

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
ESCUELA DE POSGRADO



Determinantes de la investigación para universidades del Perú

TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAGÍSTER EN ECONOMÍA

AUTOR:

Denzel Antonio Glandel Tafur

ASESOR:

José Santos Rodríguez González

Diciembre, 2019

RESUMEN

La relevancia de la generación de conocimiento para el desarrollo de las sociedades, vía el mecanismo de la mejora de productividades, es reconocida en la literatura económica. La situación del Perú en ese aspecto no es positiva, pues se encuentra entre los países con menor producción científica de la región. Este trabajo busca enriquecer la discusión sobre el tema estudiando a las universidades peruanas y haciendo uso de datos novedosos para generar evidencia empírica sobre los factores relacionados con esa producción. Se plantea que estos están ligados a características de los docentes (motivaciones, distribución del tiempo, la calidad de formación en investigación), de las instituciones a las que pertenecen, y del contexto (región, área del conocimiento). Se responden dos preguntas sobre esos factores: cuáles están relacionados a tener al menos una publicación académica indizada, y cuáles al número de esas publicaciones. Para la primera pregunta se utilizan modelos con variable dependiente binaria (Probit y Logit), mientras que para la segunda se recurre a modelos de conteo debido al tipo de distribución de la variable a explicar (Poisson y Binomial Negativo, en sus versiones básicas y otras para excesos de ceros). Entre los hallazgos más relevantes se encuentra que la producción de investigación se relaciona positivamente con tener un doctorado solo si este fue obtenido en una universidad productiva académicamente (en caso contrario la relación es inversa); y con la opinión favorable del docente sobre la importancia de la investigación en su institución. Además, el área del conocimiento es un factor significativo tanto para la probabilidad de tener al menos una publicación como para el número de estas. Así, pertenecer a áreas como Medicina o Ciencias Agrarias incrementa ambos indicadores considerablemente. Estos hallazgos tienen aplicación directa para el diseño de políticas públicas como las condiciones básicas de calidad propuestas por la SUNEDU y los lineamientos e instrumentos de política del CONCYTEC.

Índice

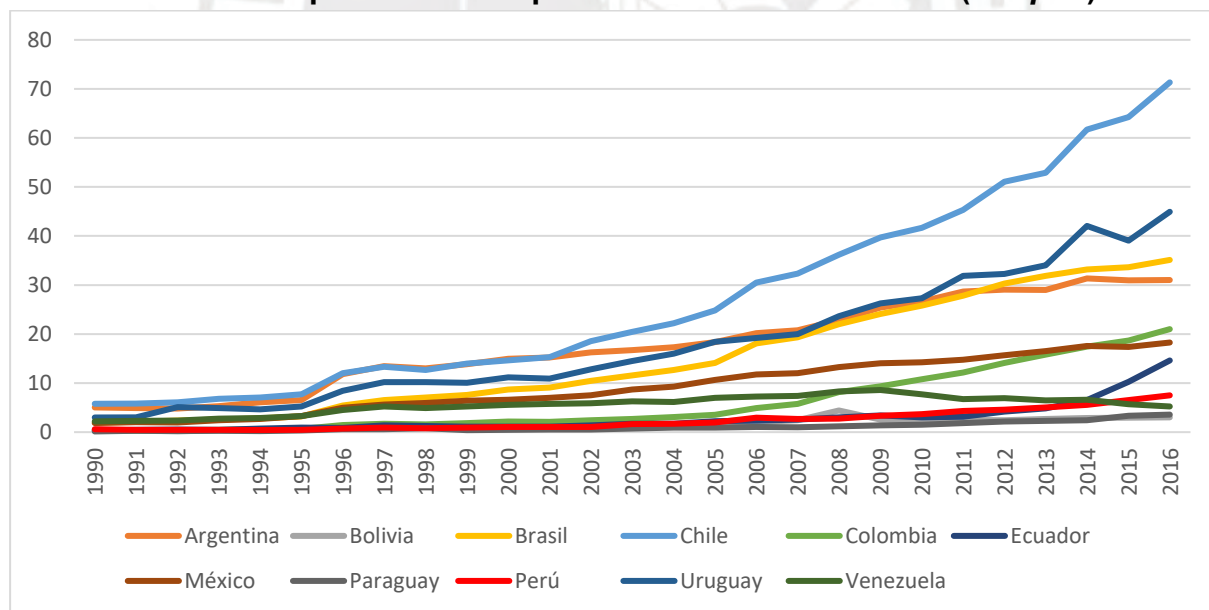
1. INTRODUCCIÓN	4
2. SOBRE LA UNIVERSIDAD EN PERÚ	7
3. MARCO TEÓRICO	14
3.1. Mecanismos entre enseñar e investigar	14
3.2. Motivaciones extrínsecas e intrínsecas	16
3.3. Sobre la medición de la productividad en investigación	17
Medición de publicaciones y tipos de universidades	18
La Ley de Lotka	18
Tipos de Universidades.....	20
A modo de síntesis	21
4. ESTUDIOS EMPÍRICOS PREVIOS	23
5. OBJETIVOS E HIPÓTESIS	32
6. CONSIDERACIONES PARA EL ANÁLISIS EMPÍRICO	34
6.1 De los datos	34
II Censo nacional universitario – 2010 (II CNU).....	36
Base de datos bibliográfica.....	41
6.2 Opciones de modelos econométricos	45
Modelos con variables dependientes binarias: Probit y Logit	46
Modelos con datos de conteo.....	47
6.3 El sesgo de agregación	51
7. ESTIMACIÓN Y RESULTADOS	52
7.1 Variables para la estimación	52
Ingreso por actividad y variables relacionadas	55
Características adscriptivas	56
Características de logro	56
Características institucionales.....	58
Variables dependientes.....	61
7.2 Resultados e interpretaciones	61
Resultados de los modelos con variable resultado dicotómica	62
Resultados de los modelos con variables de conteo sin diferencia de perfiles.....	67
Resultados de los modelos con variables de conteo considerando diferentes perfiles	76
8. CONCLUSIONES	81
9. BIBLIOGRAFÍA	85
ANEXOS	90

1. INTRODUCCIÓN

La generación de conocimiento es fundamental para el desarrollo económico. Romer (1992) resalta la importancia de las ideas como bienes económicos y, por lo tanto, enfatiza la necesidad de su estudio desde la economía para llegar a un entendimiento adecuado del desarrollo y crecimiento de las sociedades. Guellec y van Pottelsberghe (1998) utilizaron datos de 16 países miembros de la OCDE para estimar empíricamente que el incremento en 1% el gasto en investigación y desarrollo produce un incremento de alrededor de 0.15% en la productividad total de factores. A pesar de esa importancia, la producción de investigación en Perú, con respecto a otros países de Latinoamérica, se encuentra notablemente rezagada.

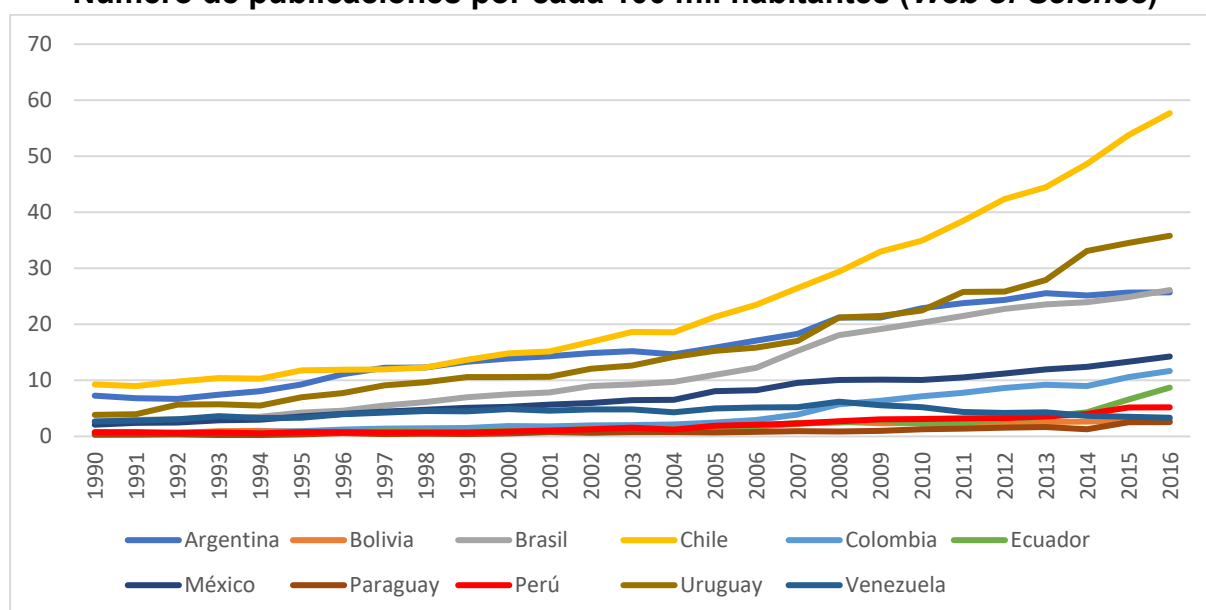
Existen múltiples maneras de medir esa producción (Toutkoushian y Webber, 2011) y una de ellas es por medio del número de publicaciones científicas que produce la sociedad, institución o el investigador analizado. Tomando el volumen de publicaciones como medida de producción, los gráficos 1.1a y 1.1b muestran la evolución de las publicaciones indizadas en *Scopus* y *Web of Science* en 11 países de Latinoamérica. Así, para 2016, Perú se encontraba en el puesto 8 de ese grupo de países.

Gráfico 1.1a
Número de publicaciones por cada 100 mil habitantes (*Scopus*)



Fuente: RICYT. Elaboración propia.

Gráfico 1.1b
Número de publicaciones por cada 100 mil habitantes (Web of Science)



Fuente: RICYT. Elaboración propia.

¿Dónde se produce ese conocimiento? Si bien la investigación tradicionalmente ha sido desarrollada dentro de contextos académicos, como los universitarios, estos no son los únicos lugares en los que se realizan esas actividades (Gibbons, et al., 1994). La teoría de la organización industrial utiliza modelos orientados a estudiar las acciones relacionadas a la Investigación y el Desarrollo que realizan las empresas buscando tener algún efecto en sus respectivas funciones de producción y así mejorar sus rentabilidades. Se estudia también las acciones desde el Estado, como el desarrollo de sistemas de registros de patentes y los beneficios tributarios que podrían tener las empresas que invierten dinero en Investigación y Desarrollo, para incentivarlas a embarcarse en esos proyectos. ¿Cuál es la situación de Perú?

Una revisión del número de publicaciones académicas indizadas en las mismas bases bibliográficas mencionadas en los párrafos anteriores muestra que, para el caso de Perú, alrededor de dos tercios de las publicaciones ha sido producida por investigadores afiliados a universidades.

Cuadro 1.1
Porcentaje de artículos según tipo de institución a la que están afiliados los autores (2008 – 2018)

TIPO DE INSTITUCIÓN	% DE ARTÍCULOS - SCOPUS	% DE ARTÍCULOS - WEB OF SCIENCE
OTRO	36%	30%
UNIVERSIDAD	64%	70%
TOTAL	100%	100%

Fuente: Scopus y Web of Science. Elaboración propia.

Así, resulta fundamental estudiar las universidades con el objetivo de entender factores en estas que están relacionados a la mayor o menor producción de conocimiento. En ese sentido, este trabajo busca aportar significativamente a entender un aspecto fundamental del desarrollo del país. Además, esto se vuelve aún más relevante en la coyuntura que el sistema universitario peruano está atravesando en este momento¹.



¹ La promulgación de la nueva Ley Universitaria 30220 el año 2014 inició el proceso de reforma del sistema universitario peruano. Ese proceso, cuya implementación continúa en la actualidad, refuerza discusiones como el sentido de las universidades y su rol como instituciones generadoras de conocimiento.

2. SOBRE LA UNIVERSIDAD EN PERÚ

El sistema educativo formal en Perú está compuesto por dos etapas, la educación básica y la superior; y por varias modalidades dentro de cada etapa. Dentro de la educación superior se encuentran la técnica, pedagógica, universitaria y técnico-productiva (Jopen, Gomez y Olivera 2014).

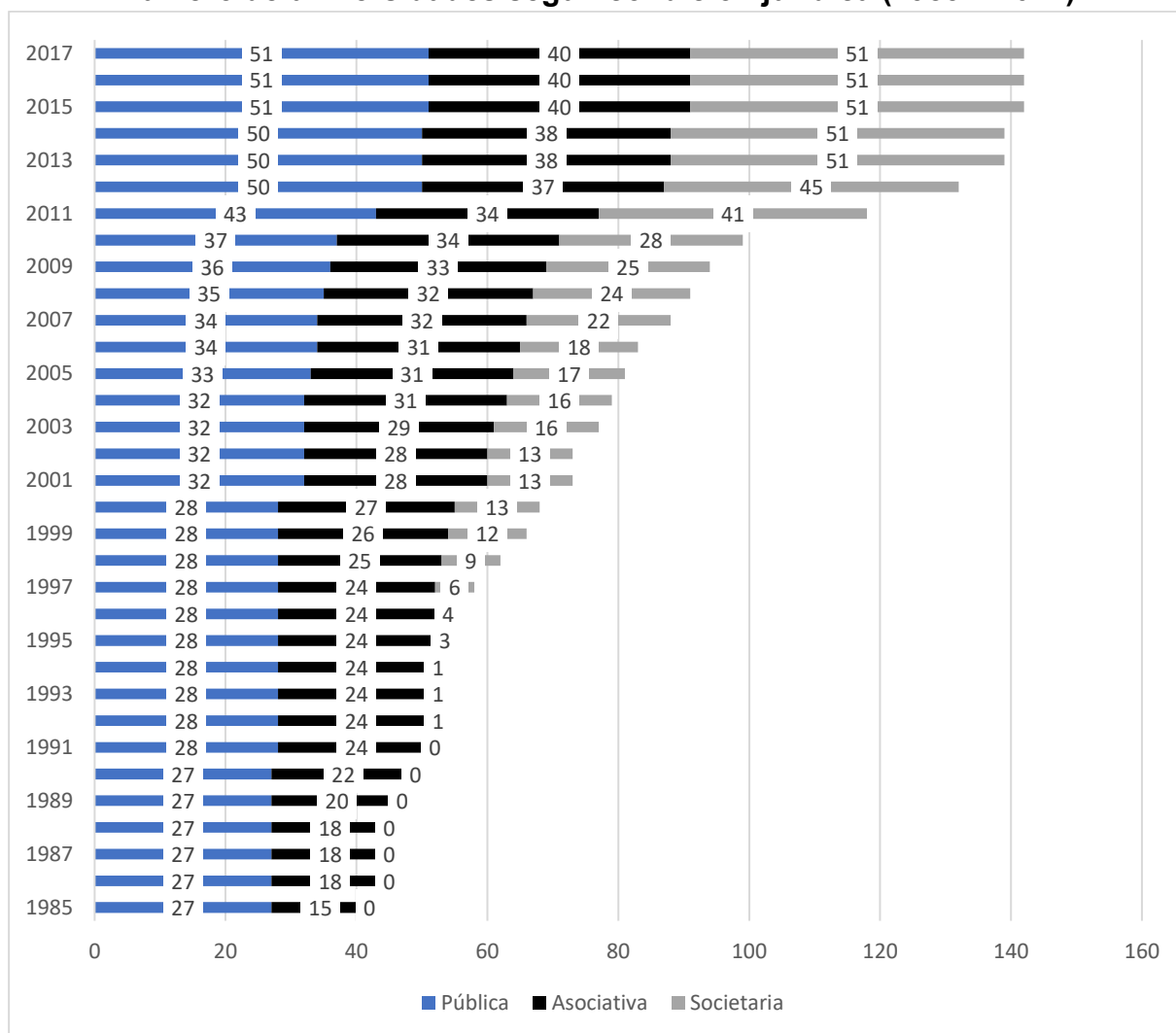
Los dos cambios recientes más importantes en el sistema de educación superior han ocurrido por: i. la liberalización de la educación en 1996 por la publicación del Decreto Legislativo 882 “Ley de Promoción de la Inversión en la Educación”²; y ii. los cambios en el sistema universitario por la nueva Ley Universitaria aprobada en 2014.

Sobre el primer evento, Cuenca y Reategui (2016) señalan que la reforma emprendida por la administración de Alberto Fujimori permitió la creación de instituciones educativas con fines de lucro con el objetivo de atraer la inversión privada al sector y así se amplíe la oferta universitaria; de manera que se incremente la cobertura y se democratice el acceso a la educación superior; además, se esperaba que la calidad del servicio educativo sea regulada por el mercado.

El cumplimiento del primer punto, incremento de cobertura a partir del crecimiento de la oferta, es evidente. El Gráfico 2.1 muestra el crecimiento del número de universidades Societarias (con fines de lucro) a partir de 1996, haciendo que el número total de universidades pase de 55 en 1995 a 142 en 2015.

² Publicado el 9 de noviembre de 1996.

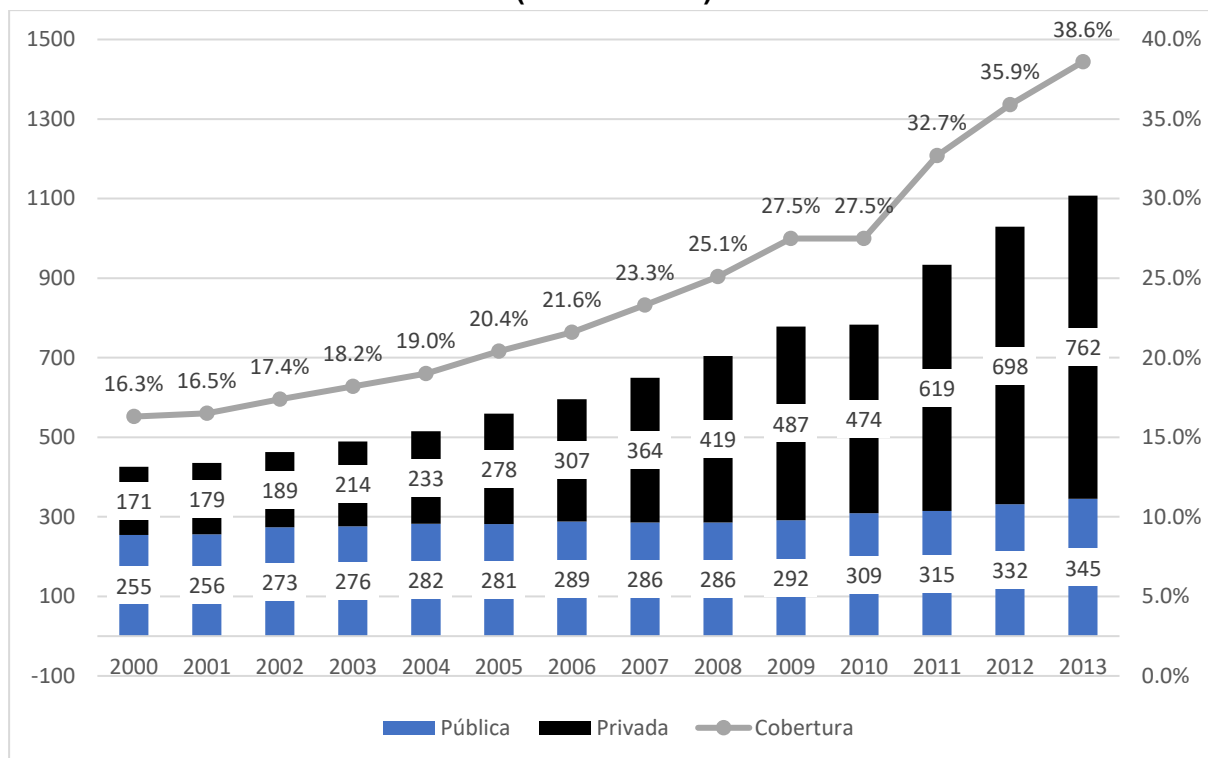
Gráfico 2.1
Número de universidades según condición jurídica (1985 – 2017)



Fuente: SUNEDU. Elaboración propia.

Ese incremento, efectivamente aumentó la cobertura. El siguiente cuadro, que reproduce datos reportados en Cuenca y Reategui (2016), muestra la variación de la cobertura (medida como el porcentaje de jóvenes entre 17 y 31 años matriculados en universidades) del sistema universitario. Así, el valor de esa tasa llegó a más de 38% en 2013. El incremento corresponde en su mayoría a las universidades privadas, que pasaron de representar al 40% de matriculados a representar casi al 70%.

Gráfico 2.2
Matrícula (en miles) según tipo de universidad y tasa de cobertura (%)
(2000 – 2013)

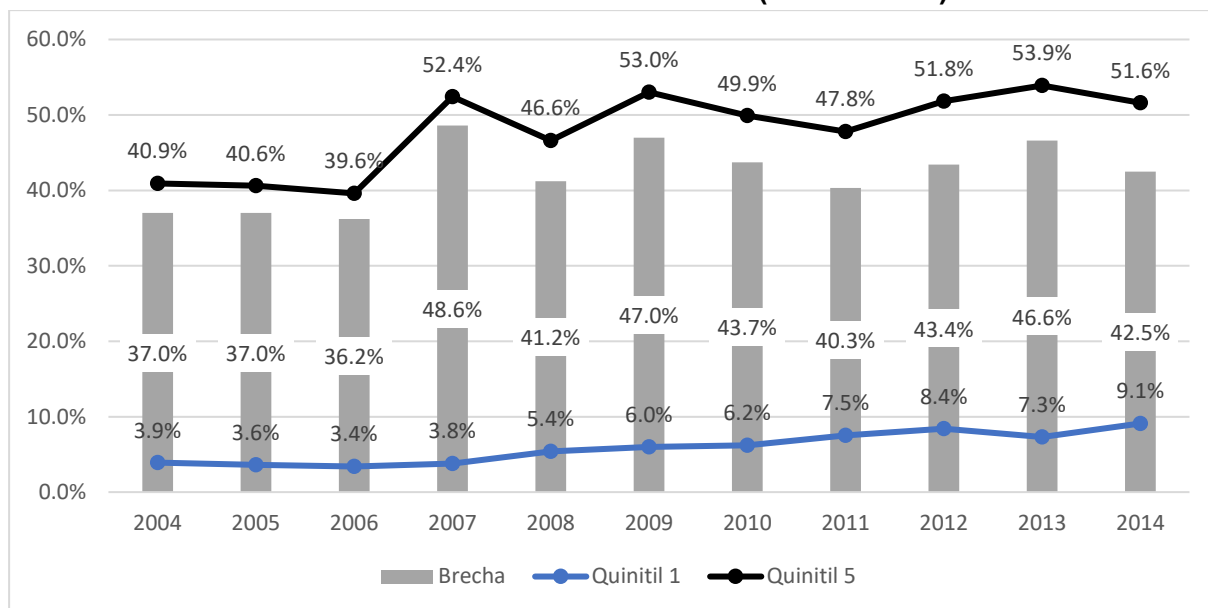


Fuente: INEI y SUNEDU reportados en Cuenca y Reategui (2016). Elaboración propia.

Nota: La cobertura fue obtenida dividiendo el número de matriculados en universidades privadas sobre el número total de jóvenes de entre 17 y 21 años.

Sin embargo, no ocurrió lo mismo con los otros dos puntos esperados: la democratización del acceso y la regulación de la calidad educativa por el mercado. Los gráficos 2.3a y 2.3b muestran que el incremento en el porcentaje de jóvenes matriculados en las universidades ha sido dispar según quintiles de ingreso familiar; y que la evolución proporción de matriculados en universidades, según lengua materna, también es diferente entre Castellano y Quechua o Aymara.

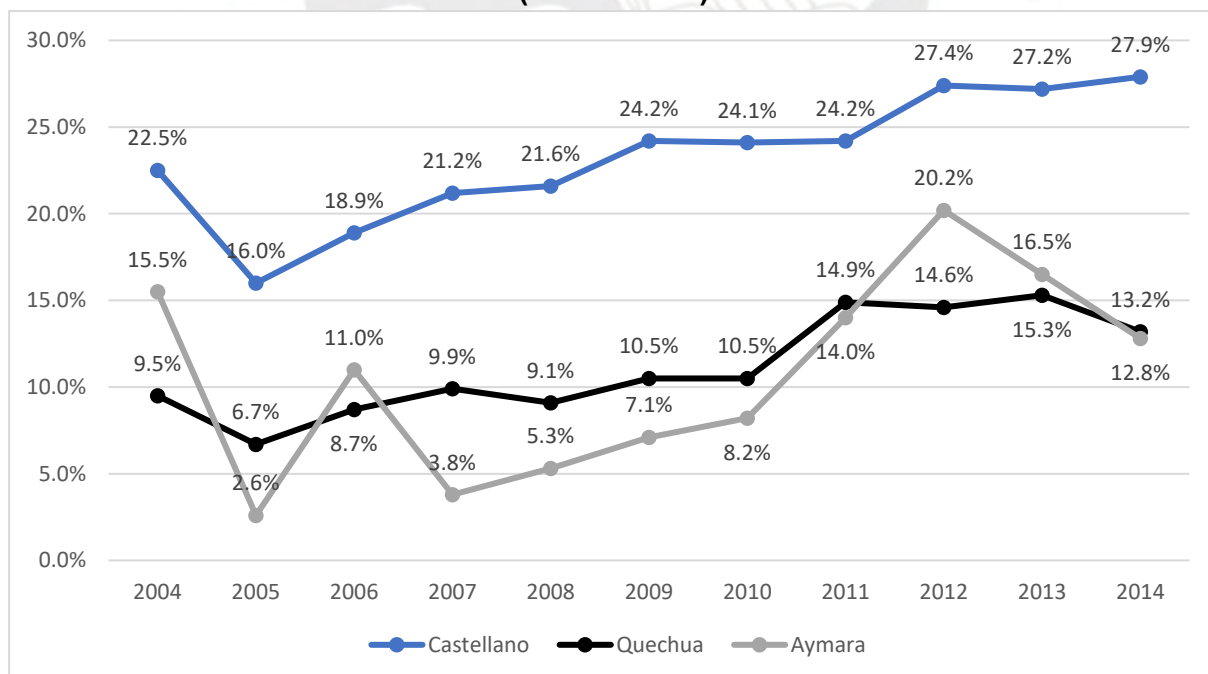
Gráfico 2.3a
Porcentaje de jóvenes pertenecientes a los quintiles (de ingreso familiar) 1 y 5 matriculados en universidades (2004 - 2014)



Fuente: ENAHO, reportado en Cuenca y Reategui (2016). Elaboración propia.

Nota: La brecha es calculada como la diferencia entre los porcentajes de matriculados que pertenecen a los quintiles de ingreso familiar 5 (más rico) y 1 (más pobre).

Gráfico 2.3b
Porcentaje de jóvenes, según lengua materna, matriculados en universidades (2004 - 2014)

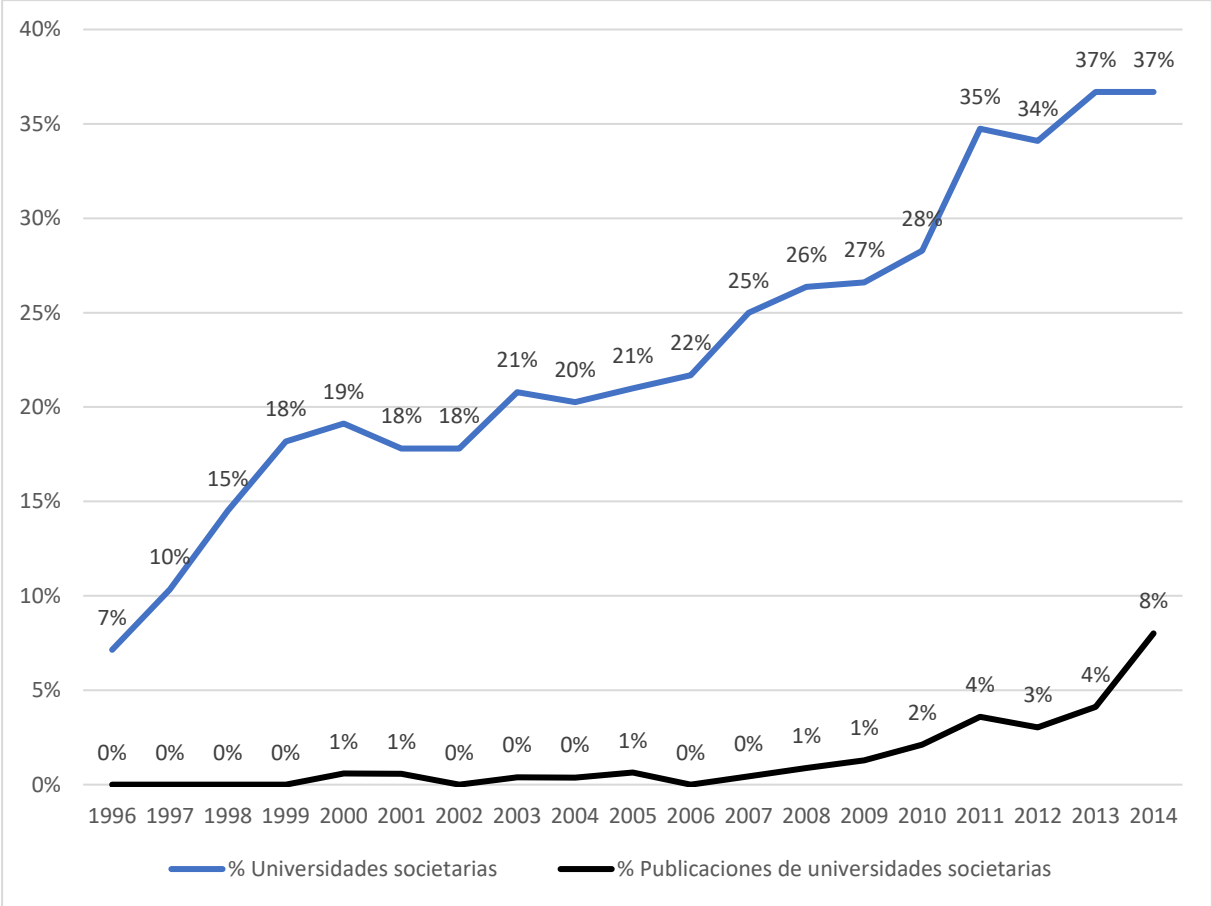


Fuente: ENAHO, reportado en Cuenca y Reategui (2016). Elaboración propia.

Sobre el tercer resultado esperable, que el mercado fuera suficiente para regular la calidad de las universidades, Yamada, Lavado y Martínez (2015) construyeron un índice de calidad a partir de los datos de los Censos Universitarios de 1998 y 2010, y determinaron que la calidad de las universidades creadas gracias al DL 882 de 1996 era significativamente inferior a la de aquellas que no fueron creadas en ese contexto, es decir, las creadas antes de 1996. Ese índice se basa en características como: ratio de profesores por estudiantes, porcentaje de estudiantes con buena opinión sobre su universidad, porcentaje de profesores cuyas investigaciones produjeron libros, etc.

