



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Riesgo de corrupción municipal en Colombia y su relación con la inversión en primera infancia: un análisis con herramientas de aprendizaje de máquinas y minería de texto

Juan Sebastian Numpaque Cano

Universidad Nacional de Colombia
Facultad de Ciencias Económicas, Escuela de Economía
Bogotá, Colombia
2022

Riesgo de corrupción municipal en Colombia y su relación con la inversión en primera infancia: un análisis con herramientas de aprendizaje de máquinas y minería de texto

Juan Sebastian Numpaque Cano

Trabajo de investigación presentado como requisito parcial para optar al título de:
Magister en Ciencias Económicas

Director:

Ph.D. Jorge Iván Bula Escobar

Línea de Investigación:

Economía del Desarrollo

Universidad Nacional de Colombia
Facultad de Ciencias Económicas, Escuela de Economía
Bogotá, Colombia
2022

A Dios y a la vida.

*A mis padres y a mi hermano, que me apoyan
en cada paso del camino.*

*A mis amigos y profesores, que llenan cada
instante de felicidad y aprendizajes.*

*A todas las personas que han dejado una
huella en mi corazón y que me han enseñado
a amar, a creer y a vivir.*

*“Vive como si fueras a morir mañana. Aprende
como si fueras a vivir para siempre”.*

Mahatma Gandhi

Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, excepto donde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando se han presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones de texto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.



Juan Sebastian Numpaque Cano

7 de agosto de 2022

Agradecimientos

Quizás este trabajo parezca un poco técnico, pero su origen no lo es tanto. Es en realidad el resultado de algo muy sencillo, pero muy poderoso: conversar.

Conversaciones con amigos, familia, profesores y hasta desconocidos fueron las que germinaron y moldearon este trabajo. Un perfecto ejemplo es el del profesor Jorge Iván Bula Escobar, profesor titular de la Escuela de Economía, quien protagoniza muy merecidamente estos agradecimientos. Siempre que me encontré estancado, bastó reunirme con él y conversar media hora para poder avanzar. Le agradezco enormemente por su buena disposición, su humanidad y su gran conocimiento de la academia, virtudes que acompañaron sin falta su constante apoyo y retroalimentación.

También agradezco mucho a Manuel Giovanni Reina Salgado, coordinador del equipo de primera infancia, infancia, adolescencia y familias del Departamento Nacional de Planeación (DNP), por todas sus enseñanzas y su orientación en torno a los temas de primera infancia. Es una dicha para Colombia tener funcionarios públicos tan apasionados e inteligentes como Manuel. Tanto me marcó su fe en la primera infancia que me inspiró a enfocar mi tesis en este tema, que hace un año era desconocido para mí.

Por último, agradezco a la profesora Sonia Esperanza Monroy Varela y al profesor Hernando Díaz Morales, profesores titulares de la Facultad de Ingeniería, por brindarme las herramientas necesarias para finalizar esta odisea. Les estaré siempre agradecido por inculcar en mí la pasión por la investigación y por toda la orientación y apoyo que me han brindado durante mi vida académica.

Resumen

Riesgo de corrupción municipal en Colombia y su relación con la inversión en primera infancia: un análisis con herramientas de aprendizaje de máquinas y minería de texto

¿Cuál es el efecto de la corrupción sobre la inversión en primera infancia en Colombia?
¿Qué mecanismos podrían adoptarse en el sistema de contratación pública para favorecer dicha inversión? El presente trabajo busca responder a estas preguntas mediante un análisis de la base de datos de la contratación pública de Colombia. Para ello, se calcula un índice de riesgo de corrupción para los municipios de Colombia y se construye un modelo para estimar su efecto sobre la proporción de recursos ejecutados para la primera infancia. Los resultados indican que los municipios con mayor riesgo de corrupción tienden a invertir una menor proporción de recursos para la primera infancia, lo que se convierte en una barrera para el desarrollo del país en el largo plazo.

Palabras clave: Gran corrupción, riesgo de corrupción, primera infancia, aprendizaje de máquinas, minería de texto.

Abstract

Municipal corruption risk in Colombia and its relationship with early childhood investment: an analysis with machine learning and text mining tools

What is the effect of corruption on investment in early childhood in Colombia? What mechanisms could be adopted in the public procurement system to increase such investment? This paper addresses these questions through an analysis of the Colombian public procurement database. To this end, a corruption risk index is calculated for Colombian municipalities and a model is constructed to estimate its effect on the proportion of resources executed for early childhood. The results indicate that municipalities with a higher risk of corruption tend to invest a lower proportion of resources for early childhood, which becomes a barrier to the long-term development of the country.

Keywords: Large-scale corruption, corruption risk, early childhood, machine learning, text mining.

Contenido

	Pág.
1. Revisión de literatura.....	3
1.1 ¿La corrupción engrasa o frena las ruedas del crecimiento?.....	3
1.2 La corrupción en la contratación pública	6
1.3 ¿Cómo medir la corrupción?	7
1.4 Inversión en primera infancia, desarrollo y crecimiento económico	8
2. Diseño metodológico.....	11
2.1 Índice de Riesgo de Corrupción (IRC).....	12
2.2 Identificación de proyectos de primera infancia	15
2.3 Estimación del modelo	18
3. Resultados.....	19
3.1 Contratación pública para la primera infancia.....	22
3.2 Efecto del riesgo de corrupción sobre la contratación pública para la primera infancia	23
4. Discusión.....	27
5. Conclusiones y recomendaciones.....	31
6. Bibliografía	33

Lista de figuras

	Pág.
Figura 1: Retorno por dólar invertido en distintas edades.....	8
Figura 2: Palabras y grupos de palabras utilizadas para identificar contrataciones para la primera infancia.....	17
Figura 3: Índice de Riesgo de Corrupción (IRC) para municipios con y sin hechos de corrupción reportados en la prensa entre 2016 y 2020.....	20
Figura 4: Distribución del Índice de Riesgo de Corrupción (IRC) en el territorio nacional.	21
Figura 5: Distribución del porcentaje del valor adjudicado para la primera infancia con respecto al total adjudicado por cada municipio.	22
Figura 6: Efecto esperado de un aumento de un punto en el IRC sobre la inversión en primera infancia en los municipios de Colombia (en puntos porcentuales).	25

Lista de tablas

	Pág.
Tabla 1: Canales de transmisión del efecto de la corrupción sobre el crecimiento económico, según la literatura.....	3
Tabla 2: Indicadores de riesgo de corrupción utilizados.....	13
Tabla 3: Documentos de política de otros sectores (distintos de primera infancia).	16
Tabla 4: Valor promedio de los indicadores de riesgo de corrupción para municipios con y sin hechos de corrupción reportados en la prensa entre 2016 y 2020.	19
Tabla 5: Resultados de la estimación del modelo lineal por MCG.....	24

Introducción

La corrupción es una problemática multidimensional de gran relevancia en el contexto social, político y económico. Dado que abarca muchos temas y actores, el concepto de corrupción es muy amplio y no existe una definición única para ella, pero una que se presenta con frecuencia es la de Transparencia-Internacional (2009), que la define como el abuso de poder utilizado para el beneficio propio y en detrimento del interés colectivo. Este trabajo se enmarca en dicha definición, aunque se enfoca en una comprensión específica de la corrupción: la gran corrupción (o corrupción a gran escala) en la contratación pública a nivel municipal. Así, el análisis aquí presentado no comprende la corrupción en el marco de sistemas legales, éticos o culturales y se centra en el ámbito político y económico.

Por sus características y su modo de operar, la corrupción se ha constituido como una de las principales barreras para alcanzar un desarrollo económico próspero y equitativo (Perry & Saavedra, 2004). Comprender sus dinámicas, sus causas y consecuencias resulta ser un problema de interés general, especialmente para los gobiernos, encargados de liderar la lucha contra la corrupción. Sin embargo, los retos en torno a la comprensión de la corrupción no son pocos: su análisis incorpora un fuerte componente cualitativo y la información cuantitativa disponible es escasa y cuenta con bajos niveles de calidad y desagregación, retos que este trabajo busca superar al incorporar el análisis de los 11,9 millones de procesos de contratación registrados en el Sistema Electrónico para la Contratación Pública (SECOP). Además, las dinámicas de la corrupción son complejas, por lo que deben analizarse a nivel agregado y no únicamente a nivel individual.

Como resultado de un análisis con modelos paramétricos y no paramétricos, el presente trabajo encuentra que la gran corrupción en Colombia reduce el crecimiento de largo plazo al reducir los recursos que las entidades territoriales destinan a la primera infancia, pues se evidencia una menor proporción de contrataciones destinadas a la primera infancia a medida que aumenta el riesgo de corrupción. Este efecto, que afecta en mayor medida a

algunos municipios, se constituye como una barrera para el desarrollo de largo plazo de Colombia, como se argumenta a lo largo del documento.

El documento se divide en cinco secciones: la primera incluye una revisión de literatura sobre la corrupción, su efecto sobre el crecimiento económico, las dinámicas que tiene en el sistema de contratación pública y las bases y desafíos metodológicos para medir la gran corrupción. También se presenta la relevancia de la inversión en primera infancia, que actúa como canal de transmisión entre la corrupción y el crecimiento económico. La segunda sección contiene el diseño metodológico, incluyendo los indicadores de riesgo de corrupción calculados, las técnicas de minería de texto utilizadas para estimar la inversión en primera infancia y los modelos construidos para estimar el efecto de la corrupción sobre la inversión en primera infancia. La tercera parte expone los resultados obtenidos; la cuarta, una discusión en torno a ellos, y la quinta, las conclusiones y recomendaciones de política que se derivan del análisis realizado.

