Three essays in Regional Economics: The case of Colombia

Julián Santiago Vásquez-Roldán

Doctoral Thesis

2020

Thesis Supervisor:

Jesús López-Rodríguez, PhD

Universidade da Coruña

Programa oficial de doctorado en *Análisis Económico y Estrategia Empresarial*

UNIVERSIDADE DA CORUÑA

Three essays in Regional Economics: The case of Colombia

DEDICATORIA y AGRADECIMIENTOS

A Dios, pues sus planes siempre son mejores que los míos.

A mi familia, que con su confianza, cariño y amor, me sabe guiar por el sendero de la vida. Su apoyo constante es mi mayor motivación y alimenta mi deseo de superación. Ellos me han enseñado que la vida cuando la vivimos con amor, así no contemos con recursos, está llena de oportunidades y satisfacciones. El corazón es un gran orientador de la inteligencia.

A mis sobrinos Salomé, Isabela y Dylan; de ellos brota mi alegría y estímulo permanente. Mi mayor satisfacción, fuera de verlos crecer felices, es que algún día lean esta tesis y se apasionen como yo del desarrollo regional, de servir con alegría y amor en el sector público.

A la gente de mis queridos municipios de Belmira y Guarne por darme raíces, identidad y unas bases firmes que me han permitido conocer desde la escasez, el amor por los temas del desarrollo regional. Vivir y formarme en esa realidad compleja es lo que motiva cada uno de mis pasos. He emprendido un camino que estoy recorriendo con el corazón para encontrar posibilidades que transformen la realidad económica y social de municipios y regiones.

A la Facultad de Economía de la Universidad Autónoma Latinoamericana y a la Universidad, en cabeza de su Rector José Rodrigo Flórez Ruiz, que es mi alma mater y el inicio de mi camino en la profesión que amo.

Al Grupo Bancolombia, el cual, a pesar de estar en la parte crucial de mi tesis me acogió y acompañó en la ruta que tracé desde lo personal y profesional. Ser Gerente Regional de la Banca Pública fue una gran satisfacción. Hacer parte del equipo de la región Antioquia hizo que mi corazón palpitara de alegría, con más energía, pues allí aprendí el significado de rodearme no solo de grandes profesionales de alto desempeño, sino de un excelente grupo humano. Mi agradecimiento a su Presidente, Juan Carlos Mora Uribe, a mis jefes Luz María Velásquez y Astrid Yepes por su apoyo, cariño y por permitirme ayudar a promover el desarrollo económico sostenible para lograr el bienestar de todos.

Al Instituto para el Desarrollo de Antioquia (IDEA) y al Gobernador Anibal Gaviria Correa por depositar su confianza en mí, por su apoyo, por inculcarme los principios de defensa de la vida y equidad. En especial, agradezco al Gobernador por permitirme entregar los mejores años de mi vida al servicio de Antioquia. Que esta tesis favorezca el impulso y el desarrollo regional de nuestra amada Antioquia.

A la Universidade da Coruña por acoger y acompañar un Colombiano que sueña posible transformar la realidad compleja e histórica que conocemos. Así mismo, agradezco a mi tutor el PhD. Jesús López Rodriguez, quien con ahínco, paciencia y hermandad siempre me enseña y seguirá enseñando. Su tolerancia, capacidad académica y humildad son mi respaldo permanente y un claro ejemplo del papel transformador que cumplen los profesores en la sociedad.

A Diego por su incondicionalidad, a mis amigos de la Universidad de Antioquia y del Banco de la República por su cariño.

A todas esas personas incondicionales que son parte de mi vida y a los que en el día a día conozco, mi dedicatoria y mi compromiso permanente, desde el corazón, con el desarrollo de Antioquia.

DEDICATION AND ACKNOWLEDGMENTS

To God, his plans are always better than mine.

To my family, who with their trust, affection, and love, knows how to guide me on the path of life. Their constant support is my greatest motivation and it fuels my desire to excel. They have taught me that life when we live it with love, even if we do not have resources, is full of opportunities and satisfactions. The heart is a great guider of intelligence.

To my nephews Salomé, Isabela and Dylan; from them springs my joy and permanent encouragement. My greatest satisfaction, apart from watching them grow happily, is that one day they will read this thesis and become passionate about regional development, that they will pursue serving with joy and love in the public sector.

To the people of my beloved municipalities of Belmira and Guarne for giving me roots, identity and firm foundations that have allowed me to know from a lack of resources, the love for regional development issues. Living and training in this complex reality is what motivates each of my steps. I have embarked on a path that I am traveling with my heart to find possibilities that transform the economic and social reality of municipalities and regions.

To the Faculty of Economics of the Universidad Autónoma Latinoamericana and to the University, headed by its President José Rodrigo Flórez Ruiz. This institution is my alma mater and it was the start of my path in the profession I love.

To the Bancolombia Group, which, despite me being in the crucial part of my thesis, welcomed me and accompanied me on the route I traced from a personal and professional point of view. Being a Regional Manager of Public Banking was a great satisfaction. Being part of the Antioquia region team made my heartbeat with joy, with more energy, because there I learned the meaning of surrounding myself not only with great high-performance professionals, but with an excellent human group. My thanks to its President, Juan Carlos Mora Uribe, to my bosses Luz María Velásquez and Astrid Yepes for their support, love and for allowing me to help promote sustainable economic development to achieve the well-being of all.

To the Institute for the Development of Antioquia (IDEA) and to the Governor of Antioquia Anibal Gaviria Correa for placing their trust in me, for their support, for instilling in me the principles of defense of life and equity. I thank the Governor for allowing me to dedicate the best years of my life to the service of Antioquia. May this thesis favor the impulse and regional development of our beloved Antioquia.

To the Universidade da Coruña, for welcoming and accompanying a Colombian who dreams of transforming the complex and historical reality we know. Likewise, I thank my tutor, Jesús López Rodriguez, who with diligence, patience and brotherhood always teaches me and will continue teaching me. His tolerance, academic ability and humility are my permanent support and a clear example of the transformative role that teachers play in society.

To Diego for his unconditionality, and my friends from the University of Antioquia and the Colombian Central Bank – Medellín – for their love.

To all those unconditional people who are part of my life and whom I know every day, my dedication, and my permanent commitment, from the heart, to the development of Antioquia.

Abstract

This doctoral thesis consists of three essays that are framed within the research fields of geographical economics and regional economics.

The first essay of the PhD thesis begins with (Chapter 1) entitled "*Market Potential and Income Disparities across Colombian Departments: Analysis for the period 1990-2015"*. This chapter focuses on the analysis of the role played by market potential in the spatial income structure observed in Colombia over the period 1990-2015. Based on the geographical economics theory we derive the so called nominal wage equation which establishes a relationship between nominal wages and a distance weighted sum of the volume of economic activities in surrounding locations which is usually known as market potential. The estimation of this equation using Colombian departamental data over the period 1990-2015 reveals that market potential plays a crucial role in explanation of the spatial distribution of per capita income. Moreover the results also show that access to markets became more important over the course of the period under analysis which is pointing to a process of spatial concentration of economic activities in Colombia where the agglomeration forces are more important than the dispersion ones. Finally, two important channels that might be affecting the shaping of the spatial distribution of income were devised, physical and human capital.

The second essay of the thesis (Chapter 2) is entitled "*The Effects of Transfers of Royalties to Subnational Governments on Regional Formal Employment in Colombia*"

Based on the aforementioned, we extend the investigation (Chapter 2) discussing the effects of the transfers of royalties to the municipalities on regional formal employment during the period 2002-2018. This chapter evaluates the effects that royalties have on formal employment in Colombia. This chapter allowed to conclude that the General SGR Royalties System allowed to redistribute the amount of royalties and use that important source of revenue, to boost regional and municipal growth, increase equity between regions given the redistribution and decrease poverty rates via formal employment. In addition, we find that the increase in the participation of royalties for the municipalities that received lesser proportions before the reform, formal employment became more dynamic.

The final essay of this thesis (Chapter 3), is entitled *"Labor Market Opportunities for the Demobilized in Colombia"*

In this essay, we address an issue of relevance in the national and international context, since, after a process of internal armed conflict of more than 50 years in Colombia, peace is signed with the oldest and richest guerrilla of the world. This chapter allows us to understand some aspects in common as well as others associated with the labor market that will mean significant efforts in public policy.

The results of this chapter of the thesis suggest that, in Colombia, given the different variables included in the model, the labor market does not generate stimuli for the incorporation of the reinserted population. In addition to this result, the essay concludes that this population is more likely fall into recidivism and engage in informal activities. This situation enables the demobilized population to be close to returning to illegality.

Resumen

Esta tesis doctoral consta de tres ensayos que se enmarcan en los campos de investigación de la economía geográfica y la economía regional. El primer ensayo (capítulo 1) que se titula "*Potencial de mercado y disparidades de ingresos en los departamentos colombianos: Análisis para el período 1990-2015*" se enmarca dentro los modelos de economía geográfica. Los otros dos ensayos titulados "*Los efectos de las transferencias de regalías a los gobiernos subnacionales en el empleo formal regional en Colombia*" (Capítulo 2) y *"Oportunidades en el mercado laboral para los desmovilizados en Colombia"* (Capítulo 3) son propios de las temáticas de análisis del campo de la economía regional en sus aspectos de análisis del mercado laboral.

El análisis de las razones de por qué unos países son ricos y otros pobres es quizás unas de las preguntas más antiguas y fundamentales de la economía. Este mismo debate se podría trasladar a una escala inferior y la pregunta sería similar, ¿por qué a nivel de país las diferentes unidades territoriales que los componen tienen unos niveles de desarrollo dispares? Diferentes dimensiones del crecimiento económico se han investigado para entender variaciones en los niveles de bienestar y crecimiento que afectan a países y a regiones dentro de cada país. En esta amplia literatura, la teoría neoclásica usa la inversión en capital físico y la educación como un elemento importante para entender las diferencias en las tasas de crecimiento entre regiones/países (Solow1956, Mankiw et al., 1992). Entre las diferentes explicaciones, los desarrollos de la teoría de la *Economía Geográfica* nos permiten construir modelos de equilibrio general para explicar no solo las razones de por qué la actividad económica se concentra en el espacio sino también la acumulación endógena de factores de producción (Krugman, 1991, Gallap et al., 1999).

La importancia del marco conceptual y teórico de la Economía Geográfica está íntimamente unida a la definición de distancia y proximidad geográfica. Aunque es posible utilizar la distancia como un factor exógeno, la teoría de la Economía Geográfica establece que otras dimensiones de la proximidad geográfica se pueden entender a través de la inclusión de otras dimensiones económicas como el comercio y el potencial económico y de renta. La literatura de la Economía Geográfica usa diferentes combinaciones de estos factores para calcular el acceso al mercado y el potencial de mercado.

Este primer ensayo de la tesis que se titula "*Potencial de mercado y disparidades de ingresos en los departamentos colombianos: Análisis para el período 1990-2015*" utiliza como marco teórico la teoría de la Economía Geográfica para dar una explicación a las importantes diferencias que existen en los niveles de desarrollo para los departamentos (regiones) en los que se estructura territorialmente la República de Colombia. Consecuentemente usa el concepto de proximidad geográfica entre los distintos departamentos colombianos a través del cálculo de los correspondientes *potenciales de mercado* departamentales para incluir otras dimensiones de la actividad económica y no solo la distancia física entre los distintos departamentos a la hora de explicar los diferenciales en los niveles de desempeño de los distintos departamentos. El potencial de mercado pasa a convertirse en una variable central a la hora de explicar cómo la geografía de segunda naturaleza, es decir, la geografía que depende de la acción humana- en contraposición a la geografía de primera naturaleza, es decir, la presencia de ríos, montañas, dotación de hidrocarburos, etc., - desempeña un papel relevante en la explicación de las disparidades en los niveles de renta per cápita para los departamentos colombianos. Con base en la teoría de la economía geográfica, derivamos la llamada ecuación nominal de salarios que establece para cada localización "i" del

espacio una relación entre los salarios nominales de esa localización "i" y una suma ponderada por la distancia del volumen de actividad económica de las localizaciones colindantes, que generalmente se conoce como potencial de mercado. La estimación de esta ecuación utilizando datos para los departamentos colombianos durante el período 1990-2015 revela que el potencial del mercado juega un papel crucial en la explicación de la distribución espacial de los niveles de renta per cápita en Colombia. Además, los resultados también muestran que el acceso a los mercados se hizo más importante a lo largo del período de análisis, lo que apunta a un proceso de concentración espacial de las actividades económicas en Colombia, donde las fuerzas de aglomeración son más importantes que las fuerzas de dispersión. Finalmente, el capítulo analiza dos canales importantes que podrían estar afectando estructura espacial de los niveles de renta en Colombia como son las dotaciones de capital físico y de capital humano.

Teniendo como base lo anteriormente mencionado, ampliamos la investigación (Capítulo 2) discutiendo los efectos de las transferencias de regalías a los municipios en el empleo formal regional durante el período 2002-2018. El segundo ensayo de la tesis se titula "*Los efectos de las transferencias de regalías a los gobiernos subnacionales en el empleo formal regional en Colombia*" se centra en el análisis del papal que desempeñan las trasferencias de regalías generadas por la explotación de recursos no renovables en la configuración de los niveles de empleo formal a partir de la creación del Sistema General de Regalías (SGR) en contraposición de los resultados del sistema anterior. Al tiempo que se analiza el papel dinamizador que pueden jugar las regalías sobre la economía y como se convierten en un movilizador de mano de obra, se estudia que rol juegan sobre la distribución o configuración de disparidades regionales como también sobre asuntos como el esfuerzo, la transparencia, la responsabilidad y la eficiencia fiscal de los entes territoriales luego de percibir estos recursos. En adición, se consideran los riesgos fiscales que pueden generar fiscales por la poca predictibilidad que puede manejarse en el presupuesto debido a que este tipo de recursos tienden a ser volátiles. La literatura ha estudiado ampliamente estos impactos y ha resaltado la existencia de un riesgo importante que ello conduzca a la llamada "maldición de los recursos" o la constitución de economías de enclave.

Por tanto, el capítulo 2 aprovecha la reforma constitucional realizada por el Gobierno y que entró en vigor en el año 2012 asociada al Sistema General de Regalías y que favoreció la redistribución de los recursos derivados de la compensación por la extracción de recursos no renovables. En este sentido, el capítulo permitió identificar el efecto que tuvo la reforma anteriormente mencionada sobre el empleo formal. En este sentido, autores como Echeverry, Alonso y García (2011), plantean que la reforma que creó el Sistema General de Regalías se fundamentó en cuatro principios: (i) la equidad social y regional, (ii) el ahorro para el futuro, (iii) la competitividad regional, y (iv) el buen gobierno. El Gobierno Nacional reemplazó el esquema anterior por el Sistema General de Regalías con el fin de utilizar esa importante fuente de recursos, para impulsar el crecimiento regional, la equidad entre regiones, disminuir los índices de pobreza y aumentar la competitividad del país. Al respecto Cárdenas (2013) sugiere que el antiguo sistema presentaba bajo impacto y resultados, esto se fundamenta en que: (i) Muchos proyectos fragmentados de bajo impacto social y económico, (ii) Deficiente planeación y ejecución, (iii) Para 2009 ninguna entidad territorial estaba certificada en todas las coberturas, y (iv) Productores: alta dependencia de regalías, baja tributación, bajo desarrollo sostenible. Por todo lo anterior este capítulo para identificar el efecto en empleo formal de las regalías utilizó un estimador de regresión discontinua, aprovechando la forma como se distribuyen las regalías del Fondo de Compensación Regional de acuerdo con la normatividad

vigente, con base en la cual las regalías distribuyen dependiendo si los municipios tienen porcentajes de su población con necesidades básicas insatisfechas, NBI, inferiores o superiores a 35%. La metodología es similar a la anteriormente aplicada en estudios de Angrist y Lavy (2001), van der Klaauw (2002), Hahn, Todd and Van der Klaauw (2001), Ferrez y Finan (2010) y en Corbi, Papaioannou y Surico (2014).

Para Colombia regalías son una importante fuente de ingresos para el desarrollo, con el potencial de financiar diferentes proyectos que, a su vez, cuentan con la capacidad de mejorar los niveles de vida de los habitantes, además de mejorar los niveles de producción en todas las regiones del país. En la Constitución Política colombiana está consignado que el Estado es el propietario del subsuelo y de los recursos naturales no renovables, los cuales le generan a favor una contraprestación económica derivada de su explotación. A partir de la Ley 141 del año 1994 se dio la creación de la Comisión Nacional de Regalías y el Fondo Nacional de Regalías (FNR), que permitieron esta contraprestación de recursos, y a su vez, establecieron las reglas para su liquidación y distribución. La distribución y ejecución de regalías a través del FNR permitió distribuir los dineros obtenidos por el Estado mediante regalías directas, hacia las entidades territoriales que poseían y de donde se explotaban los recursos, y regalías indirectas, a los entes territoriales que hacían viable el transporte de estos, pero sin tener en cuenta el resto de regiones, departamentos, municipios y veredas del país. Esto, sin lugar a dudas, constituía un serio problema de disparidad sobre todo el territorio nacional, puesto que las inversiones que se realizan con los ingresos generados por la explotación de este tipo de recursos deberían beneficiar a toda la población. Adicionalmente, de acuerdo a Cárdenas (2013), este sistema presentaba bajo impacto y resultados, con muchos proyectos fragmentados de bajo impacto social y económico, de deficiente planeación y ejecución, y generador de alta dependencia a las regalías, baja tributación, bajo desarrollo sostenible.

La reforma aprobada por el Gobierno Nacional reemplazó el FNR por el SGR en el año 2011, estaba basada en los principios de equidad social y regional, ahorro para el futuro, competitividad regional, y buen gobierno; considerando, además, la descentralización y autonomía de las entidades territoriales. Esta apuesta buscaba promover una distribución más equitativa, garantizando el uso de los recursos con eficiencia y probidad, con el fin de aportar a la transparencia y el uso efectivo de los recursos de regalías contribuyendo con ello tanto a una mayor equidad entre regiones como también a la disminución de los índices de pobreza y el aumento de la competitividad del país. La metodología usada para identificar el efecto de las regalías sobre el empleo formal regional es una regresión discontinua, aprovechando la forma como se distribuyen las regalías del Fondo de Compensación Regional (FCR o RCF en inglés) de acuerdo con la normatividad vigente, con base en la cual las regalías distribuyen dependiendo si los municipios tienen porcentajes de su población con necesidades básicas insatisfechas (NBI o UBN en inglés) inferiores o superiores a 35%. Se encuentran efectos positivos y significativos de las regalías sobre el empleo formal, principalmente en los sectores de industria; electricidad, gas y agua; actividades inmobiliarias y empresariales; e intermediación financiera. En adición, se encuentra que las regalías provenientes del FCR aumentaron la cantidad de recursos destinados a proyectos de inversión subnacionales sin reducir la cantidad de otras fuentes de financiamiento, en particular, los recursos propios de los municipios. Finalmente, los resultados también muestran que estas regalías no afectaron el esfuerzo fiscal de los municipios, y no afectaron sus ganancias tributarias iniciales.

Este capítulo permitió concluir que el SGR permitió redistribuir del monto de regalías y utilizar esa importante fuente de recursos, para impulsar el crecimiento regional y municipal, la equidad entre regiones dada la redistribución y disminuir los índices de pobreza vía empleo formal. Además, es claro que con el incremento en la participación de las regalías para los municipales que recibían menores proporciones antes de la reforma el empleo formal se dinamizó al tiempo que la reforma aumentó los recursos para entidades territoriales que no los recibían. Asimismo, dados los incipientes niveles de desarrollo de los entes territoriales, en especial los que se encuentran en sexta categoría, los recursos provenientes de regalías se convirtieron en un vehículo de disponibilidad de recursos favorables a la dinámica territorial y en favor del empleo formal. El haber utilizado la ley 1530 de 2012 y el decreto 1073 de 2012, además de la regresión discontinua fundamentada en las necesidades básicas insatisfechas NBI permitió determinar el impacto en empleo formal al encontrar un efecto positivo de las regalías en el mismo. Finalmente, se concluye que la focalización de las regalías es un problema de interés central en el país por la inmensa cuantía de estas (aproximadamente 1,6% del PIB), y porque son una fuente importante de financiación para el desarrollo regional, de forma que su adecuada distribución es fundamental para favorecer el empleo formal y dinamizar el desarrollo de las regiones.

El tercer ensayo titulado "*Oportunidades en el mercado laboral para los desmovilizados en Colombia*" plantea que la discriminación en los mercados laborales ha sido un importante objeto de estudio de las Ciencias Económicas en los últimos años. Los análisis coinciden en señalar que para hablar de discriminación se requiere que las poblaciones objeto de estudio y comparación, sean idénticas en todas las dimensiones que pueden afectar la productividad, a la luz de las consideraciones de los empleadores, (Bertrand y Mullainathan, 2004 y Altonji y Blank, 1999). de acuerdo con el propósito de este trabajo, entre los aspectos utilizados se consideran tres tareas fundamentales: (i) la caracterización de la población reinsertada de acuerdo con la información descrita anteriormente, (ii) se propone un esquema de incentivos de reinsertados pensando en el mercado laboral al que se van a enfrentar, y (iii) se estiman ecuaciones de Mincer con el fin de encontrar los determinantes que mayores efectos tienen sobre la probabilidad de conseguir empleo formal, y sobre los salarios esperados de los la población de interés en distintos escenarios.

Este capitulo final de la tesis analizó empíricamente las oportunidades del mercado laboral para los desmovilizados en Colombia. Este tercer ensayo es titulado "Oportunidades del mercado laboral para los desmovilizados en Colombia" es el resultado de una investigación realizada junto con algunos investigadores del grupo de Macroeconomía Aplicada de la Universidad de Antioquia. Aborda un tema de toda la relevancia en el contexto nacional e internacional, toda vez que, luego de un proceso de conflicto armado interno de más de 50 años en Colombia, se firma la paz con la guerrilla más antigua y rica del mundo. Para septiembre de 2017 el país contaba con 50.278 personas que ingresaron al proceso de reintegración, la tasa de desempleo para los desmovilizados era del 19,6%, casi diez puntos porcentuales superior a la nacional con presencia de una alta discriminación invisible y de género. Este es un asunto tema álgido puesto que, no es sencillo medir el grado de aceptación de la sociedad hacia los excombatientes y el sesgo percibido por ellos puede modificar su comportamiento ya sea para satisfacer a los entrevistadores, evitar sentir vergüenza o quedar bien y aumentar sus chances de ser empleados, con lo cual se disminuye la capacidad para diseñar políticas públicas que contribuyan a una reintegración exitosa. Más aún, cuando uno de los elementos fundamentales para asegurar que los desmovilizados no reincidan o logren una reintegración exitosa es que consigan un sustento económico proveniente de una actividad económica lícita. Es allí, donde este capítulo permite comprender tanto algunos aspectos en común

como otros asociados al mercado laboral que significarán esfuerzos significativos en materia de política pública.

El texto se centra pues en el cálculo de las probabilidades que un desmovilizado se incorpore laboralmente en actividades económicas formales e informales, considerando el alto riesgo de reincidencia que representa para esta población, el participar en el mercado laboral de manera informal. Parte del análisis de las características de la población desmovilizada, la cual enfrenta un alto índice de desempleo, debido a distintos factores como la falta de niveles educativos y experiencia adecuados, el bajo interés de empleabilidad, la problemática de la alta tasa de desempleo en Colombia, entre otros. Con ello, el estudio, define un perfil de los desmovilizados, empleando información de la Agencia para la Reintegración y la Normalización, y posteriormente, se emplean las ecuaciones de Mincer para realizar los respectivos análisis del mercado laboral. Los resultados de este ejercicio sugieren que, en la actualidad en Colombia dadas las diferentes variables incluidas en el modelo el mercado laboral no se genera los estímulos para la incorporación de población reinsertada. El estudio además concluye que esta población presenta mayores probabilidades de reincidir en el delito y de engancharse en actividades informales, lo cual posibilita que la población desmovilizada esté cerca de volver a la ilegalidad. Los departamentos en los que los desmovilizados tienen mayor probabilidad de encontrar una oportunidad en el mercado laboral son los departamentos de Antioquia, Bogotá, Caldas, Cundinamarca, Quindío, Risaralda, Valle del Cauca y Santander, lo cual se compagina con el potencial de mercado y el tamaño del PIB.

De otro lado, la aproximación utilizada también evidencia que para evitar la reincidencia o participación en actividades informales se requiere nivelar los años de educación, pero ello debe realizarse de manera intensiva, recursiva y acelerada. En este sentido, se requiere garantizar niveles de formación en educación mínimo de 13 años. Una limitante que encuentra el estudio para poder lograr que tengan 13 años de formación es que la mayoría de esta población se encuentra entre los 26 y los 40 años. Un análisis costo – beneficio hace inviable esta posibilidad, incluso para los que hoy cuentan con 26 años pues al terminar el ciclo de formación tendrían alrededor de 40 años lo que para ese momento les limitaría posibilidades en el mercado laboral. Por lo anterior, el estudio encuentra viable diseñar una formación a la medida tanto en la formación básica como superior que sea compatible con la generación de ingresos, así como también formación diferenciada de acuerdo con sectores económicos y en términos geográficos. Así también, encuentra viable la creación de programas tendientes a favorecer la creación de empresas (emprendimientos) que permitan la generación de ingresos. Finalmente, el estudio propone estrategias que maximicen el éxito del proceso de paz firmado entre el Gobierno Colombiano con las Fuerzas Armadas Revolucionarias de Colombia FARC, como las políticas de apoyo financiero o los beneficios tributarios.

Resumo

Esta tese doutoral consta de tres ensaios que se enmarcan dentro dos campos de investigación de Economía Xeográfica e a Economía Rexional. O primeiro ensaio (capítulo 1) titulado "*Potencial de mercado e disparidades de ingresos nos departamentos colombianos: análise para o período 1990-2015*" enmárcase dentro dos modelos de economía xeográfica. Os outros dous ensayos

titulados "*Os efectos das transferencias das regalías aos gobernos subnacionais no emprego formal rexional en Colombia* "(Capítulo 2) e "*Oportunidades no mercado laboral para os desmobilizados en Colombia* " (Capítulo 3) son específicos do campo de investigación da economía rexional no seus aspectos de análise do mercado laboral.

Analizar as razóns polas que algúns países son ricos e outros pobres é quizais unha das cuestións máis antigas e fundamentais da economía. Este mesmo debate podería transferirse a unha escala inferior e a pregunta sería semellante, por que a nivel de país as distintas unidades territoriais que os compoñen teñen diferentes niveis de desenvolvemento? Investigáronse distintas dimensións do crecemento económico para comprender as variacións nos niveis de benestar e crecemento que afectan aos países e rexións de cada país. Nesta vasta literatura, a teoría neoclásica utiliza o investimento en capital físico e educación como un elemento importante para comprender as diferenzas nas taxas de crecemento entre rexións / países (Solow1956, Mankiw et al., 1992). Entre as distintas explicacións, os desenvolvementos da teoría da Economía Xeográfica permítennos construír modelos de equilibrio xerais para explicar non só as razóns polas que a actividade económica se concentra no espazo senón tamén a acumulación endóxena de factores de produción (Krugman, 1991, Gallap et al., 1999).

A importancia do marco conceptual e teórico da economía xeográfica está intimamente ligada á definición de distancia xeográfica e proximidade. Aínda que é posible utilizar a distancia como factor esóxeno, a teoría da economía xeográfica establece que outras dimensións da proximidade xeográfica poden entenderse mediante a inclusión doutras dimensións económicas como o comercio e o potencial económico e de ingresos. A literatura de Economía Xeográfica emprega diferentes combinacións destes factores para calcular o acceso ao mercado e o seu potencial.

O primeiro ensaio da tese que leva por título "Potencial de mercado e disparidades de ingresos nos departamentos colombianos: a análise para o período 1990-2015" utiliza como marco teórico a teoría da economía xeográfica para dar unha explicación ás importantes diferenzas que existen nos niveis de desenvolvemento para os departamentos (rexións) nos que se estrutura territorialmente a República de Colombia. En consecuencia, utiliza o concepto de proximidade xeográfica entre os diferentes departamentos colombianos a través do cálculo dos potenciais mercados departamentais correspondentes para incluír outras dimensións da actividade económica e non só a distancia física entre os distintos departamentos cando explica os diferenciais no niveis de rendemento dos diferentes departamentos. O potencial do mercado convértese nunha variable central á hora de explicar como a xeografía de segunda natureza, é dicir, a xeografía que depende da acción humana, en oposición á xeografía de primeira natureza, é dicir, a presenza de ríos, montañas, dotación de hidrocarburos, etc., - xoga un papel relevante na explicación das disparidades nos niveis de renda per cápita dos departamentos colombianos. Baseándonos na teoría da Economía Xeográfica, derivamos a chamada ecuación nominal de salarios que establece para cada ubicación "i" do espazo unha relación entre os salarios nominais dese lugar "i" e unha suma ponderada pola distancia do volume de actividade económica das localidades veciñas, que se coñece xeralmente como potencial de mercado. A estimación desta ecuación empregando datos dos departamentos colombianos durante o período 1990-2015 revela que o potencial do mercado xoga un papel crucial na explicación da distribución espacial dos niveis de ingresos per cápita en Colombia. Ademais, os resultados tamén amosan que o acceso aos mercados adquiriu maior importancia ao longo do período de análise, o que apunta a un proceso de concentración espacial das actividades económicas en Colombia, onde as forzas de aglomeración son máis importantes que as forzas de dispersión.

Finalmente, o capítulo analiza dúas canles importantes que poderían afectar a estrutura espacial dos niveis de ingresos en Colombia, como son as dotacións de capital human e capital físico.

Con base no mencionado, ampliamos a investigación (capítulo 2) que discute os efectos das transferencias de dereitos aos municipios no emprego formal rexional durante o período 2002-2018. O segundo ensaio da tese (capítulo 2) leva por título "Os efectos da transferencia de dereitos aos gobernos subnacionais sobre o emprego formal en rexións en Colombia" e céntrase na análise do papel que xogan as transferencias de dereitos xeradas pola explotación de recursos non renovables na configuración dos niveis formais de emprego a partir da creación do Sistema Xeral de Royalties (SGR) en contraste cos resultados do sistema anterior. Se se analiza o papel dinámico que os dereitos de autor poden desempeñar na economía e como se converten nun mobilizador de man de obra, estúdase que papel xogan na distribución ou configuración das disparidades rexionais, así como en cuestións como o esforzo, transparencia, responsabilidade e eficiencia fiscal das entidades territoriais despois de recibir estes recursos. Ademais, considéranse os riscos fiscais que o fiscal pode xerar debido á pouca previsibilidade que se pode xestionar no orzamento porque este tipo de recursos adoitan ser volátiles. A literatura estudou extensamente estes impactos e resaltou a existencia dun risco significativo de que isto leve á chamada "maldición dos recursos" ou á constitución de economías de enclave.

Polo tanto, o capítulo 2 aproveita a reforma constitucional levada a cabo polo Goberno e que entrou en vigor en 2012 asociada ao Sistema Xeral de regalías e que favoreceu a redistribución dos recursos derivados da compensación pola extracción de recursos non renovables. Neste sentido, o capítulo permitiunos identificar o efecto que a mencionada reforma tivo no emprego formal. Neste sentido, autores como Echeverry, Alonso e García (2011) defenden que a reforma que creou o sistema xeral de dereitos foi baseada en catro principios: (i) equidade social e rexional, (ii) aforro para o futuro, (iii) competitividade rexional e (iv) bo goberno. O goberno nacional substituíu o esquema anterior polo Sistema Xeral de Royalties para utilizar esta importante fonte de recursos, promover o crecemento rexional, a equidade entre rexións, reducir as taxas de pobreza e aumentar a competitividade do país. A este respecto, Cárdenas (2013) suxire que o sistema antigo tiña un baixo impacto e resultados, baséase no seguinte: (i) Moitos proxectos fragmentados con baixo impacto social e económico, (ii) mala planificación e execución, (iii) para 2009 ningún certificouse a entidade territorial en toda a cobertura e (iv) Produtores: alta dependencia dos dereitos de autor, baixa tributación, baixo desenvolvemento sostible. Debido a todo o anterior, este capítulo utilizou un estimador de regresión discontinuo para identificar o efecto sobre o emprego formal dos dereitos, aproveitando a forma en que se distribúen os dereitos do Fondo Rexional de Compensación de acordo coa normativa vixente, en función dos dereitos. distribuído dependendo de se os concellos teñen porcentaxes da súa poboación con necesidades básicas non satisfeitas, UBN, inferior ou superior ao 35%. A metodoloxía é similar á anteriormente aplicada en estudos de Angrist e Lavy (2001), van der Klaauw (2002), Hahn, Todd e Van der Klaauw (2001), Ferrez e Finan (2010) e en Corbi, Papaioannou e Surico (2014).

Para Colombia as regalías son unha importante fonte de ingresos para o desenvolvemento, con potencial para financiar diferentes proxectos que, á súa vez, teñen a capacidade de mellorar o nivel de vida dos habitantes, ademais de mellorar os niveis de produción en todas as áreas. rexións do país. Na Constitución política colombiana afírmase que o Estado é o propietario do subsolo e dos recursos naturais non renovables, que xeran unha consideración económica derivada da súa explotación. A partir da Lei 141 de 1994 creáronse a Comisión Nacional dos dereitos e o Fondo

nacional de dereitos (FNR), que permitiron esta consideración dos recursos e, á súa vez, estableceron as regras para a súa liquidación e distribución. A distribución e execución de dereitos a través do FNR permitiu a distribución do diñeiro obtido polo Estado mediante dereitos directos, ás entidades territoriais propietarias e desde onde se explotaron os recursos, e dereitos indirectos, ás entidades territoriais que realizaron o transporte de estes, pero sen ter en conta o resto de rexións, departamentos, concellos e vilas do país. Isto, sen dúbida, constituíu un grave problema de disparidade en todo o territorio nacional, xa que os investimentos realizados cos ingresos xerados pola explotación deste tipo de recursos deberían beneficiar a toda a poboación. Ademais, segundo Cárdenas (2013), este sistema tivo un baixo impacto e resultados, con moitos proxectos fragmentados con baixo impacto social e económico, unha planificación e execución deficientes e unha alta dependencia dos dereitos de autor, baixa tributación, baixo desenvolvemento sostible .

A reforma aprobada polo goberno nacional substituíu a FNR pola SGR en 2011, baseouse nos principios de equidade social e rexional, aforrando para o futuro, competitividade rexional e bo goberno; considerando, ademais, a descentralización e autonomía das entidades territoriais. Este compromiso buscaba promover unha distribución máis equitativa, garantindo o uso dos recursos de forma eficiente e honesta, co fin de contribuír á transparencia e ao uso efectivo dos recursos de dereitos, contribuíndo así a unha maior equidade entre rexións como a a diminución das taxas de pobreza e o aumento da competitividade do país. A metodoloxía empregada para identificar o efecto dos dereitos sobre o emprego formal rexional é unha regresión discontinua, aproveitando a forma en que se distribúen os dereitos do Fondo de Compensación Rexional (FCR ou RCF en inglés) de acordo coa normativa vixente, en función de que distribúen as regalías segundo se os concellos teñan porcentaxes da súa poboación con necesidades básicas non satisfeitas (NBI ou UBN en inglés) inferior ou superior ao 35%. Hai efectos positivos e significativos das regalías sobre o emprego formal, principalmente nos sectores industriais; electricidade, gas e auga; actividades inmobiliarias e comerciais; e intermediación financeira. Ademais, compróbase que os dereitos da FCR aumentaron a cantidade de recursos destinados a proxectos de investimento subnacionais sen reducir a cantidade doutras fontes de financiamento, en particular os recursos propios dos concellos. Finalmente, os resultados tamén mostran que estas regalías non afectaron o esforzo fiscal dos concellos e non afectaron ás súas ganancias tributarias iniciais.

Este capítulo permitiu concluír que o SGR permitiu redistribuír a cantidade de dereitos e utilizar esa importante fonte de recursos, promover o crecemento rexional e municipal, a equidade entre rexións dada a redistribución e reducir as taxas de pobreza a través do emprego formal. Ademais, está claro que co aumento da cota de dereitos para os municipios que recibían proporcións máis baixas antes da reforma, o emprego formal volveuse máis dinámico mentres que a reforma aumentaba os recursos para as entidades territoriais que non os recibían. Do mesmo xeito, dados os niveis incipientes de desenvolvemento das entidades territoriais, especialmente as que se atopan na sexta categoría, os recursos procedentes dos dereitos de autor convertéronse nun vehículo para a dispoñibilidade de recursos favorables á dinámica territorial e a favor do emprego formal. Utilizando a Lei 1530 de 2012 e o Decreto 1073 de 2012, ademais da regresión discontinua baseada en necesidades básicas non satisfeitas, NBI permitiunos determinar o impacto no emprego formal atopando un efecto positivo sobre as regalías. Finalmente, conclúese que o obxectivo dos dereitos de autor é un problema de interese central no país debido á súa inmensa cantidade (aproximadamente o 1,6% do PIB) e porque son unha importante fonte de financiamento para o desenvolvemento rexional, polo que que a súa adecuada distribución é esencial para favorecer o emprego formal e estimular o desenvolvemento das rexións.

