



Munich Personal RePEc Archive

Tax base and geopolitical discontinuities

Natividad, Gabriel

Universidad de Piura

2019

Online at <https://mpra.ub.uni-muenchen.de/113169/>

MPRA Paper No. 113169, posted 31 May 2022 14:19 UTC

“Base Tributaria y Discontinuidades Geopolíticas”¹

Gabriel Natividad²

Abstract

The goal of this study is to empirically analyze the extensive and intensive margin of taxation in different geographies of Peru using publicly accessed information provided by the tax administration for the period 2005 to 2018. Because internal geographic zones in Peru differ on a number of dimensions, the study proposes an empirical method that uses similarities between different areas while it exploits the internal geopolitical borders of Peruvian regions to study the impact of socioeconomic variables on tax-related behavior.

1. Introducción

El objetivo del presente estudio es analizar econométricamente el margen extensivo e intensivo de la tributación en distintas geografías del Perú utilizando información de acceso público no reservado proporcionada en parte por la Administración Tributaria para el período 2001 a 2018. Debido a que las geografías internas peruanas difieren en un amplio espectro de dimensiones, el estudio propone un método empírico que aprovecha las similitudes entre distintas áreas peruanas a la vez que explota los límites geopolíticos internos del Perú –los de los departamentos o regiones– para estudiar impactos de variables socioeconómicas departamentales en el comportamiento tributario.

La motivación central de este análisis reside en la importancia de considerar cómo ampliar la base tributaria y cómo mejorar el cumplimiento tributario del máximo número posible de agentes económicos en el Perú. En una sociedad en la que el costo de hacer negocios incluyera el contribuir al fisco mediante el pago del impuesto a la renta, esta motivación sería obvia y estaría casi mecánicamente ligada a la creación de empresas; el caso del Perú es todavía lejano a ese estándar de sociedad. Además, como este estudio muestra, el Perú es un país centralizado geográficamente, pero a la vez un país diverso y geográficamente vasto. La identificación de factores válidos que incidan en el comportamiento tributario de individuos y empresas puede ayudar a informar la adopción de medidas que promuevan el mejor desempeño fiscal y el crecimiento del país.

El problema de estudio se deriva directamente de la observación de diferencias cuantificables en el comportamiento tributario de distintas personas y entidades. El margen extensivo de la tributación consiste en cuántas de las entidades desempeñando funciones económicas se hallan formalizadas ante la Administración Tributaria. El margen intensivo de la tributación se refleja en el comportamiento de los agentes económicos en cumplir sus obligaciones ante el Tesoro; por ejemplo, en presentar sus declaraciones de impuesto a la renta. Por tanto, el problema de estudio consiste en entender cómo varían el margen extensivo e intensivo de la

¹Informe Final, con fecha 14 de mayo de 2019.

² Departamento de Economía, Universidad de Piura, Miraflores, Lima, gabriel.natividad@udep.edu.pe

tributación en el Perú dependiendo de la geografía que se analice, y cómo algunas variables socioeconómicas departamentales pueden ser las causantes de los distintos comportamientos observados.

Las preguntas de investigación son cuatro, y cada una de ellas corresponde a una variable socioeconómica considerada geográficamente, a saber: (1) ¿Cómo influye el ambiente de educación departamental en el margen extensivo e intensivo de la tributación?, y, análogamente, ¿cómo influye (2) la inversión pública departamental, (3) la percepción de corrupción expresada en el departamento, y (4) la confianza en la SUNAT por parte de la población en el departamento?

2. Marco teórico

El marco conceptual para el presente estudio enlaza dos ramas de la literatura económica: la consideración de qué es una base tributaria en contraste con la informalidad empresarial y el rol de la geografía en el desarrollo empresarial y tributación. Luego de invocar los conceptos centrales de estas ramas, el marco conceptual incluye una síntesis para el caso concreto del Perú, enfocándose en contribuciones recientes que pueden ayudar a entender la evolución empresarial peruana. El marco conceptual termina con una síntesis de un proyecto anterior sobre la mita colonial minera peruana y las actitudes de formalización empresarial.

2.1 Base tributaria e informalidad empresarial

En esta sección se ofrece una revisión de la literatura sobre la importancia de la base tributaria en conexión con el accionar empresarial y la existencia de informalidad. La ampliación sostenida de la base tributaria y el mejoramiento del comportamiento tributario de las empresas han sido vistos como componentes teóricos fundamentales para el gasto público y el crecimiento económico (Barro, 1991). En países en desarrollo, la base tributaria ha sido caracterizada como ineficientemente pequeña, en parte debido a una economía informal prevaleciente, con diversas consecuencias negativas para la institucionalidad y el crecimiento (de Soto, 1986; Loayza, 1996; La Porta y Shleifer, 2008; De Paula y Scheinkman, 2011; La Porta y Shleifer, 2014).

Las políticas tributarias crean impactos profundos en el comportamiento de empresas (Chetty y Saez, 2005; De Paula y Scheinkman, 2010; Djankov, Ganser, McLiesh, Ramalho, y Shleifer, 2010; Baugh, Ben-David, y Park, 2016; Deng, Liu, y Wei, 2018; Zidar, Serrato, Morales, y Fajgelbaum, 2018) y de individuos (Agarwal, Liu, y Souleles, 2007). Por ejemplo, debido a regulaciones tributarias, las empresas pueden tener el incentivo a crecer poco (Zidar y cols., 2018; Garicano, Lelarge, y Van Reenen, 2016). En el contexto más amplio de los determinantes legales del crecimiento (e.g., Demircuc-Kunt y Maksimovic (1998)), el estudio de las empresas y su comportamiento tributario es fundamental para entender el desarrollo económico.

Por su parte, las consideraciones teóricas sobre la informalidad han evolucionado en el tiempo. Un punto de quiebre en las consideraciones teóricas fue la racionalización de la informalidad. Por ejemplo, las implicancias de los argumentos de de Soto (1986) sobre el valor hallado por los emprendedores en la informalidad han sido múltiples (Karas, Pyle, y Schoors, 2015), incluyendo su propia obra posterior, que enfatizó el rol del capital financiero (de Soto, 2000). La literatura ha explicado la decisión de operar sin formalización tributaria como consecuencia de los elevados costos de entrada o la baja productividad de los emprendedores que desean ingresar (de Soto, 1986; La Porta y Shleifer, 2008, 2014; Farrell, 2004; Ulyseas, 2018), con diferencias de énfasis en la reducción de costos de formalización o

la supervisión estatal (Monteiro y Assuncao, 2012; Bruhn, 2011; Woodruff, McKenzie, y de Mel, 2013; Rocha, Ulyseas, y Rachter, 2018; Meghir, Narita, y Robin, 2015; Ulyseas, 2010).

Diversos estudios internacionales han documentado la existencia de relaciones informales entre individuos y empresas, incluso sin el interés en evadir impuestos (McMillan y Woodruff, 1999; Garmaise y Moskowitz, 2003; Guirkinger, 2008; Ayyagari, Demirguc-Kunt, y Maksimovic, 2010). Esas relaciones informales son parte de un capital social que fomenta el desarrollo económico (Guiso, Sapienza, y Zingales, 2004).

Así como ha habido argumentos benévolos sobre el valor de la informalidad, también existen argumentos críticos en cuanto a la ilegalidad de no entrar a la base tributaria, o evadir impuestos si ya se forma parte de ella. Evitar al fisco ha sido parte de la lógica de la evasión y operación informal (Friedman, Johnson, Kaufmann, y Zoido-Lobaton, 2000; Slemrod y Yitzhaki, 2002; Dyreng, Hanlon, y Maydew, 2008; Scott, 2010; Beck, Lin, y Ma, 2014; Casaburi y Troiano, 2015; Artavanis, Morse, y Tsoutsoura, 2016) o de acciones específicamente diseñadas para defraudar (Marion y Muehlegger, 2008).

Independientemente del tono positivo o negativo en la discusión sobre la actividad informal entre empresarios y trabajadores, es ampliamente aceptado que sin un aumento de la base tributaria, es improbable que se pueda desarrollar adecuadamente una economía nacional. A las dificultades sobre recaudación se agrega una dimensión también crítica de la vida de un país: su dispersión geográfica.

2.2 El rol de la geografía

Una dimensión importante pero poco explorada en estudios peruanos sobre la formalización y la expansión de la base tributaria es la geografía interna del país (ver Escobal y Armas (2018) como una excepción). En general, el estudio de cómo y por qué difiere el comportamiento tributario en distintas geografías peruanas es una cuestión abierta. Estudios sobre la economía geográfica peruana han resaltado factores que condicionan el desarrollo (Escobal y Torero, 2003) y el crecimiento empresarial (Natividad, 2018a). Estos estudios constituyen un precedente útil para profundizar más en la relación entre geografía y tributación, al concluir que la geografía es una dimensión importante en el comportamiento tributario en algunas zonas del Perú.

De la literatura internacional se concluye que los condicionantes geográficos son extremadamente importantes para el crecimiento empresarial (Sorenson y Audia, 2000; Stuart y Sorenson, 2003a, 2003b; Greenstone, Hornbeck, y Moretti, 2010). Los mercados geográficos pueden ejercer distintos tipos de poder monopólico local (Hortacsu y Syverson, 2007), alejándose de los principios de competencia perfecta (Makowski y Ostroy, 2001). La falta de infraestructura pública de transporte imposibilita un mayor crecimiento de la productividad en servicios (Lagakos, 2016). La geografía también ha sido estudiada en relación con la tributación. Por ejemplo, ver Bartik (1985); Besley y Persson (2013); Bruce (2002); Bruce y Deskins (2012); Bruce y Mohsin (2006); Durantón, Gobillon, y Overman (2011); Georgellis y Wall (2006); Giroud y Rauh (2016); Leal-Ordóñez (2014); Wasylenko (1999); Moretti y Wilson (2017).

En específico, una amplia gama de estudios anteriores ha planteado que las tendencias a la formalización y al mejor comportamiento tributario pueden variar significativamente entre regiones geográficas, incluso si tales regiones se encuentran sujetas al mismo régimen tributario (Tafenu, Herwartz, y Schneider, 2010; Williams y Round, 2010; Chetty, Friedman, y Saez, 2013).

Pero la geografía captura factores subyacentes debidos a las diferencias geopolíticas entre territorios. Desde el punto de vista conceptual, si se controla por aspectos meramente geoespaciales, las diferencias entre territorios son generadas por decisiones humanas más que por un determinismo geográfico. En el presente estudio, se da especial relieve a factores departamentales tales como la educación, la inversión pública, la percepción de corrupción, y la confianza en la autoridad tributaria.

La educación suele concebirse teóricamente como un factor que conduce a mayores ingresos individuales (Card, 2001; Wantchekon, Klanja, y Novta, 2015). Sin embargo, desde el punto de vista de la tributación, ha habido hipótesis contrapuestas en lo referente al nivel de educación (Dee, 2004; Milligan, Moretti, y Oreopoulos, 2004; Buehn y Farzanegan, 2013; van der Sluis, van Praag, y Vijverberg, 2004; Maloney, 2004). Por un lado, se podría ver al empresario informal como alguien pobre y sin educación suficiente para entender las leyes que rigen la actividad económica y la tributación. Por otro lado, se concibe al emprendedor informal como alguien sagaz que rechaza la formalidad mal diseñada, abriéndose paso en el mercado con sus propias reglas informales, sin deseo de tributar al sistema formal.

La literatura ha considerado a la educación como un posible factor que lleva a la formalización empresarial. Una relación positiva entre educación y formalización indica mayor cumplimiento de deberes cívicos. Sin embargo, la relación contraria también puede ser válida si la mayor educación empeora la percepción sobre las instituciones estatales (Dee, 2004; Milligan, Moretti, y Oreopoulos, 2004; Buehn y Farzanegan, 2013). La preferencia por trabajar independientemente sumada a la relación positiva entre la educación y la propensión a emprender negocios también puede predecir una relación negativa entre el nivel de educación y la formalización de empresas (van der Sluis, van Praag, y Vijverberg, 2004; Maloney, 2004).

Por su parte, la corrupción es vista como un problema complejo, que afecta tanto a países en desarrollo (Olken y Pande, 2012) como desarrollados (Butler, Fauver, y Mortal, 2009; Do y Campante, 2014). La corrupción en distintos estamentos del Estado puede actuar en detrimento de una cultura de tributación. Las marcadas tendencias a la corrupción estatal en países en desarrollo incluso pueden exportarse a otras realidades geográficas (Fisman y Miguel, 2007; DeBacker, Heim, y Tran, 2015), o puede trascender en el tiempo mediante mecanismos de persistencia histórica (Guardado, 2018).

La inversión pública efectiva realizada en economías locales ofrece un matiz distinto aunque relacionado al de la percepción sobre corrupción. El nivel de inversión pública es un factor para la toma de decisiones tanto del gobierno central como de individuos en poblaciones locales (Goolsbee y Guryan, 2006; Levin y Tadelis, 2010), muchas veces con el objetivo de reducir la desigualdad social (Goñi, Lopez, y Servén, 2008). Existen también factores de ineficiencia que hacen que el gasto del Estado no sea satisfactorio localmente (Duggan, 2000). En particular, si los fondos recaudados por el Estado se diluyen debido a la corrupción o la ineficiencia, es posible que la inversión pública no mejore (Brollo, Nannicini, Perotti, y Tabellini, 2013; Caselli y Michaels, 2013; Hong, 2017).

Finalmente, la confianza en las instituciones juega un rol en la decisión de formalización. Los cálculos de posibles ganancias en el sector informal y las percepciones personales sobre el valor de formalizarse juegan un rol en la decisión de formalizar la empresa y entrar a la base tributaria (Hartner, Rechberger, Kirchler, y Schabmann, 2008; Kirchler, Hoelzl, y Wahl, 2008; Cummings, Martinez-Vazquez, McKee, y Torgler, 2009; Besley y Persson, 2014; Gadenne, 2017; C. C. Williams, Martinez-Perez, y Kedir, 2017). Las empresas toman decisiones de ahorro, inversión, endeudamiento y reparto de dividendos dependiendo de su

tratamiento tributario (Miller, 1977; Allen, Bernardo, y Welch, 2000; Chirinko y Wilson, 2008; Graham, 2000; Mortal y Reisel, 2013; Bennedsen y Zeume, 2017). Conforme avanza su ciclo de vida, las empresas descubren aquello en que son mejores y lanzan distintos productos y servicios al mercado (Bernardo y Chowdhry, 2002), aunque a veces les falta el capital financiero suficiente para crecer (Black y Strahan, 2002; Cabral y Mata, 2003; Beck, Demirguc-Kunt, y Maksimovic, 2005). Un factor determinante en ese crecimiento es la capacidad de sus directivos (Bloom y Van Reenen, 2010). En parte, el pensamiento directivo sobre el valor de entrar al sector formal determina la base tributaria efectiva de una localidad.

Como explica Natividad (2018a), existen muchas formas en que las empresas o individuos pueden estar mejor predispuestos a aprovechar las mejoras implementadas por otros para beneficiarse ellos mismos (Sorenson y Audia, 2000; Stuart y Sorenson, 2003a, 2003b; Greenstone, Hornbeck, y Moretti, 2010). Prominentemente, la proximidad geográfica es un factor bien documentado en la literatura sociológica y económica. Pero en muchos casos la geografía no es suficiente: hace falta un ecosistema de geografías industriales y ambientes conducentes a mayor transmisión de conocimiento dentro de la empresa. Hay externalidades positivas generadas gracias a la diseminación de información solamente dentro de una red financiera, cuestión que ha sido estudiada desde el punto de vista económico encontrando evidencia causal (Garmaise y Natividad 2016). En general, hay gran heterogeneidad en cómo las empresas pueden beneficiarse de factores geográficos en su crecimiento.

2.3 Marco conceptual y la realidad empresarial peruana

Desde el trabajo seminal de de Soto (1986), ha existido gran interés teórico y práctico en entender las causas de la informalidad empresarial en el Perú (Hernandez y de la Roca, 2006; Jackle y Li, 2006). Un argumento frecuente sobre la reducida base tributaria peruana es que las empresas muestran todavía muy baja productividad. En este sentido, el estudio de la productividad peruana ha tenido un rango amplio de contribuciones (Pagés, 2010; Céspedes, Aquije, Sánchez, y Vera-Tudela, 2014; Natividad, 2014), alguna de ellas incluso privilegiada con acceso a información confidencial de la SUNAT, pero pocas han versado específicamente sobre su relación con el comportamiento tributario.

En este sentido, una consideración pausada del panorama empresarial peruano muestra que éste depende en gran medida del funcionamiento de la banca comercial. Por un lado, los bancos y entidades financieras están sujetas a shocks positivos y negativos (Garmaise y Natividad, 2010; Schnabl, 2012; Garmaise y Natividad, 2013; Paravisini, Rappoport, Schnabl, y Wolfenzon, 2015). Los clientes bancarios, tanto de créditos de consumo como de créditos empresariales, son frágiles y sufren las consecuencias de shocks relativamente pequeños (Garmaise y Natividad, 2016, 2017). El rol del conocimiento personal en asegurar la devolución de los créditos es particularmente importante en lugares o segmentos aislados en donde la banca comercial no llega con preponderancia y deja paso a otras instituciones financieras enfocadas en esos clientes (Karlan, 2007; Frisancho, 2012; Aguilar, 2014).

Un factor estudiado en detalle en el caso peruano ha sido sobre cómo la diseminación de información sobre la situación financiera de una empresa puede beneficiar a empresas e individuos del mismo entorno geográfico. Garmaise y Natividad (2016) es el estudio que aborda directamente este tema (ver un caso indirecto en Frisancho, 2012). Dicho estudio considera esta pregunta en el contexto del Perú, estudiando el efecto de una política regulatoria y un shock crediticio a los negocios vecinos de las empresas directamente afectadas. Se muestra así que un shock que genera mayor información sobre una empresa y un mayor acceso al crédito para la empresa también conduce a un aumento de la financiación

para sus vecinos. Ante la pregunta sobre si el aumento de crédito de un vecino se deriva de la nueva información sobre la empresa y el vecindario, o es el crédito incrementado simplemente una consecuencia del nuevo préstamo a la empresa, Garmaise y Natividad (2016) proponen una serie de mecanismos para proporcionar evidencia de que la mayor provisión de préstamos a los vecinos se debe a los efectos de la información, más que a un efecto financiero.

Análisis como el de Garmaise y Natividad (2016) sugieren que las políticas de fortalecimiento de “firmas ancla” para una geografía dada, en un intento por promover su financiamiento y, por lo tanto, el financiamiento de otras firmas de la misma geografía, pueden tener cierto éxito, si como parte de este proceso también se alienta a los prestamistas a recopilar más información sobre esas empresas ancla. Los hallazgos indican la necesidad de una comprensión más compleja de los factores que influyen en los flujos de crédito locales.

En el Perú, como en otros países, los shocks negativos de créditos están asociados con serias restricciones para el acceso al crédito en el mediano plazo, incluso cuando el mecanismo no es la información. Garmaise y Natividad (2017) analizan los determinantes de esta relación, explotando las características de las regulaciones bancarias locales para identificar shocks exógenos en las clasificaciones de riesgo de algunos prestatarios. Estos shocks no tienen contenido de información, pero conducen a una reducción en el financiamiento para los consumidores afectados durante tres años. Es decir, los consumidores que experimentan una baja de calificación crediticia debido simplemente a la mala suerte están sujetos a un período prolongado de acceso financiero reducido.

Este aspecto del contexto peruano es importante. Como muestran Garmaise y Natividad (2017), los bancos en principio tienen acceso a toda la información necesaria para desentrañar la fuente aleatoria de shocks, y aun así se observa un impacto negativo en los préstamos al prestatario. Esto se relaciona con los hallazgos en otros contextos de que los eventos de calificación sin contenido de información pueden tener un impacto en los resultados financieros cuando las regulaciones están vinculadas a las calificaciones crediticias. De manera más general, las regulaciones que vinculan el historial de crédito con el estado general de la macroeconomía pueden tener un lugar en el abanico de opciones disponibles para los responsables de las políticas luego de una recesión.

