, tal parece que el PNSR tendría impactos principalmente en el margen intensivo de la oferta laboral, pero no en el margen extensivo.

Una particularidad de las intervenciones de agua y saneamiento es que tienen potenciales impactos diferenciados de género. Los resultados del ejercicio validan levemente esta hipótesis en las horas trabajadas por día. Por un lado, cuando se evalúan los impactos entre mujeres y varones, los parámetros estimados son levemente mayores para las primeras, aun cuando no llegan a ser estadísticamente diferentes. Por otro lado, cuando se identifican impactos heterogéneos según roles al interior del hogar, los impactos ocurren principalmente entre las hijas mujeres y son nulos estadísticamente para los hijos varones. Entre los padres y madres los impactos son similares.

El resto del documento está estructurado de la siguiente manera. La sección 2 ofrece una breve revisión de la literatura del impacto del agua y saneamiento, enfatizando la discusión sobre los impactos potenciales sobre variables laborales. En esta misma sección se describe el

PNSR, resaltando los detalles más importantes asociados a su regla de focalización. En la sección 3 se presenta el diseño de regresiones discontinuas para evaluar el impacto del PNSR y se hace una descripción de la base de datos construida. Más adelante, la sección 4 presenta los resultados de las estimaciones y, finalmente, en la sección 5 se presentan las conclusiones.

2 Contexto del estudio

2.1 Impacto del acceso a agua y saneamiento

Los proyectos de agua y saneamiento han sido vistos como intervenciones que en última instancia persiguen el objetivo de modificar indicadores de salud, en particular infantil. En el modelo causal presentado en Smith y Haddad (2000), por ejemplo, que a su vez se basa en el modelo de UNICEF (1990), el acceso a agua segura en el hogar es una condición que facilita un ambiente saludable para los niños, el cual a su vez es un determinante del estado de salud y nutricional. La evidencia empírica sobre este tipo de intervenciones es abundante y muestra la efectividad de la dotación de agua y saneamiento para reducir las enfermedades diarreicas agudas (Devoto et al. 2012; Ziegelhofer 2012), la desnutrición infantil (Kremer et al. 2011, Beltran y Seinfeld 2009), el bajo peso al nacer (Calzada e Iranzo 2016), la mortalidad infantil (Galiani et al, 2005), entre otros.

Sin embargo, como bien explican Mills y Cumming (2016) en una revisión de la evidencia sobre los impactos de intervenciones de Agua, Saneamiento e Higiene (WASH, por sus siglas en inglés), el acceso a esta infraestructura en tiene efectos que trascienden el campo de la salud y se proyectan sobre otras dimensiones sociales. Una de estas dimensiones se relaciona con las posibilidades que ofrece para ampliar la participación en el mercado laboral. En general, existen argumentos para sostener que es más probable que las persona enfrenten restricciones en su oferta laboral cuando el recurso agua es limitado en el hogar (Fletcher, 2018).

Son dos los mecanismos por los que el acceso a agua puede impactar sobre la oferta laboral. El primero, y más directo, se relaciona a los recursos liberados del acarreo de agua hasta el hogar. En hogares que carecen de acceso a agua domiciliaria, los miembros deben encargarse del transporte de este recurso desde la fuente de origen hasta la vivienda, lo cual, por un lado, insume tiempo, y, por otro, reduce la productividad de las labores domésticas. Mientras más alejada esté la fuente del hogar, más recursos son utilizados en el transporte de agua (tiempo, energía, entre otros). Por tanto, la dotación de agua domiciliaria permite que los miembros dejen de destinar tiempo y energía a esta actividad y lo reasignen a otras actividades, una de las cuales puede ser la oferta laboral (Sijbesma et al. 2008; Meeks 2012).

El segundo mecanismo se vincula a la reducción de la prevalencia de enfermedades por causa del consumo de agua no segura y la disposición de excretas en ambientes que facilitan la contaminación y contagio de enfermedades. En general, el agua segura que ofrecen los sistemas de potabilización del recurso (cualquiera sea la forma de tratamiento) y los servicios higiénicos adecuados reducen potencialmente la exposición a factores de riesgo que ocasionan las enfermedades diarreicas agudas. De acuerdo con Waddington et al. (2009), por ejemplo, las intervenciones que mejoran la calidad del agua (a través de la cloración, filtración, desinfección solar, etc.) proveen medios para remover los agentes patógenos que contaminan el recurso hídrico durante su transporte y almacenamiento, permitiendo que los hogares tengan a disposición agua más segura para su consumo. Además, la disponibilidad del recurso en el hogar, en mayor abundancia y con mayor calidad, permite que los miembros tengan prácticas más adecuadas de higiene en momentos clave (por ejemplo, durante la manipulación de alimentos, durante el aseo, etc.), lo que fortalece las barreras de los agentes contaminantes. De similar

modo, un sistema adecuado de saneamiento limita la contaminación fecal-oral que se produce a través de las aguas subterráneas o superficiales, los suelos contaminados con excretas y los insectos que pudieran estar en contacto con estas.

Todo lo anterior permite reducir los episodios de infecciones diarreicas de los miembros del hogar, tanto en frecuencia como en severidad, con lo cual se reducen los tiempos de convalecencia. Esto, en última instancia, permite mejorar el estado de salud de las personas y con ello incrementar el stock de capital humano (Currie y Madrian 1999). Adicionalmente, entre las personas que tienen hijos menores, la menor prevalencia de enfermedades en estos significa menor tiempo destinado a su cuidado por parte de los padres (Ivens 2008). De esta manera, bajo este segundo canal, el acceso a agua potable y saneamiento domiciliarios permite aumentar el capital humano de las personas, lo que incrementa su productividad laboral (Currie y Madrian 1999).

Una particularidad de las intervenciones de agua y saneamiento es que tienen impactos potenciales diferenciados de género. En efecto, en sociedades donde los roles socialmente aceptados asignan a las mujeres las labores de preparación de alimentos, la higiene y limpieza del hogar, el cuidado de los animales, entre otros, mientras que el varón es el proveedor de recursos, entonces las restricciones de acceso a agua domiciliaria podrían tener mayores impactos potenciales sobre ellas (Sultana, 2018). Esto ocurre porque, de acuerdo a la discusión anterior, serían ellas quienes se benefician de la liberación de recursos que el agua domiciliaria provee, generando tiempo excedentario para la realización de otras actividades.

2.2 El Programa Nacional de Saneamiento Rural

El Programa Nacional de Saneamiento Rural (PNSR) es un programa adscrito al Ministerio de Vivienda y Construcción creado en 2012 con el propósito de atender a las poblaciones rurales del país facilitándoles el acceso a agua y saneamiento básico. Desde entonces ha invertido alrededor de 1000 millones de soles y atendido a alrededor de 300,000 hogares a través de casi 1500 proyectos desplegados en todo el país¹.

En su diseño, el PNSR está compuesto por tres componentes. El primero es el componente infraestructural, que busca desarrollar como producto los sistemas de agua y saneamiento con conexiones domiciliarias. Para ello, las opciones tecnológicas ofrecidas se adaptan a las condiciones físicas, económicas, ambientales, sociales y culturales de los centros poblados priorizados para su atención (PNSR, 2013). Esto involucra la formulación inicial de los proyectos de ingeniería en cada centro poblado y la licitación de su construcción a empresas locales. Para los sistemas de agua, las obras físicas implican la construcción de infraestructura de captación del agua desde las fuentes subterráneas o superficiales, la estructura de almacenamiento y tratamiento del agua, la red de distribución hasta las viviendas y las conexiones domiciliarias. Para los sistemas de saneamiento, las soluciones implementadas son principalmente letrinas (de hoyo seco ventilado o composteras de doble cámara ventilada). Sólo en pequeñas ciudades, donde la densidad de población lo hace posible, se implementa redes de alcantarillado.

El segundo componente es de prácticas saludables y busca sensibilizar a la población respecto a los factores de riesgo asociados a la falta de agua segura y saneamiento básico. Adicionalmente, brindar capacitaciones para fortalecer las prácticas de manipulación y almacenamiento del agua, higiene en momentos críticos y cuidados de la salud. Este componente persigue el propósito de dotar de habilidades en los hogares para mejorar sus hábitos de higiene que potencian el efecto negativo de la ausencia de agua y saneamiento en los hogares.

¹ <http://pnsr.vivienda.gob.pe/portal/>

El tercer componente, llamado institucional, contempla el desarrollo y consolidación de una organización local encargada del mantenimiento de los sistemas de agua para dar continuidad a la provisión del servicio al o largo del tiempo. Con esto se busca la sostenibilidad de los sistemas. Con estos tres componentes, el PNSR no sólo busca dotar de acceso a agua y saneamiento a los hogares, sino también garantizar su uso adecuado y sostenible.

Para intervenir, el PNSR se focalizó siguiendo una sencilla regla. En cada región priorizó distritos de acuerdo a su nivel de pobreza. Luego, dentro de los distritos priorizados eligió a los centros poblados con entre 201 y 2000 habitantes. Para esto último, la información del tamaño de población de los centros poblados procedió inicialmente del Censo de Población y Vivienda elaborada por el Instituto Nacional de Estadística e Informática (INEI) el año 2007. Sin embargo, durante las labores de validación en campo el personal del PNSR pudo identificar nuevas localidades que tenían menos de 200 habitantes en 2007 pero que en el 2012 habían superado ese umbral. A estas localidades también las incorporaron al padrón de beneficiarias. Por esta razón, la regla de asignación no se cumplió estrictamente.

3 Metodología

3.1 El diseño de Regresión Discontinua

La estrategia de identificación explota la regla de asignación del programa que depende del tamaño de población observada en el Censo de Población y Vivienda de 2007. En particular, dado que, de acuerdo a la normativa del PNSR, este se asigna entre localidades que tienen entre 200 y 2000 habitantes, entonces ello genera un salto discontinuo en la probabilidad de tratamiento en estos umbrales. Sin embargo, la aplicación de la regla no ocurrió de modo determinístico, pues hubo localidades con menos de 200 habitantes que fueron atendidas. Esto crea un contexto para la implementación de un estimador de *fuzzy* de regresiones discontinuas (RD) (Hahn et al. 2001, Imbens y Lemieux 2008).

El estimador RD explota una discontinuidad en el tratamiento para identificar una relación causal. Para esto, el tratamiento debe asignarse observando un punto de corte c en una variable de asignación z_j , de modo que todas las observaciones cuyo valor de z_j es inferior (o superior) a c obtengan el tratamiento. Bajo el diseño *fuzzy* de RD se reconoce que esta regla puede no cumplirse de modo estricto. Por ello, la mecánica de estimación sigue el procedimiento de dos etapas de variables instrumentales (Angrist y Pischke 2008: cap. 6): en la primera etapa se ajusta el tratamiento contra la regla de asignación (el instrumento) y en la segunda se regresiona la variable de resultado de interés contra la variable de tratamiento ajustada en la primera etapa. La siguiente especificación plantea esto:

$$\begin{aligned} T_{ij} &= \delta + \gamma \times D_{ij} + f(z_j) + \mu_{ij} & (i) \\ y_{ij} &= \alpha + \beta \times T_{ij} + g(z_j) + e_{ij} & (ii) \\ D_{ij} &= \mathbf{1}[z_j \geq 0] & (iii) \end{aligned}$$

Donde T_{ij} es la variable binaria del estado de tratamiento, que se define como residir en una localidad beneficiaria del PNSR, y D_{ij} es la binaria de cumplimiento de la regla de asignación, que se define como residir en localidades que en 2007 tenían 200 habitantes o más. Además, $\mathbf{1}[\cdot]$ es la función índice, z_j es la variable de asignación definida como el número de personas residentes en la localidad j en el año 2007 (centrada en 200), y_{ij} es la variable de resultado de interés (estado de ocupación, horas de trabajo, etc.), $f(\cdot)$ y $g(\cdot)$ son funciones polinómicas, μ_{ij} y e_{ij} son

términos de error y δ , γ , α , β son parámetros por estimar. En este marco de estimación, este último parámetro se define como el efecto local del PNSR (Angrist y Pischke 2008). Es decir, el impacto del PNSR entre los individuos que beneficiarios del programa por residir en localidades de 200 habitantes o más.