O capítulo tres propón que a discriminación nos mercados de traballo foi un importante obxecto de estudo en Ciencias Económicas nos últimos anos. As análises coinciden en sinalar que, para falar de discriminación, as poboacións obxecto de estudo e comparación deben ser idénticas en todas as dimensións que poden afectar á produtividade, á luz das consideracións dos empresarios (Bertrand e Mullainathan, 2004 e Altonji e Blank, 1999). De acordo co propósito deste traballo, entre os aspectos empregados considéranse tres tarefas fundamentais: (i) a caracterización da poboación reinsertada segundo a información descrita anteriormente, (ii) proponse un esquema de incentivos reinsertado pensando no o mercado de traballo ao que se enfrontarán e (iii) Estímanse as ecuacións mínimas para atopar os determinantes que teñen maiores efectos sobre a probabilidade de obter emprego formal e sobre os salarios esperados da poboación de interese en diferentes escenarios.

Este último capítulo da tese analizou empíricamente as oportunidades do mercado laboral para os desmobilizados en Colombia. Este terceiro ensaio titúlase "Oportunidades do mercado laboral para os desmobilizados en Colombia" é o resultado dunha investigación realizada xunto con algúns investigadores do grupo de Macroeconomía Aplicada da Universidade de Antioquia. Aborda unha cuestión de toda relevancia no contexto nacional e internacional, xa que, despois dun proceso de conflito armado interno de máis de 50 anos en Colombia, asínase a paz coa guerrilla máis antiga e rica do mundo. En setembro de 2017, o país tiña 50.278 persoas que entraron no proceso de reinserción, a taxa de paro dos desmobilizados era do 19,6%, case dez puntos porcentuais superior á nacional con presenza de discriminacións invisibles e de xénero elevadas. Este é un tema candente xa que non é doado medir o grao de aceptación da sociedade cara aos excombatentes e o sesgo percibido por eles pode modificar o seu comportamento para satisfacer aos entrevistadores, evitar vergoña ou parecer bo e aumentar a súa posibilidades de ser empregado, diminuíndo así a capacidade para deseñar políticas públicas que contribúan a unha reinserción exitosa. Ademais, cando un dos elementos fundamentais para garantir que os desmobilizados non reincidenten ou logren unha reinserción exitosa é que obteñen un sustento económico dunha actividade económica legal. É aí, onde este capítulo nos permite comprender tanto algúns aspectos en común como outros asociados ao mercado de traballo que suporán importantes esforzos nas políticas públicas.

Por outra banda, o enfoque empregado tamén mostra que para evitar a reincidencia ou a participación en actividades informais é necesario nivelar os anos de educación, pero isto debe facerse dun xeito intensivo, recursivo e acelerado. Neste sentido, é necesario garantir un nivel mínimo de formación en educación de 13 anos. Unha das limitacións que o estudo atopa para acadar 13 anos de formación é que a maioría desta poboación ten entre 26 e 40 anos. Unha análise custo-beneficio fai inviable esta posibilidade, incluso para os que teñen actualmente 26 anos, xa que ao final do ciclo formativo terían arredor de 40 anos, o que nese momento limitaría as súas posibilidades no mercado laboral. Polo tanto, o estudo considera viable deseñar formación a medida tanto na educación básica como na superior que sexa compatible coa xeración de ingresos, así como formación diferenciada segundo os sectores económicos e en termos xeográficos. Así mesmo, considera viable a creación de programas destinados a favorecer a creación de empresas (empresas) que permitan xerar ingresos. Finalmente, o estudo propón estratexias que maximizan o éxito do proceso de paz asinado entre o goberno colombiano coas FARC das Forzas Armadas Revolucionarias de Colombia, como políticas de apoio financeiro ou beneficios fiscais.

Index of figures and tables

Introduction

El análisis de las razones de por qué unos países son ricos y otros pobres es quizás unas de las preguntas más antiguas y fundamentales de la economía. Este mismo debate se podría trasladar a una escala inferior y la pregunta sería similar, ¿por qué a nivel de país las diferentes unidades territoriales que los componen tienen unos niveles de desarrollo dispares?. Diferentes dimensiones del crecimiento económico se han investigado para entender variaciones en los niveles de bienestar y crecimiento que afectan a países y a regiones dentro de cada país. En esta amplia literatura, la teoría neoclásica usa la inversión en capital físico y la educación como un elemento importante para entender las diferencias en las tasas de crecimiento entre regiones/países (Solow1956, Mankiw et al., 1992). Por otro lado, la teoría del crecimiento endógeno enfatiza que el capital humano sirve para entender la acumulación endógena de actividad económica y el crecimiento económico (Lucas 1998, Romer 1994).

Entre las diferentes explicaciones, los desarrollos de la teoría de la *Economía Geográfica* (*Geographical economics/New Economic Geography*) nos permite construir modelos de equilibrio general para explicar no solo las razones de por qué la actividad económica se concentra en el espacio sino también la acumulación endógena de factores de producción (Krugman, 1991, Gallap et al., 1999). Concretamente Redding y Venables (2004a, 2004b) constituyen uno de los primeros intentos en formalizar los modelos de *Economía Geográfica* donde la distancia entre países se convierte en la causa real a la hora de explicar las variaciones en los niveles de renta per cápita entre países.

La importancia del marco conceptual y teórico de la Economía Geográfica está íntimamente unida a la definición de distancia y proximidad geográfica. Aunque es posible utilizar la distancia como un factor exógeno, la teoría de la Economía Geográfica establece que otras dimensiones de la proximidad geográfica se pueden entender a través de la inclusión de otras dimensiones económicas como el comercio y el potencial económico y de renta. La literatura de la Economía Geográfica usa diferentes combinaciones de estos factores para calcular el acceso al mercado y el potencial de mercado (Market Access o Market Potential en su acepción anglosajona).

El capítulo I de esta tesis doctoral utiliza como marco teórico la teoría de la Economía Geográfica para dar una explicación a las importantes diferencias que existen en los niveles de desarrollo para los departamentos (regiones) en los que se estructura territorialmente la República de Colombia. Consecuentemente usa el concepto de proximidad geográfica entre los distintos departamentos colombianos a través del cálculo de los correspondientes *potenciales de mercado* departamentales para incluir otras dimensiones de la actividad económica y no solo la distancia física entre los distintos departamentos a la hora de explicar los diferenciales en los niveles de desempeño de los distintos departamentos. El potencial de mercado pasa a convertirse en una variable central a la hora de explicar cómo la geografía de segunda naturaleza (Second Nature Geography)-es decir la geografía que depende de la acción humana- en contraposición a la geografía de primera naturaleza (First Nature Geography)-es decir la presencia de ríos, montañas, dotación de hidrocarburos, etc.,- desempeña una papel relevante en la explicación de las disparidades en los niveles de renta per cápita para los departamentos colombianos. La estimación de la denominada *ecuación nominal de salarios* de la Economía Geográfica usando datos para los departamentos colombianos para el período 1990-2015 demuestra que el potencial de mercado tiene un efecto relevante y muy

significativo desde el punto de vista económico a la hora de caracterizar la estructura espacial de los diferentes niveles de desarrollo que se observan en Colombia a lo largo de este período.

El capítulo II aprovecha la reforma constitucional realizada por el Gobierno y que entró en vigencia en el año 2012 asociada al Sistema General de Regalías y que favoreció la redistribución de los recursos derivados de la compensación por la extracción de recursos no renovables. En este sentido, el capítulo permitió identificar el efecto que tuvo la reforma anteriormente mencionada sobre el empleo formal. En este sentido, autores como Echeverry, Alonso y García (2011), plantean que la reforma que creó el Sistema General de Regalías, se fundamentó en cuatro principios: (i) la equidad social y regional, (ii) el ahorro para el futuro, (iii) la competitividad regional, y (iv) el buen gobierno. El Gobierno Nacional reemplazó el esquema anterior por el Sistema General de Regalías con el fin de utilizar esa importante fuente de recursos, para impulsar el crecimiento regional, la equidad entre regiones, disminuir los índices de pobreza y aumentar la competitividad del país. Al respecto Cárdenas (2013) sugiere que el antiguo sistema presentaba bajo impacto y resultados, esto se fundamenta en que: (i) Muchos proyectos fragmentados de bajo impacto social y económico, (ii) Deficiente planeación y ejecución, (iii) Para 2009 ninguna entidad territorial estaba certificada en todas las coberturas, y (iv) Productores: alta dependencia de regalías, baja tributación, bajo desarrollo sostenible. Por todo lo anterior este capítulo para identificar el efecto en empleo formal de las regalías utilizó un estimador de regresión discontinua, aprovechando la forma como se distribuyen las regalías del Fondo de Compensación Regional de acuerdo a la normatividad vigente, con base en la cual las regalías distribuyen dependiendo si los municipios tienen porcentajes de su población con necesidades básicas insatisfechas, NBI, inferiores o superiores a 35%. La metodología es similar a la anteriormente aplicada en estudios de Angrist y Lavy (2001), van der Klaauw (2002), Hahn, Todd and Van der Klaauw (2001), Ferrez y Finan (2010) y en Corbi, Papaioannou y Surico (2014).

El capítulo III plantea que la discriminación en los mercados laborales ha sido un importante objeto de estudio de las Ciencias Económicas en los últimos años. Los análisis coinciden en señalar que para hablar de discriminación se requiere que las poblaciones objeto de estudio y comparación, sean idénticas en todas las dimensiones que pueden afectar la productividad, a la luz de las consideraciones de los empleadores, (Bertrand y Mullainathan, 2004 y Altonji y Blank, 1999). de acuerdo con el propósito de este trabajo, entre los aspectos utilizados se consideran tres tareas fundamentales: (i) la caracterización de la población reinsertada de acuerdo con la información descrita anteriormente, (ii) se propone un esquema de incentivos de reinsertados pensando en el mercado laboral al que se van a enfrentar, y (iii) se estiman ecuaciones de Mincer con el fin de encontrar los determinantes que mayores efectos tienen sobre la probabilidad de conseguir empleo formal, y sobre los salarios esperados de los la población de interés en distintos escenarios.

Tabla de contenido

Chapter 1: Market Potential and Income Disparities across Colombian Departments: Analysis for the period 1990-2015

Abstract

This chapter tests the robustness of estimates of market potential impact on the income disparities based on the nominal wage equation of the Geographical Economics theory. The so called nominal wage equation which establishes a relationship between nominal wages and a distance weighted sum of the volume of economic activities in surrounding locations which is usually known as market potential. Results of the estimations using Colombian regional data over the period 1990-2015 indicate that once we control for physical and human capital the estimated impact of market potential decreases, but it keeps its crucial role in the explanation of the Colombian income disparities. Controlling for spatial dependence shows that around 50% of the impact of market potential on income disparities for each region can be attributed to neighbouring ones. Finally, two important channels that might be affecting the shaping of the spatial distribution of income were devised, physical and human capital.

Key words: Regional Disparities, Spatial Structure of Wages, Market Access

JEL Classification: R11, R12, R13, R14, F12, F23

1.1 Introduction

Differences in the levels of development across regions within countries is almost a natural feature. In the case of Europe, differences in regional income levels within countries is quite sizeable in many of them: for instance in Italy (Northern regions versus Southern regions), Spain (Northeastern regions versus South and East regions), Portugal (North versus South) and in most Central and Eastern European countries the Western regions of these countries are more developed than the Eastern ones. These patterns have raised considerable concerns in popular debates as well as in policy circles and have led to the establishment of policies aimed at levelling out of income differences and at allowing a catch-up of peripheral regions. In the case of Colombia a phenomena similar to the one described for the European countries is at work. Per capita income levels differ by significant amounts across the Colombian departments. A couple or three examples suffices to understand the size of these differences. In the year 2015, the ratio between the per capita income of Casanare (Oil-producing department) and the average per capita income across the Colombian departments was 2.81 or in other words, the per capita GDP of Casanare was almost three times higher than the average per capita GDP in Colombia.

If we exclude the oil-producing departments (Arauca, Casanare, Meta y Santander) the 2015 ratio between Bogota's per capita income and the average per capita income in Colombian was still 2.27 (per capita income in Bogota more than double of the country's average) . Even if we discard the capital (Bogota), the ratio between the second wealthiest department (Antioquia) and the average per capita income of the country is 1.5, which is a quite sizeable difference (50%!). There are different approaches taken by the economic theory to explain these income differences. On the one hand we can resort to the economic growth theory to deliver potential explanations for these facts which go from differences in saving rates, investment rates to problems of technology diffusion (see Barro and Sala-i-Martin, 1991, 1995 among others). On the other hand we can deliver other messages based on traditional development theories which put the emphasis on first nature geography characteristics of the locations, hours of sunshine, endowments of hydrocarbons, access to navigable rivers, etc. (see for instance Hall and Jones (1999)).

This chapter takes a different approach to the analysis of the differences in income levels across the Colombian departments. The so-called *New Economic Geography* (NEG) or *Geographical Economics* (GE)[1] (Krugman 1991a, 1991b) has provided another conceptual framework within which the geographical structure of production and income levels can be analysed explicitly. This field has experienced rapid advances in the last two decades both from theoretical as well as empirical side.

This chapter applies the geographical economics framework in an exhaustive empirical investigation of the income structure at the level of department in Colombia. Therefore it is part of the growing literature that uses the theoretical tools from the Geographical Economics to analyze the impact of distance from markets on income levels. In a more technical way, what we will do in this chapter is to test one of the main predictions of these models, the so called *nominal wage equation*, for the case of the Colombian departments over the period 1990-2015. The basic idea is that in a world where regions or countries specialize in certain

[1] See Fujita et al. (1999), Brakman et al (2009) for a comprehensive introduction to the theory of Geographical economics. Overman et al. (2003), Redding (2011) provide an overview of the existing empirical literature.

goods and export them, firms in locations which are further away from main consumer markets or input suppliers will have to pay more for shipping their goods and buy their intermediate inputs and therefore the value added left to remunerate their local factors of production, among them workers will be lower.

The chapter finds widely support to the theoretical predictions related to the nominal wage equation of the *Geographical Economics* literature, i.e, the elasticity of per capita income with regard to market potential is positive, statistically significant and economic important across the different estimations carried out in the chapter (cross-section, panel data, spatial panel data). Moreover, both in the cross section estimates of the impact of market potential on per capita income levels splitting the period under analysis in three different periods (1990-1994, 1995-2004 and 2005-2015) and in the panel data estimates controlling for difference in the slope estimators of market potential we have found evidence of an increasing role played by market potential in shaping the spatial income structure observed in Colombia. Therefore, distance from markets matters and seems to have an increasing role when looking at differences in income per capita across Colombian departments. Although it is far beyond the scope of this paper these results are also in line with a dynamic process of spatial concentration of economic activity in Colombia over the course of the year under analysis. Another important contribution of this chapter lies in disentangling the channels through which market potential affects the levels of economic development in Colombia. In particular we have include as additional controls to the baseline estimation two variables, physical capital and human capital, to capture the potential indirect effects of economic geography. The results in the non-spatial specifications show that both physical capital and human capital could be important drivers of income levels. However in the spatial specifications, human capital does not show as statistically significant.

The structure of the chapter is as follows: First in the second section, we study the dispersion of the economic development levels among the departments of Colombia, using per capita income as a proxy for the economic development levels. The results of the spatial descriptive analysis will confirm the existence of sizeable regional disparities and strong spatial dependence in the distribution of per capita income. Next we estimate the coefficient of a simple specification, which reveals the positive and significant effect of market potential on per capita income levels. The theoretical economic rationale from the geographical economics literature that support these empirical results are outline in the third section. In the fourth section we carry out several robustness checks to the baseline estimation. On the one hand we discuss the effects of controlling for regional differences in human and physical capital (section 4.1) and on the other for spatial dependence (section 4.2). Based on these arguments the original nominal wage equation of the geographical economics literature is augmented and the estimations obtained with alternative specifications are compared with those of the baseline model. In the fifth section we provide additional robustness checking including results from panel data estimators with and without spatial effects from an alternative measure of economic development levels which is closer to the theoretical magnitude in the geographical economics model. Finally, the sixth section presents the main conclusions and policy implications.

1.2 The Geography of income disparities in Colombia

Preliminary Evidence

The levels of per capita income in the Colombian *departments* differ by significant amounts (Bonet and Meisel (2006), Meisel (1993), Mora and Salazar (1994), Birchenall and Murcia (1996), Rocha and Vivas (1998), Bonet and Meisel (1999), Barón and Meisel (2003), Barón (2004) and Bonet and Meisel (2006)). Excluding from the sample the oil-producing departments the per capita income in Bogota is well above the per capita income of any other department in Colombia. The economic evidence provided in this section of the chapter was obtained from the Colombian National Institute of Statistics (DANE) using data for the thirty-three departments of Colombia for the period 1985-2015. We use the per capita income of each department as the standard measure of the level of economic development. In order to build our per capita income series from 1985 to 2015 we have to use the series interpolation method, his starts from the principle that in the base years, the national accounts are prepared in as much detail as possible, incorporating new sources and calculation methods that strengthen research. The base year remains unchanged and is used as a reference to proportionally distribute the difference that originates between the new base year and the data obtained in the previous base year (DANE, 2013). The procedure consists of four steps:

1. Nominal GDP values of two consecutive base years are identified.
2. Using a quotient, the difference between these two values is found, as follows:

$$\mathrm{D} = \frac{Vn_B^{t=B}}{Vn_b^{t=B}}$$

where **Vn** denotes nominal value, **t** denotes year under consideration. **B** is the new base and **b** is the previous base.

3. A geometric average is calculated, in order to distribute this difference between the years that separate these two bases, as follows:

$$\mathrm{r} = D^{\frac{1}{n}}$$

being ***n*** the number of observations between one base and another.

4. With this we proceed to calculate the nominal values for the years "t" between the two bases, as follows:

$$Vn_B^t = Vn_b^t * r^{(t-b)}$$

where Vn denotes nominal value, t denotes year under consideration. B is the new base, b is the previous base and r is the geometric average.

In a first step, our descriptive analysis covers the cross-sections for three years 1985, 2000 and 2015 and them we pool the data for three periods: initial (1990-1994), middle (1995-2004) and final (2005-2015). With this later division of the whole period into these three periods we want to analyze the dynamics of the relationship between per capita income and market potential avoiding potential temporal shocks that could affect the conclusions obtained from the readings of the years 1985, 2000 and 2015. However, in the last sections

of the chapter we also provide results obtained using the entire yearly panel data set, which allows controlling for unobservable department effects that are likely to affect the link between market potential and per capital income levels.

Table 1.1 shows the data of the nominal per capita income across Colombian departments for the years 1985, 2000 and 2015. The bottom part of the table provides some basic ratios which allow us to see how per capita income differences have evolved over the course of these thirty years (for the year 1985 we do not have data for the so called *Nuevos departamentos*).

Table 1.1. Per capita income across Colombia Departments (1985, 2000, 2015)

DEPARTAMENTOS	1985	2000	2015
Antioquia	202.646	5.264.144	16.942.369
Atlántico	178.431	4.082.264	13.316.873
Bogotá D. C.	260.151	7.833.795	25.603.320
Bolívar	142.963	3.638.241	15.700.750
Boyacá	142.708	4.333.273	18.586.548
Caldas	148.367	3.676.398	11.879.663
Caquetá	118.596	2.332.023	7.758.949
Cauca	84.362	2.421.236	9.797.204
Cesar	149.309	3.117.524	13.853.765
Córdoba	95.094	2.977.925	8.123.914
Cundinamarca	190.763	4.892.446	15.263.572
Chocó	**71.568**	**1.538.542**	6.366.816
Huila	174.527	3.717.620	12.458.682
La Guajira	**294.813**	2.848.754	8.314.914
Magdalena	98.116	2.341.921	8.262.278
Meta	162.734	5.630.070	34.300.254
Nariño	85.774	2.101.931	7.392.956
Norte Santander	139.687	2.825.605	9.762.595
Quindío	187.985	3.440.683	10.946.207
Risaralda	166.483	3.646.419	12.053.116
Santander	202.172	6.019.077	31.447.121
Sucre	91.110	2.130.727	7.571.211
Tolima	125.443	3.544.510	12.040.288
Valle del Cauca	198.015	5.605.757	16.497.445
Amazonas		2.576.734	6.820.298
Arauca		9.684.544	18.092.751
Casanare		**27.968.626**	**38.125.668**
Guanía		2.539.396	6.870.450
Guaviare		2.445.047	6.464.974
Putumayo		2.267.263	10.234.528
San Andrés y Providencia		5.665.891	15.959.813
Vaupés		1.972.458	**5.152.868**
Vichada		4.180.115	6.016.061
Average 1	**154.659**	**4.523.059**	**13.575.098**
Max/average1	**1,91**	**6,18**	**2,81**
Min/average1	**0,46**	**0,34**	**0,38**
Max	**294.813**	**27.968.626**	**38.125.668**
Min	**71.568**	**1.538.542**	**5.152.868**
average2*	**152.132**	**3.446.850**	**11.241.808**

*Note: Without oil-producing departments (Arauca, Casanare, Meta y Santander)
Source: Own elaboration based on DANE

Comparing the ratios of max/average for years 2000 and 2015, for which we have data for all the Colombian departments, we can conclude that income differences are very big. Though from a dynamic perspective the situation is much better in 2015 than it was in the year 2000. If we pool the data for the periods (1990-1994), (1995-2004) and (2005-2015) as it is shown in Table 1.2 the results continue to show the sizeable differences in terms of development levels across the Colombian departments. The period 1995-2004 is the period where we have the worst indicators in terms of the computed ratios and in this case the pooling of the data shows that the process of improvement from 1995-2004 period to 2005-2015 period was not as impressive as it is shown by the computations for single years (2000 and 2015). On top of that, the situation in the last period has worsen off in comparison with the one reflected for the year 2015. Overall comparing the initial period (1990-2004) with the final one (2005-2015) we assist to a process of widening the gap between in terms of per capita income differences across Colombian departments.

Table 1.2. Per capita income across Colombia Departments (1985-1994, 1995-2004, 2005-2015)

DEPARTAMENTOS	1985-1994	1995-2004	2005-2015
Antioquia	870.113	4.961.629	12.391.907
Atlántico	730.332	3.860.427	9.556.760
Bogotá D. C.	**1.085.907**	7.520.156	19.359.223
Bolívar	616.386	3.652.430	11.535.620
Boyacá	646.760	4.003.562	12.623.776
Caldas	687.184	3.665.616	8.919.007
Caquetá	505.597	2.244.126	5.487.713
Cauca	366.856	2.199.115	6.461.040
Cesar	565.524	3.336.917	10.992.903
Córdoba	452.212	2.785.394	6.540.630
Cundinamarca	778.135	4.695.636	11.524.085
Chocó	313.477	**1.511.731**	4.950.413
Huila	652.954	3.412.450	9.358.566
La Guajira	900.905	3.069.034	7.806.901
Magdalena	414.080	2.372.906	6.116.497
Meta	874.500	5.254.702	26.658.220
Nariño	354.011	2.039.700	5.241.788
Norte Santander	482.067	2.705.913	7.215.165
Quindío	699.645	3.254.584	7.834.535
Risaralda	681.533	3.559.837	9.029.497
Santander	828.615	5.805.887	20.917.072
Sucre	349.954	2.028.162	5.448.391
Tolima	588.150	3.439.160	9.057.595
Valle del Cauca	831.294	5.316.166	12.439.703
Amazonas	119.007	2.427.492	5.357.587
Arauca	440.496	7.949.647	19.718.389
Casanare	295.395	**17.131.330**	**33.084.502**
Guanía	100.663	2.083.511	4.876.989
Guaviare	331.469	2.683.196	4.845.939
Putumayo	**83.471**	1.814.622	8.040.120
San Andrés y Providencia	304.375	5.320.951	11.397.477

Vaupés	115.103	1.850.472	**3.557.343**
Vichada	159.795	3.007.640	5.229.207
Average 1	**521.999**	**3.968.609**	**10.411.350**
Max/Average1	**2,08**	**4,32**	**3,18**
Min/Average1	**0,16**	**0,38**	**0,34**
Max	**1.085.907**	**17.131.330**	**33.084.502**
Mín	**83.471**	**1.511.731**	**3.557.343**
Average 2*	**509.895**	**3.269.743**	**8.386.082**

*Note: Without oil-producing departments (Arauca, Casanare, Meta y Santander
Source: Own elaboration based on DANE

Figure 1.1 depicts three maps where the spatial distribution of the data reported in Table 1.2 can be observed. These maps confirm the existence of important differences in the economic development levels across the Colombian departments and to a certain extend a pattern of persistence over time. However, the most interesting feature is that there is a geographical pattern in distribution of per capita income levels which is reflected in a core-periphery type of structure using the terminology of the geographical economics literature. The *core* (high per capita income departments) takes in the center and Eastern parts of the country where the oil-producing departments are located (Arauca, Casanare, Meta, Santander and Bogota) and surrounding this core towards the West and far North and South we have those areas of the country that could be considered as *the periphery* in a geographical economics fashion. The further to the West and to the North and South we displace the lowest the levels of per capita income levels we find (far West-Choco, Nariño, far South-Amazonas, and far North-La Guajira).

The prediction of the core-periphery geographical economics models is that a big deal of the spatial pattern in the distribution of per capita income levels across different spatial settings has to do with the relative access to markets that the different units in the particular setting under consideration enjoy (Brakman et al., 2004, 2009; Breinlich, 2006; Bruna et al., 2015; Faiña and Lopez-Rodriguez, 2005; Fingleton, 2006; Hanson, 2005; Head and Meyer, 2006, 2011; Lopez-Rodriguez and Acevedo, 2013; López-Rodríguez and Faíña, 2006; Lopez-Rodriguez et al., 2011, Niebuhr, 2006; Overman et al., 2003; among others). This relative access to markets, or alternatively, relative remoteness, can be proxied by the so called market potential measure suggested initially by Harris (1954)[2].

[2] Harris (1954) also explains that the term was suggested by Clark (1940). The concept is analogous to that of population potential as proposed and mapped by Stewart (1947, 1948, 1952). It is an abstract index of the intensity of possible contact with markets. The concept is derived ultimately from physics, in which similar formulas are used in calculating the strength of a field, whether electrical, magnetic, or gravitational.

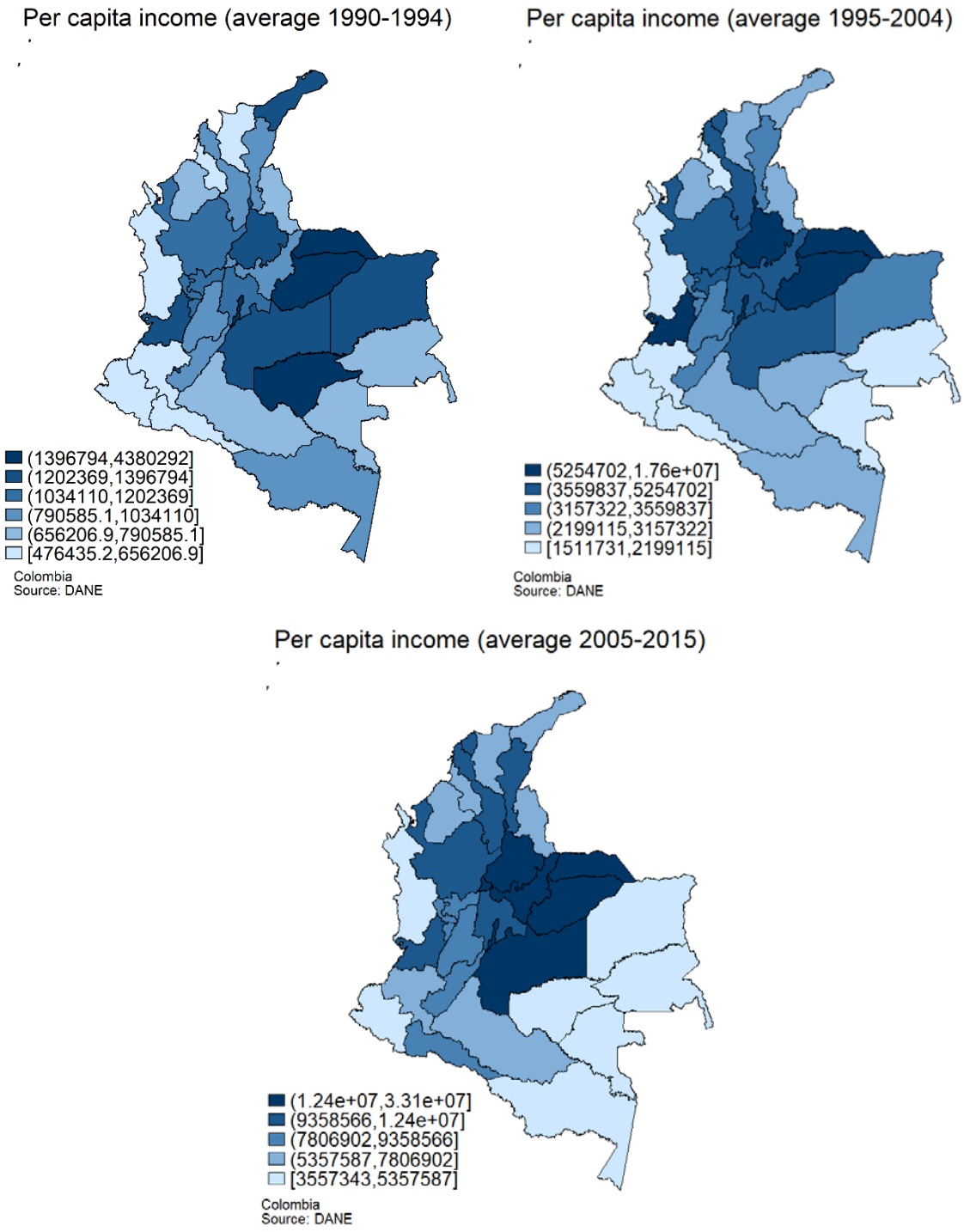


Figure 1.1. Spatial distribution of per capita income levels in Colombia.
Source: DANE and author´s own calculations

Market potential: construction and summary statistics

The Harris (1954) market potential (MP) of a Colombian department i is defined as the summation of markets accessible to that department divided by the distance between department i and the remaining ones[3]. Therefore, the market potential of a department will be positively associated with the purchasing power of the remaining departments but negatively related with the distance between each other. Mathematically, it adopts the following expression:

$$MP_{it} = \sum_{j=1}^{n} M_{jt} g(d_{ij}) \quad (1)$$

Where MPit represents the Harris (1954) market potential function for department i in period t, Mjt is a measure of the purchasing power of department j in period t (usually approximated by its income level, gross value added or population), dij is a measure of the distance between two generic departments i and j, g (·) is a decreasing function of the distance between two generic departments i and j and n is the number of departments considered. Additionally, the market potential of a given department i can be broken down into a domestic or internal component, market potential created by the department itself, (DMPit) and an external or foreign one, market potentials for that department of all remaining departments in the area under consideration, (FMPit). Approximating g (·) by the inverse of the distance between departments i and j, (1/dij), and taking into account the two components of the market potential, the mathematical expression (1) can be easily expanded to this one:

$$MP_{it} = \sum_{j=1}^{n} \frac{M_{jt}}{d_{ij}} = \frac{M_{it}}{d_{ii}} + \sum_{j \neq i}^{n-1} \frac{M_{jt}}{d_{ij}} = DMP_{it} + FMP_{it} \quad (2)$$

In making the calculations of the internal distance (dii) the standard methodology assumes that departments are circular and the internal distance is approximated by a function that is proportional to the radius of the department. The radius of a circular-shaped department "i" of size equal to "areai" is $r_i = \sqrt{area_i/\pi}$. In this chapter and following the work of Keeble et al. (1982), we will use $d_{ii} = 1/3 \cdot r_i = 0.188\sqrt{area_i}$ as the first option. On the other hand, following other authors such as CROZET (2004), Head and Mayer (2000), and Nitsch (2000) we will use $d_{ii} = 2/3 \cdot r_i = 0.376\sqrt{area_i}$ as the second option. Both formulas have been frequently used in the literature and give the average distance in a circular location under the assumption that production takes place in the centre and consumers are spread evenly across space. With regards to the variable Mjt we will use GDP and population of each department as proxies. Finally, distances will be measure both in kms between the capital cities of each department and in lorry travel times. Therefore we will build six different measures of market potential which we will label as MP1 (market potential based on GDP, physical distances and dii=1/3), MP2 (market potential based on GDP, physical distances and dii=2/3), MP3 (market potential based on GDP, travel time distances and dii=1/3), MP4 (market potential based on population, physical distances and dii=1/3) MP5

[3] The microeconomics grounds for the Harris (1954) market potential concept was first derived in the early nineties in the very influential Krugman's 1991 and 1992 papers on core-periphery geographical economics models.

(market potential based on population, physical distances and dii=2/3) and MP6 (market potential based on population, travel time distances and dii=1/3).

Tables 1.3 and 1.4 provide some information on the average composition of market potential over the periods 1990-1994, 1995-2004 and 2005-2015 to evaluate how it has changed over time. We calculate these access measures separately for the three periods by taking ten- to eleven- year averages (with the exception of the first period for which we only have five years available with the full sample of departamentos-33) with the goal of smoothing out short-run fluctuations. The total market potential has been broken down according to expression (2) into a domestic component and a foreign component and according to a weighting scheme based on a distance matrix expressed in kms (Table 1.3) and a distance matrix expressed in minutes of lorry travel times (Table 1.4). It can be seen that with the exceptions of the values obtained for the definitions of market potential MP2 and MP5, in the rest of alternative definitions (MP1, MP3, MP4 and MP6) the shares of the own market potential (domestic market potential) are around half of the total market potential of the department. When the distance matrix used in the computations of market potential is expressed in minutes of travel, the domestic market potential shares are higher than those obtained when a kms-based matrix is used[4] and in the case of MP3 the domestic component is bigger than the foreign one. Finally, in four of the definitions of market potential (MP1, MP4, MP5 and MP6) the importance of the domestic component of market potential has increased over the past 25 years (it has never decreased). This in part reflects the fact that the role of the distance is very important over the sample period.

Table 1.3. Summary statistics on Market Potential (distance matrix expressed in kms)

	1990-1994	1995-2004	2005-2015
Average fraction of Market Potential derived from own department (MP1)	47(%)	48(%)	48(%)
Average fraction of Market Potential derived from rest of departments (MP1)	53(%)	52(%)	52(%)
Average fraction of Market Potential derived from own department (MP2)	31(%)	32(%)	31(%)
Average fraction of Market Potential derived from rest of departments (MP2)	69(%)	68(%)	69(%)
Average fraction of Market Potential derived from own department (MP4)	44(%)	45(%)	45(%)
Average fraction of Market Potential derived from rest of departments (MP4)	56(%)	55(%)	55(%)

[4] Market potential definitions MP2 and MP5 are not strictly comparable due to a different method in the calculus of the domestic market potential. The domestic component in MP2 is based on the expression $d_{ii} = 2/3 \cdot r_i = 0.376\sqrt{area_i}$ whereas the domestic component in MP5 is based on $d_{ii} = 1/3 \cdot r_i = 0.188\sqrt{area_i}$

Table 1.4. Summary statistics on Market Potential (distance matrix expressed in travel times)

	1990-1994	1995-2004	2005-2015
Average fraction of Market Potential derived from own department (MP3)	51(%)	52(%)	51(%)
Average fraction of Market Potential derived from rest of departments (MP3)	49(%)	48(%)	48(%)
Average fraction of Market Potential derived from own department (MP5)	28(%)	29(%)	29(%)
Average fraction of Market Potential derived from rest of departments (MP5)	72(%)	71(%)	71(%)
Average fraction of Market Potential derived from own department (MP6)	48(%)	48(%)	49(%)
Average fraction of Market Potential derived from rest of departments (MP6)	52(%)	52(%)	51(%)

Figure 1.2 depicts the values of the log of the average market potential in the Colombian departments for the periods 1990-1994, 1995-2004 and 2005-2015. These average market potential values have been obtained first by computing year by year the market potential for each department and then averaging for the defined periods. In doing the market potential computations for these particular maps we use as the numerator (Mjt) of the market potential measure the GDP for department j at year t and as a measure of internal distance $d_{ii} = 1/3 \cdot r_i = 0.188\sqrt{area_i}$. The external distance between departmentos i and j is the distance in kilometres between the capital cities of each department.