Por otro lado, y siendo el interés de esta tesis el estudiar la producción de conocimiento, el gráfico 2.4 da una perspectiva de la importante diferencia entre el peso del número de universidades societarias (i.e. con fines de lucro) en el sistema universitario peruano y el peso de publicaciones de esas universidades en el total de publicaciones realizadas por investigadores afiliados a alguna universidad peruana. Mientras que el primer indicador llegó a ser de 37%, las publicaciones de las universidades societarias solo alcanzaron a representar el 8% en 2014.

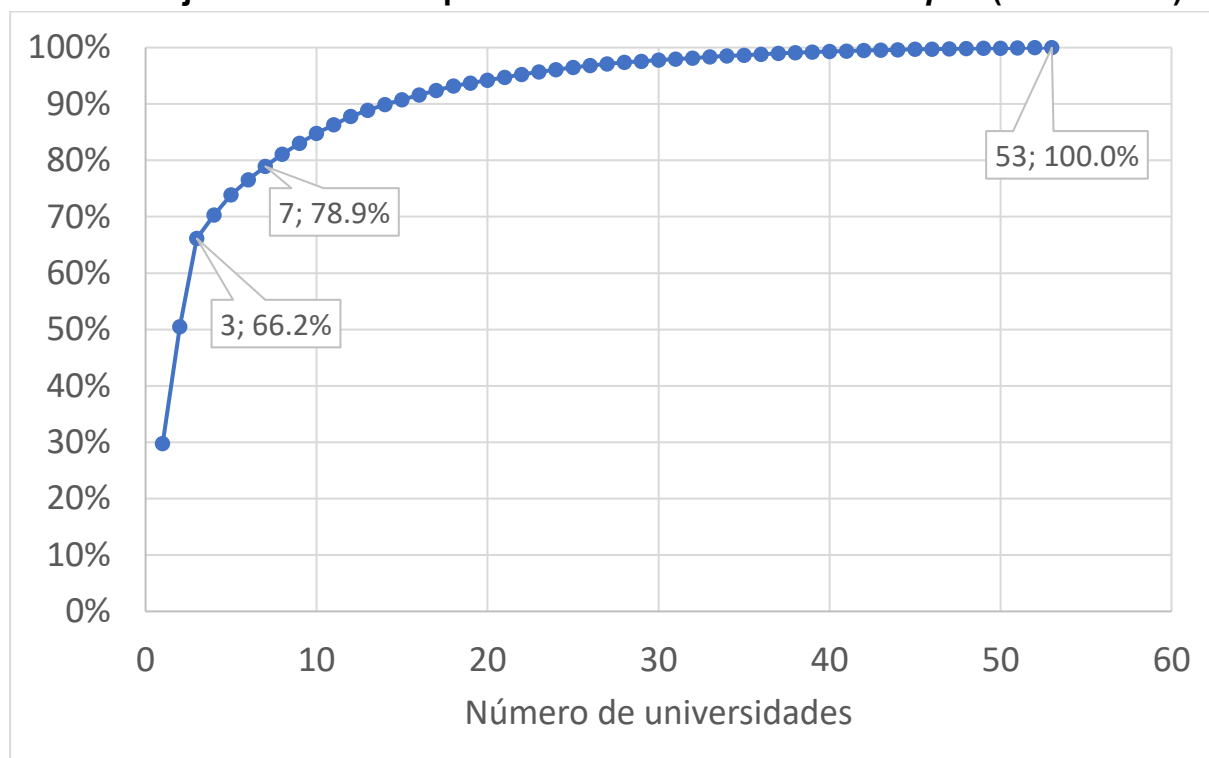
Gráfico 2.4
Porcentaje de universidades societarias y sus publicaciones, sobre el total del país, por año (1996 - 2014)



Fuente: SUNEDU y Scopus. Elaboración propia.

Por otro lado, si se analiza el número de artículos por institución, puede notarse la gran concentración de la producción de estos en pocas universidades. El gráfico 2.4 muestra que las dos terceras partes del total de publicaciones indizadas en Scopus entre 2011 y 2014 son atribuibles solo a tres universidades; que siete universidades son responsables de casi el 80% de las publicaciones; y que solo 53 de las más de 140 universidades peruanas contaban con al menos un documento indizado en esa base de datos.

Gráfico 2.4
Porcentaje acumulado de publicaciones indizadas en Scopus (2011 - 2014)



Fuente: Scopus. Elaboración propia

Es en ese contexto que el año 2014 se aprueba la nueva ley universitaria y, al año siguiente, el Ministerio de Educación del Perú (MINEDU) elabora los Lineamientos de Política para el Aseguramiento de la Calidad en la Educación Superior Universitaria. En dichos lineamientos se describen como objetivos de la universidad los siguientes puntos (MINEDU, 2015):

- *Preservar, acrecentar y transmitir de modo permanente la herencia científica, cultural y artística de la humanidad.*
- *Formar de manera integral profesionales de alta calidad y con pleno sentido de responsabilidad social.*
- *Proyectar a la comunidad sus acciones y servicios para promover su cambio y desarrollo.*

- *Colaborar de modo eficaz en la afirmación de la democracia y el estado de derecho (p. 8).*

Por su parte, el Consejo Nacional de Ciencia y Tecnología (CONCYTEC), que es el ente rector del sistema nacional de ciencia, tecnología e innovación tecnológica, señala en sus lineamientos de política nacional de ciencia y tecnología que las funciones de las universidades, dentro de ese sistema, son (CONCYTEC, 2016a):

- *Realizar investigación en ciencias y tecnología y fomentar la creación intelectual.*
- *Formar científicos y profesionales de alta calidad académica, de acuerdo con las necesidades del país.*
- *Extender su acción y sus servicios a la comunidad y promover su desarrollo integral (p. 110).*

Tanto el primero de los objetivos definidos por el MINEDU como la primera de las funciones descritas por el CONCYTEC hacen referencia a la producción intelectual o de investigación, reconociéndose a esa actividad como una de las principales funciones de las universidades.

Es así que la Superintendencia Nacional de Educación Universitaria (SUNEDU), institución creada a partir de la Nueva Ley y que posee entre sus funciones normar y supervisar condiciones básicas de calidad para las instituciones universitarias, considera como una de estas la “disponibilidad de personal docente calificado con no menos del 25% de docentes a tiempo completo” (SUNEDU, 2015). En el desarrollo de esa condición básica de calidad, la SUNEDU define nuevos perfiles de docentes, declarando que parte de estos deben estar dedicados al desarrollo de investigación.

3. MARCO TEÓRICO

La educación ha sido y es considerada como un factor fundamental del capital humano (Becker 1993; Mincer 1974). Una revisión de su importancia y su relación con la capacidad de generación de ingresos futuros, la salud, la fertilidad y otros aspectos puede verse en el trabajo de Case (2006). Además, la generación de conocimientos ha estado siempre ligada a instituciones educativas (Gibbons, et al. 1994), y estos son un insumo importante para el desarrollo de las sociedades (Romer, 1992). Por ese motivo, el estudio de los elementos del sistema educativo de un país tiene especial relevancia y relación con el desarrollo de estas.

Como se señaló antes, las universidades peruanas tienen entre sus fines el “preservar, acrecentar y transmitir de modo permanente la herencia científica, cultural y artística de la humanidad” (MINEDU, 2015). Entre las principales labores de los docentes, entonces, estarán el transmitir conocimientos e incrementarlos; es decir, enseñar e investigar. La relación entre estas dos actividades ha sido estudiada en numerosas ocasiones. A continuación, se presenta un pequeño resumen sobre los hallazgos de estos estudios.

3.1. Mecanismos entre enseñar e investigar

Hattie y Marsh (1996) llevaron a cabo una revisión de los trabajos realizados sobre la relación entre investigar y enseñar, y agruparon las teorías existentes en ocho modelos que postulaban relaciones positivas, negativas o inexistentes entre esas dos actividades. Tres modelos plantean una relación negativa entre la enseñanza y la investigación, estos son: i) de escasez de recursos, ii) de diferencias en la personalidad y iii) de premios divergentes. En un segundo grupo de modelos se señala una relación positiva: iv) de sabiduría convencional y v) el modelo “g”, que se explica abajo. Finalmente, otros tres modelos argumentan que no hay relación entre enseñar e investigar: vi) de oficio diferente, vii) de personalidades no relacionadas y viii) de la burocracia del financiamiento. El cuadro 3.1 resume los modelos mencionados y presenta un resumen de sus bases.

La contradicciones y superposiciones entre los modelos se deben a que cada uno de estos corresponde a una agrupación de diversos estudios realizados en diferentes contextos y con diferentes conclusiones³. A partir de esa revisión, los autores llevaron a cabo un meta-análisis de 58 estudios (de los que extrajeron 498 correlaciones en total) realizados de 1949 a 1992, obteniendo como resultado una relación ligeramente

³ Además de esos ocho modelos, los autores también mencionan tres modelos adicionales que plantean, más allá de encontrar si la relación entre las dos actividades es positiva, negativa o inexistente, el uso de variables moderadoras y sus interacciones

positiva entre enseñar e investigar, por lo que concluyeron que no puede aseverarse que ambas actividades estén relacionadas.

Cuadro 3.1
Modelos de la relación entre enseñar e investigar resumidos por Hattie y Marsh (1996)

Relación	Modelo ⁴	Basado en que...
Negativa	<i>Scarcity</i>	El tiempo, la energía y el compromiso que se le dedique a una de las actividades será en detrimento de la otra.
	<i>Differential personality</i>	Los tipos de personalidad que se requiere para cada una de esas actividades son contrarios (correlación negativa).
	<i>Divergent reward system</i>	Las expectativas y obligaciones que se tiene en cada rol son diferentes y conflictivas entre ellas.
Positiva	<i>Conventional wisdom</i>	La creencia general de los académicos es que la transmisión de conocimientos, el impulsar el interés por el conocimiento y las conexiones globales son características comunes a las dos actividades.
	"G"	Ambas actividades requieren de habilidades, valores y atributos similares (compromiso, creatividad, análisis crítico, etc.).
No relación	<i>Different enterprises</i>	Mientras que la investigación se orienta al descubrimiento del conocimiento, la enseñanza lo hace hacia la transmisión de este. Esto requiere cualidades no relacionadas.
	<i>Unrelated personality</i>	Los tipos de personalidad no son opuestos, como propone el modelo " <i>differential personality</i> ", sino que son no relacionados.
	<i>Bureaucratic funding</i>	Separar el financiamiento y el manejo de ambas actividades evitaría que el currículo de los estudiantes se vea influenciada por los intereses de los profesores, además de incentivar a los académicos según la práctica en la que sean mejores (enseñar o investigar).

Fuente: Hattie y Marsh (1996). Elaboración propia.

Al mejor entender del autor de esta tesis, no se ha realizado otro meta-análisis similar⁵. Sin embargo, en las últimas dos décadas sí se han publicado estudios adicionales específicos sobre el tema con resultados que no son del todo concluyentes. Por ejemplo, Marsh y Hattie (2002) utilizaron datos de una universidad en Australia para respaldar su propuesta de variables mediadoras y moderadoras (divididas en variables de fondo y recursos) buscando, con la ayuda de estas, articular la relación entre la enseñanza y la investigación. En un estudio más reciente, Palali, Elk, Bolhaar

⁴ Se mantienen los nombres en inglés utilizados por los autores para no alterar el sentido que quisieron darle.

⁵ El único estudio similar, que también es mencionado en Hattie y Marsh (1996), fue realizado en 1987 (Feldman, 1987) y encontró una ligera correlación positiva entre las dos variables.

y Rud (2018) hicieron lo propio con datos de una institución en los Países Bajos. Los resultados de esos, y otros, estudios siguen siendo no concluyentes, pues dependen en gran medida del contexto en el que se llevaron a cabo, así como de la operacionalización de las variables de interés (calidad de enseñanza e investigación).

3.2. Motivaciones extrínsecas e intrínsecas

El estar motivado puede entenderse como estar “movido a hacer algo”. La motivación y sus factores han sido ampliamente estudiados dentro de la psicología. En la teoría de la autodeterminación (Deci y Ryan 1985) se distinguen dos tipos de motivación: la intrínseca y la extrínseca. El primero de estos tipos hace referencia a realizar una actividad porque esta es inherentemente interesante o agradable para la persona, mientras que la motivación extrínseca se refiere a actuar para alcanzar un resultado separado de la acción misma (Ryan y Deci, 2000).

La teoría antes descrita se desarrolla sobre todo en el ámbito del aprendizaje y sus factores. Además, señala que los premios y castigos (incentivos, motivación extrínseca) son contraproducentes pues la repetición de estos termina socavando la motivación intrínseca, que es la más fuerte e importante entre los dos tipos de motivación (Ryan y Deci, 2000).

Sin embargo, la teoría económica considera a los incentivos, que caerían dentro del tipo de motivación extrínseca, como algo bueno para la orientación de las actividades de un agente. Esta aparente disonancia entre la teoría psicológica y económica es abordada por Bénabou y Tirole (2003). Ellos plantean un modelo para los dos tipos de motivaciones dentro de un contexto de interacción entre un principal y un agente.

El modelo propuesto postula costos ocultos para los incentivos. Una vez estos costos son tomados en cuenta, se puede determinar si los incentivos que el principal da al agente para que este último realice una tarea son o no contraproducentes. De ese análisis los autores concluyen que incluso en los casos en los que las motivaciones extrínsecas disminuyen, en lugar de incrementar, la predisposición a realizar la tarea, el agente está actuando racionalmente; es decir, no hay contradicción entre los postulados de la teoría psicológicas y los de la teoría económica.

El modelo planteado, para el caso del agente, es el siguiente:

$$U_A = [\theta(V + b) - \hat{C}_{(\sigma,b)}]e$$

Donde:

U_A : Utilidad del agente

θ : Habilidad del agente

V : Ganancia del agente si tiene éxito

b : Bono pagado por el principal al agente, si este último tiene éxito

\hat{C} : Costo para el agente. El valor esperado del costo, condicional al bono y la señal

σ : Señal que el principal envía al agente.

e : Esfuerzo que realiza el agente.

Así, el incentivo que el principal ofrece está relacionado con la señal que el agente utilizará para generarse una idea del costo que le implicaría emprender la tarea. El principal, que conoce el costo real, estará dispuesto a dar un incentivo mayor si sabe que este costo es elevado, y esto será interpretado negativamente por el agente haciendo que esté menos motivado para realizar la tarea.

Los autores también desarrollan una variante del modelo para la situación en la que el costo es conocido por ambos y la habilidad, θ , es la variable conocida por el principal y no por el agente. Este último utilizará la señal que recibe para aproximar un valor para esa variable.

Hasta este punto, se tiene que la producción de investigación de un profesor universitario tiene una relación negativa, positiva o neutra con su dedicación a la enseñanza. Además, por el estudio de las motivaciones intrínsecas y extrínsecas, esta producción dependerá de los incentivos que la universidad esté dispuesta a darle, de su habilidad como investigador, de los costos que prevé para la tarea y de su esfuerzo. Ahora se ahondará en cómo puede medirse la producción de investigación para un profesor universitario.

3.3. Sobre la medición de la productividad en investigación

El desarrollo de la medición de la producción de investigación ha estado estrechamente ligado a la bibliometría, y más específicamente a la cienciometría. El indicador para medir el nivel de investigación de una persona, una institución o un país suele ser una ponderación de diversos elementos.

Tres problemas para la medición de esta producción son discutidos por Toutkoushian y Webber (2011). El primero de estos es la limitación de información disponible, problema que no es exclusivo de la medición de producción de investigación, sino que es común a cualquier análisis cualitativo o cuantitativo. En el caso específico de la producción de investigación, no siempre es posible acceder a información fiable de variables como número de premios ganados, montos de fondos de investigación asignados, total de publicaciones realizadas, etc.

El segundo problema que mencionan los autores es el de sopesar correctamente la calidad y la cantidad. Ambos aspectos son importantes y deben ser considerados para la medición, pero este trabajo suele ser un reto, sobre todo si se toma en cuenta el

primer problema mencionado. Los intentos de afrontar este segundo aspecto se han centrado en definir criterios de calidad para los elementos que se estén contando. Por ejemplo, dar diferentes pesos a publicaciones en diferentes revistas de acuerdo con la relevancia de estas.

Finalmente, los autores abordan el inconveniente de la agregación o ponderación. Dado que se tendrán diferentes conteos de elementos esencialmente distintos, como fondos de investigación y publicaciones o patentes, la ponderación y agregación de todos estos conteos no carece de dificultad. Debe señalarse también que es imposible tener pesos universales para cada uno de los aspectos que serán medidos, y si existiesen esos pesos universales, no estarían exentos de críticas.

En el mismo estudio (Toutkoushian y Webber 2011), los autores, siguiendo lo señalado por Diamon y Graham (2000), Johnes et al. (1993) y Tan (1986), listan las siguientes cinco aproximaciones como las más utilizadas para implementar una medición de la producción de investigación: i) financiamiento para investigación, ii) conteo de publicaciones, iii) conteo de citados, iv) conteo de premios ganados, y v) conteo de patentes registradas.

En general, señalan, la medida del nivel de producción de una universidad, departamento académico, o investigador individual será uno de esos cinco indicadores, o una combinación de ellos. Sin embargo, esto debe adecuarse al contexto en estudio. Tomando en cuenta las limitaciones ya mencionadas (mediciones, calidades y ponderaciones) y la situación puntual de la organización o sociedad a analizar. Como ejemplo, en el Perú no sería adecuado tomar en cuenta el “número de premios Nobel con que cuenta la institución”.

Medición de publicaciones y tipos de universidades

La Ley de Lotka

Alfred J. Lotka (1926) analizó el número de los artículos publicados en las revistas *Chemical Abstracts* y *Auerdach's Geschichtstafel*, dedicadas a la Química y Física respectivamente, en un periodo de diez años, enfocándose en los conteos de publicaciones por autor. A partir de ese trabajo, postuló una de las leyes más famosas en bibliometría que aborda la distribución de frecuencias en la productividad científica, y que modela la concentración de las publicaciones científicas en pocos autores muy productivos según la siguiente relación:

$$y_x = \frac{6}{\pi^2 x^\alpha}$$

Donde y_x es la frecuencia relativa de autores que publican x número de artículos. Encontró, para el conjunto de información que analizó, que los valores para α se encontraba alrededor de dos. Puede notarse que esa Ley implica que el número de publicaciones por investigador estará siempre concentrado en valores bajos.

Por su parte, Bookstein (1976) plantea una versión generalizada de esa Ley, relajando el valor del coeficiente fijo que acompaña al número de publicaciones para reemplazarlo por una constante genérica como se muestra en la siguiente ecuación.

$$y_x = \frac{k}{x^\alpha}$$

Bookstein (1796) llegó a esa generalización luego de estudiar, además de la Ley de Lotka, otras distribuciones muy usadas en bibliometría (como la Ley de Zipf, que trata de la frecuencia de palabras repetidas en un texto, y la Ley de Bradford, que se ocupa de la distribución de artículos de cierto tema en un grupo de revistas). Él encuentra que:

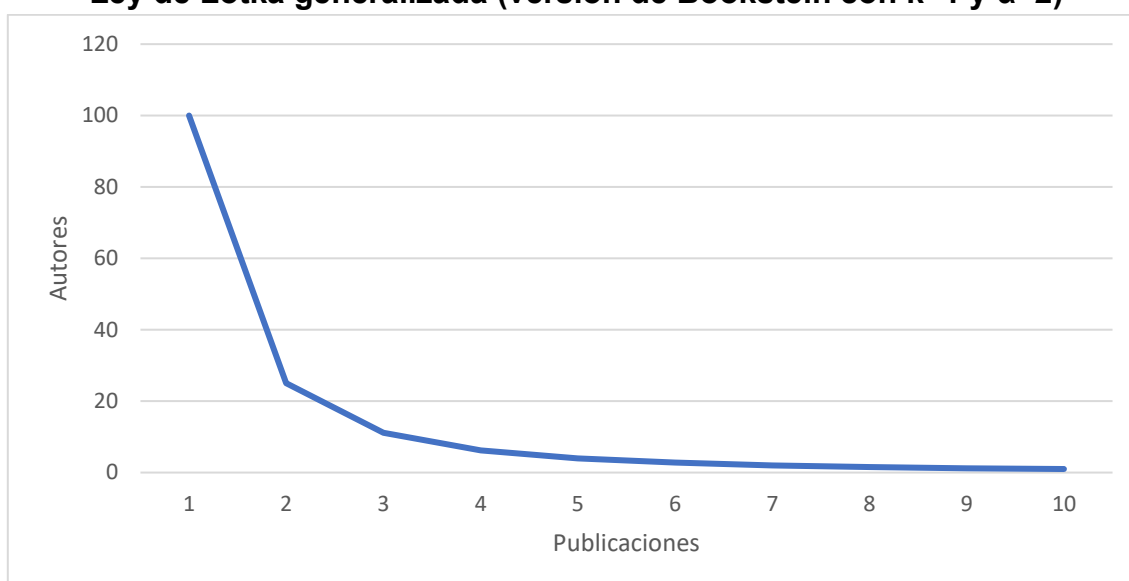
- En el campo de la bibliometría, situaciones complejas pueden ser descritas por distribuciones “elegantes”.
- Las distribuciones persisten, aun cuando se analizan grupos de datos diferentes.
- Las distribuciones son, en principio, equivalentes en el sentido que tratan de entidades (científicos, palabras, revistas) que producen ciertos eventos (publicaciones, ocurrencias) dentro de otra dimensión (tiempo, tamaño de texto).

Por esos motivos, busca unificar esas relaciones, haciendo uso de la distribución de Poisson para modelar la ocurrencia de esos eventos, llegando a la ecuación antes planteada.

Entonces, la forma de la distribución que, empíricamente, seguiría el número de publicaciones que cada investigador realiza sería similar al gráfico 2.1.

Gráfico 2.1

Ley de Lotka generalizada (versión de Bookstein con $k=1$ y $\alpha=2$)



Fuente: Elaboración propia.

Algunos artículos recientes que han utilizado la Ley de Lotka, o una variación de esta, en diversos campos del conocimiento son: Kumar y Senthikumar (2019) quienes se ocupan de la astronomía y Astrofísica en India; Sadik (2018), quien estudia la investigación sobre el Dengue; Kwiek (2018), quien explora el comportamiento de los académicos más productivos en Polonia; Miss (2018), quien evalúa la investigación en cáncer de pulmón; entre otros⁶.

Tipos de Universidades

Sam y van der Sijde (2014) realizan una revisión de los estudios hechos sobre los modelos de universidad. Los autores indican que la tradicionalmente, y a partir de las instituciones europeas, las universidades podían clasificarse en tres grupos: humboldianas, napoleónicas y anglosajonas (o “de investigación”, “de entrenamiento” y “de personalidad”). A continuación, se reseña brevemente lo expuesto por los autores:

- En el primer modelo (modelo alemán), “*the professors are not there for students, but rather both are there for science*”; así, se considera que la misión central es la creación y transmisión de conocimiento. El sistema es centralizado, pues las universidades dependen económicamente casi por completo del gobierno.

⁶ Puede revisarse el trabajo de Urbizagástegui (2004). Él aborda el tema listando varios ejemplos de trabajos realizados en diversos campos siguiendo la Ley de Lotka y describiendo algunos métodos prácticos para aplicarla.

- En el segundo modelo (modelo francés), las universidades son concebidas como entidades públicas que forman igualitariamente a sus estudiantes, de manera tal que estos, al egresar, están en similares capacidades de ejercer su profesión. Se señala también que un indicador de esto es el que los grados académicos estén “dados por la nación”, indicando así que son de similar valor sin importar la institución en la que se formó el egresado. Se da mayor importancia al “entrenamiento” en carreras vocacionales, como ingeniería, que a las humanidades y ciencias básicas.
- En el tercer modelo (modelo británico), la formación está más orientada a impartir una educación amplia que una centrada en un tema o habilidad específica. Este sistema enfatiza más el profesionalismo que el entramiento técnico, que puede ser desarrollado con estudios posteriores o mediante el trabajo en la industria. Las universidades tienen gobiernos autónomos, bajo supervisión de organismos independientes, de manera que pueden competir libremente por el reclutamiento de estudiantes.

A partir de eso, Sam y van der Sijde (2014) analizan el sistema de Estados Unidos, e indican que su modelo de universidades, si bien surgió a partir de los europeos antes descritos, se ha convertido en el más influyente en nuestros días. El modelo “angloamericano” se corresponde con la idea de “emprendedurismo”, no solo porque lo promueva entre sus estudiantes, sino porque lo utiliza para obtener recursos económicos de manera tal que pueda cumplir con su misión de producir y transmitir conocimiento.

Por otro lado, *The Carnegie Classification of Institutions of Higher Education*⁷ presenta una clasificación, que se actualiza periódicamente, de los centros de educación superior estadounidenses. Esa clasificación está basada, en su mayor parte, en el tipo de programas que ofrecen las instituciones y si éstas están orientadas a formar investigadores (para lo cual cuentan con numerosos programas de doctorado) o no. Dos ejemplos de trabajos académicos que tomaron esa clasificación para analizar temas educativos son el de Fram y Lau (1998), que analiza la percepción de las personas sobre las universidades de “investigación” y de “enseñanza”; y el de Fairweather (1995), que estudia el mercado laboral para los académicos en Estados Unidos

A modo de síntesis

En este capítulo se han revisado aspectos importantes entorno a la dinámica de la producción científica en la que están envueltos los profesores de las universidades. En primer lugar, existen distintos fundamentos que tratan la relación entre la enseñanza y la investigación para un profesor universitario. Esta relación puede ser

⁷ Puede consultarse en: <http://carnegieclassifications.iu.edu/>

positiva, negativa o neutra. Se deberán considerar variables relacionadas a ambos factores.

En segundo lugar, existe coherencia entre los resultados de investigaciones psicológicas y la teoría económica en cuanto a la motivación intrínseca y extrínseca. Incluso cuando la motivación extrínseca es contraproducente, el sujeto en cuestión está siendo racional. El modelo propuesto por Bénabou y Tirole (2003) indica que deben tomarse en cuenta la habilidad, el costo (conocido o no), el esfuerzo y el premio ofrecido por el agente. Entonces, deben introducirse características objetivas, relacionadas a motivaciones extrínsecas, como subjetivas, relacionadas a motivaciones intrínsecas.

Además, se ha señalado que la medición de la investigación realizada no carece de dificultades. Estas están relacionadas a la información disponible, la agregación de los indicadores, y la relación entre calidad y cantidad. Además, los indicadores elegidos deben guardar concordancia con el contexto que se está estudiando. Son cinco los indicadores más utilizados pero el uso de una parte de ellos o de todos dependerá de ese contexto, así como de la disponibilidad de información confiable.

Por último, se reseñaron los hallazgos hechos desde la bibliometría sobre el estudio de la productividad de los investigadores por medio de la exposición de la Ley de Lotka⁸; y los modelos de clasificación de las universidades.

⁸ Esa Ley empírica ayudará en gran medida a elegir las distribuciones que se utilizarán en los análisis econométricos en el presente trabajo.

4. ESTUDIOS EMPÍRICOS PREVIOS

En la literatura internacional puede encontrarse diversos estudios cuyos objetivos son determinar los factores asociados, y sus pesos, de la producción de investigación de instituciones, sociedades o autores. En los párrafos siguientes se presentan los hallazgos más resaltantes de algunos de esos trabajos, y en el cuadro 4.1 la lista de variables, y el signo de la correlación con la variable explicada, que resultaron estadísticamente significativas⁹.

Dundar y Lewis (1998) busca la relación entre la productividad en investigación y factores institucionales en cuatro áreas académicas: ciencias biológicas, ingeniería, ciencias físicas y matemática; y ciencias sociales y del comportamiento. Atributos individuales, atributos institucionales. 30 programas en cuatro áreas en 90 universidades de investigación de Estados Unidos. La variable dependiente es el ratio de publicaciones por profesor; para las variables independientes: tamaño del programa, concentración de profesores con publicaciones en el departamento, porcentaje de profesores “full”, gasto institucional en recursos bibliográficos, ratio de estudiantes de posgrado y profesores, porcentaje de profesores con apoyo para investigación, porcentaje de estudiantes de posgrado con trabajos de investigación, dummy para universidad pública o privada.

Print y Hattie (1997) estudiaron qué indicadores son los que toman en consideración los reclutadores de los departamentos académicos en universidades australianas para determinar si un candidato a profesor tendrá alta o baja productividad en investigación. Para esto, ellos se valieron de encuestas a estos departamentos académicos basados en 35 perfiles creados como candidatos a ocupar puestos de profesores. Encontraron que el número de artículos publicados en *journals* arbitrados o los financiamientos grandes ganados tienen mayor peso para la decisión de los reclutadores que la supervisión de estudiantes de posgrado, por ejemplo.

Por otro lado, el trabajo de Mueller, Gaus y Konradt (2016) tuvo como objetivo elaborar un modelo predictivo para el nivel de producción científica de un país en el área específica de las evaluaciones. Los autores analizaron datos para 65 países e hicieron uso de la información de 10 *peer-reviewed journals* de esa área para sus estimaciones. Encontraron que las variables que influyen en la variable dependiente fueron, en orden decreciente de influencia en esta, son: la productividad del país en el área de ciencias sociales, el PBI per cápita del país, un índice de corrupción, la antigüedad de asociaciones de evaluación¹⁰ presentes en el país; el tamaño del sector académico medido como el número de universidades en el país; un índice de

⁹ Puede consultarse también el trabajo de Hesli y Lee (2011), en el que presentan un cuadro resumen de variables que diversos estudios encontraron significativas para la producción de investigación.

¹⁰ Instituciones dedicadas a fomentar la práctica de las evaluaciones, sobre todo enfocadas en políticas públicas. En Perú se tiene a la Red Peruana de Evaluación (Eval Perú), cuya página web es: <http://www.evalperu.org>

libertades civiles; el tamaño del mercado continental de *journals* del área de evaluación; y la lengua oficial.

Hardré y otros (2011) analizaron el rol de la motivación en la producción de conocimiento con los datos obtenidos de una encuesta a académicos de 17 estados dentro de Estados Unidos. Utilizando el *path analysis*, encontraron que las variables directamente relacionadas con la dependiente final eran: el esfuerzo dedicado a la investigación (positivamente), las horas dedicadas a asesorar (positivamente) y la carga de horas de enseñanza que su departamento académico le exige (negativamente). Además de esas tres variables, consideraron otras cinco que interactuaban entre ellas y con esas tres.

Otro estudio relevante, efectuado con datos para Estados Unidos, es el de Zhang y Ehrenberg (2010). Ellos analizaron la relación entre la contratación de los profesores y el financiamiento externo captado para proyectos de investigación y desarrollo (I&D) para universidades de investigación. Clasificaron a los profesores en tres tipos: i) aquellos a tiempo completo dentro del “*tenure track*”, incluyendo los que aún no han alcanzado la categoría de “*tenured*”, ii) los de tiempo completo que no están dentro del “*tenure track*”, y iii) los profesores a tiempo parcial; para estos últimos sus labores fuera de la universidad son muy relevantes. Los autores encontraron que el número de profesores a tiempo completo está relacionado positivamente con el incremento de financiamiento externo para I&D, mientras que un incremento en la proporción de profesores a tiempo completo que no están en el “*tenure track*” disminuiría ese tipo de financiamientos. Además, el incremento del número de profesores a tiempo parcial, manteniendo constante el número de profesores a tiempo completo, también incrementa el financiamiento externo. La explicación que proponen para ese último hallazgo es que probablemente el incremento de personal a tiempo parcial ayude a descargar de horas de dictado a los profesores a tiempo completo, por lo que estos últimos tendrían más tiempo para investigar.