1. Revisión de literatura

1.1 ¿La corrupción engrasa o frena las ruedas del crecimiento?

En el día a día se habla de la corrupción como algo indeseable y con numerosos efectos negativos, sin embargo, en la literatura académica se encuentran dos hipótesis contrarias al respecto: una postula que la corrupción “frena las ruedas” del crecimiento económico (*sand the wheels*), y otra, que las “engrasa” (*grease the wheels*). Distintos autores brindan evidencia para una u otra hipótesis, lo cual depende, en general, del canal de transmisión que analizan. Los canales identificados en la literatura se presentan en la **Tabla 1**.

Tabla 1: Canales de transmisión del efecto de la corrupción sobre el crecimiento económico, según la literatura.

Canal	Hipótesis	Justificación	Autores
Inversión extranjera	Frena las ruedas	Los autores analizan que los inversores aversos al riesgo prefieren invertir en países con baja corrupción. Este análisis supone que el inversionista es independiente de la corrupción, por lo que no cobija casos como el de Odebrecht, en los que el inversionista hace parte del entramado de corrupción.	(Gründler & Potrafke, 2019; Mo, 2001)
Inflación	Frena las ruedas	Por un lado, los costos de evasión y recaudo de impuestos y la informalidad son mayores en países con alta corrupción, lo que incentiva al Estado a aumentar el “impuesto inflación” para no ver afectados sus ingresos. Por otro lado, la corrupción reduce los ingresos y aumenta los gastos del Estado, lo que genera déficit y un aumento en las tasas de interés	(Al-Marhubi, 2000; Gründler & Potrafke, 2019)

		que encarece los costos de producción. Estos se trasladan al consumidor.	
Políticas de competencia	Frena las ruedas	Las empresas sobornan a los legisladores para que las políticas de competencia desfavorezcan la entrada de nuevos competidores, lo cual reduce la innovación.	(Aghion & Howitt, 2009)
Inestabilidad política	Frena las ruedas	La corrupción beneficia más (en ingresos y oportunidades) a clases altas y aumenta la desigualdad. Junto al efecto negativo sobre la confianza en las instituciones, hay incentivos para realizar acciones ilegales o violentas como reacción a la desigualdad. Esto genera incertidumbre sobre los derechos de propiedad y dificultades en la operación de las cadenas productivas, lo que reduce la inversión y la productividad.	(Mo, 2001)
Capital social	Frena las ruedas	La corrupción afecta negativamente la confianza interpersonal y, con ella, el capital social de enlace (<i>bridging social capital</i>), que impacta positivamente el crecimiento económico. Además, en un entorno corrupto se favorece el fortalecimiento del capital social vinculante (<i>bonding social capital</i>), que afecta negativamente el crecimiento.	(Sudarsky & García-Díaz, 2020; Muringani, Fitjar, & Rodríguez-Pose, 2021)
Capital humano	Frena las ruedas	La corrupción tiene efectos sobre los determinantes del capital humano, tanto en factores observables (estado de salud, asistencia escolar y tasas de deserción escolar), como en factores inobservables, lo que disminuye la productividad.	(Abdulla, 2021; Mo, 2001)
Gasto público	Frena las ruedas	La corrupción disminuye la eficiencia del gasto público y, con ella, su impacto en el crecimiento económico.	(Fiorino, Galli, & Petrarca, 2012)
Alta regulación	Engrasa las ruedas	Los sobornos facilitan el ingreso de nuevas empresas en economías altamente reguladas, lo que favorece la innovación.	(Dreher & Gassebner, 2013)

Burocracia	Engrasa las ruedas	En países con largos procesos burocráticos, la corrupción puede permitir que proyectos favorables para el desarrollo se lleven a cabo con más rapidez, especialmente si hay afinidad política.	(Kopas, Urpelainen, & York, 2021; Méon & Weill, 2010)
------------	--------------------	--	---

Fuente: Elaboración propia.

Aunque es bastante sugestiva la evidencia sobre cada uno de los efectos presentados en la **Tabla 1**, hay varios debates abiertos. Por ejemplo, Hoa (2019) encuentra efectos tanto negativos (en eficiencia del gasto) como positivos (en capacidad de conseguir recursos) de la corrupción sobre el capital humano en provincias de Vietnam, y Mayo (2016) defiende que los sobornos no son un mecanismo eficiente para adaptarse a ambientes con alta debilidad institucional. En cuanto a los efectos a nivel de empresa, Wellalage & Thrikawala (2021) identifican que el efecto negativo de la corrupción sobre la innovación en las empresas afecta a las empresas pequeñas y medianas, pero no a las grandes, efecto que atribuyen a que los sobornos impactan una mayor proporción de los recursos en las empresas pequeñas, y a que las grandes realizan sobornos "estratégicos" y no sobornos "forzosos" (como en el caso de las pequeñas), por lo que pueden incluso beneficiarse de la corrupción.

Si bien parece haber canales de transmisión a favor de cada hipótesis, todavía no hay consenso sobre cuáles prevalecen. Los estudios más recientes favorecen la hipótesis de que la corrupción, a nivel agregado, "frena las ruedas" del crecimiento (Nur-tegin & Jakee, 2020; Uberti, 2021). Sin embargo, estos no contradicen trabajos previos que exponen casos específicos en que esto podría no cumplirse, como el presentado por (Jiang & Nie, 2014), quienes argumentan que la corrupción favoreció el milagro del crecimiento económico chino por haber ayudado a las empresas privadas a aumentar sus utilidades, al permitirles operar con mayor flexibilidad en un país altamente regulado, efecto no observado por los autores en las empresas de propiedad del Estado. En síntesis, hay un debate abierto en torno al efecto de la corrupción sobre el crecimiento y sus canales de transmisión, y aunque son numerosos los canales ya identificados, los efectos de la corrupción son tan complejos y heterogéneos entre países y sectores que no logran presentar todo el panorama del problema.

1.2 La corrupción en la contratación pública

Uno de los sistemas donde se ve reflejada la corrupción a mayor escala es el sistema de contratación pública, que compromete cerca del 12% del PIB en los países de la OCDE (OCDE, 2016). En este sistema, se suele pensar la corrupción como transacciones en que contratistas específicos ofrecen sobornos a funcionarios públicos para obtener contratos (Popa, 2019), sin embargo, Salcedo-Albaran & Garay-Salamanca (2018) argumentan que esta es una visión simplificada del problema que limita la comprensión del fenómeno de la corrupción y la posibilidad de desarrollar herramientas efectivas para combatirla. Por ello, argumentan que se debe distinguir la corrupción entre dos tipos: la gran corrupción, "que concentra su acción en la contratación estatal, en la expedición de normas y que resulta ejercida por altos funcionarios del Estado", y la pequeña corrupción, que involucra transacciones ejercidas principalmente por funcionarios de nivel medio-bajo o bajo.

Desde la escuela neoinstitucional, Boehm & Lambsdorff (2009) también han resaltado la necesidad de distinguir entre gran y pequeña corrupción, señalando además que el estudio de la corrupción debe contemplar los vínculos existentes entre diferentes grupos (políticos, económicos, familiares, entre otros), en particular en lo que respecta a corrupción en el sistema de contratación pública. En el marco de la investigación aquí propuesta, cobra mucho sentido incluir la formación de vínculos perdurables entre organizaciones públicas y privadas en el análisis, pues estos podrían aumentar el riesgo de corrupción, especialmente cuando las organizaciones públicas involucradas son de alto nivel. Allí aparece el concepto de macrocorrupción, referente a la corrupción que ha permeado instituciones, perdurado en el tiempo y que afecta el gasto público de manera sistémica y no únicamente en transacciones aisladas (Salcedo-Albaran & Garay-Salamanca, 2018).

Así, el marco conceptual es diferente cuando se busca analizar los fenómenos de corrupción a gran escala: ya no se indaga por acontecimientos esporádicos, sino por características de un sistema complejo con interacciones repetidas entre actores muy numerosos y diversos. En particular, Salcedo-Albaran & Garay-Salamanca (2018) sugieren adoptar este enfoque en el caso latinoamericano, pues argumentan que en países como Colombia se han establecido redes de macrocorrupción y de cooptación institucional que impiden caracterizar de manera adecuada la corrupción cuando esta se analiza a pequeña escala.

1.3 ¿Cómo medir la corrupción?

Una de las mayores dificultades que enfrenta la lucha contra la corrupción es su identificación. Al ser un fenómeno multidimensional con un fuerte componente de ilegalidad, resulta muy difícil rastrear la corrupción en las bases de datos existentes, lo que deriva en indicadores muy sesgados o tan poco desagregados que pierden utilidad práctica para la formulación de políticas.