From the maps it can be clearly seen that Colombian departments differ in their access to the markets. Moreover, as expected even though the time period covered in this analysis is thirty years, dispersion is persistent and there are no significant changes over the period under consideration. As for the relationship between the spatial distribution of per capita income and that of market potential the comparison of Figures 1.1 and 1.2 reveals a connection between the two magnitudes although it is far from perfect. In general high market potential departments are also departments with high per capita income levels. However, figures for some departments contradict this general statement. This is confirmed by the information provided in Figure 1.3. There is a positive relationship between market potential and per capital income levels across the three periods. However the amount of dispersion is far from negligible. It can be observed that are departments with similar low values of market potential (for instance Casanare and Arauca) but with very different per capita income levels.

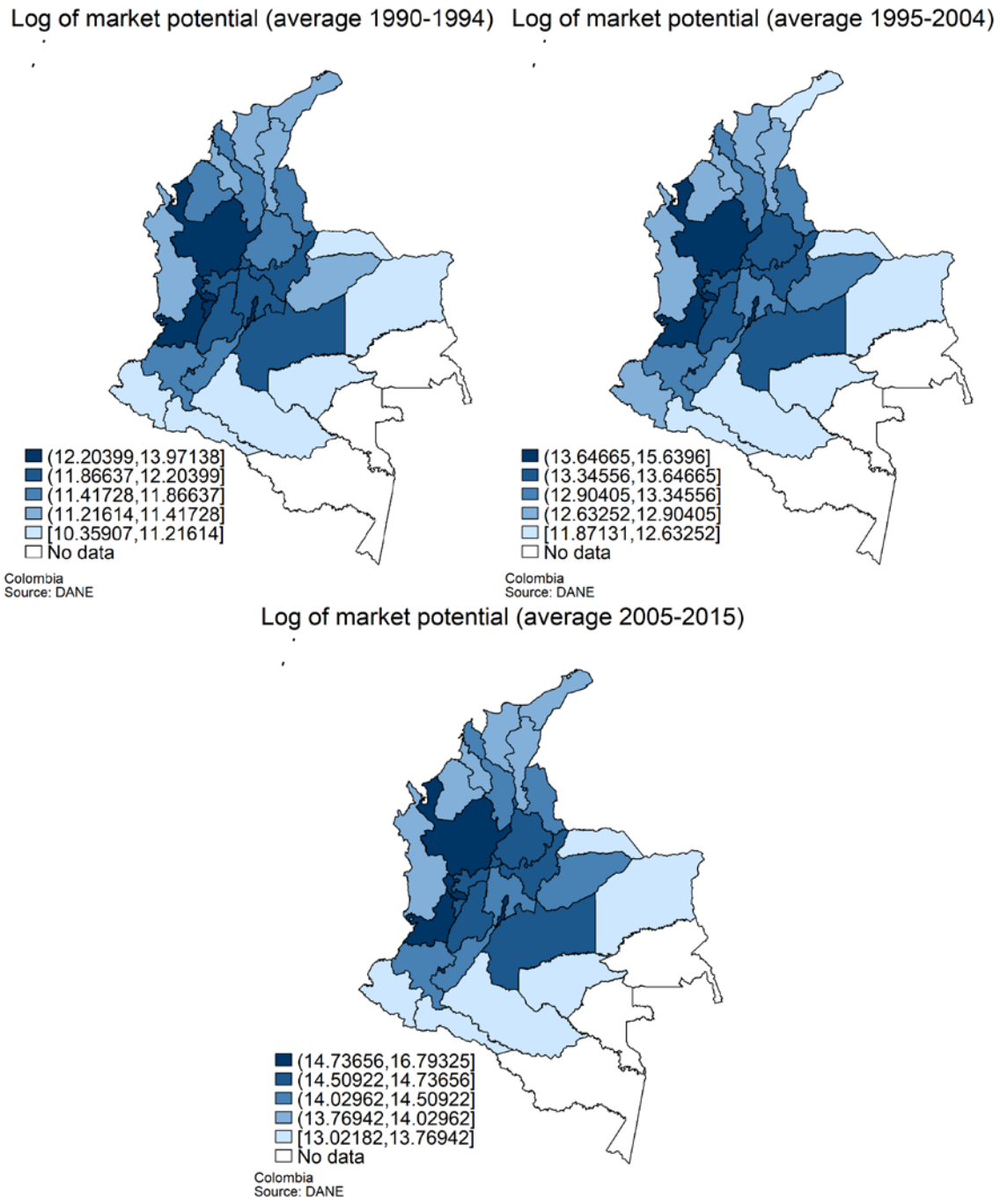
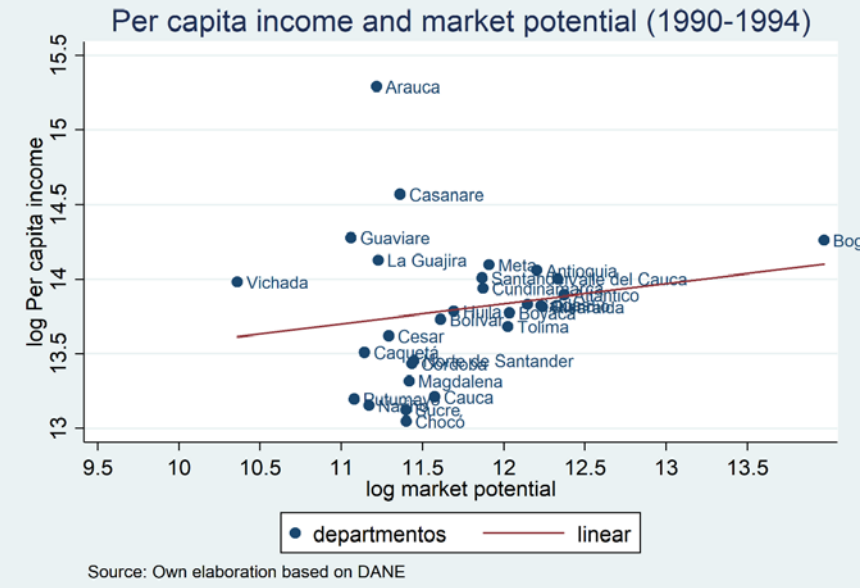
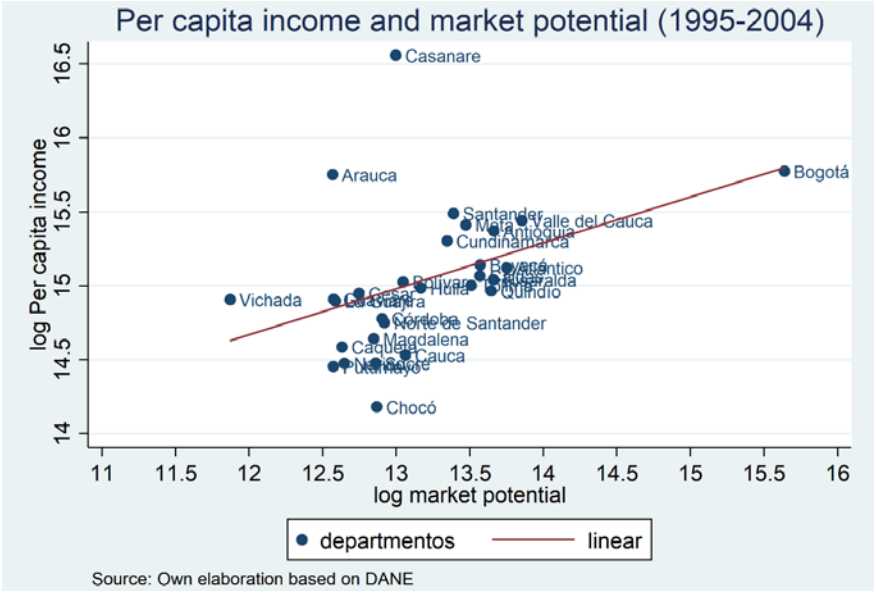


Figure 1.2. Spatial distribution of (the log of) market potential in Colombia.
Source: Colombian National Institute for Statistics (DANE) and authors' own calculations


Per capita income and market potential (1990-1994)
log Per capita income
15.5
15
14.5
14
13.5
13
Arauca
Casanare
Guaviare
La Guajira
Vichada
Meta
Antioquia
Valle del Cauca
Tolima
Cesar
Caquetá
Magdalena
Cauca
Sucre
Chocó
9.5
10
10.5
11
11.5
12
12.5
13
13.5
log market potential
departmentos
linear
Source: Own elaboration based on DANE


Per capita income and market potential (1995-2004)
log Per capita income
16.5
16
15.5
15
14.5
14
Casanare
Arauca
Bogotá
Santander
Meta
Valle del Cauca
Antioquia
Cundinamarca
Vichada
Quindío
Norte de Santander
Magdalena
Cauca
Chocó
11
11.5
12
12.5
13
13.5
14
14.5
15
15.5
16
log market potential
departmentos
linear
Source: Own elaboration based on DANE

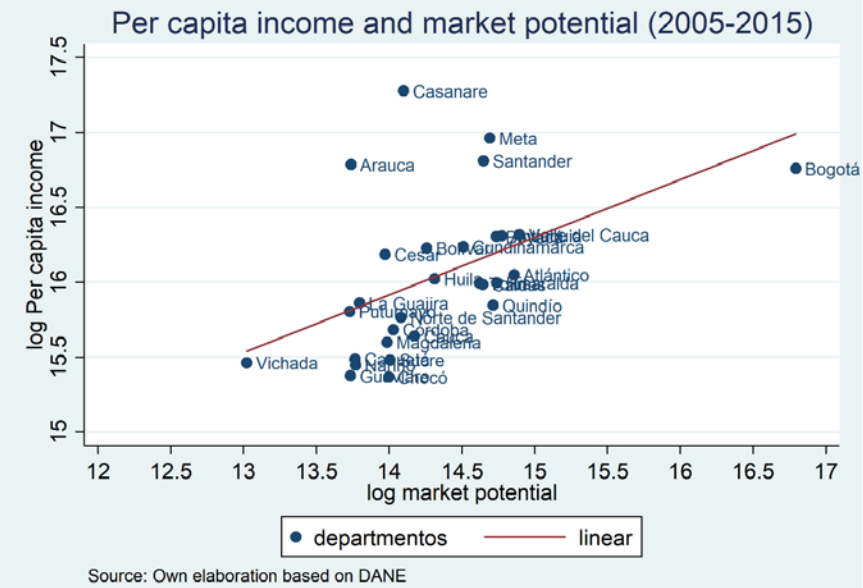
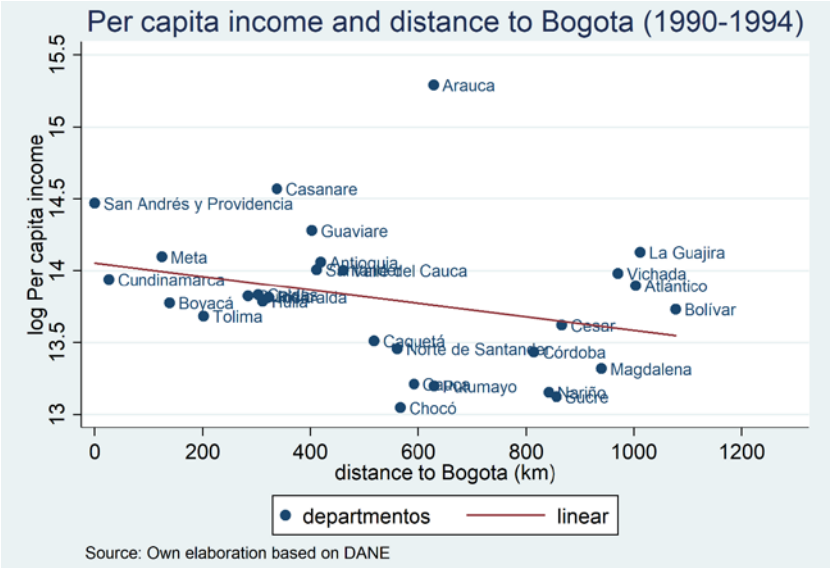
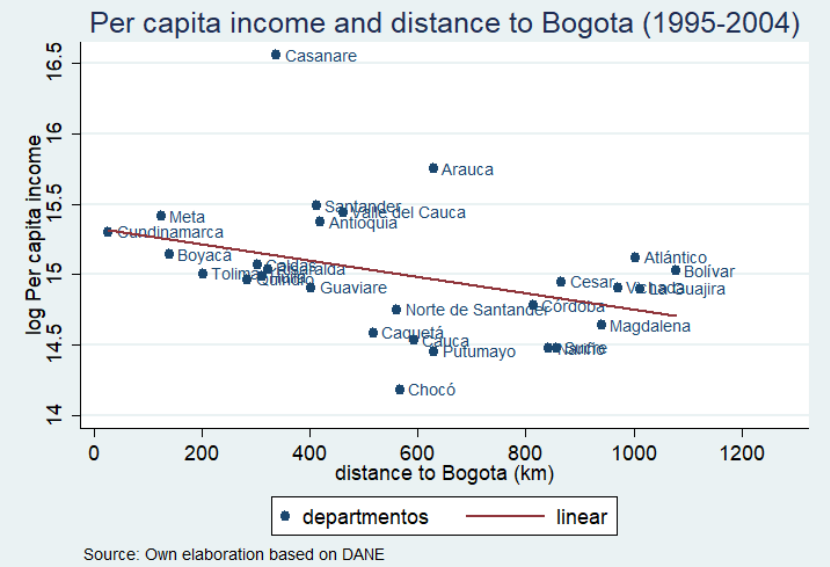


Figure 1.3. Income and market access in the Colombian departments.
Source: Colombian Institute for Statistics (DANE) and authors' own calculations.

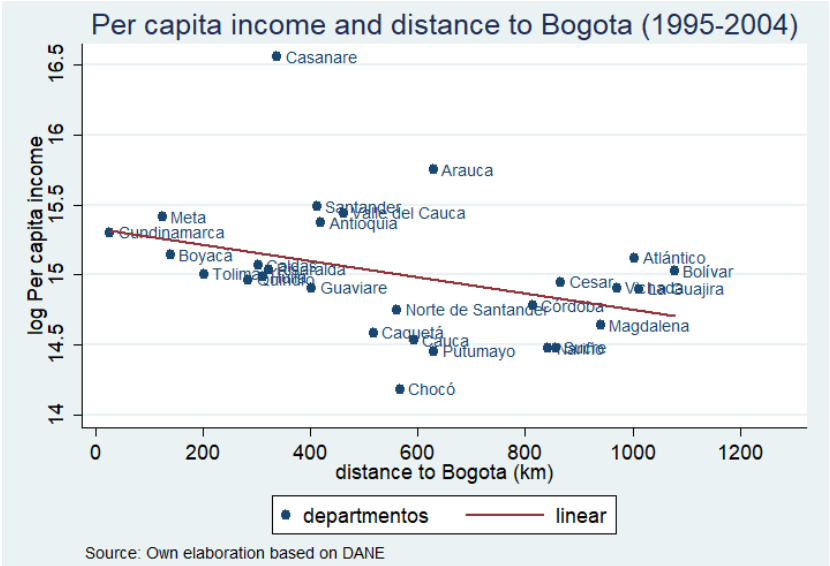
Another way of characterizing the spatial distribution of per capita income levels in Colombia is by plotting income levels against distance to Bogota. Figure 1.4 (1.5) plots the log of the average per capita income level across the periods 1990-1994, 1995-2004 and 2005-2015 against the distance to Bogota measured in kms (lorry travel times). A visual inspection to these graphs suffices to show that there is a clear per capita income gradient across the Colombian departments. The further and further away we are from Bogota the lower the average per capita income of the Colombian departments. Although it is far beyond the scope of this section, it seems that the income gradient becomes steeper in the last period 2005-2015 compared with the previous two ones. Taken into account that Bogota is more or less the geographical center of Colombia these graphs support the idea of a core-periphery spatial distribution of economic activities which is subject of research of many geographical economics models. In the next sections of this chapter, the research agenda we are going to carry out is going to be focus on the explanations of these observed per capita income differences across the Colombian departments in light of the contributions made by the geographical economics literature. More specifically our final goal will be to test one of the work horses of this literature which is the so called *nominal wage equation*.



Source: Own elaboration based on DANE



Source: Own elaboration based on DANE



Source: Own elaboration based on DANE

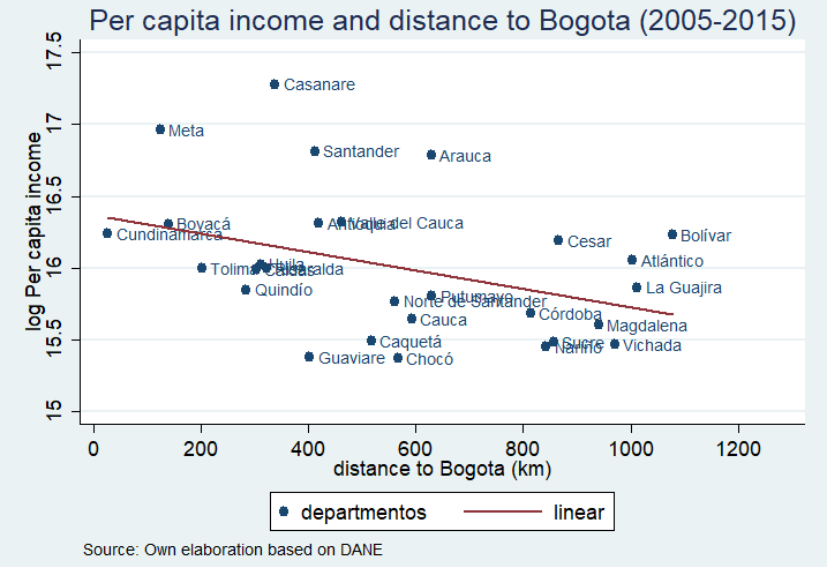
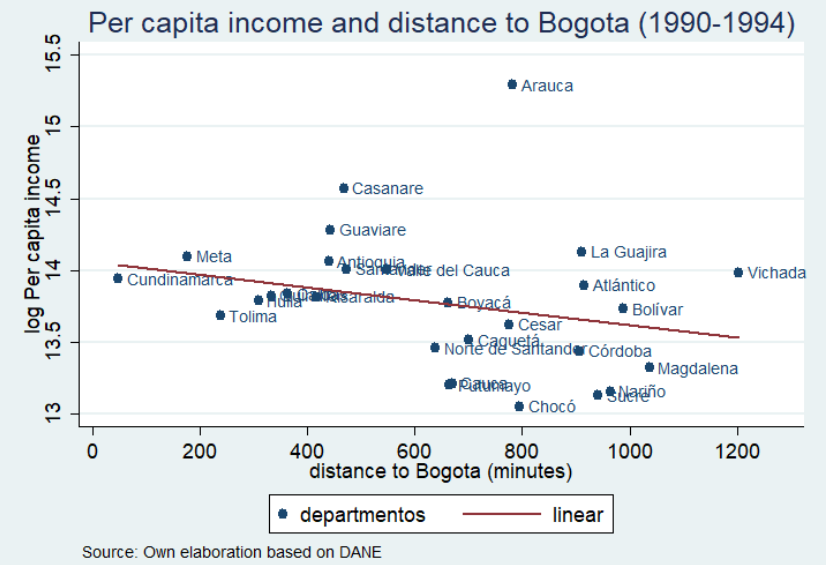
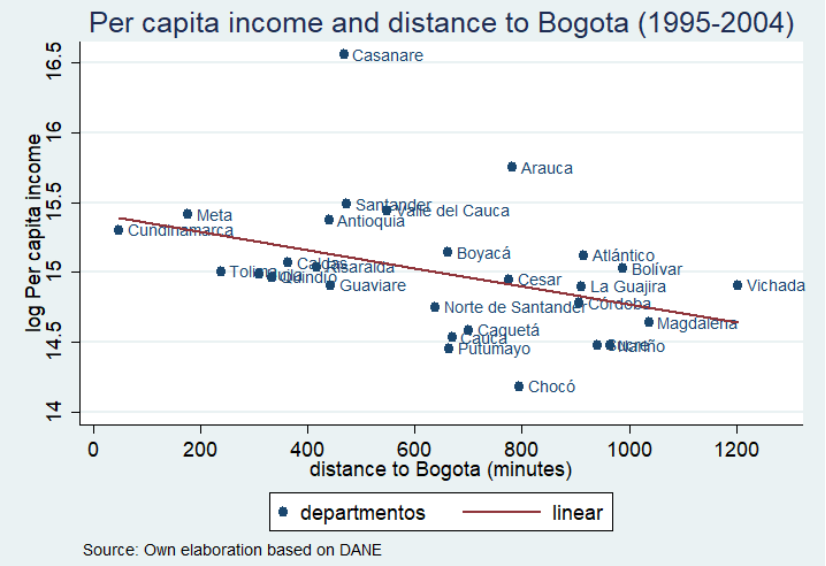


Figure 1.4. Income in the Colombian departments and distance to Bogota (km). Source: Colombian Institute for Statistics (DANE) and authors' own calculations.





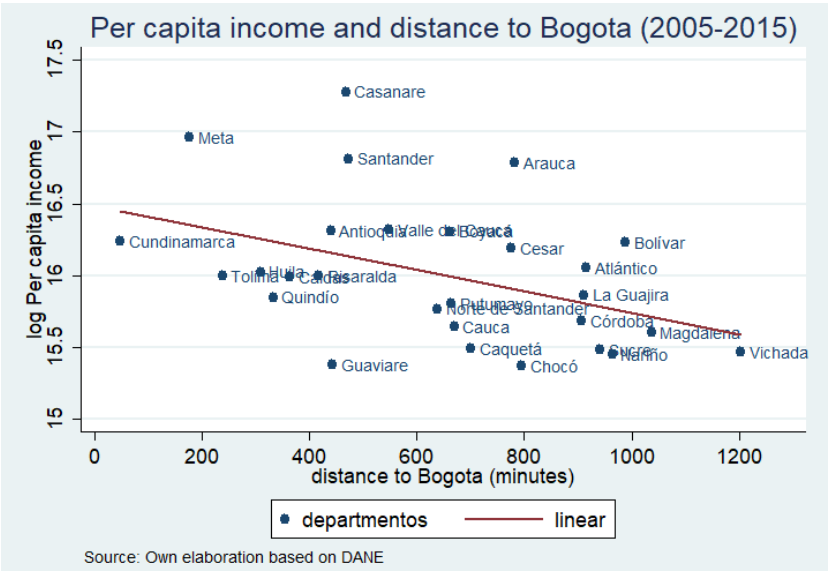


Figure 1.5. Income in the Colombian departments and distance to Bogota (minutes).
Source: Colombian Institute for Statistics (DANE) and authors' own calculations

Estimation of the Baseline Model

As a first step in our study of the robustness of the estimated impact of MP on the spatial distribution of income, we estimate a simple specification that will be used as a benchmark:

$$\ln Ypc = \alpha + \beta lnMP + \varepsilon \qquad (3)$$

where Ypc denotes the column vector with the per capita income values of the Colombian departments, MP denotes the column vector with the market potential values and ε is supposed to be (so far) a well-behaved error term. β is the parameter that captures the impact of MP on per capita income.

We begin by examining how much of the variation in Colombian "departments"[5] per capita income levels can be explained when only including information on market potential. We impose constant coefficients across the three time periods (1990–1994, 1995–2004, 2005–2015) to smooth variations introduced by short-run fluctuations in GDP. The results of the ordinary least square (OLS) estimates of the parameters in equation (3) are shown in Table 1.3. This provides the basis for our baseline estimation (OLS estimates) where we assume that the error term is uncorrelated with the explanatory variables (columns 1, 2 and 3). The results show that, with the exception of the estimates for the first period, the coefficient on market potential is both positive and statistically significant at the usual significance levels both for the period 1995-2004 and 2005-2015. In other words, market potential exerts a positive and significant impact of the expected average per capita income of the Colombian departments. Doubling market potential would increase per capita income between 31% and 38%. However, the models estimated in columns (1, 2 and 3) are marked by outlying observations as it can clearly be seen in the pictures in Figure 1.3. The outlying regions do not correspond with the spatial structure of per capita income levels determined by the majority of the observations. Outliers will seriously affect the coefficient estimates, if they are influential leverage points, i.e. outlying observations with regard to our market potential measure. In order to identify outliers we have compared the results obtained by using two statistics that measure influence (DFBETAs and Cooks distance). DFBETAs are perhaps the most direct influence measure of interest to model builders. DFBETAs focus on one coefficient and measure the difference between the regression coefficient when the ith observation is included and excluded, the difference being scaled by the estimated standard error of the coefficient.

Belsley et al. (1980) suggest observations with dfbetas > 2/Sqrt(N) should be checked as deserving special attention, but it is also common practice to use 1 (Bollen and Jackman 1990), meaning that the observation shifted the estimate at least one standard error. Cook's distance is another way of measuring influence. Cook's distance measures the aggregate change in the estimated coefficients when each observation is left out of the estimation. Values of Cook's distance that are greater than 4/N may be problematic. Therefore, in our case the observations with dfbetas greater than 2/ Sqrt(29)=0.37 or Cooks distance greater than 4/29=.13 are problematic. The results of these two statistics show that: 1) for the estimations of the period 1990-1994 Arauca (according to Cooks' distance) and Arauca and Vichada (according to dfbetas) are outliers, 2) for the period 1995-2004 Arauca and Casanare (according to Cooks' distance) and Arauca (according to dfbetas) are outliers and 3) for the period 2005-2015 Arauca and Bogota (according to dfbetas) and Arauca, Casanare and Bogota (according to Cooks distance). In order to control for the effects of the identified outlying observations we also reproduce the results of the estimations by dropping out Arauca and Vichada (1990-1994) and Arauca and Casanare (1995-2004 and 2005-2015). The results of these set of estimations are displayed in columns 4, 5 and 6 of Table 1.3. Comparing the results obtained dropping out the outliers with the ones obtained using all the observations we observe that a) the coefficient estimate on market potential for the first period under

[5]Colombia, according to the Colombian National Statistical Office (DANE) and to the 1991 Colombian constitution, is classified in a hierarchical system which divides the territory into departments, districts, municipalities and indigenous territories. The number of Colombian departments is 33, districts are 5 and municipalities and indigenous territories are 1122. This classification is known as political-administrative division (Division politico-administrativa, DIVIPOLA). For further details on the territorial classification see www.dane.gov.co and appendix A.

analysis (1990-1994) becomes positive and statistically significant; b) the elasticity estimates of per capita income with respect to market potential are higher. Now doubling market potential would increase per capita income between 27% and 46%; and c) the explanatory power of the regression has increased substantially across the three periods under study (from the range 18%-25% to 45%-50%). Looking across periods, there is a clear trend in the importance of access to markets. Access to markets seems to become more important in our last period compared with the previous two ones.

Finally, to control for the potential endogeneity between our dependent variable and the construction of the market potential variable we also present estimates using instrumental variables regression (IV estimates). These set of results are shown in Tables 1.4, 1.5 and 1.6. IV estimation is based on the existence of a set of instruments that are strongly correlated with the original endogenous variables but asymptotically uncorrelated with the error term. Furthermore, they should also be variables that are not driven by an unobservable third variable the authors suspect might be jointly affecting market potential and per capita income levels. Once these instruments are identified, they are used to build a proxy for the explanatory endogenous variables which consists of their predicted values in a regression on both the instruments and the exogenous variables. However, it is difficult to find such instruments because most socioeconomic variables are endogenous as well. In this chapter we propose to use mainly accessibility variables as instruments, since they are highly correlated with our market potential variable but also non contemporary correlated with the errors.

Following Breinlich (2006), we start by first instrumenting market potential with distance from Bogota and with the area size of the department (in km2). The first instrument capture market potential advantages of departments close to the geographic centre of Colombia. The second instrument captures the effect of large regional markets in the composition of domestic market potential. The results of the IV estimations using these two instruments are displayed in columns 1 and 2 of Tables 1.4, 1.5 and 1.6. Both distance from Bogota and area size of the department are significant in the first stage for the three periods, have the expected sign and explain between 44% and 49% of the variation in regional market potential (column 1 of Tables 1.4, 1.5 and 1.6). Turning to the second stage results (column 2 of Tables 1.4, 1.5 and 1.6), the coefficient estimate for market potential is not statistically significant for the period 1990-1994. However market potential retains its significance and its coefficient is again of similar magnitude as in the initial specifications for the periods 1995-2004 and 2005-2015. One could object to the above choice of instruments on several grounds. First, Bogota is a centroid of regional income's distribution within Colombia and distance to it might capture other determinants of income levels besides market potential. Likewise, the area size of a department could be inversely correlated with the density of regional economic activity.

Since we have an overidentified equation, with more excluded instruments than included endogenous variables, we may test whether the excluded instruments are appropriately independent of the error process. We perform the Sargan test to evaluate the validity of the instruments and the test of the model's overidentifying restrictions cannot reject the exogeneity of these variables at 1% and 5% significance levels (see column 1 of Tables 1.4, 1.5 and 1.6). However much higher p-values would be desired to avoid casting doubt on our choice of instruments. On account of the previous consideration we use two alternative

instruments. The first is average distance to all other departments (measured in km), the second is the average lorry travel times to all other departments (measured in minutes of travel times). The results of these alternative IV estimations show that when we instrument market potential with the average lorry travel times the elasticity of per capita income with respect to market potential becomes larger than when we instrument market potential with average distance measured in km no matter which period we are looking at. Moreover both approaches yield elasticities of per capita income with respect to market potential higher than both the "overidentified" IV regression estimates and the original OLS estimates. The only exception with lower elasticity of per capita income with respect to market potential estimates is the period (1995-2004) when using as instrument for market potential the average distance between departments measured in km.

Table 1.5. Results of the Estimation of the Baseline Model (OLS estimates)

Dependent variable	Log per capita income					
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors	(1990-1994)	(1995-2004)	(2005-2015)	(1990-1994)	(1995-2004)	(2005-2015)
Constant	12.22**	10.94**	10.52**	10.47**	9.88**	9.33**
	(1.64)	(1.66)	(1.84)	(1.38)	(1.01)	(1.32)
Log Market Potential (MP1) (alpm3y)	0.13	0.31**	0.38**	0.27**	0.38**	0.46**
	(0.14)	(0.12)	(0.12)	(0.11)	(0.07)	(0.09)
Period outliers	No	No	No	Yes	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.03	0.18	0.25	0.45	0.49	0.50
F-statistic	0.92 [0.34]	6.04** [0.02]	8.94** [0.00]	5.53** [0.02]	24.75** [0.00]	25.28** [0.00]
Breusch-Pagan test	5.71** [0.01]	2.25 [0.13]	0.77 [0.38]	6.09** [0.01]	1.85 [0.17]	0.84 [0.36]
Residuals Moran's I	0.245** [0.00]	0.264** [0.00]	0.195** [0.01]	0.148* [0.05]	0.086 [0.20]	-0.039 [0.99]
LM-ERR	5.11** [0.02]	5.94** [0.01]	3.23* [0.07]	1.73 [0.19]	0.58 [0.44]	0.11 [0.73]
LM-LAG	4.55** [0.03]	3.69* [0.05]	1.54 [0.21]	2.67 [0.10]	0.88 [0.34]	0.04 [0.82]
Robust LM-ERR	0.97 [0.32]	3.01 [0.08]	2.43 [0.11]	1.10 [0.29]	0.005 [0.94]	0.07 [0.78]
Robust LM-LAG	0.40 [0.52]	0.76 [0.38]	0.74 [0.39]	2.03 [1.15]	0.30 [0.58]	0.005 [0.94]
Observations	29	29	29	27	27	27

Note: Table displays coefficients and standard errors for OLS estimates. The dependent variable is the log of the average per capita income of the period (1990-1994, 1995-2004, 2005-2015). Regressors are the log of MP (MP stands for market potential which is computed as a weighted average of departments GDP values where the weighting scheme is the distance measure in kms between departments). LM-ERR Lagrange Multiplier test for spatial error dependence; LM-LAG Lagrange Multiplier test for spatial lag dependence. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

Table 1.5 also includes results for some diagnostic checks. We check for the presence of heteroskedasticity by means of the Breusch-Pagan test. Breusch-Pagan tests the null hypothesis that the error variances are all equal versus the alternative that the error variances are a multiplicative function of one or more variables. A large chi-square would indicate that

heteroskedasticity was present. In our case the chi-square values are small in the estimations of the periods 1995-2004 and 2005-2015, indicating that heteroskedasticity was not a problem. However it could be a problem for the estimations of the first period (1990-1994) where the value obtained for the statistic (5.71) and the associated p-value [0.01] indicates that the null of homoscedastic errors is rejected at the 5% significance level. The battery of spatial dependence tests reveals that the baseline per capita income market access model is likely to be (spatially) misspecified in the estimations which use the full sample of observations. The results of these spatial dependence tests will be discussed in detail in the fourth section. Before doing so, we frame the results of the baseline specification within a core-periphery geographical economics model.

Table 1.6. Results of the Estimation of the Baseline Model, 1990-1994 (IV estimates)

Dependent variable	Log MP1	Log per capita income	Log MP1	Log per capita income	Log MP1	Log per capita income
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors	(1990-1994)	(1995-2004)	(2005-2015)	(1990-1994)	(1995-2004)	(2005-2015)
Constant	15.48**	10.29**	24.39**	9.93**	26.54**	9.71**
	(0.89)	(1.66)	(2.91)	(2.04)	(3.00)	(1.91)
Log Market Potential (MP1)		0.29		0.32*		0.34**
		(0.25)		(0.17)		(0.16)
Log (dBogota)	-0.22**					
	(0.08)					
Log (a)	-0.24					
	(0.07)					
Log(dmediak)			-1.95**			
			(0.45)			
Log(dmediat)					-2.25**	
					(0.45)	
Period outliers	Yes	Yes	yes	Yes	Yes	Yes
Estimation	OLS	IV	OLS	IV	OLS	IV
R2	0.44	0.12	0.42	0.17	0.49	0.17
F-Stat (and p-value) based on exclude instruments	9.16 [0.00]		18.81[0.00]		24.17[0.00]	
Sargan test (and p-value)	2.82[0.09]		Exactly identified		Exactly identified	
Observations	26	26	27	27	27	27

Note: Table displays coefficients and t-statistics for IV estimation. The dependent variable in columns 2, 4 and 6 is the log of the average per capita income of the period (1990-1994). The independent variable in these columns is the log of market potential (MP1, see text for details of calculation). Instruments for MP used are distance to Bogota (in km) and area size of a department in km2 (column 1), average distance to other departments in km (column 3), and average lorry travel times to other departments in minutes (column 5). Columns 1, 3 and 5 display the corresponding first-stage results. For data sources see appendix A. * and ** signify statistical significance at the10% and 5% levels.

Table 1.7. Results of the Estimation of the Baseline Model, 1995-2004 (IV estimates)

Dependent variable	Log MP1	Log per capita income	Log MP1	Log per capita income	Log MP1	Log per capita income
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors						
Constant	17.39** (0.92)	10.03** (2.01)	26.41** (2.53)	10.14** (1.35)	27.29** (2.56)	9.65** (1.325)
Log Market Potential (MP1)		0.37** (0.15)		0.36** (0.10)		0.40** (0.10)
Log (dBogota)	-0.27** (0.08)					
Log (a)	-0.25** (0.07)					
Log(dmediak)			-2.04** (0.39)			
Log(dmediat)					-2.14** (0.38)	
Period outliers	Yes	Yes	yes	Yes	Yes	Yes
Estimation	OLS	IV	OLS	IV	OLS	IV
R2	0.49	0.37	0.52	0.49	0.55	0.49
F-Stat (and p-value) based on exclude instruments	11.08 [0.00]		27.25[0.00]		30.32[0.00]	
Sargan test (and p-value)	3.78[0.05]		Exactly identified		Exactly identified	
Observations	26	26	27	27	27	27

Note: Table displays coefficients and t-statistics for IV estimation. The dependent variable in columns 2, 4 and 6 is the log of the average per capita income of the period (1995-2004). The independent variable in these columns is the log of market potential (MP1, see text for details of calculation). Instruments for MP1 used are distance to Bogota (in km) and area size of a department in km2 (column 1), average distance to other departments in km (column 3), and average lorry travel times to other departments in minutes (column 5). Columns 1, 3 and 5 display the corresponding first-stage results. For data sources see appendix A. * and ** signify statistical significance at the10% and 5% levels.