Además del estimador Fuzzy de RD que se recoge en (i) a (iii), también se implementará el estimador de intención del tratamiento (también llamado *forma reducida*. Ver Angrist y Pischke 2008). Es decir, aquél que vincula la variable de resultado y_{ij} con la binaria D_{ij} bajo la siguiente especificación:

$$y_{ij} = \theta + \tau \times D_{ij} + k(z_j) + v_{ij} \quad (iv)$$

En el marco de estimación, debido a que la regla se cumple imperfectamente, el parámetro de la intención del tratamiento es inferior en magnitud a la del efecto local del tratamiento.

El supuesto que subyace a la consistencia del estimador de regresiones discontinuas es que el resultado potencial en el estado contrafactual sean continuo en el umbral de discontinuidad y en un entorno a él (Imbens y Lemieux 2008). Para que esto funcione, se demanda que los determinantes de las variables de resultado transiten suavemente en el punto de discontinuidad, con lo que se descarta que los impactos que se encuentren, en caso existan, se deban a cambios en sus determinantes (Lee y Lemieux 2010).

Este es un supuesto de identificación y no puede ser testeado (los resultados potenciales no se observan; sólo se observan los resultados realizados). Por ello, típicamente se evalúa verificando que las variables en la línea de base no cambien discontinuamente en c . Por ello, se ha verificado gráficamente y a partir de sendas pruebas estadísticas la continuidad de un conjunto numeroso de variables explicativas potencialmente correlacionadas con las variables de resultado.

Sin embargo, este supuesto podría ser violado si existiera clasificación selectiva de la población en el entorno del punto de discontinuidad (McCrary, 2008). Ello ocurriría si, por ejemplo, los hogares enterados de la regla de asignación del programa y residentes en centros poblados pequeños migraran hacia algunas localidades hasta conseguir que estas tengan el tamaño mínimo de 200 habitantes considerado por el programa para ser beneficiarios. Si ello ocurriera, las localidades "formadas" con más de 200 habitantes que sean beneficiarias se diferenciarían de las no beneficiarias en el grado de motivación de los hogares frente al programa. Luego, dado que esta es una variable no observable que podría estar correlacionada con las variables de resultado evaluadas en este estudio, entonces tendería a sesgar los parámetros. Esta, no obstante, no es una amenaza real debido a que la información de tamaño poblacional utilizado por el programa provino del Censo de Población del año 2007. Es decir, esta es información que no pudo ser manipulada por los hogares. Más adelante se presenta evidencia de esto a partir de los resultados del test de McCrary.

Una particularidad del contexto de estimación es que la variable de asignación es discreta. Por ello, seguimos las recomendaciones de Lee y Card (2008), quienes sugieren clusterizar los errores estándar a nivel de la unidad discreta, que en nuestro caso es el tamaño de población de los centros poblados.

3.2 Descripción de los datos

La base de datos principal procede de la Encuesta Nacional de Hogares (ENAH) del año 2015. Esta es una encuesta que realiza el Instituto Nacional de Estadística e Informática (INEI) de modo regular desde el año 1997 para medir el avance en la reducción de pobreza, y cuenta, entre otros, con información del acceso a servicios básicos en el hogar (agua, saneamiento, electricidad, etc.) y características de los miembros del hogar (edad, sexo, nivel educativo, empleo, etc.) de una muestra total de alrededor de 110 mil personas distribuidas entre todos los departamentos del país. Desde años recientes, el INEI distribuye la base de datos de la ENAH con información georeferenciada de los conglomerados muestrales y con códigos estandarizados de los centros poblados de residencia de las personas que responden la encuesta. Estos códigos pueden vincularse con la información del Censo de Población y Vivienda del año 2007, también realizado por el INEI. Se ha optado por utilizar la información de este año debido a que según la revisión realizada, es en 2015 en que se incrementa la muestra de localidades rurales dispersas, que son aquellas con menos de 200 habitantes. Esto permitió tener un número relativamente grande de observaciones para implementar el estimador en el entorno de 200 habitantes.

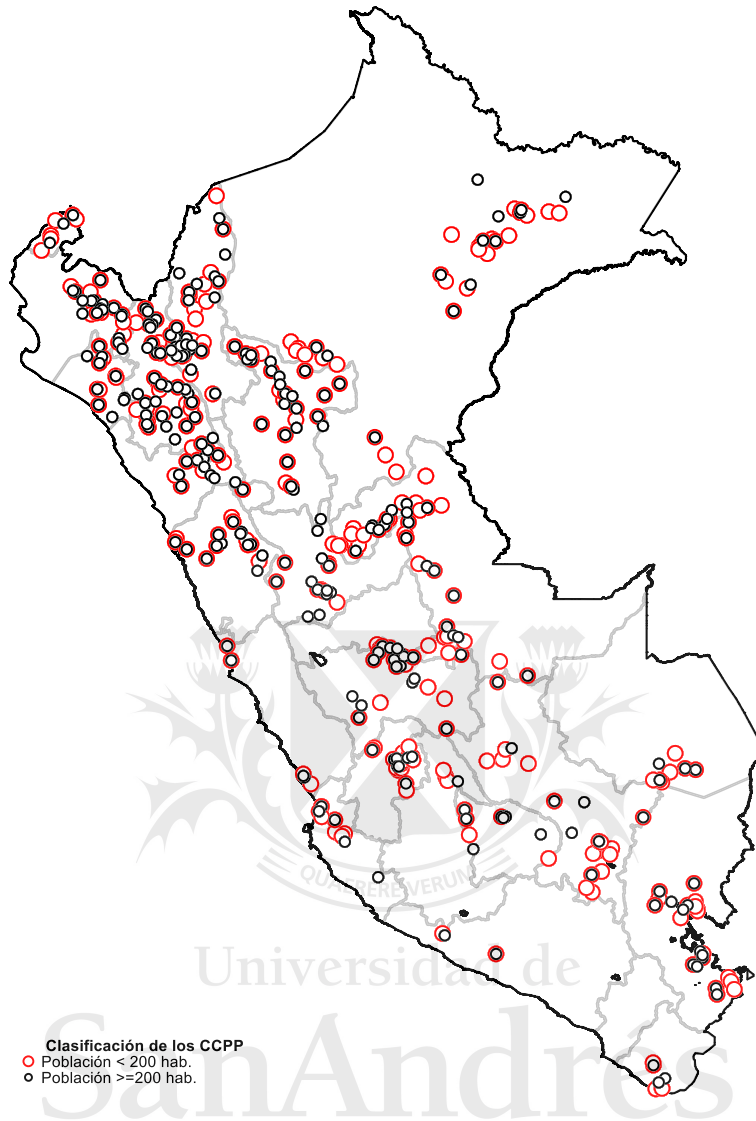
La segunda base de datos utilizada en el estudio es el Censo de Población y Vivienda del año 2007, que tiene información del total de personas en cada centro poblado visitado por los empadronadores. Esta base de datos, además de tener los códigos de identificación inequívocos por centro poblado, cuenta con las coordenadas de ubicación geográfica de estos.

La tercera base de datos utilizada es el padrón de centros poblados intervenidos por el PNSR. Este es un listado de localidades, codificadas según el identificador generado por el INEI en el Censo del 2007, con el estado de las obras de infraestructura de agua y saneamiento ejecutadas por los proyectos del programa o por ejecutar al año 2013. En total son 4269 localidades consignadas, de las cuales 57 no disponen de un código de centro poblado estandarizado, por lo que fueron retiradas de la muestra.

El pegado de las tres bases se hizo a partir del código de centro poblado. La ENAH recoge información de 3,397 centros poblados, todos en común con los del Censo de 2007. De este total, 377 son centros poblados atendidos por alguno de los proyectos del PNSR². A partir de este punto, se procedió de la siguiente manera. Primero, se tomaron únicamente los centros poblados de menos de 400 habitantes según la información del Censo 2007. De esta manera, para la evaluación del programa se tomará ventaja de la discontinuidad del tratamiento en el punto de 200 habitantes. Segundo, se mantuvo únicamente a los distritos que hubieran sido atendidos por el PNSR dentro de la muestra. Tercero, se quitaron las localidades que estuvieran a distancia mayor a 35 km de las localidades tratadas. La distribución de las poblaciones se muestra en el siguiente mapa. Como se aprecia, la muestra cubre todo el país. Más adelante, la Tabla 1 presenta estadísticas descriptivas de la muestra construida. De esta tabla llama la atención los porcentajes de hogares con acceso a agua de red domiciliaria y saneamiento básico (que incluye desagüe de red y letrinas): 46% y 20%. Estos datos revelan la alta brecha de acceso que persiste en la actualidad en el Perú.

² Se quitaron 17 localidades que eran atendidas por proyectos Frontera Norte, PRISAS y Plan Binacional.

Mapa 1. Distribución de las localidades de la muestra en el Perú



Elaboración propia

Tabla 1. Estadísticas descriptivas

	Obs	Prom.	Des.Est.	Min.	Med.	Máx.
VARIABLES A NIVEL DE HOGAR						
Agua de red	2956	0.46	0.50	0	0	1
Desagüe de red	2956	0.20	0.40	0	0	1
Índice de riqueza	2951	-2.38	1.42	-4.4	-2.7	5.8
Pared de ladrillo, piedra o adobe	2952	0.45	0.50	0	0	1
Piso diferente a tierra	2952	0.31	0.46	0	0	1
Techo de concreto, madera o tejas	2952	0.16	0.37	0	0	1
Hogar tiene conexión a electricidad	2956	0.63	0.48	0	1	1
Hogar tiene acceso a gas	2956	0.42	0.49	0	0	1
Jefe de hogar: Edad	2956	49.92	15.83	15	48	98
Jefe de hogar: Migrante	2956	0.29	0.46	0	0	1
Jefe de Hogar: Mujer	2956	0.18	0.38	0	0	1
Jefe de Hogar: Años de educación	2956	5.15	3.90	0	5	18
Nro de habitaciones	2952	2.69	1.51	1	2	13
VARIABLES GEOGRÁFICAS/A NIVEL DE CCPP						
Longitud	2956	-75.9	2.9	-80.8	-76.2	-69.1
Latitud	2956	-8.9	3.7	-18.2	-8.2	-2.9
Altitud	2956	1637	1462	17	1099	4422
Costa	2956	0.15	0.35	0	0	1
Sierra	2956	0.39	0.49	0	0	1
Selva	2956	0.46	0.50	0	0	1
Tamaño de población por localidad en 2007	2956	195	102	5	197	399
VARIABLES A NIVEL DE PERSONA (*)						
Participa en el mercado laboral	7161	0.81	0.39	0.00	1.00	1.00
Trabaja	7164	0.80	0.40	0.00	1.00	1.00
Horas de trabajo	5705	5.72	2.40	0.29	5.71	16.00
Persona: Años de educación	7164	5.73	4.16	0.00	6.00	18.00
Persona: Edad	7375	41.5	18.3	16.0	39.0	98.0
Persona: Mujer	8054	0.45	0.50	0.00	0.00	1.00
Persona: Migrante	8054	0.32	0.47	0.00	0.00	1.00
Persona: Vive en pareja	8054	0.60	0.49	0.00	1.00	1.00
Beneficiario de Juntos	8054	0.33	0.47	0.00	0.00	1.00

(*) Sólo personas de 16 años a más.

Fuente: ENAHO 2015, Censo de Población 2007, INEI.

Elaboración propia

4 Resultados

4.1 Validez del estimador

En esta sección se presentan dos tipos de ejercicios estándar para validar el estimador de regresiones discontinuas (Imbens y Lemiux 2008). En primer lugar, inspeccionamos algunas variables de control en el punto de discontinuidad. Como se indicó líneas arriba, el supuesto de identificación del estimador de regresiones discontinuas es que los resultados potenciales transiten suavemente en torno al umbral de discontinuidad. Sin embargo, debido a que esta no es una variable observable, entonces la estrategia pasa por observar las variables que estén correlacionadas con ella.