La amplia geografía peruana, sujeta a shocks naturales de distintos tipos, permite cierta flexibilidad a las empresas en cómo afrontan sus obligaciones y cómo acceden a fondos para sus inversiones. Garmaise y Natividad (2018) exploran en detalle la dimensión geográfica de la flexibilidad. Lograr la flexibilidad financiera es un objetivo clave de política empresarial. La evidencia y la teoría previa han enfatizado la importancia para las corporaciones de tener acceso al crédito con facilidad. Un aspecto de la opcionalidad en las decisiones financieras es la capacidad de las empresas que operan en múltiples áreas para enfocar estratégicamente sus préstamos en mercados locales prósperos. Garmaise y Natividad (2018) demuestran que las empresas, de hecho, optan por pedir prestado en sus lugares de la geografía peruana más prósperos, donde el crédito es abundante y relativamente barato. Sin embargo, también encuentran que la búsqueda de flexibilidad financiera geográfica puede tener un costo. Las empresas que inician los préstamos en nuevos mercados florecientes descuidan sus relaciones bancarias existentes y, como resultado, experimentan mayores riesgos de eventuales dificultades financieras. Por lo tanto, afirman Garmaise y Natividad (2018), debe entenderse que la agilidad de una empresa para acceder a financiamiento en diferentes mercados en

términos atractivos conlleva la consecuencia de degradar conexiones importantes a su actual conjunto de prestamistas.

La contribución de Garmaise y Natividad (2018) complementa investigaciones previas que describen el impacto de los shocks locales en las actividades crediticias de los bancos. A diferencia de esos trabajos previos, Garmaise y Natividad (2018) se centran en las empresas en lugar de los bancos. Documentan un efecto diferente: los estudios bancarios muestran que un shock local positivo lleva a los bancos a prestar más en otras áreas, mientras que el estudio de flexibilidad geográfica peruana muestra que estos shocks llevan a las empresas de ubicaciones múltiples a pedir menos préstamos en otros lugares. En cierto sentido, estas dos fuerzas están en tensión y tienen efectos opuestos en la ubicación de los préstamos. La interacción del banco con los préstamos y las decisiones de las empresas en diferentes geografías pueden ser bastante complejos.

Existen también shocks regulatorios o naturales en el Perú que obligan a las empresas a ser muy cuidadosas en el manejo de su capital. Natividad (2013) propone el caso del capital interno de las empresas productoras de productos basados en recursos marinos. En ese estudio, proporciona un análisis causal de los mercados internos de capital en las empresas de la industria pesquera peruana, la segunda más grande del mundo, con US\$ 2,500 millones en exportaciones anuales. La información granular sobre las operaciones de producción de todas las plantas procesadoras de pescado permite la observación directa de inversiones a corto plazo en capital de trabajo y procesos de producción. Además, los datos detallados sobre insumos, productos, geografía y opciones tecnológicas de las empresas procesadoras de pescado permiten una medición directa de la productividad divisional.

Natividad (2013) muestra que la mayor inversión en el segmento de consumo humano directo es debida a las prohibiciones de la producción de harina de pescado, incluso después de controlar por la productividad. Además, encuentra que la redistribución de la capacidad financiera en inversiones no relacionadas con la harina de pescado es particularmente más aguda cuando el financiamiento externo es más costoso y cuando las firmas son monitoreadas más de cerca por sus acreedores, sugiriendo que el financiamiento interno se administra más activamente cuando es más escaso. Una interpretación de la creación de valor de esta actividad intensificada de los mercados internos de capital es apoyada por su efecto positivo en las exportaciones. Los resultados, basados en el universo de empresas industriales del sector pesquero con las salvedades del caso, sugieren que los mercados de capital internos son propicios para la creación de valor.

La dimensión geográfica de la explotación de recursos naturales peruanos y su impacto en la productividad ha sido estudiada de diversas formas (e.g., Aragón y Rud (2013)). En particular, el estudio es potenciado con datos recolectados en tiempo real y alta frecuencia. Natividad (2014) ofrece un estudio de operaciones pesqueras. Tres temas recientemente documentados en la literatura plantean grandes retos para la investigación sobre los bordes operativos de la firma. Primero, medir el impacto de una dimensión de límites de la firma (e.g., alcance horizontal) sin considerar otras dimensiones (e.g., alcance vertical) puede conducir a conclusiones incorrectas acerca de cómo los límites de la firma impactan al desempeño. Segundo, los cambios en los límites de la firma están motivados por factores típicamente no observados externamente, evocando así preocupaciones de endogeneidad. En tercer lugar, la propiedad de activos se estudia de manera desconectada de la integración de operaciones en distintos puntos de la empresa. Natividad (2014) aborda estos desafíos empleando datos granulares en operaciones de producción y la variación plausiblemente exógena en los límites de la firma causada por adquisiciones corporativas para evaluar el

impacto causal de la integración en la productividad total de los factores (PTF). Los resultados indican un efecto positivo, estadísticamente significativo y económicamente útil de integración en la PTF.

Natividad (2014) también introduce los sistemas de información geográfica (SIG, o GIS en inglés) al estudio de la economía de la empresa peruana. Los trabajos empíricos que explotan un SIG están pasando rápidamente de áreas especializadas a la literatura más amplia, pero mucho de esta investigación se ha centrado en el corte transversal o ajustes de baja frecuencia. El estudio de la pesca en tiempo real ofrece algo nuevo al adelantar el microcosmos de límites de la firma como un área de amplio interés práctico en el que un SIG puede ayudar a abordar cuestiones fundamentales sobre la operación empresarial.

La geografía peruana también permite la coexistencia de distintos marcos regulatorios para empresas. Natividad (2016) compara el Norte y el Sur del litoral peruano al momento de un cambio en las reglas de juego de la extracción pesquera con las cuotas individuales de pesca de anchoveta. Las características clave de esta reforma regulatoria permiten estudiar el impacto causal de las cuotas. Sin duda, una reforma de cuotas es un evento, entre otros, que puede afectar la estructura de la industria, por lo que generalmente es difícil aislar el efecto causal de las cuotas sin un claro contrafactual. Natividad (2016) estudia la reforma explotando como contrafactual las operaciones de pesca de anchoveta en la región sur del país, donde se estableció el régimen de cuotas con un retraso con respecto al resto del Perú.

Natividad (2016) encuentra que varias características económicas a lo largo de las geografías en el litoral peruano parecen afectar de manera diferente los precios del mercado local después del régimen de cuotas. Si bien estos factores ayudan a sugerir un papel matizado de la heterogeneidad del mercado en la formación de precios, la reducción en el suministro diario en toda la industria causada por las cuotas de pesca individuales aparece con el peso más significativo para explicar el gran aumento de los precios en todos los mercados al implementarse el régimen. Si bien el análisis reseñado se basa en el caso del sector pesquero, la reglamentación activa del Estado peruano sugiere que no existe un sesgo particular que beneficie a algún grupo de empresas en contraste con otras.

Las empresas se acostumbran a una forma de operar, y esa inercia puede ser difícil de cambiar en el corto plazo. En el contexto peruano, Natividad y Rawley (2016) proveen evidencia econométrica sobre este mecanismo de adaptación. Bajo las condiciones adecuadas, las empresas pueden coordinar las actividades económicas de manera más eficiente que los mercados, pero también se cree que el control jerárquico crea interdependencias negativas, lo que dificulta el desempeño de las empresas con actividades múltiples. Los investigadores de estudios pasados han demostrado que las empresas pueden exhibir características de planificación central eficiente, por ejemplo, en la literatura sobre la economía de los costos de transacción, pero también que las empresas de amplio alcance operan como burocracias ineficientes. Sin embargo, a pesar de la naturaleza dual del alcance empresarial, hay poca investigación que explique cómo esas características aparentemente opuestas operan dentro de la misma empresa simultáneamente. Natividad y Rawley (2016) caracterizan las ventajas y desventajas inherentes a las decisiones de alcance empresarial mediante el estudio de las implicaciones de la interdependencia en el nivel de actividad dentro de las empresas.

Para ese análisis de la inercia de las empresas peruanas, la ventaja metodológica clave es aprovechar una prohibición imprevista de la pesca de caballa que modificó exógenamente el alcance de las actividades realizadas por las empresas pesqueras, lo que permite identificar el impacto de la interdependencia en la productividad a nivel de actividad (pesca de anchoveta).

Específicamente, Natividad y Rawley (2016) comparan los cambios en la productividad en el nivel de actividad en las empresas anteriormente con múltiples actividades (es decir, las empresas que pescaban caballa y anchoveta) en relación con los cambios en la productividad en las empresas de actividad única (es decir, las empresas que solo pescaban anchoveta históricamente). Si bien el análisis reseñado se basa en el caso del sector pesquero, puede ofrecer patrones útiles por explorar en industrias extractivas peruanas.

2.4 Antecedente sobre geografía y tributación: el impacto de la mita colonial minera

Un antecedente cercano al presente estudio es Natividad (2019), que se enfoca en los impactos de largo plazo de la mita colonial minera. Entre 1573 y 1812, los colonizadores españoles del Perú sometieron a las poblaciones masculinas locales en comunidades peruanas específicas a la mita, un régimen de trabajo forzoso que requería el despliegue obligatorio de adultos varones desde sus lugares de origen a minas recién descubiertas en el sur de Perú y Bolivia. Las autoridades coloniales determinaron los límites geográficos de la mita. Las comunidades locales ubicadas dentro de los límites asignados estaban sujetas al régimen de mita, mientras que las comunidades ubicadas fuera de los límites estaban exentas.

Natividad (2019) propone que varias características de la mita colonial alteraron radicalmente las economías locales en formas que estaban conectadas plausiblemente con sus capacidades de producción. Al mover a los varones adultos de sus lugares de origen para trabajar en ubicaciones mineras remotas y al forzar la liquidación de los activos físicos de aquellos sujetos tributarios que no estaban dispuestos a servir en las minas, la mita distorsionó gravemente los niveles de equilibrio de capital y trabajo desde el punto de vista de la función de producción.

En un diseño de regresión que modela la discontinuidad multidimensional en latitud y longitud a través de las fronteras de la mita colonial peruana, Natividad (2019) evalúa los impactos a largo plazo de la mita en las empresas dentro de la región de mita comparándolas con las empresas que se encuentran justamente afuera de los límites de la mita. Aprovechando la geocodificación de las empresas en el nivel de manzanas de censo para la configuración de discontinuidad, los modelos principales se centran en las empresas en distritos adyacentes al límite de mita e incluyen numerosos efectos fijos de segmentos de límites y controles geoespaciales. La fundamentación para la relación entre poblaciones locales y decisiones empresariales se remonta a investigaciones previas en el Perú en el estudio de la informalidad (e.g., Tello (2015)) como precedentes útiles para la investigación de Natividad (2019) sobre geografía y tributación.

Los resultados de Natividad (2019) indican un fuerte impacto discontinuo negativo de la institución de mita en el tamaño actual de la empresa. Además, los resultados indican una probabilidad estadísticamente fuerte de que las empresas en las regiones de mita en conjunto no hagan inversiones en activos fijos.

Para entender el mecanismo subyacente, Natividad (2019) utiliza datos de tres siglos de impuestos coloniales para analizar la persistencia de la mita. Usando el detalle geográfico disponible en registros de tesorería de las Cajas Reales (TePaske y Klein, 1982) para las comunidades en ambos lados del límite de mita, se examina los impuestos personales anuales. La evidencia indica que la mita causó un impacto discontinuo positivo en la tributación de áreas dentro del territorio de la mita. Además, bajo algunas especificaciones, Natividad (2019) encuentra que este aumento en la tributación de las áreas de mita se mantuvo sin cambios durante los tres siglos de transacciones registradas. Los hallazgos

ofrecen un canal plausible para los efectos de largo plazo que conecta la mita colonial con los resultados actuales, posiblemente erosionando las normas culturales de acumulación y formalidad entre las poblaciones de la mita.

Este estudio previo contribuye al debate sobre cómo las políticas interactúan con las barreras institucionales (Easterly, 2005). Este canal podría entenderse como un antecedente de siglos de desgaste de capital que es consistente con baja inversión en capital humano y subdesarrollo general, así como la erosión de la base tributaria ocasionada por la evasión.

3. Metodología

3.1 Hipótesis (acápito indicado por las bases del CIES como parte integrante de la Metodología)

El presente estudio propone analizar variables departamentales que reflejan fundamentos económicos y sociales para evaluar su impacto en el margen extensivo e intensivo de la tributación en el Perú. Bajo el supuesto de que todas las geografías son exactamente iguales en el país, la recaudación debería ser también muy similar. Sin embargo, como se ha discutido en el marco teórico, las distintas geografías regionales no muestran homogeneidad: hay factores físicos, demográficos, sociales y económicos que pueden llevar a distintos niveles de formalización, puntualidad, y cumplimiento de un buen comportamiento tributario.

Las hipótesis centrales de este estudio versan sobre cuatro variables socioeconómicas. Ello no significa que las variables geofísicas o demográficas no sean importantes. Por ejemplo, la altitud de algunas zonas de la sierra, o el gran tamaño poblacional de algunos distritos de la costa, son determinantes en la bonanza o penuria económica de los contribuyentes obligados a reportar impuestos. Justamente por su importancia, esas variables alternativas—que por sí mismas o por su influencia en otras variables deben ser propiamente consideradas—se proponen como variables de control.

La primera hipótesis del estudio versa sobre el impacto de la variable educación. Por un lado, se puede proponer que a mayor educación habrá mayor cumplimiento tanto en el margen extensivo como en el intensivo. Sin embargo, si existe una cultura de evasión en la que la mayor educación permite encontrar formas más sagaces para realizar tal evasión, es también posible la hipótesis contraria: que a mayor educación haya menor formalización y menor cumplimiento en el margen intensivo de la tributación.

La segunda hipótesis de este estudio se centra en el impacto de la inversión pública. Una visión constructiva del gasto público en las distintas regiones del país es que dichos desembolsos ayudan a crear una cultura de bienes comunes que es valorada por las poblaciones locales, incentivándolas a tributar más, teniendo así impactos positivos en el margen extensivo y en el intensivo. Sin embargo, debido a que los bienes públicos siguen financiándose incluso si hay riesgo moral o selección adversa por parte de las poblaciones locales, es probable también el caso opuesto: a más inversión pública, más tendencia a permanecer informal y a un pobre comportamiento en el margen intensivo.

La tercera hipótesis del análisis propone, de manera relacionada a la segunda hipótesis, un rol para la percepción sobre la corrupción. En la medida en que las poblaciones locales ven la corrupción como un problema, estarán menos dispuestas a contribuir con sus impuestos. Se propone, por tanto, una relación negativa entre la percepción de corrupción y la participación en el margen extensivo y en el intensivo de la tributación.

Finalmente, la cuarta hipótesis puntualiza la relación entre el nivel de confianza en la autoridad tributaria manifestado por la población local y el correspondiente comportamiento tributario. A mayor confianza depositada en la SUNAT, se puede proponer una mayor participación en la base tributaria y un mejor cumplimiento en el margen intensivo.

3.2 Datos

El repositorio de datos proviene de diversas fuentes: el Censo de Establecimientos del 2008, la Encuesta Nacional de Hogares del 2002 al 2017, el Sistema Nacional de Inversión Pública del 2001 al 2015, el registro público de la SUNAT y mapas digitales creados con información pública. La tabla 1 presenta un resumen estadístico de los datos empleados para el análisis de regresiones del margen extensivo e intensivo de la tributación en el Perú; en el encabezado de la tabla 1 se mencionan las fuentes de donde se obtuvieron las variables principales.

La tabla 1 refleja las múltiples estructuras de los datos empleados para el análisis. Algunas fuentes son de corte transversal, mientras que otras permiten observar al mismo sujeto tributario o a la misma geografía en varios períodos. Dada esta estructura múltiple de los datos, el análisis toma especificaciones diversas.

En específico, la información principal sobre el tratamiento estadísticamente representativo de las regiones peruanas (i.e., departamentos) se capturó usando observaciones en el nivel departamento-año siempre que la información estuviera disponible. En el caso de los años de educación secundaria completados y la percepción de corrupción, la información estaba disponible desde el 2005 hasta el 2018. Similarmente, aunque con menos cobertura en el tiempo, la información sobre confianza en la SUNAT se hallaba disponible desde el 2015 hasta el 2018, razón por la cual hay menos observaciones disponibles en el nivel departamento-año. Cabe resaltar que la información sobre la confianza de los hogares en la SUNAT no fue recabada en el 2007, razón por la cual no es posible analizar el margen extensivo de la tenencia de un número de RUC usando el Censo de Establecimientos (CENEC) 2008. Finalmente, la información sobre inversión pública se halla disponible de 2001 al 2015 y es calculada sobre la base de los distritos y sumada para cada departamento y año.

Toda esta información departamental, al ser un condicionante externo al comportamiento tributario, se rezaga un año con respecto a las variables dependientes. Esto es, para considerar la cuestión sobre cómo el ambiente educativo, o el ambiente de percepciones, o el ambiente de confianza, o el estado general de inversión pública del Estado en un departamento influye en el comportamiento tributario individual de las empresas estudiadas, es importante que dicho comportamiento no sea contemporáneo o concomitante con los factores propuestos.

Un argumento más fuerte para la exogeneidad de las variables departamentales es su representatividad estadística general, que va más allá del caso concreto de cada variable dependiente endógena particular a cada empresa o sujeto tributario. En concreto, gracias a que las fuentes son estadísticamente representativas en el nivel departamental (i.e., de cada región peruana), se puede hacer un análisis estadísticamente confiable. En contraste, la Encuesta Nacional de Hogares no es representativa en un nivel microgeográfico más granular que el de departamento (INEI 2017), un detalle que muchos estudios previos omiten al explicar esa fuente de datos y al intentar mostrar análisis microgeográficos. Por esa razón técnica, es estadísticamente incorrecto proponer cruces de información individual de hogares con información individual de empresas en un nivel microgeográfico o individual. El presente estudio no hace esos cruces en el nivel individual, sino que usa información individual convirtiéndola en información departamental utilizando todas las zonas de cada

departamento, haciéndola así estadísticamente representativa según la ficha técnica de las encuestas (INEI 2017), y aplica la información departamental a variables representativas del universo de empresas.

La tabla 1 registra también estadísticos sobre las empresas analizadas en el estudio. Debido a que el diseño empírico se centra en distritos del borde geográfico de los departamentos, solamente se presenta la información sobre entidades situadas en dichos distritos. El padrón RUC utilizado en el estudio considera a todos los contribuyentes, no solamente a empresas.

Son dos las fuentes empleadas para obtener información sobre empresas. Por un lado se encuentran las empresas del Censo de Establecimientos (CENEC)³; desde un punto de vista estadístico, el atractivo principal de esta fuente es que registra tanto a empresas formales como a informales, lo cual permite un análisis inédito del margen extensivo y factores departamentales; la desventaja de la información de CENEC es su antigüedad (año 2007) y su estructura de corte transversal, que impide un análisis de panel de datos. La interpretación de la tenencia de RUC como reflejo de un mero éxito administrativo de SUNAT no invalida su utilización, ya que el análisis de regresiones comparará geografías muy similares en las que el éxito administrativo de SUNAT es claramente similar, por lo cual se abstrae esa dimensión. El análisis de formalización de empresas realizado en Natividad (2019) es un precedente directo al análisis realizado aquí; el rigor científico de estudiar la decisión de obtener un número de RUC se halla así validado en la literatura económica. No existe información posterior al 2007 que sea representativa respecto a la tenencia de RUC; por ejemplo, la Encuesta Nacional de Hogares no permite calcular con fiabilidad estadística cuántos negocios son formales o informales en este sentido, ya que las preguntas de la encuesta se enfocan en el empleador de un miembro de hogar, y no en el dueño del negocio en sí; además, las identidades de las empresas no se encuentran registradas en la Encuesta Nacional de Hogares, por lo cual el problema de doble contabilidad en intentar modelar la propensión de empresas a formalizarse sería severo.

Por otro lado, este estudio cuenta con información del padrón RUC de la SUNAT para todos los contribuyentes. La ventaja de esta fuente es la ausencia de sesgo de supervivencia en su información histórica; su limitación es que solamente se enfoca en entidades que han abrazado la formalidad, al menos en parte, al obtener un RUC y registrarse con la autoridad tributaria. La información proporcionada por la SUNAT para este estudio incluye el distrito de ubicación de cada contribuyente, sus fechas de inicio de actividad y obtención o baja de número de RUC, y su código de actividad industrial. Toda esta información es aprovechada integralmente en el análisis.

Además del padrón RUC de la SUNAT, el estudio se beneficia de dos fuentes de información públicamente disponibles otorgadas por la SUNAT: las omisiones tributarias de contribuyentes (otorgada en forma de panel sin sesgo de supervivencia) y la deuda coactiva de contribuyentes (otorgada solamente como corte transversal a diciembre de 2018). El valor de estas fuentes depende de la estructura de sus datos. Los datos de omisiones tributarias son completos, mientras que los datos de deuda coactiva solamente son un corte final que no registra movimientos ni infractores pasados que rectificaron su situación. Se puede adelantar, por esto, que el potencial estadístico de una fuente es superior al de la otra, lo cual incide en el análisis.

³ El autor utiliza la información del CENEC gracias a su acuerdo con la dirección de estudios económicos de PRODUCE. En caso SUNAT deseara acceder a esa información, haría falta que SUNAT coordine y pida permiso especial de la dirección de estudios económicos de PRODUCE. Esta gestión trasciende la competencia del autor.