Table 1.8. Results of the Estimation of the Baseline Model, 2005-2015 (IV estimates)

Dependent variable	Log MP1	Log per capita income	Log MP1	Log per capita income	Log MP1	Log per capita income
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors						
Constant	18.30** (0.90)	9.65** (2.67)	27.22** (2.52)	8.81** (1.78)	28.19** (2.53)	8.18** (1.325)
Log Market Potential (MP1)		0.44** (0.18)		0.49** (0.12)		0.54** (0.12)
Log (dBogota)	-0.27** (0.08)					
Log (a)	-0.24** (0.07)					
Log(dmediak)			-1.98** (0.38)			
Log(dmediat)					-2.10** (0.38)	
Period outliers	Yes	Yes	yes	Yes	Yes	Yes
Estimation	OLS	IV	OLS	IV	OLS	IV
R2	0.48	0.44	0.52	0.50	0.54	0.49

F-Stat (and p-value) based on exclude instruments	10.48 [0.00]		26.11[0.00]		29.99[0.00]	
Sargan test (and p-value)	4.28[0.04]		Exactly identified		Exactly identified	
Observations	26	26	27	27	27	27

Note: Table displays coefficients and t-statistics for IV estimation. The dependent variable in columns 2, 4 and 6 is the log of the average per capita income of the period (2005-2015). The independent variable in these columns is the log of market potential (MP1, see text for details of calculation). Instruments for MP1 used are distance to Bogota (in km) and area size of a department in km2 (column 1), average distance to other departments in km (column 3), and average lorry travel times to other departments in minutes (column 5). Columns 1, 3 and 5 display the corresponding first-stage results. For data sources see appendix A. * and ** signify statistical significance at the10% and 5% levels.

For comparison purposes, Table 1.9, 1.10 and 1.11 reports results for three more ad hoc measures of market potential often used in geography and in the earlier economics literature. The first is a somewhat more sophisticated version of the Harris' market potential, using lorry travel times (expressed in minutes of travel) as weights instead of physical distances (we label it as MP3). These travel times between the capitals cities of the departments have been obtained from Google maps taken the option which reports the fastest route (the ones with the lowest time assigned)[6]. The other two are versions of the Harris' market potential which use as a measure of economic activity the population of each Colombian department instead of using GVA and use both travel times and physical distances as weights. We have labelled these alternative measures as MP4 (weights are physical distances) and MP6 (weights are travel times).

Table 1.9. Results of the Estimation of the Baseline Model (OLS estimates)

Dependent variable	Log per capita income					
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors	(1990-1994)	(1995-2004)	(2005-2015)	(1990-1994)	(1995-2004)	(2005-2015)
Constant	12.42**	11.32**	11.14**	10.47**	10.00**	9.65**
	(1.56)	(1.60)	(1.79)	(1.38)	(0.96)	(1.27)
Log Market Potential	0.11	0.28**	0.34**	0.26**	0.37**	0.44**
(MP2)	(0.13)	(0.12)	(0.12)	(0.11)	(0.07)	(0.08)
Period outliers	No	No	No	Yes	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.02	0.16	0.23	0.18	0.51	0.49
F-statistic	0.77	5.36**	7.46**	5.37**	26.44**	24.35**
	[0.38]	[0.02]	[0.01]	[0.02]	[0.00]	[0.00]
Breusch-Pagan test	5.71**	3.64*	1.72	6.70**	2.27	0.71
	[0.01]	[0.05]	[0.19]	[0.00]	[0.13]	[0.36]
Observations	29	29	29	27	27	27

Note: Table displays coefficients and standard errors for OLS. The dependent variable is the log of the average per capita income of the period (1990-1994, 1995-2004, 2005-2015). Regressors are the log of MP2 (MP2 stands for market potential which is computed as a weighted average of departments GDP values where the weighting scheme is the distance measure in lorry travel time between departments). Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

[6] Google maps reports several alternatives to drive between locations depending on whether we want to avoid tolls, go on highways, etc.. We have chosen the shortest distance (in travel times) between the capital cities for each pair of departments.

Table 1.10. Results of the Estimation of the Baseline Model (OLS estimates)

Dependent variable	Log per capita income					
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors	(1990-1994)	(1995-2004)	(2005-2015)	(1990-1994)	(1995-2004)	(2005-2015)
Constant	13.83**	12.54**	12.62**	11.28**	10.48**	10.63**
	(1.78)	(1.75)	(1.82)	(1.66)	(1.16)	(1.39)
Log Market Potential	-0.00	0.21	0.28*	0.20	0.38**	0.44**
(MP4)	(0.15)	(0.15)	(0.15)	(0.14)	(0.09)	(0.11)
Period outliers	No	No	No	Yes	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.00	0.07	0.11	0.08	0.37	0.37
F-statistic	0.00	2.02	3.51*	2.18	14.88**	14.70**
	[0.98]	[0.16]	[0.07]	[0.15]	[0.00]	[0.00]
Breusch-Pagan test	5.60**	3.64*	3.31*	7.29**	2.38	0.04
	[0.01]	[0.05]	[0.07]	[0.00]	[0.12]	[0.83]
Observations	29	29	29	27	27	27

Note: Table displays coefficients and standard errors for OLS. The dependent variable is the log of the average per capita income of the period (1990-1994, 1995-2004, 2005-2015). Regressors are the log of MP4 (MP4 stands for market potential which is computed as a weighted average of departments population values where the weighting scheme is the distance measure in kms between departments). Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

Looking across the coefficients in columns 1 to 6 of Tables 1.7, 1.8 and 1.9 reveal that the three measures yield qualitatively as well as quantitatively similar results. According to the reported R2s they also explain an equal proportion of income variations in the Colombian departments.

Table 1.11. Results of the Estimation of the Baseline Model (OLS estimates)

Dependent variable	Log per capita income					
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors	(1990-1994)	(1995-2004)	(2005-2015)	(1990-1994)	(1995-2004)	(2005-2015)
Constant	14.06**	13.03**	13.31**	11.54**	10.70**	11.08**
	(1.64)	(1.64)	(1.71)	(1.55)	(1.08)	(1.32)
Log Market Potential	-0.02	0.17	0.23	0.19	0.36**	0.41**
(MP6)	(0.14)	(0.14)	(0.14)	(0.13)	(0.09)	(0.11)
Period outliers	No	No	No	Yes	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.00	0.05	0.08	0.07	0.38	0.35
F-statistic	0.03	1.48	2.54	1.99	15.36**	13.52**
	[0.87]	[0.23]	[0.12]	[0.17]	[0.00]	[0.00]
Breusch-Pagan test	6.13**	5.78**	4.49**	7.43**	2.60	0.14
	[0.01]	[0.01]	[0.03]	[0.00]	[0.10]	[0.71]
Observations	29	29	29	27	27	27

Note: Table displays coefficients and standard errors for OLS. The dependent variable is the log of the average per capita income of the period (1990-1994, 1995-2004, 2005-2015). Regressors are the log of MP6 (MP6 stands for market access which is computed as a weighted average of departments population values where the weighting scheme is the distance measure in travel times between departments). Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

1.3 Geographical economics explanation: Income and Geography

The theoretical framework is a reduced version of a standard geographical economics model (multi-regional version of Krugman, 1991b). We consider a regional setting composed of R locations and we focus on the analysis of the manufacturing sector. In this sector, firms produce a great number of varieties of a homogenous differentiated good (D) under increasing returns to scale and monopolistic competition. Firms face transport costs in an iceberg form in order to receive one unit of the differentiated good at location j from location i , $T_{i,j} > 1$ units must be shipped from i, so $T_{i,j} - 1$ measures the fraction of good that is melted in transit from i to j . The manufacturing sector can produce the differentiated good in different locations.

On the demand side, the final demand in location j can be obtained via utility maximization of the corresponding CES utility function:

$$Max_{mi,j(z)}\ D_j = \left[\sum_{i=1}^{R}\int_0^{n_i} m_{i,j}(z)^{\sigma-1/\sigma}\,dz\right]^{\sigma/\sigma-1} \text{ s.t.} \qquad \sum_{i=1}^{R} n_i x_{ij}^D p_{ij} = E_j \qquad (4)$$

where D_j represents the consumption of the differentiated good in location j . D is an aggregate of industrial varieties where $m_{i,j}(z)$ means the consumption of the each available variety z in location j and produce in location i and n_i is the number of varieties produced in location i. σ represents the elasticity of substitution among the varieties of the differentiated good where $\sigma > 1$, p_{ij} ($p_{ij} = p_i T_{ij}$), is the price of varieties produced in location i and sold in j and E_j represents the total income in location j . The consumer's problem solution gives the final demand in location j for each variety produce in location i.

$$x_{ij}^D = p_{ij}^{-\sigma}\left[\sum_{n=1}^{R} n_n p_{nj}^{1-\sigma}\right]^{-1} Y_j \qquad (5)$$

If we define a price index for the differentiated goods[7] as $P_j \equiv \left[\sum_{n=1}^{R} n_n p_{nj}^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$, final demand in location j can be written as $x_{ij}^{consD} = p_{ij}^{-\sigma} P_j^{\sigma-1} E_j$. However, in order for x_{ij}^{consD} units of consumption to arrive at location j , $T_{i,j} x_{ij}^{consD}$ must be shipped. So the effective demand a firm in location i faces from a consumer in location j is given by:

$$x_{ij}^D = T_{ij} p_{ij}^{-\sigma} P_j^{\sigma-1} E_j = p_i^{-\sigma} T_{ij}^{1-\sigma} P_j^{\sigma-1} E_j \qquad (6)$$

On the supply side a typical firm in location i maximizes the following profit function:

[7] This Industrial Price Index in location j measures the minimum costs of purchasing a unit of the composed index of manufacturing goods D so it can be interpreted as an expenditure function.

$$\prod_i = \sum_{j=1}^{R} \frac{p_{ij} x_{ij}^D}{T_{i,j}} - w_i^D (F + c x_i^D) \qquad (7)$$

Technology in the increasing returns to scale manufacturing sector is given by the usual linear cost function: $l_{Dij} = F + c x_{ij}^D$, where l_{Dij}, represents the industrial workers used for the production of a variety in location i and sold in location j, F, represents a fixed cost of production, c, is the variable unit cost and x_{ij}^D is the amount of the differentiated good demanded in location j and produced in location i ($x_i^D \equiv \sum_j x_{ij}^D$ represents the total amount of output produced by the firm in location i and sold in the different j locations) and w_i^D is the nominal wage paid to the manufacturing workers in location i. First order conditions for profit maximization give the standard result that prices are set at a markup $\frac{\sigma}{\sigma - 1}$ over marginal costs. At this price profits will be $\prod_i = \left(w_i^D\right)\left[\frac{c x_i^D}{\sigma - 1} - F\right]$. Free of entry assures that in the long run firms break even, implying that $x_i^D = \bar{x} = \frac{F(\sigma - 1)}{c}$ .The price that is needed to sell this amount of output is $P_i^{\sigma} = \frac{1}{\bar{x}} \sum_{j=1}^{R} E_j P_j^{\sigma - 1} T_{i,j}^{1-\sigma}$. If we combine this expression with the fact that in equilibrium prices are a constant markup over marginal costs, the following zero-profit condition can be obtained:

$$w_i^D = \left(\frac{\sigma - 1}{\sigma c}\right)\left[\frac{1}{\bar{x}} \sum_{j=1}^{R} E_j P_j^{\sigma - 1} T_{i,j}^{1-\sigma}\right]^{1/\sigma} \qquad (8)$$

This equation is the so-called *nominal wage equation* in the Geographical economics literature. Equation (5) shows that the nominal wage level at location i depends on a weighted sum of the purchasing power of the surrounding locations where the weighted scheme is a distance function that decreases as the distance between i and j increases. If we normalize output production choosing our units in such a way that $c = \frac{(\sigma - 1)}{\sigma}$, and we set the fixed input requirement as $F = \frac{1}{\sigma}$, and define market potential in location i as $MP_i = \sum_{j=1}^{R} E_j P_j^{\sigma - 1} T_{i,j}^{1-\sigma}$, we can rewrite the *nominal wage equation* as:

$$w_i^D = \left[MP_i\right]^{1/\sigma} \qquad (9)$$

The meaning of this equation is that those firms in locations that have a good access to big markets (high market potential) will tend to remunerate their local factors of production (workers) with better salaries due to their savings in transportation costs.

1.4 Robustness Checks

Panel Data

In this section we will proceed to the estimation of the nominal wage equation using the panel data structure of our dataset. Besides the increase in the number of observations using the panel data set allows controlling for unobservable regional effects that could be shaping the spatial distribution of income across the Colombian departments. The baseline specification is the one derived in Section 2 (equation 9). We proxy wages (the price of the immobile factor) by gross value added per head (lnYrpc, measured in millions of 2005 constant pesos) again obtained from the regional database we have built. Introducing a time dimension into the estimation equation then yields the following specification:

$$\ln Yrpc_{it} = \alpha_t + \beta_1 lnMP_{it} + \varepsilon_{it} \qquad (10)$$

Table 1.12 reports the average coefficient on market potential for this baseline specification estimated by OLS for the sample pooled across all three periods (1990-2015) and using the six market potential measures defined previously (MP1, MP2, ….MP6) . Table 1.12 contains six columns where the first one corresponds to the regression of log of regional real per capita GDP against MP1, the second column corresponds to the regression of log of regional real per capita GDP against MP2 and so on until the last column (column six) which corresponds to the regression of log of regional real per capita GDP against MP6.

It can be seen that all the coefficient estimates of market potential are both significant and highly economic significant at the usual standard significant levels and therefore the results are in line with the theoretical predictions of the core-periphery geographical economics model. However there is a sizeable difference in the values of the coefficient estimates depending on the definition of market potential used in the regression. When market potential is measure using as a proxy for the economic mass of each Colombian department the figures of gross domestic product (GDP) and a distance matrix expressed in kms such as the case with MP1 and MP2 the elasticity estimates of real per capita income with respect to market potential are much higher than for the other alternative definitions of market potential (MP2, MP3…MP6). In this first two estimations (columns 1 and 2), on average doubling market potential increases real per capita income by around 36%-39%, whereas with the alternative definitions the increase is much more modest, on the range 16%-24%. Moreover MP1 and MP2 market potential definitions are more aligned with the theoretical derivation of market potential in the core-periphery geographical economics model and so this fact also explains that the explanatory power of the regressions in columns 1 and 2 is much higher than in the regressions of columns 3 to 6 where the R2 are between 10%-14%.

Table 1.12. Results of the Estimation of the Baseline Model (1990-2015) (pooled OLS estimates)

Dependent variable	Log Yrpc					
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors						
Constant	10.95** (0.24)	10.72** (0.28)	13.57** (1.71)	13.25** (0.37)	12.79** (0.43)	13.77** (0.38)
Log MP1	0.36** (0.018)					
Log MP2		0.39** (0.022)				
Log MP3			0.16** (0.028)			
Log MP4				0.20** (0.028)		
Log MP5					0.24** (0.034)	
Log MP6						0.16** (0.03)
year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.46	0.42	0.14	0.12	0.12	0.10
F-statistic	18.20 [0.00]	14.64 [0.23]	3.10 [0.12]	3.72 [0.17]	3.75** [0.00]	2.92** [0.00]
Observations	676	676	728	754	754	754

Note: Table displays coefficients and Huber-White standard errors for OLS. The dependent variable is the log of per capita income (in 2005 Colombian pesos). log of MP1 to log of MP6 are the logs of the different definitions of market potential. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

We have broken down the estimation of equation (10) for the whole period into the three sub-periods (1990-1994), (1995-2004) and (2005-2015) we have used previously and the result of the estimates are reported in Tables 1.13, 1.14 and 1.15. The estimation of the nominal wage equation by sub-periods in comparison with the estimation for the whole period shows that a change in the size (general increase) of the coefficient estimates for market potential from 1995 onwards can be observed. This fact becomes quite clear for the estimations carried out for the period 2005-2015 where the elasticity of per capita income with respect to MP1 and MP2 has moved from values in the range 0.36-0.39 (1990-2015) to values in the range 0.50-0.57 (2005-2015) which means an increase in the elasticity estimates of around 39%-46%. This trend of increase in the size of the elasticity estimates also applies to the other alternative definitions of market potential. By contrast the market potential estimates obtained for the regressions carried out for the period 1990-1994 where statistically insignificant although positive for all market potential definitions but MP1 and MP2 which are shown in Table 1.13. In this case we see that for this sub-period there is a significant drop in the market potential coefficient estimates both with respect to the average coefficient obtained for the whole period as well as for the coefficients obtained for the sub-periods 1995-2004 and 2005-2015.

Table 1.13. Results of the Estimation of the Baseline Model (1990-1994) (pooled OLS estimates)

Dependent variable	Log Yrpc	
	(1)	(2)
Regresors		
Constant	12.87**	10.72**
	(0.36)	(0.28)
Log MP1	0.21**	
	(0.02)	
Log MP2		0.23**
		(0.03)
year dummies	Yes	Yes
Estimation	OLS	OLS
R2	0.29	0.26
F-statistic	12.09	10.34
	[0.00]	[0.23]
Observations	130	130

Note: Table displays coefficients and Huber-White standard errors for OLS. The dependent variable is the log of per capita income (in 2005 Colombian pesos). log of MP1 to log of MP2 are the logs of the different definitions of market potential. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

Table 1.14. Results of the Estimation of the Baseline Model (1995-2004) (pooled OLS estimates)

Dependent variable	Log Yrpc					
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors						
Constant	10.62**	10.27**	13.01**	13.13**	12.79**	13.63**
	(0.30)	(0.33)	(0.54)	(0.55)	(0.43)	(0.58)
Log MP1	0.37**					
	(0.02)					
Log MP2		0.41**				
		(0.02)				
Log MP3			0.20**			
			(0.028)			
Log MP4				0.21**		
				(0.04)		
Log MP5					0.25**	
					(0.05)	
Log MP6						0.17**
						(0.04)
year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.50	0.45	0.14	0.12	0.06	0.10
F-statistic	29.49	28.08	3.15	2.44	2.51**	1.46**
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	260	260	280	290	290	290

Note: Table displays coefficients and Huber-White standard errors for OLS. The dependent variable is the log of per capita income (in 2005 Colombian pesos). log of MP1 to log of MP6 are the logs of the different definitions of market potential. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

Table 1.15. Results of the Estimation of the Baseline Model (2005-2015) (pooled OLS estimates)

Dependent variable	Log Yrpc					
	(1)	(2)	(3)	(4)	(5)	(6)
Regresors						
Constant	8.49** (0.30)	7.70** (0.45)	11.65** (0.59)	12.21** (0.46)	11.47** (0.54)	12.89** (0.49)
Log MP1	0.50** (0.03)					
Log MP2		0.57** (0.03)				
Log MP3			0.29** (0.04)			
Log MP4				0.29** (0.04)		
Log MP5					0.35** (0.04)	
Log MP6						0.23** (0.04)
year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.48	0.45	0.16	0.14	0.15	0.12
F-statistic	23.47 [0.00]	29.04 [0.00]	5.21 [0.00]	6.59 [0.00]	6.74** [0.00]	4.21** [0.00]
Observations	286	286	308	319	319	319

Note: Table displays coefficients and Huber-White standard errors for OLS. The dependent variable is the log of per capita income (in 2005 Colombian pesos). log of MP1 to log of MP6 are the logs of the different definitions of market potential. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

However a more robust way of testing if the effect of access to markets (market potential) has changed over a certain period of time, in our case over the course of this 35 years, is by estimating the following regression model:

$$\ln Ypc_{it} = \alpha_t + \beta_1 lnMP_{it} + \beta_2 d94 lnMP_{it} + \beta_3 d04 lnMP_{it} + \beta_4 d15 lnMP_{it} + \varepsilon_{it} \quad (11)$$

Where we interact the year dummy variables d94, d04 and d15 with our market potential index for the last year of the three periods we have defined previously (1994, 2004 and 2015). The coefficients β_2, β_3 and β_4 in equation (11) measure how a one-percent increase in market potential has changed over the 4-14 and 35- year period the impact of the average market potential estimate given by β_1. The results of the estimation of equation (11) using the definitions of market potential which are closer to the theory-based measure derived in the theoretical part of the chapter, i.e., MP1 and MP2 are shown in the following Table 1.16.

The results for the potential differential impact of market potential on per capita income levels over the three sub-periods in which we have broken down the whole time frame (1990-2015) show that when we use the market potential definition MP1, the coefficient estimate of the interaction term of the first time period 1990-1994 with market potential (MP1) is negative and statistically significant at 10% level. So this result is in line with a change in the

slope coefficient of market potential (MP1) over this period of time from the average value of 0.36 to 0.26. The coefficient estimates of the interaction terms of the other sub-periods when market potential is defined as MP1 are positive but they are not statistically significant. When the definition of market potential used to estimate equation (11) is MP2 the coefficient estimate for the interaction terms is negative and statistically significant for the first period , 1990-1994 (10% significance level) whereas it becomes positive and statistically significant at 5% significance level for the last period, 2005-2015. According to these results there are serious doubts on the constancy of the effect of market potential on per capita income levels over the period 1990-2015. It seems that the effect of a ten percent increase in market potential on the average per capita income decreases by 1% point over this 4 year period not matter which market potential definition we are using. Finally, the effect of a ten percent increase in market potential (MP2) on the average per capita increased by 2.5% over the 35 year period of our sample data set.

Table 1.16. Results of the estimation of the baseline model with year dummy interactions (1990-2015) (pooled OLS estimates)

Dependent variable	Log Yrpc	
	(1)	(2)
Regresors		
Constant	10.96**	10.73**
	(0.25)	(0.30)
Log MP1	0.36**	
	(0.019)	
Log MP2		0.39**
		(0.02)
d94MP1	-0.10*	-0.11*
	(0.019)	(0.06)
d04MP1	0.05	0.05
	(0.07)	(0.08)
d15MP1	0.20	0.25**
	(0.12)	(0.12)
year dummies	Yes	Yes
Estimation	OLS	OLS
R2	0.47	0.43
F-statistic	17.21	14.51
	[0.00]	[0.23]
Observations	676	676

Note: Table displays coefficients and Huber-White standard errors for OLS. The dependent variable is the log of per capita income (in 2005 Colombian pesos). lof of MP1 and log of MP2 are the logs of the market potential definitions MP1 and MP2. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

Robustness checks

The pooled OLS estimates of the specifications (10) and (11) are subject to a number of concerns. In order to have a consistent estimator of β_1 (elasticity of per capita income with respect to market potential) we have to assume that the many potential unobserved time-constant factors[8] which are difficult to control for in our estimations and which are affecting the level of per capita income across the Colombian departments are uncorrelated with market potential. However, holding this assumption in the context of the estimation of a relationship between per capita income and market potential is not very reasonable and therefore the pooled OLS method does not solve the omitted variables problem we are mentioning and therefore the estimates are biased and inconsistent. For instance shocks to $Yrpc_{it}$ as captured by ε_{it} are likely to be correlated across regions which in the end raises the issue that ε_{it} is also correlated with MP_{it}. Variables like institutional quality, climatic and another amenities of region, historical factors, and geographical features related to regions, etc. can be considered as additional determinants of income levels. In the case of Colombia it is well know that these factors vary across the different departments[9] and to a certain extent they are likely to be correlated across space. For these particular cases in which it is reasonable to assume the existence of unobserved regional heterogeneity in the relationship we want to estimate, having a panel data set is very useful since one of the main reasons of panel data is precisely to allow for the unobserved effects to be correlated with the explanatory variables (in our case with market potential). So the next step we are going to take in the analysis is to introduce regional (department) fixed effects into our specification (10) and carry out the estimation of equation (10) by fixed effects (FE) and first differences (FD) to obtain the fixed effect and first-differenced estimator of β_1.

The equation to be estimated is the following one:

$$\ln Ypc_{it} = \beta_0 + \beta_1 lnMP_{it} + a_i + u_{it} \qquad (12)$$

Where the variable a_i capture all unobserved time-constant factors that are affecting $Yrpc_{it}$.

As data for the variables we are using in our estimations with regard to the so called "*Nuevos departamentos*" are restricted to the departments of Vichada and Guaviare we have decided to present the subsequent set of results in this section without taking into consideration them on the premises that we cannot fully cover the whole sample of "Nuevos departamentos".

The first two columns of Table 1.17 correspond to the estimations based on a fixed effects transformation of equation (12) which leads to the estimator based on the standard least squares (within estimator or also known as fixed effects estimator). In these set of estimations we have also controlled for fixed-year effects. As expected the coefficient on market potential drops markedly but still stays both statistically and economically significant (columns 1 and 2). Since market potential is meant to capture market sizes, there is a well-acknowledged endogeneity issue in the Harris´ market potential. The market potential variable MP_{it} which

[8] The unobserved factors that change over time and affect the per capita income levels across the Colombian departments are included in the idiosyncratic error.

[9] Tabellini (2005) uses similar arguments for the analysis of the link between culture and institutions and its effects on the economic development of the European regions

in the definitions MP1 and MP2 use as a proxy for the economic activity GDP_{it} which in turn is increasing in per capita income, as captured by $Yrpc_{it}$, the dependent variable. There are several ways of dealing with this issue. First, in columns 3 and 4 we use values of MP_{it} lagged one period on the grounds that the factors that played a role in the past are uncorrelated to the factors affecting current productivity shocks in the different departments, thus avoiding problems arising from shocks linked to spatially correlated but intertemporally uncorrelated omitted variables (for instance nationwide strikes). The results of the estimations show that the coefficient estimates are positive and highly economically significant. Doubling market potential would lead to an increase in the average per capita income of around 39%-43%.

Taking longer time lags of the market potential variable helps to reduce the problems from shocks that are correlated across time. Following Boulhol and Seres (2009) we lag market potential three times. It can be seen that the estimated parameter for market potential (MP1 and MP2) which is reported in columns 5 and 6 is very significant and the magnitude is pretty much the same as with one period lags in the market potential. However, it is important to bear in mind that there are some factors that are persistent over time which are very difficult to eliminate with this approach such as for instance institutional quality, locational factors, etc.

Another way to control for this endogeneity issue is to calculate the market potential for each Colombian department by excluding the domestic component DMP_{it} in the definition of market potential contained in the expression (2) of section 2. This alternative way of computing market potential introduces measurement error as the market potential of economically quite large Colombian departments is seriously reduced (Antioquia, Valle del Cauca, Atlantico). The results of these estimations are shown in column 7. It can be seen that reestimating the baseline specification while excluding the domestic component of market potential lower the coefficient estimate for market potential from 0.39-0.43 to 0.28 but the qualitative results are unaffected. However a better treatment of this issue will be presented below with the IV estimations.

The last two columns of Table 1.17 correspond to the first-differenced estimator resulting from applying first differences to the estimation of equation (12). The results show that the β_1 coefficient estimates in the four estimations are positive and highly economic significant at the usual standard significant levels. A one percent increase in market potential leads to an increase of the average per capita income in the range of (0.39%-0.57%) depending on the definition of market potential used in the estimations. The largest effect of the market potential on the per capita income is achieved with the market potential definition (MP1).

First differences (FD) estimation procedure is based on the assumption that the first difference, Δu_{it}, of the original error u_{it} is uncorrelated over time for standard errors and t-statistics to be valid. Although sometimes this seems to be a reasonable assumption it does not follow if we assume that the original idiosyncratic errors are correlated over time. We have tested for serial correlation of the type AR(1) in the error terms of the first-differenced equations for the two estimations of columns 8 a 9. We have found negative serial correlation however the t-test of the null that the rho parameter is equal to cero cannot be rejected in the estimations where market potential is defined as MP1 although it is rejected at 5% significance level when the definition used for market potential is MP2.

Table 1.17. Robustness checks (fixed effects, lagged values, first differences)

Dependent variable Log Yrpc	Levels							First differences	
Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	13.88** (0.39)	14.65** (0.38)	10.60** (0.23)	10.22** (0.27)	10.58** (0.21)	10.16** (0.25)	12.45** (0.42)		
Log MP1	0.13** (0.03)							0.57** (0.12)	
Log MP2		0.07** (0.03)							0.39** (0.10)
Log MP1 (t-1)			0.39** (0.02)						
Log MP2 (t-1)				0.43** (0.02)					
Log MP1 (t-3)					0.39** (0.01)				
Log MP2 (t-3)						0.44** (0.02)			
Log FMP							0.28** (0.04)		
ρ[a]								-.06 (0.04)	-.08** (0.04)
year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects Department	Yes	Yes	No	No	No	No	No	Yes	Yes
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.62	0.61	0.54	0.50	0.19	0.15	0.18	0.20	0.16
Observations	624	624	600	600	650	650	624	600	600
(regions/year)	(24/26)	(24/26)	(24/25)	(24/25)	(26/25)	(26/25)	(24/26)	(24/25)	(24/25)

Note: Table displays coefficients and Huber-White standard errors for OLS. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. [a] estimate of the coefficient of AR(1) process in the error term. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

Table 1.18. Robustness checks (fixed effects, first differences) AR(1)

Dependent variable Log Yrpc	Level AR(1)		First differences AR(1)	
Regressors				
Constant	11.48** (0.79)	14.65** (0.38)		
Log MP1	0.31** (0.05)		0.67** (0.11)	
Log MP2		0.19** (0.05)		0.44** (0.09)
ρ[a]	0.821	0.821	0.14	0.08
year dummies	Yes	Yes	Yes	Yes
Fixed effects Department	Yes	Yes	No	No
Estimation	OLS	OLS	OLS	OLS
R2			0.20	0.15
Observations	624	624	600	600
(regions/year)	(24/26)	(24/26)	(24/25)	(24/25)

Note: Table displays standard errors that are robust to heteroscedasticity and contemporaneous correlation across panels. Standard errors for coefficient estimates are in parenthesis. [a] is the first-order auto-

correlation parameter. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

As Boulhol and Seres (2009) we have also estimated the model with a first-order autoregressive process in the error term and in the first differences in the variables. In the latter case, only the year-fixed effects were included as first differencing wipes out the time-invariant department effects. The results of the estimations are shown in Table 1.18 and we can see that in both specifications (Level) columns 1 and 2 and (first differences) columns 3 and 4 the estimated impact of market potential remains positive and highly economic significant at the usual standard significant levels.

In relation to the potential endogeneity bias mentioned earlier, we are going to complement the previous approaches and try to tackle the issue in a more effective way. The panel data set we have built allows us to define a set of instruments for market potential that will be used to obtain an instrumental variable estimator of the impact of economic remoteness on per capita income levels across the Colombian departments. In the earlier literature on this topic the usual approach was to use the sum of the distances of each region to Luxembourg (Breinlich, 2006) or the sum of the distances to Tokyo, Brussels and New York (Redding and Venables 2004). However Breinlich, 2006 mentions that one could have objections to this instrument for market potential on the grounds of the fact that Luxembourg is a centroid of regional income´s distribution within the EU and distance to it could be capturing other determinants of income levels besides market potential. Similarly, when Redding and Venables 2004 choose the sum of distances of each country to the three centers of global economic activity (Tokyo, Brussels and New York) as an instrument one can argue that in thirty years' time these centers no longer will be the main poles of attraction.

Therefore the choice of these three locations is in itself endogenous. In our case and following (Head and Mayer, 2006) an appealing instrument is to sum the distances of each Colombian department to all the other departments. In order to take advantage of the panel dimension of the data, we will allow the effect of this time-invariant instrument to vary through time by interacting the sum of the distances of each Colombian department to all other departments with time dummies defined for each year included in our period of analysis. The proposed instruments are $Z_{it} = h_t \frac{1}{N}\sum_{j\neq i}^{n} d_{ij}$ where N is the number of Colombian departamentos excluding the "Nuevos departamentos" and h_t are the time dummies. We will measure d_{ij} as the sum of the distances from departamento i to all other departamentos using two metrics (kilometres and lorry travel times), so we will have at our disposal two sets of instruments for market potential. Using lorry travel times allow us to control for the quality of the infrastructure.

Table 1.19. GDP per capita and economic geography (IV estimates)

Dependent variable	Log Yrpc (1)	Log Yrpc (2)	Log Yrpc (3)	Log Yrpc (4)
Regresors				
Constant	9.32** (0.40)	8.51** (0.38)	9.23** (0.42)	8.09** (0.40)
Log MP1	0.49** (0.03)	0.55** (0.03)		
Log MP2			0.52** (0.03)	0.61** (0.03)
Fixed effects Region/year	No/Yes	No/Yes	No/Yes	No/Yes
Estimation	IV	IV	IV	IV
R2	0.50	0.55	0.48	0.44
[a]Sargan´s test (and p-value)	7.09 [0.99]	14.35[0.93]	7.43[0.99]	14.48[0.93]
Observations (regions/year)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)

Note: Table displays coefficients and t-statistics for IV estimation. The dependent variable is the log of per capita income. The independent variable is the log of market potential (MP1) columns 1 and 2 and log market potential (MP2) columns 3 and 4. Instruments for MP1 and MP2 in columns 1 (2) and 3 (4) are based on the time dummies interaction with average distance to other departments in kms (lorry travel times).[a] Sargan´s overidentification test of all instruments. * and ** signify statistical significance at the10% and 5% levels. For data sources see appendix A.

The results obtained when market potential is treated as an endogenous variable and the time-varying instruments based on the average distances for each department are used are shown in Table 1.19. The estimates in the first two columns consider the definition of market potential (MP1) and the two sets of instruments defined (according to the two metrics of distances) and the estimates in the last two columns correspond the definition of market potential (MP2). It can be seen that the elasticity estimates of per capita income with respect to market potential are in both cases in line with the theoretical predictions of the model and highly economic significant. The Sargan´s test indicates that the instruments are exogenous.

1.5 Missing Links (Disentangling Channels of Influence): Physical and Human capital

Physical and Human capital

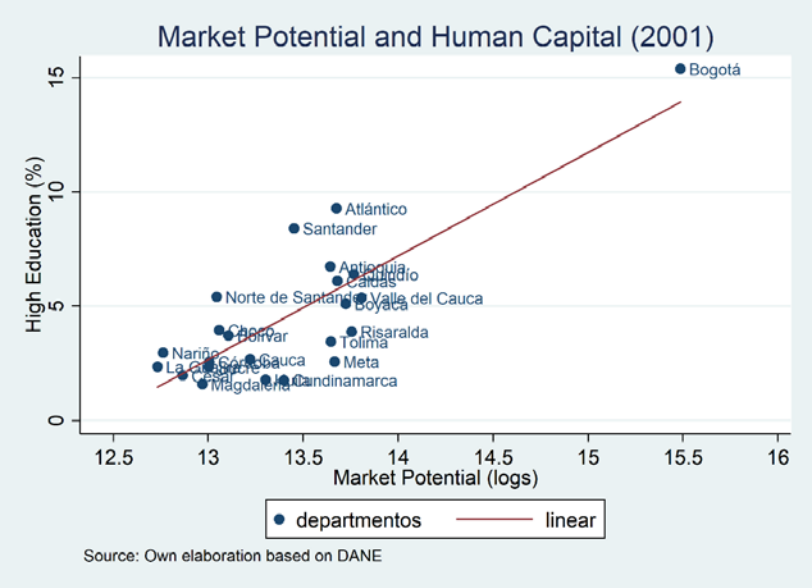
The core-periphery geographical economics model sketched in section 2 provides a theoretical framework for the empirical evidence reported in this chapter which consisted in finding a positive relationship between the level of income across the Colombian departments and their relative access to markets measured by market potential. Although we have reported some variability in the estimated elasticity of income with respect to market potential across the different empirical estimates we have carried out, the coefficient of market potential retained both economic and statistical significance. Therefore the results obtained can be considered a confirmation that the relative access of the Colombian departments to markets play an important role in shaping the income structure in Colombia. Despite this important role played by market potential with regard to the income gradient observed in Colombia it

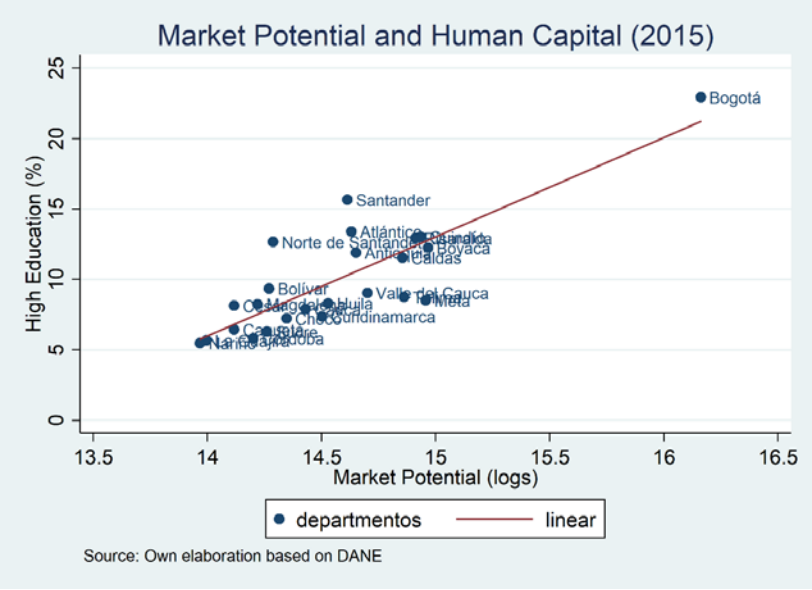
is important to clarify that the estimated baseline model does not account for other potential important determinants of the levels of income across the Colombian departments. At this point two important drivers of income levels which are related and influenced by market potential are worth mentioning: Human capital stocks and physical capital stocks. With regard to human capital on the one hand it is quite clear that locations with a better endowment of human capital (a large share of skilled workers) are locations characterized by higher income levels than locations with lower endowments of human capital.