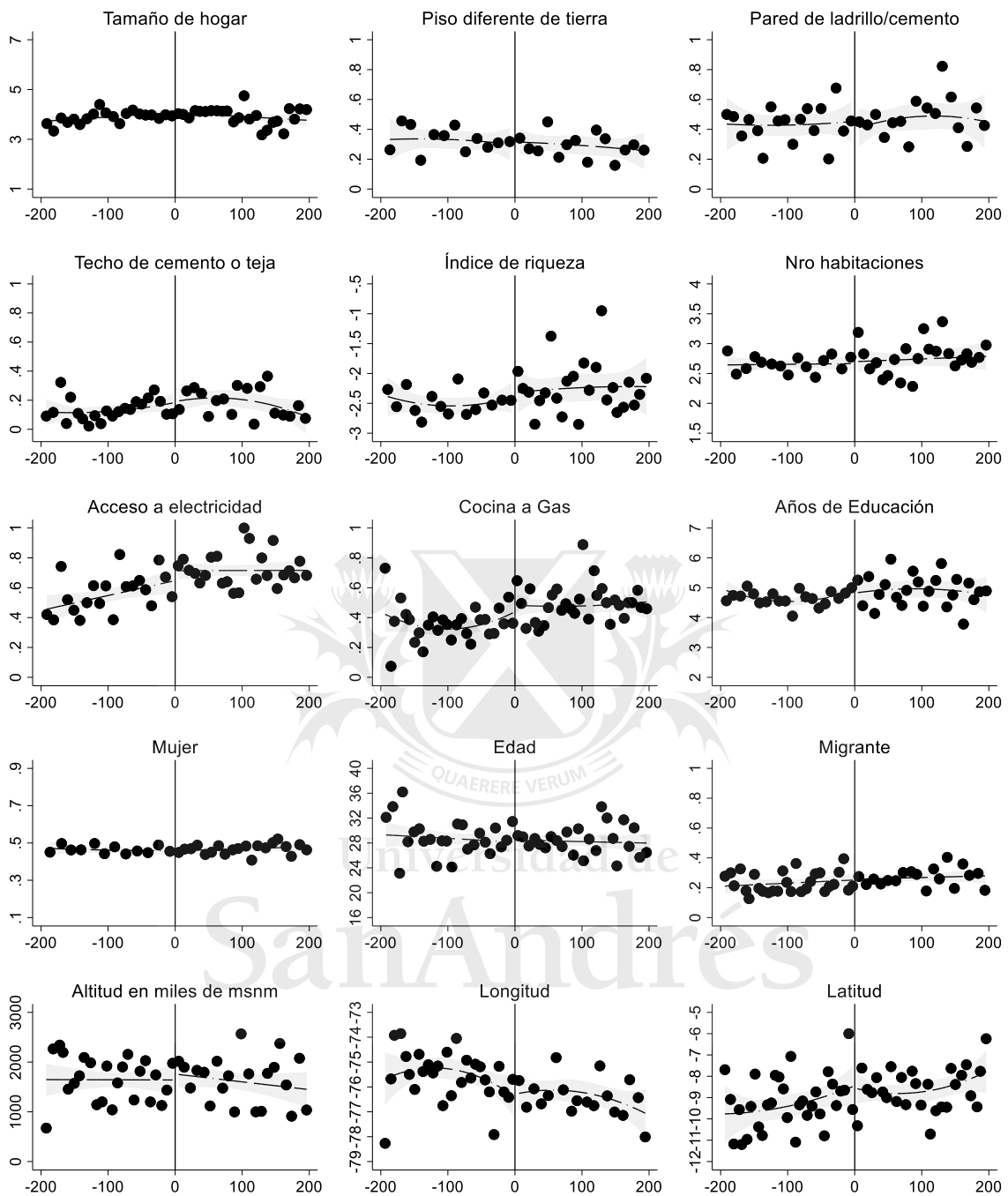
Para empezar, se presenta el **Gráfico 1**, que ofrece una prueba visual de la continuidad de un conjunto numeroso de variables de control a nivel de hogar, persona y geográficas. En cada caso, el eje horizontal se expresa como desviaciones con respecto al umbral de discontinuidad de 200 habitantes. La evidencia ofrecida no parece revelar saltos discontinuos resaltantes en ninguna variable.

Para evaluar más formalmente lo anterior, la **Tabla 2** muestra los resultados sobre las mismas variables. Las columnas (1), (5) y (9) presentan los valores promedio de cada variable entre los hogares residentes en localidades ubicadas a la izquierda del punto de corte de la regla de asignación, en radios de hasta 200 habitantes, 100 habitantes y 50 habitantes, respectivamente. Las columnas (2), (6) y (10), similarmente, presentan los valores promedio de cada variable entre los hogares residentes en localidades ubicadas a la derecha del punto de corte de la regla de asignación y dentro de los mismos radios. Las columnas (3), (7) y (11) presentan las diferencias de los valores promedio con sus errores estándar (clustereados a nivel de centros poblados) inmediatamente a la derecha. Claramente se aprecia que a medida que se acorta el rango de población alrededor del punto de corte, los promedios de las variables tienden a parecerse más entre las poblaciones ubicadas en los dos lados y las diferencias estadísticas desaparecen, con la única excepción de la variable que mide el acceso a electricidad de red.

Por otro lado, las columnas (13), (14) y (15) presentan los resultados de ajustar el modelo econométrico (i), pero tomando como dependiente cada una de las variables analizadas. En todos los casos se ha utilizado una función polinómica de tercer grado. Cuando se implementa este modelo sobre cada variable, las diferencias estadísticas en el punto de corte desaparecen, con excepción de la binaria que mide la proporción de personas residentes en la costa, que llega a ser significativa sólo al 10%. A pesar de esto, parece razonable concluir que las variables consideradas transitan de modo suave en el punto de corte.

El segundo tipo de ejercicios busca inspeccionar la continuidad de la variable de asignación para descartar su manipulación. Por ello, obtuvimos información del Censo de Población y Vivienda en el 2007 y analizamos la continuidad de la densidad del tamaño de población en el punto de discontinuidad. El panel a) del Gráfico 2 presenta el histograma de la variable de asignación, el tamaño de población por localidad, en rangos de 6 personas. Como se aprecia, este no parece mostrar saltos importantes en el punto de discontinuidad. En el panel b) se presenta el test de McCrary (2008), que es una prueba más formal de continuidad de la variable de asignación. Como se aprecia, la densidad de esta variable tampoco parece exhibir saltos estadísticamente significativos en el umbral. Con ello, se concluye que no parece haber evidencia de manipulación de la variable.

Gráfico 1. Continuidad de variables de control



Nota: En el eje horizontal se presenta el tamaño de población centrada en 200 y en el eje vertical el valor estimado de las variables de control indicadas en el título de cada gráfico. La selección de los anchos de banda se ha hecho siguiendo el procedimiento de Calónico et al. (2015). Las líneas entrecortadas representan las curvas ajustadas de regresiones realizadas de modo independiente en cada lado del gráfico. Las franjas plomas representan los intervalos de confianza de cada curva.

Fuente: ENAHO 2015, Censo de Población 2007, INEI.

Elaboración propia

Tabla 2. Pruebas de continuidad

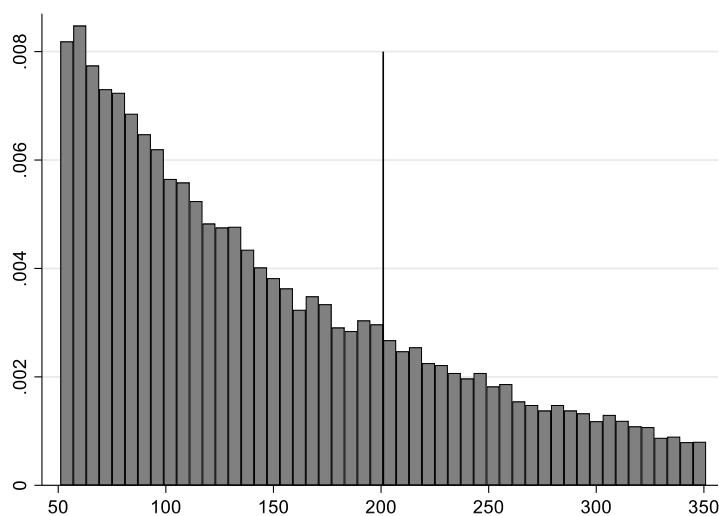
	Dentro de un radio de 200 hab.				Dentro de un radio de 100 hab.				Dentro de un radio de 50 hab.				Estimaciones RD		
	$D_{ij} = 0$ (1)	$D_{ij} = 1$ (2)	Dif. (3)	S.E. (4)	$D_{ij} = 0$ (5)	$D_{ij} = 1$ (6)	Dif. (7)	S.E. (8)	$D_{ij} = 0$ (9)	$D_{ij} = 1$ (10)	Dif. (11)	S.E. (12)	β (13)	S.E. (14)	R^2 (15)
VARIABLES A NIVEL DE HOGAR															
# miembros del hogar	3.92	3.94	0.02	(0.09)	3.95	4.03	0.09	(0.11)	3.96	4.04	0.08	(0.16)	0.28	(0.22)	0.78
Pared de adobe	0.37	0.39	0.02	(0.04)	0.39	0.37	-0.02	(0.04)	0.38	0.38	-0.01	(0.06)	-0.05	(0.08)	0.38
Piso de cemento	0.13	0.16	0.03	(0.02)*	0.11	0.16	0.04	(0.02)*	0.11	0.16	0.05	(0.03)	0.06	(0.05)	0.15
Techo de calaminas	0.59	0.65	0.05	(0.03)	0.62	0.61	-0.01	(0.04)	0.62	0.62	0.00	(0.06)	-0.01	(0.08)	0.62
Índice de riqueza	-2.50	-2.26	0.24	(0.09)**	-2.50	-2.32	0.18	(0.11)*	-2.50	-2.35	0.15	(0.14)	0.06	(0.20)	0.74
Nro de habitaciones	2.66	2.74	0.08	(0.08)	2.64	2.66	0.02	(0.10)	2.73	2.73	0.00	(0.15)	0.09	(0.20)	0.76
Hogar tiene conexión a electricidad	0.56	0.72	0.16	(0.03)***	0.60	0.70	0.09	(0.04)**	0.60	0.72	0.12	(0.06)**	0.10	(0.07)	0.65
Hogar tiene acceso a gas	0.36	0.48	0.13	(0.03)***	0.36	0.47	0.11	(0.04)**	0.38	0.45	0.08	(0.06)	0.05	(0.08)	0.43
VARIABLES A NIVEL DEL JEFE DE HOGAR															
Jefe de Hogar: Años de educación	5.10	5.20	0.09	(0.19)	5.07	5.26	0.19	(0.24)	5.07	5.05	-0.01	(0.36)	-0.21	(0.47)	0.64
Jefe de Hogar: Mujer	0.18	0.18	0.01	(0.02)	0.18	0.17	-0.01	(0.02)	0.18	0.16	-0.02	(0.03)	-0.02	(0.03)	0.18
Jefe de hogar: Edad	50.58	49.21	-1.36	(0.71)*	50.64	49.37	-1.27	(0.94)	50.65	50.08	-0.58	(1.33)	-1.10	(1.89)	0.91
Jefe de hogar: Migrante	0.27	0.32	0.06	(0.03)*	0.27	0.31	0.05	(0.04)	0.26	0.32	0.07	(0.05)	0.05	(0.07)	0.30
VARIABLES A NIVEL DE LAS PERSONAS															
Persona: Años de educación	4.81	5.09	0.29	(0.13)*	4.84	5.10	0.26	(0.18)	4.94	5.11	0.17	(0.27)	-0.01	(0.36)	0.59
Persona: Edad	29.49	29.19	-0.30	(0.61)	28.96	29.16	0.19	(0.79)	29.14	29.18	0.04	(1.23)	-1.72	(1.73)	0.63
Persona: Mujer	0.47	0.47	0.00	(0.01)	0.47	0.47	0.00	(0.01)	0.46	0.47	0.00	(0.01)	0.00	(0.02)	0.46
Persona: Migrante	0.22	0.23	0.02	(0.02)	0.24	0.23	-0.01	(0.03)	0.24	0.22	-0.02	(0.04)	-0.03	(0.05)	0.25
Persona: Vive en pareja	0.39	0.39	0.00	(0.01)	0.38	0.38	0.01	(0.01)	0.38	0.38	0.01	(0.02)	0.00	(0.02)	0.38
VARIABLES GEOGRÁFICAS															
Longitud	-75.54	-76.32	-0.79	(0.24)***	-75.64	-76.21	-0.57	(0.30)*	-76.19	-76.27	-0.08	(0.43)	0.00	(0.62)	1.00
Latitud	-9.25	-8.59	0.66	(0.30)**	-9.12	-8.62	0.49	(0.39)	-8.57	-8.70	-0.13	(0.55)	0.03	(0.77)	0.86
Altitud	1643.8	1629.3	-14.52	(117.5)	1670.6	1679.8	9.23	(149.8)	1624.8	1779.4	154.5	(200.6)	245.4	(281.0)	0.56
Costa	0.15	0.14	-0.01	(0.03)	0.14	0.12	-0.02	(0.03)	0.15	0.09	-0.05	(0.05)	-0.13	(0.07)*	0.16
Sierra	0.39	0.39	0.00	(0.04)	0.40	0.40	0.00	(0.05)	0.38	0.43	0.05	(0.07)	0.05	(0.10)	0.39
Selva	0.46	0.47	0.01	(0.04)	0.46	0.48	0.02	(0.05)	0.47	0.48	0.00	(0.07)	0.08	(0.10)	0.46