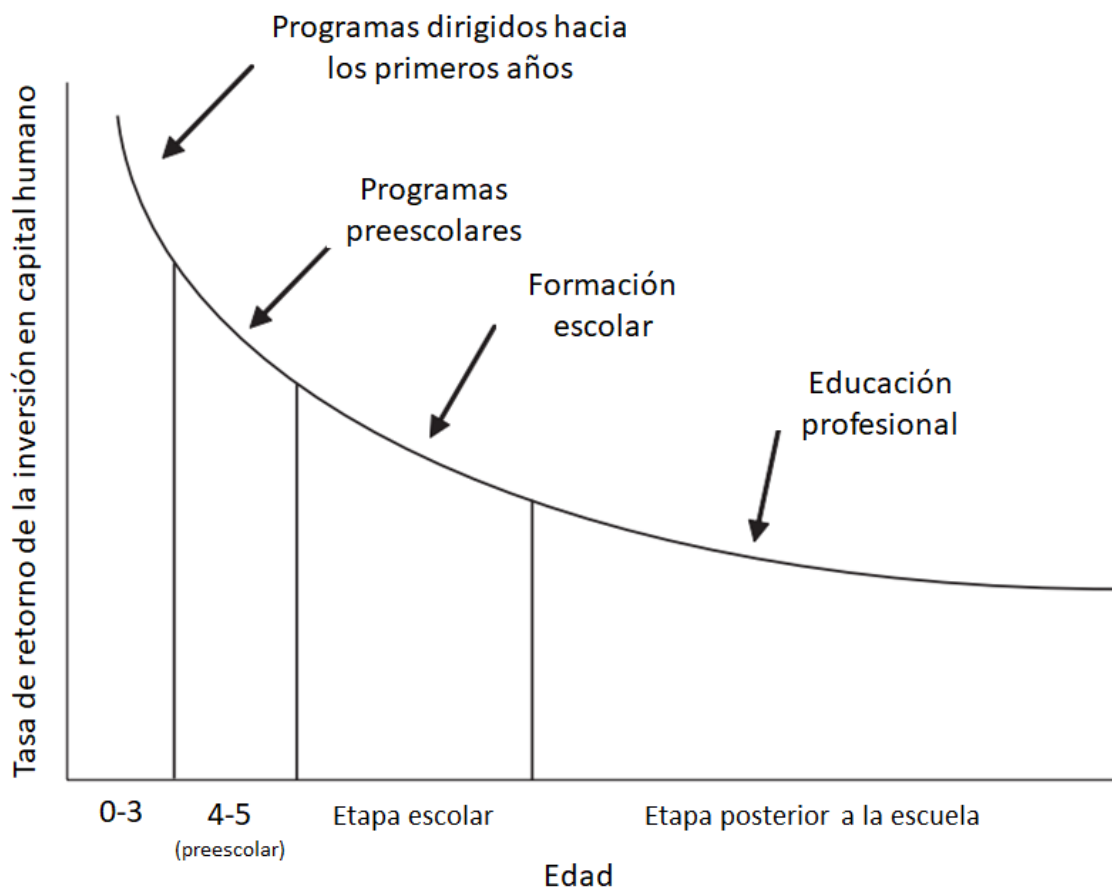
Ante esta dificultad, Fazekas, Tóth, & King (2016) sugieren estimar no la corrupción sino el riesgo de corrupción, y proponen para ello un índice con datos de contratación pública. Este se compone de indicadores objetivos que permiten estimar la “corrupción legal”, nominada así por los autores dado que se recurre a herramientas legales del sistema de contratación pública para beneficiar una red cerrada de actores. Así, el índice propuesto por Fazekas, Tóth, & King (2016) se relaciona con la macrocorrupción, en el cual las entidades se han visto permeadas a alto nivel por la corrupción, como lo describen Salcedo-Albaran & Garay-Salamanca (2018). Esto se logra al reflejar patrones que no son visibles en las contrataciones analizadas de manera individual, pues indaga por patrones a nivel de entidad que se presentan de manera sistémica y recurrente y que solo constituyen una alerta cuando se analizan las contrataciones de forma agregada, como problemas de competencia o atipicidades recurrentes en los procesos de contratación.

En el caso de Colombia, Zuleta, Ospina & Caro (2019) realizaron un ejercicio muy similar, con base en la metodología del Centro de Investigación en Política Pública del Instituto Mexicano para la Competitividad (IMCO, 2018), que es esencialmente idéntica a la de Fazekas, Tóth & King (2016). Gracias al trabajo de Zuleta, Ospina & Caro (2019), se cuenta con una serie de indicadores adaptados a las particularidades del sistema de contratación pública colombiano que permiten construir un índice de riesgo de corrupción. Si bien lo usual es calcular estos indicadores a nivel de entidad, como lo hicieron Fazekas, Tóth & King (2016) para Hungría, IMCO (2018) para México y Zuleta, Ospina & Caro (2019) para Colombia; se encuentra en la literatura que su cálculo también puede ser a nivel municipal, como lo realizan Wachs, Yasseri, Lengyel, & Kertész (2019). Este enfoque resulta muy pertinente para el análisis en el caso colombiano, pues las redes de macrocorrupción suelen ser interinstitucionales, llegando a ser intermunicipales e incluso internacionales (Salcedo-Albaran & Garay-Salamanca, 2018).

1.4 Inversión en primera infancia, desarrollo y crecimiento económico

Heckman (2008) expone la importancia de la inversión en primera infancia, determinante del desarrollo humano y el crecimiento económico. En su estudio, Heckman (2008) encuentra que las brechas de habilidades aparecen en los primeros años de vida, y que invertir en la niñez y en sus familias en etapas tempranas reduce estas brechas, producto de la mejora en la salud y en las habilidades cognitivas y socioemocionales de los niños y niñas menos favorecidos. También concluye que las intervenciones de política en los primeros años generan mayores retornos que en cursos de vida posteriores, como se presenta en la Figura 1.

Figura 1: Retorno por dólar invertido en distintas edades.



Fuente: Heckman (2008).

Por su parte, Grantham-McGregor, y otros (2007) caracterizan el efecto negativo de los riesgos a los que está expuesta la niñez (pobreza, desnutrición, malas condiciones de salud y entornos no estimulantes) sobre su desarrollo cognitivo, motor y socioemocional, lo que se convierte en un pilar fundamental para la transmisión intergeneracional de la pobreza. En conclusión, un factor que afecta negativamente las inversiones en la primera infancia es un factor que reduce el desarrollo y el crecimiento económico de un país en el largo plazo. De allí que sea relevante evaluar el efecto de la corrupción sobre la inversión en primera infancia, pues esta podría constituir un canal de transmisión no muy explorado en la literatura y que también podría explicar el efecto de la corrupción sobre el crecimiento económico.

2. Diseño metodológico

El objetivo principal de este trabajo es relacionar el riesgo de corrupción municipal con el crecimiento económico de Colombia. Esto se plantea mediante una hipótesis con visión de largo plazo, bajo la cual se formula que la primera infancia sufre un perjuicio a causa de la corrupción debido al desvío de recursos hacia otros sectores por parte de las entidades territoriales. Para evaluar dicha hipótesis, se supone una relación subyacente lineal y se plantea un modelo de regresión lineal múltiple de la forma:

$$Y_{pi} = \beta I_{RC} + \theta X + \varepsilon$$

donde Y_{pi} equivale al porcentaje de recursos contratados para la primera infancia, β es un escalar, I_{RC} es el índice (agregado) de riesgo de corrupción, X es una matriz que contiene las variables de control que se presentarán más adelante, θ es un vector de parámetros desconocidos y ε es el término de error.

Para la elección de las variables de control, se tomó como referencia el trabajo de Godfrey (2011), quien señala que la inversión en primera infancia está determinada por factores de política y factores contextuales. En factores de política se incluyen los criterios bajo los cuales se espera que ocurra la focalización de recursos destinados para la primera infancia, como lo son el Índice de Pobreza Multidimensional (IPM) de fuente censal y la brecha de atención en educación inicial, que es determinante entre las atenciones priorizadas para promover el desarrollo integral de la primera infancia. Esta focalización también se captura al incluir el presupuesto departamental destinado a las oficinas regionales del ICBF, principales ejecutoras de los programas de primera infancia. Sobre esta línea, se incluye el indicador de trabajo informal, pues parte de la financiación del ICBF proviene de los aportes parafiscales que realizan los trabajadores formales. También se incluyó el porcentaje de niños en hogares pobres multidimensionales (pobreza en la

niñez de fuente censal), pues recoge varias de las problemáticas que pueden motivar a la administración municipal a aumentar la proporción del gasto en primera infancia.

En factores contextuales se incluye el valor agregado per cápita como indicador de la disponibilidad de recursos del municipio, al igual que los puntajes del municipio en movilización y ejecución de recursos, dos dimensiones evaluadas en la Medición de Desempeño Municipal de 2020 (DNP, 2021). Como indicador de la capacidad institucional para planear, gestionar, priorizar y ejecutar apropiadamente los recursos, también se incluye el puntaje obtenido por la alcaldía del municipio en la Medición del Desempeño Institucional (DAFP, 2021). Godfrey (2011) también resalta el conocimiento de los beneficios de invertir en primera infancia entre los factores contextuales, por lo que se tomaron dos indicadores: el bajo logro educativo, como proxy del conocimiento adquirido en la escuela, y la tasa de analfabetismo en mayores de 15 años, indicador del sector educativo que representa muy bien el postulado de Heckman (2008): es más económico invertir en formación en las etapas tempranas.