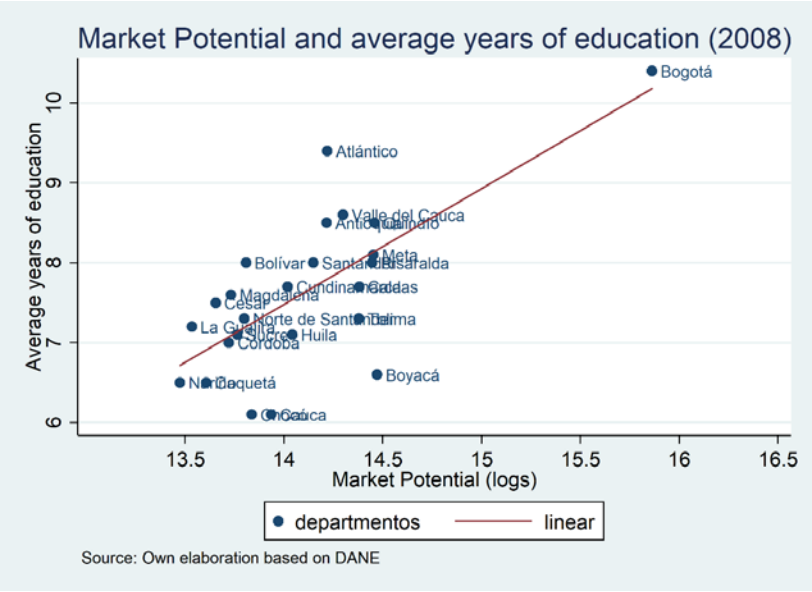
A wide range of empirical studies for developed and developing countries provide evidence that skilled or educated workers receive higher wages (see for instance Psacharopoulos, 1994). On the other hand and from the perspective of the geographical economics literature (Redding and Schott, 2003) have shown that locations with high market potential also provide more long-run incentives for human capital accumulation by increasing the premium for skilled labour. More precisely Redding and Schott (2003) result emerges from an extension of the standard two-sector (agriculture and manufacturing) Fujita et al. (1999) economic geography model to allow unskilled individuals to endogenously choose whether to invest in education. They argue that if skill-intensive sectors have higher trade costs, more pervasive input–output linkages or stronger increasing returns to scale, they show theoretically that remoteness depresses the skill premium and therefore incentives for human capital accumulation. Therefore this penalty which accrue to remote locations magnifies the effect that economic geography can have on the cross-department differences in income levels observed in Colombia. Increasing a department relative trade costs not only reduces contemporaneous factor rewards, but also lowers gross domestic product by suppressing human capital accumulation and decreasing the supply of high-income skilled workers.

Figure 1.6 shows quite clearly that the stocks of human capital are highly correlated with market potential across the Colombian departments. The first two graphs of Figure 1.6 plot the percentages of working aged (labor force) population enrolled in tertiary studies[10] against market potential for the years 2001 and 2015. The next two graphs of Figure 1.6 plot the average years of education in each Colombian department against market potential for 2008 and 2015 (see appendix A for details on the construction of these variables). No matter which variable we use to proxy human capital (percentages of high educational attainment levels or average years of education) the graphs show that the endowment of human capital is on average higher in those locations which feature high values of market potential. A remarkable feature we observe in these graphs is the outlying position of the capital of Colombia (Bogota).

[10] Tertiary studies according to the Colombian educational system includes PhD (3 years), specialization (1 year), Master (1 to 2 years), tech-professional (2 years), technological (3 years) and bachelor degree (5 years)


Market Potential and Human Capital (2001)
High Education (%)
0
5
10
15
12.5
13
13.5
14
14.5
15
15.5
16
Market Potential (logs)
Bogotá
Atlántico
Santander
Antioquia
Caldas
Norte de Santander
Valle del Cauca
Boyacá
Risaralda
Tolima
Meta
Cundinamarca
departmentos
linear
Source: Own elaboration based on DANE


Market Potential and Human Capital (2015)
High Education (%)
0
5
10
15
20
25
13.5
14
14.5
15
15.5
16
16.5
Market Potential (logs)
Bogotá
Santander
Atlántico
Norte de Santander
Boyacá
Caldas
Bolívar
Valle del Cauca
Meta
Cundinamarca
departmentos
linear
Source: Own elaboration based on DANE



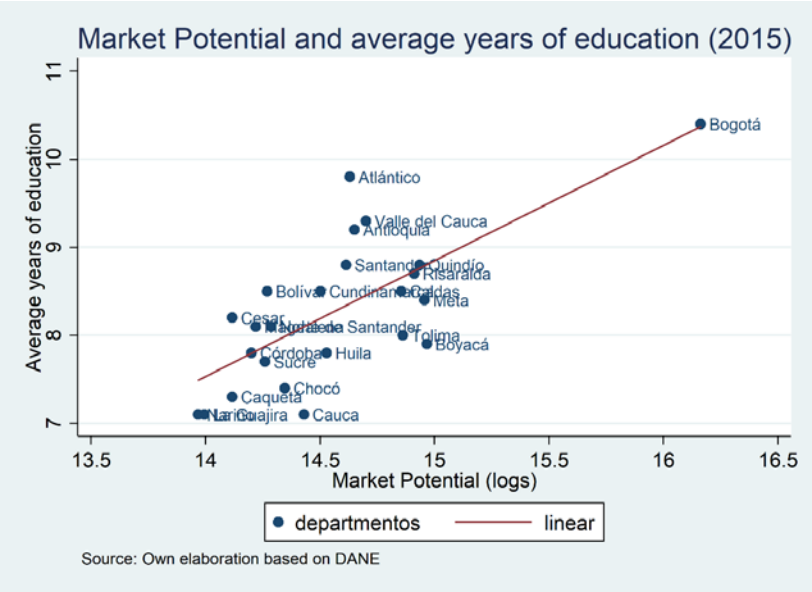


Figure 1.6. Market potential and human capital in the Colombian departments.
Source: Colombian Institute for Statistics (DANE) and authors' own calculations.

Figure 1.7 depicts two maps of the spatial distribution of human capital across the Colombian departments for the year 2015 which represent two different proxies for human capital. The upper map uses a proxy for the current state of education (percentage of labor force which is enrolled in tertiary education studies) and the lower map uses a proxy of the stock of education (average years of education for each department). A visual inspection to the maps allow us to see on the one hand the high level of correlation between the two proxies for human capital and on the other the clear core-periphery pattern depicted by the spatial distribution of both measures of human capital. The departments located along the cost (both in the pacific coast and the Caribbean sea) such as Nariño, Risaralda, Choco, Monteria, Sucre, la Guajira are the ones with the lowest endowments of human capital whereas those located

in the central parts of the country (Andean area) such as Antioquia, Santander, Cundinamarca, Tolima are the ones better endowed with human capital.

A priori, it seems reasonable that the conditions that prompt human capital accumulation hold for capital intensive goods and therefore centrality might also have a positive impact on physical capital accumulation.

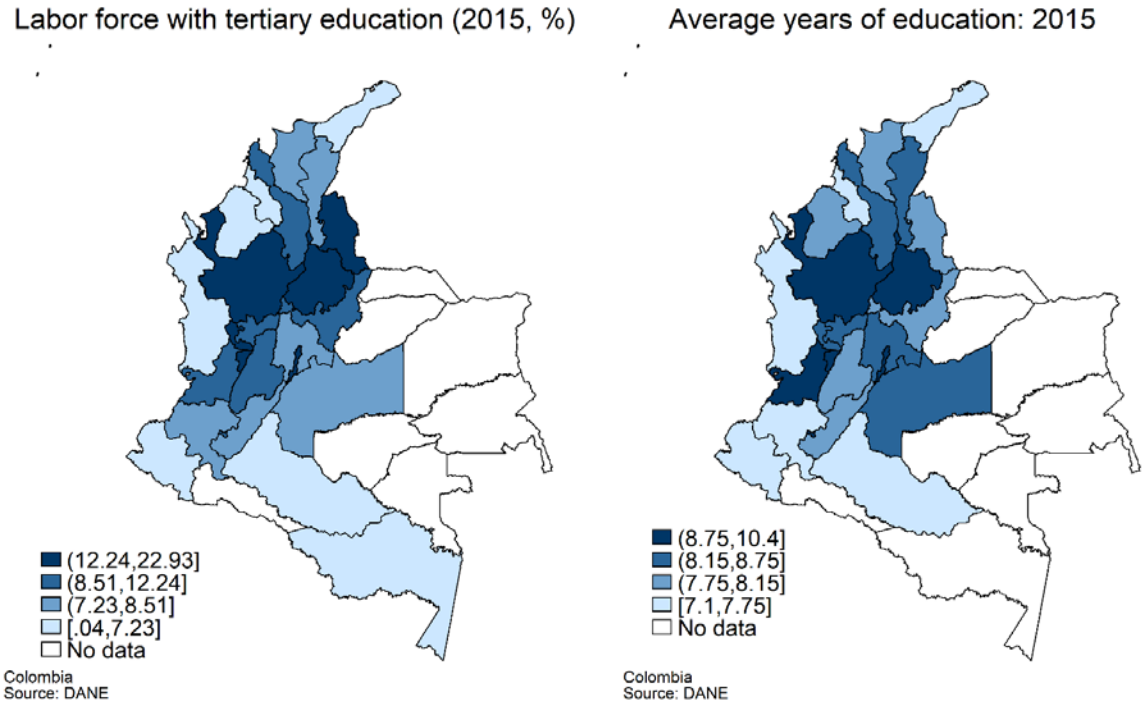
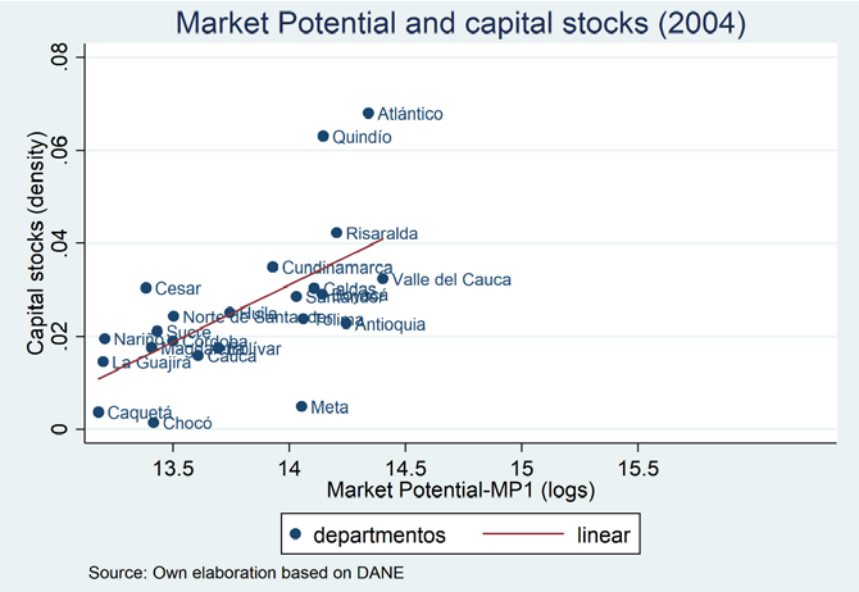


Figure 1.7. Human capital in the Colombian departments (Left (% Labor force with tertiary education), Right (average years of education)).
Source: Colombian Institute for Statistics (DANE) and authors' own calculations.



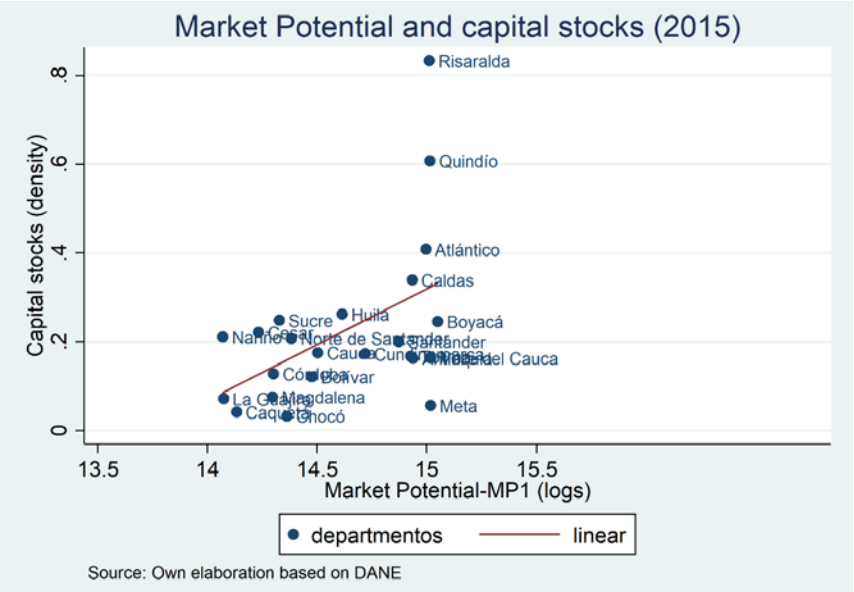


Figure 1.8. Market potential and physical capital in the Colombian departments.
Source: Colombian Institute for Statistics (DANE) and authors' own calculations.

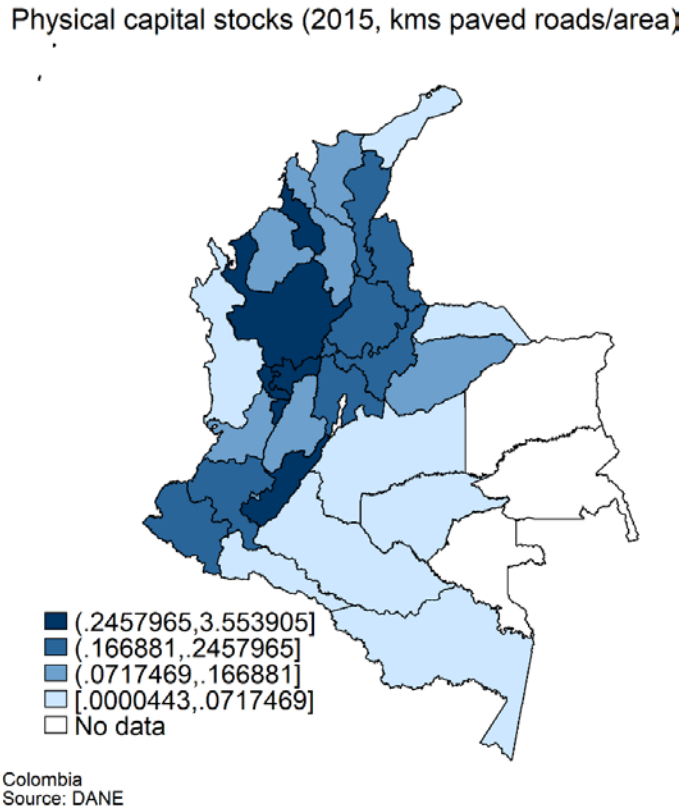


Figure 1.9. Physical capital stocks in the Colombian departments.
Source: Colombian Institute for Statistics (DANE) and authors' own calculations.

Figure 1.8 plots the kilometers of paved roads relative to the area of each Colombian department (Kilometers of paved roads/area) against market potential for the years 2004 and 2015 (see appendix A for details on the construction of these variables). The two graphs show quite clearly the fact that the Colombian departments with high market potential are those better endowed with physical capital. Figure 1.9 depicts the distribution of the physical capital proxy (kms of paved roads/area) across the Colombian departments. Again the core and central locations of Colombia are the best endowed in terms of physical capital stocks.

The regression results reported in Table 1.20 confirm the relationship depicted in Figures 1.6 and 1.8 on the link between human and physical capital and market potential. It can be seen that the coefficient associated with market potential is positive and statistically significant in the six set of regressions shown in Table 1.20. Therefore there is a positive correlation between market potential and physical and human capital stocks. Of course there are many more determinants of physical and human capital accumulation besides market potential which are far beyond the scope of this chapter but these findings are at least corroborating a long run impact of access to markets.

Table 1.20. Market potential and human and physical capital

Dependent variable	Log (capital stock)		Highed		Log (average years education)	
Regresors	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-24.61**	-24.83**	-73.45**	-75.48**	-0.37**	-0.54**
	(1.60)	(1.69)	(3.21)	(3.91)	(0.13)	(0.11)
Log MP1	1.55**		5.59**		0.17**	
	(0.11)		(0.22)		(0.009)	
Log MP2		1.58**		5.80**		0.18**
		(0.12)		(0.27)		(0.008)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS
R2	0.43	0.41	0.63	0.55	0.65	0.56
F-statistic	191.42	171.06	614.58	433.86	339.79**	520.13**
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	253	253	359	359	192	192

Note: Table displays coefficients and Huber-White standard errors for OLS. The dependent variable is the log of capital stocks (columns 1 and 2), percentage of labor force with tertiary education (columns 3 and 4) and log average years of education (columns 5 and 6). Log MP1 and Log MP2 are the logs of market potential definitions MP1 and MP2. Standard errors for coefficient estimates are in parenthesis. p-values for the statistics are in brackets. For data sources see text and appendix A. * and ** mean statistical significance at 10% and 5% respectively.

So assuming that the accumulation of physical and human capital across the Colombian departments are pretty much influenced by their relative access to markets, a natural way of testing the importance of market potential as a key factor in explaining the spatial distribution of income levels in Colombian is by incorporating both physical capital stocks and human capital stocks as additional regressors in the baseline specification estimated earlier.

Table 1.21. GDP per capita and economic geography: Disentangling channels of influence (IV estimates)

Dependent variable	Log Yrpc							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regresors								
Constant	6.15** (1.38)	5.57** (1.15)	3.92** (1.39)	3.13** (1.01)	6.49** (1.31)	5.98** (1.18)	4.31** (1.34)	2.93** 1.10
Log MP1	0.42** (0.12)	0.56** (0.08)	0.62** (0.12)	0.73** (0.07)				
Log MP2					0.36** (0.10)	0.51** (0.08)	0.54** (0.10)	0.73** (0.07)
Av. years of education	0.15** (0.05)		0.09* (0.04)		0.21** (0.04)		0.17** (0.04)	
Highed		0.002 (0.009)		-0.01 (0.03)		0.01 (0.009)		-0.00 (0.008)
lcapstock	0.29** (0.04)	0.27** (0.03)	0.27** (0.04)	0.26** (0.03)	0.31** (0.04)	0.29** (0.03)	0.30** (0.04)	0.29** (0.03)
Fixed effects Region/year	No/Yes	No/Yes	No/Yes	No/Yes	No/Yes	No/Yes	No/Yes	No/Yes
Estimation	IV	IV	IV	IV	IV	IV	IV	IV
R2	0.54	0.56	0.55	0.58	0.52	0.50	0.53	0.52
[a]Sargan´s test (and p-value)	0.21 [1.00]	0.32 [1.00]	0.18 [1.00]	0.69 [1.00]	0.24 [1.00]	0.23 [1.00]	0.20 [1.00]	0.39 [1.00]
Observations (regions/year)	184 (23/8)	253 (23/11)	184 (23/8)	253 (23/11)	184 (23/8)	253 (23/11)	184 (23/8)	253 (23/11)

Note: Table displays coefficients and t-statistics for IV estimation. The dependent variable is the log of per capita income. The independent variable is the log of market potential (MP1) columns 1, 2, 3 and 4 and log market potential (MP2) columns 5, 6, 7 and 8. Instruments for MP1 and MP2 in columns 1, 2, 5 and 6 (3, 4, 7 and 8) are based on the time dummies interaction with average distance to other departments in kms (lorry travel times).[a] Sargan´s overidentification test of all instruments. * and ** signify statistical significance at the10% and 5% levels. For data sources see appendix A

Table 1.21 reports the results of the extended regression. In columns 1 to 4 (5 to 8) we regress per capita income against log of MP1 (log of MP2), log of capital stocks and the two proxies for human capital (average years of education -columns 1 (5) and 3 (7)- and labor force enrolled in tertiary education -columns 2 (6) and 4 (8)). The coefficient estimates for market potential are in all regressions positive and highly statistically significant and remain economically significant. Instrumenting market potential with the interaction of the time dummies with average distance to other departments in kms reports higher values for the elasticity estimates of income with respect to market potential than when the instruments for market potential are defined as interaction of the time dummies with average distance to other departments measured in travel times. The values for the elasticity estimates of income with regard to MP1 (MP2) compared vis-a-vis are 0.42 vs 0.62 (regressions where we proxy human capital by the average years of education) (0.36 vs 0.54) and when we proxy human capital by percentage of labor force with tertiary studies the values are 0.56 vs 0.73 (0.51 vs 0.73).

With respect to the coefficients associated with physical and human capital proxy by the average years of education, the signs are in line with the theoretical expectations. Of course the impact of human capital on income levels is much lower than the impact of capital stocks.

Doubling the capital stock of a region increases income per capita by about 26%-30%, while increasing the average years of education increases income per capita by about 9%-21%.

When we proxy human capital by the percentage of labor force currently enrolled in tertiary education the coefficient estimates for human capital are not statistically significant and even in some regressions are not in line with theoretical expectations (although they are not significant).

Spatial dependence

A potential concern in the previous estimations of the nominal wage equation is related with the spatial dependence. We have addressed the issue of global spatial autocorrelation by computing the (1950) Moran's I (see for instance, Anselin 1993) which is defined as:

$$\mathrm{I} = \frac{n}{S_0}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(y_i-\bar{y})w_{ij}(y_j-\bar{y})}{\sum_{i=1}^{n}(y_i-\bar{y})2} \quad (13)$$

Where n stands for the number of Colombian departamentos, y_i is the variable under the analysis, $\bar{y}$ the sample average and $S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} = 1'W1$ being 1 a $(nx1)$ vector of ones is the summation of all the elements in the weight matrix and w_{ij} is the generic element of W, a spatial row-standardized weight matrix defined as:

$$W = \begin{pmatrix} 0 & \cdots & k_N w_{1,N} \\ \vdots & \ddots & \vdots \\ k_N w_{N,1} & \cdots & 0 \end{pmatrix} \qquad \text{where} \qquad k_i = \left(\sum_{j=1}^{N} w_{ij}\right)^{-1}$$

The first two moments of the (1950) Moran's I under the assumption that y is normally distributed are:

$$E[I] = -\frac{1}{n-1} \quad (14)$$

$$V[I] = \frac{(3S_0^2 + S_1 n^2 - nS_2)}{S_0(n+1)(n-1)} - \frac{1}{(n-1)^2} \quad (15)$$

With $S_1 = \left({}^{1}/_{2}\right)\sum_{i=1}^{n}\sum_{j=1}^{n}(w_{ij}+w_{ji})^2$ and $S_2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ji}\right)^2$. Hence under the normality hypothesis the moments of the statistic are uniquely determined by the elements of the spatial matrix. Its asymptotic distribution is normal:

$$\sqrt{n}\left[I - E[I]\right] \sim \mathcal{N}[0; V(I)] \qquad (16)$$

When the Moran's I test takes positive values there is positive autocorrelation which implies that regions with high (low) values have neighbours with high (low) values with respect to the average. On the contrary negative values for the Moran's I imply negative autocorrelation where regions with high (low) values are surrounded by neighbours with low (high) values.

For our particular application of the Moran's I test to the per capita income and market potential of the Colombian departments we have on the one hand excluded the so called

"Nuevos departamentos" and on the other the Islands of San Andres and Providencia. The reasons for the exclusions of these *spatial units* are quite different. In the case of "Nuevos departamentos" we do not have data for the whole period under analysis and this would cause problems with the spatial econometric exercise we will carry out later in this section. In the case of the islands of San Andres and Providencia the reasons of the exclusion are based on the types of weight matrices we are going to use in our exploratory spatial analysis. Technically, if the W matrix contains islands the element S_0 will be wrongly computed.

Our exploratory spatial analysis uses three different matrices of spatial weights. First, a k-nearest neighbour weight matrix has been used by considering the 5 nearest neighbours (k=5) where $w_{ij} = 1$ for the 5 nearest neighbours to each departmento and $w_{ij} = 0$ otherwise. Second a contiguity weight matrix, where $w_{ij} = 1$ if departamentos i and j are neighbours and $w_{ij} = 0$ otherwise. And third, an inverse weight matrix with elements defined by:

$$w_{ij} = \frac{1}{d_{ij}^2} \quad (17)$$

Where d_{ij} is the distance between the centroids of departamentos i and j.

Table 1.22. Results of the Global Spatial Autocorrelation Test (Moran's I)

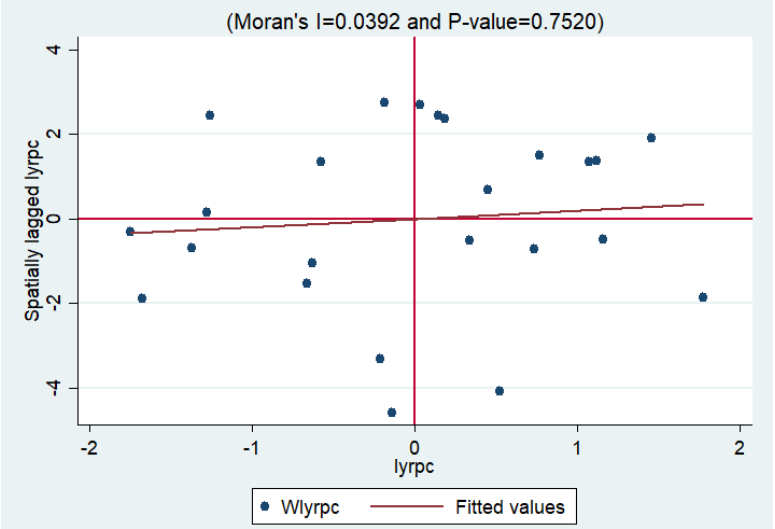
	5-nearest neighbor	First-order contiguity	Inverse distance
Log of per capita income 1990 (real terms)	0.039 (0.103)	0.079 (0.135)	0.063 (0.090)
Log of per capita income 2000 (real terms)	0.075 (0.101)	0.122 (0.133)	0.110*** (0.089)
Log of per capita income 2015 (real terms)	0.070 (0.101)	0.133 (0.132)	0.087 (0.089)
Log of market potential, 1990 (MP1)	-0.047 (0.094)	0.004 (0.124)	0.019 (0.083)
Log of market potential, 2000 (MP1)	0.091 (0.086)	0.129 (0.112)	0.084*** (0.075)
Log of market potential, 2015 (MP1)	0.181*** (0.085)	0.183*** (0.111)	0.122*** (0.074)
Log of market potential, 1990 (MP2)	-0.022 (0.093)	0.032 (0.081)	0.032 (0..081)
Log of market potential, 2000 (MP2)	0.196*** (0.086)	0.196*** (0.113)	0.128*** (0.076)0.
Log of market potential, 2015 (MP2)	0.305*** (0.090)	0.186*** (0.079)	0.186*** (0.079)

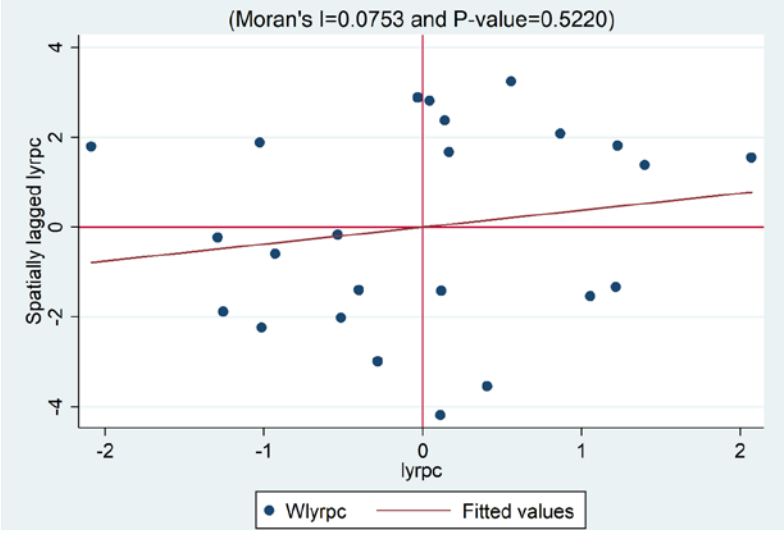
Note: *** represents significance at 1 per cent. Standard errors are in parentheses

Table 1.22 reports the results of the Moran's I test of global spatial correlation for the variables log of per capita income and log of market potential (in this later case using the definitions of market potential MP1 and MP2) in three different years, 1990, 2000 and 2015 and using three alternative types of spatial weights matrices. With regards to the results of the test on the variable log of per capita income it seems that per capita income across the Colombian departments is not spatially correlated. The only statistically significant result is for the year 2000 when the spatial weights matrix is defined as the inverse of the square of

distance between departments. A different conclusion is obtained when the variable under scrutiny is market potential. With the exception of the results of the test for the year 1990 where we did not find any significant spatial autocorrelation for the market potential variable, the results for the years 2000 and 2015 indicate that the Moran's I test clearly rejects its null hypothesis of absence of spatial dependence in both years for the three weight matrices. The only exceptions of the test are for the log MP1 in the year 2000 with for the weight matrices 5-nearest neighbour and first-order contiguity.

A visual inspection of the results of the Moran´s I test can be seen through the Moran scatterplots. Figures 1.10 to 1.12 show the Moran scatterplots for the real per capita income disparities in the years 1990, 2000 and 2015 respectively using the three spatial weights matrices we have defined (5-nearest neighbors in Figure 1.10, Contiguity matrix in Figure 1.11 and inverse of the distance in Figure 1.12).





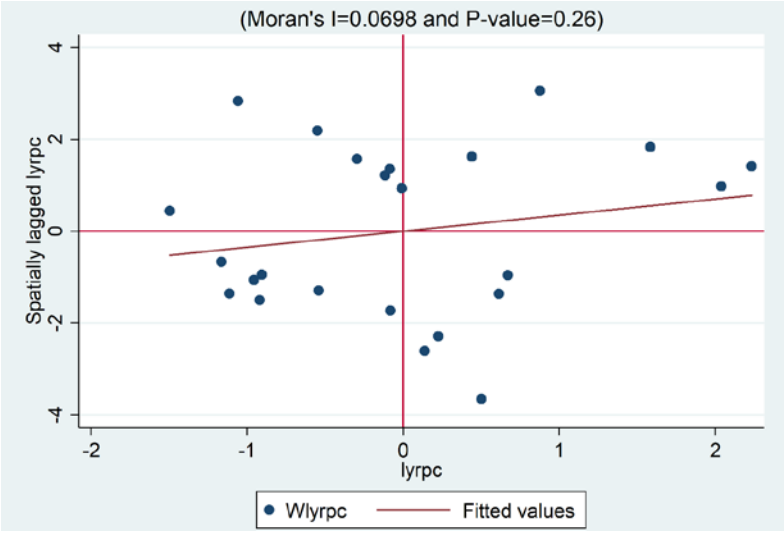
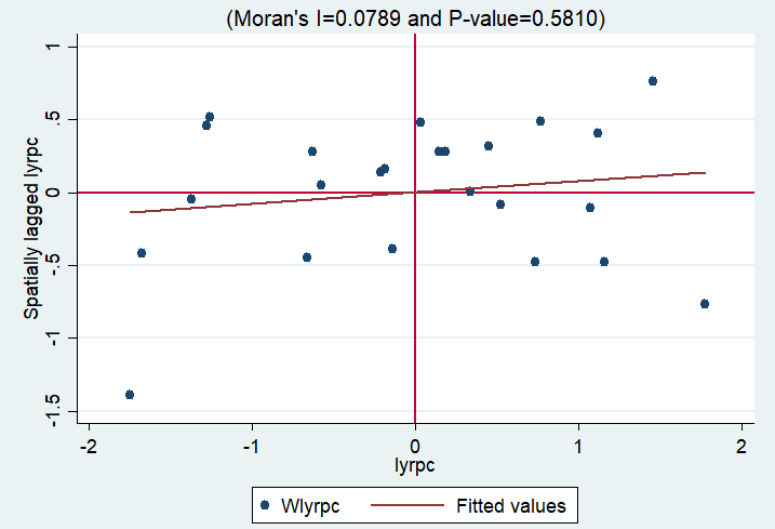


Figure 1.10. Moran scatterplots (W: 5-nearest neighbours) for the income disparities in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.



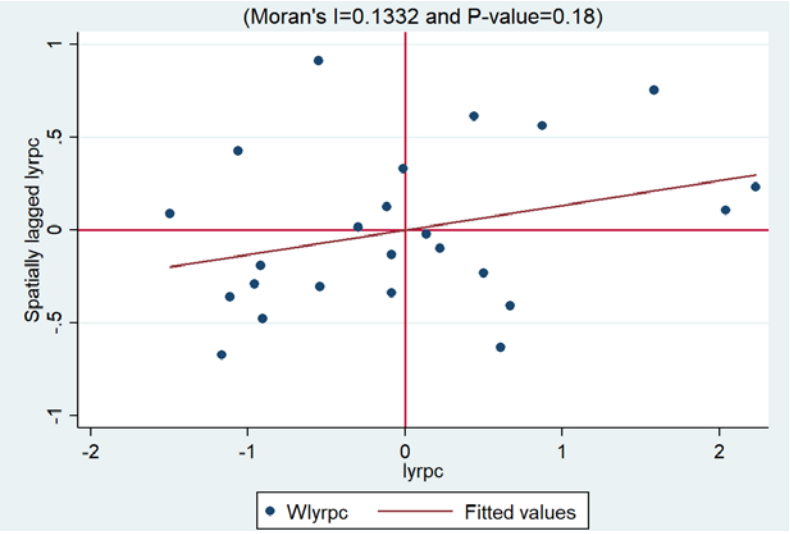
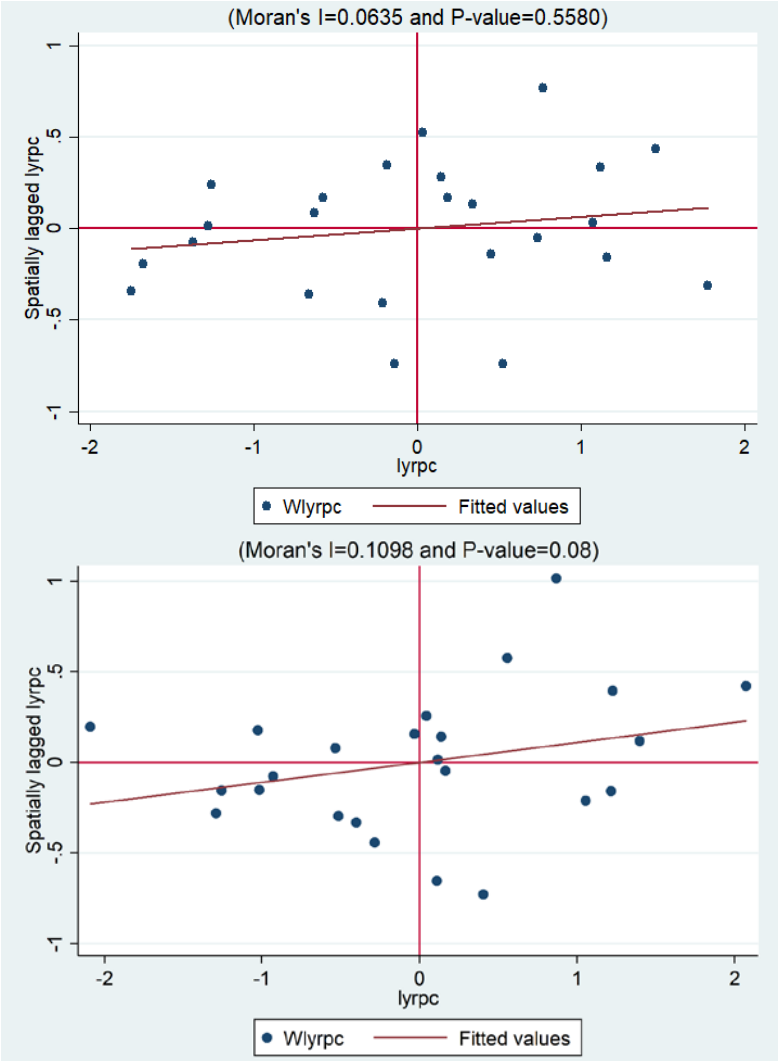


Figure 1.11. Moran scatterplots (W: contiguity) for the income disparities in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.