En Asia, Shin y Cummings (2010) estudiaron la producción académica y las características de los profesores de 99 universidades de Corea del Sur. Los autores tomaron como motivación el cambio en las políticas públicas de varios países de la región en cuanto al desarrollo de la ciencia; haciendo uso de métodos de regresión y bases de datos de encuestas, encontraron que las variables como la preferencia de los profesores por investigar, tiempo para investigar, colaboración internacional, la experiencia y la orientación institucional hacia objetivos tienen efectos positivos en la producción; también que disciplina a la que se dedica el profesor es relevante, mientras características como el sexo y el número de hijos no lo son. Por otro lado, Nurhudatiana y Anggreini (2015) realizaron un estudio con el fin de predecir la producción, entendida como publicaciones, de los profesores de una universidad en Indonesia. Haciendo uso de técnicas de árboles de decisión, concluyeron que las características relevantes para predecir si un profesor publicará o no un artículo en

cierto año son: el grado académico, el campo de estudio y publicaciones antes de trabajar en la universidad son importantes para profesores con dos años o menos en la universidad; las publicaciones en los dos años anteriores, grado académico y número de años de servicio son importantes para profesores con entre tres y siete años en la universidad; y las publicaciones en dos años anteriores, el grado académico y el campo de estudio para aquellos que tienen más de siete años como profesores.

El caso de los trabajos empíricos realizados para Noruega es especial porque, en palabras de Rørstad y Aksnes (2015): *“Norway is well suited as a case for analysing this issue. Unique datasets are available not only on the research output but also on the research staff”* (p.2). Algunos de los estudios recientes para este país, además del citado de Rørstad y Aksnes, son el de Piro, Rørstad y Akneses (2016) y el de Hvide y Jones (2018). En estos se concluye que características como la edad, el género, la posición académica y el área del conocimiento están asociados a la productividad de los investigadores. Además, que incentivos como la propiedad sobre los productos de las investigaciones (como las patentes o empresas fundadas en base a los resultados de los proyectos) también son factores relevantes para el investigador.

También en Europa, van Arensbergen, van der Weijden y van den Besselaar (2012) estudiaron la relación entre la producción de investigación y el sexo en investigadores de Países Bajos. Si bien no utilizaron técnicas econométricas para su trabajo, sí realizaron comparaciones de las medianas de publicaciones de investigadores hombres y mujeres¹¹. Encontraron que sí existe diferencia en el número de publicaciones realizadas y el número de citaciones recibidas, en ambos casos con ventaja para los hombres; sin embargo, esa diferencia desaparece cuando solo se considera a los investigadores que se encuentran en la etapa inicial de su carrera.

En Latinoamérica, Ruedas y Rodenes (2015) llevaron a cabo un estudio para Colombia. Los datos que utilizaron fueron los conseguidos con una encuesta a 223 grupos de investigación del país registrados en un portal administrativo del gobierno. Su análisis aplicó la técnica de regresión lineal múltiple y utilizó como variables independientes explicativas de la producción científica la externalización, la dotación tecnológica, la cultura motivadora, y el tiempo del investigador. Cada una de esas variables es una composición de otras menos complejas. Las variables que encontraron significativas para explicar a la producción científica fueron: cultura motivadora, dotación tecnológica, externalización y tiempo para investigación.

Por otro lado, Teodorescu (2000) pone en evidencia que la opción de generalizar modelos teóricos para explicar la motivación docente es un reto debido a que las características culturales y sociales de cada región difieren ampliamente. Es por ese motivo que la mayoría de los estudios se concentra en el análisis de un país en

¹¹ Para esto, hicieron uso de pruebas de Mann-Whitney, por tener distribuciones no paramétricas.

particular, haciendo que el modelo a utilizar tenga que ajustarse. Un ejemplo de ese reto es que en su estudio encuentra una variación importante en el coeficiente de determinación para diferentes países utilizando el mismo conjunto de indicadores. Así, los indicadores usados explican un 40% de la varianza en Japón mientras que solo un 20% en Chile. Una de las implicancias relevantes de su trabajo es que, por las diferencias considerables entre los países estudiados, la mayoría de los resultados encontrados en estudios para naciones desarrolladas no pueden ser importados a naciones en vías de desarrollo.

En el Perú no se han realizado trabajos cuantitativos de este tipo. Si bien existen rankings de universidades y dentro de ellos se asigna un “puntaje de investigación” a estas, no se ha estudiado en profundidad cuáles son los determinantes, y sus pesos, de la investigación. El CONCYTEC señaló dentro de su Memoria Institucional del 2014 que estaba llevado a cabo el estudio “Determinantes de la Producción Científica en las Universidades Peruanas”, sin embargo, en el 2016 informó a la Presidencia del Consejo de Ministros (PCM) que el estudio no pudo ser terminado por falta de información, en el mismo informe presentó el avance del proyecto (CONCYTEC 2016b). Este avance se basó en entrevistas a profundidad a docentes de diversas universidades, quienes listaron las siguientes dimensiones como las relevantes en la producción de investigación en sus instituciones:

- i) Dimensión colaborativa: las redes que logran crear los docentes son por esfuerzos propios, sobre todo durante sus estudios de posgrado. Pocas universidades privadas brindan apoyo a sus profesores para que estos puedan participar en redes de investigación.
- ii) Dimensión demográfica: la edad y el género influyen en la producción. El periodo después de terminar el posgrado es el más productivo. La maternidad es percibida como negativa para el desarrollo de investigación.
- iii) Trayectoria: las habilidades del docente, así como el manejo de otros idiomas son importantes para el desarrollo de investigación.
- iv) Dimensión económica: el bajo nivel de salarios obliga a los profesores, sobre todo a los de universidades nacionales, a realizar actividades complementarias para generar más ingresos.
- v) Dimensión institucional: la mayoría de las universidades nacionales no cuenta con políticas que incentiven la investigación en sus instituciones. Algunas universidades privadas cuentan con incentivos como la reducción de horas de dictado de clases.

Puede notarse que los factores identificados por el estudio de CONCYTEC coinciden en gran medida con los que se suelen tomar en cuenta en los trabajos empíricos realizados en otros países. Esto servirá de guía para seleccionar las características del docente a considerar para el presente estudio.

Cuadro 4.1
Resumen de estudios empíricos, sus variables y hallazgos

ESTUDIO	ÁREA DEL CONOCIMIENTO ABORDADA	VARIABLE DEPENDIENTE	VARIABLES INDEPENDIENTES SIGNIFICATIVAS	RELACIÓN CON VARIABLE DEPENDIENTE	DATOS CONSIDERADOS
CONCYTEC (2016) ¹²	Varios	Producción científica	Colaborativa Demográfica Económica Institucionalidad Trayectoria	(+) (+) (+) (+) (+)	Perú, 26 universidades
Dundar y Lewis (1998)	Varios	Ratio de publicaciones por profesor	Es institución pública Gastos en recursos bibliográficos Número de profesores Porcentaje de estudiantes de posgrado como RAs Porcentaje de full professors Porcentaje de profesores con apoyos para investigación Porcentaje de profesores con publicaciones Ratio estudiantes de posgrado a profesores	(-) (+) (+) (+) (+) (+) (+) (+)	Estados Unidos, 90 universidades de investigación
Hardré y otros (2011)	Varias disciplinas	Documentos arbitrados publicados; participación en conferencias	Apoyo de su departamento a la investigación Carga de asesoría Carga de enseñanza Carga de servicio	(+) (+) (-) (-)	Estados Unidos, 781 profesores de 28 universidades

¹² Como se mencionó, a diferencia de los otros estudios reseñados, el de CONCYTEC es cualitativo.

			Eficacia para investigar	(+)	
			Esfuerzo (%) a investigar	(+)	
			Interés intrínseco para investigar	(+)	
			Valor que le asigna a investigar	(+)	
Hvide y Jones (2018)	Varios	Número de startups creadas; número de patentes	Tenencia de derechos de propiedad intelectual sobre investigaciones	(+)	Noruega
Mueller y otros (2016)	Evaluación	Documentos arbitrados publicados	Antigüedad de sociedades de evaluación	(+)	65 países
			Índice de control de corrupción	(+)	
			Índice de libertades civiles	(+)	
			Inglés como lengua oficial	(+)	
			Número de universidades	(+)	
			PBI per cápita	(+)	
			Productividad de investigación en Ciencias Sociales	(+)	
			Tamaño del mercado continental de journals	(+)	
Print y Hattie (1997)	Educación	Preferencia de los reclutadores	Artículos en conferencias	(+)	Australia
			Capítulos en libros	(+)	
			Fondos de investigación grandes	(+)	
			Fondos de investigación menores	(+)	
			Libros peer reviewed	(+)	
			Monografías	(+)	
			Publicaciones en journals arbitrados	(+)	
			Publicaciones en journals no arbitrados	(+)	
			Reportes	(+)	
			Supervisión de estudiantes de doctorado	(+)	

			Supervisión de estudiantes de maestría	(+)	
			Textos	(+)	
Rørstad y Aksnes (2015)	Varias disciplinas, por separado	Artículos arbitrados publicados	Categoría en la institución (ascendente)	(+) (+) (+) ¹³	Noruega, 4 universidades de (12400 investigadores)
			Edad	(-) (+) (+)	
			Edad al cuadrado	(-) (-) (-)	
			Es mujer	(-) (-) (-)	
Ruedas y Rodenes (2015)	Varios	Índice de producción de nuevos conocimientos	Cultura motivadora	(+)	Colombia, 347 grupos de investigación
			Dotación tecnológica	(+)	
			Externalización	(+)	
			Tiempo para investigación	(+)	
Teodorescu (2000)	Varios	Artículos arbitrados publicados	Años en su institución actual	(+) / 1 país	10 países, por separado
			Calidad de estudios de posgrado	(+) / 1 país	
			Categoría en la institución (ascendente)	(+) / 2 países	
			Edad	(-) / 1 país	
			En tenure track	(+) / 1 país	
			Énfasis institucional a la investigación	(+) / 2 países	
			Envolvimiento en actividades profesionales internacionales	(+) / 2 países	
			Estímulos extrínsecos percibidos	(+) / 1 país	
			Evaluación de la investigación por la institución	(+) / 1 país	
			Fondos de investigación recibidos	(+) / 7 países	

¹³ El primer signo corresponde a Ciencias Naturales, el segundo a Ingeniería y Tecnología, y el tercero a Medicina.

			Género	(-) / 1 país	
			Horas de dictado de clases	(-) / 1 país	
			Horas de trabajo administrativo	(+) / 1 país	
			Horas dedicadas a investigar	(+) / 4 países	
			Número de asociaciones internacionales a las que pertenece	(+) / 4 países	
			Número de asociaciones nacionales a las que pertenece	(+) / 2 países	
			Número de conferencias internacionales a las que asistió	(+) / 7 países	
			Número de conferencias nacionales a las que asistió	(+) / 5 países	
			Posee otro trabajo académico	(+) / 1 país	
			Presión de la institución para hacer investigación	(-) / 2 países	
			Salario	(+) / 2 países	
Zhang y Ehrenberg (2010)	Varias disciplinas	Fondos externos para I%D	Financiamiento propio a proyectos de I&D el año anterior	(+)	Estados Unidos, Universidades
			Matricula de estudiantes	(+)	
			Número de profesores a tiempo completo	(+)	
			Proporción de profesores tiempo completo fuera de Tenure track	(-)	
			Proporción de profesores tiempo parcial tiempo completo fuera Tenure track	(+)	
			Valor ponderado de financiamiento de agencias federales	(+)	
Shin y Cummings (2010)	Varios	Libros y artículos publicados	Edad	(-)	Corea del Sur, 99 universidades
			Tiempo para investigación	(+)	
			Preferencia por investigación	(+)	
			Colaboración internacional	(+)	

Área diferente a humanidades	(+)
Años desde obtención del PhD	(+)
Experiencia posdoctoral	(+)
Participación en proyectos de investigación	(+)
Soporte institucional	(-)
Orientación institucional hacia objetivos	(+)
Universidad de investigación/doctoral	(+)

Fuente: Elaboración propia.



5. OBJETIVOS E HIPÓTESIS

Para afrontar la situación del bajo nivel de la producción de investigación en las universidades peruanas, primero deben entenderse adecuadamente sus causas. Por ese motivo, la pregunta de investigación de este trabajo es:

¿Cuáles son los determinantes de la producción de investigación¹⁴ en las universidades peruanas?

Al responder esa pregunta, se logrará el siguiente objetivo:

Aportar a la discusión sobre la calidad universitaria en el Perú estudiando los determinantes de la producción académica¹⁵ y cuantificando su aporte. En particular, revisando los factores relacionados con las características de los docentes universitarios y con la institución.

Teodorescu (2000), basado en Fox (1985), propone que el nivel de producción de investigación en las universidades está relacionado con características “adscriptivas”, “de logro”, e “institucionales”. Las características adscriptivas son aquellas que no son controladas por los individuos, como el sexo; las de logro son aquellas que informan sobre el desarrollo académico de los individuos, como el grado académico alcanzado. Por otro lado, las características institucionales describen el contexto institucional en el que los académicos desarrollan su labor, como comodidades o beneficios que su institución puede darle.

En esa línea, la hipótesis de este trabajo es que: *las características más relevantes para explicar la producción de investigación en las universidades peruanas están relacionadas con lo revisado en el marco teórico, esto es, i. el uso del tiempo entre enseñar e investigar, ii. las motivaciones intrínsecas o extrínsecas que puede tener el docente, y iii. el tipo de institución en el que se encuentra.* Además, dado el bajo nivel de desarrollo de la investigación en el país, el bajo financiamiento y la baja proporción de profesores adecuadamente entrenados para desarrollar investigación, se propone *iv. que el nivel de financiamiento y v. la tenencia de doctorado también serán especialmente importantes.*

Entonces, la función de producción a analizar será:

$$Y_i = f(T_i, M_i, I_i, S_i, D_i)$$

Donde:

¹⁴ La producción se entiende como las publicaciones académicas realizadas por los docentes universitarios.

¹⁵ Ídem.

T: Tiempo para enseñar e investigar

M: Motivaciones

I: Tipo de institución

S: Financiamiento

D: Estudios de doctorado

i: Indicador de la unidad de análisis

Cada uno de los términos podrá referirse a una o más variables.



6. CONSIDERACIONES PARA EL ANÁLISIS EMPÍRICO

En esta sección se da cuenta de los factores, relacionados a los datos necesarios y modelos econométricos disponibles, considerando lo presentado en las secciones anteriores, que se tomaron en cuenta para el análisis empírico realizado para este trabajo.

En el resto de esta sección se detallan los puntos expuestos en el siguiente orden: se comienza por definir qué fuentes de información se utilizarán, explicando brevemente su elección. Luego se las describe y, con eso, se expone la argumentación para el uso de las técnicas econométricas elegidas para este estudio; a continuación, se reseñan brevemente esas técnicas usando como fuentes principales los trabajos de Cameron y Trivedi (2010) y Wooldridge (2002) para los modelos Probit y Logit, y los de Cameron y Trivedi (1998, 2010) para la familia de modelos con datos de conteo. Finalmente, se aborda brevemente el problema del sesgo de agregación y se plantea una posible manera de afrontarlo en el contexto de esta tesis.

6.1 De los datos

En la sección anterior se explicó que, siguiendo a Teodorescu (2000), el grupo de variables independientes a tomar en cuenta se dividirá en tres: i) atributivas; ii) de logro; e iii) institucionales. Es decir, esas variables deberán ser características de los investigadores y su contexto (instituciones en las que trabajan). Sin embargo, la identificación de los profesores de las universidades en Perú es compleja pues no existe un registro nacional de estos.

La información, más reciente y uniforme, sobre las características de docentes y universidades peruanos es la recopilada por la Superintendencia Nacional de Educación Universitaria (SUNEDU) durante al proceso de licenciamiento obligatorio en el que las universidades peruanas han tenido que participar; no obstante, esos datos no son públicos.

Por otro lado, algunas universidades peruanas publican información sobre sus profesores en sus páginas web institucionales. Los problemas con esas fuentes son: que no se tiene información sobre la fecha de su actualización; y que la información proveniente de esas páginas web es no uniforme, pues esta depende de la institución que la publicó, es decir, de cada universidad, facultad, departamento académico o escuela profesional.

Una tercera alternativa que se consideró como fuente para las características de docentes e instituciones fue el Censo Nacional de Investigación y Desarrollo a Centros de Investigación que se llevó a cabo el año 2016. La unidad de observación de ese censo es el "Centro de Investigación", que puede estar adscrito a otra unidad de

investigación más grande o a una facultad en una universidad (CONCYTEC, 2016), sin embargo, la información pública de ese censo es insuficiente para poder construir perfiles de los investigadores por institución.

Además de las fuentes antes descritas, existen dos Censos Nacionales Universitarios, llevados a cabo en los años 1996 y 2010. Si bien esos censos son anónimos a nivel individual, es posible identificar el número y características de docentes por cada institución, así como clasificar su labor dentro de un área del conocimiento. Además, cuenta con un módulo para las características institucionales de las universidades. Por lo tanto, se considera que, para las variables independientes en este estudio, la fuente más adecuada de la cual obtener características es el II Censo Nacional Universitario.

Por otro lado, dada la función de producción planteada, la variable de resultado, o dependiente, debe ser una medida de “producción” de la investigación realizada por los docentes dentro de una universidad. Las posibilidades abordadas en el marco teórico para llegar a una medición de esa variable son: i) medida de financiamiento obtenido para investigación, ii) conteo de publicaciones, iii) conteo de citas de publicaciones, iv) conteo de premios ganados, y v) conteo de patentes registradas; o una combinación de las anteriores. De estas opciones, la mayor parte de los trabajos empíricos reseñados optan por estudiar algún tipo de índice cuantitativo, siendo el principal de estos el número de las publicaciones de los investigadores. El motivo para eso es que la medición de un indicador de esa clase es relativamente menos compleja de implementar dado que puede recurrirse a bases de datos bibliométricas, como Scopus, Web of Science y Google Scholar, que registran información como número de publicaciones e índices de citas.

En Perú, el CONCYTEC implementó dos plataformas en las que se puede encontrar la producción de los investigadores: el Directorio Nacional de Investigadores e Innovadores (DINA) y el Registro de Investigadores en Ciencia y Tecnología del Sistema Nacional de Ciencia, Tecnología e Innovación Tecnológica (REGINA). Si bien ambas plataformas son de acceso libre y abierto, el registro en estas no siguió ningún proceso censal o administrativo masivo, sino que su uso depende por completo de cada investigador. Además, en ambos casos la orientación es hacia investigadores en ciencia y tecnología, por lo que no se cubriría el universo de las áreas del conocimiento. Los repositorios institucionales constituyen una segunda opción, pero estos también adolecen del carácter universal y estandarizado que se busca para este estudio.

Considerando una vez más el II CNU, se tiene que en este hay una sección específica para el registro de la producción intelectual de los docentes. A pesar de que el II CNU también se basó en autoreportes, sí tiene la característica de haberse aplicado

universalmente y, además, de guardar relación con las unidades de análisis ya consideradas para las variables independientes.

Finalmente, también se considerarán las bases de datos bibliométricas antes mencionadas, aunque será necesario seguir algunos procesos adicionales para no tener problemas con las unidades de análisis. A continuación, se describen las fuentes de información mencionadas.

II Censo nacional universitario – 2010 (II CNU)

El II CNU se constituye como la principal fuente de información para el presente trabajo. Como se ha mencionado, a partir de los microdatos de ese estudio estadístico, se construirán tanto las variables independientes (las características de los profesores y sus instituciones) para todos los análisis que se implementaran; como la variable dependiente para parte de esos análisis.

Según su Plan Directriz (INEI 2010), el método de recojo de información empleado para este Censo fue el auto empadronamiento, y este fue realizado por los alumnos, profesores y personal administrativo de la universidad, según corresponda. Como el trabajo de campo se realizó en 2010, las respuestas recogidas corresponden a aquellos alumnos, profesores y administrativos que estuvieron activos entre marzo y mayo de ese año. Cada una de las universidades fue responsable, con la asesoría o apoyo del INEI de ser necesario, de implementar una plataforma virtual para tal fin. No se tiene registro de que se haya realizado algún proceso de verificación posterior de la información obtenida.

Las bases de datos del II CNU incluyen seis módulos: del alumno de pregrado, del alumno de posgrado, del docente, del personal administrativo, del campus universitario, y de la sede central. Para los fines de este trabajo, se utilizarán los módulos 3 (contiene las características docentes a utilizarse como variables independientes y el reporte de número de publicaciones, que se usará como variable dependiente) y 6 (contiene las características institucionales de las universidades, que se usarán como variables independientes).

Módulo 3 del II CNU:

La unidad de análisis en este módulo es el “docente en una universidad” y recoge las respuestas de estos con respecto a sus características generales y académicas; sus situaciones administrativa y académica dentro en la universidad; sus actividades económicas no universitarias; sus actividades en investigación en los últimos dos años; y su opinión sobre la universidad. Siendo esa la unidad, los docentes que enseñan en más de una universidad tuvieron la posibilidad de enviar tantas cédulas como el número de universidades en las que trabajan. De los 59085 registros

presentes en la base de datos pública, 53739 son docentes únicos. El cuadro 6.1 muestra la distribución de frecuencias del número de cédulas enviadas por cada docente.

Cuadro 6.1
Número de docentes en las universidades activos en las universidades peruanas entre marzo y mayo de 2010, según el número cédulas enviadas

Número de cédulas distintas enviadas	Número de docentes	Porcentaje
1	49063	83.04%
2	4092	13.85%
3	515	2.61%
4	58	0.39%
5	8	0.07%
6	2	0.02%
9	1	0.02%
TOTAL	53739	100.00%

Fuente: II CNU. Elaboración propia.

Cabe señalar que, si bien el cuadro muestra que solo 4676 profesores enviaron más de una cédula censal, en total fueron 10780 profesores distintos los que contestaron afirmativamente a la pregunta: “¿trabaja usted en otra universidad?”. Además, el II CNU clasificó a los profesores universitarios según:

- Su condición laboral: profesor contratado, profesor ordinario, jefe de práctica, profesor extraordinario.
- Su régimen de dedicación a la docencia: dedicación exclusiva, tiempo completo, tiempo parcial.

En el cuadro 6.2 se muestra la distribución cruzada, de la base de datos total, con respecto a esas dos variables.

Cuadro 6.2
Número de docentes¹⁶ en las universidades activos en las universidades peruanas entre marzo y mayo de 2010, según condición laboral y dedicación

Condición laboral / Régimen de dedicación	Dedicación exclusiva	Tiempo completo	Tiempo parcial	TOTAL
Profesor ordinario	7839	7112	5535	20486
Profesor extraordinario	36	51	281	368
Profesor contratado	710	4136	29576	34422
Jefe de Práctica	262	673	2874	3809
TOTAL	8847	11972	38266	59085

Fuente: II CNU. Elaboración propia.

¹⁶ Al considerarse la totalidad de registros, los datos e el cuadro pueden ser entendidos como “números de puestos”, más que como “número de personas”.

Los cuadros 6.3 y 6.4 presentan una revisión conjunta del régimen de dedicación declarada por los docentes en el censo; el número de formularios que enviaron, que se interpretaría como el número de universidades en las que laboran; y sus respuestas a las preguntas: “*además de la ocupación que tiene en la universidad ¿realiza usted alguna otra ocupación no universitaria remunerada?*” y “*¿trabaja usted en otra(s) universidad(es)?*”. Además, el gráfico 6.1 muestra la distribución del total de horas reportadas por los docentes.

Cuadro 6.3
Número de docentes según cédulas enviadas y dedicación

Número de cédulas enviadas / Régimen de dedicación	Exclusiva	Tiempo completo	Tiempo parcial	TOTAL
1	8,530	10,236	30,297	49,063
2	285	1,514	6,385	8,184
3	27	199	1,319	1,545
4	5	22	205	232
5	0	1	39	40
8	0	0	12	12
9	0	0	9	9
TOTAL	8,847	11,972	38,266	59,085

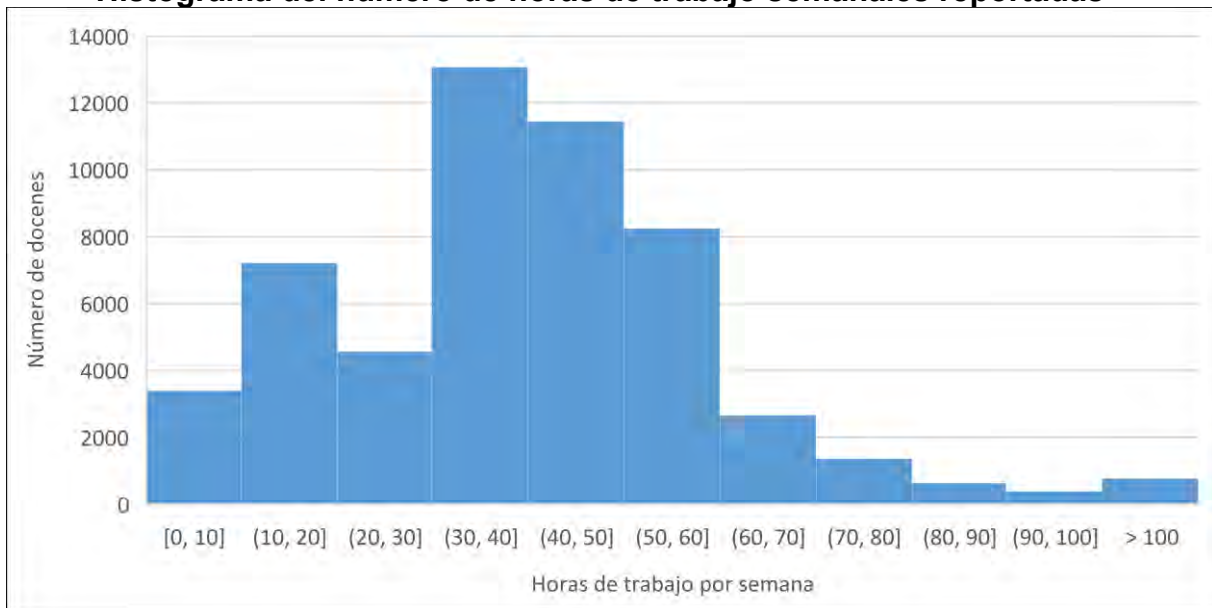
Fuente: II CNU. Elaboración propia.

Cuadro 6.4
Número de docentes según dedicación y respuestas a preguntas sobre labores en otras instituciones

Régimen de dedicación	Trabaja en otra universidad		Otra actividad remunerada		Total
	Sí	No	Sí	No	
Exclusiva	340	8,507	891	7,956	8,847
Tiempo completo	2,442	9,530	3,879	8,093	11,972
Tiempo parcial	12,589	25,677	24,082	14,184	38,266
TOTAL	15,371	43,714	28,852	30,233	59,085

Fuente: II CNU. Elaboración propia.

Gráfico 6.1
Histograma del número de horas de trabajo semanales reportadas¹⁷



Fuente: II CNU. Elaboración propia.

Los cuadros anteriores muestran ciertas inconsistencias en los datos, por ejemplo, profesores que reportaron tener dedicación exclusiva también reportaron trabajar en otras instituciones educativas o enviaron más de una cédula; profesores que reportaron números de horas de trabajo semanal mayores a 100 o iguales a cero. Las consideraciones tomadas al respecto se encuentran en el cuadro 6.5.

Por otro lado, para este trabajo, se considerará como población a los docentes que cumplan con las siguientes características:

- Profesores Ordinarios, sin importar su dedicación.
- Profesores Contratados, a tiempo completo o dedicación exclusiva.

Por lo descrito, antes de implementar los análisis econométricos, se realizaron los procesos resumidos en el cuadro 6.5, a partir de la base de datos pública inicial.

¹⁷ El cálculo del número total de horas de trabajo semanal se hizo considerando todas las actividades remuneradas, en universidades y fuera de estas, declaradas en todas las cédulas reportadas por los profesores

Cuadro 6.5**Resumen de los procesos realizados para construir la base de datos a analizar**

Proceso realizado	Número de registros que	
	Salen	Quedan
1) Base inicial	0	59085
2) Se retiran las cédulas con régimen de dedicación exclusiva, pero que enviaron más de una cédula	317	58768
3) Se retiran las cédulas con régimen de dedicación exclusiva y que contestaron que sí trabajan en otra universidad	164	58604
4) Se mantienen cédulas con $0 < \text{horas de trabajo semanal} \leq 80$ (sumando todas las horas de todas las cédulas del profesor)	3154	55450
5) Se retiran las cédulas de Jefes de Práctica	3677	51773
6) Se retiran las cédulas de Profesores Contratados a tiempo parcial	27517	24256
7) Se retiran las cédulas de Profesores Extraordinarios-Eméritos y Extraordinarios-Honorarios	103	24153
8) Se retiran todas las cédulas que no pudieron ser clasificadas en una a de las áreas OCDE	217	23936
9) Se retiran todas las cédulas de la Facultad Pontificia Civil de Lima	20	23916
10) Se retiran las cédulas sin información de región	3	23913
11) Se retiran las cédulas con errores en año de nacimiento	99	23814
12) Se retiran las cédulas sin información de país de pregrado	2	23812
13) Se retiran las cédulas sin información de país de posgrado	1	23811
14) Se retiran las cédulas con error en número de personas que dependen económicamente del profesor	1	23810
15) Se retiran las cédulas de docentes presentes en más de una universidad, pero indicaron NO trabajar en otra universidad	211	23599
16) Se retiran cédulas de docentes con cero horas de dedicación a todas las actividades	1	23598
17) Se retiran cédulas de docentes cuyas universidades no pueden clasificarse como públicas, privadas asociativas o privadas societarias	64	23534 ¹⁸

Fuente: II CNU. Elaboración propia.

De esa manera, el universo de docentes que serán considerados para implementar los análisis econométricos pasó a ser 23670. Con relación a la información sobre investigación, se tuvo en cuenta (o se empleó) las siguientes restricciones:

¹⁸ En 72 casos, los docentes fueron clasificados en dos áreas del conocimiento y fueron considerados en ambas, por lo que el número final de observaciones se incrementó a 23606.

- Un mismo docente puede haber reportado una investigación en más de una ocasión, asociando ese producto a más de una universidad, consignado el mismo título en todas las cédulas entregadas.
- Un mismo docente puede haber reportado una investigación en más de una ocasión, asociando ese producto a más de una universidad, consignando títulos distintos en cada cédula entregada.
- No en todas las cédulas en las que los docentes señalaron realizar investigación se consignó el detalle de esta (nombre, tipo, medio en el que se publicó, etc.).

Por ese motivo, y considerando la naturaleza de las otras fuentes de información a utilizar, se contará como investigación realizada a toda aquella que haya sido reportada, sin importar si esta fue informada por el mismo docente en más de una cédula. Además, se contarán solo las investigaciones reportadas como “finalizadas”.