Si bien el análisis se enfoca en variables dependientes (comportamiento tributario) y sus factores determinantes (aspectos educativos, de percepción o de inversión), también hace uso de variables de control para reducir la posibilidad de que las relaciones establecidas sean espurias. En específico, la elevación del terreno en que se encuentra un distrito, la pendiente promedio de ese distrito, y la población distrital en miles para cada año analizado son factores considerados en el análisis. Debido a que distintas especificaciones utilizan la información con distinta frecuencia, la tabla 1 resume las variables de control dependiendo de cómo se las presenta en el análisis. Las diferencias en los valores estadísticos de la misma variable (e.g., elevación en el panel B vs. elevación en el panel C de la tabla 1) obedecen a la distinta composición de la muestra según sea el caso.

3.3 Patrones observables en el resumen estadístico

Una revisión del resumen estadístico de la tabla 1 puede ayudar a visualizar los principales conceptos por analizar. Interesa especialmente considerar el nivel de variación de las variables independientes y dependientes para el análisis.

Los factores departamentales a los que distintas empresas se hallan sujetas según el lado de la frontera departamental en que se sitúan tienen un nivel importante de variación. Como se distingue en la tabla 1, por ejemplo, el nivel educativo de las poblaciones de los departamentos varía desde 2.3 años hasta 3.8 años, con una desviación estándar de 0.37. Esto implica que, aplicando la regla estadística de dos desviaciones estándar arriba o abajo del promedio, hay un rango normal de casi 1.5 años de variación en la educación secundaria con respecto al promedio nacional de 2.3 años; muy pocos valores se ubican en los extremos dada esta distribución. Esta variación asegura que el análisis captura un nivel sustancial de posibles diferencias en el ambiente educativo que influye en las decisiones tributarias.

Los otros factores departamentales, cuya información estadística se resume en la tabla 1, también varían significativamente. La percepción de corrupción como problema en la sociedad tiene un promedio de 0.31 y una desviación estándar de 0.15. Por su parte, la confianza en la SUNAT muestra un promedio de 0.23, variando en un rango de 0.07 a 0.34 dependiendo del departamento y del año. Asimismo, la inversión pública muestra una gran variación también, con un valor promedio de 478 millones de soles y una desviación estándar de 690.6. El análisis de cómo impactan estos determinantes departamentales se beneficiará de estos niveles de variación.

Las variables dependientes principales en el análisis se encuentran descritas en los paneles B, C, D, y E de la tabla 1. La primera dimensión del análisis versa sobre el margen extensivo de la base tributaria. La tenencia de RUC por parte de empresas en el corte transversal de diciembre 2007 tiene un promedio de 0.56, y lógicamente, la desviación estándar de esta binaria no está definida.

Por otro lado, la variable binaria sobre si el contribuyente inscribe el RUC en sus primeros treinta días de actividad tiene un valor promedio de 0.988, sugiriendo un alto nivel de prontitud (declarada) en la inscripción de actividad económica; algo de variación tiene esta variable, dejando campo para el análisis econométrico del margen extensivo bajo la interpretación de que la prontitud de los contribuyentes registrados en registrarse antes o durante su actividad económica refleja su propensión a formar parte de la base tributaria. Debido a que la fecha de inicio de actividades declarada por las empresas *ex post* no puede ser verificada por la SUNAT, esta variable depende de la veracidad de las instituciones inscritas; aun así, algo de valor estadístico se puede desprender de esta declaración.

Además del margen extensivo, la información disponible permite abordar el margen intensivo de la tributación. La omisión de declaraciones de impuesto a la renta por parte de sujetos tributarios se captura con una variable binaria que varía para cada contribuyente cada año. Esta variable tiene un valor promedio muy bajo (solamente 0.5% de las observaciones en cuestión muestran una omisión de declaración). Similarmente, el logaritmo del número anual de omisiones de declaración del impuesto a la renta por parte de cada contribuyente muestra un valor relativamente pequeño en promedio. Estas variables capturan el buen o mal comportamiento de los sujetos tributarios en cuanto a su reporte de impuestos y dan una idea del margen intensivo de la base tributaria.

De manera más limitada, debido a la naturaleza de datos de corte transversal otorgados por SUNAT, la información sobre deuda coactiva al fin del año 2018 es utilizada para examinar el margen intensivo. La dificultad consiste que, al ser información sobre el balance de deuda coactiva, lo que interesa primero es entender cuál es la muestra relevante para comparar a esos contribuyentes con balance de deuda al final del 2018. Por ejemplo, hay empresas con deuda coactiva en el 2018 que corresponde a ejercicios bastante anteriores (e.g., 2014, 2015); muchas de esas empresas pueden haber dejado de existir, y no pueden compararse con empresas aún activos en el 2018. Por tanto, la muestra resumida en el Panel E de la tabla 1 se basa en todos aquellos contribuyentes situados en distritos de borde de departamento que estaban activas (i.e., no de baja) en el año 2018. Debido a que la deuda coactiva originada en períodos tributarios de trimestres anteriores al cuarto trimestre del 2018 puede haber sido pagada antes de diciembre de 2018, sería inexacto considerar en la variable dependiente dichos períodos tributarios. Por tanto, para generar la variable dependiente, solamente se considera aquella deuda coactiva cuyo ejercicio de generación de deuda es octubre 2018, o noviembre 2018, o diciembre 2018. Con este diseño empírico, se compara correctamente a todos los contribuyentes activos para estimar quiénes de entre ellos pueden estar mostrando mal comportamiento tributario al generar deuda coactiva.

La variable sobre deuda coactiva a diciembre de 2018 generada en períodos del cuarto trimestre del 2018 es modelada en forma binaria y tiene un valor de 0.004. Por su parte, el logaritmo de este monto de deuda coactiva (más uno) muestra un valor promedio de 0.028, con un gran nivel de variación, como refleja la desviación estándar de 0.45. En general, hay variación sustancial tanto en las variables dependientes como en las independientes.

3.4 Descripción gráfica de tratamiento

Antes de detallar la especificación econométrica, es útil revisar el mapa geográfico del Perú en las dimensiones departamentales de mayor interés. La figura 1 presenta un mapa de los 24 departamentos del Perú y de las provincias de Lima y Callao⁴. En colores oscuro y claro se refleja la primera variable binaria (de tratamiento y control, respectivamente) consistente en la dicotomización de la variable sobre años de educación secundaria completados capturada con información de hogares del 2007 según se encuentren por encima de la mediana departamental de educación secundaria promedio.

El procedimiento para definir los niveles educativos en los departamentos es sencillo gracias a la disponibilidad de la data. La información de la Encuesta Nacional de Hogares se encuentra en el nivel individual. Tomando esos datos, se colapsa la información en el promedio departamental de la variable “años de educación secundaria completada”. El promedio captura los niveles de educación secundaria de distintos departamentos, que

⁴ En adelante, cualquier mención a departamentos considerará a Lima-Provincias y Lima como un solo departamento, y analizará al Callao como una provincia distinta.

muestran valores bastante diferentes en la geografía peruana. Lo que se busca en esta descripción panorámica de los niveles de educación en los departamentos peruanos es abordar la cuestión sobre si los distintos niveles educativos de geografías peruanas influyen en el margen extensivo de si las empresas deciden entrar al sistema tributario o no. Nótese que el diseño empírico utiliza información *departamental* estadísticamente representativa para caracterizar diferencias de tratamiento entre *departamentos distintos*.

La información sobre niveles de educación secundaria (años completados) es significativa estadísticamente en el nivel departamental. La ficha técnica (INEI 2017) indica que la representatividad estadística se obtiene precisamente en el nivel departamental. Esta es una justificación adecuada para que la caracterización de la educación como variable explicativa se realice en departamentos geográficos, y no en zonas más granulares. En general, todas las variables departamentales usadas en el diseño empírico son calculadas en el nivel de departamento, no en niveles más granulares, para así respetar la representatividad estadística de la información recolectada originalmente.

Además, es razonable convertir en dicotómica la variable sobre años completados de educación secundaria en grupos alta educación (tratamiento) vs. baja educación (control). En general, la reducción a una variable dicotómica ayuda a aclarar el impacto del tratamiento para que se noten diferencias entre grupos con valores significativamente distintos.

La figura 2, de manera análoga a la figura 1, mapea los departamentos peruanos clasificándolos en los grupos de tratamiento y control, según el valor promedio departamental de la confianza de los hogares en la SUNAT sea alto o bajo con respecto a la mediana de los promedios departamentales de la variable de confianza. La información estadística sobre el nivel de confianza en SUNAT en el nivel departamental es estadísticamente representativa justamente en el nivel departamental. Nuevamente, nótese que el diseño empírico utiliza información *departamental* estadísticamente representativa para caracterizar diferencias de tratamiento entre *departamentos distintos*.

Al comparar los mapas de educación y confianza en la SUNAT, un primer patrón visible es la relación negativa entre los altos niveles educativos de la población y los bajos niveles de confianza de hogares en la SUNAT. Esto sugiere que los patrones de desconfianza de los hogares respecto a la autoridad tributaria no necesariamente implican un menor nivel de educación, e incluso puede indicar una mayor prevalencia de hogares altamente educados en zonas de mayor desconfianza.

Un segundo patrón de interés en la figura 1 es la concentración geográfica en los niveles de mayor educación: hay grandes zonas adyacentes del Perú en que la educación se encuentra por encima de la mediana, mientras que otras zonas adyacentes del Perú se encuentran casi ininterrumpidamente por debajo de la mediana departamental de años de educación secundaria. Este patrón indica que las comparaciones entre grupos de tratamiento y control utilizando bordes departamentales necesariamente se restringe a algunos departamentos, ya que muchos departamentos de tratamiento se encuentran junto a otros que son también de tratamiento. Este patrón se aplica solamente a la especificación que modela el tratamiento de manera binaria. En el caso distinto en que la especificación define tratamiento de manera continua usando los promedios de educación secundaria departamental, la comparación entre departamentos adyacentes es más amplia y no se centra solamente en los contrastes marcados del caso binario.

Aunque esta descripción gráfica sugiere alguna relación entre distintas variables socioeconómicas (medidas en distintos años), la relación de interés no se enfoca en la

relación de estas variables entre sí sino en el impacto independiente que puedan tener en el comportamiento tributario de las empresas en los territorios correspondientes.

3.5 Especificación

Una dificultad econométrica al analizar diferencias geográficas y diferencias socioeconómicas como la tributación es el de la endogeneidad. Si uno ve a una empresa en una localidad Z tributando bien, y ve a otra empresa en la localidad K tributando mal, existe la posibilidad de que la ubicación Z sea mejor que la ubicación K, pero también existe un sinnúmero de explicaciones alternativas no observadas en el análisis que estén dificultando la inferencia, por ejemplo mediante variables omitidas o mediante selección.

Dada la potencial endogeneidad de las diferencias geográficas y la tributación, la especificación empírica propuesta en este estudio consiste en fijar la atención en territorios adyacentes al límite entre dos circunscripciones con influencia geopolítica distinta. Volviendo al ejemplo de las ubicaciones Z y K: si son seleccionadas de manera tal que Z y K estén una muy cerca de la otra, con la misma latitud, la misma zona del Perú, y controlando por muchos factores, dejando que simplemente varíe entre ellas el límite geopolítico entre el departamento 1 y el departamento 2 que contiene a esas localidades, es posible que el efecto de las variables departamentales de 1 y 2 tenga una influencia medible con cierta confiabilidad. Las variables departamentales de interés tienen que estar medidas de manera confiable. Afortunadamente, estas variables se obtienen ya sea de fuentes universales (no muestras sino poblaciones) o de fuentes muestrales con representatividad estadística en el nivel departamental. Se garantiza así que el tratamiento de la variable departamental es representativo.

Esta técnica viene reforzada por varias otras ayudas. Por ejemplo, la inclusión de efectos fijos para numerosos segmentos de las fronteras que separan a los departamentos del Perú, así como el empleo de una ventana de observaciones reducida a cada lado de la frontera (distritos del borde) ayuda a una estimación más efectiva.

La muestra para el análisis son todos los sujetos tributarios que se encuentran ubicados en un “distrito borde”. Un distrito borde es aquél que se encuentra adyacente a los límites geopolíticos departamentales. La ventaja de emplear solamente estas geografías consiste en hacer la comparación lo más cercana posible, de tal forma que la variación de interés se dé efectivamente en el tratamiento correspondiente a cruzar la frontera que existe entre un departamento y el departamento contiguo. Por su parte, todos aquellos distritos del interior (es decir, que no bordean los límites del departamento) son excluidos del análisis econométrico, y los resultados no necesariamente deben extrapolarse a ellos (Lee y Lemieux 2010). Un precedente para esta especificación es el de Natividad (2019) en el estudio de la mita colonial minera y sus impactos de largo plazo en empresas.

La especificación propuesta para el análisis es:

$$y_i = \alpha + \beta \text{tratamiento}_d + \gamma X_{id} + g(\text{ubicación}_d) + \lambda_j + \phi_s + \mu_i \quad (1)$$

donde la variable dependiente refleja ya sea el margen extensivo o el margen intensivo de la tributación. Para modelar el margen extensivo se utiliza primero la tenencia de RUC en diciembre de 2007, y luego, la decisión de inscribir el RUC relativamente temprano en la vida operativa de aquellos contribuyentes que efectivamente se registran eventualmente con la SUNAT. Posteriormente, para modelar el margen intensivo de la tributación, se utiliza como variable dependiente las omisiones tributarias en la declaración del impuesto a la renta, ya sea como una variable binaria con valor igual a uno si sucedió, o como logaritmo de uno

más el número de omisiones ocurridas. Finalmente, para modelar el margen intensivo de la tributación usando deuda coactiva con el Tesoro Público (cualquiera sea su origen) en el último trimestre de 2018 se utiliza una variable binaria igual a uno cuando el contribuyente tiene deuda coactiva, o también el logaritmo de uno más el monto de la deuda coactiva en soles.

En la ecuación (1), *tratamiento* es una variable dicotómica igual a uno para cuando la variable departamental es mayor o igual al valor de la mediana nacional, y cero de otra manera. En todos los casos en que la variable departamental está disponible en el nivel departamento-año, la especificación permite usar estos valores distintos cada año, así como usar efectos fijos de año. La firma i en el distrito d , cuyo segmento de frontera más cercano es s , es analizada usando polinomios que siguen la función g de su ubicación geográfica, tal como es estipulado en modelos de regresión discontinua (Lee y Lemieux, 2010). Aunque hay cierto margen de libertad para elegir la función polinómica, siguiendo a Natividad (2019), la función de polinomios g preferida aquí es cúbica, aunque por robustez se estima también el modelo con polinomios de grado dos o grado cuatro en un análisis suplementario no incluido en este informe parcial y se verifica que los resultados son similares.

Las variables de control en el nivel de empresa y distrito X_{id} son: elevación y pendiente del distrito, y efectos fijos de edad de la empresa (en años desde el inicio de operaciones) y de industria. Es importante controlar por estas variables geográficas en vista de argumentos teóricos y evidencia empírica sobre las diferencias de tributación en zonas con distinta altitud (e.g., Scott 2009). Los errores estándar son calculados empleando *clusters* robustos a nivel de distrito. El análisis se enfoca en distritos colindantes con la frontera de los departamentos del Perú.

El diseño de este estudio, estimando la ecuación (1), cubre toda la geografía nacional enfocándose en diferencias entre departamentos. La variable geográfica de distancias consiste en la distancia al límite departamental. Los efectos fijos de segmento del borde departamental cubren porciones de 50 kilómetros cada una para toda la red de fronteras departamentales peruanas.

El comparar distritos que se encuentran a uno y otro lado de los bordes geopolíticos de un departamento requiere controlar efectivamente por diferencias geográficas que podrían distorsionar el análisis si los distritos no están uno muy cerca del otro. La especificación de la ecuación (1), para hacer efectiva esta idea, incluye efectos fijos correspondientes a los segmentos de 50 kilómetros de longitud en que se dividen todos los bordes departamentales del Perú.

Adicionalmente, para estimar efectos entre empresas muy distintas cuando se mide su comportamiento tributario se añaden las variables de control en el nivel de cada empresa: la edad y la industria de actividad. Estas variables están efectivamente capturadas como efectos fijos incluidos en la ecuación (1).

4. Resultados definitivos

4.1 Patrones observables en la evolución del registro de contribuyentes

Una limitación de cualquier base de datos oficial sobre el número de empresas registrado es que no considera la actividad de las entidades informales, que deben quedar así forzosamente sobreentendidas. Una ventaja de este estudio es proporcionar también modelos con información del CENEC sobre empresas informales, que ayudarán a entender el margen extensivo de la tributación, como se verá en el reporte de modelos de regresión.

A pesar de sus limitaciones, el registro de contribuyentes puede dar una idea general sobre cómo operan los negocios en el país. El análisis gráfico del padrón del RUC proporcionado por SUNAT muestra, para el período de 1994 a 2018, un crecimiento sostenido de la inscripción de entidades con actividad económica acompañado de salidas del registro tributario mediante la baja del RUC. La figura 3 detalla cómo estas inscripciones o altas han tenido una tendencia alcista sostenida, con pendiente similar a la de las bajas aunque con una magnitud mucho mayor.

Sin embargo, la inscripción de empresas no es indiscriminada: es el sector servicios el que más sostenidamente ha crecido en el período analizado. El segundo panel de la figura 3 refina la tendencia, distinguiendo el crecimiento entre servicios y manufactura, siendo ésta última una familia de actividades económicas con crecimiento menos sostenido y con un volumen de empresas más pequeño. En el tercer panel de la figura 3 se distingue un patrón mixto para las empresas del sector extractivo. Al considerar el crecimiento generalizado de la inscripción de negocios en el Perú, cabe anotar a qué sector de la economía pertenecen.

El éxito de una entidad que se inscribe en el registro de contribuyentes de la SUNAT puede evaluarse en el breve plazo de un año luego de dicha inscripción. La figura 4 detalla cómo la fracción de contribuyentes que son dados de baja dentro de los primeros 12 meses desde su inscripción al RUC ha variado entre menos de 0.01 hasta casi 0.06, sin revelar un patrón temporal directamente apreciable. En la presente década, la cifra se había estabilizado alrededor de 0.04, aunque una notable baja en la mortalidad de los contribuyentes recién inscritos caracterizó al 2018. El analizar esta muestra temprana de éxito ayuda a entender—a falta de información sobre ventas o ganancias—cómo va cambiando el conjunto general de negocios.

Precisamente el dinamismo de las entradas y salidas de empresas puede impactar en la experiencia acumulada por aquellas empresas que subsisten en el mercado. Una forma de analizar esa experiencia es midiendo la edad de los contribuyentes presentes en distintos puntos del tiempo de la evolución del mercado de contribuyentes en el Perú. La figura 5 muestra gráficos de densidad kernel para la edad de los contribuyentes en los años 1994, 2000, y 2018. Claramente, se aprecia un envejecimiento —o madurez— de las empresas activas en el mercado con el paso del tiempo. Considerando la situación del total de contribuyentes en el país en el 2018, es digno de mención que la etapa del ciclo de vida en la que se encuentran es muy diferente que el total de contribuyentes hace dos o tres décadas, por lo cual las políticas sectoriales diseñadas para generar impactos en contribuyentes hoy no pueden ser las mismas que hace veinte años.

Un condicionante relevante para la supervivencia de los negocios es la población local. Como demostración simple de cómo se relaciona el número de negocios con la población de cada departamento, la figura 6 muestra cuatro instantáneas distintas, para los años 2001, 2006, 2012, y 2018, respectivamente. En cada una de las imágenes se muestra el número de contribuyentes registrados en cada departamento así como la población de ese departamento. (Lima y Callao son excluidas del gráfico, pues rompen el esquema dado su gran tamaño). La relación se muestra como notablemente positiva para todos los años, con un fuerte crecimiento en el número acumulado de contribuyentes activas hacia el final de la muestra temporal.

Reiterando la limitación de la información públicamente disponible para este estudio, cabe recordar que no se conocen los ingresos o la tributación de cada contribuyente. Sin embargo, para aquel segmento obligado a reportar el número de empleados y prestadores de servicios,

existe información anual sobre estas dos series, dando así una buena pista sobre el volumen de la empresa.

En específico, la figura 7 muestra cómo, para los contribuyentes que han operado en el mercado tanto en el año 2008 (inicio de la serie con empleados) como en el 2018 (fin de la serie), ha venido variando el promedio anual de empleados y prestadores de servicios. Se distingue hacia la primera mitad de la muestra una relación positiva en las series que comparan el promedio de estos dos tipos de trabajadores; pero en la segunda mitad de la muestra, el número promedio de empleados crece asintóticamente hasta 45, mientras que el número de prestadores de servicios decrece drásticamente. Este patrón es particularmente revelador del enfriamiento en el crecimiento de las empresas.