Nota: Las columnas (1), (2), (5), (6), (9) y (10) presentan los valores promedio de cada variable en cada lado del punto de corte, dentro del rango indicado. Las columnas (3), (7) y (11) presentan las diferencias entre los promedios de cada grupo, mientras que las columnas (4), (8) y (12) muestra los errores estándar clustereados a nivel de tamaño de población. La columna (13) presenta el parámetro de la regresión discontinua asociada a cada variable tomando una función polinómica de tercer grado, la columna (14) es el error estándar respectivo y la columna (15) el coeficiente de ajuste. Las estrellas *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Errores estándar robustos, clustereados a nivel de tamaño de población por centro poblado.

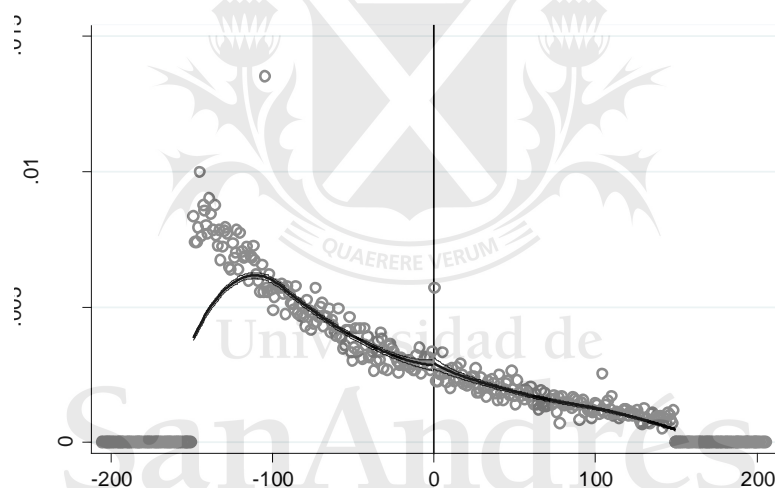
Fuente: ENAHO 2015, Censo de Población 2007, INEI
Elaboración propia

Gráfico 2. Pruebas de continuidad de la variable de asignación

a) Histograma de la variable de asignación



b) Test de McCrary



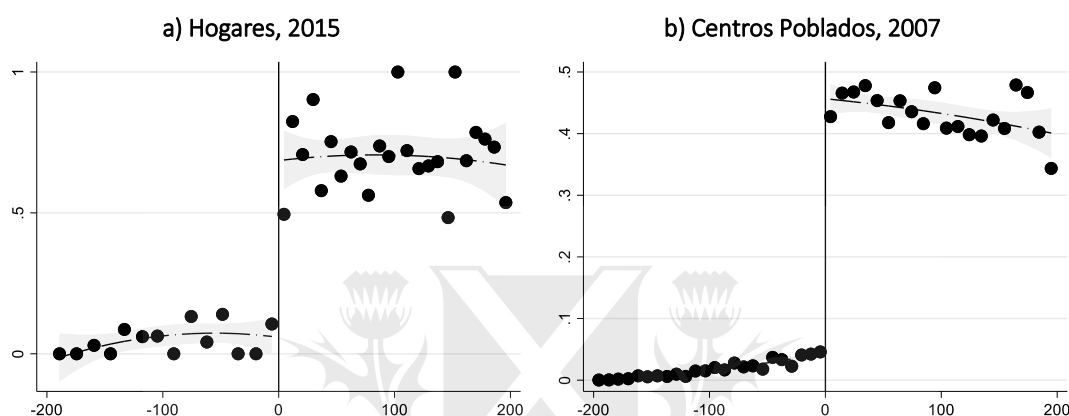
Test de discontinuidad = 0.03, S.E.=0.05
Elaboración propia

4.2 Relevancia del instrumento

Empezamos mostrando la relación entre la participación en el PNSR y el tamaño de población. De acuerdo a la regla, los hogares residentes en localidades de 200 habitantes a más deberían tener una mayor probabilidad de ser beneficiarios que aquellos en localidades más pequeñas. Para verificar esto, inicialmente graficamos la probabilidad de que un hogar sea beneficiario del programa en función del tamaño de población de la localidad, tomando información de la muestra de hogares construida a partir de la ENAHO 2015. Esto se presenta en el panel a) del **Gráfico 3**. En el eje horizontal se representa el tamaño de población de las localidades de residencia de los hogares como desvíos de 200, el punto de discontinuidad. En el eje vertical se presenta la proporción de hogares beneficiarios. Se han añadido curvas ajustadas de modo independiente con los puntos estimados de la derecha e izquierda del punto de discontinuidad, así como sus intervalos de confianza correspondientes. Como se aprecia, el gráfico reporta un salto de los datos en 0.

Adicionalmente, para validar el resultado anterior pero esta vez con datos censales, se volvieron a graficar los puntos que replican la proporción de beneficiarios según el tamaño de población de las localidades, pero esta vez con datos del Censo de Población de 2007. Los resultados, que se presentan en el panel b) del **Gráfico 3**, indican que la discontinuidad se mantiene en el umbral de asignación del tratamiento. Es decir, la información censal revela que dentro de los distritos priorizados para ser intervenidos, aquellos que en 2007 tenían 200 habitantes o más son lo que tienen una mayor probabilidad de ser parte de listado de localidades intervenidas por el PNSR en comparación con los que tienen menos de 200 habitantes. Este resultado se replica también con la información muestral que se utiliza para hacer los ejercicios de evaluación del PNSR.

Gráfico 3. Participación en el PNSR versus tamaño de población



Nota: En el eje horizontal se presenta el tamaño de población centrado en 200 y en el eje vertical el valor estimado de las variables de control indicadas en el título de cada gráfico. La selección de los anchos de banda se ha hecho siguiendo el procedimiento de Calonico et al. (2015). Las líneas entrecortadas representan las curvas ajustadas de regresiones realizadas de modo independiente en cada lado del gráfico. Las franjas plomas representan los intervalos de confianza de cada curva.

Fuente: ENAHO 2015, Censo de Población 2007, INEI.

Elaboración propia

Para evaluar más formalmente el resultado gráfico anterior, la **Tabla 3** presenta las estimaciones de la regresión que vincula la probabilidad de que un hogar sea beneficiario del PNSR con la regla de asignación. Para esto, se implementa la especificación (i) de la sección 3.1 utilizando un polinomio de tercer grado como aproximación de $f(z_j)$. De acuerdo al modelo *Fuzzy* de RD, esta regresión representa la primera etapa del modelo econométrico.

La columna (1) de la **Tabla 3** presenta el parámetro asociado a la variable binaria D_{ij} , que define el cumplimiento de la regla, sin incluir ningún regresor adicional al polinomio de tercer grado. La magnitud estimada es de alrededor de 0.6, significativo al 1%. En la columna (2) se replica la estimación anterior, pero esta vez añadiendo efectos fijos regionales para controlar la influencia de la política de saneamiento que implementan los gobiernos regionales. El parámetro se mantiene en el mismo valor y preserva significancia. La columna (3) de la tabla añade a la especificación anterior controles a nivel de hogar, para aislar particulares socioeconómicas que pudieron influir en la asignación del tratamiento. El parámetro se mantiene en magnitud y significancia. Finalmente, la columna (4) añade a la especificación anterior controles geográficos, como la altura del centro poblado, las coordenadas (longitud, latitud) y dos *dummies* que indican si el hogar reside en sierra o selva. El parámetro estimado continúa manteniéndose estadísticamente significativo.

Tabla 3. Probabilidad de que un hogar sea beneficiario, 2015

	(1)	(2)	(3)	(4)
D_{ij}	0.623*** (0.070)	0.626*** (0.073)	0.632*** (0.073)	0.632*** (0.072)
Obs	2,956	2,956	2,951	2,951
R2	0.444	0.501	0.508	0.519
Efectos fijos regionales	No	Si	Si	Si
Controles de hogar	No	No	Si	Si
Controles geográficos	No	No	No	Si

Nota: Se utiliza una función polinómica de tercer grado. Los controles de hogar incluyen la edad, los años de educación, el sexo y la condición de migrante del jefe de hogar; un índice de acceso a servicios en el hogar construido por componentes principales; una variable binaria que indica si el hogar tiene acceso a electricidad y una variable binaria que indica si la vivienda del hogar tiene un piso diferente de tierra. Los controles geográficos incluyen la longitud, latitud y altitud del centro poblado de residencia de los hogares. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población por centro poblado. *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: EN AHO 2015, Censo de Población 2007, INEI. Elaboración propia

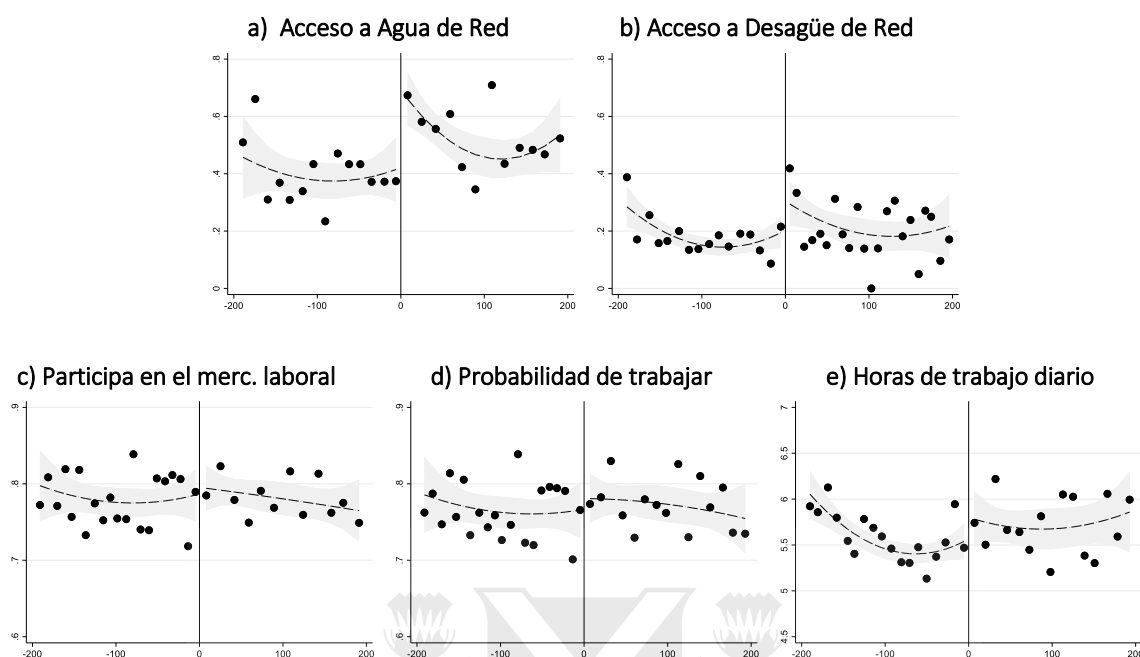
4.3 Impacto en el acceso a agua y la oferta laboral

Se presenta ahora los resultados del ejercicio de evaluación de impacto, para lo cual se muestra inicialmente el **Gráfico 4**, que recoge el diseño de regresiones discontinuas. Aquí también los puntos representan los promedios de las variables dependientes entre hogares que caen dentro de distintos rangos de población. En cada gráfico se ha añadido una curva que ajusta los puntos de cada lado, así como áreas grises que señalan los intervalos de confianza respectivos al 90%. La línea vertical centrada en cero representa el punto de discontinuidad.