El cuadro 1A (ver anexo) muestra el total de investigaciones, según tipo, por universidad; mientras que en el cuadro 2A (ver anexo) se tiene el número de investigaciones reportadas como difundidas en revistas indexadas, también por universidad. Las variables que considerar para cada análisis serán descritas en la sección 7.

Base de datos bibliográfica

Inicialmente, la única fuente grande de datos bibliométricos en base a los cuales se construían indicadores era la del *Institute for Scientific Information* (ISI, actualmente con el nombre de *Web of Science* -WoS-). En noviembre de 2004 se lanzaron dos nuevas bases de datos bibliográficas: por un lado, la compañía editora Reed Elsevier lanzó *Scopus*, y la empresa Google presentó *Google Scholar* (Bar-Ilan, 2008). Cabe señalar que WoS y Scopus nacieron, y se mantienen, como productos restringidos para sus suscriptores, mientras que Google Scholar es un servicio de acceso gratuito y abierto. La fortaleza de los dos primeros sistemas sobre el tercero es el cuidado que tienen al indizar revistas, libros, resúmenes, editoriales, etc. (Harzing y Alakangas, 2016). Uno de los mecanismos que las dos plataformas restringidas utilizan para asegurar la calidad de las fuentes que indizan es, además de revisiones administrativas, comprobar que los procesos de selección de artículos llevados a cabo por esas fuentes sean arbitrados.

Se han llevado a cabo múltiples estudios cuyo objetivo ha sido comparar esas bases de datos; si bien la mayoría de estos se han enfocado solo en un área o una universidad, o en un grupo limitado de investigadores o revistas, también hay algunos que han analizado conjuntos más amplios (Harzing y Alakangas, 2016; Martin-Martin y otros, 2018). Una de las razones para esto puede haber sido la relativa complejidad que supone obtener los microdatos de esas fuentes. Cabe señalar también que tanto

Scopus como WoS, además de ser plataformas de acceso restringido para sus suscriptores institucionales, tienen herramientas adicionales de análisis bibliométrico, InCites en el caso de WoS y SciVal para Scopus, que requieren el pago de suscripciones institucionales adicionales.

El uso de esas fuentes no está libre de críticas. Una de las más importantes es que existen sesgos al momento de indizar un *journal*. Por ejemplo, Chavarro y otros (2018) hacen uso de la regresión logística para calcular la probabilidad de que un *journal* con ciertas características resulte indizado por WoS, encontrando que tanto el idioma como el país o la disciplina son características relevantes al momento de calcular esa predicción; sin embargo, tomarlas como referencia al momento de realizar evaluaciones cuantitativas de la producción científica es el estándar actual cuando no se cuenta con fuentes más adecuadas (los autores mencionan casos como el de Noruega u Holanda donde se han implementado sistemas cuantitativos propios para evitar esos sesgos al momento de evaluar la producción científica de sus investigadores).

Por lo expuesto, se decidió no considerar a Google Scholar como una de las alternativas para elegir la fuente a utilizar, quedando como opciones Scopus y WoS. La elección de una de esas dos plataformas se realizó tomando en cuenta la situación descrita en las tablas 6.6, 6.7 y el gráfico 6.2.

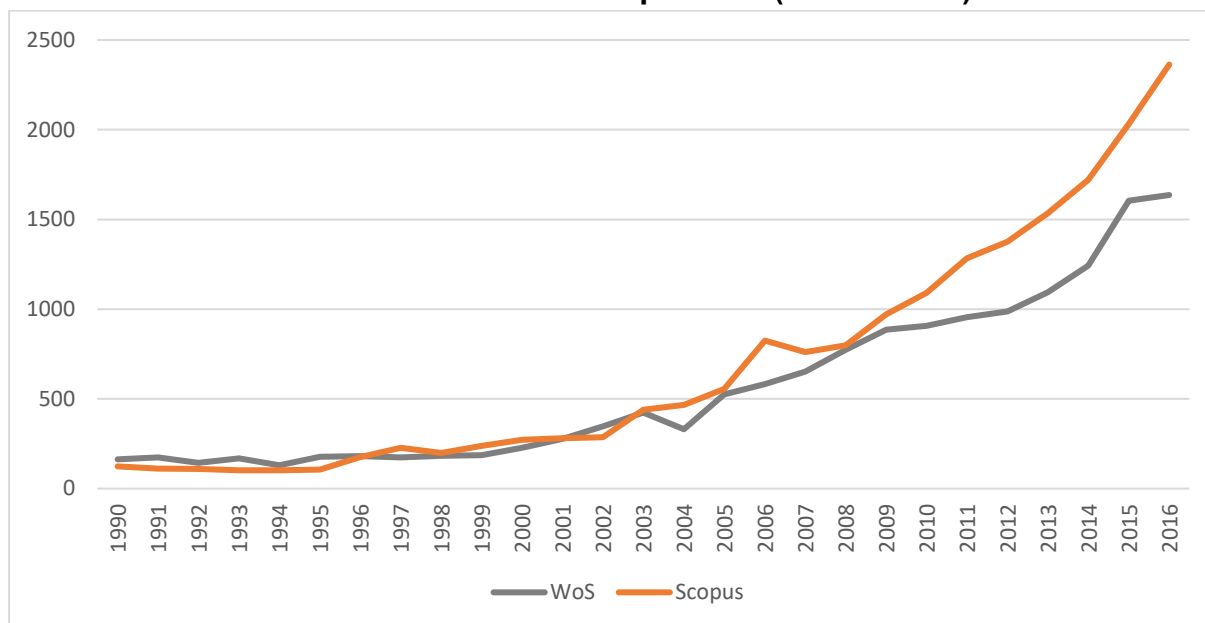
Cuadro 6.6
Criterios para elección de base de datos bibliométrica

Criterio	Scopus	Web of Science
Acceso (a la plataforma y/o a herramientas de análisis)	Solo a plataforma general	Solo a plataforma general
Capacidad de descarga de metadatos de las publicaciones	Sí, mediante búsquedas individuales	Sí, mediante búsquedas individuales
Identificación de instituciones	Con número ID asignado	Por nombre estandarizado
Número de universidades peruanas presentes	59	53

Fuente: Elaboración propia.

El gráfico 6.2 muestra la evolución del número de publicaciones con filiación a instituciones peruanas en ambas plataformas desde 1990 hasta 2016. Puede notarse que la relación entre ambas series de datos es alta (el cociente de correlación de Pearson para esos 27 años es 0.99), y que, a partir del 2007, la tasa de crecimiento del número de artículos en Scopus es superior a la de WoS. Por otro lado, el cuadro 6.7 muestra en detalle, para los años 2007 – 2016, el ratio entre ambas series. Para ese periodo, Scopus ha registrado 30% más artículos, con instituciones peruanas en las filiaciones, que WoS.

Gráfico 6.2
Número de artículos registrado en WoS y Scopus con al menos un autor afiliado a una institución peruana (1990 – 2016)



Fuente: RICYT. Elaboración propia.

Cuadro 6.7
Ratio entre número de documentos con filiación peruana en Scopus y WoS (2007 - 2016)

Año	WoS	Scopus	Ratio ¹⁹
2007	651	760	1.17
2008	773	799	1.03
2009	885	971	1.10
2010	907	1090	1.20
2011	955	1283	1.34
2012	987	1376	1.39
2013	1094	1535	1.40
2014	1242	1721	1.39
2015	1605	2032	1.27
2016	1636	2363	1.44
TOTAL	10735	13930	1.30

Fuente: RICYT. Elaboración propia.

Entonces, tomando en consideración que la correlación entre los números de publicaciones hechas por autores afiliados a instituciones peruanas presentes en WoS y Scopus es muy cercana a uno, y que esta última plataforma ha tenido sistemáticamente una mayor cobertura de estos; se decidió optar por Scopus como la fuente que se utilizará para construir la base de datos de producción de investigación por universidad.

Para cumplir ese objetivo, se siguieron los siguientes pasos:

¹⁹ El ratio es igual al cociente de dividir los documentos en Scopus sobre los documentos en WoS.

1. Se identificaron cuáles de las 100 universidades presentes en el II CNU tienen al menos un documento en Scopus que la incluya en las filiaciones institucionales de sus autores.
2. Se descargaron todos los registros para cada una de esas las universidades. Esto incluyó todos los tipos de publicaciones y todos los años²⁰.
3. Se eligieron para el conteo solo los registros del tipo *article* (artículos) o *conference paper* (artículos en conferencias). El motivo para esto es que son esos los tipos de documentos que pueden ser clasificados en algún área del conocimiento²¹.
4. Los datos resultantes fueron clasificados temáticamente con base en:
 - Artículos: Tomando en cuenta el área, según la base de datos del *Scimago Journal & Country Rank*²² – plataforma web asociada a Scopus-, tomando el área del conocimiento que se le asignó a la revista en la que se publicó el artículo.
 - Artículos en conferencias: Según los filtros disponibles en la plataforma de Scopus.
5. La clasificación resultante, que incluía 26 temas distintos en total, fue adaptada por el autor a la clasificación de seis áreas usada por la OCDE según el cuadro 4A (ver anexo).
6. Se eligió el periodo de evaluación: de 2011 a 2014.
 - Debido a que el trabajo de campo para el II CNU fue llevado a cabo el año 2010, y por lo tanto las características docentes que fueron recogidas en ese censo corresponden a ese año; y que la Nueva Ley Universitaria fue aprobada el año 2014, por lo que es razonable asumir que la mayoría de universidades peruanas tuvieron, a partir de 2014, un incentivo fuerte para implementar cambios significativos en sus planas docentes. El cuadro 6.8 permite ver la distribución de los artículos a considerar.
7. Finalmente, se realizó el conteo simple del número de publicaciones que cada área dentro de una universidad (en adelante universidad-área) posee.

Cuadro 6.8
Distribución del número de artículos publicados entre 2011 y 2014 según régimen de universidad y área OCDE

Área OCDE	Régimen de la universidad			Total	
	Asociativa	Pública	Societaria	Suma	%
1_Ciencias Naturales	673	472	34	1179	26%

²⁰ Se realizó *web scraping* a la plataforma de Scopus. Un resumen del proceso, que incluye el algoritmo, utilizado puede verse en el diagrama 3A (ver anexo).

²¹ Para este trabajo se utilizó la clasificación de la Organización para la Cooperación y el Desarrollo Económicos (OCDE 2007).

²² Disponible en: <https://www.scimagojr.com/journalrank.php>

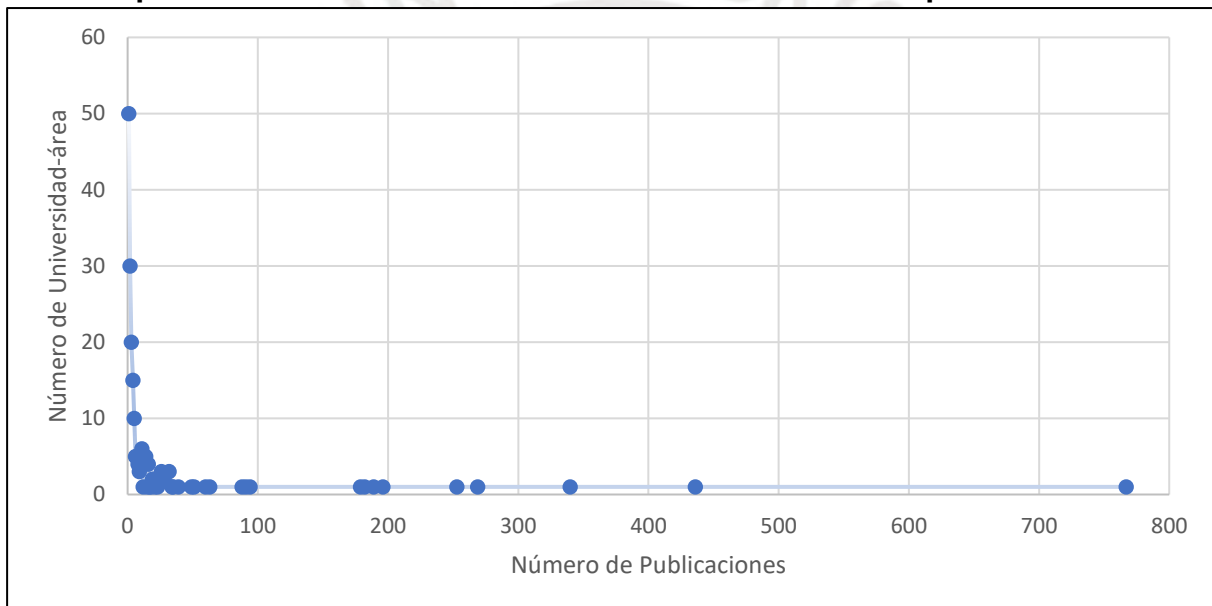
2_Ingeniería y tecnología	258	179	13	450	10%
3_Medicina	893	402	114	1409	31%
4_Ciencias Agrarias	266	758	23	1047	23%
5_Ciencias Sociales	306	62	27	395	9%
6_Humanidades	57	21	7	85	2%
Total	2453	1894	218	4565	100%

Fuente: Scopus. Elaboración propia.

Como resultado del proceso descrito, se obtuvo una base de datos con 202²³ registros de universidad-área y el número de publicaciones, artículos o artículos de conferencia, con al menos un autor afiliado a esa institución, en Scopus. El gráfico 6.3 muestra, como era de esperarse por la ley de Lotka, una gran concentración de universidades-área con pocas publicaciones.

Gráfico 6.3

Dispersión: número de universidad-área vs número de publicaciones



Fuente: Scopus. Elaboración propia.

6.2 Opciones de modelos econométricos

Dada la información disponible descrita en las secciones anteriores, las opciones para las estimaciones econométricas del modelo de producción propuesto se resumen en el cuadro 6.9.

²³ Los 202 registros no incluyen aquellas universidad-área sin publicaciones en Scopus (estas vendrían a ser los “ceros” de la variable de conteo).

Cuadro 6.9
Opciones para el análisis econométrico

Variable dependiente²⁴	Variabes independientes	Fuentes de información	Análisis a implementar
Dicotómica: Publicó / No publicó	Características individuales de los docentes y universidades	II CNU	Probit / Logit
Conteo: Número de publicaciones	Características promedio, para cada universidad-área, de los docentes y universidades	II CNU, Scopus	Poisson, Binomial Negativo, Zero inflated, Zero truncated

Fuente: Elaboración propia.

El utilizar modelos con variable dependiente discreta y de conteo, será posible responder a las preguntas: ¿produce investigación? y ¿cuánto?, pues los modelos con variable dependiente binaria servirán para medir la “propensión a tener al menos una publicación indizada” de cada uno de los docentes; mientras que los modelos de conteo buscarán medir la intensidad de producción.

Modelos con variables dependientes binarias: Probit y Logit

Pertencen a la familia de modelos paramétricos y, siguiendo a Cameron y Trivedi (2010), se utilizan cuando la variable dependiente y toma dos valores según el siguiente comportamiento:

$$y = \begin{cases} 1 & \text{con probabilidad } p \\ 0 & \text{con probabilidad } 1 - p \end{cases}$$

Por lo tanto, la función de distribución de esta probabilidad puede expresarse de la siguiente manera:

$$\Pr(y) = p^y(1 - p)^{1-y}$$

Los dos primeros momentos para y pueden generalizarse y son:

$$\begin{aligned} E[y] &= p \\ \text{Var}[y] &= p(1 - p) \end{aligned}$$

²⁴ Debe resaltarse que, al ser unidades de análisis diferentes, los conjuntos de variables explicativas a utilizar para las dos opciones para la variable dependiente, dicotómica o de conteo, también serán distintos.

El objetivo es modelar la probabilidad p a partir de las variables explicativas relevantes haciendo uso de alguna forma funcional adecuada. En el contexto de este trabajo, estas variables explicativas serían las características de profesores e instituciones. Si se define al vector de características como x y al vector de los coeficientes que resulten de la regresión como β , la probabilidad condicional a las características para un y_i en particular tendrá la siguiente forma:

$$p_i = \Pr(y_i = 1|X) = F(x'_i\beta)$$

Al ser $F(.)$ una función de distribución, el valor de esa probabilidad condicional estará restringido a $0 \leq p \leq 1$.

La diferencia entre los modelos Probit y Logit está en la forma funcional que asumen para $F(.)$. Los modelos Probit toman la forma de una distribución normal estándar, mientras que los Logit, la de una distribución logística. La selección de cuál de los dos resultados seleccionar se realizará, siguiente nuevamente a Cameron y Trivedi (2010), comparando el nivel de verosimilitud alcanzado por cada uno de ellos.

Se debe tener en cuenta que los coeficientes obtenidos por estos métodos no corresponden a los efectos marginales de las variables explicativas correspondientes sobre la variable dependiente, estos efectos podrían ser obtenidos derivando, resultando valores distintos para cada individuo con características x_i .

Modelos con datos de conteo

La familia de modelos con datos de conteo aborda adecuadamente el comportamiento de variables dependientes con las siguientes características:

- Enteros no negativos.
- En su mayoría, con valores iguales o cercanos a cero.

Esta familia de modelos suele usar las distribuciones Poisson o Binomial Negativa para explicar el comportamiento de la variable de conteo. Además, pueden realizarse etapas de análisis auxiliares cuando el proceso generador de esa variable dependiente tiene comportamientos específicos de manera tal que no se consideren las observaciones para las cuales tiene el valor de cero, o que se considere solo una parte de estas. En el siguiente cuadro se muestra un resumen de seis de los modelos de conteo más usuales²⁵.

²⁵ Para todos ellos es posible también utilizar técnicas de estimación robusta de la varianza para violaciones del supuesto de equidispersión, como se explica luego.

Cuadro 6.10
Resumen de modelos de conteo, distribuciones estadísticas que usan y sus características

		Distribución de variable dependiente	
		Poisson	Binomial negativo
Modelo	Simple	Considera todos los datos Usa distribución Poisson para explicar dependiente	Considera todos los datos Usa distribución Binomial negativa para explicar dependiente
	Zero - truncated	No considera ceros Usa distribución Poisson para explicar dependiente	No considera ceros Usa distribución Binomial negativa para explicar dependiente
	Zero - inflated	Considera solo parte de los ceros Usa distribución Poisson para explicar dependiente	Considera solo parte de los ceros Usa distribución Binomial negativa para explicar dependiente

Fuente: Elaboración propia.

Modelos con distribución de Poisson

Para que la variable explicada o dependiente (Y) siga la distribución de Poisson debe cumplir con las siguientes propiedades:

- Equidispersión: Los dos primeros momentos, el valor esperado y la varianza, de la variable deben ser similares. De no ser el caso, la variable presentaría sobredispersión o subdispersión y no seguiría el proceso de Poisson.
- No exceso de ceros: De tenerse un número excesivo de ceros, debe recurrirse a los modelos de “truncamiento” o de “inflación”.

Y seguir el siguiente proceso:

$$p(Y = y) = \frac{e^{-\mu} \mu^y}{y!}$$

Para $y = 0, 1, 2 \dots$; donde el parámetro de intensidad o *rate* es:

$$\mu = \exp(X'_i \beta)$$

Siendo X el vector de variables explicativas de la dependiente. Por la forma de la función que define a Y , μ también es igual a sus dos primeros momentos, es decir, el parámetro de intensidad es igual al conteo esperado. En este punto, debe hacerse notar que, siendo la varianza igual a la exponencial en la ecuación anterior, el modelo de Poisson es intrínsecamente heterocedástico, pues el valor varía según los valores dentro del vector de características β (Deb y otros 2017).

Violación del supuesto de equidispersión

De una manera más general, puede tomarse que la varianza no es necesariamente igual a la media, sino que sigue la siguiente forma funcional:

$$\text{var}(y|x) = \mu + \alpha\mu^2$$

Según el valor de α , la variable presentará:

- $\alpha = 0$, equidispersión. Este caso corresponde a la distribución Poisson.
- $\alpha > 0$, se tiene sobredispersión.
- $\alpha < 0$, se tiene subdispersión.

El valor de α es determinado por una prueba de sobredispersión cuyas hipótesis son las siguientes:

$$H_0: \alpha = 0$$

$$H_1: \alpha \neq 0$$

Se han desarrollado diversos métodos para abordar los casos en los que se rechaza la hipótesis nula anterior. En este trabajo se hará uso de: la distribución Binomial Negativa y la distribución Poisson con estimación robusta de varianzas, siguiendo lo recomendado por Cameron y Trivedi (2009).

Para lidiar con el otro posible problema, el exceso de ceros, es posible utilizar cualquiera de las dos alternativas a los modelos base que se presentaron en el cuadro resumen: el modelo *Zero – truncated* y el modelo *Zero – inflated*.

Modelo zero - truncated

Considera que las observaciones cuyas características corresponden a un $y = 0$ no responden a un proceso de generación diferente a las observaciones que tienen $y > 0$. Para analizar ese comportamiento, asume dos funciones diferentes, f_1 y f_2 , como las generadoras de los datos observados. La primera de esta es responsable de los conteos iguales a cero y la segunda, de los positivos. El proceso total, $g(y)$, será determinado como sigue:

$$g(y) = \begin{cases} f_1(0); y = 0 \\ \frac{1 - f_1(0)}{1 - f_2(0)} f_2(-1); y \geq 1 \end{cases}$$

Puede notarse que los modelos base son el resultado de asumir que esos dos procesos corresponden a una única función, es decir, $f_1 = f_2$. Para la estimación de la primera etapa, en la que se termina la función de probabilidad responsable de los ceros, puede utilizarse los modelos Probit o Logit.

Modelo zero - inflated

En este modelo también se consideran dos posibles funciones generadoras para los datos observados. La diferencia con el modelo *Zero – truncated* es que, en este caso, la función responsable de las observaciones positivas lo es también de parte de los ceros observados. El proceso total de la generación puede escribirse como:

$$g(y) = \begin{cases} f_1(0) + (1 - f_1(0))f_2(0); y = 0 \\ (1 - f_1(0))f_2(x); y \geq 1 \end{cases}$$

Entonces, los ceros observados tienen dos fuentes generadoras, como resultado de un proceso con resultados binarios, representado por $f_1(\cdot)$ en el esquema anterior, y como valor de cuenta igual a cero para un proceso de conteo, representado por $f_2(\cdot)$.

Para la interpretación, tanto para los modelos que utilizan la distribución Poisson como para los que usan la distribución Binomial Negativa, se parte de la relación señalada para el valor esperado condicional:

$$E(Y_i/X_i) = \exp(X'_i\beta)$$

De manera que, para una variable explicativa *dummy*, β representa el logaritmo natural del número de veces por el que se multiplicaría el valor esperado condicional $E(Y_i/X_i)$ si X_i valiese 1 en lugar de 0. Si la variable explicativa fuese continua en lugar de *dummy*, el coeficiente se interpretaría como una semielasticidad: 100β viene a ser el porcentaje en el que el valor esperado condicional cambia ante el incremento en una unidad de la variable independiente.

A propósito de las interpretaciones señaladas en el párrafo anterior, también suele definirse como un indicador al “ratio de incidencia”, que es igual a $\exp(\beta)$, por lo que su interpretación directa es: el número de veces por el que se multiplica $E(Y_i/X_i)$.

Por otro lado, es posible calcular los efectos marginales de la siguiente manera:

$$\frac{\partial E(y|x)}{\partial x_j} = \beta_j \exp(X'_i\beta)$$

Es decir, “una unidad de incremento en X incrementa a la media de Y en el porcentaje igual al coeficiente”.

6.3 El sesgo de agregación

En este trabajo se hará uso de datos individuales agregados a nivel de Universidad - Área de conocimiento para los análisis de conteo, pues no se tiene información suficiente para identificar qué publicaciones indizadas en Scopus pertenecen a cada profesor dentro del II CNU. Las variables que se utilizarán para la estimación de esos modelos se construirán como promedios de los datos individuales dicotómicos (esto también puede interpretarse como porcentaje).

Debido a eso, es posible caer dentro del llamado “problema de agregación de individuos” si se pretende modelar comportamientos individuales a partir de esos datos. Clark y Avery (1976) sugieren hacer lo posible por tomar un sistema agregado en el que se pierda lo menos posible la información individual, esto es, agrupar a los individuos de la forma más homogénea posible.

Por otro lado, Stoker (1993) también aborda esa situación y propone algunas vías posibles para lidiar con ella. Menciona que pueden seguirse tres aproximaciones al tratar con modelos de datos agregados: i) modelamiento de solo datos agregados, ii) microsimulaciones y iii) modelamiento con datos agregados e individuales. Si bien su artículo se enfoca en aplicaciones a modelos macroeconómicos, se considera que los análisis y recomendaciones que presenta son relevantes para el presente estudio.

En el primer tipo de aproximación que Stoker realiza, la de los modelos con agentes representativos, se asume que las decisiones de un grupo son tomadas por un agente que se comporta como todos ellos. Los resultados no deberían ser asumidos como válidos a nivel desagregado, pues de hacerlo se estaría implicando que las decisiones de todos los individuos son similares porque no se está tomando en cuenta la heterogeneidad dentro de los grupos. Debe cuidarse que las conclusiones de esos modelos se interpreten solo a nivel agregado.

En cuanto a las microsimulaciones, estas plantean el modelamiento del comportamiento de cada uno de los individuos en los datos. Estos comportamientos individuales serán luego agregados para construir el comportamiento del conjunto. Si bien esta aproximación al problema es más rica y detallada que la agregación simple, tiene en contra que su implementación es compleja, no solo por la disponibilidad de información, sino también porque este se requiere definir un conjunto de variables relevantes para cada individuo por separado, para luego buscar alguna manera de agregar esos análisis individuales. Finalmente, la tercera aproximación que se aborda propone mezclar modelos individuales dentro de los grupos, con un modelo que considere a los grupos como unidad.

7. ESTIMACIÓN Y RESULTADOS

Se implementaron todos los análisis resumidos en el cuadro 6.9 y en esta sección se presentarán los resultados, luego de explicar la elección y construcción de las variables para los modelos.

En primer lugar, se tomó en cuenta la disponibilidad de datos de características individualizadas por docente que podían obtenerse del II CNU. Dado que la variable dependiente disponible en esa fuente es un auto-reporte de la producción de investigación, medida como cantidad de publicaciones, sin ningún proceso adicional de verificación posterior, y no por revisión de bases de datos bibliométricas, como es el estándar; se consideró que, en ese contexto, la mejor alternativa son los modelos con variable dependiente binaria y se realizaron las estimaciones utilizando los modelos Probit y Logit considerando a la condición “publicó al menos un artículo en revista indizada” como la variable dependiente.

En segundo lugar, se recurrió a información obtenida de Scopus para determinar la producción, medida como el conteo de publicaciones, de cada área-universidad²⁶ desde 2011 hasta 2014. A eso se agregaron las características promedio de los docentes y de las universidades disponibles en el II CNU. La base de datos resultante se utilizó para implementar estimaciones con los modelos de conteo descritos. Se incluyeron análisis adicionales tomando en cuenta la diferencia entre perfiles de docentes con base en su formación académica.

Si bien la variable dependiente, que mide el nivel de producción de investigación, ya sea en una escala dicotómica de publicó o no publicó, o como un conteo de los artículos indizados en Scopus, ha quedado ya correctamente definida y justificada en el capítulo anterior, aún no se ha hecho lo propio con las variables independientes.

7.1 Variables para la estimación

La elección de las variables independientes se realizó tomando en cuenta el objetivo de este trabajo: contribuir a la discusión sobre la calidad universitaria en el Perú. Además, de acuerdo con las hipótesis de investigación, se consideraron variables que midan características pertenecientes a los tres grupos utilizados por Teodorescu (2000): atributivas, de logro, e institucionales.

Por otro lado, gracias a la revisión de estudios empíricos previos, pudo identificarse un conjunto de factores asociados a la producción de investigación resumidos en el

²⁶ En cada área-universidad se asume como producción el total de artículos indizados a Scopus que presenten al menos un autor con filiación institucional a la universidad en cuestión. Las características docentes que se tomarán en cuenta como variables explicativas son el promedio de esas características en cada área-universidad (esto es equivalente a calcular el porcentaje de profesores que cumplen cierta característica).

cuadro 3.1. A partir de esos resultados, en el cuadro 6.1 se presentan las variables presentes (y disponibles) en el II CNU que miden aspectos similares a los abordados en esos estudios, además de clasificarlos en cada uno de los tres grupos. Cabe señalar que los trabajos reseñados utilizaron diversas unidades de análisis, dependiendo del enfoque y alcance que buscaban alcanzar, por lo que la matriz de variables del II CNU es una adaptación para este estudio.

Cuadro 7.1
Variables disponibles en el II CNU

Grupo	Variables en II CNU	Valores en II CNU
Atributivas	Número de personas dependientes económicamente	Número entero
Atributivas	Es mujer	Dummy, 1 = Sí
Atributivas	Edad en años	Número entero
Atributivas	Es casado / conviviente	Dummy, 1 = Sí
Atributivas	Pertenece a Lima o Callao	Dummy, 1 = Sí
Atributivas	Pertenece al Área OCDE 1	Dummy, 1 si pertenece al Área 1
Atributivas	Pertenece al Área OCDE 2	Dummy, 1 si pertenece al Área 2
Atributivas	Pertenece al Área OCDE 3	Dummy, 1 si pertenece al Área 3
Atributivas	Pertenece al Área OCDE 4	Dummy, 1 si pertenece al Área 4
Atributivas	Pertenece al Área OCDE 5	Dummy, 1 si pertenece al Área 5
Atributivas	Pertenece al Área OCDE 6	Dummy, 1 si pertenece al Área 6
De logro	Sabe inglés	Dummy, 1 = sí
De logro	Estudio fuera de Perú	Dummy, 1 = Sí
De logro	Tiene Doctorado	Dummy, 1 = Sí
De logro	Desempeña cargo administrativo	Dummy, 1 = Sí
De logro	Trabaja en otra universidad	Dummy, 1 = Sí
De logro	Realiza actividades remuneradas no universitarias	Dummy, 1 = Sí
De logro	Tiempo completo o dedicación exclusiva	Dummy, 1 = Sí
Institucionales	Ingreso por la universidad	Número
Institucionales	Ingreso por actividades fuera de la universidad	Número
Institucionales	Dispone de ambiente de trabajo en universidad	Dummy, 1 = Sí
Institucionales	Cree que investigación de su universidad contribuye a sociedad	Dummy, 1 = Sí
Institucionales	Tiene perspectiva positiva sobre su universidad	Dummy, 1 = Sí
Institucionales	Horas totales de trabajo en universidad	Número entero

Institucionales	Horas haciendo investigación en universidad	Número entero
Institucionales	Horas enseñando en universidad	Número entero
Institucionales	Horas de trabajo fuera de universidad	Número entero
Institucionales	Universidad tiene plan estratégico	Dummy, 1 = Sí
Institucionales	Financiamiento total	Número entero
Institucionales	Universidad tiene convenios de investigación	Dummy, 1 = Sí

Fuente: II CNU. Elaboración propia.

Debe mencionarse que las variables institucionales que provienen del Módulo 6 del II CNU no tienen información para cuatro universidades con profesores que enviaron información para el Módulo 3, por lo que al incluirse estas disminuirá el número de observaciones, ya sean estas docentes o unidades de Universidad-área del conocimiento.