Al considerar cuán distribuido está el personal laboral en el territorio peruano, cabe preguntar sobre su concentración en distintos departamentos. La figura 8 analiza el índice de Herfindahl basado en la participación de cada departamento en el número total de empleados o de prestadores de servicios por separado y muestra niveles relativamente altos de concentración. Esta descripción se complementa con la figura 9, que muestra la fracción de empleados y prestadores de servicios en Lima y Callao comparada con el resto del Perú. La tendencia en empleados es muy estable, alrededor del 65% de todo el país. En contraste, la tendencia en prestadores de servicios es hacia la menor preponderancia de la capital en años más recientes.

Pero no todos los sectores económicos tienen la misma demanda de personal. La creencia de que algunos sectores económicos han visto disminuir sus demanda por trabajadores en años recientes se confirma en la figura 10. Al distinguir la fracción de empleados totales por sector económico, se distingue una menor importancia de manufactura, una marcha relativamente estable de servicios, y un incremento reciente en industrias extractivas. La figura 11, capturando análogamente la realidad para prestadores de servicios, muestra tendencias más mixtas.

Así como las tendencias en la cantidad de personal han cambiado en los últimos años, es posible explorar una conexión, a manera de ilustración, con la desconcentración de industrias. En específico, es posible que el dinamismo reciente del mercado no haya ayudado necesariamente a la liberalización de todas las industrias, sea por falta de atractivo en algunas, sea por una mayor presencia de importaciones que compiten con producción sin compartir el mismo código industrial que refleje esa mayor competencia. Al medir la concentración de empleados en entidades dentro de cada industria, la figura 12 muestra la evolución de las que eran industrias altamente concentradas en el mercado laboral de empleados en el año 2008. Observando a estas empresas en los siguientes diez años se muestra que algunas mantienen su concentración, mientras que algunas la disminuyen.

Así como el personal empleado o contratado es un correlato del volumen empresarial, las omisiones en la declaración del impuesto a la renta por parte de los contribuyentes brindan una idea de diferencias fundamentales en el desempeño de los contribuyentes. La figura 13 brinda una panorámica para tres sectores. La fracción de omisiones debida a manufactura ha ido disminuyendo, mientras que la fracción de servicios ha ido aumentando. Esta relación puede asociarse meramente al enorme crecimiento del sector servicios. Sin embargo, no se confirma con extracción: a pesar de haber crecido en número de empleados proporcionalmente, el sector extracción no se ha disparado en su participación en omisiones tributarias.

Habiendo considerado la evolución de las entidades con información pública no reservada disponible en el registro de contribuyentes, el estudio reporta a continuación los resultados de

tres conjuntos de análisis de regresiones. Primero, los modelos sobre el margen extensivo y sus determinantes departamentales. Segundo, los modelos sobre el margen intensivo capturado por las omisiones en declaraciones de impuesto a la renta, que gozan de la ventaja de contar con un panel de datos de omisiones. Tercero, los modelos sobre el margen intensivo explorando datos limitados del cierre del año 2018 en un corte transversal sujeto a potenciales sesgos de supervivencia y de variables no observadas.

4.2 Resultados de regresiones sobre el margen extensivo

4.2.1 Educación

La tabla 2 reporta los resultados del análisis de regresiones estimando la ecuación (1) con el objetivo de estudiar formalmente cómo las diferencias departamentales entre los grupos de tratamiento y control para la variable educacional se relacionan con diversas dimensiones del margen extensivo de la base tributaria.

El primer modelo reportado emplea como variable dependiente la tenencia de RUC en diciembre de 2007; comparando a empresas que operan en el mercado, algunas formales, otras no, el modelo captura cómo el nivel educativo del departamento al que pertenece esa empresa de distrito de borde impacta en el margen extensivo. La primera columna de resultados en la tabla 2 reporta los estimados de la ecuación (1) con la variable dependiente sobre tenencia de RUC. El coeficiente de interés es el de la variable binaria *alta educación (1/0)*. Este coeficiente tiene un valor de -0.487 y un estadístico *t* de -9.78, indicando una alta significancia estadística y económica para la educación en explicar si las empresas de un distrito borde tienen RUC. Aquellos departamentos con alto nivel educativo tienden a albergar empresas con mucho menor propensión a registrarse con la autoridad tributaria. El coeficiente negativo de -0.487 representa un valor enorme comparado con el valor promedio de 0.56 para la variable dependiente.

El diseño empírico permite que la atribución del efecto de la educación sea marginal, pues la regresión controla las posibles influencias paralelas de otros factores incluidos como variables de control y efectos fijos en la ecuación (1). En concreto, la elevación del terreno en el distrito, la pendiente geográfica promedio del distrito y la población del distrito en cuestión se encuentran incluidas en la estimación. Incluso controlando por esas posibles diferencias geoespaciales y sociales, la educación se relaciona negativamente con la formalización empresarial para constituir la base tributaria peruana. Además, la diferencia entre territorios a un lado y al otro de los límites geopolíticos departamentales se halla controlada por un polinomio cúbico flexible de la distancia del distrito al punto más cercano del borde geográfico. Por tanto, son las diferencias que se originan al pertenecer a un departamento peruano con más educación o menos educación generan un impacto discontinuo en la base tributaria peruana.

El segundo modelo reportado en la tabla 2 examina el impacto de la educación en la decisión de inscribir el RUC relativamente temprano en la vida operativa de aquellos contribuyentes que efectivamente se registran eventualmente con la SUNAT. La variable dependiente es la inscripción del RUC dentro de los primeros 30 días de la actividad económica del contribuyente. El coeficiente de interés es el de *alta educación (1/0)* y tiene un valor de -0.004 y un estadístico *t* de -3.19, confirmando así el impacto negativo de la educación en el margen extensivo de la base tributaria peruana. La magnitud económica de este impacto es razonable: el pertenecer a un departamento con alto nivel educativo representa un 30% del cierre de la brecha entre registrarse con SUNAT en los primeros 30 días de actividad vs. no registrarse en esos días.

La diferencia entre los modelos presentados en la tabla 2 debe precisarse. Ambos son modelos que toman solamente una observación por entidad; el modelo que analiza la tenencia de RUC mira a las empresas de distintas edades que fueron consideradas por el CENEC en diciembre de 2007; el modelo que analiza la inscripción en los primeros 30 días mira a todos los contribuyentes del padrón RUC examinando una decisión en la etapa inicial de la vida de estas entidades. Por estas diferencias de diseño, el primero modelo debe controlar por efectos fijos de edad de la empresa y el segundo modelo no debe controlar, pues mide una decisión de la entidad naciente. Asimismo, el segundo modelo debe controlar por efectos fijos del año en que el contribuyente inicia operaciones según el padrón RUC, mientras que el primer modelo mira a todas las empresas de diciembre de 2007 y al controlar por su edad no necesita controlar por su año de inicio de actividades.

En general, los resultados sobre el nivel de educación y la base tributaria en geografías peruanas reportados aquí son nuevos para la literatura sobre tributación en el Perú, y sugieren una posible actitud oportunista del emprendedor situado en geografías con mayor nivel educativo para operar al margen de la formalidad (ver de Soto 1987). Al ser un primer conjunto de resultados en la literatura, y con las limitaciones de información estadística del caso, será importante contar con futuras investigaciones que ahonden en el tema. En un trabajo previo, Natividad (2019) encuentra impactos de la mita colonial minera en un mayor nivel de asistencia escolar y un menor nivel de formalización de las empresas en territorios adyacentes al borde geográfico de esa institución colonial. Es sugerente, por tanto, ese precedente en cuanto a un posible signo negativo en la asociación. Sin embargo, el precedente de Natividad (2019) es diferente, ya que relaciona un factor colonial ya abolido (la mita) en su influencia independiente sobre dos resultados contemporáneos (la educación y la formalización empresarial). En el presente estudio, los dos factores contemporáneos se relacionan econométricamente, aprovechando un diseño geográficamente más amplio en el territorio nacional completo.

4.2.2 Inversión pública

La tabla 3 amplía la investigación sobre los determinantes departamentales del margen extensivo de la base tributaria en empresas situadas en distritos borde, enfocándose esta vez en las diferencias en el nivel de inversión pública realizada por el Estado en todos los departamentos del Perú. Esta variable, proveniente del SNIP, captura la idea de que quizás la decisión individual de entrar a la base tributaria mediante la formalización está condicionada por cuánto gasto público reciben las localidades en cuestión.

El primer modelo de la tabla 3 reporta cómo una situación de alta inversión pública en el departamento en el año 2006 tiene un impacto negativo en la tenencia de RUC en el 2007, con un coeficiente de -0.393 y un estadístico t de -5.74 . Este impacto negativo es significativo económica y estadísticamente, indicando que aquellos emprendedores situados en departamentos comparativamente más beneficiados por el gasto del Estado son precisamente los más ausentes de la base tributaria peruana. El modelo controla efectivamente por muchos otros factores geográficos y sociales, revelando una influencia marginal fuertemente negativa de la inversión pública a la formación de la base tributaria. Nuevamente, al ser una investigación pionera en la relación entre estas variables, y con las salvedades del caso, sus resultados abren el camino para futuras investigaciones que ahonden en la cuestión.

La tabla 3 también reporta, en el segundo modelo, los resultados sobre cómo la inversión pública departamental impacta en la prontitud a inscribirse en el registro de SUNAT. La ventaja de este segundo modelo es que utiliza todos los años del padrón

RUC para los cuales la información de gasto público está disponible. Al analizar cómo es impactada la variable binaria sobre inscripción al RUC en los primeros 30 días de actividad, el coeficiente de alta inversión pública es -0.002 , con un estadístico t de -2.37 . Los resultados del segundo modelo de la tabla 3 complementan así la relación empírica descubierta en el primer modelo: a más inversión pública, menos prontitud a formar parte de la base tributaria. Utilizando datos bastante distintos, una especificación con más períodos y una dimensión relacionada pero no igual a la de la tenencia de RUC, los modelos de la tabla 3 dan una visión coherente.

Aunque puede parecer sorpresiva la relación negativa entre inversión pública y las actitudes individuales frente a la tributación, una lógica económica de *free riding* puede explicar dicha contraposición, complementando el caso de la educación. En específico, si un potencial contribuyente ve que ya existe gran inversión pública en su zona, es posible que no vea su tributación individual como determinante para la continuación de esa provisión de gasto público. Se da, de esta forma, un efecto *crowding out* típico de los incentivos pobremente alineados y con impactos colaterales.

4.2.3 Percepción de corrupción

Una posibilidad diferente sobre la influencia geográfica en los departamentos es analizada en la tabla 4: la percepción de alta corrupción en una zona geográfica podría reducir la confianza del emprendedor en la autoridad pública. Si hay una interpretación subjetiva de alta corrupción por parte de la población de una región, es posible que dicho condicionamiento impacte negativamente en la decisión de entrar a la base tributaria.

Para investigar esta conjetura, los modelos reportados en la tabla 4 utilizan las mismas variables dependientes de tablas anteriores, enfocándose esta vez en la explicación potencial de alta corrupción percibida por la población. En la primera columna de la tabla 4 se observa el impacto de alta corrupción en la tenencia de RUC; el coeficiente de interés tiene un valor de -0.437 y un estadístico t de -7.17 . La significancia es sustancial tanto estadística como económicamente en explicar la decisión de entrar a formar parte de la base tributaria peruana.

Por su parte, el segundo modelo de la tabla 4 analiza el impacto de la alta percepción de corrupción en la inscripción del RUC en los primeros 30 días de actividad. El coeficiente de interés es -0.001 y el estadístico t es -1.49 . No se encuentra significancia estadística para la alta percepción de corrupción en la dimensión temporal de cuándo entrar a la base tributaria. Los resultados emergen de un diseño empírico y un repositorio de datos únicos en la literatura hasta ahora, con las limitaciones de información estadística del caso, por lo cual será importante contar con futuras investigaciones que ahonden en el tema.

4.2.4 Confianza en SUNAT

Un último factor geográfico relevante para el estudio es la prevalencia de un alto grado de confianza poblacional en la SUNAT en la decisión de entrar a la base tributaria peruana. Debido a que esta variable de confianza solamente se halla disponible en la muestra de hogares para años posteriores al 2008, en que se realizó el CENEC, la única variable dependiente que se puede analizar para entender el margen extensivo de la tributación es la decisión de inscribir el RUC hasta los treinta primeros días de actividad económica.

La columna única de la tabla 5 reporta los resultados del impacto de la alta confianza en la prontitud para obtener el RUC. El coeficiente de interés es -0.001 y el estadístico t es -1.12 , indicando así una falta de impacto estadístico de la confianza en SUNAT en el margen extensivo.

Es importante reconciliar la falta de impacto de la confianza en SUNAT en la inscripción temprana de RUCs con el fuerte impacto negativo tanto de la educación como de la inversión pública. Por un lado, la racionalidad de los agentes económicos parece versar más sobre cómo operar fuera de formalidad gracias a que el nivel de conocimientos es mayor o a que el nivel de infraestructura es mayor; esta racionalidad no considera en sí un rol del prestigio suscitado por SUNAT en la decisión de formar parte de la base tributaria.

La prontitud para inscribirse en la SUNAT capturada por la segunda variable dependiente del análisis de márgenes extensivos solo modela indirectamente el margen extensivo. Todos los sujetos estudiados en su decisión de prontitud eventualmente se inscriben en SUNAT en algún momento y el padrón RUC registra eso. Pero no se observa a todos aquellos emprendedores que nunca se inscriben en el RUC. Por tanto, la falta de información sobre confianza en los años del CENEC, y la falta de información sobre informales en los años de la variable de confianza en SUNAT son dos ausencias que pueden jugar un rol en el poder estadístico de las pruebas realizadas.

4.3 Resultados de regresiones sobre el margen intensivo usando omisiones tributarias

4.3.1 Educación

El comportamiento tributario de los contribuyentes en diversas geografías peruanas, una vez que ya forman parte de la base tributaria, se debe analizar desde el punto de vista del margen intensivo de su accionar tributario. Con este fin, la tabla 6 modela como variables dependientes las omisiones tributarias en la declaración del impuesto a la renta, si es que las hubiera tenido, tanto como una dicotómica con valores cero y uno, o como el logaritmo de uno más el número de omisiones de declaración tributaria. Las variables de control son las mismas que en casos anteriores; debido a que ahora los modelos de margen intensivo son todos de panel de datos, a cada contribuyente se le observa varias veces (es decir, en varios años), lo cual permite incluir efectos fijos tanto de edad del contribuyente como de año calendario. Los modelos, por tanto, se conforman a un diseño robusto, el cual ayuda a interpretar la influencia de variables departamentales como algo genuinamente ligado al concepto que representan.

La variable departamental explicativa es la dicotómica para alta educación. El primer modelo reporta una relación positiva. El coeficiente de interés es 0.001 y el estadístico t es 2.35, lo cual implica un alto nivel de significancia estadística. Dos contribuyentes muy similares situadas en distritos adyacentes al límite geopolítico departamental tienen comportamientos tributarios distintos: aquella entidad en un ambiente de mayor educación es más propensa a caer en una omisión en la declaración del impuesto a la renta ese año.

Yendo más allá del hecho de si el contribuyente comete una omisión o no, el segundo modelo de la tabla 6 analiza como variable dependiente la cantidad de omisiones tributarias en cada año para cada contribuyente. Es decir, el análisis ya no se enfoca solamente en el hecho binario de la omisión sino en la frecuencia o incidencia de omisiones cada año, capturada en el logaritmo de uno más el número de esas omisiones. La segunda columna de la tabla 6 muestra un coeficiente de 0.001 para la variable de interés, con un t estadístico de 1.81, confirmando así su significancia. Este resultado es consistente con el del primer modelo y sugiere que el mayor nivel educativo de una geografía peruana es perjudicial para el margen intensivo de la tributación. En zonas con mayor educación hay peor comportamiento en omisiones tributarias. Los resultados emergen de una metodología y un repositorio de datos pioneros en la literatura hasta ahora, con las limitaciones de información estadística del caso,

por lo cual será importante contar con futuros estudios que profundicen en las relaciones de interés.

4.3.2 Inversión pública

La tabla 7 estudia el impacto de la alta inversión pública departamental en las omisiones tributarias de los contribuyentes cada año. El primer modelo de la tabla 7 reporta resultados altamente significativos al usar la inversión pública para explicar la variable binaria de si hubo omisiones ese año para cada contribuyente en cuestión. El coeficiente es -0.001 y el estadístico t es -3.48 , indicando así que a mayor inversión pública, menor propensión a omitir la declaración al impuesto a la renta.

El segundo modelo de la tabla 7 propone el logaritmo de las omisiones de declaraciones de renta como variable dependiente. Nuevamente, la alta inversión pública impacta negativamente en esa incidencia de omisiones; el coeficiente es -0.002 y el estadístico t , -3.38 .

Esta influencia benéfica de la inversión pública en reducir las omisiones tributarias de los contribuyentes resalta cuando se contrapone a la influencia perjudicial de la educación en las omisiones tributarias. Este contraste sugiere que mientras la educación y la inversión pública afectan alineadamente a la decisión de entrar a la base tributaria, una vez que la empresa ya es contribuyente, resulta que la educación y la inversión pública tienen impactos opuestos en el comportamiento tributario. Los resultados emergen de la metodología y el repositorio de datos empleados, y será útil contar con futuros estudios que amplíen la investigación. El diseño empírico de distinguir por un lado el margen extensivo y por otro el intensivo, incluso con las mismas variables departamentales de influencia, permite ofrecer conclusiones diferenciadas.

4.3.3 Percepción de corrupción

Una condición departamental que ofreció resultados mixtos en el análisis del margen extensivo es la alta percepción de corrupción en diversos departamentos peruanos. Es posible que los resultados mixtos (significativos vs. no significativos) en el margen extensivo sirvan como precedente para evaluar un impacto más claro de la percepción de corrupción en el margen intensivo. Si hay regiones en que la corrupción es un problema más notorio, ello puede llevar a impactos en la forma directa en que cada contribuyente se comporta al reportar sus impuestos al fisco.

La tabla 8 analiza formalmente la existencia y frecuencia de omisiones de declaración del impuesto a la renta. Sin embargo, en este caso la variable explicativa de interés, alta percepción de corrupción, no tiene un impacto estadísticamente significativo en el margen intensivo de tributación. Si bien el mecanismo teórico de corrupción generalizada podría suponer un impacto positivo, éste no se refleja en los resultados del diseño empírico estimado. Los resultados emergen de la metodología y el repositorio de datos empleados, y convendrá continuar investigando las relaciones de interés en futuros estudios.

Considerando la posibilidad de que la percepción de corrupción no influya independientemente en un modelo de omisiones tributarias sino que dependa del nivel de inversión pública presente en la región, en un modelo de regresión adicional, no tabulado aquí para ahorrar espacio, se incluyó ambas variables separadamente (inversión pública y corrupción) en el análisis, así como su interacción. Sin embargo, luego de esta ampliación, el

impacto de la percepción de corrupción en la omisión de la declaración del impuesto a la renta continuó resultando no significativo estadísticamente tanto de manera independiente como en la interacción de variables.

4.3.4 Confianza en SUNAT

El estudio ha documentado anteriormente que la alta confianza en la autoridad tributaria por parte de las poblaciones departamentales no es un factor en la decisión de formar parte de la base tributaria. Retomando esta variable departamental, esta vez como posible explicación del margen intensivo, la tabla 9 reporta los resultados del análisis. Nuevamente, no se halla el poder explicativo usual para el coeficiente de alta confianza en SUNAT en el primer y segundo modelo de la tabla.

4.4 Resultados de regresiones sobre el margen intensivo usando deuda coactiva

4.4.1 Educación

Hasta este punto, los modelos han hecho uso de información completa, tanto en forma de corte transversal como en estructuras de panel de datos. A continuación, para terminar el análisis del margen intensivo de la tributación, los modelos utilizan data de estructura más limitada: aquella proveniente del corte transversal de balances de deuda coactiva impagos a fin de diciembre 2018 generados en los períodos fiscales del último trimestre de 2018. La dificultad de analizar esta muestra es que no registra toda aquella deuda coactiva generada en el último trimestre del 2018 pero pagada a la SUNAT antes de fin de diciembre del 2018, por ejemplo, entre el 1 de octubre y el 30 de diciembre.

La importancia de evaluar los balances de deuda coactiva reside en encontrar otra dimensión del margen intensivo de la tributación usando información pública. El primer modelo de la tabla 10 evalúa el impacto de la variable departamental de alta educación en una dicotómica para si el contribuyente tiene deuda coactiva al final del 2018. El coeficiente tiene un valor de 0 y un estadístico t de -0.19; por esto, no hay significancia estadística de la educación departamental en explicar la existencia de balance de deuda coactiva.