También se incluyó el porcentaje de abstención municipal en las elecciones de gobernador de 2019 como proxy de la veeduría ciudadana. Esto implica un supuesto que se plantea desde el nivel individual: si una persona considera relevante su rol en la comunidad, tendrá una mayor propensión a votar y una mayor propensión a realizar veeduría ciudadana. De allí que se espere una alta correlación entre la participación y la veeduría, la cual es relevante dado que puede evitar que el riesgo de corrupción se materialice en actos efectivos de corrupción. Finalmente, se incluyeron el porcentaje de procesos sin objeto contractual detallado, el porcentaje de contratos sin cuantía, y una variable de puntos de acceso a internet fijo per cápita, pues estas variables pueden explicar posibles subregistros en los procesos de contratación del municipio, en general. También se controla por la población del municipio y por el total de procesos de contratación registrados.

2.1 Índice de Riesgo de Corrupción (IRC)

Para construir un índice de riesgo de corrupción, se procesaron las bases de datos de SECOP I y SECOP II y se realizó el cálculo de los indicadores presentados en la Tabla 2. La mayoría de ellos son tomados de Zuleta, Ospina & Caro (2019).

Tabla 2: Indicadores de riesgo de corrupción utilizados.

Símbolo	Indicador	Valores	Dimensión	Definición
N_1	Número de procesos de contratación por contratista.	1 a ∞	Competencia	Número de contratos adjudicados, dividido entre el número de contratistas únicos a los que ha sido adjudicado al menos un proceso en el municipio.
C_2	Porcentaje del monto adjudicado mediante modalidades no competitivas.	0% a 100%	Competencia	Monto total adjudicado mediante modalidades no competitivas, dividido entre el total de recursos adjudicados.
C_3	Porcentaje de procesos adjudicados mediante modalidades no competitivas.	0% a 100%	Competencia	Número de procesos adjudicados mediante modalidades no competitivas, dividido entre el total de procesos adjudicados.
C_4	Índice Herfindahl-Hirschman.	0% a 100%	Competencia	Suma de los cuadrados de las cuotas de mercado que posee cada contratista. La cuota corresponde al porcentaje de recursos adjudicados al contratista, con respecto al monto total adjudicado por el municipio.
C_5	Índice C_4 .	0% a 100%	Competencia	Suma de las cuatro mayores cuotas de mercado que poseen los contratistas del municipio. La cuota corresponde al porcentaje de recursos adjudicados al contratista, con respecto al monto total adjudicado por el municipio.
C_6	Porcentaje de procesos sin objeto	0% a 100%	Transparencia	Procesos de contratación cuyo objeto contractual no se registra en el SECOP o se registra

	contractual detallado.			utilizando 5 o menos palabras, dividido entre el número total de procesos.
C_7	Porcentaje de procesos de contratación con valor inverosímil o sin reportar.	0% a 100%	Transparencia	Procesos de contratación cuya cuantía es menor que \$10.000 pesos colombianos o no se reporta, dividido entre el número total de procesos.
C_8	Porcentaje del monto adjudicado a contratistas sancionados.	0% a 100%	Anomalías	Monto total adjudicado a contratistas sancionados dividido entre el monto total adjudicado por el municipio.

Fuente: Elaboración propia a partir de Zuleta, Ospina & Caro (2019).

Finalmente, al indicador de “número de procesos de contratación por contratista” (N_1) se aplicó la siguiente transformación:

$$C_1 = 1 - \frac{1}{N_1}$$

Con ello, todos los indicadores resultaron con un rango de 0% a 100%, donde un porcentaje más alto indica un mayor riesgo de corrupción. Gracias a esta característica, se utiliza una agregación por iguales ponderaciones para obtener un Índice de Riesgo de Corrupción de la siguiente manera:

$$I_{RC} = \frac{1}{8} \sum_{i=1}^8 C_i$$

No se incluyeron indicadores sobre procesos con un solo oferente o cuya duración representa atipicidades en el proceso de contratación, pues esta información estaba presente únicamente en el SECOP II, base de datos poco representativa de las contrataciones que realizan las entidades territoriales que llevaría a obtener indicadores sesgados.

2.2 Identificación de proyectos de primera infancia

La base de datos del SECOP no contiene una variable que permita distinguir las contrataciones que se realizan para la primera infancia, por lo cual se recurrió a identificar estos proyectos utilizando minería de texto. Para ello, se tomó como referencia la metodología de DNP (2018). Esta toma como insumo dos grupos de documentos:

- Documentos de política que hacen referencia a la primera infancia.
- Documentos de política que hacen referencia a otros temas.

El grupo de documentos (en adelante, corpus) de primera infancia corresponde a:

- Ley 1804 de 2016: Por la cual se establece la política de Estado para el Desarrollo Integral de la Primera Infancia de Cero a Siempre y se dictan otras disposiciones.
- Documento CONPES social 109: Política pública de primera infancia “Colombia por la primera infancia”.
- Atención Integral: Prosperidad para la Primera Infancia (estrategia “de Cero a Siempre”).
- Documento CONPES 3887: Distribución de los recursos del sistema general de participaciones para la atención integral de la primera infancia, vigencia 2016, y orientaciones para su inversión.
- Documento CONPES 3887: Distribución de los recursos del sistema general de participaciones para la atención integral de la primera infancia, vigencia 2017.

Por otro lado, el corpus de “otros temas” estaba conformado por 33 documentos CONPES de 11 temas diferentes de política pública, resumidos en la Tabla 3. Se excluyeron documentos de los sectores de educación, recreación y deporte, dado que muchos de los términos relevantes para dichos sectores podían ser también claves para los temas de primera infancia.

Tabla 3: Documentos de política de otros sectores (distintos de primera infancia).

Sector o temática	Documentos Conpes		
Transporte	3916	3900	3963
Cultura	3659	3803	3783
Vivienda	3897	3859	3848
Agua potable y saneamiento	3798	3780	3715
Agricultura	3811	3763	3675
Paz	3867	3850	3932
Ambiente	3716	3700	3697
Salud	3456	3415	3337
Minas y energía	3873	3510	3347
Telecomunicaciones	3898	3854	3769
Comercio, industria y turismo	3866	3771	3709

Fuente: Elaboración propia.

A todos los documentos se les realizó un procesamiento estándar: lectura automática de los textos, transformación del texto a minúsculas, eliminación de números y signos no alfanuméricos, eliminación de *stopwords* (conectores, preposiciones y palabras que no agregan significado al texto) y una homologación de palabras similares en significado (por ejemplo: “participar”, “participado” y “participamos”, se convierten en “participar”). Para ejemplificar este procedimiento, puede pensarse en una cadena de texto como “Prestar servicios en 5 municipios para la atención de la primera infancia”, que quedaría “Prestar servicio municipio atención primera infancia”. Finalmente, se utilizó el modelo de bolsa de palabras (*bag of words - BoW*) para obtener una representación numérica de los textos, pues dicho modelo permite obtener una matriz cuyas columnas corresponden a todas las palabras que aparecen en todos los documentos, y cuyas filas corresponden a las páginas de los 38 documentos analizados. En esta matriz *BoW* se representa el número de veces que aparece cada palabra en cada página.

Para determinar las palabras más representativas de primera infancia (P) frente a los otros temas (O), se aplica el siguiente procedimiento:

- Se separa la matriz *BoW* en dos: la correspondiente a las páginas de documentos de primera infancia (BoW_p) y la correspondiente a páginas de otros temas (BoW_o).
- Aplicando análisis de componentes principales, se obtiene el espacio base BoW_o (otros temas), con correspondiente matriz de proyección V . Se puede interpretar este espacio como el de “documentos de política pública”.

- A la matriz BoW_p (primera infancia) se le resta la componente de BoW_p que se proyecta en la base de BoW_o (otros documentos). Esto corresponde a calcular $X = BoW_p - (BoW_p * V)V^T$. El valor promedio de cada columna de la matriz resultante (X) se puede interpretar como una medida de qué tan útil es la palabra correspondiente a dicha columna para distinguir un documento de primera infancia de un documento de política sobre otro tema.

Luego, se tomaron las 500 palabras con mayor puntuación y se etiquetaron manualmente aquellas que efectivamente resultan relevantes para identificar proyectos de primera infancia. Dado que se encontraron palabras como “educación” e “inicial” que no son representativas por sí solas de la primera infancia, se construyeron grupos de palabras clave para la identificación de la inversión en primera infancia. Estas fueron: “primera infancia”, “cero a siempre”, “cero a cinco”, “cero a seis”, “atención integral de la infancia”, “atención integral de la niñez”, “atenciones priorizadas”, “educación inicial”, “ruta integral de atenciones” y “entorno protector”. Las palabras y grupos de palabras resultantes de dicho ejercicio se presentan en la Figura 2.

Figura 2: Palabras y grupos de palabras utilizadas para identificar contrataciones para la primera infancia.



Fuente: Elaboración propia.

Así, se lograron clasificar como “contrataciones relacionadas con primera infancia” todas aquellas para las cuales se identificó en el objeto contractual alguno de los grupos de palabras descritos anteriormente, o 3 o más palabras clave de primera infancia.