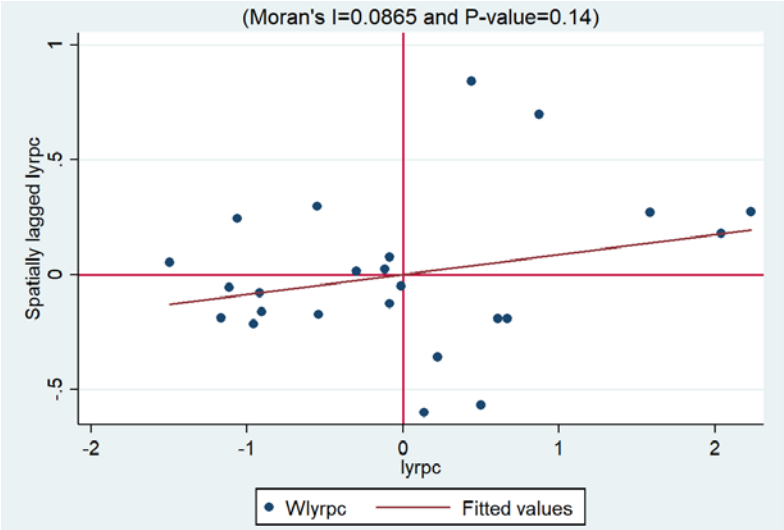
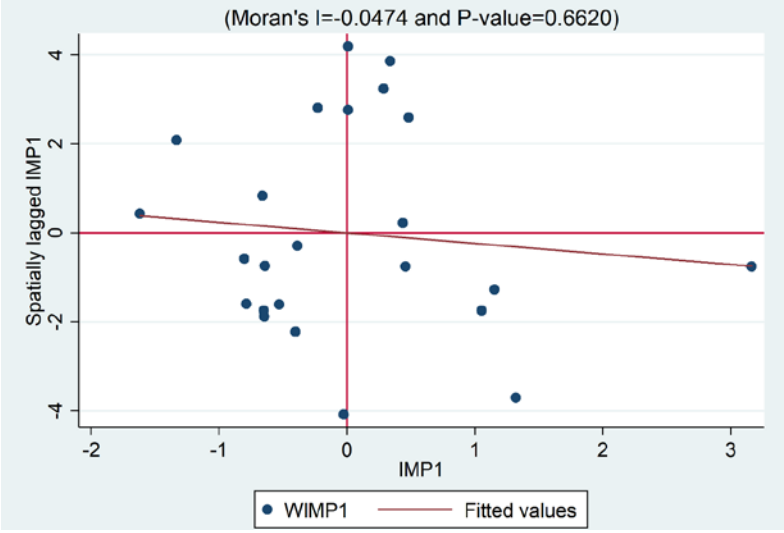


Figure 1.12. Moran scatterplots (W: inverse distance) for the income disparities in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.

It can be seen that the only case in which a statistically significant positive spatial autocorrelation is found is for the year 2000 using an inverse of distance weights matrix. Figures 1.13 to 1.15 (1.16 to 1.18) depict the Moran scatterplots for market pontential-MP1 (MP2) along the years 1990, 2000 and 2015 and with the same set of spatial weights matrix used of per capita income. For the case of market potential, a statistically significant positive spatial autocorrelation can be observed for the years 2000 and 2015. As a matter of example, using data on MP2 we have plotted a map of clusters for the year 2015 (Figure 1.19). It can be visualized the three clusters of high-high (hot spots) which correspond to the departments of Meta and Tolima, low-low (cold spots) corresponding to La Guajira and Cesar and low-high (light blue) corresponding to Caquetá.



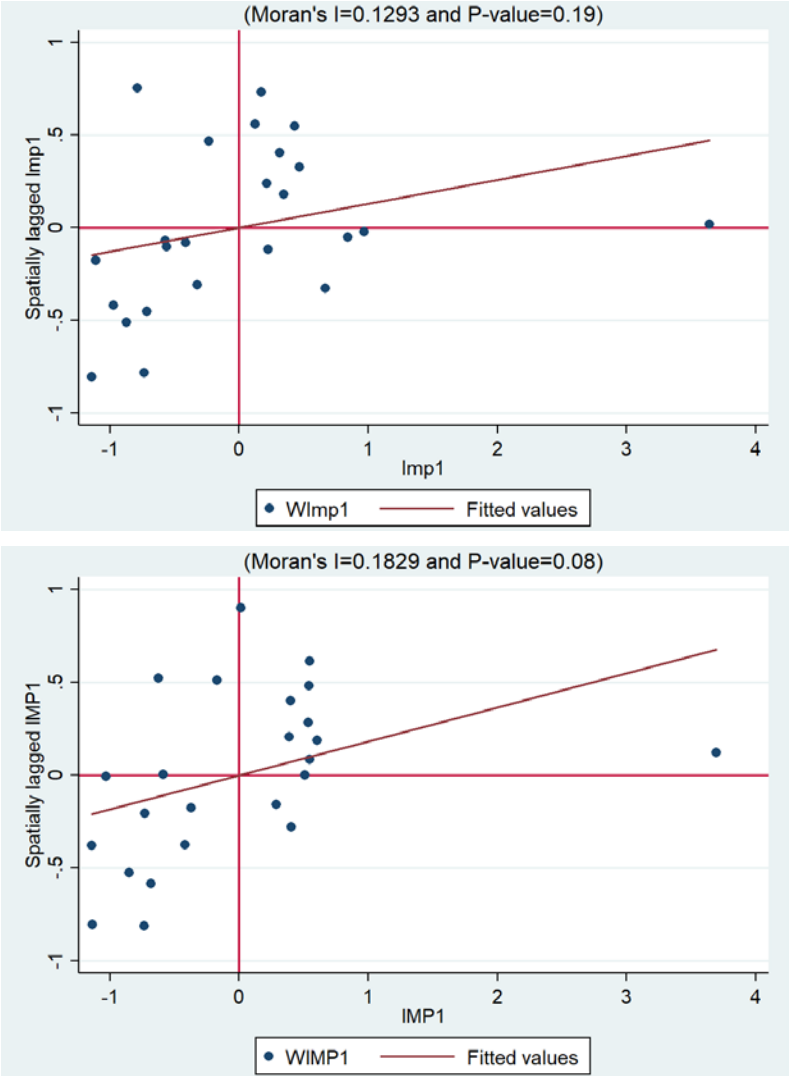
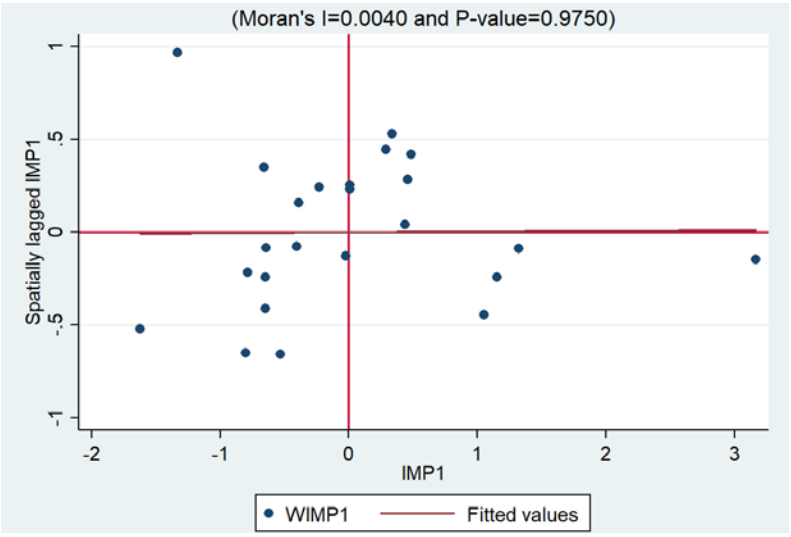


Figure 1.13. Moran scatterplots (W: 5-nearest neighbours) for the market potential (MP1) in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.



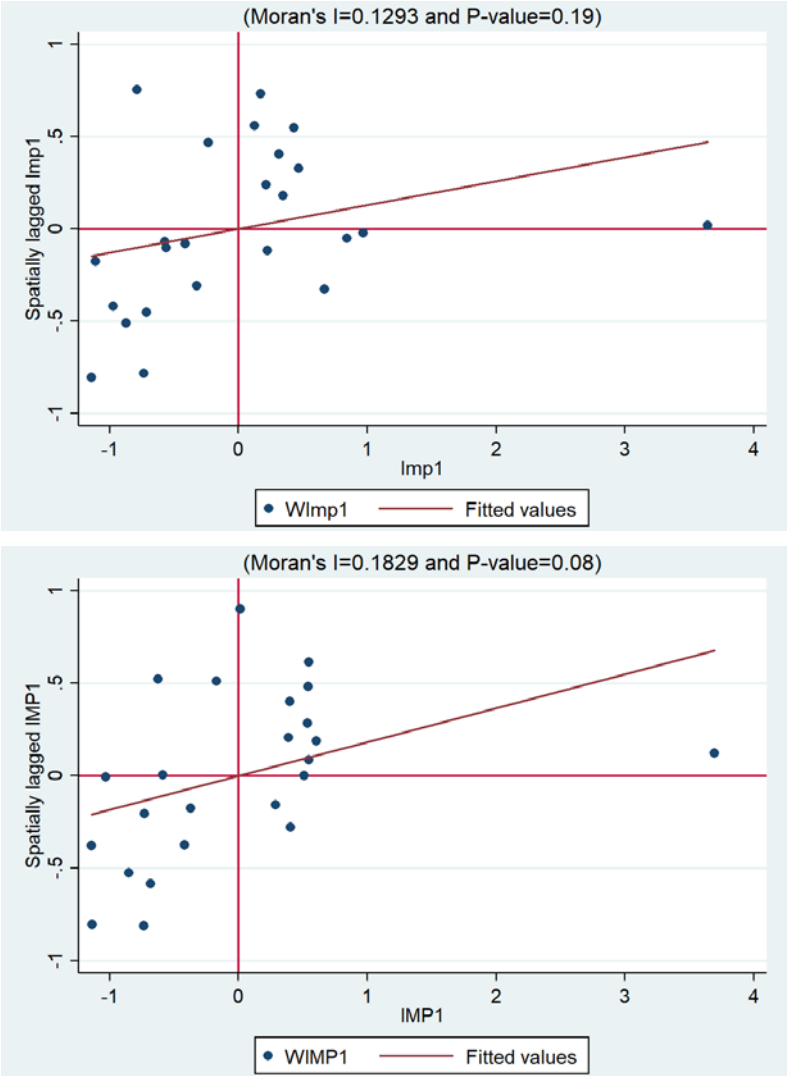
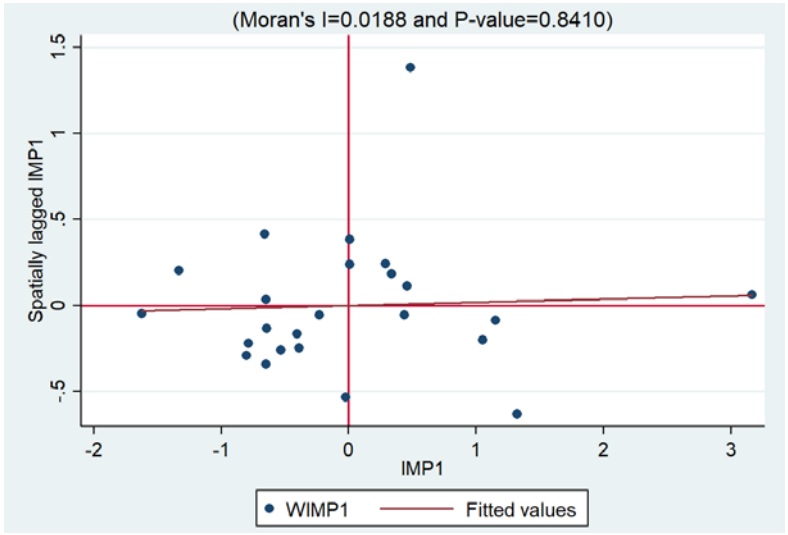


Figure 1.14. Moran scatterplots (W: contiguity) for the market potential (MP1) in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.



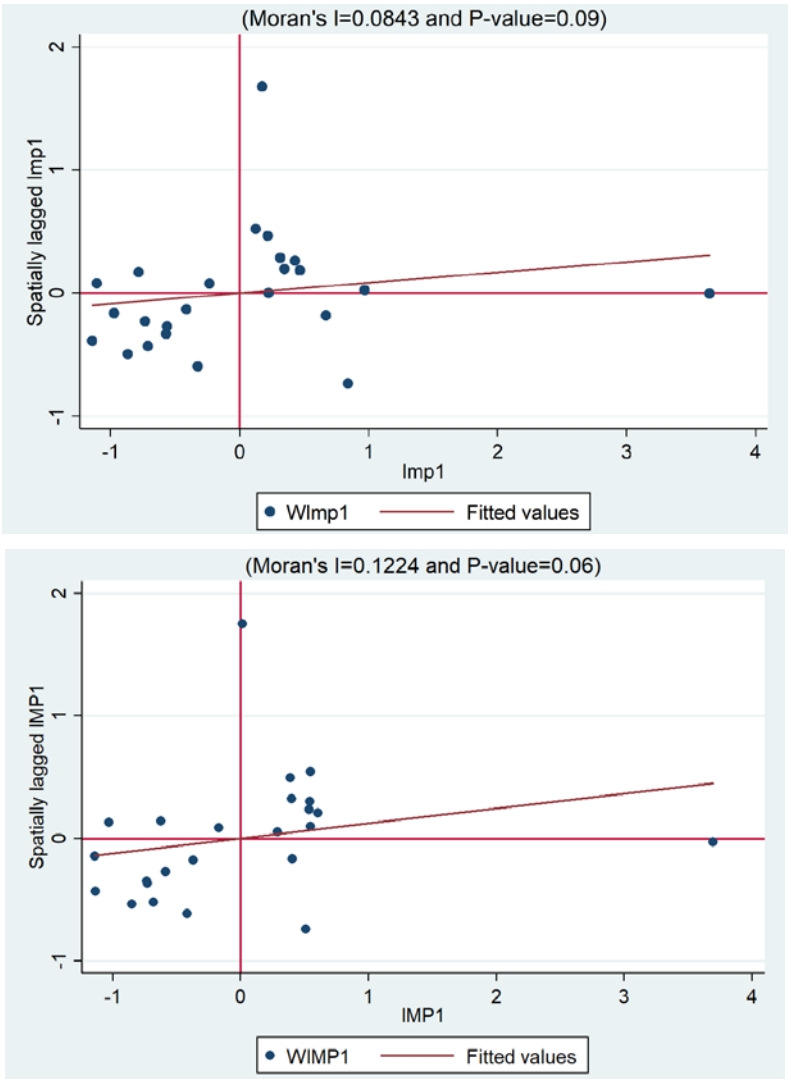
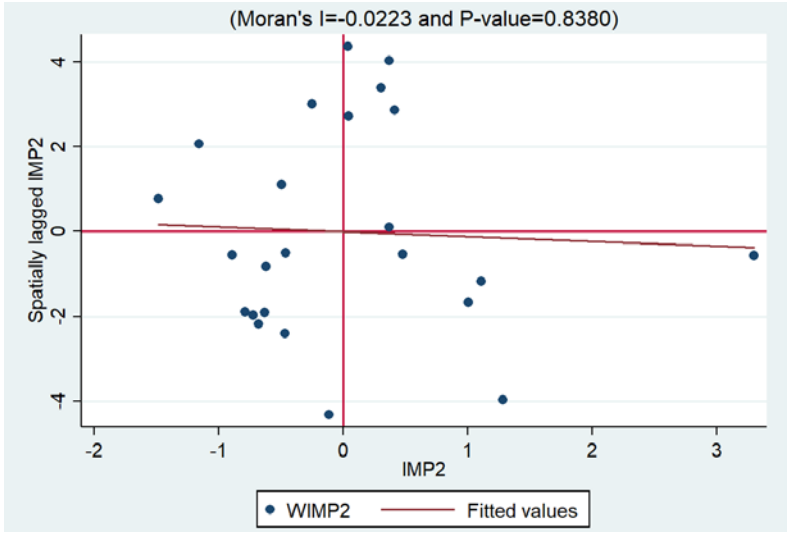


Figure 1.15. Moran scatterplots (W: inverse distance) for the market potential (MP1) in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.



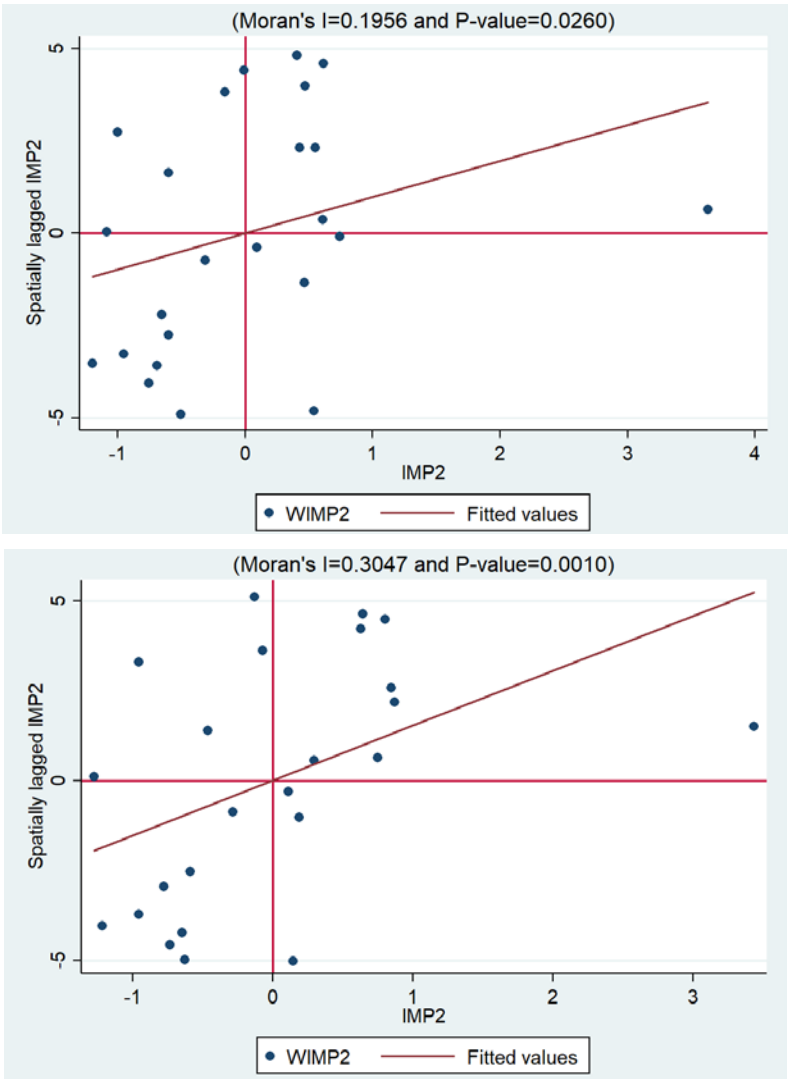
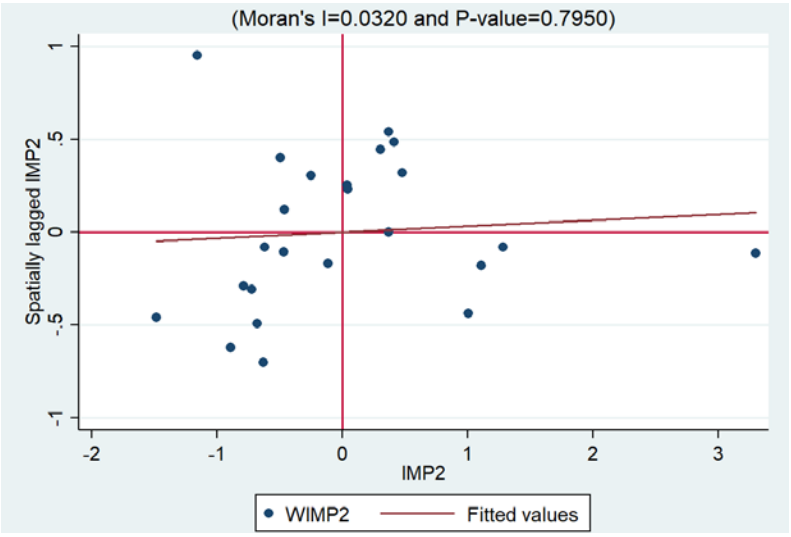


Figure 1.16. Moran scatterplots (W: 5-nearest neighbours) for the market potential (MP1) in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.



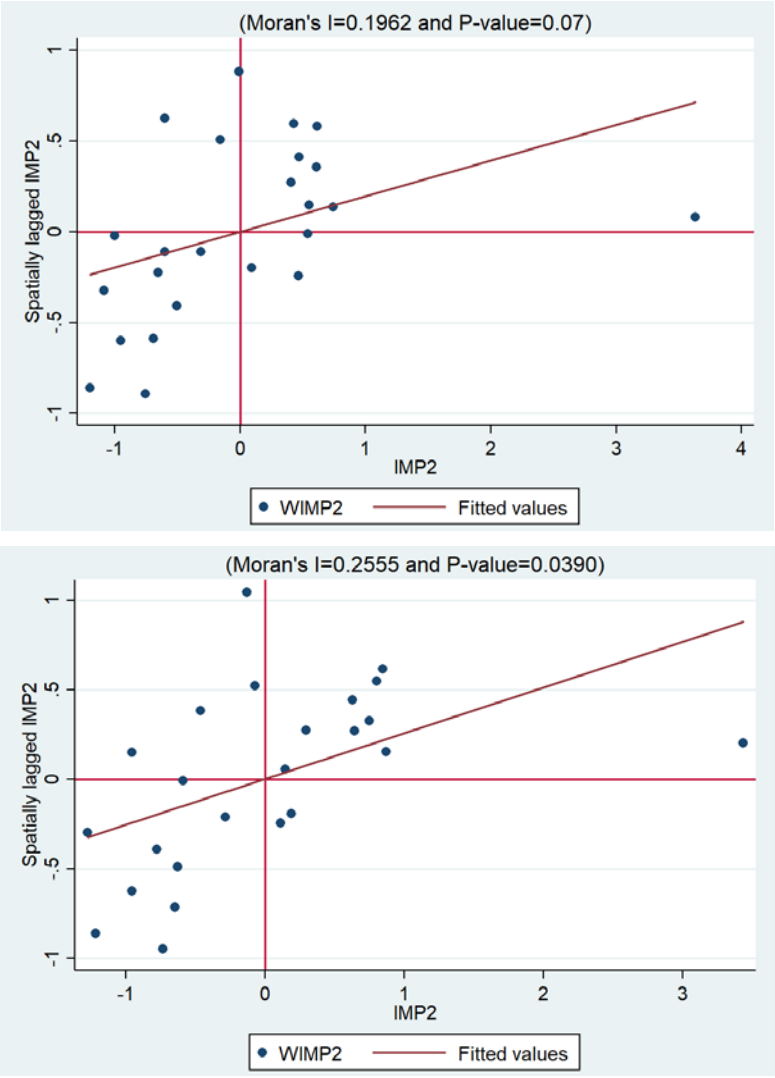
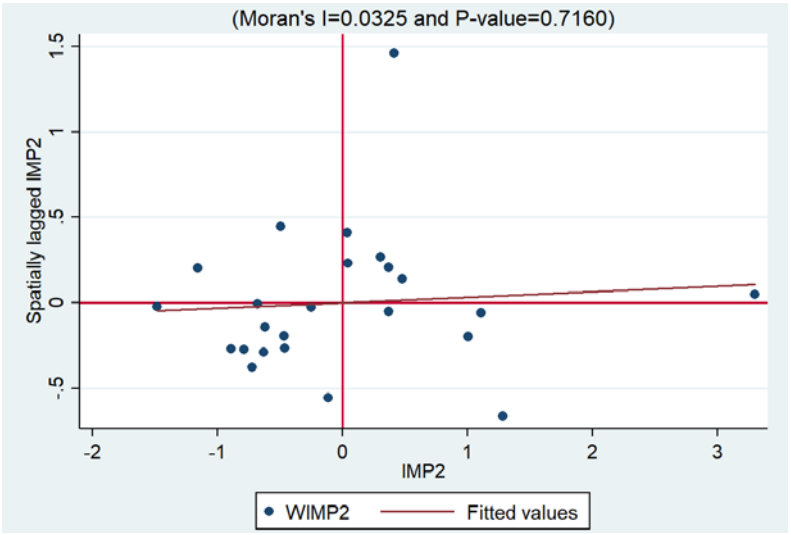


Figure 1.17. Moran scatterplots (W: contiguity) for the market potential (MP2) in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.



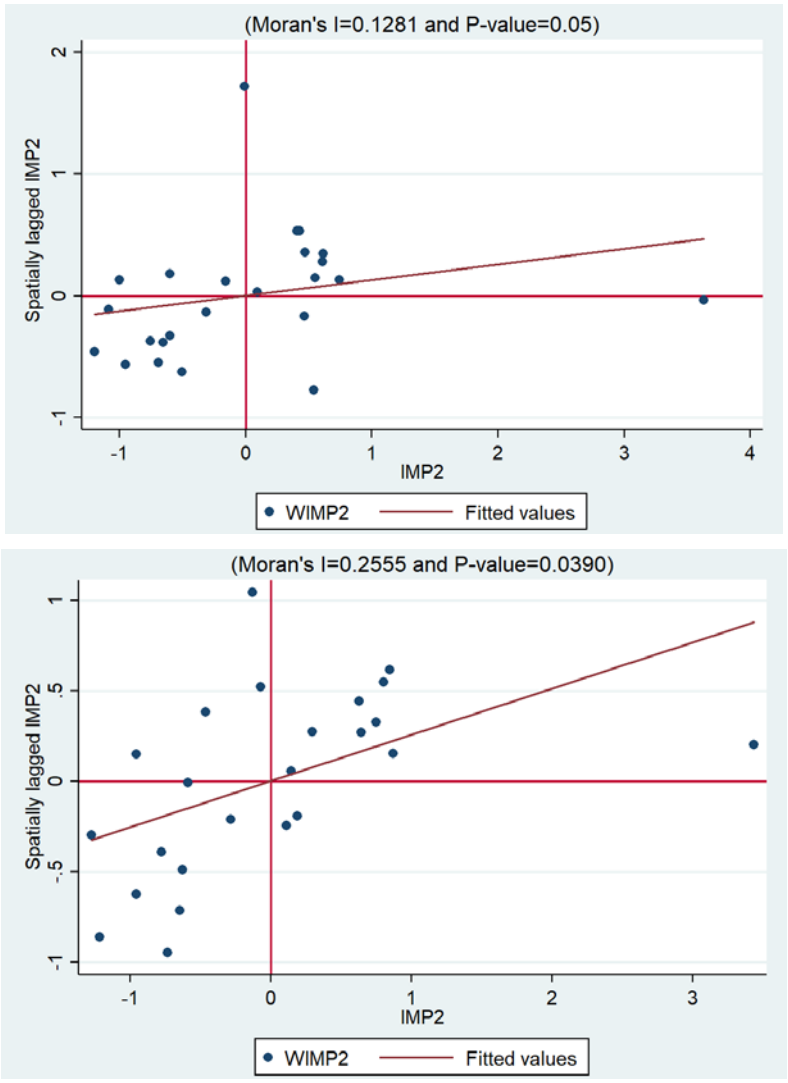


Figure 1.18. Moran scatterplots (W: inverse distance) for the market potential (MP2) in Colombia. Upper figure 1990, middle figure 2000 lower figure 2015. Source: DANE and authors' own calculations.

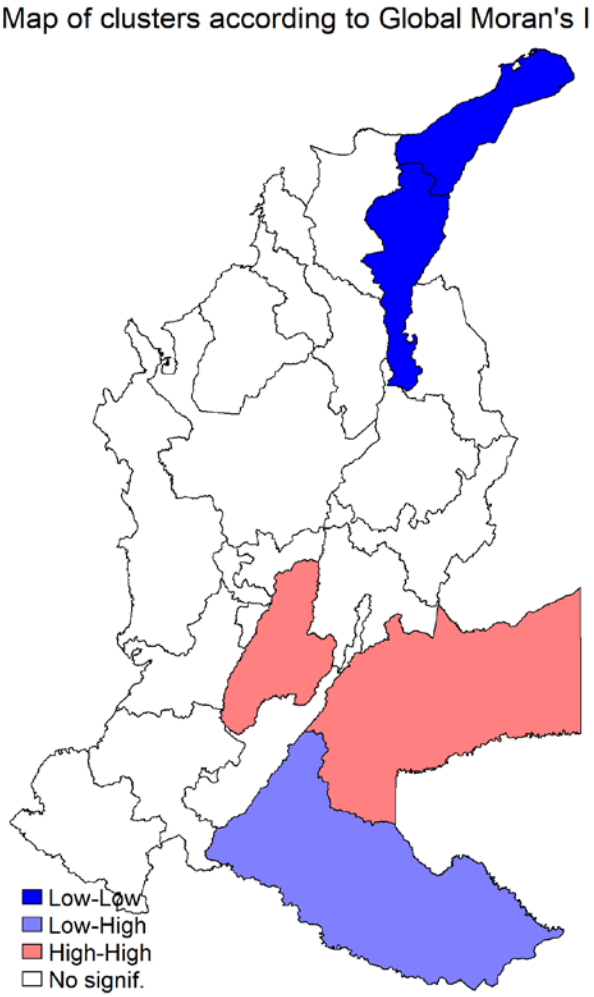


Figure 1.19. Cluster map Global Moran's I (W: inverse distance) for the market potential (MP2) in year 2015 in Colombia excluding "Nuevos departamentos". Source: DANE and authors' own calculations.

Spatial Empirical Specification

The previous results partially point to the idea that we must expect some sort of pattern of spatial dependence in our data which will need to be controlled for by applying spatial econometrics techniques in the estimation or our relationship between per capita income and market potential. This is confirmed by the results of the Moran's I and the battery of Lagrange Multiplier (LM) tests of spatial dependence shown in Table 1.23 in columns 1 and 2. The tests conclude in favor of significant residual spatial dependence which can be either nuisance of substantive. Failure of controlling for this spatial dependence will lead to results based on the OLS estimator that will be inefficient and biased.

The estimated results corresponding to these three specifications are reported in columns 3 to 5 (6 to 8) when the market potential is defined as MP1 (MP2). In the six spatial panel estimations we have carried out control for unobserved department time invariant fixed effects have been taken into account. In all estimations, the results of the Hausman tests of the null hypothesis of no systematic difference in coefficients between random and fixed

effects estimator is rejected and therefore the models have been estimated by fixed effects. The estimation results of the three spatial models show that the spatial parameter is strongly significant in all cases and large in magnitude. As for the effect of market potential, the results continue to support the hypotheses of the nominal wage equation of the core-periphery geographical economics model, i.e, the positive impact of market potential in shaping the spatial income structure observed across the Colombian departments. However, it can be seen that when the estimations are carried out controlling for spatial autocorrelation, by means of either SAR, SEM or SDM specifications, the drop in the magnitude of the coefficient estimates for market potential is quite sizeable.

Table 1.23. Results of the estimation of the spatial panel fixed effects model based on 5-nearest neighbour spatial weight matrix

	Levels (lYrpc)							
Dependent variable	Ordinary least square (OLS)		Spatial lag (SAR)	Spatial error (SEM)	Spatial Durbin (SDM)	Spatial lag (SAR)	Spatial error (SEM)	Spatial Durbin (SDM)
Regresors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	11.46*** (0.18)	11.93*** (0.19)						
Log MP1	0.30** (0.013)		0.11*** (0.01)	0.18*** (0.01)	0.11*** (0.03)			
WLog MP1					-0.00 (0.03)			
Log MP2		0.27*** (0.01)				0.08*** (0.01)	0.14*** (0.01)	0.05** (0.02)
WLog MP2								0.03 (..03)
rho			0.49*** (0.04)		0.49*** (0.04)	0.53*** (0.04)		0.52*** (0.04)
lambda				0.50*** (0.05)			0.53*** (0.05)	
Residuals Moran's I	13.431*** [0.000]	13.828*** [0.000]						
LM-ERR	174.988*** [0.000]	186.144*** [0.000]						
LM-LAG	149.236*** [0.000]	215.130*** [0.000]						
Robust LM-ERR	10.132*** [0.000]	11.324*** [0.001]						
Robust LM-LAG	14.685*** [0.000]	40.310*** [0.000]						
Hausman test			17.73*** [0.000]	16.50*** [0.000]	11*** [0.01]	20.73*** [0.000]	9.76*** [0.000]	16.83*** [0.000]
COMFAC test					8.50*** [0.003]			13.03*** [0.000]
Fixed effects Region/year	No/No	No/No	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
Yes/No	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)
624 (24/26)	0.45	0.37	0.46	0.52	0.46	0.38	0.49	0.26
Estimation	OLS	OLS	MLE					

Note: *, **, *** represent significance at 10 percent, 5 percent and 1 percent respectively. P Values for the statistics are in brackets. LM-ERR== Lagrange Multiplier test for spatial error dependence; LM-LAG==Lagrange Multiplier test for spatial lag dependence. Spatial weights matrix based on 5-nearest neighbors. Standard errors for coefficient estimates are in parentheses

Therefore, three specifications have been considered to control for spatial dependence: The spatial autorregresive model (SAR):

$$\ln Ypc_{it} = \rho W lnYpc_{it} + \beta_1 lnMP_{it} + u_{it} \quad (18)$$

The spatial error model (SEM):

$$\ln Ypc_{it} = \beta_1 lnMP_{it} + \lambda W u_{it} + \varepsilon_{it} \quad (19)$$

and the spatial Durbin model (SDM):

$$\ln Ypc_{it} = \rho W lnYpc_{it} + \beta_1 lnMP_{it} + \theta_1 W lnMP_{it} + u_{it} \quad (20)$$

It is important to bear in mind that though the spatial Durbin model (SDM) has a more complex structure than the spatial lag model (SAR) it can be reduced to a spatial error model (SER). In order to test this hypothesis we apply the test of common factors (COMFAC test). The COMFAC statistic is based on the following null and alternative hypothesis:

$$H_0 = \theta_1 + \rho\beta_1 = 0$$
$$H_1 = \theta_1 + \rho\beta_1 \neq 0$$

Which under the null means that the model can be merged into a SEM.
The value of the COMFAC statistic when we run our regression of per capita income against market potential defined as MP1 is 8.50 (p-value 0.0035) which means that we reject the null and therefore the correct spatial specification would be a SDM. However the results of the estimation of the SDM show that the spatial lag of market potential does not show up as statistically significant and therefore we conclude that the best spatial specification corresponds to a SAR. The same conclusions are obtained when our market potential variable is defined as MP2. The value of the COMFAC statistic is 13.03 (p-value 0.000) and therefore the test points to a SDM. However the spatial lag of market potential is not statistically significant leading us to a SAR-type of spatial specification.

.Table 1.24. Direct and Indirect effects (Average), W (5-nearest neighbours)

	SAR		
VARIABLES	Direct	Indirect	Total
Log MP1	0.11***	0.10***	0.21***
	(0.01)	(0.01)	(0.01)
Log MP2	0.087***	0.087***	0.17***
	(0.01)	(0.01)	(0.01)

Standard errors between brackets ***p<0.01, **p<0.05,*p<0.1.
Source: Own Elaboration

The econometric results of Table 1.23 are complemented by the computation of the marginal effects (Table 1.24) for the spatial lag (SAR) and spatial Durbin model (SDM) which are broken down into direct, indirect and total effects following the methodology of Drukker et al. (2010).

The direct effects of the market potential on per capita income levels can be attributed to the self-effect that a shock in terms of market potential of department "j" generates in the expected per capita income of department "j". In contrast, the indirect effects of market potential which in this case are global-type of spatial effects that are dynamically dispersed throughout the system represent the impact that a shock in the market potential of department "j" generates in the expected per capita income of department "j" via their impact on the expected per capita income of department "i≠j".

The results in Table 1.24 show that both the direct and indirect effects of market potential MP1 and MP2 on the per capita income levels of the Colombian departments are statistically significant at the standard significance levels and economically important. Moreover the results show that the spillover effects are of the same importance than the direct effects. Ceteris Paribus, doubling the market potential of a department increases the expected per capita income of the department by between 17%-21%. Half of this increment in expected per capita income can be attributed to the spillover effects.