La segunda columna de la tabla 10 presenta un análisis similar, esta vez usando el logaritmo del monto de la deuda coactiva en soles. El coeficiente es -0.001, con un estadístico t de -0.25, indicando la falta de significancia estadística para esta variable departamental. Estos resultados indican que, si bien la educación tiene impactos en el margen extensivo e intensivo de la tributación medidos, esos impactos no son capturados en el balance de la deuda coactiva a final del 2018.

4.4.2 Percepción de corrupción

De manera análoga al análisis precedente, la tabla 11 presenta modelos reflejando el impacto de la alta percepción de corrupción en la existencia e incidencia de deuda coactiva al final del 2018. El coeficiente de interés tiene estadísticos t iguales a 1.10 y 1.15, indicando así su falta de significancia estadística. No existe evidencia aceptable, por tanto, de un impacto positivo de la percepción de la corrupción en generación de deuda coactiva.

4.4.3 Confianza en SUNAT

Un último factor departamental analizado en las secciones anteriores es la confianza en la autoridad tributaria. La tabla 12 muestra los resultados de cómo impacta esta variable departamental en la existencia de deuda coactiva o en su monto. Nuevamente, los estadísticos t del coeficiente de interés (0.79 y 0.98) se encuentran debajo del mínimo aceptable para

significación estadística. Se concluye que la confianza en SUNAT no impacta en el comportamiento de generación o pago de deuda coactiva a fines del 2018.

4.5 Resultados en el nivel departamental trascendiendo las zonas de distritos borde

El diseño econométrico seguido en todos los modelos de regresiones presentados restringe el análisis a los distritos del borde de cada departamento. Ese diseño, justificado en la sección sobre especificación (3.5), sigue a la literatura metodológica (Lee and Lemieux, 2010). Si bien no es recomendable indagar sobre los resultados en distritos que no sean del borde, ya que no se aborda en ellos el problema de la endogeneidad al ser poco comparables con distritos de otros departamentos, puede ser interesante ver algunos de los resultados principales en una ventana más amplia de observación.

En este sentido, los principales resultados sobre el margen extensivo se refieren a la educación. La tabla 13 implementa los mismos modelos que en el diseño principal, pero permitiendo que entren a la muestra todos los distritos de todos los departamentos, no solamente los distritos borde. Se aprecia que los resultados son muy similares a los obtenidos en la especificación recomendada.

4.6 Resultados de regresiones de robustez de la especificación de polinomios cúbicos

El diseño empírico de la ecuación (1) propone controlar por un polinomio cúbico de distancia al borde departamental. La argumentación de este diseño es tan defendible como otras opciones, por lo cual es legítimo variar la especificación para indagar su robustez.

En modelos no tabulados, todas las regresiones con resultados significativos discutidos en las secciones precedentes fueron rehechas con un polinomio cuadrático. Los resultados hallados en esas regresiones de robustez fueron los mismos que los reportados aquí. Debido a que los polinomios de menor orden son una verificación enfática de que el impacto de la variable departamental no se debe meramente a una distancia al borde departamental, esos resultados cuadráticos son importantes para proponer la robustez de los modelos.

Análogamente, en modelos no tabulados para ahorrar espacio, las regresiones tomaron una especificación con polinomios de cuarto grado. Nuevamente, todos los resultados significativos reportados en las secciones anteriores mantuvieron su significancia e interpretación luego de esta revisión de robustez.

5. Conclusiones

El presente estudio ha reportado el análisis de diferencias en el margen extensivo e intensivo entre todos aquellos contribuyentes peruanos localizados en lados opuestos de los bordes geopolíticos de los departamentos del Perú para el período 2001 al 2018. Para el análisis se ha utilizado información reservada del investigador e información pública proporcionada por la SUNAT del padrón de contribuyentes.

El análisis empieza delineando el contexto empírico en el cual se realiza la recaudación. Un análisis gráfico considera la evolución en el tiempo de las inscripciones y bajas de contribuyentes, la edad de los contribuyentes en el mercado, el tamaño empresarial medido en personal, la concentración geográfica de entidades y la concentración de la distribución geográfica y sectorial del personal empleado o de servicios.

Es en este marco en el que se desenvuelve la Autoridad Tributaria para llevar a cabo sus objetivos de recaudación eficiente. Pero no es un escenario fijo o exógeno a la actividad tributaria: este escenario económico es posiblemente influido y moldeado por políticas

tributarias, como sugiere la figura 7 al considerar la convergencia asintótica a 45 empleados en promedio en las empresas. Este contexto además influye en el análisis al poner de relieve la importancia de controlar en los modelos por dimensiones socioeconómicas y geográficas que pueden tener una influencia paralela en las variables de interés.

Más allá de descubrir patrones sugerentes, el estudio ha implementado un enfoque econométrico que permite obtener algunas conclusiones robustas proponiendo una metodología que se enfoca en modelos causales válidos para distritos situados en el borde de departamentos contiguos, siendo precisamente esos distritos los que permiten abordar metodológicamente el problema de endogeneidad. Dichas conclusiones se reseñan a continuación.

Primero, la educación departamental tiene un efecto inequívocamente negativo en el margen extensivo de la tributación y un efecto inequívocamente negativo en el margen intensivo de la tributación. A mayor educación departamental, menor formalización de empresas y peor comportamiento de reporte del impuesto a la renta entre los contribuyentes ya formalizados. Estas conclusiones se desprenden directamente de la metodología y el repositorio de datos empleados, y convendrá continuar investigando las relaciones de interés en futuros estudios.

Segundo, la inversión pública departamental tiene un efecto inequívocamente negativo en el margen extensivo de la tributación. Los resultados emergen de un diseño empírico pionero, y convendrá continuar investigando las relaciones de interés en otros análisis. Asimismo, la inversión pública tiene un efecto positivo en el reporte adecuado de las declaraciones de impuesto a la renta de los contribuyentes ya formalizados. Este resultado contrasta con el de educación y sugiere un mecanismo en el que la mayor presencia del Estado mediante su gasto público influye de manera diferenciada en aquellos aún considerando formalizarse versus aquellos ya formalizados.

Tercero, la percepción de corrupción tiene cierto efecto negativo en el margen extensivo, aunque este impacto depende de cómo se mida dicho margen. En específico, si el margen se mide usando la decisión de formalización, el impacto es claramente negativo, interpretándose como el deseo de las entidades de permanecer fuera de un sistema formal percibido como corrupto. Pero si el margen extensivo se mide como la prontitud en que los que ya han decidido formalizarse lo hacen con respecto a su inicio de operaciones, la percepción de corrupción no tiene impacto. En general, no se encuentra evidencia de un impacto de la percepción de corrupción en el margen intensivo de la tributación.

Cuarto, no hay evidencia de que la confianza en SUNAT impacte en la prontitud en registrarse entre aquellos que ya decidieron formalizar su negocio mediante la inscripción en el registro de contribuyentes. Tampoco hay evidencia de que la confianza en SUNAT afecte el margen intensivo de la tributación.

Todas las conclusiones se conectan directamente con la metodología y los datos empleados, con amplia evidencia de robustez y transparencia en el análisis. Estas conclusiones deberán contrastarse con aquellas de futuros análisis con metodologías alternativas.

6. Recomendaciones de política

6.1 Políticas de investigación tributaria

Este estudio ofrece resultados de investigación, y al descubrir estos nuevos resultados, es natural preguntarse por más. En este sentido, es apropiado afirmar que la investigación tributaria en el Perú tiene un largo camino por recorrer. La literatura revisada para el caso

peruano es exigua y no brinda un marco inicial adecuado para construir un nuevo repositorio de conocimientos sobre cómo, cuánto, cuándo, dónde y por qué tributan las entidades peruanas. Esta falta de investigación puede ser resultado de las fuerzas de oferta y demanda.

Ya es un paso en la dirección adecuada recurrir a canales como el de los “proyectos cortos tributarios” en alianza con el Consorcio de Investigación Económica y Social; pero parece un paso todavía preliminar. Cabe recomendar una política de mayor apoyo y recursos a la investigación para que la administración tributaria, dentro de la figura más eficiente para sus fines, incluya la cooperación de investigadores experimentados por los canales adecuados.

De manera relacionada, así como un recurso necesario para la investigación tributaria es la experiencia y tiempo del investigador, ese recurso debe venir acompañado de información estadística relevante. En este sentido, siguiendo las normativas de la protección del secreto tributario, la administración tributaria solamente entrega a investigadores externos aquella información que no califica como reservada de acuerdo con el código tributario, y esto es entendible. Una recomendación respecto a esta política de acceso a la información para investigación en el Perú consiste en reconsiderar casos especiales dentro de la normativa de protección de datos, para que investigadores externos puedan estudiar problemas complejos y brindar nuevas alternativas de solución. En un ejemplo para el caso de Estados Unidos, Chetty, Friedman y Saez (2013) utilizan, bajo contrato de confidencialidad, información del Internal Revenue Service de Estados Unidos para su investigación luego publicada en una revista científica. La presente recomendación exhorta a que futuros desarrollos de investigación en el Perú se enriquezcan con el acceso a ese tipo de información, con las normativas especiales del caso, a investigadores bajo contrato, para beneficio de la administración tributaria.

6.2 Políticas de regulación al crecimiento empresarial

Este estudio muestra patrones en el crecimiento de la fuerza laboral promedio dentro de la empresa que sorprenden y no parecen estar motivados por la dinámica normal de las empresas peruanas. Por ejemplo, el estancamiento del número promedio de empleados en menos de 45 personas por empresa en promedio y el declive en el número promedio de prestadores de servicios de 8 a 6 por empresa luego de haber mostrado una tendencia anual alcista son notorios y suceden alrededor del año 2013. Esta tendencia debe ser del conocimiento de la administración tributaria y recuerda patrones endógenos en otros países (e.g., Garicano et al., 2016). Una recomendación potencial de política que se desprende de este patrón es el considerar la flexibilización de regulaciones laborales y tributarias que pueden estar limitando el crecimiento de las empresas más eficientes. Con la regulación en mente, las empresas evitan su crecimiento formal y buscan canales alternativos, evitando así la regulación.

Una recomendación de política distinta se obtiene del fuerte incremento en el número de contribuyentes registrados después del 2006. Las poblaciones locales crecen de manera orgánica, pero el número de contribuyentes surge vertiginosamente en algunos de estos lugares. Es por tanto recomendable definir las causas del aumento en el registro nominal de entidades de forma que ese crecimiento venga acompañado de una presión tributaria equilibrada y de amplia cobertura. Hay más contribuyentes nominales, pero no está claro si la presión tributaria está también más diseminada.

El estudio muestra que el dinamismo de las empresas que empezaron hace más de dos décadas ha llevado a muchas de ellas a madurar; la distribución de edades empresariales ha cambiado significativamente en veinte años. Una recomendación de política directamente

relacionada a este patrón versa sobre la distinción en el ciclo de vida de las empresas, proponiendo medidas que sean acordes con su estado de crecimiento y evolución. Sin necesidad de crear incentivos tributarios, se puede considerar de manera diferente a las empresas más jóvenes versus las maduras, llevándolas a crecer sus negocios de manera robusta y a considerar al fisco como un factor más en su forma de hacer negocios dependiendo del ciclo de vida.

6.3 Políticas educativas

El resultado de que la mayor educación en un departamento peruano lleva a menos formalización, menos prontitud en inscripción, y mayor número de omisiones tributarias sugiere políticas claras de mejor educación cívica y tributaria de la población en distintas regiones del Perú.

Por un lado, la decisión de formalización aparece típicamente temprano en el ciclo de vida de un emprendedor o profesional independiente, por lo cual la educación escolar y universitaria respecto a concientización tributaria es pertinente y crucial para una mejora de la actitud. La revisión de los currículos y planes de estudios en instituciones educativas estatales y privadas, y las campañas de difusión de la administración tributaria en esos segmentos de la población, interesan para un impacto efectivo.

Por otro lado, los resultados de puntualidad y respeto a las normas tributarias del contribuyente son relevantes para aquellas entidades que ya están contribuyendo al fisco, por lo cual las políticas institucionales de la administración tributaria dirigidas al contribuyente mediante campañas, servicio de apoyo, recordatorios y flexibilización de procesos llevarán a una cultura de mejor desempeño. Así como un buen empleador no se olvida de pagar la planilla a fin de mes, tampoco debe olvidarse de sus obligaciones tributarias.

Estas campañas en su fase inicial no deben tener un contenido necesariamente de penalización. La educación de la población y la educación del contribuyente pueden mostrar los beneficios de una base tributaria más amplia y más respetuosa de las cláusulas fiscales. Si esos beneficios están claros para la administración tributaria, deben ser expresados de esa forma al público general.

6.4 Políticas de inversión pública

Es imperativo reconciliar el mayor flujo de fondos públicos a localidades del Perú con una mayor exigencia tributaria por parte de los beneficiados. Una recomendación de política, para lograr esa coordinación, es diseñar mecanismos de control y condicionamiento que exijan a las poblaciones locales a cumplir metas de desempeño de las cuales dependa el gasto público futuro. No es razonable que aquellos que reciben más son los que menos desean entrar en la base tributaria, pero ese es un resultado de esta investigación que debe informar la toma de decisiones de organismos de gasto público en el Estado.

Por otra parte, la relación positiva entre inversión pública y mejor cumplimiento de los contribuyentes es reveladora de que una política que enfatice la conexión entre el buen comportamiento en los reportes tributarios y la mayor recepción de fondos para inversión estatal puede reforzar esta buena dinámica. Si los contribuyentes saben que sus aportes se usan efectivamente en sus regiones, ello generará una retroalimentación útil para los fines de recaudación.

6.5 Políticas de imagen institucional

Aunque con menor robustez, los resultados de este estudio encuentran un rol para la percepción de corrupción que puede llevar a la administración tributaria a diferenciarse de otras entidades con problemas de imagen. En específico, reforzar la buena imagen de la SUNAT ante la población y los contribuyentes solidificará la actitud de contribuir efectivamente al fisco.

6.6 Plan de incidencia

Queda en potestad de la administración tributaria el hacer uso de los resultados de este estudio para su organización interna y sus políticas de difusión.

Las bases para los criterios de selección de este proyecto corto tributario indicaron que un plan de incidencia no se aplica, razón por la cual no se propone nada aquí en ese sentido. Debido a que el presente proyecto corto tributario es patrocinado por la SUNAT, queda en la SUNAT proponer de manera coordinada ideas sobre incidencia que sean commensurables y proporcionadas a un proyecto corto como el requerido por las bases.

Referencias

- Agarwal, S., Liu, C., y Souleles, N. (2007). The Reaction of Consumer Spending and Debt to Tax Rebates—Evidence from Consumer Credit Data. *Journal of Political Economy*, 115(6), 986–1019.
- Aguilar, G. (2014). Microfinanzas y Crecimiento Regional en el Perú. *Economía*, 36(72), 143–173.
- Allen, F., Bernardo, A. E., y Welch, I. (2000, Dec.). A theory of dividends based on tax clienteles. *Journal of Finance*, 55(6), 2499–2536.
- Aragón, F., y Rud, J. P. (2013). Natural resources and local communities: Evidence from a Peruvian gold mine. *American Economic Journal: Economic Policy*, 5(2), 1–25.
- Artavanis, N., Morse, A., y Tsoutsoura, M. (2016). Measuring income tax evasion using bank credit: evidence from Greece. *Quarterly Journal of Economics*, 131, 739–798.
- Ayyagari, M., Demirguc-Kunt, A., y Maksimovic, V. (2010). Formal versus informal finance: evidence from China. *Review of Financial Studies*, 23(8), 3048–3097.
- Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics*.
- Bartik, T. J. (1985). Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States. *Journal of Business & Economic Statistics*, 3(1), 14–22.
- Baugh, B., Ben-David, I., y Park, H. (2016). Can taxes shape an industry? evidence from the implementation of the “amazon tax”. *Journal of Finance*, forthcoming.
- Beck, T., Demirguc-Kunt, A., y Maksimovic, V. (2005). Financial and legal constraints to growth: Does firm size matter? *Journal of Finance*, 60(1), 137–177.
- Beck, T., Lin, C., y Ma, Y. (2014, apr). Why do firms evade taxes? The role of information sharing and financial sector outreach. *Journal of Finance*, 69(2), 763–817.
- Bennedsen, M., y Zeume, S. (2017). Corporate Tax Havens and Transparency. *Review of Financial Studies*, forthcoming.
- Bernardo, A. E., y Chowdhry, B. (2002). Resources, real options, and corporate strategy. *Journal of Financial Economics*, 63(2), 211–234.
- Besley, T., y Persson, T. (2013). Taxation and development. En M. F. Alan Auerbach Raj Chetty y E. Saez (Eds.), *Handbook of public economics* (pp. 51–110). Amsterdam: Elsevier.
- Besley, T., y Persson, T. (2014). Why Do Developing Countries Tax So Little? *Journal of Economic Perspectives*, 28(4), 99–120.
- Black, S., y Strahan, P. (2002). Entrepreneurship and bank credit availability. *Journal of Finance*, 57(6), 2807–2833.
- Bloom, N., y Van Reenen, J. (2010). Why do management practices differ across firms and countries? *Journal of Economic Perspectives*, 24(1), 203–224. 10

Brollo, F., Nannicini, T., Perotti, R., y Tabellini, G. (2013). The Political Resource Curse. *American Economic Review*.

Bruce, D. (2002). Taxes and Entrepreneurial Endurance : Evidence from the Self-Employed. *National Tax Journal*, 55(1), 5–24.

Bruce, D., y Deskins, J. (2012). Can state tax policies be used to promote entrepreneurial activity? *Small Business Economics*, 38(4), 375–397.

Bruce, D., y Mohsin, M. (2006). Tax Policy and Entrepreneurship: New Time Series Evidence. *Small Business Economics*, 26(5), 409–425.

Bruhn, M. (2011). License to sell: the effect of business registration reform on entrepreneurial activity in Mexico. *Review of Economics and Statistics*, 93, 382–386.

Buehn, A., y Farzanegan, M. R. (2013). Impact of education on the shadow economy: Institutions matter. *Economics Bulletin*.

Butler, A. W., Fauver, L., y Mortal, S. (2009, jul). Corruption, Political Connections, and Municipal Finance. *Review of Financial Studies*, 22(7), 2873–2905.

Cabral, L. M., y Mata, J. (2003). On the evolution of the firm size distribution: facts and theory. *American Economic Review*, 93(4), 1075–1090.

Card, D. (2001). Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica*, 69(5), 1127–1160.

Casaburi, L., y Troiano, U. (2015). Ghost-house busters: The electoral response to a large anti tax evasion program. *Quarterly Journal of Economics*.

Caselli, F., y Michaels, G. (2013, jan). Do Oil Windfalls Improve Living Standards? Evidence from Brazil. *American Economic Journal: Applied Economics*, 5(1), 208–238.

Céspedes, N., Aquije, M., Sánchez, A., y Vera-Tudela, R. (2014). Productividad sectorial en el Perú : Un análisis a nivel de firmas. *Revista Estudios Económicos*, 28, 9–26.

Chetty, R., Friedman, J., y Saez, E. (2013). Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review*, forthcoming.

Chetty, R., y Saez, E. (2005). Dividend Taxes and Corporate Behavior: Evidence From the 2003 Dividend Tax Cut. *Quarterly Journal of Economics*, 120(3), 791–833.

Chirinko, R., y Wilson, D. (2008). State investment tax incentives: A zero-sum game? *Journal of Public Economics*, 92(12), 2362–2384.

Cummings, R. G., Martinez-Vazquez, J., McKee, M., y Torgler, B. (2009). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior and Organization*, 70(3), 447–457.

DePaula, Á., y Scheinkman, J.A. (2010, oct). Value-Added Taxes, Chain Effects, and Informality. *American Economic Journal: Macroeconomics*, 2(4), 195–221.

DePaula, Á., y Scheinkman, J.A. (2011, may). The informal sector: an equilibrium model and some empirical evidence from Brazil. *Review of Income and Wealth*, 57(May), S8–S26.