El panel a) del **Gráfico 4** muestra las tasas de acceso a agua de red domiciliar de los hogares en función del tamaño de las localidades en 2007. El detalle más resaltante de este gráfico es el salto abrupto en la curva ajustada en el punto inmediatamente a la derecha de la recta vertical. El panel b) del mismo gráfico muestra las tasas de acceso a desagüe de red y letrinas domiciliarias. Aquí, claramente las tasas son mucho menores que las de acceso a agua de red, pero continúa observándose un pequeño salto en la línea vertical. Ambas variables representan los resultados inmediatos del PNSR por tratarse de los productos que el programa entrega a los hogares beneficiarios. Los saltos discontinuos, por tanto, reflejarían los impactos de corto plazo del programa.

Los paneles c), d) y e) del **Gráfico 4** muestran resultados a nivel de las variables de mercado laboral planteadas para personas de 16 años a más. La participación en el mercado laboral (panel c) se mide como el conjunto de personas que ofrecen mano de obra en el mercado, independientemente de su condición de ocupación. Se mide como las personas que trabajan o se encuentra desempleadas. El gráfico no parece mostrar un salto significativo en el umbral de discontinuidad. La población que trabaja (panel d) se mide como el conjunto de personas que en las dos semanas previas a la encuesta contaban con algún trabajo, ya sea por cuenta propia o en condición de dependiente de algún empleador. Aquí tampoco parece visualizarse un salto significativo en el umbral de discontinuidad. Las horas trabajadas (panel e), finalmente, se definen como el número promedio de horas por día que cada persona trabajó en las dos semanas previas a la fecha de la encuesta. Esta variable se calcula únicamente entre personas que reportaron tener algún trabajo. Aquí sí se aprecia un pequeño salto, pero que no parece ser significativo debido a la amplitud del intervalo de confianza.

Gráfico 4. Ilustración de los resultados del diseño RD



Nota: En el eje horizontal se presenta el tamaño de población centrada en 200 y en el eje vertical el valor estimado de las variables de control indicadas en el título de cada gráfico. La selección de los anchos de banda se ha hecho siguiendo el procedimiento de Calonico et al. (2015). Las líneas entrecortadas representan las curvas ajustadas de regresiones realizadas de modo independiente en cada lado del gráfico. Las franjas plomas representan los intervalos de confianza de cada curva.

Fuente: ENAHO 2015, Censo de Población 2007, INEI.

Elaboración propia

Los estimados formales sobre las variables de producto se presentan en la **Tabla 4**. En todos los casos la función $f(z_j)$ se ha aproximado como un polinomio de tercer grado. El panel a) de la tabla muestra los resultados de la intención del tratamiento, que recoge el impacto de la regla de asignación. Es decir, se ajusta regresiones que vinculan como variables dependientes el acceso a agua de red y el acceso a desagüe de red contra D_{ij} . Como se aprecia, en el caso del acceso a agua el parámetro de interés estimado toma un valor de 0.24 que es robusto a la inclusión de diversas variables de control, como efectos fijos departamentales, controles de hogar y controles geográficos. En el caso del acceso a saneamiento básico, el parámetro estimado sin ningún control carece de significancia, pero inmediatamente en las estimaciones siguientes este se hace positivo y significativo en poco más de 0.10, robusto también a la inclusión de todos los controles considerados. En la última fila del panel a) se ha incluido el valor promedio de la variable dependiente en el grupo de control. Por ejemplo, entre los controles el 39.4% de los hogares acceden a agua de red. Siendo así, entre los hogares residentes en localidades de 200 habitantes a más pero ubicadas cerca al punto de discontinuidad esta magnitud se incrementa en alrededor de 24 puntos porcentuales adicionales.

En el panel b) de la **Tabla 4** se presentan los estimados del diseño *fuzzy* de regresiones discontinuas, que corrige por las diferencias en las tasas de participación inducidas por el cumplimiento de la regla (Angrist y Pischke 2008). Esta vez, en el caso del acceso a agua de red en el hogar el impacto ha sido estimado en cerca de 0.38 en la especificación sin controles, siendo el parámetro robusto a la introducción de efectos fijos regionales, controles de hogar y controles geográficos. En la tabla además se ha añadido el valor del estadístico F de la regresión de la primera etapa para aproximar su validez. Dado que en todas las regresiones se ha calculado errores estándar robustos, clustereados a nivel del tamaño poblacional, entonces este parámetro es igual al estadístico de Klinderber y Paap (2006). Staiger y Stock (1997) recomiendan asegurarse

de que este estadístico tenga un valor superior 10 para descartar problemas de instrumentos débiles³. Como se aprecia, en todas las estimaciones el valor de la prueba F de la primera etapa supera el umbral de 10, por lo cual se concluye que no hay evidencia de instrumentos débiles.

Las columnas (5) a (8) del panel b) de la **Tabla 4** muestran los impactos sobre el acceso a saneamiento básico. Como en el caso de la intención del tratamiento, la especificación sin controles (columna 5) reporta un parámetro positivo pero no significativo. Sin embargo, la añadidura de controles incrementa rápidamente la precisión de los estimados y con ello la significancia estadística. En la especificación que controla por efectos fijos departamentales, controles de hogar y controles geográficos, el parámetro de impacto es de 0.17, el cual es similar al valor del grupo de control. Es decir, el acceso a saneamiento se habría duplicado entre beneficiarios por causa del PNSR.

Tabla 4. Resultados del estimador de regresiones discontinuas a nivel de productos

	El hogar tiene agua de red				El hogar tiene saneamiento básico			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
a) Intención del tratamiento								
D_{ij}	0.239*** (0.080)	0.268*** (0.072)	0.218*** (0.068)	0.219*** (0.067)	0.097 (0.062)	0.134** (0.056)	0.111** (0.056)	0.109** (0.055)
Obs.	2,956	2,956	2,951	2,951	2,956	2,956	2,951	2,951
R2	0.022	0.175	0.253	0.268	0.008	0.121	0.168	0.173
b) Fuzzy RD								
T_{ij}	0.383*** (0.132)	0.427*** (0.124)	0.345*** (0.115)	0.347*** (0.113)	0.156 (0.102)	0.214** (0.099)	0.176* (0.095)	0.173* (0.094)
Obs.	2,956	2,956	2,951	2,951	2,956	2,956	2,951	2,951
R2	-0.073	-0.110	0.028	0.045	-0.032	-0.065	0.012	0.020
F-stat 1ra Etapa	79.21	72.60	75.28	76.20	79.21	72.60	75.28	76.20
Var. Dep. Prom.	0.394				0.175			
EF Departamentales	No	Si	Si	Si	No	Si	Si	Si
Controles de hogar	No	No	Si	Si	No	No	Si	Si
Controles geográficos	No	No	No	Si	No	No	No	Si

Nota: Se utiliza una función polinómica de tercer grado. Los controles de hogar incluyen la edad, los años de educación, el sexo y la condición de migrante del jefe de hogar; un índice de acceso a servicios en el hogar construido por componentes principales; una variable binaria que indica si el hogar tiene acceso a electricidad. Los controles geográficos incluyen la longitud, latitud y altitud del centro poblado de residencia de los hogares, así como *dummies* que identifican observaciones en la sierra y selva. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población por centro poblado. *,** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia

En la **Tabla 5** se presentan los resultados del impacto del PNSR sobre las variables laborales, también utilizando una función polinómica de tercer grado. Las columnas (1) a (4) reportan los impactos sobre el indicador de participación en el mercado laboral. En el panel a) se presentan los resultados de la estimación de intención del tratamiento, mientras que en el panel b) se exhiben los impactos estimados del diseño *fuzzy* RD. La tabla parece indicar que el impacto sobre esta variable es nulo, pues, aun cuando el parámetro es positivo, no resulta significativo en ningún caso, bajo ninguna especificación.

Las columnas (5) a (8) reportan los impactos sobre la variable que indica si el individuo trabaja. En este caso, los resultados son similares a los de la variable anterior, pues los impactos parecen nulos, tanto si se mide la intención del tratamiento, como el impacto local del tratamiento recogido por el *fuzzy* RD.

³ Este umbral sólo es válido cuando se trata de un modelo exactamente identificado con un regresor endógeno y un instrumento excluido, como es el caso del estudio. En modelos sobreidentificados es necesario comparar los valores de Klinderber y Paap (2006) con las tablas de Stock y Yogo (2005).

Tabla 5. Resultados del estimador de regresiones discontinuas a nivel de variables laborales

	Participación en el mercado laboral				Probabilidad de trabajar				Nro de horas de trabajo diario					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)		
a) Intención del tratamiento														
D_{ij}	0.013 (0.025)	0.016 (0.025)	0.021 (0.025)	0.023 (0.025)	0.023 (0.025)	0.023 (0.026)	0.029 (0.025)	0.030 (0.026)	0.340* (0.191)	0.438** (0.196)	0.419** (0.205)	0.414** (0.190)		
Obs.	6,915	6,915	6,906	6,906	6,917	6,917	6,908	6,908	5,569	5,569	5,564	5,564		
R2	0.000	0.018	0.087	0.089	0.000	0.019	0.093	0.095	0.004	0.078	0.142	0.146		
b) RD Fuzzy														
T_{ij}	0.021 (0.041)	0.027 (0.042)	0.035 (0.042)	0.037 (0.041)	0.037 (0.041)	0.037 (0.043)	0.047 (0.043)	0.049 (0.043)	0.551* (0.314)	0.715** (0.318)	0.680** (0.325)	0.670** (0.300)		
Obs.	6,915	6,915	6,906	6,906	6,917	6,917	6,908	6,908	5,569	5,569	5,564	5,564		
R2	-0.000	-0.000	0.069	0.071	-0.000	-0.000	0.074	0.076	-0.005	-0.010	0.061	0.066		
F-Stat 1ra Etapa	71.08	64.32	67.11	68.48	71.11	64.32	67.11	68.47	63.61	57.78	59.90	60.55		
Var.Dep.Prom.		0.816					0.805				5.699			
FE departamental	No	Si	Si	Si	No	Si	Si	Si	No	Si	Si	Si		
Controles de Hogar	No	No	Si	Si	No	No	Si	Si	No	No	Si	Si		
Controles geográficos	No	No	No	Si	No	No	No	Si	No	No	No	Si		

Nota: Se utiliza una función polinómica de tercer grado. Los controles de hogar incluyen la edad, los años de educación, el sexo, el sexo al cuadrado y la condición de migrante del jefe de hogar; una variable binaria que indica si el hogar tiene acceso a electricidad. Los controles geográficos incluyen la longitud, latitud y altitud del centro poblado de residencia de los hogares, así como *dummies* que identifican observaciones en la sierra y selva. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población por centro poblado. *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia

San Andrés

Las columnas (9) a (12) de la **Tabla 5** presentan los resultados sobre la variable que recoge el número de horas de trabajo diario. En este caso, en la especificación sin regresores adicionales el parámetro de impacto ha sido estimado en 0.34 bajo el estimador de intención del tratamiento y en 0.55 bajo el estimador *Fuzzy RD*. Cuando se añaden efectos fijos departamentales (columna 10) los parámetros se incrementan levemente pero mantienen significancia, al igual que cuando se añaden controles que buscan aislar la influencia de la educación, el sexo y la edad de las personas, así como la participación en Juntos por parte del hogar, la condición migratoria del jefe de hogar, etc. (columna 11). En este último caso, los parámetros se mantienen en 0.41 y 0.68 bajo los estimados de la intención del tratamiento y *Fuzzy RD*, respectivamente. Los parámetros estimados tampoco cambian significativamente cuando se añaden regresores geográficos para controlar por la ubicación geográfica, altitud y región natural de residencia (columna 12). Es decir, independientemente de la especificación, los impactos del PNSR sobre las horas trabajadas se ubican entre 0.55 y 0.67 horas (diseño de RD), que equivalen a alrededor de 33 y 40 minutos por día.