Por otro lado, se presenta a continuación cuáles de las variables mostradas están relacionadas a lo planteado en la hipótesis de investigación:

- Uso del tiempo (T): Desempeña cargo administrativo, trabaja en otra universidad, es docente a tiempo completo o con dedicación exclusiva, número total de horas (que le dedica a cada actividad).
- Motivaciones (M): Ingreso por la universidad (extrínseca), dispone de un ambiente de trabajo en la universidad (extrínseca), Cree que la investigación de su universidad contribuye a sociedad (intrínseca), Tiene perspectiva positiva sobre su universidad (intrínseca).
- Tipo de institución (I): si es pública, privada asociativa o privada societaria; si pertenece a Lima o Callao o a otra región del país²⁷.
- Financiamiento (S): Financiamiento total que tuvo la universidad para investigación.
- Entrenamiento para investigación (D): Obtuvo el doctorado.

Las demás variables (edad, sexo, manejo de inglés, área del conocimiento) servirán como controles.

Debe indicarse también que las escalas de medición no se afectarán cuando se estime los modelos Probit y Logit, pues estos tomarán como unidad de análisis a cada docente, pero pasarán a ser promedios (interpretados como porcentajes, en el caso de las variables dicotómicas) cuando se implementen los modelos de conteo, dado que la unidad de análisis pasará de ser el profesor en cada universidad a ser el conjunto de profesores dentro de una Universidad-área definida. En los siguientes párrafos se presentan algunas consideraciones, así como los resultados esperados, y su justificación, para cada variable.

²⁷ Esta información se obtuvo del portal de SUNEDU: <https://www.sunedu.gob.pe/>

Ingreso por actividad y variables relacionadas

Para aproximarse a la medición del nivel de actividades económicas no universitaria se cuenta con tres variables: Ingreso monetario por labores no universitarias (como valor absoluto o como el logaritmo de este); horas de trabajo no universitarias; y realiza actividades remuneradas no universitarias (dicotómica). Las tres variables, como es de suponerse, están altamente correlacionadas:

Cuadro 7.2
Matriz de correlaciones entre variables de medición de actividades fuera de la universidad

Correlación	Ingresos por actividades fuera de la universidad	Log del ingreso por actividades fuera de la universidad	Horas de trabajo fuera de universidad	Realiza actividades remuneradas no universitarias
Ingresos por actividades fuera de la universidad	1			
Log del ingreso por actividades fuera de la universidad	0.7172	1		
Horas de trabajo fuera de universidad	0.6594	0.8288	1	
Realiza actividades remuneradas no universitarias	0.6149	0.9675	0.7952	1

Fuente: II CNU. Elaboración propia.

De entre esas tres variables, la que contiene más información y precisión es el número de horas de trabajo fuera de la universidad. En el caso de la variable “ingreso por actividades fuera de la universidad”, los valores base están dados como intervalos (se tomaron las marcas de clase de cada intervalo para las estimaciones en este trabajo); mientras que la variable “realiza actividades remuneradas no universitarias” es dicotómica²⁸.

²⁸ También se implementaron modelos incluyendo las otras variables para probar robustez, los resultados se mantienen.

Características adscriptivas

Como se mencionó, son aquellas sobre las que el individuo no tiene control, o al menos no uno relacionado con el estudio en cuestión, pero que pueden estar afectar mediante algún mecanismo a la variable resultado. En este caso se tiene: el número de personas dependientes económicamente, el sexo, la edad, el estado civil, encontrarse en Lima o Callao o no, y el área del conocimiento al que dedica su actividad académica²⁹.

En cuanto a la variable sexo, Rørstad y Aksnes (2015) y Teodorescu (2000) encontraron que el ser mujer está relacionado negativamente con la producción de artículos científicos, y van Arensbergen y otros (2012) señalan que esa diferencia estaría reduciéndose para académicos más jóvenes. Sobre la edad del investigador, Rørstad y Aksnes (2015) encuentra resultados mixtos según la disciplina en cuestión, mientras que Teodorescu (2000) encuentra que sus efectos son negativos; Rørstad y Aksnes (2015) también consideran el valor de esa variable al cuadrado, encontrando coeficientes negativos para todas las áreas.

Sobre las dos variables relacionadas a la situación familiar, el estudio de Shin y Cummings (2010) mostraba que el número de hijos no resultó ser una variable significativa para las publicaciones realizadas por profesores surcoreanos. Sin embargo, puede argumentarse que ambas variables tienen relación con la demanda familiar del tiempo del investigador, para el caso de las personas dependientes, estas requerirían parte del tiempo del investigador, mientras que tener una pareja puede constituir un factor que reduciría ese tiempo necesario, al compartir responsabilidades con una persona más.

Por otro lado, si bien el encontrarse laborando en Lima o Callao (o la capital del país, según sea el caso) no es una variable directamente analizada en los estudios reseñados, esta se relaciona con varias características abordadas por Mueller y otros (2016), Ruedas y Rodenes (2015) como el número de universidades, tamaño del mercado, cultura motivadora, etc. Para todos los casos, los efectos encontrados fueron positivos, por lo que se esperará lo mismo en este trabajo.

Características de logro

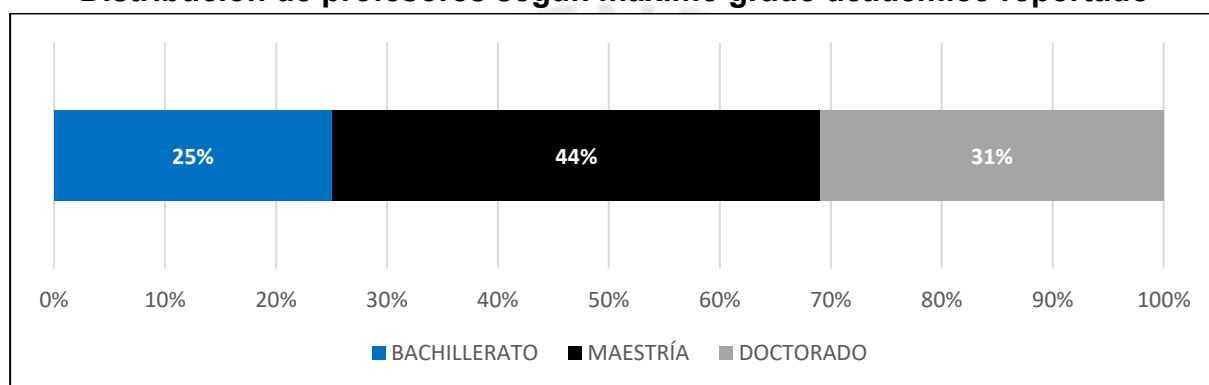
Se trata de características relacionadas con el “*desenvolvimiento del docente en la vida académica*” (Teodorescu 2000). Así, se empieza considerando el manejo del inglés, dado que existe un sesgo a favor de ese idioma al momento de indizar revistas en bases bibliográficas universales (Chavarro y otros 2018), se espera que su

²⁹ Esa última variable se considera exógena porque, si bien el investigador ha elegido a qué área dedicarse, él no elige el comportamiento cuantitativo de esa área.

conocimiento tenga un efecto positivo en la publicación de investigaciones. Por otro lado, haber realizado estudios fuera del país ayuda a crear y fortalecer redes de colaboración, por lo que también se espera que esa característica se relacione positivamente con la variable resultado (Concytec, 2016, Teodorescu 2000).

El grado académico no es considerado en los estudios revisados pues, al haber sido realizados en países con sistemas de educación superior más desarrollados en comparación con Perú, se asume que todos los investigadores o profesores universitarios cuentan con un doctorado. La situación en Perú es diferente. El gráfico 7.1 muestra la distribución de profesores (considerados para este estudio) según grado académico.

Gráfico 7.1
Distribución de profesores según máximo grado académico reportado



Fuente: II CNU. Elaboración propia.

Se considera la tenencia de doctorado, y no de maestría, como un factor relevante porque es durante la obtención de ese grado cuando el entrenamiento en investigación debería ser más importante, por lo que se esperaría encontrar un efecto positivo en la producción de investigación.

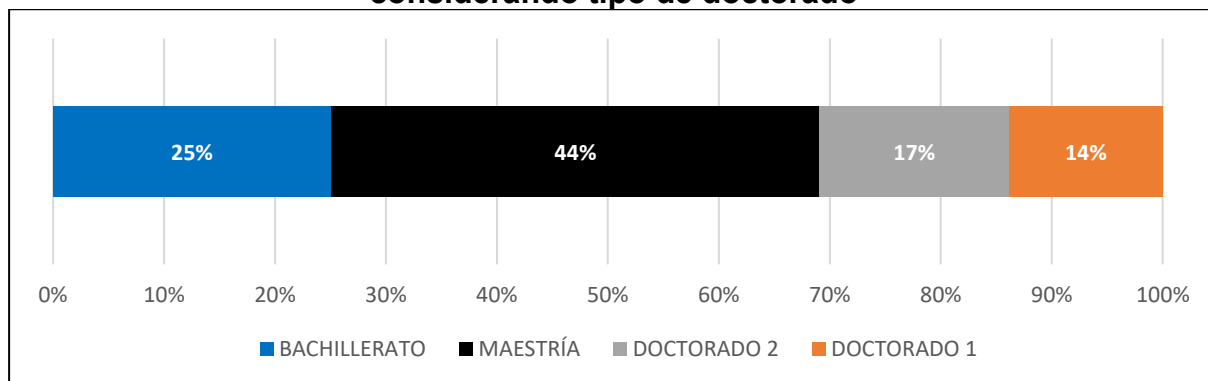
Por otro lado, se decidió tomar en cuenta la universidad en la que el docente obtuvo el doctorado, según esta tenga alta o baja producción académica. La clasificación tomó como base el nivel de investigación del país o la universidad, según los estudios hayan sido llevados fuera o dentro de Perú, que otorgó el título. Los criterios específicos fueron:

- Son de Tipo 1, aquellos doctorados obtenidos en países que hayan sido clasificados mejor que Perú según el *Scimago Journal & Country Ranking*³⁰.
- Son de Tipo 1 aquellos títulos otorgados por alguna de las primeras universidades que concentran el 80% de las publicaciones indizadas en Scopus para el periodo analizado.
- Para todos los otros casos, el doctorado fue clasificado como de Tipo 2.

³⁰ Disponible en <https://www.scimagojr.com/>

El resultado agregado³¹ de esa clasificación puede verse en el gráfico 7.2. Los profesores que manifestaron tener doctorados “tipo 1” representan solo al 14% del total, mientras que aquellos con grados de “tipo 2” llegan a ser al 17%.

Gráfico 7.2
Distribución de profesores según máximo grado académico reportado, considerando tipo de doctorado



Fuente: II CNU. Elaboración propia.

Entonces, en cuanto al grado académico, se tendrán tres tipos de docentes: profesores sin grado de doctor, profesores con doctorado tipo 2, y profesores con doctorado tipo 1, esperándose obtener efectos positivos para los dos últimos tipos, y un mayor efecto para el último.

El efecto de trabajar en otra universidad no es claro. Al estar considerándose únicamente a los profesores ordinarios, y a aquellos contratados con dedicación a tiempo completo en alguna universidad, puede argumentarse que trabajar en otra institución resta tiempo de dedicación en su institución principal (aquella en la que es profesor ordinario o a tiempo completo), por lo que tendría un efecto negativo; sin embargo, puede argumentarse también que el mantener más de una actividad académica refuerza su actividad investigadora al posiblemente tener acceso a recursos y contactos en más de una institución. En Concytec (2016) se considera la segunda posibilidad expuesta, y Teodorescu (2000) encuentra que ese efecto es positivo.

Finalmente, con respecto al tipo de dedicación del profesor, en referencia al tiempo de dedicación a la universidad, se espera que, a más dedicación, se encuentre mayor producción académica.

Características institucionales

³¹ El detalle de la clasificación por países y universidades puede revisarse en el Anexo.

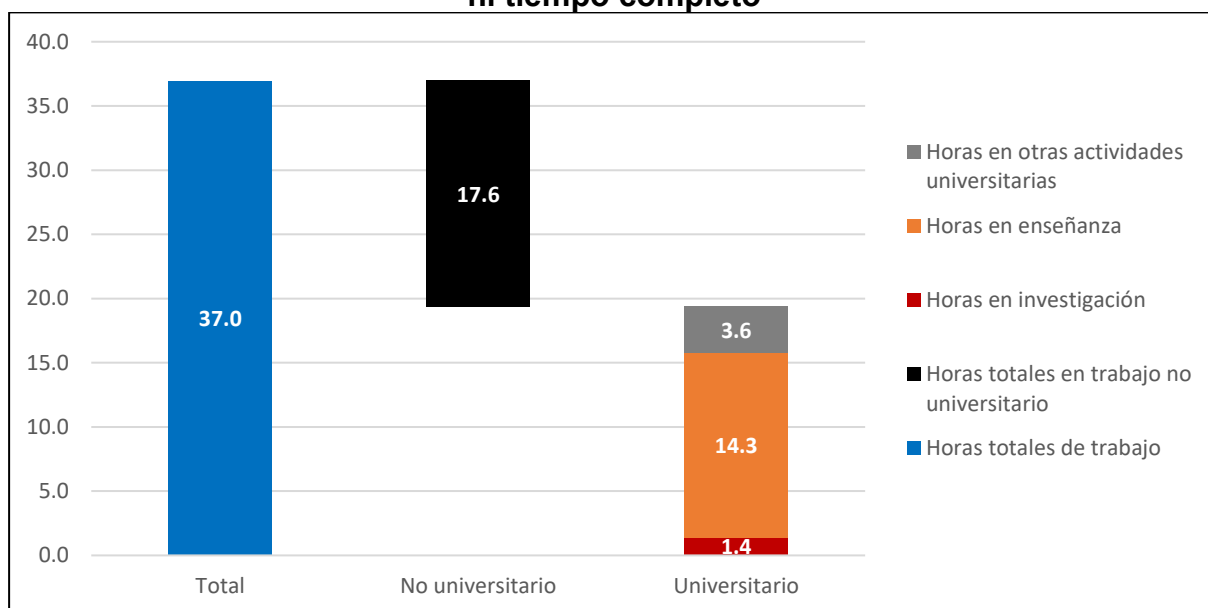
El tercer grupo de variables que serán incluidas en la estimación está conformado por aquellas características que influyen en el resultado, la producción de investigación, pero que son propias del ambiente en el que se desenvuelve el investigador (Teodorescu 2000).

Tanto en Concytec (2016) como en Teodorescu (2000) se menciona que los salarios por la labor en la universidad son un factor positivo para la producción de conocimiento. En ambos casos, ese salario se comportaría como un incentivo al docente. Debe recordarse que, como se mencionó al inicio de este capítulo, los datos de ingresos monetarios en el II CNU están disponible por intervalo, por lo que los análisis para este estudio se realizaron tomando como referencia las marcas de clase. Otra de las características institucionales de la que se tiene información en el II CNU es la condición de desempeñar, o no, un cargo administrativo en la universidad. Nuevamente, si se recurre al marco conceptual de que la cantidad de tiempo de la que dispone cada persona es limitada, la dedicación a actividades administrativas debería tener un efecto negativo en la producción de investigación. Entre la evidencia empírica revisada, Teodorescu (2000) no encontró que la condición de tener un cargo administrativo esté relacionada con la producción de investigación para ninguno de los 10 países que analizó.

En cuanto a la distribución del tiempo del profesor entre horas de trabajo fuera de la universidad, horas realizando investigación, horas enseñando en la universidad, horas realizando trabajos no académicos en la universidad (la suma de estas tres últimas variables viene ser el total de horas de trabajo en la universidad), se espera que en general, el tiempo dedicado a la investigación tenga un efecto positivo en la generación de publicaciones, mientras que el tiempo dedicado a otras actividades se relacione negativamente con esa variable resultado (Ruedas y Rodenas 2015, Teodorescu 2000, Zhang y Ehrenberg 2010). Los gráficos 7.3a y 7.3b muestran el comportamiento de los promedios de horas de dedicación reportados por actividad, según los profesores tengan dedicación exclusiva o a tiempo completo, o no.

Gráfico 7.3a

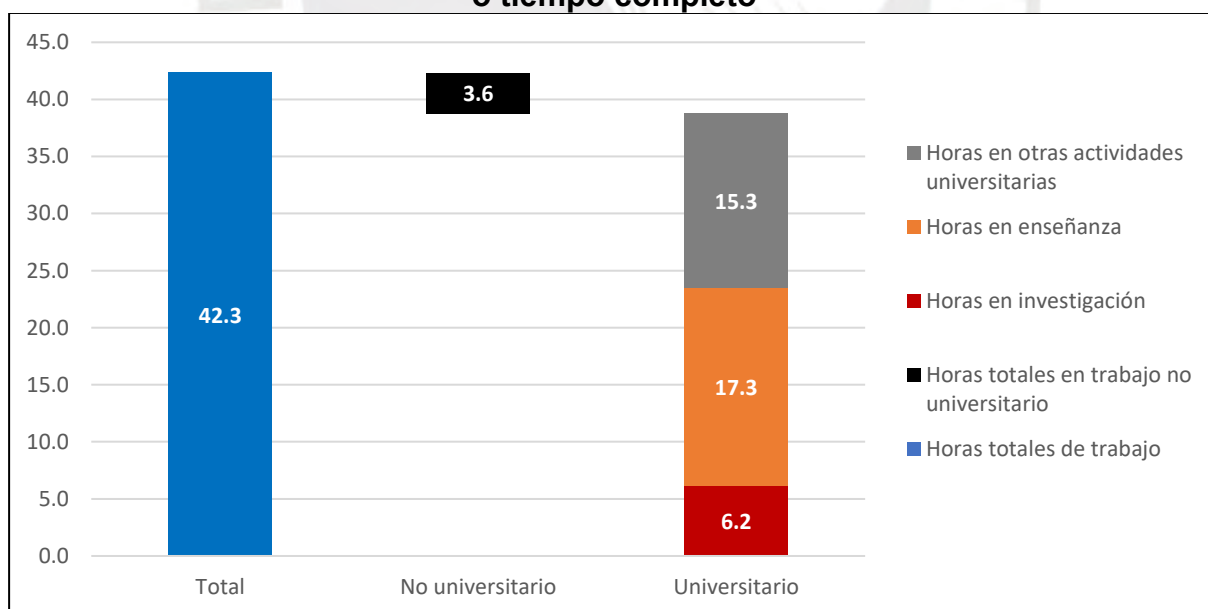
Promedio de horas de trabajo reportadas según este sea en la universidad o fuera de esta, y por el tipo de labor. Para profesores sin dedicación exclusiva ni tiempo completo



Fuente: II CNU. Elaboración propia.

Gráfico 7.3b

Promedio de horas de trabajo reportadas según este sea en la universidad o fuera de esta, y por el tipo de labor. Para profesores con dedicación exclusiva o tiempo completo



Fuente: II CNU. Elaboración propia.

Por otro lado, si bien no existe para Perú una clasificación similar a la realizada por *The Carnegie Classification of Institutions of Higher Education*, para tomar en cuenta de alguna manera el “tipo” de universidad, se utilizará la condición de ser: pública,

privada asociativa (sin fines de lucro) o privada societaria (con fines de lucro). Las universidades pertenecientes al tercer grupo serían las más orientadas a la enseñanza.

De acuerdo con Hardré y otros (2011), Ruedas y Rodenes (2015) y Teodorescu (2000), la opinión que el investigador tiene sobre su institución también es relevante. En este trabajo se considera ese factor al incluir en el análisis las respuestas a las preguntas ¿cree que la investigación de su universidad contribuye a la sociedad? y ¿cuál es su perspectiva sobre su universidad? Se espera que la opinión favorable respecto a la institución, recogida en ambas variables, esté relacionada positivamente con la producción de conocimiento.

En último término, se tienen también tres variables, recogidas del módulo de universidades del II CNU, que serán tomadas como aproximaciones a la institucionalidad y a la agencia de las universidades. Para el primer caso, se toma el tener o no un plan estratégico; mientras que para el segundo, se considera el financiamiento total para investigación que tiene la universidad y si esta tiene convenios de investigación. En los tres casos se espera una relación positiva con la variable resultado.

Variables dependientes

Como se ha propuesto, la variable dependiente dependerá del modelo en cuestión. Para los modelos con variables dicotómicas, la variable dependiente será el haber publicado o no un artículo en una revista o un congreso que esté dentro de una base de datos indizada, según las respuestas registradas en el módulo 3 del II CNU; mientras que, para los modelos de conteo, la variable será el número de publicaciones en Scopus, para el periodo 2011 – 2014, que tiene cada universidad-área del conocimiento.

Es decir, el subíndice i de la función de producción $Y_i = f(T_i, M_i, I_i, S_i, D_i)$ identificará a cada profesor o cada universidad-área del conocimiento según sea el caso.

7.2 Resultados e interpretaciones

Se implementaron todos los modelos descritos en la sección anterior, esto es: regresiones Probit y Logit tomando como unidad de análisis al docente universitario y con todas las características personales e institucionales disponibles; modelos con variable dependiente de conteo (Poisson, Poisson con estimación robusta de la varianza por incumplimiento de equidispersión, Binomial Negativa; tanto los modelos base como las variantes con truncamiento e inflación de ceros) con el mismo set de características anteriores; y finalmente modelos con variable dependiente de conteo

(los antes mencionados), pero tomando diferentes perfiles docentes, y por lo tanto con un set de características reducido.

Resultados de los modelos con variable resultado dicotómica

En el cuadro 7.3 se describe la base de datos utilizada para estos modelos. El promedio se interpreta como el porcentaje de ocurrencia para todas las variables *dummy*; así, es evidente la baja tasa de resultados positivos, “tener una publicación indizada”, para la variable dependiente. Además, por los órdenes de magnitud para las variables “ingreso por la universidad” y “cuadrado de la edad”, se considerará para las regresiones a la primera dividida por mil y a la segunda por cien. Por otro lado, para explorar la posible presencia de multicolinealidad, se calcularon los factores de inflación de varianza (VIF) para todas las variables independientes (ver cuadro 7A en anexo³²).

Cuadro 7.3
Descripción de los datos para los análisis con dependiente binaria

Variable	Observaciones	Promedio	Desv. Est.	Mín.	Máx.
Tiene publicación indizada	23,606	0.1	0.3	0.0	1.0
Sabe inglés	23,606	0.8	0.4	0.0	1.0
Ingreso por la universidad	23,606	2622.4	2407.4	250.0	20000
Número de personas dependientes económicamente	23,606	3.0	1.7	0.0	10.0
Es mujer	23,606	0.3	0.4	0.0	1.0
Edad en años	23,606	51.7	11.3	21.0	90.0
Cuadrado de la edad	23,606	2797.8	1156.9	441.0	8100.0
Es casado / conviviente	23,606	0.7	0.5	0.0	1.0
Pertenece a Lima o Callao	23,606	0.4	0.5	0.0	1.0
Estudió fuera de Perú	23,606	0.0	0.1	0.0	1.0
Tiene Doctorado del tipo 1	23,606	0.1	0.3	0.0	1.0
Tiene Doctorado del tipo 2	23,606	0.2	0.4	0.0	1.0
Desempeña cargo administrativo	23,606	0.4	0.5	0.0	1.0
Dispone de ambiente de trabajo en universidad	23,606	0.8	0.4	0.0	1.0
Trabaja en otra universidad	23,606	0.2	0.4	0.0	1.0
Tiempo completo o dedicación exclusiva	23,606	0.8	0.4	0.0	1.0
Cree que investigación de su universidad contribuye a sociedad	23,606	0.8	0.4	0.0	1.0
Tiene perspectiva positiva sobre su universidad	23,606	0.8	0.4	0.0	1.0
Horas totales de trabajo en universidad	23,606	34.5	12.6	0.0	80.0

³² Todas las variables presentaron VIF menores a 10, salvo edad y su cuadrado. Se esperaba una alta correlación entre ambas variables; y por ser edad al cuadrado una variable de control, no es trascendente para los resultados.

Horas haciendo investigación en universidad	23,606	5.1	5.7	0.0	80.0
Horas enseñando en universidad	23,606	16.6	7.8	0.0	80.0
Horas de trabajo fuera de universidad	23,606	6.7	12.8	0.0	70.0
Pertenece al Área OCDE 1	23,606	0.1	0.3	0.0	1.0
Pertenece al Área OCDE 2	23,606	0.2	0.4	0.0	1.0
Pertenece al Área OCDE 3	23,606	0.2	0.4	0.0	1.0
Pertenece al Área OCDE 4	23,606	0.1	0.3	0.0	1.0
Pertenece al Área OCDE 5	23,606	0.4	0.5	0.0	1.0
Pertenece al Área OCDE 6	23,606	0.1	0.2	0.0	1.0
Universidad privada asociativa	23,606	0.2	0.4	0.0	1.0
Universidad pública	23,606	0.7	0.5	0.0	1.0
Universidad privada societaria	23,606	0.1	0.3	0.0	1.0

Fuente: II CNU, SUNEDU. Elaboración propia.

El cuadro 7.4 muestra los coeficientes y los errores estándar resultantes de los modelos Probit y Logit. Siguiendo a Cameron y Trivedi (2010), entre esos dos modelos estimados con el mismo número de variables es preferible aquel con mayor *log likelihood*, por lo que, en este caso, se escoge el modelo Probit³³. Por otro lado, el ajuste de ambos modelos resultó bajo: el pseudo R^2 para el modelo Probit fue 0.1056 y 0.1038 para el Logit. Esto se debe, en parte, a que la proporción de casos exitosos se encuentra alrededor de 0.1³⁴.

Cuadro 7.4
Coeficientes, significancia y errores estándar resultantes para modelos Probit y Logit

VARIABLES	Probit Publicó	Logit Publicó
Sabe inglés	0.214*** (0.036)	0.451*** (0.077)
Ingreso monetario / 1000	0.011* (0.006)	0.021* (0.012)
Número de personas dependientes económicamente	-0.005 (0.008)	-0.011 (0.017)
Es mujer	0.062** (0.030)	0.118** (0.060)
Edad en años	0.037*** (0.010)	0.079*** (0.020)
Edad al cuadrado / 100	-0.039*** (0.009)	-0.083*** (0.019)
Es casado / conviviente	0.028	0.055

³³ Los valores resultantes de log likelihood fueron -5697 para el modelo Probit, y -5709 para el Logit.

³⁴ En los gráficos 8A del anexo puede verse la distribución de la probabilidad de éxito, notándose una sobreposición considerable.

	(0.029)	(0.057)
Pertenece a Lima o Callao	0.041	0.087
	(0.030)	(0.059)
Estudio fuera de Perú	-0.329***	-0.583**
	(0.126)	(0.247)
Tiene Doctorado del tipo 1	0.512***	0.967***
	(0.037)	(0.070)
Tiene Doctorado del tipo 2	0.225***	0.443***
	(0.032)	(0.064)
Desempeña cargo administrativo	0.104***	0.200***
	(0.028)	(0.055)
Dispone de ambiente de trabajo en universidad	0.066*	0.132*
	(0.035)	(0.071)
Trabaja en otra universidad	-0.011	-0.029
	(0.039)	(0.077)
Tiempo completo o dedicación exclusiva	-0.091*	-0.176*
	(0.047)	(0.094)
Cree que investigación de su universidad contribuye a sociedad	0.118***	0.232***
	(0.036)	(0.072)
Tiene perspectiva positiva sobre su universidad	-0.030	-0.053
	(0.032)	(0.063)
Horas totales de trabajo en universidad	0.002	0.006
	(0.002)	(0.004)
Horas haciendo investigación en universidad	0.035***	0.063***
	(0.003)	(0.005)
Horas enseñando en universidad	0.001	0.002
	(0.002)	(0.004)
Horas de trabajo fuera de universidad	-0.007***	-0.015***
	(0.001)	(0.003)
Pertenece al Área OCDE 1	0.245***	0.486***
	(0.069)	(0.136)
Pertenece al Área OCDE 2	0.010	0.013
	(0.067)	(0.135)
Pertenece al Área OCDE 3	0.415***	0.830***
	(0.067)	(0.134)
Pertenece al Área OCDE 4	0.134*	0.262*
	(0.073)	(0.146)
Pertenece al Área OCDE 5	-0.046	-0.085
	(0.064)	(0.129)
Universidad privada asociativa	0.305***	0.707***
	(0.075)	(0.171)
Universidad pública	0.412***	0.925***
	(0.074)	(0.168)
Financiamiento a investigación (en millones)	0.032***	0.062***
	(0.004)	(0.008)
Constante	-3.495***	-6.847***
	(0.263)	(0.550)

Errores estándares entre paréntesis

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Se obtuvieron resultados inesperados para las variables “estudió fuera del Perú”, “desempeña cargo administrativo” y “tiempo completo o dedicación exclusiva”. En el primer caso, el efecto negativo de esa variable podría explicarse porque más de la mitad de docentes que manifestaron haber estudiado fuera del país no cuentan con un doctorado, que es el grado en el que se está considerando se recibe un entrenamiento riguroso en investigación de calidad, o por alguna posible asociación entre los valores de esa variable con el de la tenencia de doctorados de tipo 1 o 2³⁵. La explicación para las otras dos variables es similar, la relación entre la tenencia de doctorado y el desempeño de algún cargo es positiva, mientras que la relación entre la tenencia de doctorado y el estar a tiempo completo o dedicación exclusiva es negativa³⁶.

Siguiendo con la interpretación de los signos de los coeficientes, los resultados obtenidos para las variables edad y edad al cuadrado indicarían que también en cuanto a “probabilidad de tener publicaciones indizadas” existen rendimientos marginales decrecientes. Por otro lado, las variables *dummy* introducidas para señalar el área del conocimiento al que pertenece el docente, medidas en relación al área 6 (Humanidades), muestran que solo el pertenecer a las áreas 1 (Ciencias Naturales), 3 (Medicina) o 4 (Ciencias Agrarias) tienen efectos diferenciados (positivos), mientras que el pertenecer al área 2 (Ingeniería y Tecnología) o al área 5 (Ciencias Sociales) no generaría ninguna diferencia. Además, se comprueba que poseer un doctorado incrementa la probabilidad de contar con publicaciones indizadas. Además, pertenecer a una universidad públicas o a una privadas asociativa, también incrementa la probabilidad de contar con publicaciones indicadas.

En cuanto a la dedicación de tiempo, como era de esperarse, el número de horas dedicadas a la investigación afecta positivamente a la probabilidad de la variable dependiente, mientras que el número de horas dedicadas a actividades no universitarias lo hace negativamente. También se obtuvieron relaciones positivas con la variable resultado para las variables “saber inglés” y “ser mujer”; en el primer caso, puede deberse a que ese idioma es en la actualidad de uso estándar en la academia, mientras que para el resultado de la segunda variable, es posible que, como solo se consideraron docentes ordinarios o contratados a tiempo completo, las mujeres en ese universo hayan tenido que pasar por pruebas más rigurosas que sus pares

³⁵ El estadístico Tau b de Kendall para el cruce de “haber estudiado en el extranjero” y “tener un doctorado tipo 1” es positivo y significativo, mientras que para el cruce de “haber estudiado en el extranjero” y “tener un doctorado tipo 2” es negativo y significativo. Para ambos casos los valores de Tau son menores a 0.1.