- de Soto, H. (1986). *El otro sendero*. Lima: Editorial El Barranco.
- de Soto, H. (2000). *The mystery of capital: Why capitalism triumphs in the West and fails everywhere else*. Basic Books.
- DeBacker, J., Heim, B. T., y Tran, A. (2015). Importing corruption culture from overseas: Evidence from corporate tax evasion in the United States. *Journal of Financial Economics*, 117, 122–138.
- Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics*, 88(9-10), 1697–1720.
- Demirguc-Kunt, A., y Maksimovic, V. (1998). Law, finance, and firm growth. *Journal of Finance*, 53(6), 2107–2137. doi: 10.1111/0022-1082.00084
- Deng, Y., Liu, X., y Wei, S.-J. (2018, may). One fundamental and two taxes: when does a Tobin tax reduce financial price volatility? *Journal of Financial Economics*, 40, 1–76.
- Djankov, S., Ganser, T., McLiesh, C., Ramalho, R., y Shleifer, A. (2010, jul). The Effect of Corporate Taxes on Investment and Entrepreneurship. *American Economic Journal: Macroeconomics*, 2(3), 31–64.
- Do, Q.-A., y Campante, F. R. (2014). Isolated Capital Cities, Accountability and Corruption: Evidence from US States. *American Economic Review*, forthcoming.
- Duggan, M. G. (2000). Hospital ownership and public medical spending. *Quarterly Journal of Economics*, 115(4), 1343–1373.
- Duranton, G., Gobillon, L., y Overman, H. G. (2011, sep). Assessing the effects of local taxation using microgeographic data. *The Economic Journal*, 121(555), 1017–1046.
- Dyreng, S. D., Hanlon, M., y Maydew, E. L. (2008). Long-Run Corporate Tax Avoidance. *The Accounting Review*, 83(1), 61–82.
- Easterly, W. (2005). National policies and economic growth: A reappraisal. En P. Aghion y S. N. Durlauf (Eds.), *Handbook of economic growth* (Vols. 1, Part A, p. 1015 - 1059). Elsevier.
- Escobal, J., y Armas, C. (2018). *Inclusión económica y tributación territorial: el caso de las exoneraciones altoandinas*.
- Escobal, J., y Torero, M. (2003). *Adverse geography and differences in welfare in Peru*.
- Farrell, D. (2004). The hidden dangers of the informal economy. *The McKinsey Quarterly*(4), 36–47.
- Fisman, R., y Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6), 1020–1048.
- Friedman, E., Johnson, S., Kaufmann, D., y Zoido-Lobaton, P. (2000). Dodging the grabbing hand: the determinants of unofficial activity in 69 countries. *Journal of Public Economics*, 76, 459–493.
- Frisancho, V. (2012). Signaling creditworthiness in Peruvian microfinance markets: The role of information sharing. *B.E. Journal of Economic Analysis & Policy*, 12(1).

- Gadenne, L. (2017, jan). Tax Me, but Spend Wisely? Sources of Public Finance and Government Accountability. *American Economic Journal: Applied Economics*, 9(1), 274–314.
- Garicano, L., Lelarge, C., y Van Reenen, J. (2016, nov). Firm size distortions and the productivity distribution: Evidence from France. *American Economic Review*, 106(11), 3439–3479. 12
- Garmaise, M. J., y Moskowitz, T. J. (2003). Informal financial networks: Theory and evidence. *Review of Financial Studies*, 16(4), 1007–1040.
- Garmaise, M. J., y Natividad, G. (2010). Information, the cost of credit and operational efficiency: An empirical study of microfinance. *Review of Financial Studies*, 23(6), 2560–2590.
- Garmaise, M. J., y Natividad, G. (2013). Cheap credit, lending operations, and international politics: The case of global microfinance. *Journal of Finance*, 68(4), 1551–1576.
- Garmaise, M. J., y Natividad, G. (2016). Spillovers in local banking markets. *Review of Corporate Finance Studies*, 5(2), 139–165.
- Garmaise, M. J., y Natividad, G. (2017). Consumer default, credit reporting, and borrowing constraints. *The Journal of Finance*, 72(5), 2331–2368.
- Garmaise, M. J., y Natividad, G. (2018). *Financial flexibility: at what cost?* (Working paper)
- Georgellis, Y., y Wall, H. J. (2006). Entrepreneurship and the Policy Environment. *Federal Reserve Bank of St. Louis Review*(April), 95–112.
- Giroud, X., y Rauh, J. (2016). State taxation and the reallocation of business activity: Evidence from establishment-level data. *Journal of Political Economy*, forthcoming.
- Gonii, E., Lopez, H., y Servén, L. (2008). Fiscal redistribution and income inequality in Latin America. *World Development*, 39(January), 1558–1569.
- Goolsbee, A., y Guryan, J. (2006). The impact of internet subsidies in public schools. *Review of Economics and Statistics*, 88(2), 336–347.
- Graham, J. (2000). How big are the tax benefits of debt? *Journal of Finance*, 55(5), 1901–1941.
- Greenstone, M., Hornbeck, R., y Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118, 536–598.
- Guardado, J. (2018). Office-Selling, Corruption, and Long-Term Development in Peru. *American Political Science Review*, 112, 971–995.
- Guirking, C. (2008). Understanding the coexistence of formal and informal credit markets in Piura, Peru. *World Development*, 36(8), 1436–1452.
- Guiso, Luigi, Sapienza, Paola, y Zingales, Luigi. (2004, jun). The role of social capital in financial development. *The American Economic Review*, 94(3), 526–556.

- Hartner, M., Rechberger, S., Kirchler, E., y Schabmann, A. (2008). Procedural fairness and tax compliance. *Economic Analysis and Policy*, 38(1), 137–152.
- Hernandez, M., y de la Roca, J. (2006). Evasión tributaria e informalidad en el Perú. *Economía y Sociedad* (62), 65–73.
- Hong, J. Y. (2017). How Natural Resources Affect Authoritarian Leaders' Provision of Public Services: Evidence from China. *The Journal of Politics*.
- Hortacsu, A., y Syverson, C. (2007). Cementing relationships: Vertical integration, foreclosure, productivity, and prices. *Journal of Political Economy*, 115(2), 250–301.
- Jackle, A. E., y Li, C. A. (2006). Firm dynamics and institutional participation: A case study on informality of micro enterprises in Peru. *Economic Development and Cultural Change*, 54(3), 557–578.
- Karas, A., Pyle, W., y Schoors, K. (2015). A “de Soto Effect” in Industry? Evidence from the Russian Federation. *Journal of Law and Economics*, 58(2), 451–480. doi: 10.2139/ssrn.1973509
- Karlan, D. (2007). Social connections and group banking. *Economic Journal*, 117, F52–F84.
- Kirchler, E., Hoelzl, E., y Wahl, I. (2008). Enforced versus voluntary tax compliance: The “slippery slope” framework. *Journal of Economic Psychology*, 29(2), 210–225.
- La Porta, R., y Shleifer, A. (2008). The unofficial economy and economic development. *Brookings Papers on Economic Activity*, 2008(2), 275–363.
- La Porta, R., y Shleifer, A. (2014). Informality and development. *The Journal of Economic Perspectives*, 28(3), 109–126.
- Lagakos, D. (2016, apr). Explaining Cross-Country Productivity Differences in Retail Trade. *Journal of Political Economy*, 124(2), 579–620. doi: 10.1086/684718
- Leal-Ordóñez, J. C. (2014). Tax collection, the informal sector, and productivity. *Review of Economic Dynamics*, 17(2), 262–286.
- Lee, D., y Lemieux, T. (2010). Regression discontinuity designs in Economics. *Journal of Economic Literature*, 48, 281—355.
- Levin, J., y Tadelis, S. (2010). Contracting for government services: theory and evidence from US cities. *The Journal of Industrial Economics*, 58(3), 507–541.
- Loayza, N. (1996). The economics of the informal sector: a simple model and some empirical evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45, 129–162.
- Makowski, L., y Ostroy, J. M. (2001). Perfect competition and the creativity of the market. *Journal of Economic Literature*, 39, 479–535.
- Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7), 1159–1178.

- Marion, J., y Muehlegger, E. (2008). Measuring Illegal Activity and the Effects of Regulatory Innovation: Tax Evasion and the Dyeing of Untaxed Diesel. *Journal of Political Economy*, 116(4), 633–666.
- McMillan, J., y Woodruff, C. (1999). Interfirm relationships and informal credit in Vietnam. *Quarterly Journal of Economics*, 114(4), 1285–1320.
- Meghir, C., Narita, R., y Robin, J.-M. (2015). Wages and informality in developing countries. *American Economic Review*, 105(4), 67.
- Miller, M. (1977). Debt and taxes. *Journal of Finance*, 32(2), 261–275.
- Milligan, K., Moretti, E., y Oreopoulos, P. (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics*, 88(9-10), 1667–1695.
- Monteiro, J. C., y Assuncao, J. J. (2012). Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises. *Journal of Development Economics*, 99(1), 105–115.
- Moretti, E., y Wilson, D. (2017, apr). The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists. *American Economic Review*, 107, 1858–1903.
- Mortal, S., y Reisel, N. (2013). Capital allocation by public and private firms. *Journal of Financial and Quantitative Analysis*, 48(1), 77-103. 14
- Natividad, G. (2013). Financial capacity and discontinuous investment: Evidence from emerging market multibusiness firms. *Review of Financial Studies*, 26(9), 2375–2410.
- Natividad, G. (2014). Integration and productivity: Satellite-tracked evidence. *Management Science*, 60(7), 1698–1718.
- Natividad, G. (2016). Quotas, productivity and prices: The case of anchovy fishing. *Journal of Economics & Management Strategy*, 25(1), 220–257.
- Natividad, G. (2018a). *Innovando para exportar: el rol de la geografía en las industrias*. INEI.
- Natividad, G. (2019). Stunted firms: The long-term impacts of colonial taxation. *Journal of Financial Economics*, <https://doi.org/10.1016/j.jfineco.2019.05.003>
- Natividad, G., y Rawley, E. (2016). Interdependence and performance: A natural experiment in firm scope. *Strategy Science*, 1(1), 12–31.
- Olken, B. A., y Pande, R. (2012). Corruption in Developing Countries. *Annual Review of Economics*, 4(1), 479–509.
- Pagés, C. (2010). *La era de la productividad: Cómo transformar las economías desde sus cimientos*. Banco Interamericano de Desarrollo.
- Paravisini, D., Rappoport, V., Schnabl, P., y Wolfenzon, D. (2015). Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. *Review of Economic Studies*, 82, 333–359.

- Rocha, R., Ulyseas, G., y Rachter, L. (2018). Do lower taxes reduce informality? Evidence from Brazil. *Journal of Development Economics*, 134, 28–49.
- Schnabl, P. (2012). The international transmission of bank liquidity shocks: Evidence from an emerging market. *The Journal of Finance*, 67(3), 897–932.
- Scott, J. C. (2010). *The art of not being governed: An anarchist history of upland Southeast Asia*. Yale University Press.
- Slemrod, J., y Yitzhaki, S. (2002). Tax avoidance, evasion, and administration. En A. J. Auerbach y M. Feldstein (Eds.), *Handbook of public economics* (Vol. 3, p. 1423 - 1470). Elsevier.
- Sorenson, O., y Audia, P. G. (2000, Sep.). The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940-1989. *American Journal of Sociology*, 106(2), 424–461.
- Stuart, T., y Sorenson, O. (2003a). The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy*, 32(2), 229–253.
- Stuart, T., y Sorenson, O. (2003b). Liquidity events and the geographic distribution of entrepreneurial activity. *Administrative Science Quarterly*, 48, 175–201.
- Tafenu, E., Herwartz, H., y Schneider, F. (2010). Regional estimates of the shadow economy in Europe. *International Economic Journal*, 24(4), 629–636.
- Tello, M. (2015). ¿Es la informalidad laboral una decisión voluntaria en el Perú? *Economía y Sociedad*, 87, 19–25.
- TePaske, J., y Klein, H. (1982). *The Royal Treasuries of the Spanish Empire in America*. Durham: Duke University Press.
- Ulyseas, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, 91(1), 87–99.
- Ulyseas, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108(8), 2015–2047.
- van der Sluis, J., van Praag, M., y Vijverberg, W. (2004). *Entrepreneurship Selection and Performance: A Meta-analysis of the Impact of Education in Less Developed Countries*.
- Wantchekon, L., Klanja, M., y Novta, N. (2015, may). Education and Human Capital Externalities: Evidence from Colonial Benin. *The Quarterly Journal of Economics*, 130(2), 703–757.
- Wasylenko, M. (1999). Taxation and economic development: The state of the economic literature. *Public Administration and Public Policy*, 72, 309–328.
- Williams, C., y Round, J. (2010). Spatial Variations in the Character of Off-the-Books Entrepreneurship: Lessons from a Study of Contrasting Districts in Moscow. *Entrepreneurship and Small Business*, 10(2).

Williams, C. C., Martinez-Perez, A., y Kedir, A. M. (2017). Informal Entrepreneurship in Developing Economies: The Impacts of Starting Up Unregistered on Firm Performance. *Entrepreneurship: Theory and Practice*, 41(5), 773–799.

Woodruff, C., McKenzie, D., y de Mel, S. (2013). The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5(2), 122–150.

Zidar, O., Serrato, J. C., Morales, E., y Fajgelbaum, P. (2018). State taxes and spatial misallocation. *Review of Economic Studies*.

Tabla 1: Resumen Estadístico

El panel A de esta tabla muestra información en el nivel departamento-año para cada uno de los departamentos (regiones) del Perú entre el 2005 y el 2018 para el caso de la ENAHO, y del 2001 al 2015 para la inversión pública, proveniente del SNIP. Los años de secundaria completados, la percepción de corrupción y la confianza en SUNAT—definida como valores de suficiente y bastante a la pregunta sobre cuánto confía un individuo en la autoridad tributaria—se obtienen de todos los años en que están disponibles en la ENAHO. La inversión pública se define como el gasto público en un departamento en cada año.

La información del panel B al panel E solamente incluye distritos del borde de cada departamento. El panel B resume información de corte transversal en el nivel de empresa para las empresas del CENEC 2008. Tiene RUC (1/0) es una variable binaria. Elevación es calculada en el nivel promedio del distrito usando información satelital. Pendiente es calculada en el nivel promedio del distrito utilizando información satelital. La población en miles es calculada en el nivel distrito-año por el INEI. La edad en años de la empresa proviene del CENEC.

El panel C de esta tabla resume información en el nivel de contribuyente proveniente del padrón RUC de SUNAT. Inscribe RUC en primeros 30 días de actividad es una variable binaria. Elevación y pendiente se calculan con información satelital para el promedio distrital. La población es de INEI. El año es el de inicio de actividades de cada contribuyente registrado en el padrón RUC.

El panel D presenta información en el nivel contribuyente-año para los contribuyentes del padrón RUC. Omitió declaración de impuesto a la renta es una binaria. El logaritmo del número de omisiones de declaración de impuesto a la renta suma uno al número de omisiones para que el logaritmo esté bien definido incluso para aquellos contribuyentes que en un año dado no tengan ninguna omisión.

El panel E resume información para todos los contribuyentes activos en el año 2018 solamente. Tiene deuda coactiva es una binaria igual a uno si en el corte transversal de deuda coactiva de diciembre de 2018, el contribuyente aparece con deuda coactiva con el Tesoro Público correspondiente a los períodos de octubre 2018, noviembre 2018 o diciembre 2018. El logaritmo del monto de la deuda coactiva se toma sobre el valor de uno más el monto para que el logaritmo esté bien definido incluso para aquellos contribuyentes sin deuda coactiva.

	Núm.Obs.	Promedio	Desv.Est.	Mín.	Máx.
<i>Panel A: departamento-año</i>					
Promedio de años de secundaria completados	350	3.10	0.37	2.30	3.80
Promedio de percepción de corrupción	350	0.31	0.15	0.05	0.75
Promedio de confianza en SUNAT	100	0.23	0.06	0.07	0.34
Inversión pública	375	478.31	690.64	12.59	5043.40
Año	350	2012	4.04	2005	2018
<i>Panel B: empresas CENEC</i>					
Tiene RUC (1/0)	279,849	0.56	<i>n.a.</i>	0.00	1.00
Elevación	279,849	705.31	1219.20	3.11	4680.50
Pendiente	279,849	4.06	3.64	1.09	37.21
Población en miles	278,749	162.89	161.87	0.19	582.90
Edad en años	279,849	7.14	7.85	1.00	174.00
<i>Panel C: contribuyentes SUNAT</i>					
Inscribe RUC en primeros 30d actividad (1/0)	3,658,221	0.988	<i>n.a.</i>	0.00	1.00
Elevación	3,658,221	598.87	1139.12	3.11	4680.50
Pendiente	3,658,221	3.79	3.41	1.09	37.21
Población en miles	3,658,221	186.08	189.06	0.18	700.18
Año	3,658,221	2012	5	2002	2018
<i>Panel D: contribuyentes-años SUNAT</i>					
Omitió declaración de IR (1/0)	35,863,030	0.005	<i>n.a.</i>	0.00	1.00
Log (1+núm.omisiones de dec.IR)	35,863,030	0.006	0.10	0.00	2.89
<i>Panel E: contribuyentes SUNAT 2018-4to.trim.</i>					
Tiene deuda coactiva (1/0)	3,790,474	0.004	<i>n.a.</i>	0.00	1.00
Log (1+monto de deuda coactiva)	3,790,474	0.028	0.45	0.00	14.86

Tabla 2: Educación y Margen Extensivo de Base Tributaria

Las observaciones están en el nivel empresa o contribuyente utilizando las fuentes descritas en la Tabla 1. El primer modelo introduce información solamente del CENEC 2008. El segundo modelo introduce información del padrón RUC de SUNAT para todos los contribuyentes. Alta educación es una variable dicotómica igual a uno para todos aquellos departamentos-año con un valor de promedio de años de secundaria completados superior a la mediana nacional para ese año. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Tiene RUC (1/0)	Inscribe RUC en primeros 30 días de actividad
	(2.1)	(2.2)
Alta educación (1/0)	-0.487*** (-9.78)	-0.004*** (-3.19)
Elevación	-0.000** (-2.42)	0.000 (0.54)
Pendiente	-0.008*** (-2.73)	-0.000 (-0.54)
Población	-0.000** (-2.45)	0.000*** (3.47)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	No
Efectos fijos de año	No	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.20	0.01
Núm. observaciones	278749	3234688
Núm. clusters (distritos)	749	751

Tabla 3: Inversión Pública y Margen Extensivo de Base Tributaria

Las observaciones están en el nivel empresa o contribuyente utilizando las fuentes descritas en la Tabla 1. El primer modelo introduce información solamente del CENEC 2008. El segundo modelo introduce información del padrón RUC de SUNAT para todos los contribuyentes. Alta inversión pública es una variable dicotómica igual a uno para todos aquellos departamentos-año con un valor de promedio de inversión pública superior a la mediana nacional para ese año. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Tiene RUC (1/0)	Inscribe RUC en primeros 30 días de actividad
	(3.1)	(3.2)
Alta inversión pública (1/0)	-0.393*** (-5.74)	-0.002** (-2.37)
Elevación	-0.000** (-2.40)	0.000 (0.23)
Pendiente	-0.008*** (-2.77)	-0.000 (-0.55)
Población	-0.000** (-2.45)	0.000*** (3.10)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	No
Efectos fijos de año	No	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.20	0.01
Núm. observaciones	278749	2646805
Núm. clusters (distritos)	749	751

Tabla 4: Percepción de Corrupción y Margen Extensivo de Base Tributaria

El primer modelo introduce información solamente del CENEC 2008. El segundo modelo introduce información del padrón RUC de SUNAT para todos los contribuyentes. Alta percepción de corrupción es una variable dicotómica igual a uno para todos aquellos departamentos-año con un valor de promedio de percepción de corrupción superior a la mediana nacional para ese año. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Tiene RUC (1/0) (4.1)	Inscribe RUC en primeros 30 días de actividad (4.2)
Alta percepción de corrupción (1/0)	-0.437*** (-7.17)	-0.001 (-1.49)
Elevación	-0.000** (-2.42)	0.000 (0.53)
Pendiente	-0.008*** (-2.74)	-0.000 (-0.52)
Población	-0.000** (-2.45)	0.000*** (3.47)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	No
Efectos fijos de año	No	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.20	0.01
Núm. observaciones	278749	3234688
Núm. clusters (distritos)	749	751

Tabla 5: Confianza en SUNAT y Margen Extensivo de Base Tributaria

El modelo introduce información del padrón RUC de SUNAT para todas las empresas. (No es posible analizar la información del CENEC en el contexto de la confianza en SUNAT debido a que las encuestas sobre confianza en SUNAT empezaron después del CENEC 2008). Alta confianza en SUNAT es una variable dicotómica igual a uno para todos aquellos departamentos-año con un valor de promedio de confianza en SUNAT superior a la mediana nacional para ese año. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variable dependiente:
	Inscribe RUC en primeros 30 días de actividad
	(5.1)
Alta confianza en SUNAT (1/0)	-0.001 (-1.12)
Elevación	0.000 (0.44)
Pendiente	0.000 (0.24)
Población	0.000*** (2.96)
Polinomio cúbico	Sí
Efectos fijos de segmento geográfico	Sí
Efectos fijos de año	Sí
Efectos fijos de industria	Sí
R^2	0.01
Núm. observaciones	1231718
Núm. clusters (distritos)	751