2.3 Estimación del modelo

Conceptualmente, el modelo tiene alto riesgo de heterocedasticidad, pues el efecto del riesgo de corrupción sobre la inversión en primera infancia tiene mayor variabilidad en los municipios con mayor riesgo de corrupción. Esto se debe a que, en los municipios con alto riesgo de corrupción, será amplia la brecha de corrupción entre los municipios en los que se materialice el riesgo y los municipios en los que no. Sin embargo, esta brecha será menor entre municipios con bajo riesgo de corrupción. Por ello, el modelo lineal se estimó utilizando mínimos cuadrados generalizados (MCG) (Aitken, 1935), que permite obtener mejores estimaciones en caso de heterocedasticidad.

En primera instancia, se formuló el modelo lineal incluyendo todas las variables de control. Luego se utilizó el factor de inflación de la varianza (*FIV*) para eliminar variables que generaran problemas de colinealidad (se presume colinealidad con $FIV \geq 5$). Después se estimó un modelo en el que se eliminaron todas las variables no significativas. Para contrastar los anteriores, se utilizó el algoritmo de regresión paso a paso (*stepwise*) con eliminación bidireccional para obtener un modelo con un mejor balance entre número de variables y el nivel de ajuste. Finalmente, se utilizó la técnica de bosques aleatorios causales (*casual random forest*) para validar que los resultados fueran robustos al adoptar un enfoque no paramétrico, que no supone una relación lineal subyacente (Athey, Tibshirani, & Wager, 2019). En todos los casos se identificaron y eliminaron los municipios con valores atípicos en el IRC o en la inversión en primera infancia (Y_{pi}) utilizando el método de 1,5 veces el rango intercuartílico.

3. Resultados

Se utilizaron 11,9 millones de registros (procesos de contratación) para calcular los indicadores de interés para todos los municipios de Colombia. Se excluyeron 7 municipios que registraron menos de 50 procesos para evitar estimaciones poco confiables. Para validar que los indicadores capturaran información relevante para medir la corrupción, se tomó la base de datos del Monitor Ciudadano (Transparencia por Colombia, 2021). Esta base registra los hechos de corrupción reportados en la prensa entre 2016 y 2020 y relaciona el municipio de ocurrencia del hecho. Así, se evaluó si existían diferencias entre el valor promedio de cada indicador entre aquellos municipios con y sin hechos de corrupción registrados, obteniendo los resultados que se presentan en la **Tabla 4**. Allí se observa que, en promedio, los municipios con hechos de corrupción tienen un peor desempeño en todos los indicadores calculados, diferencia que es significativa en la mitad de los indicadores, lo cual es consistente con el resultado esperado.

Tabla 4: Valor promedio de los indicadores de riesgo de corrupción para municipios con y sin hechos de corrupción reportados en la prensa entre 2016 y 2020.

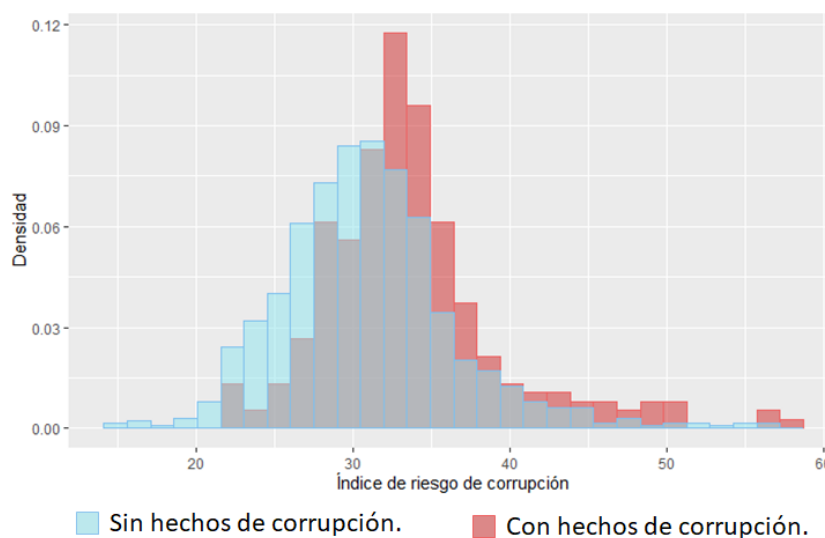
Indicador	Municipios con hechos de corrupción	Municipios sin hechos de corrupción	Diferencia estadísticamente significativa (95%)
Número de procesos de contratación por contratista.	3,43	3,15	Sí
Porcentaje del monto adjudicado mediante modalidades no competitivas.	75,02	66,35	Sí
Porcentaje de procesos adjudicados mediante modalidades no competitivas.	80,50	72,20	Sí
Índice Herfindahl-Hirschman.	8,76	7,40	No
Índice C4.	33,33	32,70	No

Porcentaje de procesos sin objeto contractual detallado.	2,84	2,26	No
Porcentaje de procesos de contratación con valor inverosímil o sin reportar.	0,30	0,23	Sí
Porcentaje del monto adjudicado a contratistas sancionados.	0,16	0,12	No

Fuente: Cálculos propios a partir de SECOP I y SECOP II. Hechos de corrupción tomados de Monitor Ciudadano (Transparencia por Colombia, 2021). Para la significancia estadística se usaron las pruebas t de Student y U de Mann-Whitney con un nivel de confianza del 95%.

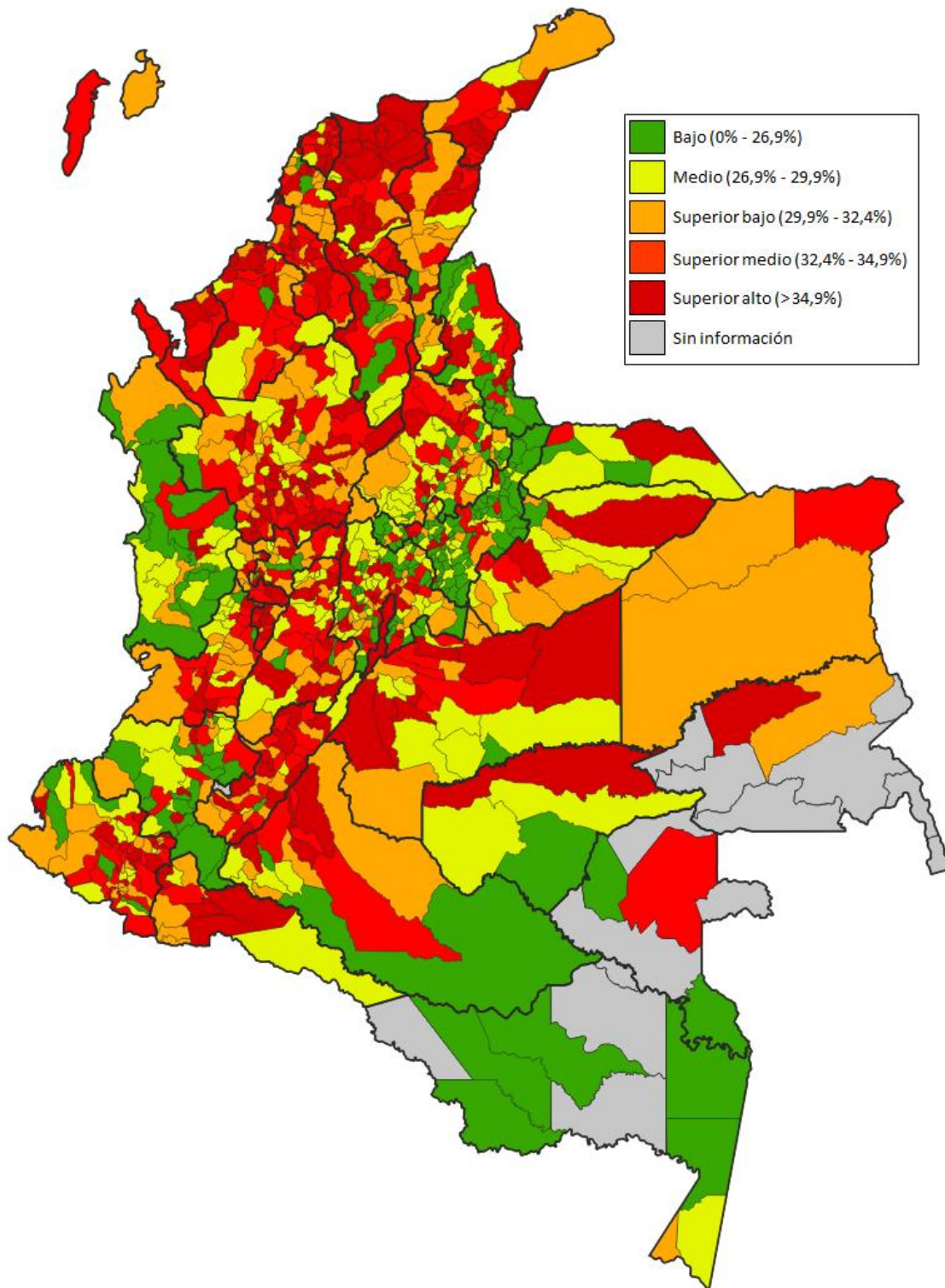
Realizada esta validación, se procedió a realizar la agregación de los 8 indicadores en un único índice. Dado el orden de magnitud de los distintos indicadores, algunos influyen más en el valor final del índice agregado, sin embargo, se decidió no aplicar ninguna corrección dado que los efectos observables a mayor escala son, precisamente, los relacionados con el tipo de corrupción de interés para el presente trabajo. El valor del Índice de Riesgo de Corrupción (IRC) resultante, para municipios con y sin hechos de corrupción, se presenta en la Figura 3, y su distribución en el territorio nacional se presenta en la Figura 4. El valor promedio del IRC fue de 33,7 y 30,9 en municipios con y sin riesgo de corrupción, diferencia estadísticamente significativa.