En la **Tabla 5** también se ha añadido el estadístico F de la primera etapa de estimación del modelo *Fuzzy RD* en cada especificación utilizada. En todos los casos el estadístico supera el umbral de 10, por lo cual se descarta que existan problemas de instrumentos débiles que estén sesgando los parámetros estimados.

Para poner en perspectiva los resultados obtenidos en la **Tabla 5**, resulta útil compararlos con lo observado en la literatura revisada. En principio, la ausencia de impactos en la probabilidad de ocuparse es consistente con los hallazgos Devoto et al. (2012), quienes no encuentran que la conexión a agua incremente la probabilidad de que los jefes de hogar tengan una actividad generadora de ingresos en Marruecos. Asimismo, los impactos observados sobre las horas diarias de trabajo son consistentes con los hallazgos de Meeks (2012) en Kirguistán, quien encuentra un impacto del acceso a agua domiciliaria en el trabajo productivo de hasta 60 minutos diarios, y Zhou y Turvey (2017) en China, quienes encuentran que la falta de acceso a agua inhibe la oferta laboral en el margen intensivo en hasta 6.5%.

Por otro lado, también resulta útil comparar los estimados de impacto con los tiempos destinados al acarreo de agua en el hogar. En Perú, no existe encuesta de uso de tiempo del año 2015. La encuesta más reciente es la del año 2010. Según esta base, en promedio una persona de 16 años o más sin acceso a agua en el hogar en el ámbito rural dedicaba 30.3 minutos diarios al acarreo de agua (2.12 horas semanales). Los impactos estimados en las horas de trabajo diario mostrados en la **Tabla 5** son de entre 33 y 40 minutos. Por tanto, el acceso a agua y saneamiento básico domiciliarios estaría permitiendo liberar no sólo el tiempo de acarreo del agua, sino también el de otras actividades domésticas relacionadas a la higiene, aseo y limpieza fuera del hogar (como el lavado de ropa, que se suele realizar, por ejemplo, en la fuente de agua), entre otros, para destinarlos al trabajo productivo. Este mismo resultado fue también encontrado por Meeks (2012).

4.4 Robustez de las estimaciones

Para validar los hallazgos anteriores, en esta sección se implementan dos ejercicios de verificación de la robustez de los parámetros. Primero, en las estimaciones de la **Tabla 4** y la **Tabla 5** se utilizan arbitrariamente polígonos de tercer grado para aproximar las funciones $f(z_j)$ y $g(z_j)$. Por ello, se verifica la robustez de los estimados ante polígonos de diverso grado. Segundo, los estimados dependen crucialmente de la discontinuidad del tratamiento en el umbral de 200 habitantes. Si este umbral se modificara arbitrariamente, entonces debería observarse que los parámetros se tornan nulos estadísticamente. Por ello, también se implementa una prueba de falsificación que

muestra los cambios en los parámetros de impacto antes variaciones artificiales en el punto de discontinuidad de asignación del tratamiento.

La siguiente tabla presenta las estimaciones adicionales para diversos grados de las funciones polinómicas. Estas se han hecho para las tres variables de resultado laboral utilizando el estimador de intención del tratamiento y RD *fuzzy*. En el anexo del documento se presentan ejercicios de sensibilidad similares para las variables de acceso a agua y saneamiento básico domiciliarios.

Tabla 6. Sensibilidad de las estimaciones de impacto en variables laborales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Participación en el mercado laboral								
Intención del Tratamiento								
D_{ij}	0.009	0.012	0.008	0.011	0.013	0.023	0.015	0.024
	(0.020)	(0.019)	(0.020)	(0.019)	(0.025)	(0.025)	(0.024)	(0.025)
Obs.	6,915	6,906	6,915	6,906	6,915	6,906	6,915	6,906
RDD Fuzzy								
T_{ij}	0.014	0.020	0.012	0.018	0.021	0.037	0.024	0.040
	(0.032)	(0.031)	(0.032)	(0.032)	(0.041)	(0.041)	(0.040)	(0.041)
Obs.	6,915	6,906	6,915	6,906	6,915	6,906	6,915	6,906
F-Stats 1ra Etapa	119.2	125	128.9	131.5	71.08	68.48	75.13	70.47
Probabilidad de trabajar								
Intención del Tratamiento								
D_{ij}	0.016	0.020	0.015	0.019	0.023	0.030	0.025	0.032
	(0.020)	(0.020)	(0.020)	(0.020)	(0.025)	(0.026)	(0.024)	(0.025)
Obs.	6,917	6,908	6,917	6,908	6,917	6,908	6,917	6,908
RDD Fuzzy								
T_{ij}	0.026	0.032	0.025	0.031	0.037	0.049	0.040	0.052
	(0.032)	(0.032)	(0.032)	(0.033)	(0.041)	(0.043)	(0.040)	(0.043)
Obs.	6,917	6,908	6,917	6,908	6,917	6,908	6,917	6,908
F-Stats 1ra Etapa	119.2	125.1	128.9	131.6	71.11	68.47	75.15	70.46
Horas trabajadas								
Intención del Tratamiento								
D_{ij}	0.376**	0.377***	0.403**	0.396***	0.340*	0.414**	0.323*	0.411**
	(0.157)	(0.146)	(0.157)	(0.145)	(0.191)	(0.190)	(0.188)	(0.189)
Obs.	5,569	5,564	5,569	5,564	5,569	5,564	5,569	5,564
RDD Fuzzy								
T_{ij}	0.606**	0.609***	0.655**	0.643***	0.551*	0.670**	0.523*	0.668**
	(0.258)	(0.231)	(0.259)	(0.232)	(0.314)	(0.300)	(0.307)	(0.300)
Obs.	5,569	5,564	5,569	5,564	5,569	5,564	5,569	5,564
F-Stats 1ra Etapa	109.2	114.1	117.9	120.1	63.61	60.55	67.60	62.54
Grado del polígono	Lineal		Cuadrática		Cúbica		Cuarta	
Controles	No	Sí	No	Sí	No	Sí	No	Sí

Nota: Los controles incluyen efectos fijos regionales, la edad, los años de educación, el sexo, el sexo al cuadrado y la condición de migrante del jefe de hogar; una variable binaria que indica si el hogar tiene acceso a electricidad. Además se incluyen controles geográficos, como la longitud, latitud y altitud del centro poblado de residencia de los hogares, así como *dummies* que identifican observaciones en la sierra y selva. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población. *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia

Los resultados de la **Tabla 6** son sugerentes y replican lo mostrado en la **Tabla 5**. Es decir, los parámetros asociados a la participación en el mercado laboral y el indicador que indica si el individuo trabaja no resultan significativos bajo ninguna especificación. A diferencia, las estimaciones del número de horas trabajadas por día sí muestran un impacto positivo y significativo bajo cualquier especificación. En el caso del estimador de la intención del tratamiento, los parámetros estimados toman valores de entre 0.32 y 0.41, significativos en todos

los casos, mientras que bajo el estimador RD *Fuzzy* los parámetros se ubican entre 0.55 y 0.66, también significativos en todos los casos. Es decir, los impactos del PNSR sobre las horas de trabajo diario son robustos en magnitud y significancia al orden de los polinomios seleccionados y a la inclusión de regresores potencialmente explicativos.

Por otro lado, en la **Tabla 7** se presentan los resultados de las pruebas de falsificación. En este caso, artificialmente se ha modificado el punto de corte a 80, 120, 160, 200, 240, 280 y 320, de modo que se evalúa cómo cambian los parámetros de impacto en cada caso. Bajo este ejercicio, debido a que un supuesto de identificación es que las variables de resultado transitan suavemente fuera del punto de discontinuidad (Imbens y Lemieux 2008), entonces los parámetros de impacto estimados no deberían ser significativos en los puntos de discontinuidad artificialmente creados. Si esto no se cumpliera, entonces existirían elementos de juicio para sospechar de la validez de los estimados anteriores.

Los resultados son sugerentes. Por ejemplo, bajo el estimador de la intención del tratamiento, la participación en el mercado laboral y el indicador de trabajo no resultan significativos en ningún punto de discontinuidad. En cambio, el impacto sobre las horas trabajadas sólo es significativo en el punto de discontinuidad de 200, pero no fuera de él. Los resultados del diseño fuzzy de RD son muy similares. Es decir, el indicador de participación laboral y el indicador de trabajo son no significativos en ningún caso. En cambio, el indicador de horas trabajadas diarias es significativo en el punto de discontinuidad de 200, pero no en los demás puntos considerados. Más aún, la mayor relevancia del instrumento en la primera etapa de estimación sólo se sostiene cuando el umbral de discontinuidad es 200, que es cuando el estadístico F supera valores de 60. En los demás puntos el estadístico F cae por debajo del valor referencial de 10, lo que indica que en esos casos el instrumento no resulta relevante. Sólo el umbral de 120 el estadístico F de la primera etapa supera levemente el valor referencial de 10, pero los parámetros de impacto están imprecisamente estimados y no llegan a ser significativos, y tienen signo diferente de los parámetros del estimador de la intención del tratamiento.

En suma, los resultados de la **Tabla 7** sugieren que las variables de resultado sólo exhiben saltos discontinuos en el punto de corte de 200 habitantes, que es el umbral que determina la asignación del PNSR. Fuera de este punto las variables de resultado parecen transitar suavemente, en línea con el supuesto de identificación del estimador.