Tabla 6: Educación y Margen Intensivo de Omisiones Tributarias

Las observaciones están en el nivel contribuyente-año, según se describe en la Tabla 1. En paréntesis se reportan los estadísticos *t*. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Omitió declaración de impuesto a la renta (1/0)	Logaritmo de 1+número de omisiones declaración imp.renta.
	(6.1)	(6.2)
Alta educación (1/0)	0.001** (2.35)	0.001* (1.81)
Elevación	-0.000** (-2.00)	-0.000** (-1.97)
Pendiente	-0.000 (-0.74)	-0.000 (-0.34)
Población	-0.000 (-1.27)	-0.000 (-1.38)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	Sí
Efectos fijos de año	Sí	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.01	0.01
Núm. observaciones	31729121	31729121
Núm. clusters (distritos)	751	751

Tabla 7: Inversión Pública y Margen Intensivo de Omisiones Tributarias

Las observaciones están en el nivel contribuyente-año, según se describe en la Tabla 1. En paréntesis se reportan los estadísticos *t*. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Omitió declaración de impuesto a la renta (1/0)	Logaritmo de 1+número de omisiones declaración imp.renta.
	(7.1)	(7.2)
Alta inversión pública (1/0)	-0.001*** (-3.48)	-0.002*** (-3.38)
Elevación	-0.000** (-2.09)	-0.000** (-2.02)
Pendiente	0.000 (0.37)	0.000 (1.10)
Población	-0.000 (-1.49)	-0.000 (-0.88)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	Sí
Efectos fijos de año	Sí	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.01	0.01
Núm. observaciones	25246699	25246699
Núm. clusters (distritos)	751	751

Tabla 8: Percepción de Corrupción y Margen Intensivo de Omisiones Tributarias

Las observaciones están en el nivel contribuyente-año, según se describe en la Tabla 1. En paréntesis se reportan los estadísticos *t*. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Omitió declaración de impuesto a la renta (1/0)	Logaritmo de 1+número de omisiones declaración imp.renta.
	(8.1)	(8.2)
Alta percepción de corrupción (1/0)	-0.000 (-1.24)	-0.000 (-0.88)
Elevación	-0.000** (-1.98)	-0.000* (-1.96)
Pendiente	-0.000 (-0.75)	-0.000 (-0.35)
Población	-0.000 (-1.27)	-0.000 (-1.38)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	Sí
Efectos fijos de año	Sí	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.01	0.01
Núm. observaciones	31729121	31729121
Núm. clusters (distritos)	751	751

Tabla 9: Confianza en SUNAT y Margen Intensivo de Omisiones Tributarias

Las observaciones están en el nivel contribuyente-año, según se describe en la Tabla 1. En paréntesis se reportan los estadísticos *t*. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Omitió declaración de impuesto a la renta (1/0)	Logaritmo de 1+número de omisiones declaración imp.renta.
	(9.1)	(9.2)
Alta confianza en SUNAT (1/0)	-0.000 (-0.16)	-0.000 (-0.01)
Elevación	-0.000* (-1.82)	-0.000** (-2.26)
Pendiente	-0.000** (-2.02)	-0.000** (-2.05)
Población	0.000*** (3.50)	0.000*** (3.86)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	Sí
Efectos fijos de año	Sí	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.01	0.01
Núm. observaciones	13190522	13190522
Núm. clusters (distritos)	751	751

Tabla 10: Educación y Margen Intensivo de Deuda Coactiva (4to. Trimestre de 2018)

Las observaciones son de corte transversal para contribuyentes activos en el 2018, según se describe en la Tabla 1. En paréntesis se reportan los estadísticos *t*. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Tiene deuda coactiva (1/0)	Logaritmo de 1+monto de deuda coactiva
	(10.1)	(10.2)
Alta educación (1/0)	-0.000 (-0.19)	-0.001 (-0.25)
Elevación	0.000 (0.34)	0.000 (0.25)
Pendiente	-0.000 (-0.11)	-0.000 (-0.07)
Población	-0.000*** (-3.59)	-0.000*** (-3.85)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.01	0.01
Núm. observaciones	3685532	3685532
Núm. clusters (distritos)	751	751

Tabla 11: Percepción de Corrupción y Margen Intensivo de Deuda Coactiva (4to. Trimestre de 2018)

Las observaciones son de corte transversal para contribuyentes activos en el 2018, según se describe en la Tabla 1. En paréntesis se reportan los estadísticos *t*. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Tiene deuda coactiva (1/0)	Logaritmo de 1+monto de deuda coactiva
	(11.1)	(11.2)
Alta percepción de corrupción (1/0)	0.001 (1.10)	0.009 (1.15)
Elevación	0.000 (0.34)	0.000 (0.25)
Pendiente	-0.000 (-0.12)	-0.000 (-0.07)
Población	-0.000*** (-3.59)	-0.000*** (-3.85)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.01	0.01
Núm. observaciones	3685532	3685532
Núm. clusters (distritos)	751	751

Tabla 12: Confianza en SUNAT y Margen Intensivo de Deuda Coactiva (4to. Trimestre de 2018)

Las observaciones son de corte transversal para contribuyentes activos en el 2018, según se describe en la Tabla 1. En paréntesis se reportan los estadísticos *t*. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Tiene deuda coactiva (1/0)	Logaritmo de 1+monto de deuda coactiva
	(12.1)	(12.2)
Alta confianza en SUNAT (1/0)	0.001 (0.79)	0.009 (0.98)
Elevación	0.000 (0.35)	0.000 (0.26)
Pendiente	-0.000 (-0.13)	-0.000 (-0.08)
Población	-0.000*** (-3.59)	-0.000*** (-3.85)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.01	0.01
Núm. observaciones	3685532	3685532
Núm. clusters (distritos)	751	751

Tabla 13: Educación y Margen Extensivo de Base Tributaria en Todos los Distritos del Departamento

Las observaciones están en el nivel empresa o contribuyente utilizando las fuentes descritas en la Tabla 1. El primer modelo introduce información solamente del CENEC 2008. El segundo modelo introduce información del padrón RUC de SUNAT para todos los contribuyentes. Alta educación es una variable dicotómica igual a uno para todos aquellos departamentos-año con un valor de promedio de años de secundaria completados superior a la mediana nacional para ese año. Los símbolos *, **, y *** representan el nivel de 10 %, 5 %, y 1 % en significación estadística, respectivamente.

	Variables dependientes:	
	Tiene RUC (1/0)	Inscribe RUC en primeros 30 días de actividad
	(13.1)	(13.2)
Alta educación (1/0)	-0.436*** (-8.02)	-0.004*** (-4.82)
Elevación	-0.000*** (-3.85)	0.000 (0.44)
Pendiente	-0.011*** (-6.53)	0.000 (1.51)
Población	-0.000** (-2.37)	0.000*** (6.42)
Polinomio cúbico	Sí	Sí
Efectos fijos de segmento geográfico	Sí	Sí
Efectos fijos de edad	Sí	No
Efectos fijos de año	No	Sí
Efectos fijos de industria	Sí	Sí
R^2	0.16	0.01
Núm. observaciones	705018	8644565
Núm. clusters (distritos)	1828	1831

Figura 1: Definición de tratamiento por alta educación

La figura presenta los departamentos del Perú según sean catalogados como de alta educación o no. Por simplicidad, el gráfico se basa en un promedio de la variable anual de educación.



Figura 2: Definición de tratamiento por alta confianza en SUNAT

La figura presenta los departamentos del Perú según sean catalogados como de alta confianza en SUNAT o no. Por simplicidad, el gráfico se basa en un promedio de la variable anual de confianza.



Figura 3: Inscripciones y bajas

El primer panel muestra las inscripciones y bajas en el RUC. El segundo y tercer panel muestran las inscripciones por sector económico.

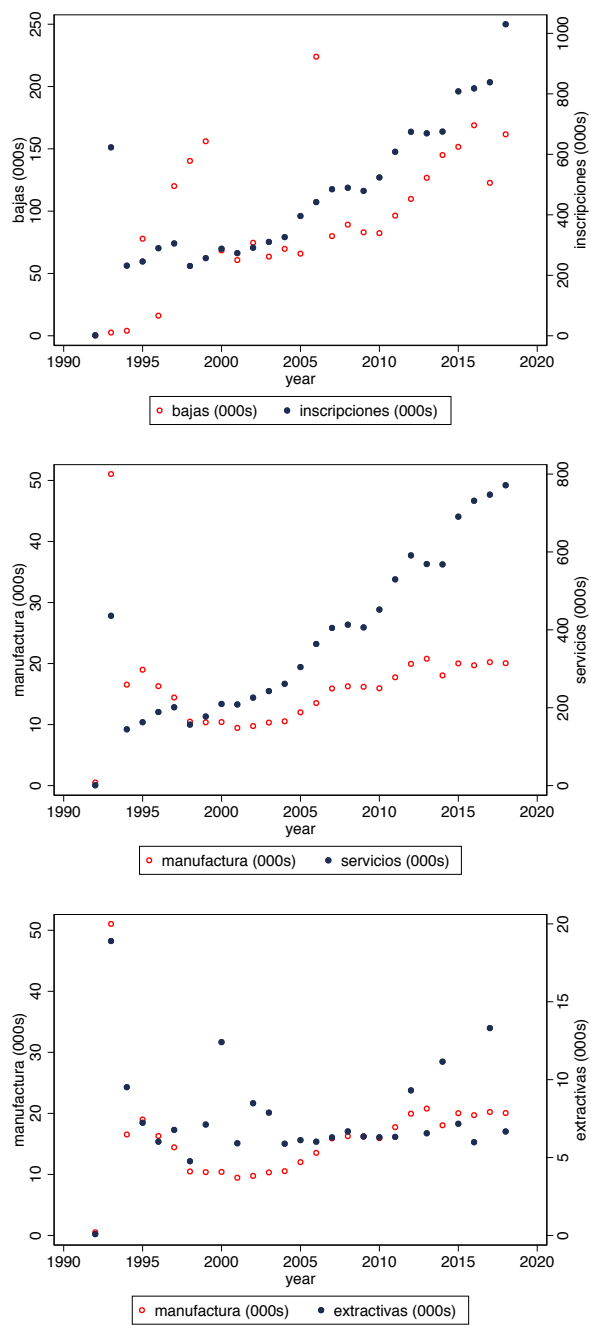


Figura 4: Bajas dentro de los primeros 12 meses

La figura grafica la fracción de contribuyentes inscritos cuyo RUC es dado de baja dentro de los 12 meses transcurridos desde su inscripción.

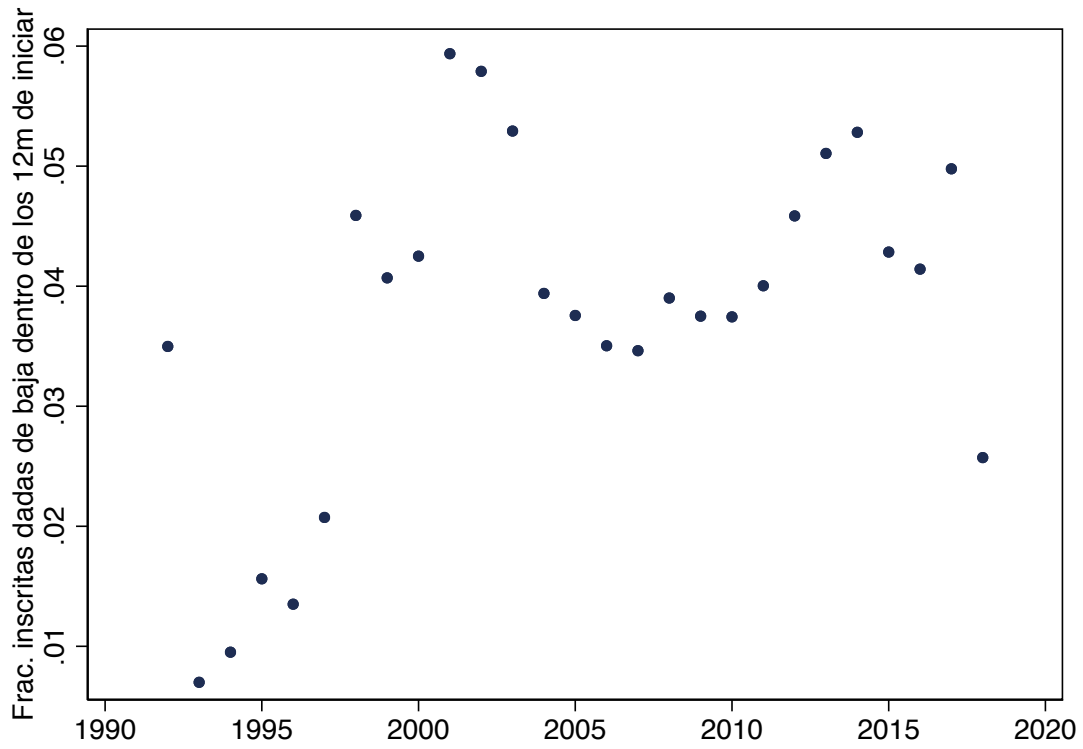


Figura 5: Distribución de edades en 1994, 2000, y 2018

La edad de cada entidad contribuyente está medida en años continuos desde la fecha de inicio de actividades reportada a SUNAT; se omiten edades mayores a 50 años en ambos gráficos para facilidad de lectura

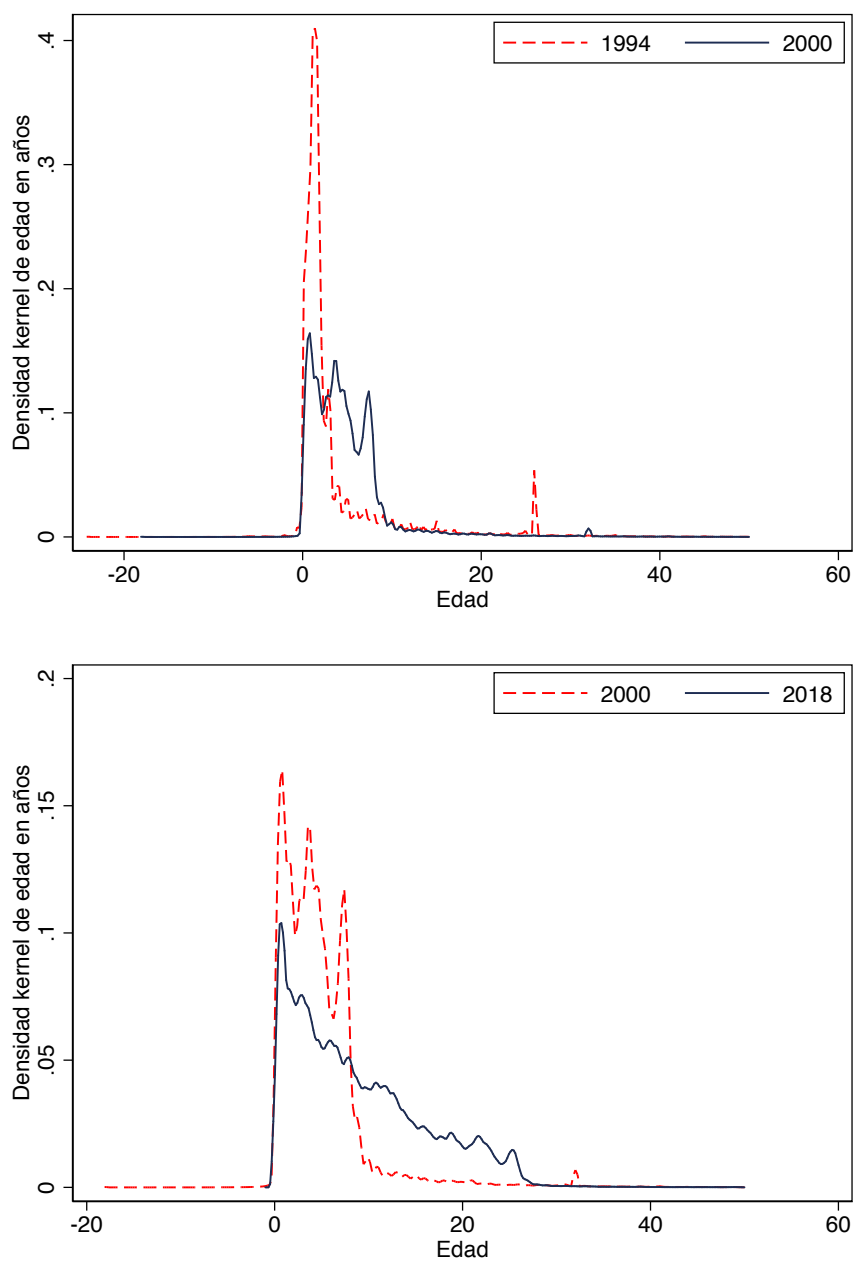


Figura 6: Población y contribuyentes activos excluyendo Lima y Callao

La figura presenta el índice de Herfindahl de la participación en la contratación de empleados para cada industria considerando solamente a las seis industrias más concentradas en el 2008.

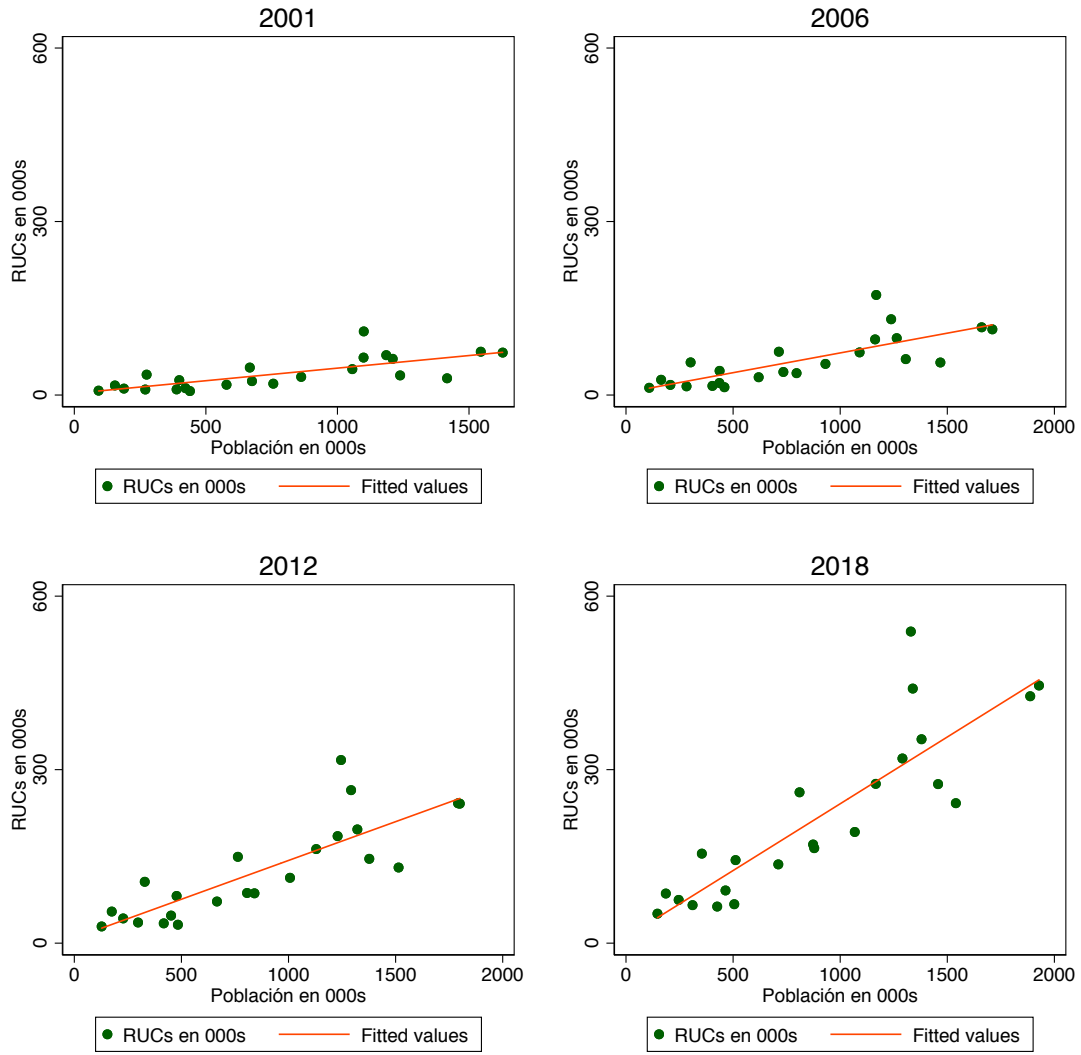


Figura 7: Promedio de empleados y prestadores en contribuyentes que subsisten desde el 2008

La muestra para el cálculo de personal es exclusivamente aquellos contribuyentes con información de empleados o prestadores de servicios que se hallan presentes tanto en el 2008 como en el 2018.