³⁶ Relaciones también calculadas con el estadístico Tau b de Kendall. En ambos casos resultados significativos y pequeños.

hombres. Al respecto, Alcázar, Balarín y Román (2018) usan datos de una fuente propia y muestran que, en las Ciencias Sociales en Perú, la diferencia en la producción académica está en la cantidad de publicaciones que los profesores señalaron tener.

Sobre la importancia de las motivaciones, los resultados muestran que la opinión del profesor sobre si la investigación de su universidad contribuye a la sociedad tiene una relación positiva con la probabilidad de contar con publicaciones en Scopus, lo cual puede tomarse como “motivación intrínseca”; mientras que los ingresos monetarios, que vendrían a ser “motivación extrínseca”, también muestran relación positiva, aunque menos significativa, con la variable dependiente. La variable “dispone de ambiente de trabajo”, que sería una manera de aproximarse a la “comodidad” física de los docentes en las universidades, también resultó positiva y estadísticamente significativa.

Como se sabe, los coeficientes para los análisis de regresión Probit y Logit no se interpretan de forma similar a los obtenidos con modelos como el de Mínimos Cuadrados Ordinarios, es decir, no representan el valor de los efectos marginales para las variables. Debe recurrirse a estimaciones adicionales para calcular esos efectos, las dos opciones disponibles son: estimar el *average marginal effect* (AME) o el *marginal effect at mean* (MEM). No sería adecuado interpretar el MEM para este estudio puesto que se tiene una cantidad elevada de variables explicativas dicotómicas, lo que causaría que los valores promedio sobre los que se calcula ese efecto no sean parte representativa del universo analizado (Cameron y Trivedi, 2010). Por lo tanto, en el cuadro 7.5 se muestran los AME para el modelo Probit.

Cuadro 7.5
Promedio de efectos marginales para el modelo Probit

VARIABLES	dy/dx	Std. Err.
Sabe inglés	0.0271	0.0048
Ingreso monetario / 1000	0.0021	0.0008
Número de personas dependientes económicamente	-0.0003	0.0011
Es mujer	0.0091	0.0040
Edad en años	0.0041	0.0013
Edad al cuadrado / 100	-0.0046	0.0013
Es casado / conviviente	0.0042	0.0038
Pertenece a Lima o Callao	0.0035	0.0040
Estudio fuera de Perú	-0.0528	0.0182
Tiene Doctorado del tipo 1	0.0643	0.0050
Tiene Doctorado del tipo 2	0.0295	0.0043
Desempeña cargo administrativo	0.0137	0.0037
Dispone de ambiente de trabajo en universidad	0.0090	0.0046
Trabaja en otra universidad	-0.0009	0.0051
Tiempo completo o dedicación exclusiva	-0.0030	0.0063
Cree que investigación de su universidad contribuye a sociedad	0.0149	0.0047

Tiene perspectiva positiva sobre su universidad	-0.0037	0.0042
Horas totales de trabajo en universidad	0.0003	0.0002
Horas haciendo investigación en universidad	0.0045	0.0003
Horas enseñando en universidad	0.0002	0.0003
Horas de trabajo fuera de universidad	-0.0009	0.0002
Pertenece al Área OCDE 1	0.0289	0.0091
Pertenece al Área OCDE 2	0.0006	0.0089
Pertenece al Área OCDE 3	0.0534	0.0089
Pertenece al Área OCDE 4	0.0198	0.0097
Pertenece al Área OCDE 5	-0.0031	0.0085
Universidad privada asociativa	0.0344	0.0100
Universidad pública	0.0540	0.0097
Financiamiento a investigación (en millones)	0.0042	0.0005

Fuente: Elaboración propia.

En cuanto a las magnitudes de los efectos, la variable más relevante es la “tenencia de doctorado”, pero con diferencias importantes según la “calidad” de este. Contar con un doctorado obtenido en una universidad con un buen nivel de investigación incrementa la probabilidad de tener al menos una publicación indizada en 6.43%, lo cual representa más del doble que el incremento si el doctorado fue obtenido en una universidad con un bajo nivel de investigación (2.95%). En segundo lugar en tamaño, vienen los efectos por pertenecer a áreas del conocimiento diferentes a Humanidades; así, los incrementos en la probabilidad de ocurrencia de la variable respuesta están entre 1.98% y 5.34%, el efecto mayor es para los que pertenecen al área médica, mientras que el menor para los de las ciencias agrarias. A continuación, se encuentra el tipo de universidad al que pertenece el docente: con respecto a pertenecer a una universidad privada societaria, el pertenecer a una universidad pública incrementa la probabilidad en 5.40%, mientras que el pertenecer a una universidad privada asociativa, en 3.44%.

El saber inglés, por su parte, incrementa la probabilidad de interés en 2.71%, mientras que tener una idea positiva sobre la contribución de la investigación de la universidad en la sociedad la incrementa en 1.49% y contar con un cargo administrativo en 1.37%. Los demás efectos positivos no llegan al 1% de incremento en la probabilidad estudiada. Los resultados para las variables no dicotómicas (horas, ingresos e inversión en miles) tienen valores menores, pero debe considerarse que eso también depende del orden de magnitud de los valores de la variable.

Resultados de los modelos con variables de conteo sin diferencia de perfiles.

Las características de la base de datos construida para los análisis de esta sección pueden ser revisadas en el cuadro 7.6. También en este caso, para evaluar multicolinealidad, se calcularon los VIF para todas las variables independientes, con resultados similares a los de la sección anterior (ver cuadro 9A en el anexo).

Cuadro 7.6**Descripción de los datos para los análisis con variable dependiente de conteo**

Variable	Observaciones	Promedio	Desv. Est.	Mín.	Máx.
Número de publicaciones en Scopus	381	11.3	55.6	0	767
Sabe inglés	381	0.8	0.2	0	1
Ingreso monetario por la universidad	381	2546.9	1578.3	250	12500
Número de personas dependientes económicamente	381	3.0	1.0	0	9
Es mujer	381	0.3	0.2	0	1
Edad en años	381	48.8	7.2	27	77
Cuadrado de la edad	381	2438.4	709.0	729	5929
Es casado / conviviente	381	0.6	0.2	0	1
Estudio fuera de Perú	381	0.0	0.1	0	1
Tiene Doctorado del tipo 1	381	0.1	0.2	0	1
Tiene Doctorado del tipo 2	381	0.2	0.2	0	1
Desempeña cargo administrativo	381	0.5	0.2	0	1
Dispone de ambiente de trabajo en universidad	381	0.8	0.2	0	1
Trabaja en otra universidad	381	0.2	0.2	0	1
Tiempo completo o dedicación exclusiva	381	0.8	0.3	0	1
Cree que investigación de su universidad contribuye a sociedad	381	0.9	0.1	0	1
Tiene perspectiva positiva sobre su universidad	381	0.8	0.2	0	1
Horas totales de trabajo en universidad	381	33.8	9.1	2	68
Horas haciendo investigación en universidad	381	4.3	3.4	0	30
Horas enseñando en universidad	381	16.0	4.7	0	38
Horas de trabajo fuera de universidad	381	5.9	6.2	0	40
Número de profesores	381	61.0	87.2	1	682
Pertenece al Área OCDE 1	381	0.2	0.4	0	1
Pertenece al Área OCDE 2	381	0.2	0.4	0	1
Pertenece al Área OCDE 3	381	0.2	0.4	0	1
Pertenece al Área OCDE 4	381	0.1	0.3	0	1
Pertenece al Área OCDE 5	381	0.2	0.4	0	1
Pertenece al Área OCDE 6	381	0.1	0.3	0	1
Lima/Callao vs otras regiones	381	0.4	0.5	0	1
Universidad tiene plan estratégico	381	0.9	0.3	0	1
Financiamiento total (en miles)	381	1639.2	2851.7	0	13000
Universidad tiene convenios de investigación	381	0.6	0.5	0	1
Universidad privada asociativa	381	0.3	0.5	0	1
Universidad pública	381	0.4	0.5	0	1
Universidad privada societaria	381	0.2	0.4	0	1

Fuente: II CNU, SUNEDU, Scopus. Elaboración propia.

En la sección de lineamientos metodológicos se describieron numerosas combinaciones posibles para modelos con variables de conteo. Para elegir el modelo más adecuado en este contexto, se utilizan los siguientes criterios:

- Prueba de dispersión, para elegir entre distribución Poisson, Poisson con cálculo robusto de varianza o Binomial Negativa: De acuerdo con lo señalado en la sección de opciones de modelos econométricos, el uso de la distribución está limitado para datos que muestren “equidispersión”, por lo que de detectarse que se incumple ese criterio, debe preferirse la distribución Poisson con cálculo robusto de varianza o la distribución Binomial Negativa.
- Análisis del comportamiento histórico, para elegir entre el modelo base, el *zero truncated* o el *zero inflated*: El objetivo es detectar si los “ceros” presentes en el conteo de las publicaciones pertenecen, total o parcialmente, o no al mismo proceso generador de datos que los conteos positivos. De pertenecer completamente, se optaría por el modelo base; en caso solo parte de los “ceros” sigan el mismo proceso que los valores positivos, el modelo a tomar sería el *zero inflated*; mientras que si no se cumple ninguna de las dos condiciones anteriores, se optará por el modelo *zero truncated* pues el proceso generador de los “ceros” sería diferente al que genera las observaciones positivas.

Dispersión: El resultado de la prueba de sobredispersión arrojó parámetros significativamente distintos de cero³⁷, esto indica la ausencia de “equidispersión” en los datos, por lo que se elige la distribución Binomial Negativa.

Comportamiento histórico: Se realizaron dos análisis; en primer lugar, se calculó la correlación entre el número de publicaciones por universidad para el periodo 2001 – 2010 (periodo 1, toda la década anterior al censo) y el periodo 2011 – 2014 (periodo 2, años posteriores al censo y previos a la promulgación de la Nueva Ley Universitaria). En segundo lugar, se calcularon las probabilidades condicionales de contar con publicaciones en el periodo 2 dado que se tuvo o no se tuvo publicaciones en el periodo 1³⁸. Como resultado, se tiene que la correlación entre el número de publicaciones en el periodo 1 y el periodo 2 es 0.977; por otro lado, la probabilidad de tener publicaciones en el periodo 2 dado que se tuvo publicaciones en el periodo 1 es 0.979; y la probabilidad de tener publicaciones en el periodo 2 dado que no se tuvo publicaciones en el periodo 1 es 0,021. Por lo tanto, se considera para este trabajo que todos los “ceros” en la variable de conteo son generados por un proceso diferente a los valores enteros positivos, esto puede interpretarse como “aquellas áreas-universidad que figuran con cero publicaciones en el periodo de análisis pueden considerarse como unidades que no producen ese tipo productos de investigación”.

³⁷ En todos los casos, “alpha” fue diferente de cero con una significancia inferior al 1%.

³⁸ Las probabilidades condicionales fueron calculadas usando la siguiente relación: $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Entonces, entre el modelo base, el *zero inflated* o el *zero truncated*, se elige a este último como el más adecuado.

Se implementaron todos los modelos de conteo descritos en la sección de lineamientos metodológicos. Debe recordarse que, para estos modelos, la unidad de análisis no es el docente, sino el área-universidad, por lo que el número de observaciones se reduce significativamente. Además, como se mencionó en la sección 6, la base de datos utilizada es el resultado de unir los datos del II CNU con los datos generados utilizando técnicas de *web scraping* para el portal de Scopus. Los valores de los coeficientes, sus niveles de significancia y los errores estándares para seis³⁹ de los modelos implementados pueden verse en el cuadro 7.7; mientras que en el cuadro 7.8 se encuentran los valores de los *log likelihood* y pseudo R^2 resultantes para los mismos modelos.

Cuadro 7.7
Coeficientes, significancia y errores estándar resultantes para modelos de conteo

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Publicaciones	Publicaciones	Publicaciones	Publicaciones	Publicaciones	Publicaciones
Sabe inglés	4.065*** (0.287)	4.065*** (1.264)	1.802** (0.814)	3.452*** (0.338)	3.452** (1.448)	2.153 (1.398)
Ingreso monetario / 1000	0.263*** (0.020)	0.263*** (0.086)	0.531*** (0.119)	0.062*** (0.024)	0.062 (0.109)	0.163 (0.157)
Número de personas dependientes económicamente	-0.566*** (0.058)	-0.566** (0.249)	-0.562** (0.221)	-0.972*** (0.079)	-0.972** (0.416)	-1.173** (0.482)
Es mujer	-0.644*** (0.170)	-0.644 (0.718)	0.564 (0.752)	-0.130 (0.218)	-0.130 (0.932)	0.251 (1.379)
Edad en años	0.742*** (0.080)	0.742* (0.404)	0.041 (0.214)	0.088 (0.087)	0.088 (0.444)	-0.002 (0.410)
Cuadrado de la edad	-0.007*** (0.001)	-0.007* (0.004)	-0.001 (0.002)	-0.000 (0.001)	-0.000 (0.004)	0.001 (0.004)
Es casado / conviviente	-2.075*** (0.247)	-2.075* (1.130)	-0.117 (0.886)	-0.104 (0.327)	-0.104 (1.461)	-0.853 (1.895)

³⁹ No se muestran los resultados con las regresiones *zero truncated*, por haberse considerado como más adecuadas a las *zero truncated*. Los resultados de los modelos base se muestran para fines comparativos.

Estudio fuera de Perú	2.012***	2.012	-0.517	-0.272	-0.272	0.322
	(0.475)	(1.890)	(2.028)	(0.907)	(4.370)	(4.904)
Tiene Doctorado del tipo 1	-0.203	-0.203	2.137*	-0.185	-0.185	0.688
	(0.197)	(0.817)	(1.154)	(0.255)	(0.920)	(1.773)
Tiene Doctorado del tipo 2	-4.446***	-4.446***	-1.326*	-6.448***	-6.448***	-1.258
	(0.310)	(1.572)	(0.772)	(0.381)	(1.872)	(1.241)
Desempeña cargo administrativo	0.635***	0.635	-1.306	-0.954***	-0.954	-1.615
	(0.200)	(0.726)	(0.811)	(0.240)	(0.907)	(1.283)
Dispone de ambiente de trabajo en universidad	-1.426***	-1.426*	0.444	-1.672***	-1.672	0.839
	(0.232)	(0.754)	(0.899)	(0.287)	(1.051)	(1.402)
Trabaja en otra universidad	-1.266***	-1.266**	-0.818	-1.106***	-1.106	-0.865
	(0.189)	(0.645)	(0.888)	(0.299)	(1.046)	(1.856)
Tiempo completo o dedicación exclusiva	-0.566***	-0.566	-0.081	-0.512**	-0.512	0.725
	(0.164)	(0.690)	(0.798)	(0.223)	(0.927)	(1.151)
Cree que investigación de su universidad contribuye a sociedad	1.958***	1.958*	-1.668	3.939***	3.939***	1.825
	(0.319)	(1.017)	(1.424)	(0.372)	(1.342)	(2.686)
Tiene perspectiva positiva sobre su universidad	0.515**	0.515	2.424*	-0.707**	-0.707	-0.591
	(0.246)	(0.809)	(1.293)	(0.291)	(0.927)	(2.168)
Horas totales de trabajo en universidad	-0.027**	-0.027	-0.002	0.067***	0.067	0.007
	(0.011)	(0.045)	(0.032)	(0.014)	(0.071)	(0.048)
Horas haciendo investigación en universidad	0.023**	0.023	0.139**	-0.010	-0.010	0.226**
	(0.009)	(0.038)	(0.057)	(0.011)	(0.052)	(0.088)
Horas enseñando en universidad	-0.080***	-0.080*	-0.030	-0.196***	-0.196***	-0.099*
	(0.011)	(0.041)	(0.037)	(0.015)	(0.060)	(0.055)
Horas de trabajo fuera de universidad	0.011	0.011	0.057*	0.067***	0.067**	0.087*
	(0.007)	(0.029)	(0.032)	(0.009)	(0.034)	(0.047)
Pertenece al Área OCDE 1	3.158***	3.158***	2.714***	3.303***	3.303***	3.522***
	(0.166)	(0.648)	(0.521)	(0.168)	(0.511)	(0.739)
Pertenece al Área OCDE 2	1.988***	1.988**	1.976***	2.128***	2.128***	2.591***
	(0.173)	(0.781)	(0.535)	(0.173)	(0.586)	(0.769)
Pertenece al Área OCDE 3	4.251***	4.251***	3.296***	3.671***	3.671***	3.661***
	(0.169)	(0.766)	(0.569)	(0.178)	(0.604)	(0.947)

Pertenece al Área OCDE 4	3.513***	3.513***	3.551***	3.090***	3.090***	3.869***
	(0.173)	(0.653)	(0.578)	(0.180)	(0.562)	(0.835)
Pertenece al Área OCDE 5	2.169***	2.169***	1.527***	1.671***	1.671***	1.970***
	(0.159)	(0.656)	(0.497)	(0.153)	(0.400)	(0.735)
Lima/Callao vs otras regiones	1.191***	1.191***	0.436	0.235**	0.235	-0.117
	(0.080)	(0.328)	(0.339)	(0.110)	(0.520)	(0.542)
Universidad tiene plan estratégico	1.182***	1.182	2.158***	0.475	0.475	1.608
	(0.390)	(0.961)	(0.626)	(0.438)	(0.733)	(1.276)
Financiamiento a investigación (en millones)	0.201***	0.201***	0.195***	0.122***	0.122***	0.117**
	(0.007)	(0.031)	(0.038)	(0.008)	(0.033)	(0.056)
Universidad tiene convenios de investigación	0.908***	0.908**	0.342	0.018	0.018	0.180
	(0.071)	(0.363)	(0.264)	(0.085)	(0.495)	(0.375)
Universidad privada asociativa	0.248**	0.248	0.240	0.855***	0.855	
	(0.098)	(0.393)	(0.400)	(0.123)	(0.588)	
Universidad pública	2.319***	2.319***	2.293***	1.319***	1.319*	
	(0.136)	(0.569)	(0.527)	(0.175)	(0.799)	
Inalpha			0.856***			0.920***
			(0.113)			(0.306)
Constante	-23.811***	-23.811**	-7.294	-5.843***	-5.843	-5.260
	(2.020)	(10.563)	(5.132)	(2.215)	(11.365)	(8.973)
Observaciones	381	381	381	165	165	165

Errores estándares entre paréntesis.

*** p<0.01, ** p<0.05, * p<0.1

Cuadro 7.8
Valores de *log likelihood* y pseudo R^2 obtenidos para modelos de conteo

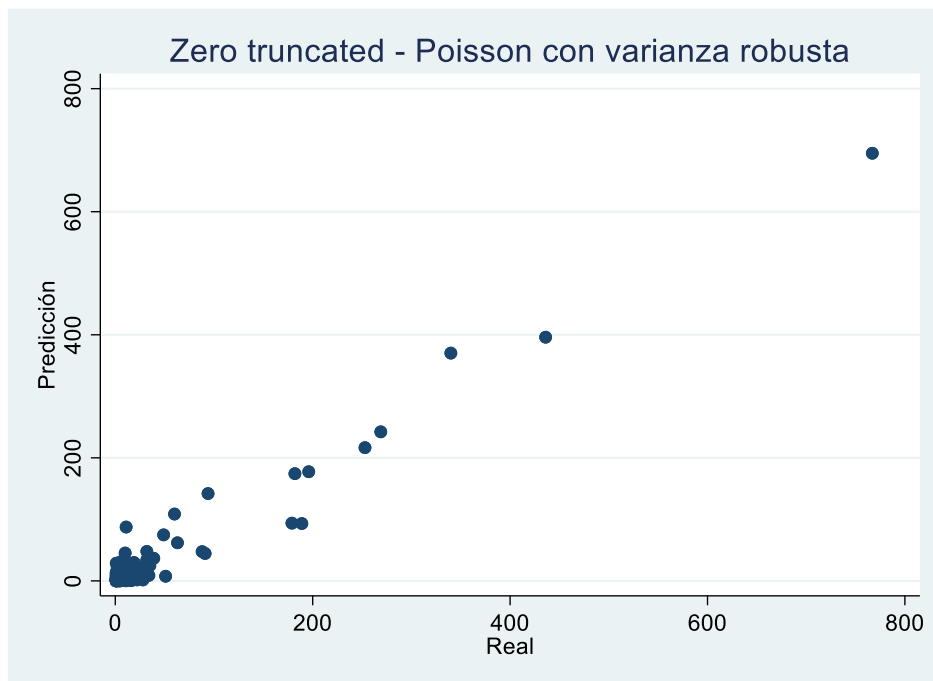
Número	Modelo	Distribución	Log likelihood	Pseudo R2
1	Base	Poisson	-1856.53	0.83
2	Base	Poisson con estimación robusta de varianza	-1856.53	0.83
3	Base	Binomial negativa	-685.08	0.18
4	Zero truncated	Poisson	-1231.23	0.83
5	Zero truncated	Poisson con estimación robusta de varianza	-1231.23	0.83
6	Zero truncated	Binomial negativa	-497.73	0.12

Fuente: Elaboración propia.

De acuerdo con lo considerado, entonces, los modelos más adecuados al contexto en estudio son el 5 y el 6. De estos, el que mejor ajuste con los datos tiene, utilizando como criterio la comparación de pseudo R^2 es el número 5.

Gráfico 7.4a

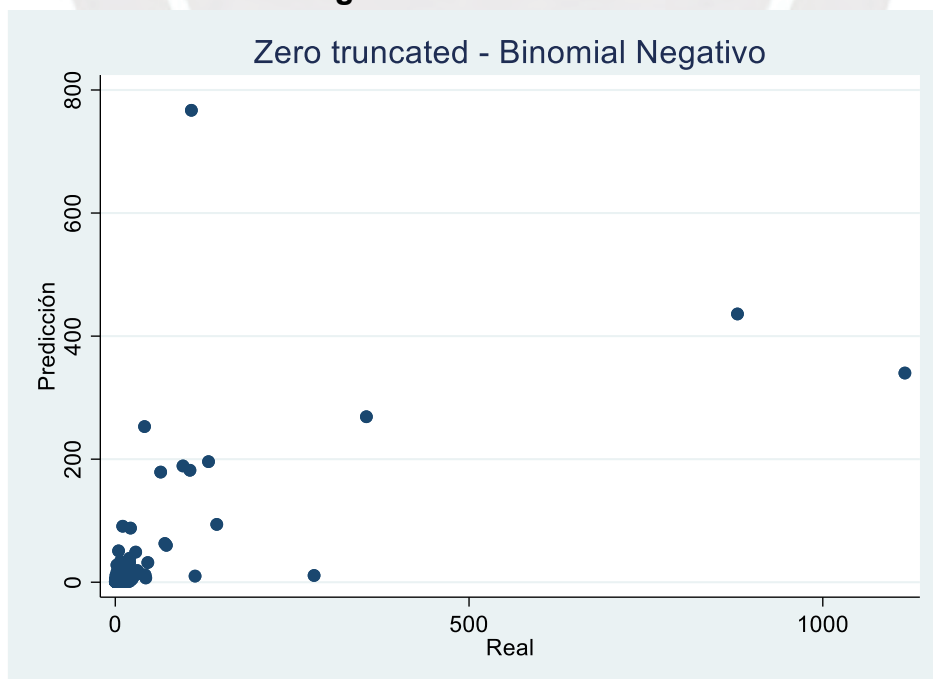
Dispersión entre conteo real y conteo predicho con el modelo Poisson con varianza robusta *zero truncated*



Fuente: Elaboración propia.

Gráfico 7.4b

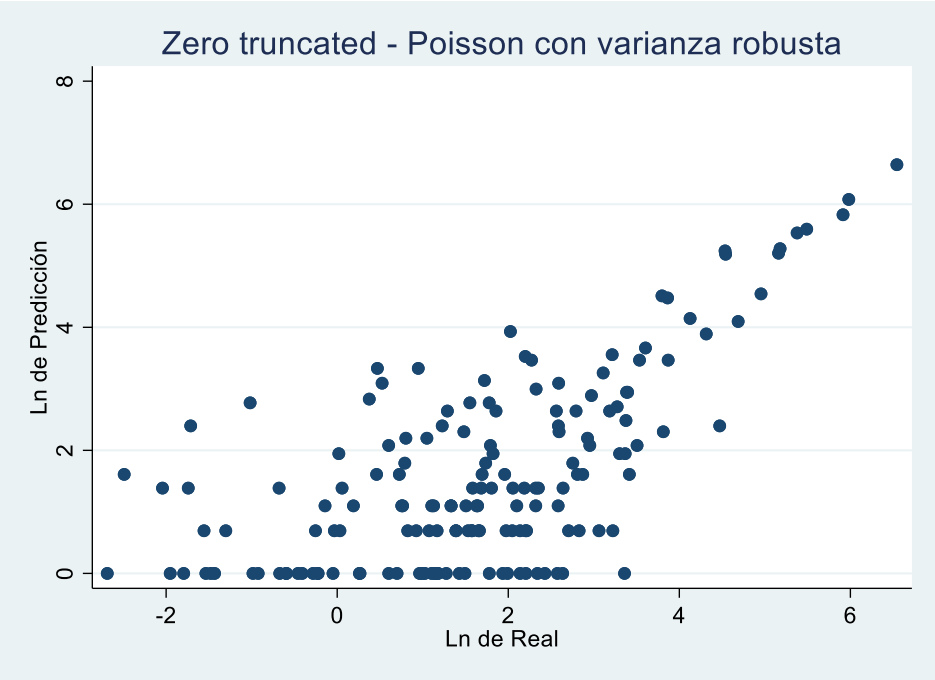
Dispersión entre conteo real y conteo predicho con el modelo Binomial Negativo *zero truncated*



Fuente: Elaboración propia.

Esto puede complementarse observando el comportamiento de los datos predichos; en los gráficos 7.4a y 7.4b puede verse su dispersión con respecto a los datos reales y resulta evidente que el modelo predice mejor para los valores “grandes”. En los gráficos 7.5a y 7.5b se hace uso del logaritmo natural de las mismas variables para observar mejor el comportamiento en todo el rango de valores.

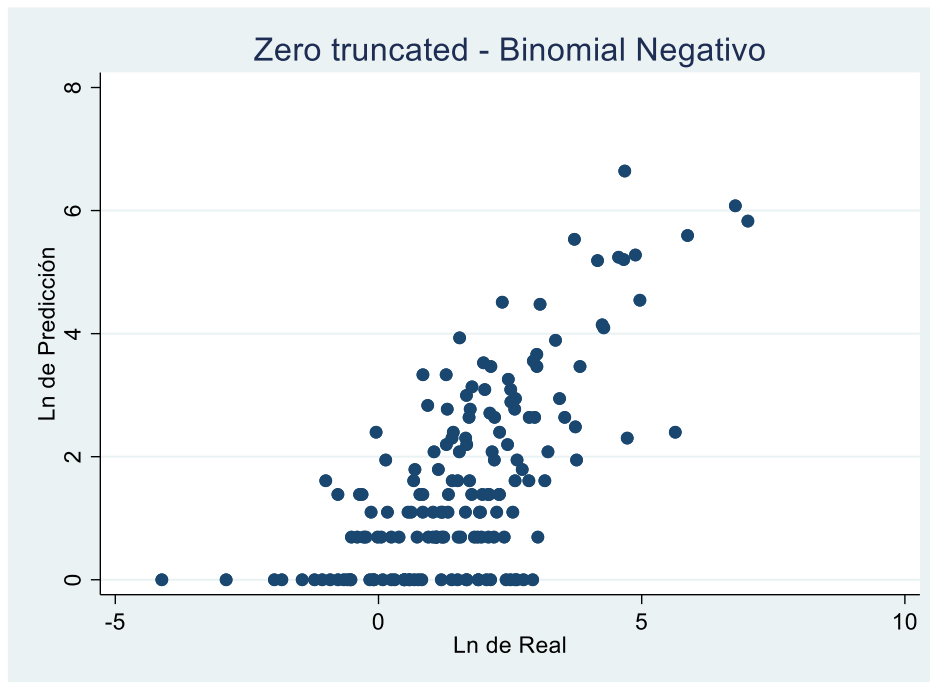
Gráfico 7.5a
Dispersión entre logaritmo del conteo real y del conteo predicho con el modelo Poisson con varianza robusta *zero truncated*



Fuente: Elaboración propia.



Gráfico 7.5b
Dispersión entre logaritmo del conteo real y del conteo predicho con el Binomial Negativo *zero truncated*



Fuente: Elaboración propia.

Considerando entonces al modelo de distribución Poisson con estimación de varianza robusta – *Zero Truncated*, las variables estadísticamente significativas que se relacionan positivamente con número de publicaciones de un área-universidad en Perú son: el saber inglés, creer que la investigación que se desarrolla en su universidad, el área del conocimiento (pertenecer a cualquier área, con respecto a Humanidades), el financiamiento a la investigación de la universidad, y si la universidad es pública; mientras que se encontró una relación negativa para: el número de personas dependientes, la tenencia de un doctorado de una universidad con bajo nivel de investigación, el número de horas de enseñanza, y el número de horas de trabajo fuera de la universidad.

Si bien, en general, esto refuerza lo encontrado al implementar las regresiones Probit y Logit, debe mantenerse en perspectiva que estos últimos resultados corresponden a una medida de la “intensidad de publicación”, solo entre aquellas áreas-universidad que sí han tenido al menos una publicación indizada en Scopus en el periodo de análisis, pues se está truncando todos los ceros de la base de datos. Debe considerarse también que los valores para las variables provenientes de las *dummy* (como “ser mujer” o “trabajar en otra universidad”) se convierten, en realidad, en la proporción o porcentaje de docentes que cumple con esa característica dentro del total de profesores de su área-universidad.

En cuanto a las magnitudes de los efectos, la interpretación para las variables provenientes de *dummy* (como las mencionadas en el párrafo anterior) es: “el incremento de la proporción de docentes que cumplen con esa característica en el área-universidad en 1%, produce una variación porcentual en el promedio del número de publicaciones indizadas igual al valor del coeficiente de esa variable”⁴⁰. Así, la mayor relación positiva con el número de publicaciones indizadas viene del porcentaje de profesores que creen que la investigación de su universidad tiene un efecto positivo en la sociedad, por cada 1% adicional de docentes con esa idea, 3.94%. La segunda característica más relevante vendría ser el saber inglés, pues cada incremento en 1% en la proporción de profesores que conocen ese idioma produciría un incremento en 3.45% en el promedio del número de publicaciones indizadas.

Fuera de las variables que miden proporción de profesores, puede notarse que por cada hora adicional promedio dedicada a la enseñanza, se produciría un decremento del promedio de publicaciones de casi 20%, mientras que por cada millón de soles adicionales en financiamiento se obtendría un incremento de 12.2%.

Por otro lado, pertenecer a cualquier área diferente a Humanidades tiene un efecto positivo respecto a esa; ese efecto se encuentra entre 1.67% para Ciencias Sociales y 3.67% para Medicina. Otra característica contextual relevante es el tipo de universidad en el que se encuentra el profesor: pertenecer a una universidad pública generaría un incremento en el promedio de publicaciones del área-universidad en 1.32%, con respecto a pertenecer a una universidad privada societaria (con fines de lucro).