Figura 3: Índice de Riesgo de Corrupción (IRC) para municipios con y sin hechos de corrupción reportados en la prensa entre 2016 y 2020.



Fuente: Cálculos propios a partir de SECOP I y SECOP II. Hechos de corrupción tomados de Monitor Ciudadano (Transparencia por Colombia, 2021).

Figura 4: Distribución del Índice de Riesgo de Corrupción (IRC) en el territorio nacional.

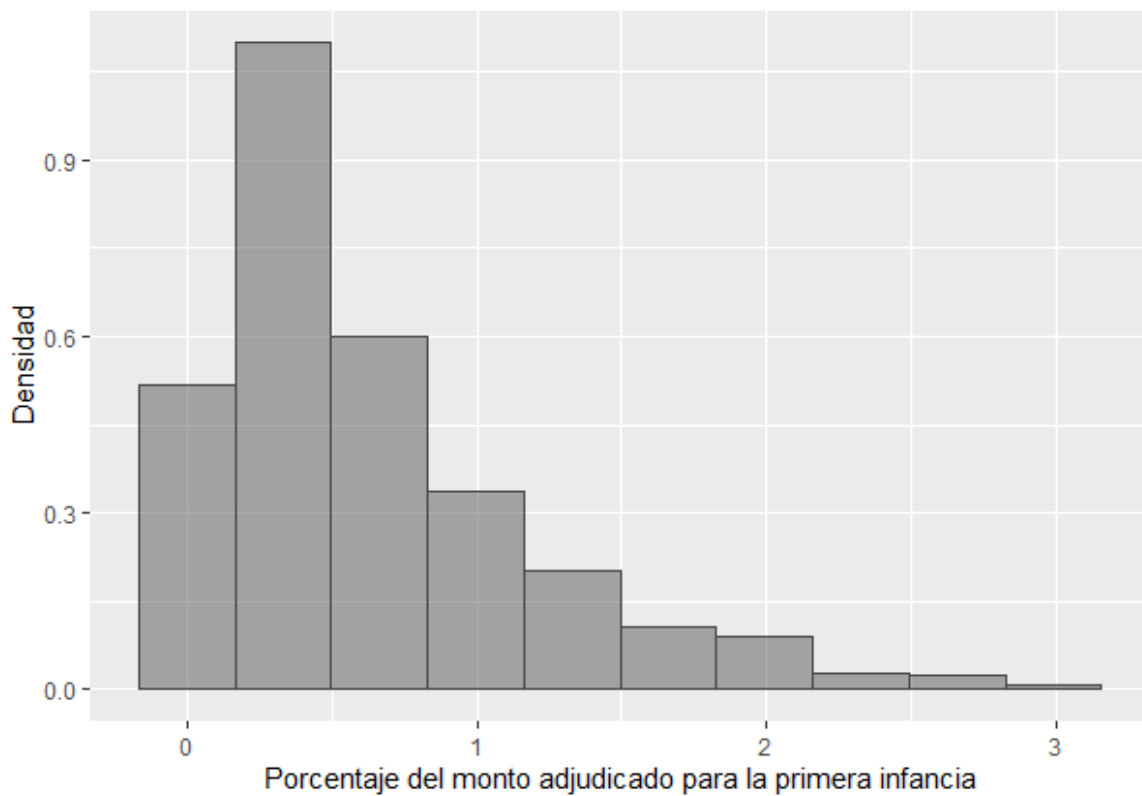


Fuente: Cálculos propios a partir de SECOP I y SECOP II.

3.1 Contratación pública para la primera infancia

Se identificaron 134.081 procesos de contratación relacionados con la primera infancia, correspondientes al 1,12% del total de procesos. A la luz de este indicador, el 95% de los municipios de Colombia ejecutan entre el 0,01% y el 3,8% de sus recursos en contrataciones destinadas a la primera infancia. La distribución de esta variable a nivel municipal se presenta en la Figura 5.

Figura 5: Distribución del porcentaje del valor adjudicado para la primera infancia con respecto al total adjudicado por cada municipio.



Fuente: Cálculos propios a partir de SECOP I y SECOP II. Nota: Incluye 1066 municipios (se excluyen atípicos).

3.2 Efecto del riesgo de corrupción sobre la contratación pública para la primera infancia

Los resultados de la estimación por MCG se presentan en la Tabla 5. En línea con lo esperado, el IRC reduce la proporción del gasto destinada a la primera infancia. En todos los casos, agregar el IRC al modelo aumenta notablemente el R^2 ajustado. También se encuentra que la proporción del gasto destinado a la primera infancia es mayor en los departamentos en los que el ICBF tiene un mayor presupuesto, lo que puede deberse a una mayor capacidad institucional para la gestión y ejecución de proyectos para la primera infancia en dichos departamentos. En cambio, el presupuesto ajustado por población (per cápita) disminuye dicha proporción, lo que puede deberse a la interacción con la variable de población, pues una mayor población puede crear más incentivos políticos a invertir en temas populares como la primera infancia, pero a su vez (por definición) disminuye el presupuesto per cápita.

También resulta interesante el efecto del trabajo informal, pues puede pensarse que el hecho de financiar al ICBF con aportes parafiscales actúa como un incentivo para que dicha entidad destine menos recursos (o haga un menor esfuerzo por garantizar su ejecución) en los departamentos con mayor informalidad. Sin embargo, este efecto no es concluyente, dado que puede existir una fuerte correlación (desde el punto de vista conceptual, pues desde el estadístico no se evidencia) entre el trabajo informal, el riesgo de corrupción, la débil capacidad institucional y la baja ejecución de recursos, lo que hace que este efecto sea más difícil de identificar. Los demás efectos son consistentes con las hipótesis planteadas al inicio de la sección III.

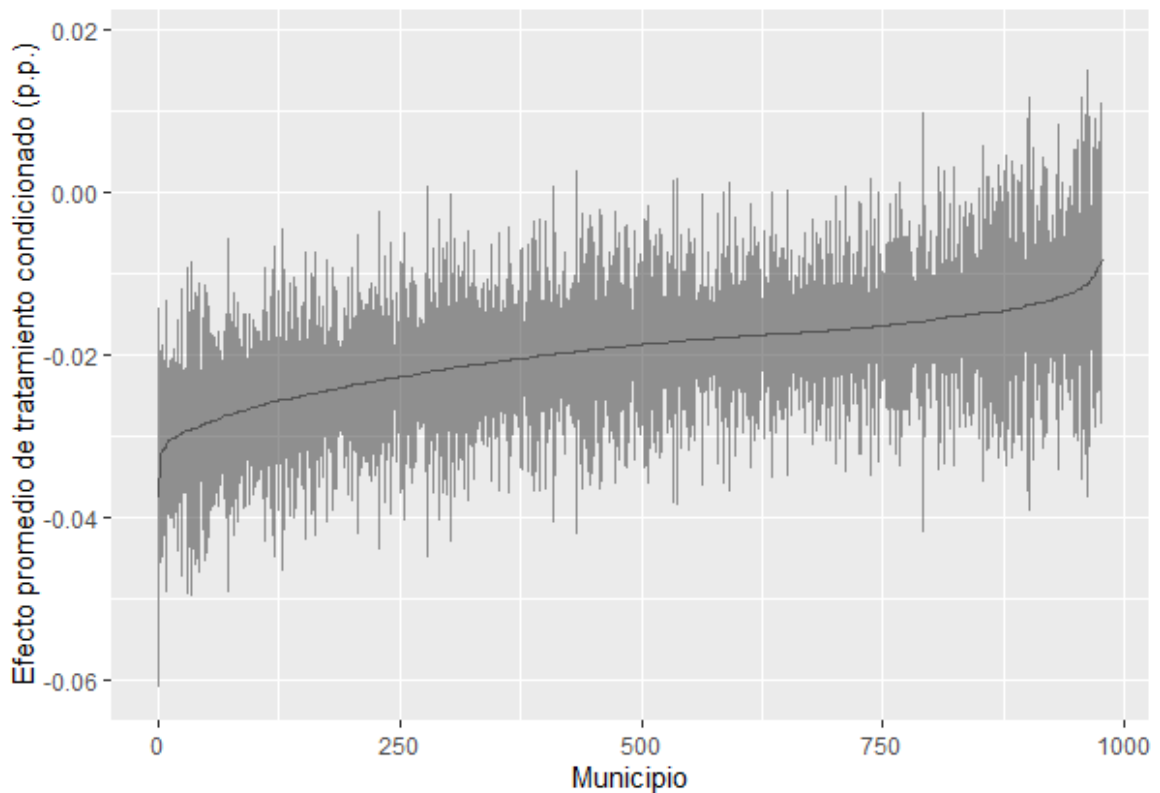
Tabla 5: Resultados de la estimación del modelo lineal por MCG.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IRC		-0.019*** (0.004)		-0.018*** (0.004)		-0.018*** (0.003)		-0.018*** (0.003)
Presupuesto ICBF dpto.	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Presupuesto ICBF per cápita	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Trabajo informal	-0.009** (0.004)	-0.009** (0.004)	-0.009** (0.004)	-0.009** (0.004)	-0.012*** (0.003)	-0.011*** (0.003)	-0.014*** (0.003)	-0.013*** (0.003)
Analfabetismo	0.004 (0.003)	0.007** (0.003)	0.004 (0.003)	0.006* (0.003)	0.005* (0.003)	0.007** (0.003)	0.004 (0.003)	0.006* (0.003)
log(Población en miles)	0.056** (0.025)	0.061** (0.024)	0.058** (0.025)	0.064*** (0.024)	0.026 (0.016)	0.053*** (0.017)	0.045*** (0.013)	0.074*** (0.014)
IPM	0.000 (0.006)	-0.004 (0.006)	0.006*** (0.002)	0.005** (0.002)			0.005*** (0.002)	0.003** (0.002)
Bajo logro educativo	-0.003 (0.003)	-0.005* (0.003)	-0.002 (0.002)	-0.003 (0.002)	-0.004* (0.002)	-0.005** (0.002)		
% NNA en hogares IPM	0.006 (0.006)	0.010 (0.006)			0.006*** (0.002)	0.005*** (0.002)		
log(Total procesos)	-0.047* (0.027)	-0.013 (0.028)	-0.047* (0.027)	-0.014 (0.028)				
Movilización de recursos	0.002* (0.001)	0.001 (0.001)	0.002** (0.001)	0.001 (0.001)				
Ejecución de recursos	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)				
% contratos sin cuantía	-0.012 (0.040)	-0.013 (0.039)	-0.012 (0.040)	-0.013 (0.039)				
% contratos sin objeto	0.000 (0.003)	0.004 (0.003)	0.000 (0.003)	0.003 (0.003)				
Abstención en elecciones	-0.001 (0.002)	-0.002 (0.002)	-0.001 (0.002)	-0.001 (0.002)				
% NN en educación inicial	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)				
Desempeño institucional	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)				
Accesos fijos a internet	0.002 (0.004)	0.002 (0.004)	0.002 (0.004)	0.002 (0.004)				
Valor agregado per cápita	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)				
Intercepto	1.353*** (0.381)	1.744*** (0.384)	1.330*** (0.381)	1.690*** (0.383)	1.454*** (0.250)	1.922*** (0.260)	1.372*** (0.245)	1.787*** (0.255)
Observaciones	977	977	977	977	977	977	977	977
R ²	0.110	0.133	0.109	0.131	0.109	0.130	0.100	0.125
R ² ajustado	0.092	0.115	0.092	0.114	0.099	0.122	0.093	0.118