Tabla 7. Pruebas de falsificación

	Puntos de corte						
	<i>c</i> = 80	<i>c</i> = 120	<i>c</i> = 160	<i>c</i> = 200	<i>c</i> = 240	<i>c</i> = 280	<i>c</i> = 320
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intención del Tratamiento							
Participación en el mercado laboral							
<i>D_{ij}</i>	0.021 (0.030)	-0.016 (0.023)	-0.003 (0.023)	0.023 (0.025)	-0.013 (0.025)	0.026 (0.024)	0.011 (0.033)
Obs.	6,906	6,906	6,906	6,906	6,906	6,906	6,906
Probabilidad de trabajar							
<i>D_{ij}</i>	0.015 (0.032)	-0.020 (0.024)	-0.005 (0.025)	0.030 (0.026)	-0.012 (0.025)	0.023 (0.025)	-0.000 (0.035)
Obs.	6,908	6,908	6,908	6,908	6,908	6,908	6,908
Horas trabajadas							
<i>D_{ij}</i>	0.005 (0.214)	-0.234 (0.186)	0.177 (0.174)	0.414** (0.190)	-0.180 (0.184)	-0.159 (0.224)	-0.303 (0.274)
Obs.	5,564	5,564	5,564	5,564	5,564	5,564	5,564
Fuzzy RDD							
Participación en el mercado laboral							
<i>T_{ij}</i>	-0.313 (0.541)	0.095 (0.138)	0.078 (0.698)	0.037 (0.041)	0.063 (0.117)	-0.126 (0.127)	-0.137 (0.412)
Obs.	6,906	6,906	6,906	6,906	6,906	6,906	6,906
F-Stata 1ra etapa	1.195	14.25	0.218	68.48	5.548	5.182	0.526
Probabilidad de trabajar							
<i>T_{ij}</i>	-0.222 (0.529)	0.116 (0.145)	0.143 (0.771)	0.049 (0.043)	0.058 (0.117)	-0.110 (0.129)	0.002 (0.414)
Obs.	6,908	6,908	6,908	6,908	6,908	6,908	6,908
F-Stata 1ra etapa	1.195	14.29	0.215	68.47	5.548	5.183	0.527
Horas trabajadas por día							
<i>T_{ij}</i>	-0.072 (3.314)	1.342 (1.052)	-6.943 (21.991)	0.670** (0.300)	0.816 (0.847)	0.741 (1.142)	2.622 (3.615)
Obs.	5,564	5,564	5,564	5,564	5,564	5,564	5,564
F-Stata 1ra etapa	1.00	13.88	0.115	60.55	5.598	5.195	0.931

Nota: Se utilizan polígonos de tercer grado. Todas las regresiones incluyen los siguientes controles: efectos fijos regionales, la edad, los años de educación, el sexo, el sexo al cuadrado y la condición de migrante del jefe de hogar; una variable binaria que indica si el hogar tiene acceso a electricidad. Además se incluyen controles geográficos, como la longitud, latitud y altitud del centro poblado de residencia de los hogares, así como *dummies* que identifican observaciones en la sierra y selva. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población. *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia

4.5 Impactos heterogéneos

Los resultados anteriores parecen sugerir que el PNSR habría logrado impactar en el margen intensivo de la oferta laboral al incrementar el número de horas trabajadas, pero no en el margen extensivo, relacionado a la participación en el mercado de trabajo y en la probabilidad de trabajar. Ahora se analiza la existencia de impactos heterogéneos del acceso a agua y saneamiento sobre los indicadores laborales.

Inicialmente se investiga heterogeneidades a partir del acceso a agua de red en las localidades en el año 2007. Intuitivamente, dado que las personas en localidades donde la mayor proporción de hogares carecía de acceso a agua en el período pre-tratamiento son las que más se beneficiaron del PNSR, entonces deberían ser también las que experimentan los mayores impactos del programa⁴. El panel a) de la **Tabla 8** presenta estos resultados. La clasificación de las localidades

⁴ El argumento se construye porque en el estudio se evalúa principalmente el impacto del acceso a y no necesariamente el de calidad. Según conversaciones con personas del PNSR, hubo hogares en localidades que en el período pre-tratamiento tenían agua entubada en la vivienda, pero que no era sometida a ningún tipo de tratamiento para su

fue realizada a partir de la información del Censo de Población 2007. Las localidades sin acceso son aquellas que en 2007 tenían hasta 5% de las viviendas con acceso a agua de red, mientras que las localidades con acceso son las que tenían una proporción de hogares con acceso a agua mayor a este valor. Como se aprecia, los impactos sobre la participación en el mercado laboral y la probabilidad de tener un trabajo son nulos, como anteriormente se había visto, pero sí se torna positivo y significativo sobre el número de horas semanales trabajadas entre hogares residentes en localidades que en 2007 carecían de acceso a agua. En el resto de localidades, el impacto parece ser nulo.

Seguidamente, se evalúa heterogeneidades en función del sexo de las personas. De acuerdo a diversa literatura (Ivens 2008; Sijbesma 2008; Sultana 2018), las intervenciones en agua y saneamiento tienen potenciales dimensiones de género al ser las mujeres quienes realizan mayormente las actividades domésticas vinculadas al acarreo de agua. Por tanto, serían estas las que experimenten los mayores beneficios del acceso a este recurso en el hogar. El panel b) de la **Tabla 8** es sugerente en este punto, pues si bien el parámetro asociado a las mujeres en la variable que recoge el número de horas diarias trabajadas es levemente mayor (0.754 versus 0.584 de los varones), no es estadísticamente diferente. En efecto, se implementó un test de Wald para evaluar la significancia de la diferencia entre los parámetros de varones versus mujeres (Cameron y Trivedi 2005: 7.2) y el *p-value* obtenido fue de 0.358, como se muestra en la tabla.

En el panel c) de la tabla se explora más este punto, pero esta vez se analizan heterogeneidades en función del rol de las personas al interior del hogar, en la línea de Zhou y Turvey (2017). Para esto, se han considerado cuatro roles: padre, madre, hijo e hija. Los resultados de este ejercicio sugieren que el impacto entre padres es igual independientemente del sexo, pero resulta mayor entre las hijas. En los hijos varones el impacto es nulo. Este resultado valida las implicancias de género de este tipo de infraestructura.

En el panel d) se consideran impacto a partir del tamaño de los hogares. Intuitivamente, en hogares más grandes las labores de acarreo de agua pueden distribuirse entre más miembros, con lo cual los tiempos por persona destinadas a esta actividad son menores que en hogares más pequeños. Por tanto, los impactos potenciales serían mayores en estos últimos. El panel d) de la **Tabla 8** valida esta hipótesis al mostrar que los impactos en hogares con hasta 3 miembros son positivos y significativos, tanto en la participación laboral, como en la probabilidad de trabajar y en el número de horas trabajadas.

Por último, en el panel c) se exploran heterogeneidades distributivas en función de los niveles iniciales de pobreza de los distritos. Para ello, se obtuvo las tasas de pobreza distrital estimadas por el INEI del año 2009 y con ellas se clasificó a la muestra en dos grupos: por debajo de la mediana (distrito menos pobre) y por encima de la mediana (distrito más pobre). Los resultados sugieren que aun cuando los parámetros estimados de personas en distritos más pobres son puntualmente mayores que los de personas en distritos menos pobres, estas diferencias no son estadísticamente significativas.

consumo seguro. Por tanto, cuando el programa intervino sobre estos hogares generó impactos en términos de calidad, principalmente, pero no de acceso porque ya contaban con este recurso en la vivienda con antelación.

Tabla 8. Impactos heterogéneos

	Participación en el mercado laboral	Probabilidad de Trabajar	# Horas trabajadas diarias
	(1)	(2)	(3)
a) Según acceso inicial a agua en 2007			
$T_{ij} \times$ Localidades sin acceso a agua [A]	0.029 (0.033)	0.040 (0.034)	0.780*** (0.229)
$T_{ij} \times$ Localidades con acceso a agua [B]	0.047 (0.038)	0.062 (0.039)	0.379 (0.257)
R2	0.155	0.159	0.118
P-value [A]=[B]	0.385	0.298	0.006
b) Según sexo de los informantes			
$T_{ij} \times$ Varones [A]	0.048 (0.035)	0.059 (0.036)	0.584** (0.238)
$T_{ij} \times$ Mujeres [B]	0.019 (0.036)	0.033 (0.037)	0.754*** (0.256)
R2	0.155	0.159	0.119
P-value [A]=[B]	0.272	0.331	0.358
c) Según rol en el hogar			
$T_{ij} \times$ Padre [A]	0.051 (0.036)	0.060 (0.038)	0.673** (0.292)
$T_{ij} \times$ Hijo [B]	-0.003 (0.050)	0.018 (0.054)	-0.039 (0.325)
$T_{ij} \times$ Madre [C]	0.011 (0.042)	0.027 (0.043)	0.675** (0.314)
$T_{ij} \times$ Hija [D]	0.032 (0.061)	0.041 (0.063)	0.933** (0.407)
R2	0.153	0.157	0.126
P-value [A]=[C]	0.184	0.289	0.992
P-value [B]=[D]	0.538	0.692	0.003
d) Según tamaño del hogar			
$T_{ij} \times$ Hogar con hasta 3 miembros [A]	0.078** (0.040)	0.091** (0.041)	0.837*** (0.318)
$T_{ij} \times$ Hogar con más de 3 miembros [B]	0.015 (0.038)	0.027 (0.040)	0.565** (0.283)
R2	0.154	0.159	0.118
P-value [A]=[B]	0.001	0.001	0.102
e) Según nivel inicial de pobreza distrital			
$T_{ij} \times$ Distrito menos pobre [A]	0.017 (0.039)	0.031 (0.040)	0.629** (0.298)
$T_{ij} \times$ Distrito más pobre [B]	0.047 (0.039)	0.058 (0.041)	0.676** (0.310)
R2	0.155	0.159	0.119
P-value [A]=[B]	0.155	0.198	0.834
Obs	6,906	6,908	5,564

Nota: Se utiliza una función polinómica de tercer grado. Los controles de hogar incluyen la edad, los años de educación, el sexo, el sexo al cuadrado y la condición de migrante del jefe de hogar; una variable binaria que indica si el hogar tiene acceso a electricidad y una variable binaria que indica si la vivienda del hogar tiene un piso diferente de tierra. Los controles geográficos incluyen la longitud, latitud y altitud del centro poblado de residencia de los hogares. En paréntesis se presentan errores estándar robustos, clustereados a nivel de centro poblados. *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia

Como ejercicio final, en la siguiente tabla se descomponen los impactos en las horas trabajadas según sector de ocupación: agrícola y no agrícola. Cada indicador representa la totalidad de horas trabajadas en cada sector económico considerando todas las actividades realizadas por la persona, tanto en la actividad primaria como en la actividad secundaria si la tuviese, así como en

todas las ocupaciones que realice en condición de dependiente de algún empleador y/o por cuenta propia. Los resultados son sugerentes, pues indican que el mayor tiempo liberado es utilizado para ocuparse en actividades agrícolas. El impacto en actividades no agrícolas es nulo, lo cual es consistente con los hallazgos de Zhou y Turvey (2017).

Al descomponer estos impactos entre varones y mujeres, los resultados muestran pequeñas diferencias en los valores puntuales, que resultan levemente más grandes para los primeros, pero que no llegan a ser estadísticamente diferentes. Así, desde la mirada agregada, este resultado es consistente con el panel b) de la tabla anterior.

Tabla 9. Impactos según Composición de las horas trabajadas

	Actividades agrícolas (1)	Actividades no agrícolas (2)
Total		
T_{ij}	0.818** (0.360)	-0.157 (0.330)
Obs	5,520	5,520
R2	0.147	0.100
Según sexo		
$T_{ij} \times \text{Varón [A]}$	0.873** (0.359)	-0.284 (0.329)
$T_{ij} \times \text{Mujer [B]}$	0.742* (0.410)	0.018 (0.390)
Obs	5,520	5,520
R2	0.147	0.100
P-value [A]=[B]	0.599	0.253
Promedio de var. dependiente	4.364	1.347

Nota: Se utiliza una función polinómica de tercer grado. Todas las regresiones incluyen efectos fijos regionales y controles de hogar como la edad, los años de educación, el sexo, el sexo al cuadrado y la condición de migrante del jefe de hogar; una variable binaria que indica si el hogar tiene acceso a electricidad y *dummies* que permiten identificar si la observación procede de la sierra o selva. También se incluyen controles geográficos, como la longitud, latitud y altitud del centro poblado de residencia de los hogares. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población del centro poblados. *,** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia

5 Conclusiones

En este estudio se ha hecho una evaluación de impacto del Programa Nacional de Saneamiento Rural sobre indicadores laborales de la población beneficiaria. Por la naturaleza de la intervención, que entrega agua potable y saneamiento hasta las viviendas de los hogares, el PNSR tiene el potencial de liberar recursos de tiempo que los individuos pueden trasladar al trabajo productivo. A pesar de que esta hipótesis es sugerente, las investigaciones que intentan validarla no son numerosos. Por ello, este estudio brinda nueva evidencia que contribuye con esta literatura.

Los resultados encontrados son sugerentes y van en línea de esta hipótesis. Primero, el PSNR habría logrado incrementar el acceso a agua potable y saneamiento en los hogares de localidades beneficiarias. De acuerdo a las estimaciones, como consecuencia del programa el PNSR habría incrementado el acceso a agua a los hogares de localidades beneficiarias en alrededor de 35 puntos porcentuales, equivalentes a casi 90% con respecto al valor estimado del grupo de control. Asimismo, habría logrado incrementar el acceso a saneamiento básico, que incluye servicios

sanitarios conectados a red pública y letrinas, en 17 puntos porcentuales, equivalentes a 100% del valor base del grupo de control.