Resultados de los modelos con variables de conteo considerando diferentes perfiles

Para profundizar en lo ya encontrado en las secciones anteriores con respecto a dos de los elementos de la función de producción descrita en la hipótesis, se decidió implementar un grupo adicional de modelos de conteo (Poisson, Poisson con estimación robusta de la varianza, y Binomial Negativa; tanto en sus versiones base como las *zero truncated* y *zero inflated*), considerando distintos “tipos” de profesores, según su entrenamiento para investigar⁴¹. Los tres “tipos” de profesores considerados son:

- Profesor 1: No tiene doctorado
- Profesor 2: Obtuvo doctorado en universidad poco productiva⁴²

⁴⁰ Puede demostrarse a partir de la relación $E(Y_i/X_i) = \exp(X'_i\beta)$, considerando que X'_i son porcentajes.

⁴¹ También en este caso, los modelos más adecuados son los *zero truncated* que consideren sobredispersión.

⁴² La clasificación entre universidades poco productivas y productivas es la misma que se utilizó para la clasificación de doctorados “tipo 1” y “tipo 2”.

- Profesor 3: Obtuvo doctorado en universidad productiva

La descripción de la base de datos para estos análisis se muestra en el cuadro 7.9, y los resultados para seis de los nueve modelos implementados se encuentran en los cuadros 7.10 (coeficientes, significancia y errores estándar) y 7.11 (valores de los *log likelihood* y pseudo R^2). Los gráficos 7.6 y 7.7 sirven como un apoyo adicional para tener una idea de la bondad de ajuste⁴³.

Cuadro 7.9
Descripción de los datos para los análisis con variable dependiente de conteo y diferencias en los tipos de profesores

Variable	Observaciones	Promedio	Desv. Est.	Mín.	Máx.
Número de publicaciones en Scopus	381	11.3	55.6	0	767
Número de profesores 1	381	42.2	58.4	0	480
Número de profesores 2	381	10.6	19.4	0	156
Número de profesores 3	381	8.2	22.1	0	215
Pertenece al Área OCDE 1	381	0.2	0.4	0	1
Pertenece al Área OCDE 2	381	0.2	0.4	0	1
Pertenece al Área OCDE 3	381	0.2	0.4	0	1
Pertenece al Área OCDE 4	381	0.1	0.3	0	1
Pertenece al Área OCDE 5	381	0.2	0.4	0	1
Pertenece al Área OCDE 6	381	0.1	0.3	0	1
Lima/Callao vs otras regiones	381	0.4	0.5	0	1
Universidad tiene plan estratégico	381	0.9	0.3	0	1
Financiamiento total (en miles)	381	1639.2	2851.7	0	13000
Universidad tiene convenios de investigación	381	0.6	0.5	0	1
Universidad privada asociativa	381	0.3	0.5	0	1
Universidad pública	381	0.4	0.5	0	1
Universidad privada societaria	381	0.2	0.4	0	1

Fuente: Elaboración propia.

Cuadro 7.10
Coefficientes, significancia y errores estándar resultantes para modelos de conteo con diferencia en los perfiles

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Publicacion es	Publicacion es	Publicacion es	Publicacion es	Publicacion es	Publicacion es
Número de profesores 1	0.003*** (0.000)	0.003 (0.003)	0.003 (0.003)	0.003*** (0.000)	0.003 (0.002)	0.001 (0.004)
Número de profesores 2	-0.025*** (0.002)	-0.025* (0.014)	-0.003 (0.006)	-0.027*** (0.002)	-0.027** (0.012)	-0.004 (0.008)
Número de profesores 3	0.010*** (0.001)	0.010* (0.005)	0.030*** (0.009)	0.009*** (0.001)	0.009* (0.005)	0.036** (0.016)

⁴³ Se construyeron gráficos solo para el modelo *zero truncated* con distribución Poisson y cálculo de varianza robusta, por ser ese el modelo que considerado como el más adecuado.

Pertenece al Área OCDE 1	2.439*** (0.127)	2.439*** (0.440)	2.359*** (0.454)	2.423*** (0.134)	2.423*** (0.426)	3.659*** (0.756)
Pertenece al Área OCDE 2	1.735*** (0.135)	1.735*** (0.568)	1.460*** (0.463)	1.714*** (0.143)	1.714*** (0.560)	2.495*** (0.820)
Pertenece al Área OCDE 3	2.476*** (0.133)	2.476*** (0.524)	2.621*** (0.469)	2.591*** (0.141)	2.591*** (0.504)	3.582*** (0.765)
Pertenece al Área OCDE 4	3.073*** (0.128)	3.073*** (0.484)	3.134*** (0.501)	2.980*** (0.137)	2.980*** (0.489)	4.269*** (0.786)
Pertenece al Área OCDE 5	1.166*** (0.138)	1.166** (0.503)	0.658 (0.458)	1.199*** (0.145)	1.199** (0.479)	1.803** (0.753)
Lima/Callao vs otras regiones	1.788*** (0.052)	1.788*** (0.378)	1.195*** (0.271)	1.628*** (0.057)	1.628*** (0.501)	1.251** (0.559)
Universidad tiene plan estratégico	2.560*** (0.380)	2.560*** (0.903)	2.289*** (0.581)	0.929** (0.406)	0.929 (0.716)	1.494 (1.670)
Financiamiento a investigación (millones)	0.181*** (0.004)	0.181*** (0.026)	0.111*** (0.038)	0.136*** (0.004)	0.136*** (0.030)	0.039 (0.069)
Universidad tiene convenios de investigación	0.628*** (0.061)	0.628** (0.292)	0.519** (0.263)	0.295*** (0.068)	0.295 (0.409)	0.226 (0.543)
Universidad privada asociativa	0.998*** (0.080)	0.998** (0.452)	0.066 (0.358)	1.021*** (0.088)	1.021** (0.470)	0.475 (0.859)
Universidad pública	1.168*** (0.086)	1.168** (0.531)	0.202 (0.378)	0.958*** (0.088)	0.958* (0.554)	-0.192 (0.766)
Inalpha			1.140*** (0.110)			15.853 (0.000)
Constante	-5.627*** (0.406)	-5.627*** (1.107)	-4.501*** (0.775)	-2.942*** (0.434)	-2.942*** (1.140)	-18.235*** (1.972)
Observaciones	381	381	381	165	165	165

Errores estándares entre paréntesis

*** p<0.01, ** p<0.05, * p<0.1

Cuadro 7.11

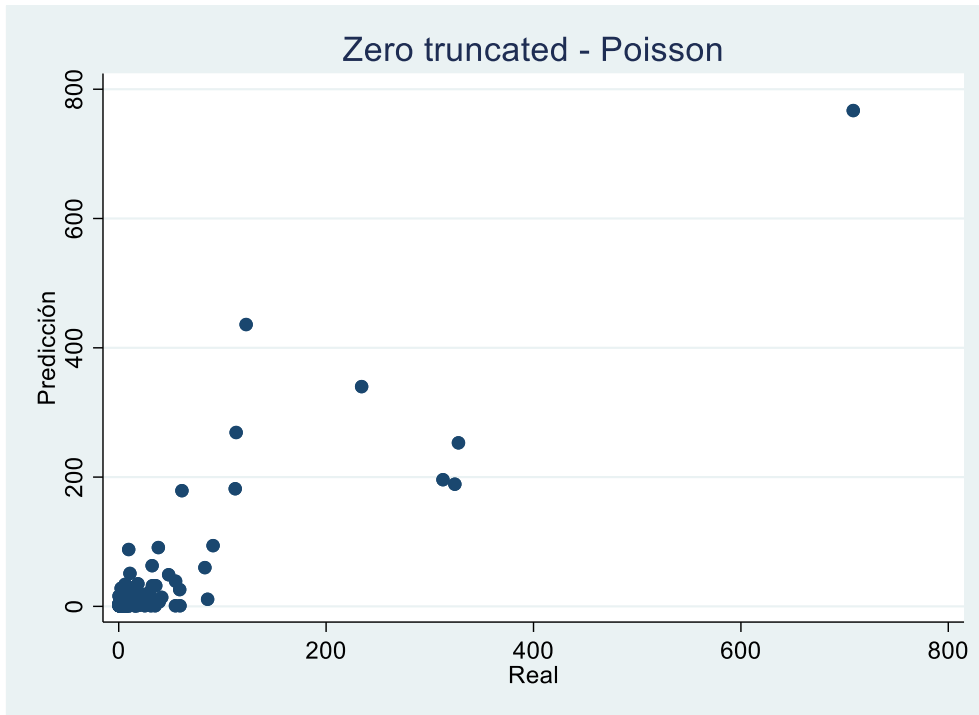
Valores de *log likelihood* y pseudo R^2 obtenidos para modelos de conteo

Número	Modelo	Distribución	Log likelihood	Pseudo R2
1	Base	Poisson	-2614.55	0.76
2	Base	Poisson con estimación robusta de varianza	-2614.55	0.76
3	Base	Binomial negativa	-719.11	0.14
4	Zero truncated	Poisson	-1883.22	0.75
5	Zero truncated	Poisson con estimación robusta de varianza	-1883.22	0.75
6	Zero truncated	Binomial negativa	-512.65	0.09

Fuente: Elaboración propia.

Gráfico 7.6

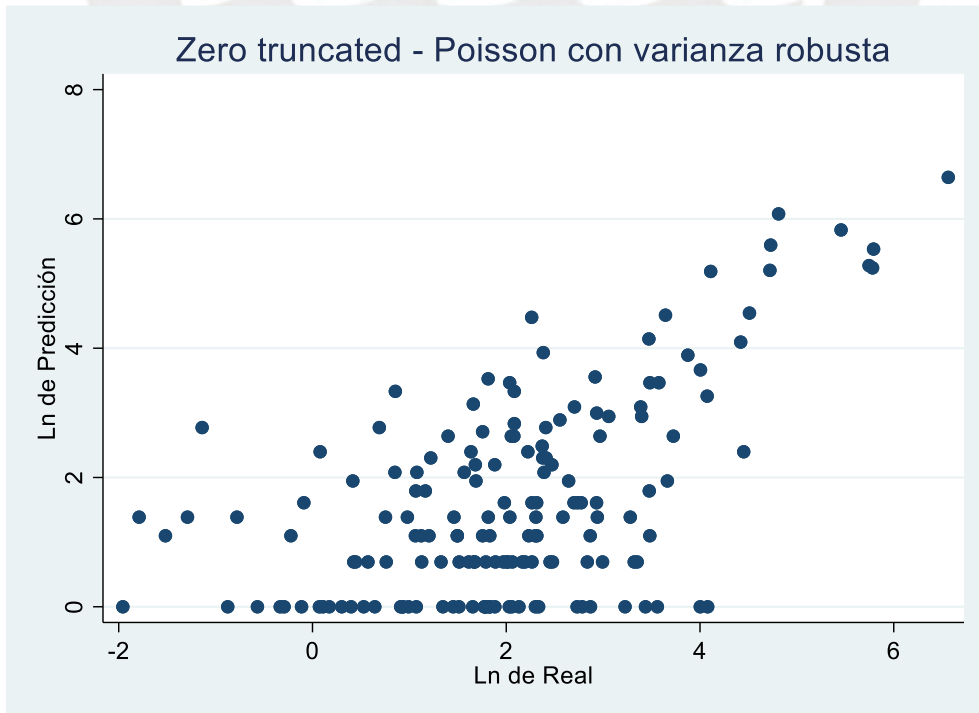
Dispersión entre conteo real y del conteo predicho con el modelo Poisson con varianza robusta *zero truncated*



Fuente: Elaboración propia.

Gráfico 7.7

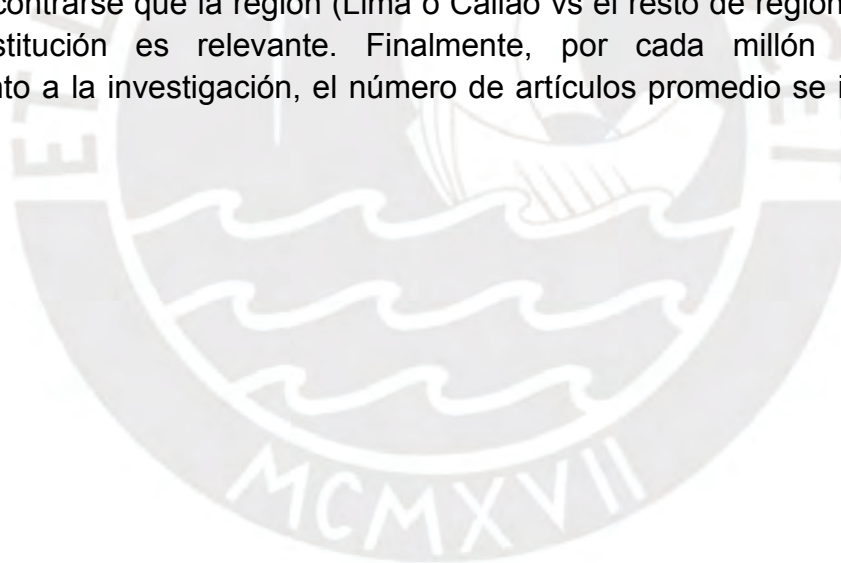
Dispersión entre logaritmo del conteo real y del conteo predicho con el modelo Poisson con varianza robusta *zero truncated*



Fuente: Elaboración propia.

Los resultados para los números de profesores según su perfil (sin doctorado, con doctorado en universidad de poca investigación, con doctorado en universidad de alta investigación) muestran que la cantidad de profesores sin doctorado no es relevante para el número de publicaciones, mientras que cada profesor adicional con doctorado en una universidad de alta investigación incrementa el número promedio de publicaciones en 1%, y que el efecto de los docentes con el otro tipo de doctorado es negativo (-2.7%).

Además, en línea con lo encontrado en las otras dos familias de regresiones, el área del conocimiento al que se dedica el docente es relevante. Pertenecer a Ciencias Agrícolas incrementaría el promedio del número de publicaciones en 2.98% con respecto a Humanidades, mientras que pertenecer a Medicina lo haría en 2.59%, y pertenecer a Ciencias Naturales en 2.42%; las áreas con los menores incrementos son Ingeniería y Tecnología (incremento de 1.71%) y Ciencias Sociales (incremento de 1.2%). De manera similar, se corrobora lo ya encontrado para el “tipo” de universidad: pertenecer a una universidad privada asociativa incrementa las publicaciones en 1.2% con respecto a pertenecer a una privada societaria, mientras que el incremento por pertenecer a una pública es ligeramente inferior a 1%. También vuelve a encontrarse que la región (Lima o Callao vs el resto de regiones) en la que está la institución es relevante. Finalmente, por cada millón adicional en financiamiento a la investigación, el número de artículos promedio se incrementaría en 13.6%.



8. CONCLUSIONES

En el planteamiento de la hipótesis se propusieron cinco factores relevantes para la generación de conocimiento dentro de las universidades peruanas, enmarcados dentro de la clasificación de tres grupos realizada por Teodorescu (2000). Los métodos aplicados hicieron posible medir la importancia esos de factores. Los modelos con variable dependiente binaria sirvieron para analizar la probabilidad de “haber publicado” un artículo, mientras que los modelos con variable dependiente de conteo fueron útiles para analizar la “intensidad” de esa producción.

En cuanto a la propensión a tener un artículo publicado e indizado, los resultados para los factores considerados en la hipótesis fueron:

- Uso del tiempo (T): el desempeñar cargo en la universidad podría generar un incremento en la probabilidad de publicar de 1.37%; ser docente a tiempo completo o con dedicación exclusiva disminuye la probabilidad en 0.3%; cada hora dedicada a investigar incrementa la probabilidad en 6.3%, mientras que cada hora dedicada a actividades no universitarias disminuyen la misma probabilidad en 1.5%.
- Motivaciones (M): por cada mil soles de incremento en el ingreso, la probabilidad de tener publicaciones se incrementa en 0.2%; el disponer de un ambiente incrementa la probabilidad en 0.9%; el creer que la investigación de su universidad contribuye a sociedad incrementa la probabilidad en 1.5%.
- Tipo de institución (I): si es pública 5.4%, privada asociativa 3.4%⁴⁴.
- Financiamiento (S): por cada millón de soles para investigación, la probabilidad se incrementa en 6.2%.
- Entrenamiento para investigación (D): Tener un doctorado tiene efectos positivos en la probabilidad de publicar. Si el grado se obtuvo en una universidad con alta investigación, la probabilidad llega a ser 6.4%, mientras que si fue obtenido en una universidad con baja investigación, será de 2.9%.

Sobre la “intensidad” de publicación, se encontraron las siguientes variables significativas:

- Uso del tiempo (T): si el promedio de horas de enseñanza se incrementa en una hora, el promedio de publicaciones disminuye en 19.6%.
- Motivaciones (M): el incremento en 1% de la proporción de profesores que creen que la investigación de su universidad contribuye a sociedad se relaciona con un incremento de 3.9% del promedio de publicaciones.

⁴⁴ Esta información se obtuvo del portal de SUNEDU: <https://www.sunedu.gob.pe/>

- Tipo de institución (I): existe diferencia significativa en el número de publicaciones de acuerdo con el tipo de gestión de la institución, la ventaja es de la universidad pública o privada asociativa sobre la societaria.
- Financiamiento (S): el efecto del financiamiento se encontró entre 12.2% y 13.6% de incremento en el promedio de publicaciones por cada millón de soles adicionales.
- Entrenamiento para investigación (D): considerar la diferencia en los doctorados reportados hizo posible encontrar diferencias en los efectos de tener profesores con doctorado en universidades con baja investigación o en universidad con alta investigación. La relación encontrada entre los primeros y el número de publicaciones fue que, por cada uno de esos doctores, el promedio de publicaciones disminuye en -6.5% o -2.7% (dependiendo del modelo utilizado), mientras que por uno de los doctores del segundo tipo, se incrementaría en 1% el promedio de publicaciones.

Se destaca también lo encontrado en cuanto a las diferencias entre las áreas de conocimiento. Existe una ventaja significativa de todas las áreas sobre las Humanidades, tanto en la probabilidad de tener al menos una publicación como en el número de publicaciones; además, esa diferencia no es homogénea. Así, el pertenecer a Medicina incrementa la probabilidad de publicar en más de 5%, y el número de publicaciones en alrededor de 3.5%, mientras que el efecto de pertenecer a Ciencias Sociales o Ingeniería es mucho menor.

De cara a las políticas públicas, la relevancia de los hallazgos puede ejemplificarse con los siguientes puntos:

- Entre las “Condiciones básicas de calidad” (CBC) propuestas por la SUNEDU, el hecho de exigir determinados grados académicos se debe a que se desea tener personal adecuado para transmitir y crear conocimiento, sin embargo, debe tenerse en cuenta que las características de la institución en la que se obtuvo ese grado académico son importantes pues se ha mostrado que el tipo de universidad en la que se obtuvo el doctorado está relacionado con el aporte a la producción de conocimiento.
- Mantener un porcentaje mínimo de profesores a tiempo completo también es un requisito de las CBC de la SUNEDU. No obstante, esa intervención solo incrementaría la investigación de la universidad si esta permite que dediquen parte importante de ese tiempo a investigar, y no solo los recarga con actividades de enseñanza o labores administrativas.
- Por otro lado, uno de los instrumentos de política del CONCYTEC es el otorgamiento de becas de posgrado “a través de FENDECYT y FINCYT en universidades peruanas y en el exterior” (CONCYTEC 2016b, p.113). El objetivo de esa política es aportar al desarrollo de la Ciencia y Tecnología en

Perú. Nuevamente, el instrumento debe considerar la diferencia en la formación recibida según cuál haya sido la universidad.

- Además, este año se publicó el “Reglamento de calificación, clasificación y registro de los investigadores del sistema nacional de ciencia, tecnología e innovación tecnológica⁴⁵”, en el que se detalla los criterios que debe cumplir un investigador para ser catalogado en determinado nivel. Sin embargo, se ha mostrado que existen diferencias significativas en la productividad científica según el tema de estudio, pero el reglamento no contempla eso.

Por lo tanto, los resultados de esta tesis cumplen con el objetivo de “aportar a la discusión de sobre la calidad universitaria en el Perú”.

Limitaciones

Debe señalarse también que este trabajo muestra la necesidad de contar con un sistema de información actualizado y pertinente sobre las universidades, sus características institucionales, sus profesores y la producción académica de estos. De esa manera, el monitoreo y evaluación de la situación de la producción académica en el país podría realizarse de mejor manera.

A propósito de lo mencionado en el párrafo anterior, hay varios aspectos de este trabajo que pueden ser mejorados. Algunas sugerencias para esto son:

- Afinar los criterios para contabilizar las publicaciones (considerar, por ejemplo: diferencias entre tipos de publicaciones, citas, número de autores por publicación, etc.), con la posibilidad de construir algún índice compuesto, por ejemplo (El Gibari, Gómez y Ruiz, 2018).
- Utilizar bases de datos con información actualizada para las características de las instituciones y los profesores.
- De no contarse con datos administrativos universales (del tipo del II CNU), centrarse solo en una universidad o un grupo de estas, o solo en un área del conocimiento, tomando como criterio la información respecto a los investigadores y sus características disponibles.
- Explorar los efectos de cambios en el contexto en el que se desenvuelven las universidades. Por ejemplo, la dación de la Nueva Ley Universitario o la creación de Vicerrectorados de Investigación dentro de esas instituciones generarían cambios en los perfiles de los profesores, y estos cambios a su vez generarían comportamientos distintos en la producción científica.

⁴⁵ Disponible en

https://portal.concytec.gob.pe/images/noticias/Propuesta_del_nuevo_Reglamento_del_investigador.pdf

Las políticas públicas elaboradas con base en evidencia tienden a tener mejores resultados porque están elaboradas con más información, por lo que se espera que este trabajo sirva para enriquecer la discusión sobre el tipo de universidad, y dentro de ella el tipo de profesor, que el Perú necesita para incrementar su producción de conocimiento.



9. BIBLIOGRAFÍA

- Bar-Ilan, J. (2008). Which h-index? – A comparison of WoS, Scopus and Google Scholar. *Scientometrics*. 74(2). 257-271.
- Becker, G. (1993), *Human capital: a theoretical and empirical analysis, with special reference to education*, Chicago, United States: The university of Chicago Press.
- Bénabou, R. y Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of economic studies*, 70(3), 489 – 520.
- Bookstein, A. (1976). The bibliometric distributions. *Library Quarterly*, 46(4), 416–423.
- Cameron, A. y Trivedi, P. (2010). *Microeconometrics using Stata*. College Station, Texas: Stata Press.
- Cameron, A. y Trivedi, P. (1998). *Regression analysis of count data*. Cambridge, United Kingdom: Cambridge University Press.
- Casse, A. (2006). The primacy of education. Understanding poverty. *Oxford University Press*.
- Chavarro, D., Rafols, I. y Tang, P. (2018). To what extent is inclusion in the Web of Science and indicator of journal 'quality'?. *Research Evaluation*, 27(2), 106 – 118.
- Coxe, S., West, S. y Aiken, L. (2009). The analysis of count data: a gentle introduction to Poisson Regression and its alternatives. *Journal of Personality Assessment*, 91:2, 121-136.
- Creamer, E. G. (1998). Assessing faculty publication productivity: issues of equity. *ASHE-ERIC Tiger Education Report*, 26(2), 1-91.
- CONCYTEC (2016a). I Censo Nacional de Investigación y Desarrollo a Centros de Investigación. Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica.
- CONCYTEC (2016b). Política Nacional para el Desarrollo de la Ciencia, Tecnología e Innovación Tecnológica – CTI.

- Cuenca, R. y Reátegui, L. (2016). *La (incumplida) promesa universitaria en el Perú*. Recuperado de <http://repositorio.iep.org.pe/handle/IEP/9>
- Deb, P, Norton, E., y Manning, W. (2017). *Health Econometrics Using Stata*. Stata Press.
- Diamond, N., y Graham, H. (2000). How Should We Rate Research Universities?. *Change: The Magazine of Higher Learning*, 32(4), 20-33.
- Dundar, H. y Lewis, D. (1998). Determinants of research productivity in higher education. *Research in higher education*. 39(6), 607 – 631.
- Samira El Gibari, Trinidad Gómez, Francisco Ruiz. (2018) Evaluating university performance using reference point based composite indicators. *Journal of Informetrics*. Volume 12, Issue 4, 1235-1250.
- Elsevier (2017). Scopus content coverage guide.
https://www.elsevier.com/__data/assets/pdf_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf
- Fram, E. H., & Lau, G. H. (1996). Research universities versus teaching universities - public perceptions and preferences. *Quality Assurance in Education: An International Perspective*, 4(3), 27.
- Feldman, K. (1987). Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: A review and exploration. *Research in Higher Education*, 26, 227-298.
- Fox, M. F. (1985). Publication, performance, and reward in science and scholarship. En Smart J. C. (ed.), *Higher Education: Handbook of Theory and Research*, Vol. 1. New York: Agathon Press.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. y Trow, M. (1994), *New Production of Knowledge: The dynamics of science and research in contemporary societies*, London, Great Britain: Sage publications.
- Guellec, D. y Pottelsberghe. B. (2001), “R&D and Productivity Growth: Panel Data Analysis of 16 OECD Countries”, *OECD Science, Technology and Industry Working Papers*, (2001)3.
- Hattie, J. y Marsh, W. (1996). The Relationship between Research and Teaching: A Meta-analysis. *Review of educational research*. 66(4),507-542.

- Hardré, P., Beesley, A., Miller, R. y Pace, T. (2011). Faculty motivation to do research: across disciplines in research-extensive universities. *The journal of the Professoriate*.
- Harzing, A. y Alakangas, S. (2016). Google Scholar, Scopus and Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*. 106:787-804
- Hesli, V. L., y Lee, J. M. (2011). Faculty Research Productivity: Why Do Some of Our Colleagues Publish More than Others? *PS: Political Science & Politics*, 44(2), 393.
- Hvide, H. y Jones, B. (2018). University Innovation and the Professor's Privilege. *American Economic Review*. 108(7): 1860 – 1898.
- Johnes, A., Taylor, J., y Francis, B. (1993). The Research Performance of UK Universities: A Statistical Analysis of the Results of the 1989 Research Selectivity Exercise. *Journal Of The Royal Statistical Society. Series A (Statistics In Society)*, (2), 271.
- Jopen, G., Gómez, W. y Olivera, H. (2014). Departamento de Economía PUCP. Sistema educativo peruano: balance y agenda pendiente (379). Recuperado de <http://files.pucp.edu.pe/departamento/economia/DDD379.pdf>
- Kumar, S., y Senthilkumar, R. (2019). Applicability of lotka's law in astronomy & astrophysics research of india. *Library Philosophy and Practice*, 2129.
- Kwiek, M. (2018). High research productivity in vertically undifferentiated higher education systems: Who are the top performers? *Scientometrics*, 115(1), 415-462.
- Lotka, A. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12), 317-323.
- Marsh, H. y Hattie, J. (2002). The Relation between Research Productivity and Teaching Effectiveness: Complementary, Antagonistic, or Independent Constructs? *The Journal of Higher Education*, 73(5), 603.
- Martin-Martin, A., Orduna-Malea, E., Thelwall, M. y Lopez-Cozar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citation in 252 subject categories. *Journal of Informetrics*. 12 (2018): 1160-1177.
- Ministerio de Educación del Perú (MINEDU). (2015). Lineamientos de política para el aseguramiento de la calidad en la educación superior universitaria: aportes y

recomendaciones.

Recuperado

de

<http://repositorio.minedu.gob.pe/handle/123456789/3727>

- Miss, S. P. S., & Anandhalli, G. B. (2018). Authorship trend and collaborative research in lung cancer: A time series analysis study. *Library Philosophy and Practice*, 2018
- Mueller, C., Gaus, H., Konradt, I. (2016). Predicting Research Productivity In International Evaluation Journals Across Countries. *Journal Of MultiDisciplinary Evaluation*. 12(27).
- Organización para la Cooperación y el Desarrollo Económicos (2007). Revised Field of Science and Technology (FOS) Classification in the Frascati.
- Piro, F., Rørstad, K. y Aksnes, D. (2016). How does prolific professors influence on the citation impact of their university departments?. *Scientometrics*. 107:941 – 961.
- Print, M. y Hattie, J. (1997). Measuring quality in universities: An approach to weighting research productivity. *Higher Education*. 33 ,453 – 469.
- Raymond A. K. Cox, y Kee H. Chung. (1991). Patterns of Research Output and Author Concentration in the Economics Literature. *The Review of Economics and Statistics*, 73(4).
- RICYT Red Iberoamericana de Ciencia y Tecnología.
<http://dev.ricyt.org/ui/v3/comparative.html?indicator=CSCI>
- Rørstad, K. y Aksnes, D. (2015). Publication rate expressed by age, gender and academic position – A large-scale analysis of Norwegian academic staff. *Journal of Informetrics*. 317 – 333.
- Sadik Batcha, M. (2018). Lotka's applicability on global dengue research publication: A scientometric study. *DESIDOC Journal of Library and Information Technology*, 38(4), 266-270.
- Sam, C. y van der Sijde, P. (2014). Understanding the concept of the entrepreneurial university from the perspective of higher education models. *Higher Education*, 68(6), 891.
- Shin, J., & Cummings, W. (2010). Multilevel analysis of academic publishing across disciplines: research preference, collaboration, and time on research. *Scientometrics*, 85(2), 581–594.