Errores estándar entre paréntesis. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$. El modelo (2) incluye todas las variables, el (4) excluye % NNA en hogares pobres multidimensionales para evitar colinealidad, el (6) incluye las escogidas con el algoritmo de regresión paso a paso (*stepwise*) y el (8) excluye las variables no significativas en (4). Los modelos (1), (3), (5) y (7) tienen una formulación idéntica a los anteriores, respectivamente, pero se excluye el IRC. Fuente: Elaboración propia.

Los resultados del modelo de bosques aleatorios causales indican que los resultados anteriores son robustos. Con dicho modelo se estimó que, en promedio (y *ceteris paribus*), un punto en el IRC reduce la inversión en primera infancia en 0,021 puntos porcentuales (error estándar = 0,003 p.p.). Este efecto se estima como $E \left[\frac{Cov[I_{RC}, Y_{pi}|X]}{Var[Y_{pi}|X]} \right]$ de manera agregada, y como $\frac{Cov[I_{RC}, Y_{pi}|X]}{Var[Y_{pi}|X]}$ de manera individual, para estimar el efecto esperado en cada municipio. Los resultados de las estimaciones en los municipios, ordenados de mayor a menor efecto (en magnitud absoluta) se presentan en la Figura 6. Para el 88,9% de los municipios, el efecto estimado es estadísticamente diferente de cero.

Figura 6: Efecto esperado de un aumento de un punto en el IRC sobre la inversión en primera infancia en los municipios de Colombia (en puntos porcentuales).



Fuente: Elaboración propia. Nota: Intervalos de confianza del 95%.

4. Discusión

En este trabajo se utilizó información del sistema de contratación pública de Colombia para indagar sobre el efecto causal de la centralidad de los contratistas en la red de contratación pública sobre el riesgo de corrupción a nivel municipal. Hasta donde el autor tiene conocimiento, este es el primer trabajo en:

- Estimar índices de corrupción municipal en la contratación pública de Colombia.
- Utilizar herramientas de minería de texto para estimar los recursos contratados para un sector específico (en este caso, primera infancia).
- Identificar los determinantes de la inversión municipal en primera infancia en Colombia.
- Estimar el efecto del riesgo de corrupción sobre la inversión municipal para la primera infancia, que afecta a su vez el desarrollo de largo plazo de los países.

Los resultados obtenidos son completamente consistentes con los hallazgos e hipótesis de Salcedo-Albaran & Garay-Salamanca (2018) y Wachs, Yasseri, Lengyel, & Kertész (2019). En particular, se soporta la hipótesis de que hay patrones en el estudio de la corrupción que solo pueden visibilizarse a nivel macro, y no analizando transacciones de manera individual. Estos resultados sirven como base para la toma de decisiones y el diseño de estrategias en la lucha contra la corrupción que se planteen a nivel macro, centradas en analizar la red de contratación como un todo.

Los hallazgos presentados también complementan el trabajo de Fiorino, Galli & Petrarca (2012). Ellas encuentran que la corrupción reduce la efectividad del gasto público, impactando negativamente el crecimiento económico, conclusión muy similar a la de Meier & Stiglitz (2000), quienes argumentan la corrupción actúa como un impuesto a la inversión, con fuertes implicaciones en la persistencia de la pobreza y la desigualdad. Si bien este

trabajo no aborda los temas de crecimiento, pobreza y desigualdad, sí evidencia que, para el caso de la primera infancia, la gran corrupción no solo deriva en menor eficiencia del gasto, sino también en una menor proporción del gasto invertida en el sector, y este menor nivel de recursos puede alimentar la transmisión intergeneracional de la pobreza y frenar la movilidad social, lo que reforzaría los postulados de Meier & Stiglitz (2000). Igualmente, estos resultados son consistentes con lo planteado por Perry & Saavedra (2004), quienes argumentan que los países que controlan mejor la corrupción tienden a crecer a tasas mayores en el mediano y largo plazo.

Inevitablemente, hay que reconocer distintas limitaciones que tiene el enfoque aquí presentado. Los indicadores de riesgo de corrupción no capturan información intermunicipal ni interdepartamental, y en muchos casos las redes de macrocorrupción se generan a esta escala. Además, algunos indicadores son sensibles a fenómenos que no necesariamente implican riesgo de corrupción, como la adjudicación de contratos a un mismo contratista debida a la contratación de productos o servicios especializados, o a la experiencia misma del contratista en el municipio. Igualmente, el IRC se validó únicamente con la información reportada por Monitor Ciudadano, que se limita a hechos de corrupción mapeados y denunciados entre 2016 y 2020. Tampoco es posible garantizar que el método de agregación utilizado para construir el IRC sea óptimo, por lo que podrían explorarse otras técnicas en un trabajo futuro.

Finalmente, es clave resaltar que la evidencia únicamente sugiere (no asegura) la existencia de un efecto causal del riesgo de corrupción sobre la inversión en primera infancia. El sustento de la relación causal es principalmente conceptual, pues en todos los modelos construidos puede existir un sesgo de variable omitida. A pesar de haber incluido las variables de control identificadas en la literatura, más otras conceptualmente pertinentes, es posible que haya variables adicionales que expliquen la proporción de recursos destinados a la primera infancia, y el valor del R^2 ajustado en los modelos de regresión es un indicio de ello. Igualmente, una menor inversión en primera infancia puede crear en el largo plazo condiciones sociales favorables para la corrupción, lo que podría causar un problema de endogeneidad en el modelo.

A pesar de estas limitaciones, es justo y necesario expresar que los resultados obtenidos son muy valiosos para la formulación de políticas públicas. Este enfoque basado en datos

públicos y con una perspectiva macro brinda nueva evidencia que soporta la hipótesis de que la corrupción es un fenómeno estructural, que puede abordarse con instrumentos de política y que los efectos negativos que tiene sobre el crecimiento son mayores que los positivos. Los resultados, consistentes con la perspectiva teórica de *sand the wheels*, visibilizan la pertinencia de incluir en los sistemas de contratación pública una serie de “alertas” para identificar anomalías no solo a nivel de proceso (como se ha venido abordando), sino también a nivel de entidad, de municipio e incluso intermunicipal. Así, este trabajo sirve como orientación para el diseño de herramientas tecnológicas que apoyen la lucha contra la corrupción y la materialización del riesgo de corrupción en patrones que afecten a la primera infancia y, con ella, al desarrollo del país en el largo plazo.