Table 1.25. Results of the estimation of the spatial panel fixed effects model based on First-order contiguity spatial weight matrix

	Levels (lYrpc)							
Dependent variable	Ordinary least square (OLS)		Spatial lag (SAR)	Spatial error (SEM)	Spatial Durbin (SDM)	Spatial lag (SAR)	Spatial error (SEM)	Spatial Durbin (SDM)
Regresors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	11.46*** (0.18)	11.93*** (0.19)						
Log MP1	0.30** (0.013)		0.11*** (0.01)	0.17*** (0.01)	0.17*** (0.02)			
WLog MP1					-0.07*** (0.03)			
Log MP2		0.27*** (0.01)				0.08*** (0.01)	0.14*** (0.01)	0.12** (0.03)
WLog MP2								-0.05* (0.03)
rho			0.42*** (0.04)		0.45*** (0.04)	0.45*** (0.04)		0.47*** (0.04)
lambda				0.46*** (0.04)			0.48*** (0.04)	
Residuals Moran's I	12.409*** [0.000]	12.547*** [0.000]						
LM-ERR	150.508*** [0.000]	154.235*** [0.000]						
LM-LAG	136.329*** [0.000]	180.606*** [0.000]						
Robust LM-ERR	26.845*** [0.000]	6.120*** [0.01]						
Robust LM-LAG	12.666*** [0.000]	32.491*** [0.000]						
Hausman test			15.07*** [0.000]	11.59*** [0.000]	8.39*** [0.04]	17.00*** [0.000]	8.10*** [0.000]	16.72*** [0.000]
COMFAC test					0.02 [0.90]			0.43 [0.51]

Dependent variable	Ordinary least square (OLS)		Spatial lag (SAR)	Spatial error (SEM)	Spatial Durbin (SDM)	Spatial lag (SAR)	Spatial error (SEM)	Spatial Durbin (SDM)
Fixed effects Region/year	No/No	No/No	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No
Yes/No	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)	624 (24/26)
624 (24/26)	0.45	0.37	0.50	0.52	0.52	0.45	0.49	0.49
Estimation	OLS	OLS	MLE					

Note: *, **, *** represent significance at 10 percent, 5 percent and 1 percent respectively. P Values for the statistics are in brackets. LM-ERR== Lagrange Multiplier test for spatial error dependence; LM-LAG==Lagrange Multiplier test for spatial lag dependence. Spatial weights matrix based on 5-nearest neighbors. Standard errors for coefficient estimates are in parentheses

We have repeated the estimations carried out in Table 1.23 (and the associated direct and indirect effects -Table 1.24) in Table 1.25 using a first-order contiguity type of spatial weights matrix. Broadly speaking the results are very similar. The key different between the results in Tables 1.23 and 1.24 is related to whether or not the spatial Durbin model can potentially be merged into a spatial error model (SER). In this case, when a first-order contiguity spatial weights matrix is used, the value of the COMFAC statistic does not allow us to reject the null which means that our correct spatial specification is a spatial error model (SER). The conclusions regarding the important role played by market potential are unaltered.

Extended Spatial Empirical Specification

In this section we have extended the previous spatial empirical specifications along the lines follow in section 4.1. However our data restrictions for carried out this extended estimation were quite severe. On the one hand we do not have a fully balanced panel data on physical capital and with regard to human capital we were able to collect data on average years of education in each Colombian department as a fully balanced panel only for the period 2008-2015. Therefore the extended spatial empirical specifications will be carried out for the period 2008-2015 and only controlling for human capital.

Table 1.26. Results of the estimation of the extended spatial panel fixed effects model based on 5-nearest neighbour spatial weight matrix

	Levels (lYrpc)							
Dependent variable	Ordinary least square (OLS)		Fixed effects	Fixed effects	Spatial lag (SAR)	Spatial error (SEM)	Spatial lag (SAR)	Spatial error (SEM)
Regresors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	9.75*** (0.82)	9.69*** (0.89)	8.63*** (0.48)	9.23*** (0.47)				
Log MP1	0.36** (0.07)		0.50*** (0.03)		0.59*** (0.07)	0.51*** (0.06)		
AYE	0.10*** (0.05)	0.15*** (0.04)			-0.007 (0.03)	-0.01 (0.03)		0.02 (0.03)

Dependent variable	Ordinary least square (OLS)		Fixed effects	Fixed effects	Spatial lag (SAR)	Spatial error (SEM)	Spatial lag (SAR)	Spatial error (SEM)
Log MP2		0.35*** (0.08)		0.46*** (0.03)			0.50*** (0.08)	0.43** (0.06)
Rho					-0.20 (0.13)		0.17 (0.14)	
lambda						-0.07 (0.17)		-0.11 (0.17)
Residuals Moran's I	0.728 [0.467]	1.26 [0.207]						
LM-ERR	0.226 [0.635]	0.977 [0.323]						
LM-LAG	0.053 [0.817]	0.354 [0.552]						
Robust LM-ERR	0.199 [0.656]	0.660 [0.001]						
Robust LM-LAG	0.026 [0.871]	0.036 [0.849]						
Pesaran test			4.342 [0.000]	4.305 [0.000]				
Hausman test					2.82 [0.42]	2.04 [0.56]	5.28 [0.15]	7.98** [0.05]
Fixed effects Region/year	No/No	No/No	Yes/No	Yes/No	Yes/ No	Yes/ No	Yes/No	Yes/No
Number obs. (región/year)	192 (24/8)	192 (24/8)	192 (24/8)	192 (24/8)	192 (24/8)	192 (24/8)	192 (24/8)	192 (24/8)
R2	0.44	0.43	0.41	0.38	0.41	0.41	0.40	0.39
Estimation	OLS	OLS	FE	FE	MLE			

Note: *, **, *** represent significance at 10 percent, 5 percent and 1 percent respectively. P Values for the statistics are in brackets. LM-ERR== Lagrange Multiplier test for spatial error dependence; LM-LAG==Lagrange Multiplier test for spatial lag dependence. Spatial weights matrix based on 5-nearest neighbors. Standard errors for coefficient estimates are in parentheses

Prior to properly estimate our extended spatial panel model, columns 1 and 2 of Table 1.24 show the results of estimating an extended spatial pooled OLS model which includes as an additional control the stock of human capital (average years of education). The results of the estimation show that after controlling for human capital, the elasticity estimate of per capita income with regard to market potential is positive and statistically significant. As for the impact of human capital, the results show that it exerts a positive impact of the per capita income levels being the estimated coefficient statistically significant. We have also checked for spatial dependence in the estimations of the spatial pooled OLS model. The battery of spatial dependence tests (Moran's I, LL-LAG, LL-ERR and their robust versions) reveal that this extended spatial specification with the incorporation of human capital is not spatially misspecified. However, the estimation of the baseline spatial panel data fixed effects models reveals just the opposite. The results of the Pesaran's tests in columns 3 and 4 which correspond to the baseline estimations under the definitions of market potential MP1 (column 3) and MP2 (column 4) show that the null of lack of cross sectional independence is strongly rejected. In other words, one reason for rejecting the null could be the potential presence of spatial autocorrelation. To check this premise we proceed with the estimation of the following extended spatial panel data models:

The spatial autoregressive model (SAR):

$$\ln Ypc_{it} = \rho W lnYpc_{it} + \beta_1 lnMP_{it} + \beta_2 AYE_{it} + u_{it} \quad (21)$$

And the spatial error model (SEM):

$$\ln Ypc_{it} = \beta_1 lnMP_{it} + \beta_2 AYE_{it} + \lambda W u_{it} + \varepsilon_{it} \quad (22)$$

The results of the estimation of models (21) and (22) are reported in columns 5 (7) and 6 (8) for the definition of market potential MP1 (MP2). As it can be seen in the results, neither the estimated spatial lags (rho and lambda) nor the human capital variable are statistically significant. In other words, the lack of significance of the spatial lag parameters rho and lambda jointly with the Pesaran's test results is pointing out to a kind of correlation across the units which does not have a spatial structure.

With regard to the market potential variable it continues to show its relevance in the explanation of the spatial income structure observe in Colombia (positive, statistical and economic significant) and the size of the coefficient estimates is in the range 0.43-0.59.

We have repeated the estimations of models (21) and (22) in columns 3, 4 (5, 6) of Table 1.27 using a first-order contiguity weights matrix. While in column 3 and 4 we use the definition of market potential as MP1 in column 5 and 6 we use MP2. The results of the Hausman test do not reject the null of difference in coefficient not systematic when the definition of market potential used in the estimations is MP2 whereas it is rejected under the definition MP1. Again the results show that human capital is not statistically significant in any of the estimations. Market potential keeps its statistical significance and economic importance across the different spatial panel model specifications.

Table 1.27. Results of the estimation of the extended spatial panel fixed effects model based on first-order contiguity spatial weight matrix

	Levels (lYrpc)					
Dependent variable	Ordinary least square (OLS)		Spatial lag (SAR)	Spatial error (SEM)	Spatial lag (SAR)	Spatial error (SEM)
Regresors	(1)	(2)	(3)	(4)	(5)	(6)
Constant	9.75***	9.69***				
	(0.82)	(0.89)				
Log MP1	0.36**		0.54***	0.53***		
	(0.07)		(0.07)	(0.06)		
AYE	0.10***	0.15***	-0.01	-0.018	0.01	0.01
	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Log MP2		0.35***			0.44***	0.44**
		(0.08)			(0.07)	(0.06)
rho			-0.04		-0.009	
			(0.09)		(0.09)	
lambda				0.07		0.04
				(0.11)		(0.11)
Dependent variable	Ordinary least square (OLS)		Spatial lag (SAR)	Spatial error (SEM)	Spatial lag (SAR)	Spatial error (SEM)

Residuals	0.598	0.935				
Moran's I	[0.550]	[0.35]				
LM-ERR	0.166	0.540				
	[0.684]	[0.462]				
LM-LAG	0.645	0.354				
	[0.422]	[0.552]				
Robust LM-ERR	0.088	0.008				
	[0.767]	[0.927]				
Robust LM-LAG	0.568	0.548				
	[0.451]	[0.459]				
Pesaran test						
Hausman test			2.37	2.30	8.70**	9.26**
			[0.49]	[0.51]	[0.03]	[0.03]
Fixed effects Region/year	No/No	No/No	Yes/No	Yes/No	Yes/No	Yes/No
Number obs.	192	192	192	192	192	192
(región/year)	(24/8)	(24/8)	(24/8)	(24/8)	(24/8)	(24/8)
R2	0.44	0.43	0.41	0.40	0.39	0.39
Estimation	OLS	OLS	MLE			

Note: *, **, *** represent significance at 10 percent, 5 percent and 1 percent respectively. P Values for the statistics are in brackets. LM-ERR== Lagrange Multiplier test for spatial error dependence; LM-LAG==Lagrange Multiplier test for spatial lag dependence. Spatial weights matrix based on 5-nearest neighbors.
Standard errors for coefficient estimates are in parentheses

1.6 Conclusions

This chapter analysed the spatial distribution of per capita income across the Colombian departments over the period 1990-2015 according to the postulates of the core-periphery geographical economics models. More specifically, we have analysed the existence of a spatial income structure in Colombia by testing the so called *nominal wage equation* of the geographical economics models. The nominal wage equation (see equation 8 and 9 in section 3) implies that income or wages will be higher in or near economic centers. Therefore given the stylized facts about the spatial distribution of per capita income in Colombia (core-periphery pattern, sizeable income disparities and a strong income-gradient with regard to distance from Bogota), we have carried out an empirical exercise to validate or refute the theoretical predictions regarding to the nominal wage equation. The nominal wage equation relates income or wages in a location with a weighted sum of the volume of economic activity in the surrounding locations where the weighted scheme is a function of the inverse of the distance measure either in kilometers or travel times between locations. This weighted sum is usually known as market potential or market access. Therefore our main goal was to test the role of market potential in shaping the spatial income structure observed in Colombia. To do so, using our panel data set we have estimated a number of different specifications regressing per capita income on market potential. Our first approach to the empirical exercise was to carried out a pooled OLS estimation of the impact of market potential on per capita income over the whole sample period (1990-2015) and also by sub-periods (1990-1994, 1995-2004 and 2005-2015) using 6 different alternatives for the definition of market potential which we labelled as MP1, MP2, MP3, MP4, MP5 and MP6. The results of this baseline estimation of the nominal wage equation were in line with the theoretical predictions of the

model, i.e., market potential no matter which definition is used in the regressions exerts a positive and economically significant impact in the expected or average income of the department. The estimated slope parameters of market potential using the full sample vary depending on market potential definition used but they were in the range 0.16-0.36. In the analysis carried out by sub-periods we have also found a positive and economically significant effect of market potential on per capita income.

By dividing the whole period into these three sub-periods we have also tested if the effect of access to markets (market potential) had changed over the period 1990-2015. The results of the estimations showed that the effect of market potential was not the same across the three sub-periods. Whereas in the initial period the estimated slope parameter was in the range (0.26-0.29) in the last period was in the range (0.56-0.64). This result is strongly supportive of the fact that the process of agglomeration of economic activities in Colombia has increased over the last twenty five years fostering therefore a process of widening the income gap across departments. Our next step in the empirical exercise was to carry out some robustness checks in the baseline estimation. Using the panel data character of our data we introduce regional (department) fixed effects in order to control for many potential unobserved time-constant factors and estimated the baseline specification by fixed effects (FE) and first differences (FD).

The results are again in favour of the hypothesis of the nominal wage equation. In order to preserve the interpretation of t-statistics and to produce valid standard errors the previous FE and FD estimations were estimated with a first-order autoregressive process in the error term (AR(1)). These alternative estimations showed that market potential remains positive and highly economic significant at the usual standard significant levels. The estimated slope parameter for market potential was in the range (0.19-0.31) for the FE-AR(1) and in the range 0.44-0.67 in the FD-AR(1). Our next concern was about the potential endogeneity of market potential. We have approached to this issue by instrumenting market potential with two different sets of instruments which consisted of interacting the sum of the distances of each Colombian department to all other departments with time dummies defined for each year. The results of the IV estimations delivered slope parameter estimates for market potential which were in the range 0.49-0.55 when the definition of market potential was MP1 and 0.52-0.61 when the definition employed for market potential was MP2.

Another important step we took in the estimation of the baseline nominal wage equation was to disentangle the potential effects of physical and human capital since there are reasons both from a theoretical side as well as from an empirical one that the accumulation of both physical and human capital across the Colombian departments is pretty much influenced by their relative access to markets. The results of the extended baseline estimations controlling for both human and physical capital showed that market potential keeps its importance.

Our final concern in relation with our empirical estimates of the nominal wage equation was related to the spatial dependence. The initial exploratory spatial analysis we did using three types of weight matrices (5-nearest neighbour, contiguity and inverse of distance) showed that market potential was spatially correlated. To take this issue into account we estimated three specifications to control for spatial dependence, a spatial autorregresive model (SAR), a spatial error model (SEM) and a spatial Durbin model (SDM) using two alternatives for the

spatial weight matrix (5-nearest neighbour and contiguity). The results of the COMFAC statistic using as spatial weight matrix the 5-nearest neighbours were in favour of a SDM. However the spatial lag of market potential in this spefication was not statistically significant leading us to conclude in favour of a SAR-type of spatial specification. The estimated direct impact of market potential in the SAR model was pretty much of the size of the indirect impact (0.11) suggesting that the spillover effects are quite important. Using a first-order contiguity spatial weight matrix lead us to use conclude in favour of a SER type of specification. The estimated impact of market potential was in the range 0.14-0.17.

Finally, we have extended the spatial empirical specification controlling for human capital. In this case the estimated extended spatial panel fixed effects models continue to validate the theoretical predictions of the nominal wage equation.

1.7 Appendix

Appendix A. Source of data and dictionary

1. **dp**: código DIVIPOLA (División política administrativa) de los departamentos de Colombia. "DANE, Codificación de la división Político-Administrativa de Colombia (DIVIPOLA) http://geoportal.dane.gov.co:8084/Divipola/".
2. **departamento:** nombre del departamento
3. **ndpto:** dumy creada con valor 0 si departamento es viejo departamento y 1 si departamento es nuevo departamento
4. **capital:** nombre de la capital del departamento
5. **d1:** distancia interna dentro de cada departamento tomada como 1/3raizcuadra (área/pi)
6. **d2:** distancia interna dentro de cada departamento tomada como 2/3raizcuadra (área/pi)
7. **nm:** número de municipios y corregimientos departamentales de cada departamento: "DANE, Codificación de la división Político-Administrativa de Colombia (DIVIPOLA) http://geoportal.dane.gov.co:8084/Divipola/". Última actualización 30/09/2016
8. **a:** superficie de cada departamento en KM2. Fuente: "Instituto Geográfico Agustín Codazzi - IGAC-, Diccionario Geográfico http://www.igac.gov.co/digeo/app/index.html". "Cundinamarca y Bogotá no vienen separados, si se trabaja Bogotá como independiente debe quitarse su extensión del total. El área del Archipiélago de San Andrés se tomó de la página de la gobernación del mismo, en el IGAC es de sólo 44 km². Última actualización 2016.
9. **pmd3pxxxx** (xxxx inicia en 1951 y termina en el 2016; representa el potencial de mercado doméstico calculado a partir de los datos de población de cada departamento. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.33. **Fuentes**: Cálculos propios. Datos: 1951-1984, DNP, Estadísticas Históricas de Colombia. https://cutt.ly/stdWQSO; 1985-2016, DANE, Estadísticas de Demografía y Población. https://cutt.ly/ptdWbLL.
10. **pmdpxxxx** (xx inicia en 1951 y termina en el 2016; representa el potencial de mercado doméstico calculado a partir de los datos de población de cada departamento. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.66. **Fuentes**: Cálculos propios. Datos: 1951-1984, DNP, Estadísticas Históricas de Colombia. https://cutt.ly/stdWQSO; 1985-2016, DANE, Estadísticas de Demografía y Población. https://cutt.ly/ptdWbLL.
11. **pmd3yxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB corriente de cada departamento. Para el

cálculo de la distancia interna entre departamentos se usa el valor 0.33. **Fuentes:** Cálculos propios. Datos de DANE: Retropolación: https://cutt.ly/ltdWf9y; Cuentas Nacionales Departamentales: https://cutt.ly/NtdWxku.

12. **pmdyxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB corriente de cada departamento. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.66. **Fuentes**: Cálculos propios. Datos de DANE: Retropolación: https://cutt.ly/ltdWf9y; Cuentas Nacionales Departamentales: https://cutt.ly/NtdWxku.
13. **pmfpxxxx** (xx inicia en 1951 y termina en el 2016; representa el potencial de mercado exterior calculado a partir de los datos de población de cada departamento y usando como indicador de distancia la distancia expresada en kms. **Fuentes:** Cálculos propios. Datos: 1951-1984, DNP, Estadísticas Históricas de Colombia. https://cutt.ly/stdWQSO; 1985-2016, DANE, Estadísticas de Demografía y Población. https://cutt.ly/ptdWbLL. Distancias obtenidas de http://es.distancias.himmera.com/[11], entre otras.
14. **pmfyxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado exterior calculado a partir de los datos de PIB corriente de cada departamento.
15. **Fuentes:** Cálculos propios. Datos de DANE: Retropolación: https://cutt.ly/ltdWf9y; Cuentas Nacionales Departamentales: https://cutt.ly/NtdWxku. Distancias obtenidas de http://es.distancias.himmera.com/, entre otras.
16. **yxxxx**: representa PIB en términos corrientes; xxxx empieza en 1980 y termina en 2015, Millones de pesos. **Fuente**: Cálculos propios. Datos de DANE: Retropolación: https://cutt.ly/ltdWf9y; Cuentas Nacionales Departamentales: https://cutt.ly/NtdWxku.
17. **pxxxx** (xxxx inicia en 1985 – y termina en 2020): representa la población anual de cada departamento. **Fuentes**: "DANE, Estadísticas de Demografía y Población, Proyecciones de Población https://cutt.ly/otdWHYB". Proyecciones basadas en el Censo 2005. Última actualización 12/05/2011
18. **intotxxxx**: Ingresos totales (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012): (millones de pesos). **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio, https://cutt.ly/4tdWKSz.
19. **intribxxx**: Ingresos tributarios (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012): (millones de pesos). **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio, https://cutt.ly/4tdWKSz.
20. **incapxxxx**: Ingresos capital (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012): (millones de pesos). **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio, https://cutt.ly/4tdWKSz.
21. **inregaxxxx**: Ingresos regalías (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012): (millones de pesos). **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio, https://cutt.ly/4tdWKSz.
22. **fbkxxxx** (xxxx inicia en 1990 y termina en 2012**:** formación bruta de capital fijo). (millones de pesos). **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio, https://cutt.ly/4tdWKSz.

[11] Se analizaron varias páginas y calculadoras de tiempo para los cálculos de distancias por carretera con el objeto de ser lo más precisos posibles. Así, es posible encontrar las diversas rutas posibles para hacer el recorrido entre dos capitales y se optó por usar la ruta que implicaba menor tiempo de desplazamiento, aun cuando éstas requirieran un recorrido en Km más extenso. Así, esta es una cuestión asociada a la geografía y las diferentes páginas usadas para revisar y asegurar las rutas, distancias y tiempos suelen ofrecerte la misma información.

23. **htxxxx** (xxx inicia en 2000 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario como suma de DOCTORADO (3 años), ESPECIALIZACION (1 año), MAESTRIA (1-2 años), TECNICA PROFESIONAL (2 años), TECNOLOGICA (3 años) y UNIVERSITARIA (5 años) **Fuentes**: "Ministerio de Educación, Sistema Nacional de Información de la Educación Superior (SNIES), Matriculados https://cutt.ly/ZtdWVdF". Última actualización 16/05/2016.
24. **hdxxxx** (xxx inicia en 2000 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de DOCTORADO (3 años) **Fuentes**: "Ministerio de Educación, Sistema Nacional de Información de la Educación Superior (SNIES), Matriculados https://cutt.ly/ZtdWVdF". Última actualización 16/05/2016
25. **hexxxx** (xxx inicia en 2000 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de ESPECIALIZACION (1 años) **Fuentes**: "Ministerio de Educación, Sistema Nacional de Información de la Educación Superior (SNIES), Matriculados https://cutt.ly/ZtdWVdF". Última actualización 16/05/2016
26. **hmxxxx** (xxx inicia en 2000 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de MAESTRIA (1-2 años) **Fuentes**: "Ministerio de Educación, Sistema Nacional de Información de la Educación Superior (SNIES), Matriculados https://cutt.ly/ZtdWVdF". Última actualización 16/05/2016
27. **htpxxxx** (xxx inicia en 2000 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de TECNICA PROFESIONAL (2 años) **Fuentes**: "Ministerio de Educación, Sistema Nacional de Información de la Educación Superior (SNIES), Matriculados https://cutt.ly/ZtdWVdF". Última actualización 16/05/2016
28. **htcxxxx** (xxx inicia en 2000 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de TECNOLOGICA (3 años) **Fuentes**: "Ministerio de Educación, Sistema Nacional de Información de la Educación Superior (SNIES), Matriculados https://cutt.ly/ZtdWVdF". Última actualización 16/05/2016
29. **huxxxx** (xxx inicia en 2000 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de UNIVERSITARIA (5 años) **Fuentes**: "Ministerio de Educación, Sistema Nacional de Información de la Educación Superior (SNIES), Matriculados https://cutt.ly/ZtdWVdF". Última actualización 16/05/2016
30. **hgtxxxx** (xxx inicia en 2001 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario como suma de DOCTORADO (3 años), ESPECIALIZACION (1 año), MAESTRIA (1-2 años), TECNICA PROFESIONAL (2 años), TECNOLOGICA (3 años) y UNIVERSITARIA (5 años) **Fuentes**: "Observatorio Laboral para la Educación, Graduados y Estudiantes, Sistema de Información. https://cutt.ly/TtdW8jx".
31. **hgdxxxx** (xxx inicia en 2001 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de DOCTORADO (3 años) **Fuentes**: "Observatorio Laboral para la Educación, Graduados y Estudiantes, Sistema de Información. https://cutt.ly/TtdW8jx"
32. **hgexxxx** (xxx inicia en 2001 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de ESPECIALIZACION (1 años) **Fuentes**: "Observatorio Laboral para la Educación, Graduados y Estudiantes, Sistema de Información. https://cutt.ly/TtdW8jx"
33. **hgmxxxx** (xxx inicia en 2001 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de MAESTRIA (1-2 años) **Fuentes**: "Observatorio Laboral para la Educación, Graduados y Estudiantes, Sistema de Información. https://cutt.ly/TtdW8jx"
34. **hgtpxxxx** (xxx inicia en 2001 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de TECNICA PROFESIONAL (2 años) **Fuentes**: "Observatorio Laboral para la Educación, Graduados y Estudiantes, Sistema de Información. https://cutt.ly/TtdW8jx"

35. **hgtcxxxx** (xxx inicia en 2001 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de TECNOLOGICA (3 años) **Fuentes**: "Observatorio Laboral para la Educación, Graduados y Estudiantes, Sistema de Información. https://cutt.ly/TtdW8jx"
36. **hguxxxx** (xxx inicia en 2001 y termina en 2015): total de personal- hombres y mujeres- con capital humano terciario de UNIVERSITARIA (5 años) **Fuentes**: "Observatorio Laboral para la Educación, Graduados y Estudiantes, Sistema de Información. https://cutt.ly/TtdW8jx"
37. **pmd3tpxxxx** (xx inicia en 1951 y termina en el 2016; representa el potencial de mercado doméstico calculado a partir de los datos de población de cada departamento y usando como indicador de distancia la distancia expresada en minutos. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.33 y se expresa en minutos. Si consideramos una velocidad de crucero de 60km/h, la distancia interna dentro de cada departamento expresada en minutos se corresponde a la distancia en kms (pmd3tpxxxx=pmd3pxxx)
38. **pmdtpxxxx** (xx inicia en 1951 y termina en el 2016; representa el potencial de mercado doméstico calculado a partir de los datos de población de cada departamento y usando como indicador de distancia la distancia expresada en minutos. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.66 y se expresa en minutos. Si consideramos una velocidad de crucero de 60km/h, la distancia interna dentro de cada departamento expresada en minutos se corresponde a la distancia en kms (**pmdtpxxxx=pmdpxxx**)
39. **pmd3tyxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB corriente de cada departamento y usando como indicador de distancia la distancia expresada en minutos. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.33 y se expresa en minutos. Si consideramos una velocidad de crucero de 60km/h, la distancia interna dentro de cada departamento expresada en minutos se corresponde a la distancia en kms (**pmd3tyxxxx=pmd3yxxx**)
40. **pmdtyxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB corriente de cada departamento y usando como indicador de distancia la distancia expresada en minutos. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.66 y se expresa en minutos. Si consideramos una velocidad de crucero de 60km/h, la distancia interna dentro de cada departamento expresada en minutos se corresponde a la distancia en kms (**pmdtyxxxx=pmdyxxxx**)
41. **pmftpxxxx** (xx inicia en 1951 y termina en el 2016; representa el potencial de mercado exterior calculado a partir de los datos de población de cada departamento y usando como indicador de distancia la distancia expresada en minutos. **Fuentes:** Cálculos propios. Datos: 1951-1984, DNP, Estadísticas Históricas de Colombia. https://cutt.ly/stdWQSO; 1985-2016, DANE, Estadísticas de Demografía y Población. https://cutt.ly/ptdWbLL. Distancias obtenidas de http://es.distancias.himmera.com/, entre otras.
42. **pmftyxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado exterior calculado a partir de los datos de PIB corriente de cada departamento y usando como indicador de distancia la distancia expresada en minutos. **Fuentes:** Cálculos propios. Datos de DANE: Retropolación: https://cutt.ly/ItdWf9y; Cuentas Nacionales Departamentales: https://cutt.ly/NtdWxku. Distancias obtenidas de http://es.distancias.himmera.com/, entre otras.
43. **peaxxxx** (xxxx inicia en 2001 y termina en 2015) representa la población económicamente activa. **Fuente:** DANE, Mercado Laboral por Departamentos. http://www.dane.gov.co/files/investigaciones/boletines/ech/ml_depto/anexo_dep_16.xls
44. **petxxxx** (xxxx inicia en 2001 y termina en 2015) representa la población en edad de trabajar. **Fuente:** DANE, Mercado Laboral por Departamentos. http://www.dane.gov.co/files/investigaciones/boletines/ech/ml_depto/anexo_dep_16.xls

45. **viasxxxx** (xxxx inicia en 1949-1998, 2003-2004, 2006 y termina en 2015) representa la red vial pavimentada total. (en kilómetros) **Fuentes:** Memorias de Obras Públicas 1949,1960,1966,1970 y 1974. Ministerio del Transporte, Documentos, Estadísticas. https://www.mintransporte.gov.co/Documentos/documentos_del_ministerio/Estadisticas
46. **viasaxxxx** (xxxx inicia en 1949-1998, 2003-2004, 2006 y termina en 2015) representa la red vial afirmada total. (en kilómetros) **Fuentes:** Memorias de Obras Públicas 1949,1960,1966,1970 y 1974. Ministerio del Transporte, Documentos, Estadísticas. https://www.mintransporte.gov.co/Documentos/documentos_del_ministerio/Estadisticas
47. **yrxxxx**: representa PIB en términos reales; xxxx empieza en 1980 y termina en 2015, Millones de pesos constantes del año base 2005. **Fuente**: Cálculos propios. Datos de DANE: Retropolación: https://cutt.ly/ItdWf9y; Cuentas Nacionales Departamentales: https://cutt.ly/NtdWxku.
48. **pmd3yrxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB real (año base 2005) de cada departamento. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.33. **Fuentes**
49. **pmdyrxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB real (año base 2005) de cada departamento. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.66. **Fuentes**
50. **pmfyrxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB real (año base 2005) de cada departamento
51. **pmd3tyrxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB real (año base 2005) de cada departamento y usando como indicador de distancia la distancia expresada en minutos. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.33 y se expresa en minutos. Si consideramos una velocidad de crucero de 60km/h, la distancia interna dentro de cada departamento expresada en minutos se corresponde a la distancia en kms (pmd3tyrxxxx=pmd3yrxxx).
52. **pmdtyrxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado doméstico calculado a partir de los datos de PIB real (año base 2005) de cada departamento. y usando como indicador de distancia la distancia expresada en minutos. Para el cálculo de la distancia interna entre departamentos se usa el valor 0.66 y se expresa en minutos. Si consideramos una velocidad de crucero de 60km/h, la distancia interna dentro de cada departamento expresada en minutos se corresponde a la distancia en kms (pmdtyxxxx=pmdyxxxx)**.**
53. **pmftyrxxxx** (xx inicia en 1985 y termina en el 2015; representa el potencial de mercado exterior calculado a partir de los datos de PIB real (año base 2005) de cada departamento y usando como indicador de distancia la distancia expresada en minutos.
54. **intotrxxxx**: Ingresos totales en términos reales (año base 2005) (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012). En millones de pesos, deflactado con el Deflactor implícito de la serie retropolada por DANE. **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio. https://cutt.ly/4tdWKSz."
55. **intribrxxx**: Ingresos tributarios en términos reales (año base 2005) (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012). En millones de pesos, deflactado con el Deflactor implícito de la serie retropolada por DANE. **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio. https://cutt.ly/4tdWKSz."
56. **incaprxxxx**: Ingresos capital en términos reales (año base 2005) (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012). En millones de pesos, deflactado con el Deflactor implícito

de la serie retropolada por DANE. **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio. https://cutt.ly/4tdWKSz."

57. **inregarxxxx**: Ingresos regalías en términos reales (año base 2005) (de la hoja Eje.Dptales) (xxxx inicia en 1984 y termina en 2012). En millones de pesos, deflactado con el Deflactor implícito de la serie retropolada por DANE. **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio. https://cutt.ly/4tdWKSz."
58. **fbkrxxxx** (xxxx inicia en 1990 y termina en 2012**:** (formación bruta de capital fijo en términos reales (año base 2005). En millones de pesos, deflactado con el Deflactor implícito de la serie retropolada por DANE. **Fuentes**: "DNP, Ejecuciones Presupuestales de cada Departamento y Municipio. https://cutt.ly/4tdWKSz."
59. **ypxxxx:** representa PIB del sector primario en términos corrientes; xxxx empieza en 2000 y termina en 2016, Millones de pesos. **Fuente**: Cálculos propios. Datos DANE, Cuentas Departamentales. http://www.dane.gov.co/files/investigaciones/pib/departamentales/B_2005/Ramas_actividad_2015provisional.xls
60. **ysxxxx:** representa PIB del sector primario en términos corrientes; xxxx empieza en 2000 y termina en 2016, Millones de pesos. **Fuente**: Cálculos propios. Datos DANE, Cuentas Departamentales. http://www.dane.gov.co/files/investigaciones/pib/departamentales/B_2005/Ramas_actividad_2015provisional.xls
61. **ytxxxx:** representa PIB del sector primario en términos corrientes; xxxx empieza en 2000 y termina en 2016, Millones de pesos. **Fuente**: Cálculos propios. Datos DANE, Cuentas Departamentales. http://www.dane.gov.co/files/investigaciones/pib/departamentales/B_2005/Ramas_actividad_2015provisional.xls
62. **yprxxxx:** representa PIB del sector primario en términos reales; xxxx empieza en 2000 y termina en 2016, Millones de pesos constantes del año base 2005. **Fuente**: Cálculos propios. Datos DANE, Cuentas Departamentales. http://www.dane.gov.co/files/investigaciones/pib/departamentales/B_2005/Ramas_actividad_2015provisional.xls
63. **ysrxxxx:** representa PIB del sector secundario en términos reales; xxxx empieza en 2000 y termina en 2016, Millones de pesos constantes del año base 2005. **Fuente**: Cálculos propios. Datos DANE, Cuentas Departamentales. http://www.dane.gov.co/files/investigaciones/pib/departamentales/B_2005/Ramas_actividad_2015provisional.xls
64. **ytrxxxx:** representa PIB del sector terciario en términos reales; xxxx empieza en 2000 y termina en 2016, Millones de pesos constantes del año base 2005. **Fuente**: Cálculos propios. Datos DANE, Cuentas Departamentales. http://www.dane.gov.co/files/investigaciones/pib/departamentales/B_2005/Ramas_actividad_2015provisional.xls

Chapter 2 - The Effects of Transfers of Royalties to Subnational Governments on Regional Formal Employment in Colombia

Abstract

We assess the effects on formal employment, of a subset of the royalties that are transferred to Colombian municipalities that seek to progressively redistribute among all municipalities, a share of the royalties generated from the exploitation of nonrenewable resources. The targeting of these royalties allows us to identify its effect on employment by means of a regression discontinuity methodology. We find positive and significant effects of these royalties on formal employment, mainly in the sectors of financial intermediation, real estate, the union of industry and commerce, and to a lesser extent, on agriculture. A one percentage point increase in the level of royalties would increase in 0.78-1.0% formal employment. This is equivalent to say that investing nearly 2 monthly minimum wages per year of these royalties, would leverage investment resources of nearly 8 minimum wages per year per new formal employment generated. These results are robust to several specifications regarding the nonparametric estimation and the control variables included. Furthermore, we find that royalties coming from the RCF increased the amount of resources from the RCF supporting subnational investment projects without reducing the amount of other sources of funding, in particular, of the municipalities' own resources. Finally, we find that these royalties did not affect the municipalities' fiscal effort, keeping their initial tax earnings unaffected.

JEL CODES: C21; C30; C90; J23

KEYWORDS: Royalties, Formal Employment, Regression Discontinuity.

2.1 Introduction

The collection of non-renewable royalties provides a flexible fiscal policy tool for national and sub-national governments in order to promote regional growth, equity between regions, lower poverty rates, and increase competitiveness (Oates 1972, 1993; Tanzi, 1995). The role of royalties from non-renewable resources became particularly important in early 2000s when the world experienced a boom in commodity prices that led to a period of substantial increases in commodities' exports from many developing economies (Van der Ploeg, 2011). Recent studies have provided convincing evidence on the importance of the royalties on economic and social outcomes and fiscal effort.[12] Notably, literature has emphasized on the so-called "resource curse" and how it can be exacerbated by such extraordinary booms (Brunnschweiler, 2008; Van der Ploeg, 2011; Dude and Vargas, 2013; Caselli and Micheals, 2013; Warner 2015; Dauvina and Guerreiro, 2017; Santos, 2018). Recently, literature in economics have focused on the role institutions play to alleviate or aggravate the resource curse (Atkinson and Hamilton, 2003; Mehlun et al. 2006; Robinson, Torvik and Verdierc, 2006; Bjorvatn, et al. 2012; Boschini, Pettersson, and Roine, 2012; De Medeiros and Dos Santos, 2013; Robinson et al. 2014; Kanna, 2017).

A key institution in this set up is the set of rules that incorporated non-renewable royalties into the general fiscal regime of the governments that in turn provides the main mechanism to redistribute all resources between national and sub-national governments. After the boom, several Latin-American countries (Brosio and Jimenez, 2012), Colombia among them, became aware that those extraordinary economic resources could serve other purposes such as contributing to a more equitable allocation of the additional rents across sub-national governments.