- Stoker, T. (1993). Empirical Approaches to the Problem of Aggregation Over Individuals. *Journal of Economic Literature*.
- SUNEDU (2015). El Modelo de Licenciamiento y su Implementación en el Sistema Universitario Peruano.
- Tan, D. (1986). The Assessment of Quality in Higher Education: A Critical Review of the Literature and Research. *Research In Higher Education*, (3), 223.
- Teodorescu, D. (2000). Correlates of faculty publication productivity: A cross national analysis. *Higher Education*. 39, 201-222.
- Toutkoushian, R. y Webber, K. (2011). Measuring the research performance of postsecondary institutions. En Cheol, J. y otros (editores). *University rankings. Theoretical basis, methodology and impacts on global higher education*.
- Van Arensbergen, P., van der Weijden, I., y van den Besselaar, P. (2012). Gender differences in scientific productivity: a persisting phenomenon? *Scientometrics*, 93(3), 857–868.
- Vitanov, N. (2016). *Science dynamics and research production. Indicators, indexes, statistical laws and mathematical models*. Springer International Publishing.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: The MIT Press.
- Zhang, L. y Ehrenberg, R. (2010). Faculty Employment and R&D Expenditures at Research Universities. *Economics of Education Review*. 29(2010) 329-337

ANEXOS

Cuadro 1A

Número de investigaciones según tipo, por universidad

UNIVERSIDAD	TIPO DE INVESTIGACIÓN			TOTAL
	Aplicada	Básica	Tecnológica	
UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS	1015	798	131	1944
UNIVERSIDAD NACIONAL FEDERICO VILLARREAL	545	593	120	1258
UNIVERSIDAD ALAS PERUANAS	595	501	128	1224
PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ	583	517	53	1153
UNIVERSIDAD NACIONAL DE TRUJILLO	570	521	48	1139
UNIVERSIDAD DE SAN MARTÍN DE PORRES	607	369	66	1042
UNIVERSIDAD NACIONAL SAN LUIS GONZAGA	468	429	93	990
UNIVERSIDAD NACIONAL DEL ALTIPLANO	462	366	101	929
UNIVERSIDAD NACIONAL DEL CENTRO DEL PERU	434	349	145	928
UNIVERSIDAD NACIONAL DE SAN AGUSTÍN	426	319	135	880
UNIVERSIDAD NACIONAL PEDRO RUÍZ GALLO	392	300	92	784
UNIVERSIDAD NACIONAL DE EDUCACIÓN ENRIQUE GUZMAN Y VALLE	361	297	90	748
UNIVERSIDAD PRIVADA CÉSAR VALLEJO	341	276	39	656
UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD	253	256	139	648
UNIVERSIDAD NACIONAL JOSÉ FAUSTINO SÁNCHEZ CARRIÓN	274	255	102	631
UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA	331	238	46	615
UNIVERSIDAD NACIONAL DE PIURA	306	223	64	593
UNIVERSIDAD NACIONAL HERMILIO VALDIZÁN	305	143	48	496
UNIVERSIDAD PERUANA CAYETANO HEREDIA	319	157	16	492
UNIVERSIDAD NACIONAL DE CAJAMARCA	199	246	26	471
UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN	150	249	68	467
UNIVERSIDAD NACIONAL DE INGENIERÍA	280	109	74	463
UNIVERSIDAD NACIONAL DE LA AMAZONÍA PERUANA	205	170	66	441
UNIVERSIDAD NACIONAL AGRARIA LA MOLINA	269	82	69	420
UNIVERSIDAD CATÓLICA SANTA MARÍA	201	157	54	412
UNIVERSIDAD NACIONAL SANTIAGO ANTÚNEZ DE MAYOLO	236	139	32	407
UNIVERSIDAD NACIONAL DEL CALLAO	183	148	62	393
UNIVERSIDAD RICARDO PALMA	198	151	40	389

UNIVERSIDAD ANDINA NESTOR CÁCERES VELÁSQUEZ	187	158	32	377
UNIVERSIDAD INCA GARCILASO DE LA VEGA	190	133	33	356
UNIVERSIDAD NACIONAL DANIEL ALCIDES CARRIÓN	167	94	62	323
UNIVERSIDAD PERUANA LOS ANDES	137	132	43	312
UNIVERSIDAD CATÓLICA LOS ÁNGELES DE CHIMBOTE	128	153	17	298
UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO	139	116	23	278
UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS	158	78	26	262
UNIVERSIDAD PRIVADA SAN PEDRO	132	111	18	261
UNIVERSIDAD DE LIMA	120	115	15	250
UNIVERSIDAD NACIONAL DE HUANCAMELICA	106	97	26	229
UNIVERSIDAD DE PIURA	111	94	19	224
UNIVERSIDAD NACIONAL DEL SANTA	85	88	48	221
UNIVERSIDAD NACIONAL AGRARIA DE LA SELVA	102	79	17	198
UNIVERSIDAD PRIVADA ANTONIO ORREGO	88	89	17	194
UNIVERSIDAD TECNOLÓGICA DEL PERU	107	53	34	194
UNIVERSIDAD ANDINA DEL CUSCO	95	83	9	187
ASOCIACIÓN UNIVERSIDAD PRIVADA SAN JUAN BAUTISTA	74	95	14	183
UNIVERSIDAD NACIONAL DE SAN MARTÍN	91	44	32	167
UNIVERSIDAD PRIVADA SEÑOR DE SIPAN	73	73	19	165
UNIVERSIDAD NACIONAL DE UCAYALI	92	49	13	154
UNIVERSIDAD PERUANA UNIÓN	74	51	27	152
UNIVERSIDAD PRIVADA SAN IGNACIO DE LOYOLA	74	60	10	144
UNIVERSIDAD DE TACNA	73	54	13	140
UNIVERSIDAD PRIVADA NORBERT WIENER	79	52	5	136
UNIVERSIDAD CIENTÍFICA DEL SUR	66	56	6	128
UNIVERSIDAD NACIONAL TORIBIO RODRIGUEZ DE MENDOZA DE AMAZONAS	53	60	10	123
UNIVERSIDAD DEL PACÍFICO	93	24	3	120
UNIVERSIDAD FEMENINA DEL SAGRADO CORAZÓN	52	52	3	107
UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURÍMAC	46	38	22	106
UNIVERSIDAD NACIONAL AMAZÓNICA DE MADRE DE DIOS	38	42	24	104
UNIVERSIDAD NACIONAL DE TUMBES	55	43	5	103
UNIVERSIDAD CATÓLICA SAN PABLO	31	45	23	99
UNIVERSIDAD CIENTÍFICA DEL PERÚ	52	36	10	98
UNIVERSIDAD CONTINENTAL DE CIENCIA E INGENIERÍA	51	34	7	92
UNIVERSIDAD PRIVADA DEL NORTE	44	26	6	76
UNIVERSIDAD TECNOLÓGICA DE LOS ANDES	27	42	4	73

UNIVERSIDAD PRIVADA DE HUÁNUCO	39	24	10	73
UNIVERSIDAD CATÓLICA SEDES SAPIENTIAE	36	30	1	67
UNIVERSIDAD JOSÉ CARLOS MARIÁTEGUI	31	19	8	58
UNIVERSIDAD ANTONIO RUÍZ DE MONTOYA	20	34	1	55
UNIVERSIDAD PRIVADA ANTONIO GUILLERMO URRELO	16	30	4	50
UNIVERSIDAD DE CIENCIAS Y HUMANIDADES	15	25	2	42
UNIVERSIDAD DE ADMINISTRACIÓN DE NEGOCIOS ESAN	23	18		41
UNIVERSIDAD PERUANA DE LAS AMÉRICAS	20	17	2	39
UNIVERSIDAD JAIME BAUSATE Y MEZA	16	21		37
UNIVERSIDAD PERUANA DE CIENCIAS E INFORMÁTICA	17	12	6	35
UNIVERSIDAD NACIONAL TECNOLÓGICA DEL CONO SUR DE LIMA	11	18	3	32
UNIVERSIDAD NACIONAL INTERCULTURAL DE LA AMAZONÍA	21	8	2	31
UNIVERSIDAD CATÓLICA DE TRUJILLO	12	17		29
UNIVERSIDAD PERUANA DE INTEGRACIÓN GLOBAL	10	14	3	27
UNIVERSIDAD PRIVADA SERGIO BERNALES	11	13	1	25
UNIVERSIDAD PRIVADA SAN CARLOS	11	11	2	24
UNIVERSIDAD NACIONAL JOSÉ MARÍA ARGUEDAS	12	7	4	23
UNIVERSIDAD PRIVADA JUAN MEJÍA BACA	13	7	3	23
UNIVERSIDAD PRIVADA MARCELINO CHAMPAGNAT	10	8	1	19
UNIVERSIDAD PRIVADA DE TRUJILLO	8	9		17
UNIVERSIDAD PARTICULAR DE CHICLAYO	6	8	2	16
UNIVERSIDAD ADA A BYRON	5	8	3	16
FACULTAD DE TEOLOGÍA PONTIFICIA Y CIVIL DE LIMA	7	8		15
UNIVERSIDAD PRIVADA DE LAMBAYEQUE	4	6	1	11
UNIVERSIDAD AUTÓNOMA DEL PERÚ	7	4		11
UNIVERSIDAD PERUANA SIMÓN BOLIVAR	7	2	2	11
UNIVERSIDAD NACIONAL DE MOQUEGUA	2	5	2	9
UNIVERSIDAD PERUANA DEL ORIENTE	4	3	1	8
UNIVERSIDAD PRIVADA TELESUP	5	1	1	7
UNIVERSIDAD LE CORDON BLEU	4	1		5
UNIVERSIDAD PRIVADA DE HUANCAYO	2	2	1	5
UNIVERSIDAD PRIVADA DE PUCALLPA	4	1		5
UNIVERSIDAD PERUANA DEL CENTRO	2	1		3
UNIVERSIDAD PRIVADA ARZOBISPO LOAYZA	1	2		3
UNIVERSIDAD PARA EL DESARROLLO ANDINO	1	1	1	3
UNIVERSIDAD PRIVADA DE ICA	1			1
TOTAL	14212	11402	2909	28523

Fuente: II CNU. Elaboración propia.



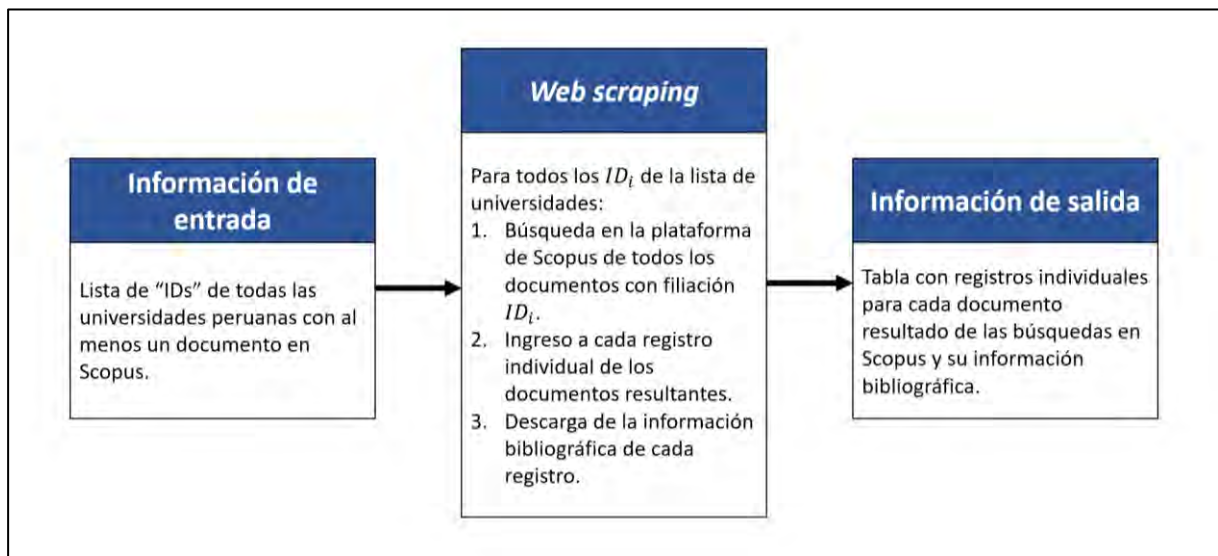
Cuadro 2A:**Número de investigaciones difundidas en revistas indexadas, por universidad**

UNIVERSIDAD	TOTAL
UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS	633
UNIVERSIDAD PERUANA CAYETANO HEREDIA	209
UNIVERSIDAD NACIONAL DE TRUJILLO	187
UNIVERSIDAD DE SAN MARTÍN DE PORRES	172
UNIVERSIDAD NACIONAL HERMILIO VALDIZÁN	167
UNIVERSIDAD NACIONAL FEDERICO VILLARREAL	159
PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ	155
UNIVERSIDAD NACIONAL DE CAJAMARCA	124
UNIVERSIDAD PRIVADA CÉSAR VALLEJO	115
UNIVERSIDAD NACIONAL SANTIAGO ANTÚNEZ DE MAYOLO	113
UNIVERSIDAD NACIONAL AGRARIA LA MOLINA	106
UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA	97
UNIVERSIDAD NACIONAL DE INGENIERÍA	91
UNIVERSIDAD RICARDO PALMA	88
UNIVERSIDAD CATÓLICA SANTA MARÍA	86
UNIVERSIDAD NACIONAL DE PIURA	73
UNIVERSIDAD ALAS PERUANAS	70
UNIVERSIDAD NACIONAL DEL CENTRO DEL PERU	68
UNIVERSIDAD NACIONAL DE SAN AGUSTÍN	50
UNIVERSIDAD PRIVADA ANTENOR ORREGO	45
UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD	43
UNIVERSIDAD NACIONAL DEL ALTIPLANO	42
UNIVERSIDAD PERUANA LOS ANDES	42
UNIVERSIDAD NACIONAL DE LA AMAZONÍA PERUANA	39
UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS	39
UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN	38
UNIVERSIDAD NACIONAL SAN LUIS GONZAGA	38
UNIVERSIDAD CIENTÍFICA DEL SUR	37
UNIVERSIDAD NACIONAL DANIEL ALCIDES CARRIÓN	36
UNIVERSIDAD NACIONAL PEDRO RUÍZ GALLO	32
UNIVERSIDAD INCA GARCILASO DE LA VEGA	32
UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO	30
UNIVERSIDAD DE PIURA	26
UNIVERSIDAD DE LIMA	24
UNIVERSIDAD NACIONAL DE HUANCVELICA	24
UNIVERSIDAD DEL PACÍFICO	23
ASOCIACIÓN UNIVERSIDAD PRIVADA SAN JUAN BAUTISTA	22
UNIVERSIDAD NACIONAL JOSÉ FAUSTINO SÁNCHEZ CARRIÓN	21
UNIVERSIDAD TECNOLÓGICA DEL PERU	21
UNIVERSIDAD NACIONAL DE EDUCACIÓN ENRIQUE GUZMAN Y VALLE	21
UNIVERSIDAD NACIONAL DE UCAYALI	20
UNIVERSIDAD ANDINA NESTOR CÁCERES VELÁSQUEZ	20
UNIVERSIDAD FEMENINA DEL SAGRADO CORAZÓN	20
UNIVERSIDAD CATÓLICA LOS ÁNGELES DE CHIMBOTE	19

UNIVERSIDAD NACIONAL DE TUMBES	19
UNIVERSIDAD PRIVADA SAN IGNACIO DE LOYOLA	19
UNIVERSIDAD NACIONAL AMAZÓNICA DE MADRE DE DIOS	18
UNIVERSIDAD NACIONAL DEL CALLAO	17
UNIVERSIDAD PRIVADA NORBERT WIENER	17
UNIVERSIDAD DE TACNA	17
UNIVERSIDAD PRIVADA SAN PEDRO	15
UNIVERSIDAD PERUANA UNIÓN	14
UNIVERSIDAD PRIVADA SEÑOR DE SIPAN	14
UNIVERSIDAD ANDINA DEL CUSCO	13
UNIVERSIDAD CATÓLICA SAN PABLO	13
UNIVERSIDAD DE ADMINISTRACIÓN DE NEGOCIOS ESAN	12
UNIVERSIDAD NACIONAL DEL SANTA	12
UNIVERSIDAD NACIONAL AGRARIA DE LA SELVA	11
UNIVERSIDAD NACIONAL DE SAN MARTÍN	10
UNIVERSIDAD CIENTÍFICA DEL PERÚ	10
UNIVERSIDAD PRIVADA DEL NORTE	9
UNIVERSIDAD NACIONAL TORIBIO RODRIGUEZ DE MENDOZA DE AMAZONAS	9
UNIVERSIDAD PRIVADA DE HUÁNUCO	8
UNIVERSIDAD PERUANA DE CIENCIAS E INFORMÁTICA	8
UNIVERSIDAD CATÓLICA SEDES SAPIENTIAE	6
UNIVERSIDAD NACIONAL TECNOLÓGICA DEL CONO SUR DE LIMA	6
UNIVERSIDAD ANTONIO RUÍZ DE MONTOYA	6
UNIVERSIDAD CONTINENTAL DE CIENCIA E INGENIERÍA	6
UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURÍMAC	6
UNIVERSIDAD PRIVADA ANTONIO GUILLERMO URRELO	5
UNIVERSIDAD NACIONAL INTERCULTURAL DE LA AMAZONÍA	5
UNIVERSIDAD PRIVADA MARCELINO CHAMPAGNAT	4
UNIVERSIDAD JOSÉ CARLOS MARIÁTEGUI	4
UNIVERSIDAD NACIONAL JOSÉ MARÍA ARGUEDAS	3
UNIVERSIDAD AUTÓNOMA DEL PERÚ	3
FACULTAD DE TEOLOGÍA PONTIFICIA Y CIVIL DE LIMA	2
UNIVERSIDAD PRIVADA SAN CARLOS	2
UNIVERSIDAD DE CIENCIAS Y HUMANIDADES	2
UNIVERSIDAD TECNOLÓGICA DE LOS ANDES	2
UNIVERSIDAD PERUANA DE LAS AMÉRICAS	2
UNIVERSIDAD PRIVADA DE TRUJILLO	2
UNIVERSIDAD NACIONAL DE MOQUEGUA	1
UNIVERSIDAD PERUANA SIMÓN BOLIVAR	1
UNIVERSIDAD JAIME BAUSATE Y MEZA	1
UNIVERSIDAD PRIVADA DE LAMBAYEQUE	1
UNIVERSIDAD LE CORDON BLEU	1
TOTAL	3853

Diagrama 3A

Diagrama resumen del proceso utilizado para la construcción de la base de datos bibliométrica a partir de Scopus



Elaboración propia.



Cuadro 4A
Correspondencia entre áreas del conocimiento consideradas por la OCDE y Scopus

SCOPUS		OCDE
Área general	Área específica	
Health Sciences	Medicine	Medicina y ciencias de la salud
Health Sciences	Nursing	Medicina y ciencias de la salud
Health Sciences	Veterinary	Ciencias agrarias
Health Sciences	Dentistry	Medicina y ciencias de la salud
Health Sciences	Health Professions	Medicina y ciencias de la salud
Life Sciences	Agricultural and Biological Sciences	Ciencias agrarias
Life Sciences	Biochemistry, Genetics and Molecular Biology	Ciencias naturales
Life Sciences	Immunology and Microbiology	Ciencias naturales
Life Sciences	Neuroscience	Ciencias naturales
Life Sciences	Pharmacology, Toxicology and Pharmaceutics	Medicina y ciencias de la salud
Physical Sciences	Chemical Engineering	Ingeniería y tecnología
Physical Sciences	Chemistry	Ciencias naturales
Physical Sciences	Computer Science	Ciencias naturales
Physical Sciences	Earth and Planetary Sciences	Ciencias naturales
Physical Sciences	Energy	Ingeniería y tecnología
Physical Sciences	Engineering	Ingeniería y tecnología
Physical Sciences	Environmental Science	Ciencias naturales
Physical Sciences	Materials Science	Ingeniería y tecnología
Physical Sciences	Mathematics	Ciencias naturales
Physical Sciences	Physics and Astronomy	Ciencias naturales
Social Sciences	Arts and Humanities	Humanidades
Social Sciences	Business, Management and Accounting	Ciencias sociales
Social Sciences	Decision Sciences	Ciencias naturales
Social Sciences	Economics, Econometrics and Finance	Ciencias sociales
Social Sciences	Psychology	Ciencias sociales
Social Sciences	Social Sciences	Ciencias sociales

Elaboración propia.



Cuadro 5A

Países cuyos doctorados fueron considerados como de “Tipo 2” y universidades peruanas cuyos doctorados fueron considerados como de “Tipo 1”

Países con doctorados "tipo 2"	Universidades peruanas con doctorados "tipo 1"
Angola Burkina Faso Ecuador Gambia Kiribati Mozambique Puerto Rico	Pontificia Universidad Católica del Perú Universidad de San Martín de Porres Universidad Nacional Agraria la Molina Universidad Nacional de Ingeniería Universidad Nacional de San Agustín Universidad Nacional de San Antonio Abad del Cusco Universidad Nacional Mayor de San Marcos Universidad Peruana Cayetano Heredia Universidad Peruana de Ciencias Aplicadas



Cuadro 5A**Número de publicaciones por universidad en Scopus (2008 y 2009)**

UNIVERSIDAD	2008	2009	TOTAL
UNIVERSIDAD PERUANA CAYETANO HEREDIA	155	170	325
UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS	86	104	190
PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ	71	63	134
UNIVERSIDAD NACIONAL AGRARIA LA MOLINA	25	28	53
UNIVERSIDAD NACIONAL DE INGENIERÍA	11	19	30
UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD	8	21	29
UNIVERSIDAD NACIONAL DE SAN AGUSTÍN	10	11	21
UNIVERSIDAD NACIONAL DE TRUJILLO	9	9	18
UNIVERSIDAD RICARDO PALMA	10	6	16
UNIVERSIDAD NACIONAL DE LA AMAZONÍA PERUANA	5	10	15
UNIVERSIDAD DE SAN MARTÍN DE PORRES	6	8	14
UNIVERSIDAD DE PIURA	4	6	10
UNIVERSIDAD CATÓLICA SAN PABLO	4	5	9
UNIVERSIDAD NACIONAL SAN LUIS GONZAGA	3	5	8
UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS		5	5
UNIVERSIDAD NACIONAL FEDERICO VILLARREAL	5		5
UNIVERSIDAD NACIONAL DE PIURA	5		5
UNIVERSIDAD CATÓLICA SANTA MARÍA	4		4
TOTAL	421	470	891

Cuadro 6A**Número de publicaciones por universidad en Web of Science (2008 y 2009)**

UNIVERSIDAD	2008	2009	TOTAL
UNIVERSIDAD PERUANA CAYETANO HEREDIA	186	202	388
UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS	102	89	191
PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ	66	60	126
UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD	15	24	39
UNIVERSIDAD NACIONAL AGRARIA LA MOLINA	26		26
UNIVERSIDAD NACIONAL DE INGENIERÍA	10	10	20
UNIVERSIDAD RICARDO PALMA	10	7	17
UNIVERSIDAD NACIONAL DE LA AMAZONÍA PERUANA	3	12	15
UNIVERSIDAD NACIONAL DE TRUJILLO	8	5	13
UNIVERSIDAD NACIONAL DE SAN AGUSTÍN	9	4	13
UNIVERSIDAD DE PIURA	5	6	11
UNIVERSIDAD NACIONAL DE PIURA	6	5	11
UNIVERSIDAD CATÓLICA SANTA MARÍA	4	6	10
UNIVERSIDAD NACIONAL SAN CRISTOBAL DE HUAMANGA	10		10
UNIVERSIDAD DE SAN MARTÍN DE PORRES	6	4	10
UNIVERSIDAD NACIONAL FEDERICO VILLARREAL	3	6	9
UNIVERSIDAD NACIONAL DEL ALTIPLANO	3	4	7
UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS		6	6
UNIVERSIDAD DEL PACÍFICO	2	3	5
UNIVERSIDAD NACIONAL DE SAN MARTÍN	4		4
UNIVERSIDAD DE ADMINISTRACIÓN DE NEGOCIOS ESAN		3	3
UNIVERSIDAD DE LIMA	3		3
UNIVERSIDAD CATÓLICA SAN PABLO	2		2
UNIVERSIDAD PRIVADA SAN IGNACIO DE LOYOLA		2	2
UNIVERSIDAD ALAS PERUANAS	2		2
UNIVERSIDAD NACIONAL DEL CALLAO		2	2
TOTAL	485	460	945

Cuadro 7A**VIF de todas variables independientes para modelos con variable dependiente binaria.**

Variable	VIF	1/VIF
Edad en años	61.92	0.02
Cuadrado de la edad	60.44	0.02
Pertenece al Área OCDE 5	5.16	0.19
Pertenece al Área OCDE 2	4.16	0.24
Pertenece al Área OCDE 3	4.00	0.25
Universidad pública	3.80	0.26
Universidad privada asociativa	3.34	0.30
Pertenece al Área OCDE 1	2.90	0.34
Horas totales de trabajo en universidad	2.87	0.35
Pertenece al Área OCDE 4	2.58	0.39
Tiempo completo o dedicación exclusiva	2.05	0.49
Horas haciendo investigación en universidad	1.69	0.59
Horas de trabajo fuera de universidad	1.48	0.68
Horas enseñando en universidad	1.47	0.68
Ingreso por la universidad	1.44	0.69
Pertenece a Lima o Callao	1.30	0.77
Desempeña cargo administrativo	1.27	0.79
Número de personas dependientes económicamente	1.21	0.82
Tiene Doctorado del tipo 1	1.21	0.83
Tiene perspectiva positiva sobre su universidad	1.19	0.84
Es mujer	1.19	0.84
Trabaja en otra universidad	1.15	0.87
Es casado / conviviente	1.14	0.87
Tiene Doctorado del tipo 2	1.13	0.88
Cree que investigación de su universidad contribuye a sociedad	1.13	0.88
Dispone de ambiente de trabajo en universidad	1.10	0.91
Sabe inglés	1.07	0.93
Estudió fuera de Perú	1.02	0.98

Gráficos 8A

Gráfico de la densidad de probabilidad estimada con función kernel el modelo Probit.

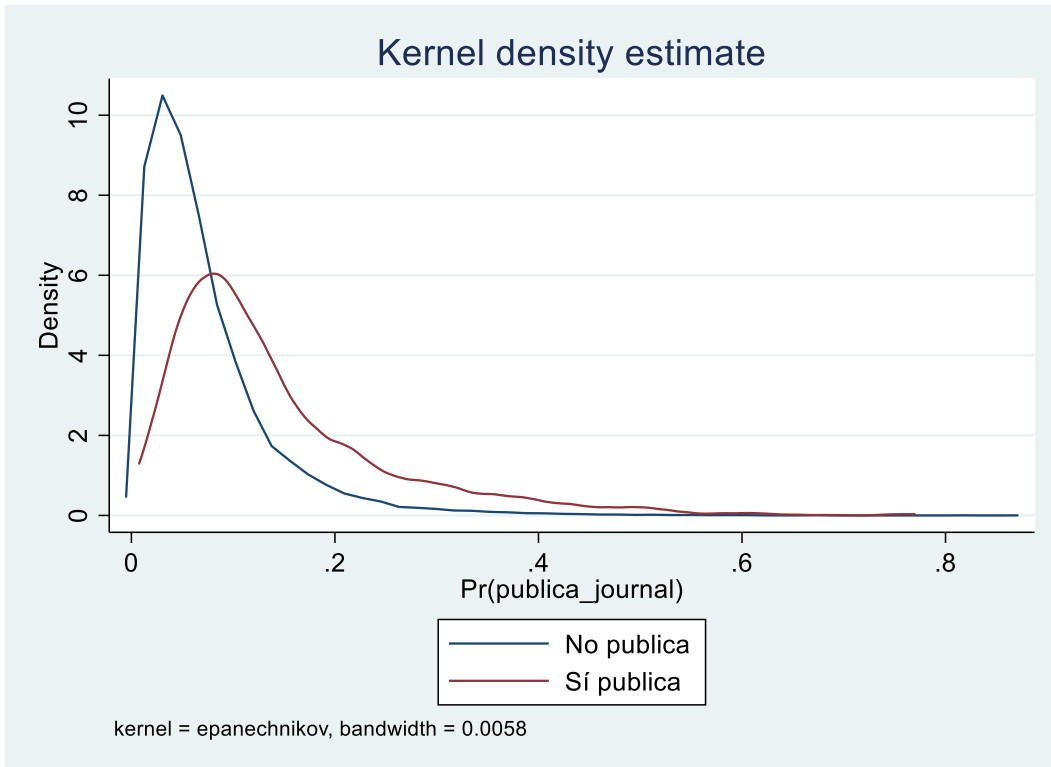
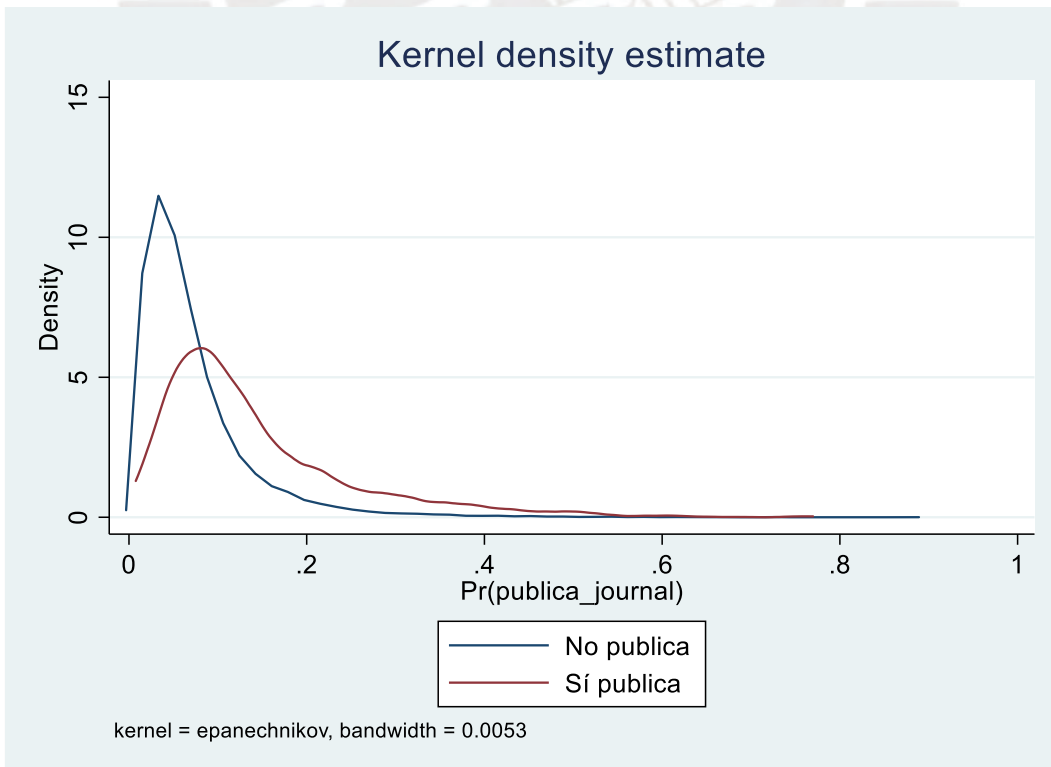


Gráfico de la densidad de probabilidad estimada con función kernel el modelo Logit.



Cuadro 9A**VIF de todas variables independientes para modelos con variable dependiente de conteo.**

Variable	VIF	1/VIF
Edad en años	90.63	0.011034
Cuadrado de la edad	86.73	0.01153
Horas totales de trabajo en universidad	4.96	0.201534
Universidad pública	4.17	0.239818
Pertenece al Área OCDE 2	2.91	0.343845
Pertenece al Área OCDE 5	2.9	0.345112
Tiene perspectiva positiva sobre su universidad	2.67	0.37504
Horas haciendo investigación en universidad	2.57	0.388595
Pertenece al Área OCDE 1	2.52	0.396876
Pertenece al Área OCDE 3	2.51	0.398275
Tiempo completo o dedicación exclusiva	2.31	0.433785
Universidad privada asociativa	2.24	0.446483
Pertenece al Área OCDE 4	2.14	0.468143
Horas enseñando en universidad	1.98	0.505665
Desempeña cargo administrativo	1.94	0.516048
Tiene Doctorado del tipo 1	1.89	0.529238
Lima/Callao vs otras regiones	1.83	0.546625
Horas de trabajo fuera de universidad	1.8	0.556781
Ingreso monetario por la universidad	1.76	0.566715
Es casado / conviviente	1.72	0.580317
Número de personas dependientes económicamente	1.68	0.594539
Cree que investigación de su universidad contribuye a sociedad	1.68	0.595059
Es mujer	1.65	0.607564
Número de profesores	1.64	0.608135
Dispone de ambiente de trabajo en universidad	1.56	0.639303
Tiene Doctorado del tipo 2	1.53	0.654187
Trabaja en otra universidad	1.39	0.719827
Sabe inglés	1.38	0.725528
Universidad tiene convenios de investigación	1.34	0.746454
Financiamiento total	1.27	0.789553
Estudio fuera de Peru	1.26	0.795174
Universidad tiene plan estratégico	1.24	0.805782