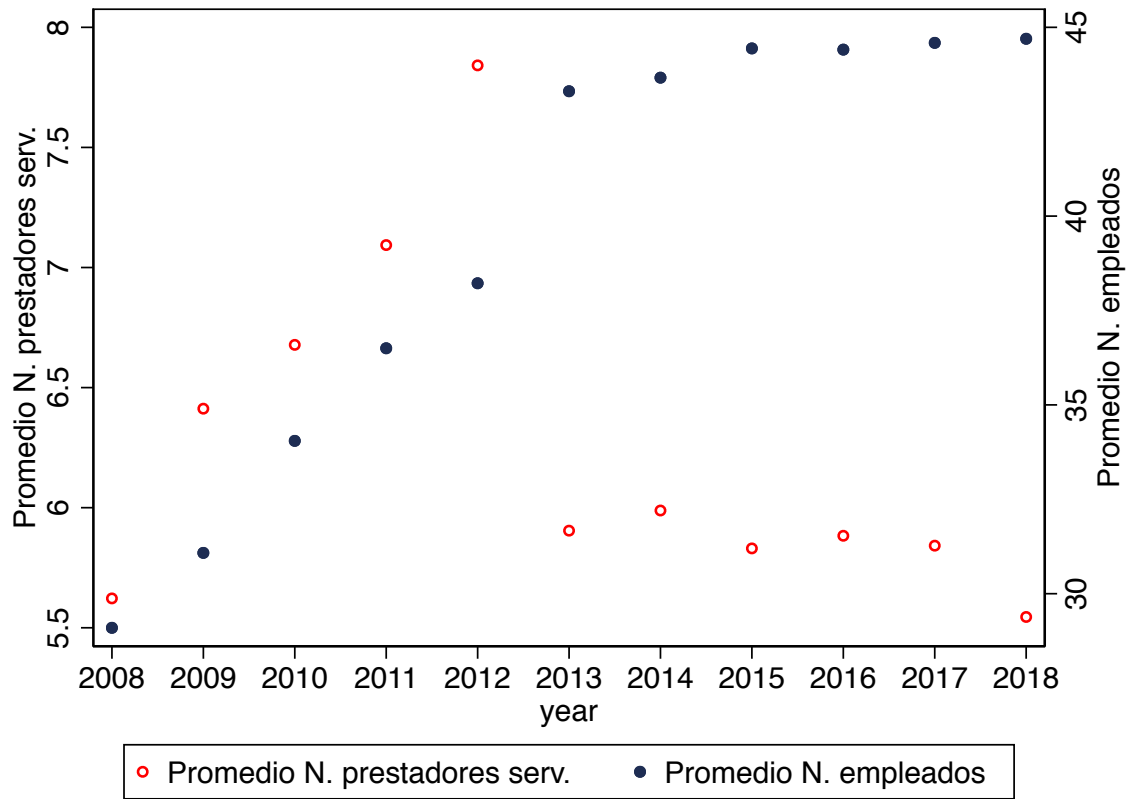


Figura 8: Concentración de personal en departamentos

La figura presenta el índice de Herfindahl anual para la participación de las 25 regiones del Perú (24 departamentos y el Callao) en el número de empleados y prestadores de servicios.

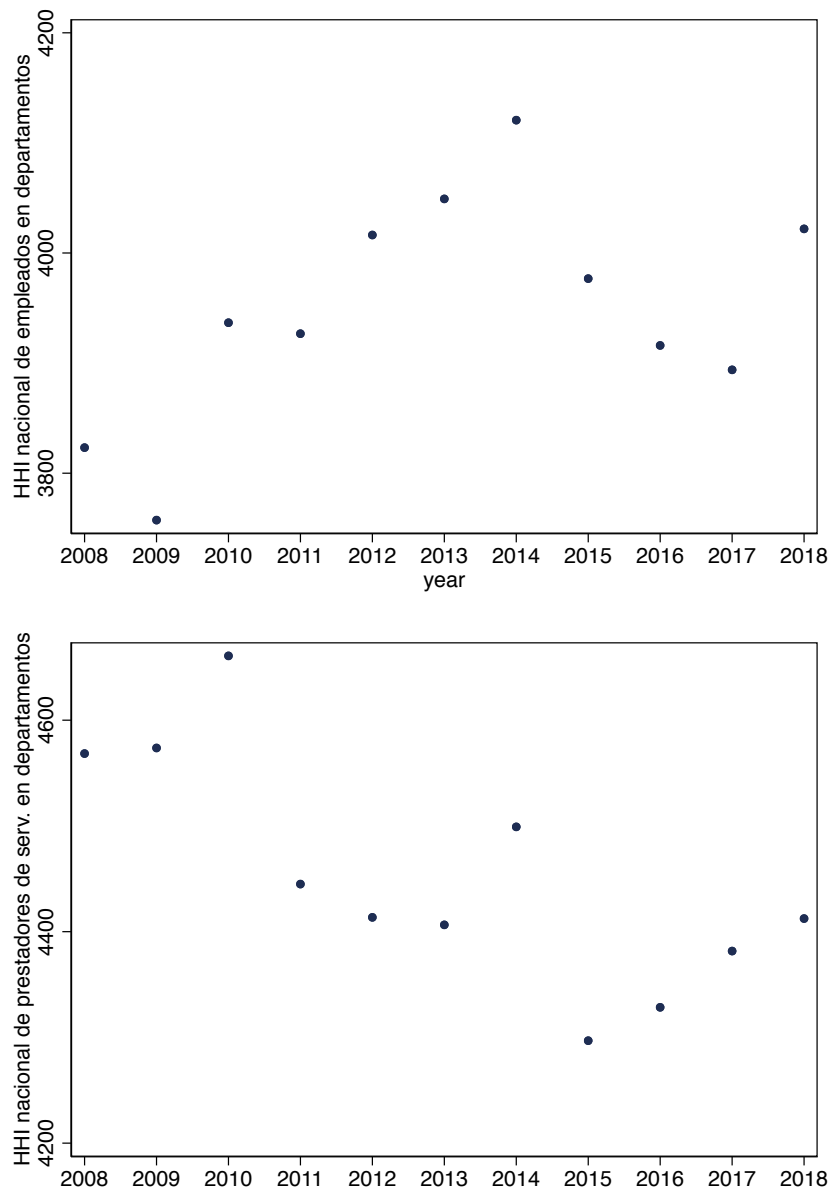


Figura 9: Personal en Lima y Callao como fracción del Perú

La figura presenta la participación de Lima y Callao en el total nacional de empleados y prestadores de servicios.

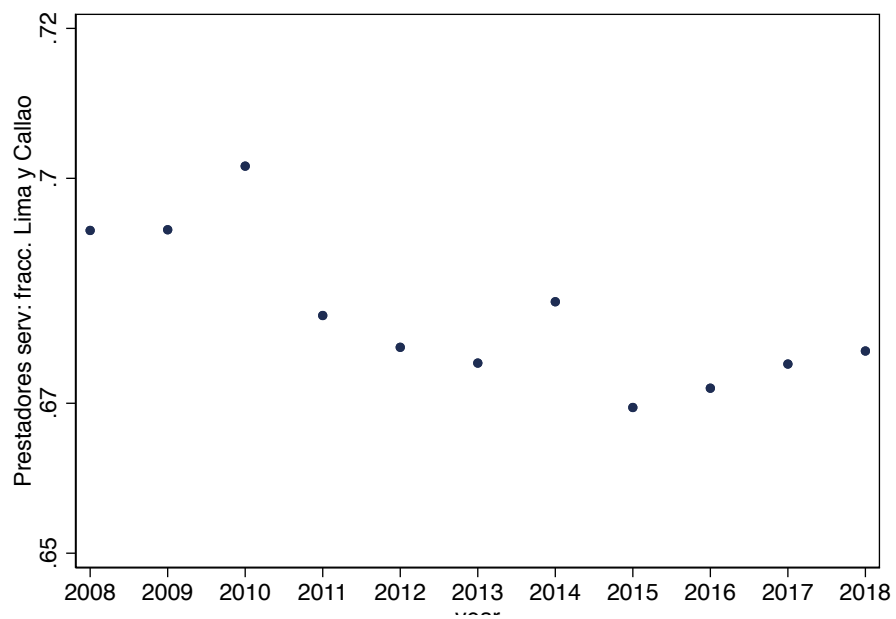
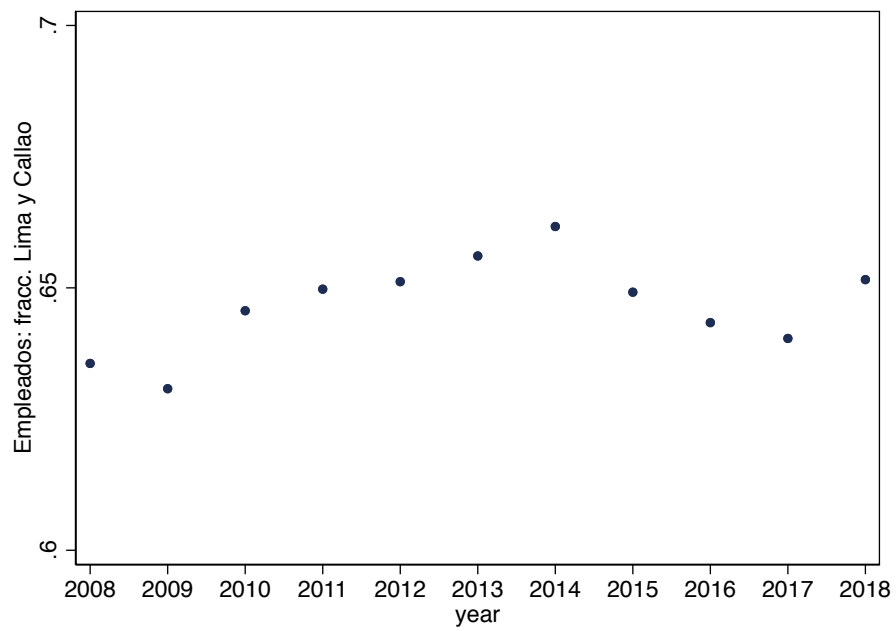


Figura 10: Empleados por sector económico

La figura presenta la participación del total de empleados en los sectores de manufactura, servicios, y actividades extractivas.

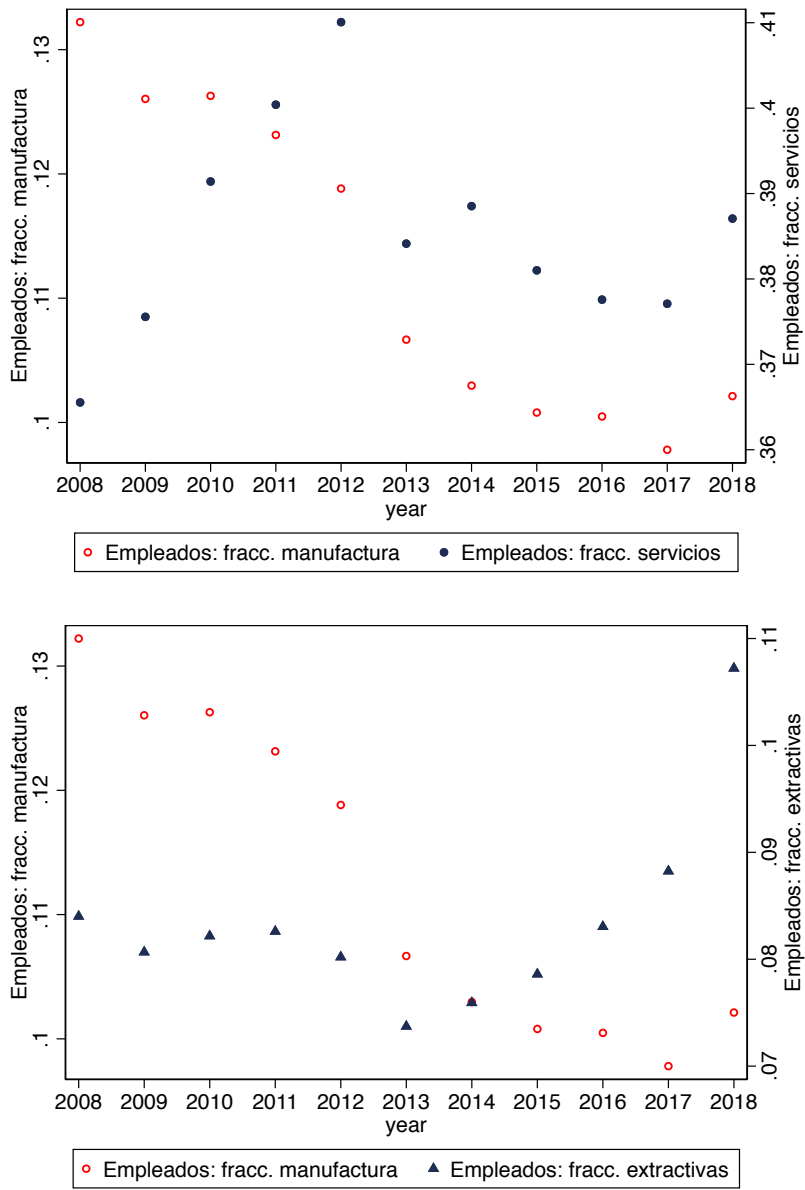


Figura 11: Prestadores de servicios por sector económico

La figura presenta la participación del total de prestadores de servicios en los sectores de manufactura, servicios, y actividades extractivas.

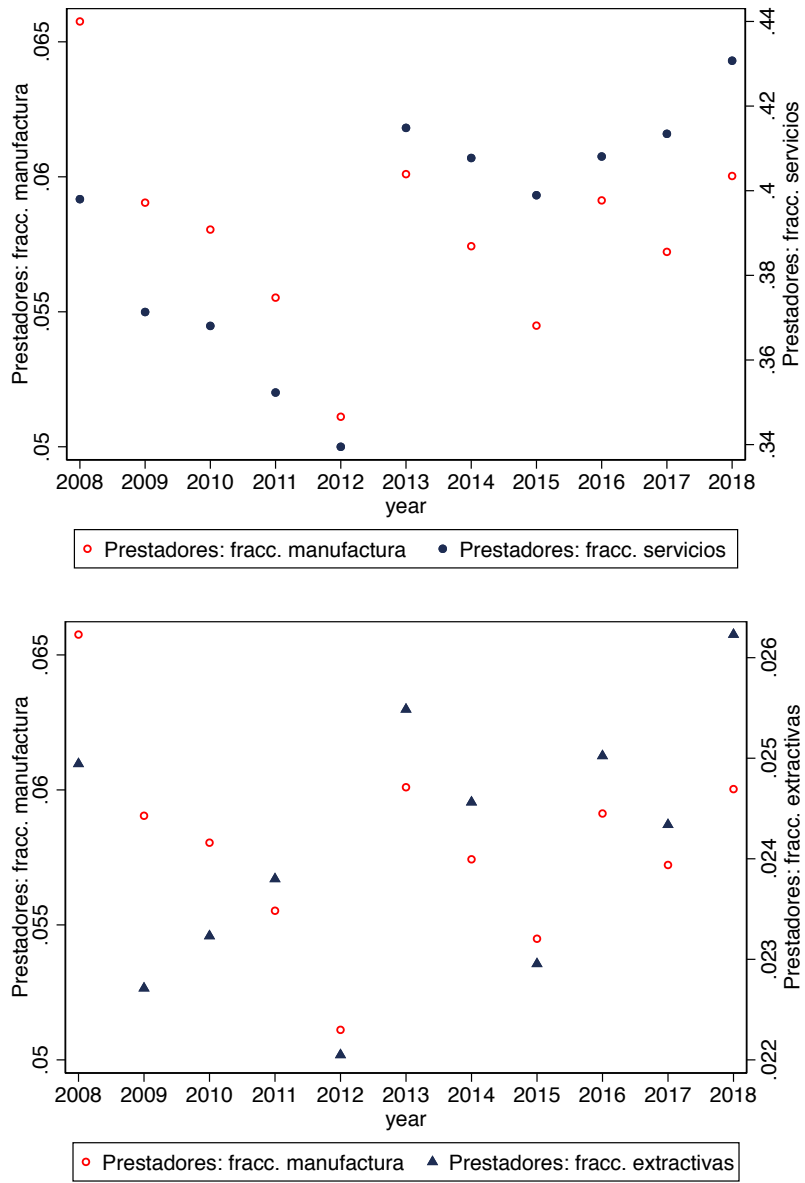


Figura 12: Concentración de empleados en seis industrias más concentradas en 2008

La figura presenta el índice de Herfindahl de la participación en la contratación de empleados para cada industria considerando solamente a las seis industrias más concentradas en el 2008.

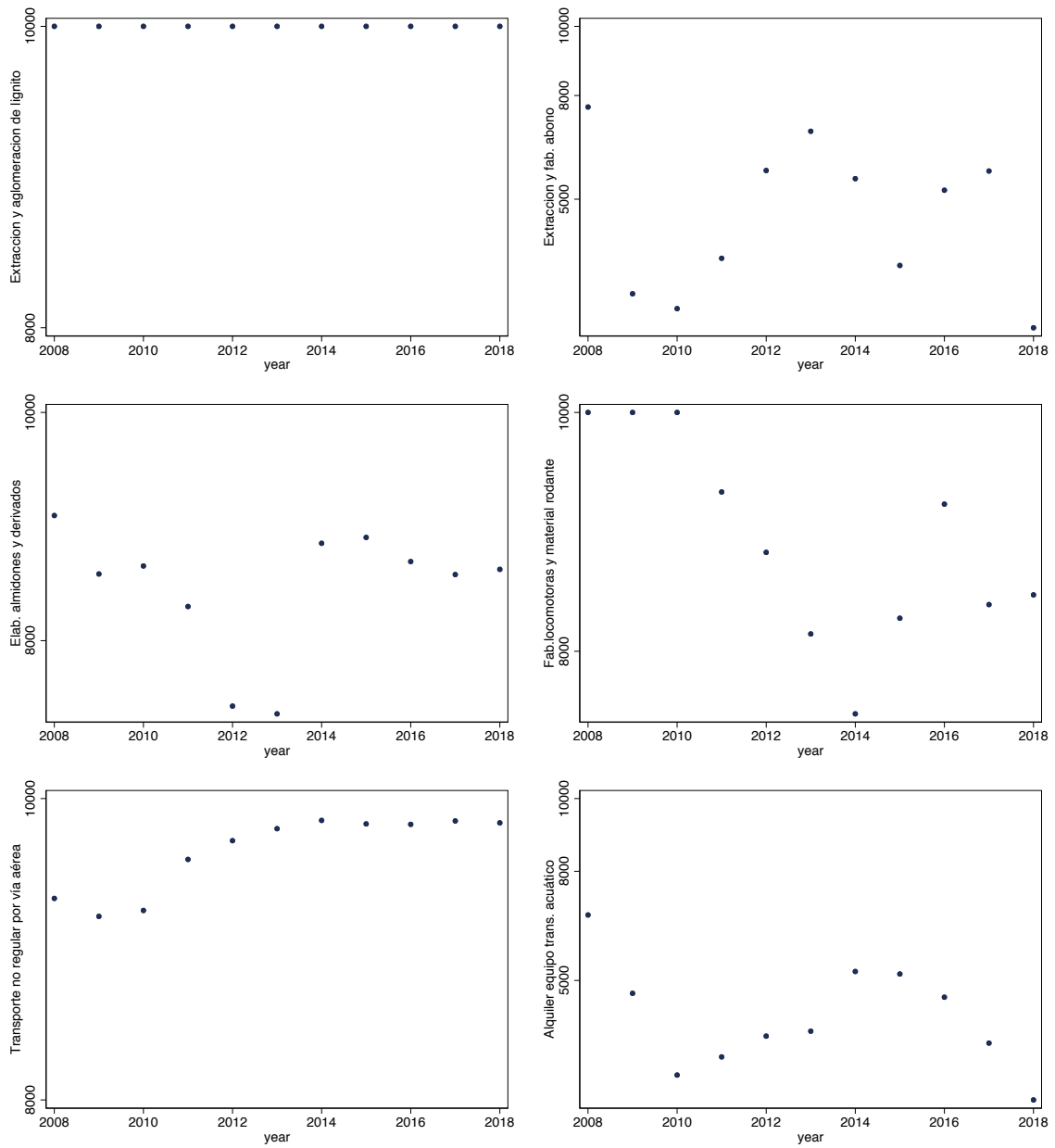


Figura 13: Omisiones tributarias en sectores económicos

La figura presenta la fracción de todas las omisiones anuales en declaración de impuesto a la renta que corresponde los sectores de manufactura, servicios y extracción.