Segundo, se ha encontrado evidencia de que el PNSR habría impactado en el margen intensivo de la oferta laboral, al incrementar las horas trabadas en cerca de entre 33 y 40 minutos diarios. Sin embargo, en promedio los impactos en el margen extensivo serían nulos, pues no habría logrado incrementar la oferta de mano de obra en el mercado laboral ni la probabilidad de trabajar. Las estimaciones han sido sometidas a diversas pruebas de robustez y falsificación y los resultados se mantienen en todos los casos, lo cual brinda credibilidad a las magnitudes estimadas. Estos estimados son levemente mayores al tiempo que destina una persona de 16 años a más al transporte de agua desde la fuente hasta la vivienda. Por tanto, el acceso a agua y saneamiento básico domiciliarios estaría permitiendo liberar no sólo el tiempo de acarreo del agua, sino también el de otras actividades domésticas relacionadas a la higiene, aseo y limpieza fuera del hogar (como el lavado de ropa, que se suele realizar, por ejemplo, en la fuente de agua), entre otros, para destinarlos al trabajo productivo.

Tercero, se ha realizado un conjunto de ejercicios que buscaron identificar impactos heterogéneos. Los hallazgos muestran que el PNSR habría tenido mayor efectividad en localidades que en el período pre-tratamiento carecían de acceso a agua. Además, aun cuando intuitivamente el agua tiene impactos potenciales diferenciados de género, los resultados muestran que los impactos entre varones y mujeres serían iguales en promedio. Sin embargo, sí se encuentra evidencia de que los impactos en la oferta laboral fémina de las hijas es mayor que en el caso de los hijos varones. En estos últimos el impacto es nulo. También se ha encontrado que el impacto es mayor entre hogares pequeños, con hasta 3 miembros. En estos último, incluso, se visualizan impactos en el margen extensivo e intensivo.

Bibliografía

- Beltran, A. y J. Seinfeld (2005), Desnutrición Crónica Infantil en el Perú Un problema persistente. Centro de Investigación de la Universidad del Pacífico, documento de trabajo DD/09/14.
- Calonico, S.; M. Cattaneo and R. Titiunik (2015), "Optimal Data-Driven Regression Discontinuity Plots", *Journal of the American Statistical Association* 110(512): 1753-1769.
- Calzada, J. y S. Iranzo (2016), Neither Public nor Private: The Effects of Communal Water Provision on Child Health in Peru. University of Barcelona, mimeo.
- Currie, J. y B. Madria (1999), Health, Health Insurance and the Labor Market. En *Handbook of Labor Economics*, vol. 3, ch. 50.
- Devoto, F.; E. Duflo; P. Dupas W. Parienté y V. Pons (2012) "Happiness on the tap. Water adoption in urban Morocco", *American Economic Journal: Economic Policy*, vol. 4, pp.68-99.
- Fletcher, A. (2018) "More than Women and Men: A Framework for Gender and Intersectionality Research on Environmental Crisis and Conflict". En *Water Security Across the Gender Divide pp 35-58*
- Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for Life: The Impact of the Privatization of Water Services on Child Mortality. *Journal of Political Economy*, 113(1), 83-120. doi:10.1086/426041
- Hahn, J., van der Klaauw, W., Todd, P., 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69 (1), 201-209.
- Imbens, G., and Lemieux, T. (2008), "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142, 615-635.
- Ivens, S. (2008) "Does Increased Water Access Empower Women?". *Development*, 51, (63-67). Doi:10.1057/palgrave.development.1100458
- Kleibergen, F. y R. Paap (2007). "Generalized reduced Rank tests using singular value decomposition". *Journal of Econometrics* 133, 97-126.
- Kremer, M.; J. Leino; E. Miguel y A. Zwane (2011), "Spring cleaning: Rural water impacts, valuation, and property rights institutions". *Quarterly Journal of Economics*, 126, 145-205.
- McCrary (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test". *Journal of Econometrics*. 142 (2): 698-714. Angrist y Pischke 2008
- Meek, R. (2012), Water Works: The economic impact of water infrastructure. Harvard Kennedy School, Discussion Paper 12-35
- Mills, J. y O. Cumming (2016) *The impact of water, sanitation and hygiene on key health and social outcomes. Review of evidence.*
- Ministerio de Economía y Finanzas (2015), *Guía para la identificación, formulación y evaluación social de proyectos de inversión pública de servicios de saneamiento básico urbano a nivel de perfil.*
- Moreano, O. (2010), *Lineamientos para la elaboración de proyectos/perfiles en saneamiento básico rural con adaptación al cambio climático.* Organismo Panamericano de la Organización Mundial para la Salud
- Naciones Unidas (2015) Objetivos de desarrollo del milenio. Informe 2015.
- Organización Mundial de la Salud (2017), "Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines". Disponible en:

<http://apps.who.int/iris/bitstream/handle/10665/258617/9789241512893-eng.pdf;jsessionid=12578AC18BE307D5B262A90BD57817FB?sequence=1>

Organización Mundial de la Salud (2017), *Progres on drinking wáter, sanitation and hygiene*.

PNSR (2013), Guía de opciones técnicas para abastecimiento de agua potable y saneamiento para Centros Poblados del ámbito rural.

PNSR, 2013

Sijbesma, C.; J. Verhagen; R. Nanavaty y A. James (2008), "Impact of domestic water supply on gender and income: results from a participatory study in drought prone region in Gujarat, India". *Water policy*, 1-11

Smith, L. y L. Hahhad (2000), *Explaining child malnutrition in developing countries. A cross country analysis*. International Food por Policy Research Institute.

Staiger, Douglas and James H. Stock (1997). "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65, 557-586.

Stock, James H. and Motohiro Yogo (2005), "Testing for Weak Instruments in Linear IV Regression." Ch. 5 in J.H. Stock and D.W.K. Andrews (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Cambridge University Press.

Sultana, F. (2018) Gender and Water in a Changing Climate: Challenges and Opportunities. En *Water Security Across the Gender Divide pp 35-58*

UNICEF (United Nations Children's Fund) (1990), "Strategy for improved nutrition of children and women in developing countries". New York.

Waddington, H., B. Snilstveit, H. White y L. Fewtrell (2009), "Water Sanitation and hygiene interventions to combat childhood diarrhea in developing countries". 3ie Synthetic Review 001.

Zhou y Turkey (2017), "Drinking wáter and off-farm labor supply: between gender and within gender bias". *The australian journal of agricultural and resource economics*, 62, pp. 103-120.

Ziegelhofer , Z. (2012), Down with diarrhea: Using fuzzy regression discontinuity design to link communal water supply with health. Graduate Institute of International and Development Studies Working Paper, No. 05/2012

Anexo 1. Robustez de las estimaciones de acceso a agua y saneamiento básico

Tabla A1. Sensibilidad de las estimaciones a nivel de participación en el PNSR y Acceso a agua y saneamiento

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Beneficiario del PNSR								
D_{ij}	0.617***	0.625***	0.613***	0.622***	0.623***	0.632***	0.624***	0.631***
	(0.055)	(0.054)	(0.052)	(0.053)	(0.070)	(0.072)	(0.068)	(0.071)
Obs.	2,956	2,951	2,956	2,951	2,956	2,951	2,956	2,951
Acceso a agua de red								
Intención del Tratamiento								
D_{ij}	0.222***	0.189***	0.219***	0.186***	0.239***	0.219***	0.227***	0.210***
	(0.066)	(0.054)	(0.065)	(0.053)	(0.080)	(0.067)	(0.078)	(0.066)
Obs.	2,956	2,951	2,956	2,951	2,956	2,951	2,956	2,951
RDD Fuzzy								
T_{ij}	0.359***	0.302***	0.356***	0.300***	0.383***	0.347***	0.363***	0.333***
	(0.108)	(0.089)	(0.107)	(0.088)	(0.132)	(0.113)	(0.126)	(0.110)
Obs.								
F-Stata 1ra Etapa	2,956	2,951	2,956	2,951	2,956	2,951	2,956	2,951
Intención del Tratamiento	127.8	132.9	137.1	139.8	79.21	76.20	84.34	78.62
Acceso a saneamiento								
Intención del Tratamiento								
D_{ij}	0.122**	0.107**	0.123**	0.107**	0.097	0.109**	0.093	0.106*
	(0.048)	(0.043)	(0.048)	(0.043)	(0.062)	(0.055)	(0.062)	(0.055)
Obs.	2,956	2,951	2,956	2,951	2,956	2,951	2,956	2,951
RDD Fuzzy								
T_{ij}	0.198**	0.171**	0.200**	0.172**	0.156	0.173*	0.148	0.167*
	(0.081)	(0.073)	(0.081)	(0.073)	(0.102)	(0.094)	(0.100)	(0.093)
Obs.	2,956	2,951	2,956	2,951	2,956	2,951	2,956	2,951
F-Stata 1ra Etapa	127.8	132.9	137.1	139.8	79.21	76.20	84.34	78.62
Grado del polígono	Lineal		Cuadrática		Cúbica		Cuarta	
Controles	No	Si	No	Si	No	Si	No	si

Nota: Los controles incluyen efectos fijos regionales y controles de hogar, como la edad, los años de educación, el sexo y la condición de migrante del jefe de hogar; un índice de acceso a servicios en el hogar construido por componentes principales; una variable binaria que indica si el hogar tiene acceso a electricidad. También se incluyen controles geográficos, como la longitud, latitud y altitud del centro poblado de residencia de los hogares, así como *dummies* que identifican observaciones en la sierra y selva. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población por centro poblado. *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia

Tabla A2. Falsificación de las estimaciones a nivel de acceso a agua y saneamiento

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	c=80	c=120	c=160	c=200	c=240	c=280	c=320
Intención del tratamiento							
Acceso a agua de red							
D_{ij}	0.026	0.035	0.038	0.219***	-0.076	-0.166**	-0.172
	(0.078)	(0.067)	(0.068)	(0.067)	(0.077)	(0.080)	(0.115)
Obs.	2,951	2,951	2,951	2,951	2,951	2,951	2,951
Acceso a saneamiento básico							
D_{ij}	0.033	-0.004	0.014	0.109**	-0.067	-0.030	0.090
	(0.067)	(0.046)	(0.050)	(0.055)	(0.057)	(0.060)	(0.077)
Obs.	2,951	2,951	2,951	2,951	2,951	2,951	2,951
Fuzzy RDD							
Acceso a agua de red							
T_{ij}	-0.426	-0.195	-1.806	0.347***	0.378	0.923	1.785
	(1.416)	(0.388)	(7.353)	(0.113)	(0.387)	(0.641)	(2.269)
Obs.	2,951	2,951	2,951	2,951	2,951	2,951	2,951
F-Stata 1ra etapa	0.986	15.77	0.0874	76.20	4.830	3.931	0.728
Acceso a saneamiento básico							
T_{ij}	-0.540	0.023	-0.677	0.173*	0.331	0.165	-0.930
	(1.235)	(0.256)	(3.495)	(0.094)	(0.334)	(0.361)	(1.324)
Obs.	2,951	2,951	2,951	2,951	2,951	2,951	2,951
F-Stata 1ra etapa	0.986	15.77	0.0874	76.20	4.830	3.931	0.728

Nota: Se utilizan polígonos de tercer grado. Todas las regresiones incluyen efectos fijos regionales y controles de hogar, como la edad, los años de educación, el sexo y la condición de migrante del jefe de hogar; un índice de acceso a servicios en el hogar construido por componentes principales; una variable binaria que indica si el hogar tiene acceso a electricidad. También se incluyen controles geográficos, como la longitud, latitud y altitud del centro poblado de residencia de los hogares, así como *dummies* que identifican observaciones en la sierra y selva. En paréntesis se presentan errores estándar robustos, clustereados a nivel de tamaño de población por centro poblado. *, ** y *** indican, respectivamente, significancia al 1%, 5% y 10%. Fuente: ENAHO 2015, Censo de Población 2007, INEI. Elaboración propia