5. Conclusiones y recomendaciones

- El estudio de la corrupción en Colombia debe utilizar información agregada, pues existen dinámicas que solo pueden visibilizarse a nivel macro, especialmente cuando se han desarrollado estructuras corruptas que permean las instituciones más grandes de la administración. En este caso, se analizó el agregado de las instituciones a nivel municipal, pero analizar el nivel departamental, regional o nacional puede derivar en información de gran valor para la lucha contra la corrupción.
- La corrupción reduce la proporción de los recursos que se invierte para la primera infancia, lo cual brinda evidencia para la hipótesis de que la corrupción "frena las ruedas" del crecimiento. Si bien el escenario ideal consiste en disminuir la corrupción, se puede plantear un camino alternativo (mas no excluyente) en armonía con la perspectiva de que la corrupción "engrasa las ruedas" del crecimiento, por ejemplo, ideando alternativas para evitar que los intereses de actores corruptos vayan en contravía de los de la primera infancia, incluso cuando el sea difícil evitar la ocurrencia de transacciones corruptas. Para ello, es necesario identificar si esta menor proporción de recursos para la primera infancia se relaciona con una mayor proporción de recursos en otros sectores específicos, ejercicio que puede realizarse con la misma metodología aquí presentada. Si los resultados arrojan, por ejemplo, que el sector de infraestructura resulta favorecido, se puede fomentar la construcción de Centros de Desarrollo Integral, o si se identifica el sector vivienda, se puede favorecer la construcción de viviendas que actúen como entornos protectores para los niños y niñas. Aunque estas soluciones no reducen la pérdida de recursos debida a la corrupción, sí permiten mitigar su efecto negativo sobre el desarrollo de largo plazo al evitar que se reduzcan las

inversiones en primera infancia, curso de vida durante el cual los retornos a la inversión son mayores.

- El efecto del riesgo de corrupción sobre la inversión en primera infancia es muy heterogéneo a lo largo de los distintos municipios de Colombia. Se recomienda estudiar los municipios en los que este efecto es menor o no significativo, indagando por características que puedan orientar el diseño de políticas para evitar la materialización del riesgo de corrupción en el sector de la primera infancia e incluso en otros sectores.
- La información de la contratación pública puede apoyar la implementación y seguimiento de las políticas que hay en Colombia para la niñez, como la política De Cero a Siempre y la Política Nacional de Infancia y Adolescencia, pues permite un alto nivel de desagregación, aun con los desafíos en cuanto a subregistro, estructura y calidad de la información que todavía deben superarse.
- Se recomienda indagar con mayor profundidad por el efecto del trabajo informal sobre la financiación de la primera infancia. El modelo construido sugiere un efecto negativo, aunque no concluyente, que podría relacionarse con que el diseño del mercado laboral impide que los trabajadores informales realicen aportes parafiscales al ICBF. Esto podría desincentivar la inversión en primera infancia en municipios con mayor informalidad, pues es dicha inversión la que constituye una “retribución” del aporte que realiza el trabajador formal. Hasta donde el autor tiene conocimiento, este efecto no se ha explorado en la literatura y puede brindar valiosas recomendaciones de política en torno a la financiación del ICBF.

6. Bibliografía

- Abdulla, K. (2021). Corrosive effects of corruption on human capital and aggregate productivity. *Kyklos - International review for social sciences*, 74, 445-462.
- Aghion, P., & Howitt, P. (2009). *The economics of growth*. Cambridge, Massachusetts: Massachusetts Institute of Technology.
- Aitken, A. C. (1935). On Least Squares and Linear Combinations of Observations. *Proceedings of the Royal Society of Edinburgh*, 55, 42-48.
- Al-Marhubi, F. (2000). Corruption and inflation. *Economics Letters*, 66, 199-202.
- Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized Random Forests. *Annals of Statistics*, 47, 2.
- Boehm, F., & Lambsdorff, J. (2009). Corrupción y anticorrupción: una perspectiva neo-institucional. *Revista de Economía Institucional*, 11-21.
- DAFP. (2021). *Resultados Medición del Desempeño Institucional 2020*. unpublished. Obtenido de <https://www.funcionpublica.gov.co/web/mipg/resultados-medicion>
- DNP. (2018). *Análisis de los Planes Nacionales de Desarrollo a través del tiempo*. unpublished. Obtenido de https://colaboracion.dnp.gov.co/CDT/Desarrollo%20Digital/Big%20Data/2018/10_PND_en_el_tiempo/PND_en_el_tiempo_Informe.pdf
- DNP. (2021). *Medición de Desempeño Municipal*. unpublished. Obtenido de <https://www.dnp.gov.co/programas/desarrollo-territorial/Estudios-Territoriales/Indicadores-y-Mediciones/Paginas/desempeno-integral.aspx>
- Dreher, A., & Gassebner, M. (2013). Greasing the wheels? The impact of regulations and corruption on firm entry. *Public Choice*, 155, 413-432.
- Fazekas, M., Tóth, I., & King, L. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research*, 22, 369-397.
- Fiorino, N., Galli, N., & Petrarca, I. (2012). Corruption and Growth: Evidence from the Italian Regions. *European Journal of Government and Economics*, 1.
- Godfrey, E. (2011). *Determinants of public investment in early childhood development within the education sector at national and local levels in Uganda*. Nairobi: Department of Early Childhood Studies.
- Grantham-McGregor, S., Cheung, Y., Cueto, S., Glewwe, P., Richter, L., Strupp, B., & Group, I. C. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet - Child Development*, 60-70.
- Gründler, K., & Potrafke, N. (2019). *Corruption and Economic Growth: New Empirical Evidence*. Tech. rep., Cesifo working papers, Paper No. 7816.
- Heckman, J. (2008). Schools, skills and synapses. *Economic Inquiry*, 46, 289-324.
- Hoa, T. (2019). The effects of corruption on the human capital accumulation process: Evidence from Vietnam. *Economics of Transition and Institutional Change*, 28, 69-88.

- IMCO. (2018). *Índice de Riesgos de Corrupción: El sistema mexicano de contrataciones públicas*. unpublished.
- Jiang, T., & Nie, H. (2014). The stained China miracle: Corruption, regulation, and firm performance. *Economics Letters*, 123, 366-369.
- Kopas, J., Urpelainen, J., & York, E. (2021). *Greasing the Wheels: the Politics of Environmental Clearances in India*. Studies in Comparative International Development.
- Mayo, R. (2016). Does Bribery Grease the Wheels of Economic Growth? *Munich Personal RePEc Archive, MPRA Paper No, 98433*.
- Meier, G., & Stiglitz, J. (2000). *Frontiers of Development Economics*. New York: World Bank and Oxford University Press.
- Méon, P., & Weill, L. (2010). Is Corruption an Efficient Grease? *World Development*, 38, 244-259.
- Mo, P. (2001). Corruption and Economic Growth. *Journal of Comparative Economics*, 29, 66-79.
- Muringani, J., Fitjar, R., & Rodríguez-Pose, A. (2021). Social capital and economic growth in the regions of Europe. *EPA: Economy and Space*, 53, 1412-1434.
- Nur-tegin, K., & Jakee, K. (2020). Does corruption grease or sand the wheels of development? New results based on disaggregated data. *The Quarterly Review of Economics and Finance*, 75, 19-30.
- OCDE. (2016). *Towards Efficient Public Procurement in Colombia: Making the Difference*. Obtenido de <http://www.oecd.org/governance/towards-efficient-public-procurement-in-colombia-9789264252103-en.htm>
- Perry, G., & Saavedra, V. (2004). Corrupción y crecimiento económico. En L. Villar, & D. Álvarez (Edits.), *Lucha integral contra la corrupción en Colombia: reflexiones y propuestas* (págs. 17-46). Fedesarrollo.
- Popa, M. (May de 2019). Uncovering the structure of public procurement transactions: Business and Politics. *Cambridge Core*. Obtenido de <https://www.cambridge.org/core/journals/business-and-politics/article/uncovering-the-structure-of-public-procurement-transactions/9278E0C88F3B11027053B8E1C3FEDEF2>
- Salcedo-Albaran, E., & Garay-Salamanca, L. (2018). *El gran libro de la corrupción en Colombia*. Bogotá: Planeta.
- Sudarsky, J., & García-Díaz, D. (2020). *La brecha entre la Sociedad, la Política y el Estado: La Cuarta Medición del Capital Social de Colombia*. Bogotá: Puntoaparte Bookvertising.
- Transparencia por Colombia. (2021). *Radiografía de la corrupción 2016-2020*. unpublished. Obtenido de <https://www.monitorciudadano.co/hechos-corrupcion/>
- Transparencia-Internacional. (2009). *Guía de lenguaje claro sobre lucha contra la corrupción*, Berlín.
- Uberti, L. (2021). *Corruption and growth: Historical evidence, 1790–2010*. *Journal of Comparative Economics*.
- Wachs, J., Yasseri, T., Lengyel, B., & Kertész, J. (2019). Social capital predicts corruption risk in towns. *Royal Society open science*, 6, 3. Obtenido de <https://royalsocietypublishing.org/doi/10.1098/rsos.182103>
- Wellalage, N., & Thrikawala, S. (2021). Does bribery sand or grease the wheels of firm level innovation: evidence from Latin American countries. *Journal of Evolutionary Economics*, 31, 891-929.
- Zuleta, M., Ospina, S., & Caro, C. (2019). *Índice de riesgo de corrupción en el sistema de compra pública colombiano a partir de una metodología desarrollada por el Instituto Mexicano para la Competitividad*. Bogotá: Banco Interamericano de Desarrollo. Obtenido de https://www.fedesarrollo.org.co/sites/default/files/fedesarrollo_cpbd.pdf