In 2011, Colombia Government replaced the National Royalties Fund for the General System of Royalties (SGR) in order to improve regional growth, equity among regions, reduce poverty rates and increase competitiveness (Bahl, 2000). The SGR was proposed given the poor results associated with the previous system (Echeverry, Alonso and García, 2011; Urban Economics Ltda- National Consulting Center, 2012; Cardenas, 2013) and its regressive distribution (Bonet and Urrego, 2014).[13]

In this chapter, we shed light on the effect of the reform on formal employment and fiscal effort through the new set of rules that govern the allocation of royalties. In particular, we aim to identify whether the rise in the share of royalties, for municipal entities receiving smaller proportions before the reform, has boosted formal employment and reduce the fiscal effort. To identify this effect, we use a regression discontinuity design by exploiting the way the in which royalties are distributed between municipalities after the reform. [14]

[12] Sachs and Warner (1995) provides a more general analysis of the relationship between natural resource exports and GDP.
[13] According to Cardenas (2013), the old system had low impact and results. His diagnostic is based on: (i) the existence of many fragmented projects of low social and economic impact, (ii) poor planning and execution, (iii) as of 2009, no local authority was certified in all coverages?, and (iv) producers: high dependency on royalties, low levels of tax collection, low levels of sustainable development.
[14] The methodology is similar to that previously applied in studies done by Angrist and Lavy (1999), Van der Klaauw (2002), Hahn, and Van der Klaauw Todd (2001), and Ferraz and Finan (2010) and by Corbi, Papaioannou and Surico (2014).

Before the reform, 80% of the royalties were received by territorial entities that represented 17% of the population, and with the reform is expected that the same 80% of royalties would benefit 70% of the population. The new SGR dramatically changed the regional distribution of royalties. Between 2002 and 2010, only two departments, Casanare and Meta, received most of the royalties, and fewer than 500 municipalities benefit from these resources. In contrast, the SGR foresees that all Colombian territories should receive royalties, regardless of whether they extracted natural resources or not. In addition, the reform created the Regional Compensation Fund (RCF), which would transfer royalties to 1.088 municipalities, with greater resources to the poorest municipalities. We exploit the fact that regulation allocates royalties to municipalities depending on whether a municipality's population has their unsatisfied basic needs index (UBN) lower or higher than 35% on municipalities with 100 thousand inhabitants or less (or municipalities in categories 4 to 6 in Colombian regulations).

We estimate the effect of royalties on formal employment using the Calonico et al. (2014) non-parametric regression discontinuity (RD) model, comparing municipalities above and below the 35% UBN threshold. The key assumptions are that there is no manipulation of the forcing variable, UBN, and baseline characteristics vary smoothly at the 35% UBN cutoff. We show that these assumptions hold and then we interpret our estimates as causal effects. We also show a clear discontinuity in the amount transferred as royalties at the UBN cutoff. The first stage estimation suggests that municipalities around the cutoff whose proportion of population with UBN exceeded 35% leads to receive 53-62% less royalties in reference to municipalities below the threshold.

Using administrative records of the mandatory contributions to health, pensions and disability insurance from the Ministry of Social Protection, we assess the effects of royalties on formal employment. Our main result shows that increases in royalties, associated with discontinuity on the 35% UBN threshold, lead to a substantial and statistically significant rise of formal employment. In particular, one percentage point increase in the level of royalties explain an increase of 0,78-1,0% on formal employment. Since all formal worker in Colombia must pay for their access to the health and pensions system, this result also implies a higher access to these social services as a results of the reform. Interesting, the results are not homogenous by economic sector. The sectors that would have benefited the most of the new system are financial intermediation, real state, manufacturing and electricity. Nonetheless, we do not observe crowding-out effect between sectors.

In addition, we explore the effects of royalties on the fiscal effort of the municipalities. We expected that municipalities receiving additional royalty transfers might either complement the additional resources with their own taxes, which would require them to increase their fiscal effort, or substitute their own resources with the additional royalty transfers, lowering their fiscal effort. To test our hypothesis, we first use administrative records from the Territorial Single Form (FUT by its acronym in Spanish), which contains the accounting balance of all sub-national governments (states, municipalities, etc.). We estimate the effects of the royalty transfers on the fiscal effort municipalities exert on their taxes earnings, and in particular, on their property tax and the tax to industries and commerce. Although the point estimates are negative, the effects are not statistically significant at conventional levels.

Second, we use administrative records from the Department of National Planning (DNP), which contains the census of all investment projects registered in its Bank of Investment Projects (BPIN by its acronym in Spanish) in years 2013, 2014 and 2015. The data contains all investment projects by year and entity (states, municipalities, private entities, public firms, etc.) in Colombia.[15] Using information at the municipality level, we estimate the effect of the royalties transferred to municipalities from the RCF on the logarithm of the amount funded with resources of the RCF, and on the logarithm of the amount funded with their own resources. We estimate a negative coefficient on the logarithm of the amount funded with their own resources, although the effects are not statistically significant at conventional levels. Conversely, we found a positive coefficient on the logarithm of the amount funded with resources of the RCF, although the effects are significant at the 10% level. Overall, we do not find evidence of change on the fiscal effort of the municipalities.

Our findings contribute mainly to three literatures. First, we contribute to the existing literature about the role of the institutions to overcome the resource curse by exploiting an exogenous change in the set of rules that distribute non-renewable royalties between national and sub-national governments. Colombia's reform creates a new set of rules that transfer more royalties to municipalities with more unsatisfied basic needs. This change allows the beneficiaries municipalities to increase the investment, without affect the fiscal effort, and increase the formal employment. We show that using a targeting that take in account the needs of the municipality may lead to a reversal effect on the resource curse. Our results complement the findings of Robinson, Torvik, and Verdierc (2006), Bjorvatn, et al. (2012) and Medeiros and Dos Santos (2013) that show that government with stronger institutions and higher accountability may help to overcome the resource curse. We also contribute to the more general literature on resource curse.

The remainder of the paper presents background information and the institutional framework; it then presents the empirical strategy to identify the effects of royalties on formal employment, the data, and the empirical results, to finally provide some concluding remarks.

2.2 Background

According to Echeverry, Alonso and García (2011), the reform that created the General System of Royalties, was based on four principles: (i) social and regional equity, (ii) savings for the future, (iii) regional competitiveness, and (iv) good governance. The Government replaced the previous scheme with the General System of Royalties in order to use this important resource to promote regional growth, equity among regions, reduce poverty rates and increase competitiveness, all of those very reasonable and acceptable goals (Bahl, 2000).

The new General Royalties System (SGR) dramatically changed the regional distribution of royalties. As shown in Figure 2.1, whereas between 2002 and 2010 only two departments,

[15] The municipalities in our sample account for a total of 6,657 projects, 4,622 of which were at least partly funded with resources of the Regional Compensation Fund.

Casanare and Meta, received royalties for more than $3,2 billion each, it was expected that between 2012 and 2020, 8 departments were to receive at least that amount: La Guajira, Cesar, Bolivar, Cordoba, Antioquia, Casanare, Meta and Valle del Cauca. Some of these changes are due to the new allocation of the SGR, and partly due to the increase that is was expected to be given from then on in resources by concepts of royalties. Echeverry et al. (2011) projected that they could grow between 2010 and 2020 up to 80 percent. As shown on the map, while between 2002 and 2010 there were 19 departments receiving less than $0,8 billion, the expectation with the reform was that between 2012 and 2020 all departments but Quindío, would receive a higher amount.

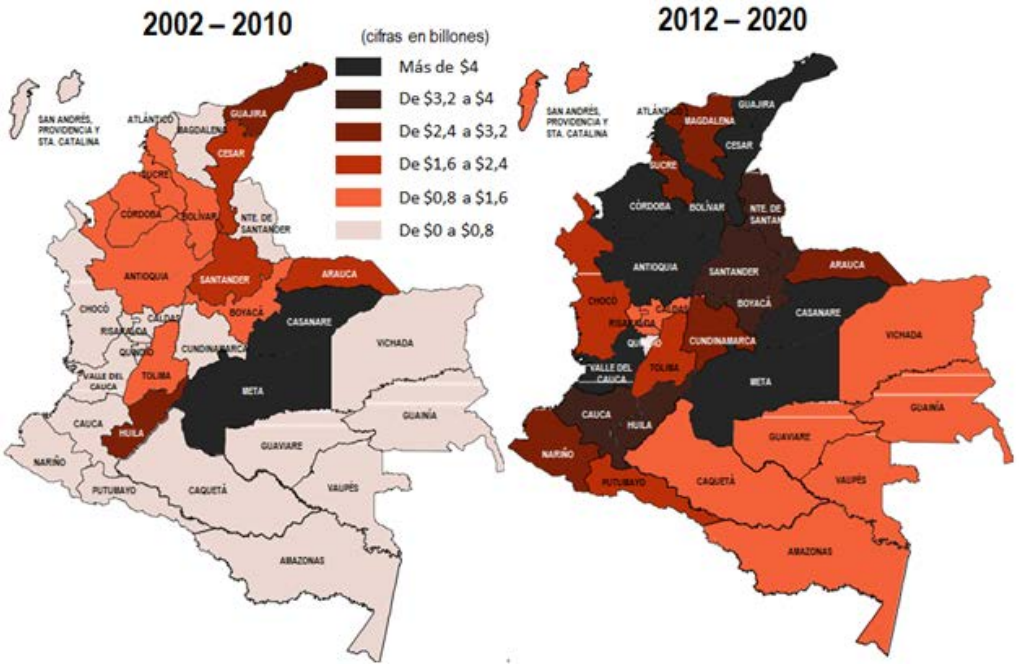


Figure 2.1. Total Royalties department: before and after the reform**.** Source: DNP (2013)

Another way to illustrate the changes in the amount of royalties received by the different regions is in terms of the per capita allocation for each municipality before and after the reform. Figure 2.2 illustrates the location of each municipality according to per capita royalties received between 2010 and 2011 (dark diamonds), and according to those received between 2011 and 2012 (clear diamonds). The graph on the left includes all municipalities, and the one on the right only includes those who received less than $1+ million per capita, in order to better illustrate what happened within them. The first set of points (2010-2011) shows that on average, per capita royalties increased for most municipalities, and in particular for municipalities that were producers, i.e. those receiving direct royalties, among which the figure distinguishes the members of the Association of Oil, Mining and Port Municipalities of Colombia (AMPET of Colombia) with a cross. On the other hand, the points that illustrate the allocation for the years 2011 and 2012 show a clear decrease in the amount of royalties per capita. This decrease is due in part to the fact that in 2012, the total amount of royalties had a significant decrease.

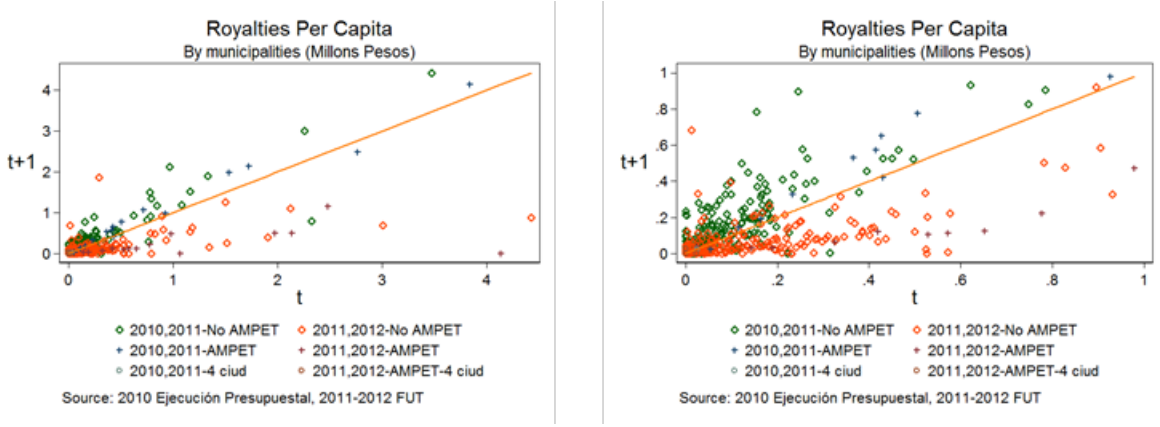


Figure 2.2. Royalties per capita by municipality. Source: budget execution 2011-2012

To better illustrate the changes in distribution within each year, Figure 2.3 shows the shares of royalties for each municipality. It is important to point out that the participation of municipalities in the AMPET are on a steeper diagonal between 2011 and 2012 (clear diamonds) than between 2010 and 2011 (dark diamonds), which means that their participation was reduced with the SGR.

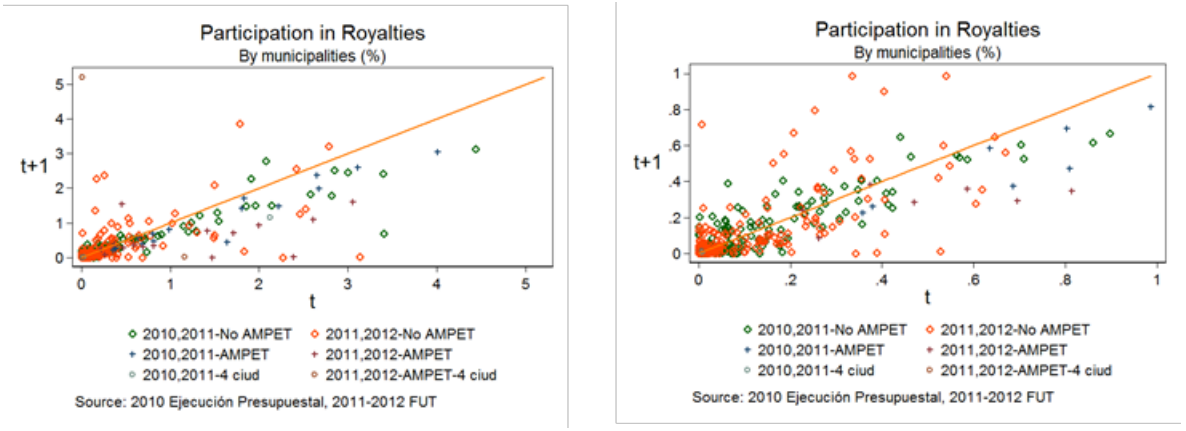


Figure 2.3. Participation of municipalities in royalties. Source: budget execution 2011-2012

To illustrate the relationship between employment and royalties, Figure 2.4 presents the performance of employment in the 80 municipalities that receive direct royalties and most of the other municipalities. The graph on the left shows the number of employees, and the one on the right shows the number of total jobs (each employee can have more than one job). The figure suggests that the royalties could have an impact on higher employment rates in municipalities that do not belong to the select group of the 80 municipalities that receive more direct royalties, and incidentally, would benefit most by the SGR.

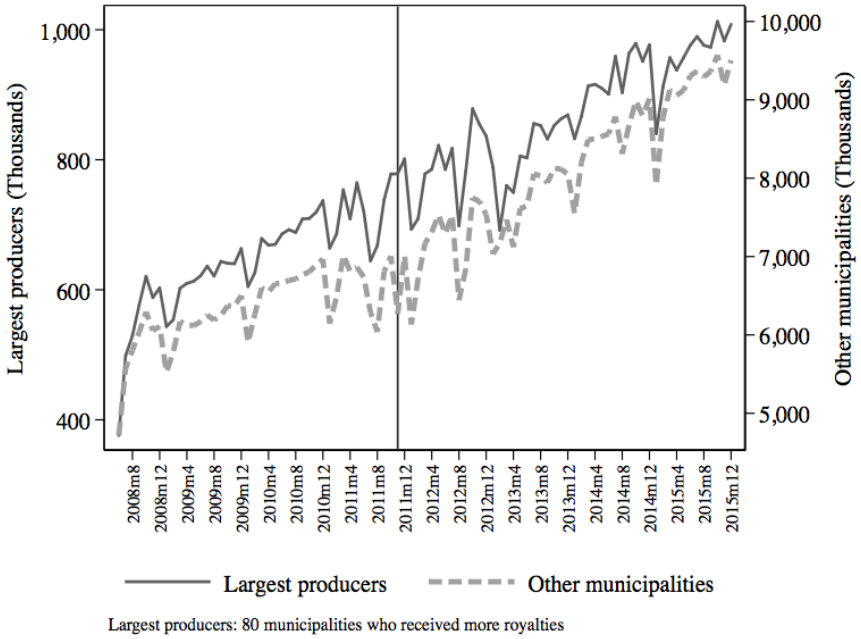


Figure 2.4. Employment trends: producing versus non-producing municipalities. Source: Integrated Contribution Liquidation Form, MSPS.

Then, we analyze the effect of royalties under the post-reform redistribution, on the formal labor market. We seek to identify whether municipalities having an increase in the amount of royalties received, increased their formal employment.

2.3 Institutional Framework[16]

Colombia's Constitution is clear in stating: (i) the resources of the subsoil, and non-renewable resources, belong to the State, (ii) anyone that exploits a nonrenewable natural resource will cause, in favor of the State, an economic payment by way of a royalty, and (iii) the departments, districts and municipalities, where nonrenewable natural resources were exploited, as well as municipalities and districts with sea and river ports where such resources or byproducts are transported through, shall be entitled to participate of those royalties and compensations and directly invest those resources.[17]

This constitutional mandate, and the respective legislation that supplements it, led the country to transfer significant resources to the regions in which non-renewable resources were exploited, while other regions, with precarious levels of development, had no opportunity to benefit from the mining and energy booms that have occurred. To correct this situation, the government promoted a reform that sought to ensure the responsible and efficient management of royalties, to save some of those resources for times in which the resources generated by this sector were substantially reduced, and ensure a more equitable distribution

[16] This section is based on DNP (2013)
[17] Articles 332, 360 and 361 of the Colombian Constitution.

of those resources across regions to promote the balanced development and growth of the country and its regions.

Before the reform, 80% of the royalties were received by territorial entities that represented 17% of the population, and initially, it was estimated that with the reform, that same 80% of royalties would benefit 70% of the population. Additionally, the reform created the Regional Compensation Fund (RCF), which would transfer royalties to 1,088 municipalities starting from 2012, with greater resources to the poorest municipalities, compared with fewer than 500 municipalities receiving royalties under the previous scheme.

As shown in Figure 2.5, the reform provides resource savings for pensions (10%), science, research and development (10%), and additionally, provides resource savings for the regions and for a Stabilization Fund (25-30%), in such a way that it expected that in the future, 40% of the resources of the SGR would be earmarked to savings, and 60% to investment. It also foresees that all Colombian territories could receive royalties, regardless of whether they extracted natural resources or not.

In this document, we use the targeting criteria of the Regional Compensation Fund, RCF, which in 2014 paid out approximately COP$ 1,25 trillion in royalties. While 60% of the resources of the RCF is distributed among the departments based on their population, unemployment and poverty, the remaining 40% is distributed among municipalities, 30% of which are distributed on the basis of their poverty, and the remaining 10% based on their municipal category, poverty and population.[18] In particular, 30% of the resources are distributed according to the population of the municipalities, provided that the fraction of their population with unsatisfied basic needs, UBN, is greater than 35%. The remaining 10% is distributed among the municipalities in categories 4, 5 and 6 that had not received resources from the first 30%, and with UBN $\leq$ 35%, based on both, their population and UBN.

[18] Law 617 of the year 2000 defines seven categories of municipalities, a "*Especial*" category for municipalities with more than 2 million inhabitants, and the categories 1-6, where 1 refers to the most populated municipalities and those with higher levels of own collected local fiscal revenue and 6 refers to the least populated municipalities, and those with lowest fiscal revenues.

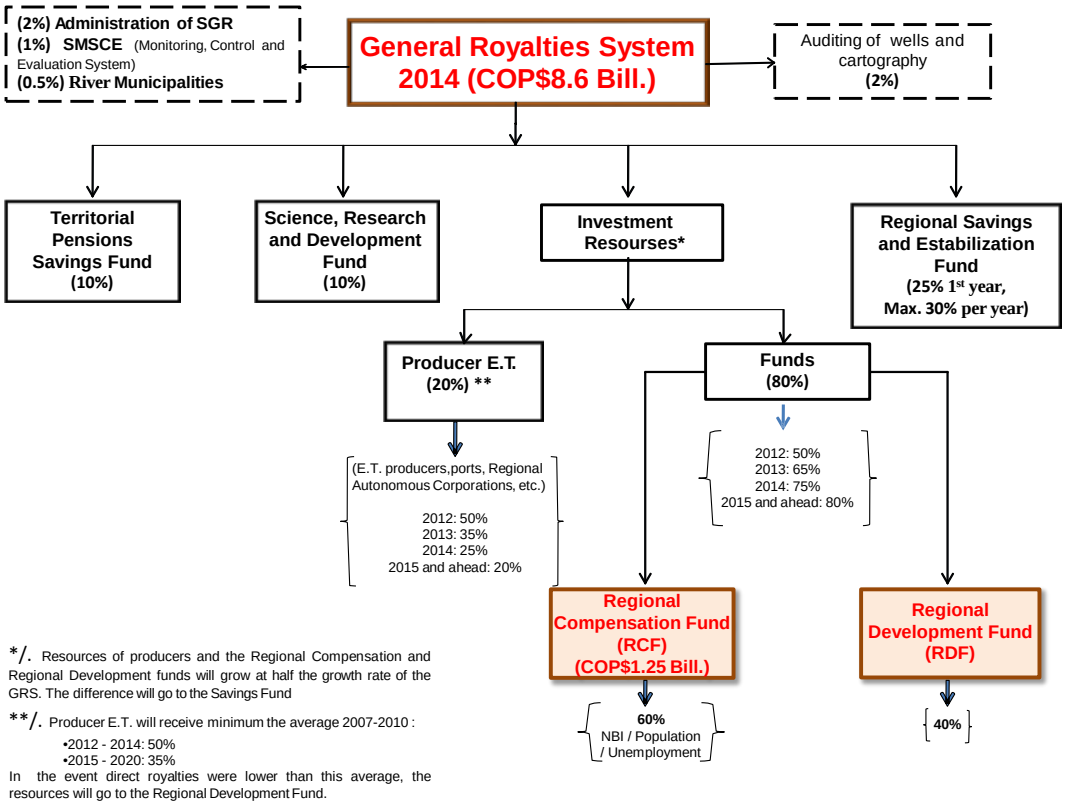


Figure 2.5. Distribution of Resources of the General System of Royalties, 2014

2.4 Empirical Strategy

The amount of royalties received by the municipalities is plausibly correlated with a wide range of socio-economic attributes, and thus, a naive comparison of employment across municipalities according to the level of royalties may be confounded by those attributes. Particularly, unobservable variables are expected to be correlated with royalties. For example, it may be possible that in some cases, decision makers seek to facilitate the payments of resources to municipalities with large vulnerabilities, providing support in the formulation and presentation of investment projects. Moreover, to the extent that municipalities with greater institutional capacity are more likely to formulate and present the best projects, it becomes more likely for them to receive higher transfers of royalties while simultaneously it is likely that greater institutional capacity is positively correlated with formal employment of the municipality, causing the omission of this institutional capacity variable to overestimate the effect of royalties on employment. We would also be overestimating this effect if less institutional capacity would lead to suspending payments to the municipality (receiving fewer royalties) for violating the requirements.[19]

[19] Decree 416 of the year 2007 orders the suspension of Royalty payments when the municipality fails to mobilize resources six months after their first payment, if the municipality suspends projects for more than six months during the last 12 months, if the municipality spends resources from royalties with a different destination than permitted by law, etc.

To solve this endogeneity problem, we exploit the fact that regulation allocates royalties to municipalities discontinuously at the UBN threshold of 35%. In particular, the law orders the RCF: (i) to transfer 30% of its resources to municipalities with a UBN index above the 35% threshold based on their population, and (ii) to transfer 10% of its resources, to municipalities that did not receive transfers according to (i), which belonged to municipal categories 4th, 5th or 6th, and had a UBN ≤ 35%, transferring 60% of these resources based on the population of these municipalities, and 40% of them based on their poverty measured by the UBN.[20]

Given these rules, we assess and exploit the existence of a discontinuous change in the royalty transfers around the 35% UBN threshold, based on which implement a regression discontinuity approach to evaluate the impact of royalties on formal employment. Although the law assigns a higher share of this 40% of resources of the RCF to municipalities above the 35% UBN cutoff, the way resources are distributed among eligible municipalities on both sides of the cutoff varies. First, there are only 85 municipalities in categories *Especial*, 1, 2 and 3: 46 with UBN ≤ 35% and 39 with UBN > 35%, and only 3 (8) of them in a bandwidth of ±5% (±10%), which seems too few for robustly allowing us to learn of the effects of royalties on them. In addition, among those with UBN ≤ 35%, only municipalities in categories 4, 5 and 6 are eligible for royalties. Given this, we focus the estimation of the effects of royalties on the subset of municipalities in categories 4, 5 and 6. Secondly, municipalities in categories 4, 5 and 6, represent 53% percent of the population with UBN > 35%, while they represent 29% of those with UBN ≤ 35%, meaning that according to these shares, royalties are on average distributed among a smaller population on the left hand side of the UBN cutoff.

Additionally, once we focus on the ±10% bandwidth, that is, on municipalities with 25% ≤ UBN ≤ 45%, the eligible population in categories 4, 5 and 6, and in the 35-45% bandwidth, represents 13% of the total population above the cutoff, that is, the share of the 30% of resources of the RCF assigned to them is just 13% of that 30%, while the eligible population in the 25-35% bandwidth represents 48.7% of the total population below the cutoff. This implies that municipalities in the 25-35% UBN bandwidth would receive about $383 per capita per each $100 millions of royalties, while those in the 35-45% UBN bandwidth would receive about $351, that is, a smaller amount per capita.[21] This unintended by the legislators, consequence of the law, is basically due to the fact that all municipalities above the 35% UBN cutoff were eligible, while only those in categories 4, 5 and 6 were, which ended up benefitting in per capita terms municipalities on the left of the cutoff (those relatively better off) in relation to those above it.

Our empirical analysis is based on a two-stage regression. The first stage is estimated according to:

$$ln(R_m) = \beta_0 + \beta_1 I[UBN_m \geq \overline{UBN}] + \beta_2 f(UBN_m) + \mu_m (1)$$

[20] See Law 1530 of the year 2012 and Decree 1073 of the year 2012.

[21] They would actually receive even less than $351, since resources transferred to municipalities above the 35% UBN cutoff are also distributed based on the UBN, and municipalities closer to the cutoff would receive a smaller share that the average municipality in that UBN range.

Where $ln(R_m)$ represents the logarithm of royalties transferred to municipality m and $I[UBN_{dm} \geq \overline{UBN}]$ is a dummy variable equals one if its *UBN* is greater than or equal to the cutoff level $\overline{UBN}$. $f(UBN_m)$ is a continuous piecewise polynomial in the *UBN*, and μ_m is an idiosyncratic error term.[22] The second stage is based on:

$$ln(E_m) = \alpha_0 + \alpha_1 E[ln(R_m)|UBN_i] + \alpha_2 f(UBN_m) + \varepsilon_i \ (2)$$

Where $ln(E_w)$ is the logarithm of formal employment in municipality m and $E[ln(R_m)|UBN_i]$ is the predicted value of $ln(R_m)$ estimated in the first stage (Van der Klauuw, 2002).[23]

We estimate α_1, our parameter of interest, both parametrically and non-parametrically. In both cases, the choice of bandwidth involves a trade-off between bias and efficiency. To deal with this trade-off, baseline estimates report the optimal bandwidth, bias correction, and robust standard errors proposed by Calonico, Cattaneo and Titiunik (2014a). These estimates are a refinement of the non-parametric local polynomial estimators usually employed.

The Royalties that are used in the estimation are the 40% of total paid royalties coming from the Regional Compensation Fund which are described in article 34.2 of law 1530 of 2012, as reported by the Department of National Planning, DNP, and received by each municipality after the 2011 reform.[24] Formal employment is obtained based on the Unified Register of Contributions (PILA for its acronym in Spanish), from the Ministry of Health and Social Protection, MSPS. The PILA contains the formal census of workers in the country who monthly contribute to health and pensions.

[22] Following Gelman and Imbens (2014), we include only linear and quadratic polynomials, so, $f(UBN_i)$ is a continuous piecewise (linear or quadratic) polynomial in the municipality's *UBN*, that is, $f(UBN_m) = \sum_{k=0}^{M} \varphi_{ok} UBN_i^k + \sum_{k=1}^{M} \varphi_{1k} (UBN_i - \overline{UBN})^k 1[UBN_i \geq \overline{\text{UBN}}]$ with M = 1, 2 for the linear and quadratic cases respectively.

[23] For our main specification we use the logarithm of both employment and royalties (plus one, to allow for zero values) instead of including the levels, since this reduces the skewness in both variables, which (measured in levels) have a skewness of 13.74 and 8.17 respectively. These variables in levels also have a standard deviation that is larger than their means (See descriptive statistics in Table 2.1).

[24] Note that we do not include the 30% of the RCF that is transferred **to departments** with municipalities with an UBN > 35%, which are described in article 34.1.2 of law 1530 of 2012, since those resources are transferred to the departments rather than to their municipalities.

Table 2.1. Summary Statistics

	Mean	*Median*	*SD*	*Min*	*Max*
Panel A. Dependent variables					
Employment	996.9	222.6	3,415.6	7.3	76,098.4
Log of (one plus) employment	5.686	5.410	1.366	2.120	11.240
... in 2013	5.562	5.306	1.357	2.552	11.122
... in 2015	5.719	5.442	1.387	2.148	11.303
... in Agriculture	2.535	2.294	1.800	0.000	8.103
... in Mining	1.451	0.811	1.645	0.000	7.777
... in Manufacturing	1.811	1.266	1.755	0.000	8.356
... in Electricity	0.939	0.000	1.344	0.000	5.741
... in Construction	2.147	1.969	1.740	0.000	8.746
... in Commerce	2.836	2.657	1.888	0.000	9.000
... in Transport	1.771	1.266	1.688	0.000	8.216
... in Financial intermediation	2.263	2.026	1.425	0.000	8.391
... in Real State	3.974	3.722	1.601	0.000	10.179
... in Social and personal services	4.698	4.527	1.207	0.969	10.164
Panel B. Treatment and control variables					
Royalties FCR (millions of pesos)	853.9	555.9	1,302	22.8	20,325
Log of Royalties FCR	20.147	20.136	0.858	16.942	23.735
... in 2013	20.497	20.479	0.822	18.000	24.056
... in 2015	19.888	19.871	0.849	17.348	23.455
Direct royalties (millions of pesos)	393.4	0.0	2,428	0.0	58,923
Log of (one plus) direct royalties	5.347	0.000	8.375	0.000	24.800
Prob(Direct royalties>0)	0.301	0.000	0.459	0.000	1.000
UBN	44.958	43.166	18.081	8.944	98.814
Revenues per capita	604.4	565.9	526.0	0.0	6,259
Property tax per capita	22.875	10.703	34.625	0.000	584.972
Transferences per capita	455.3	419.7	408.6	0.0	4,265
Population	19,501	12,393	27,166	1,004	409,542
Category 4, 5 and 6	0.990	1.000	0.099	0.000	1.000

Sources: formal employment is obtained from the PILA fo the MSPS, royalty figures are provided by the DNP, the UBN from DANE and the fiscal figures from the Banco de la República and the DNP. The figures correspond to 1019 municipalities.

2.5 Results

We begin by presenting evidence that the main assumptions required by RDD are met in our data. First, we assess the existence of a discontinuity in the amount transferred as royalties at the UBN cutoff. According to Figure 2.6, this assumption is fulfilled by our design. The figure reports the local average level of royalties for municipalities around the cutoff, their 95% confidence intervals, and a linear (Panel A) or quadratic (Panel B) polynomial fit. Each dot represents the average of the royalties variable within bins of equal size, selected so that there are 6 bins at each side of the cutoff. A jump in the level of royalties at the threshold, consistent with the rule of thumb average reduction in per capita royalties previously estimated, is evident in both figures.

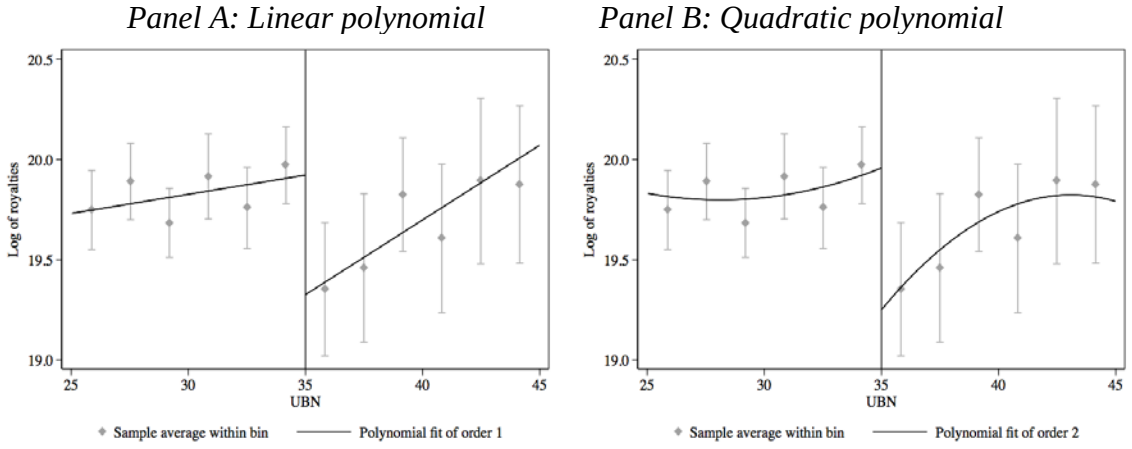


Figure 2.6. Royalties discontinuity in UBN under New Regulation

Secondly, we assess the potential manipulation of the forcing variable, *UBN*, which would violate our identification strategy. If *UBN* was manipulated, then municipalities could influence their amount of royalties received. In our context this is unlikely, since the *UBN* considered by the law to target royalty transfers was built at the municipal level based on the Colombian Population Census which was collected in 2005, long before the Legislative Act that took place in 2011, and amended the constitution to make way for the SGR new law that was approved in 2012. However, following Lee and Lemieux (2010), we also test manipulation by testing for sorting around the threshold. Figure 2.7 reports the histogram illustrating the distribution of municipalities by *UBN*, showing that the distribution of municipalities at $\overline{UBN} = 0{,}35$, does not register an abrupt change that could suggest that some municipalities had been able to somehow underestimate, or overestimate, their *UBN* index in order to become eligible to receive larger transfers. Also, we check the distribution of our forcing variable around the winning threshold using McCrary (2008) test. Figure 2.8 shows the results, and reports the statistic of the null hypothesis of no jump in the distribution. A discontinuous jump in either direction would indicate that is municipalities systematically are on the right hand side of the cutoff. The estimate is 0,023 with a standard error of 0,145, confirming there is no jump in the density at the threshold.

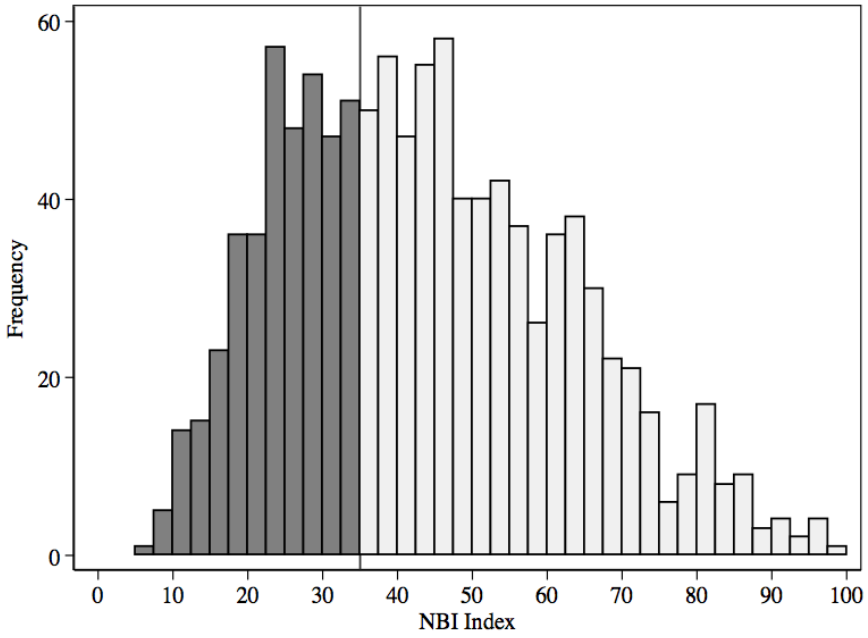


Figure 2.7. Distribution of municipalities according to UBN

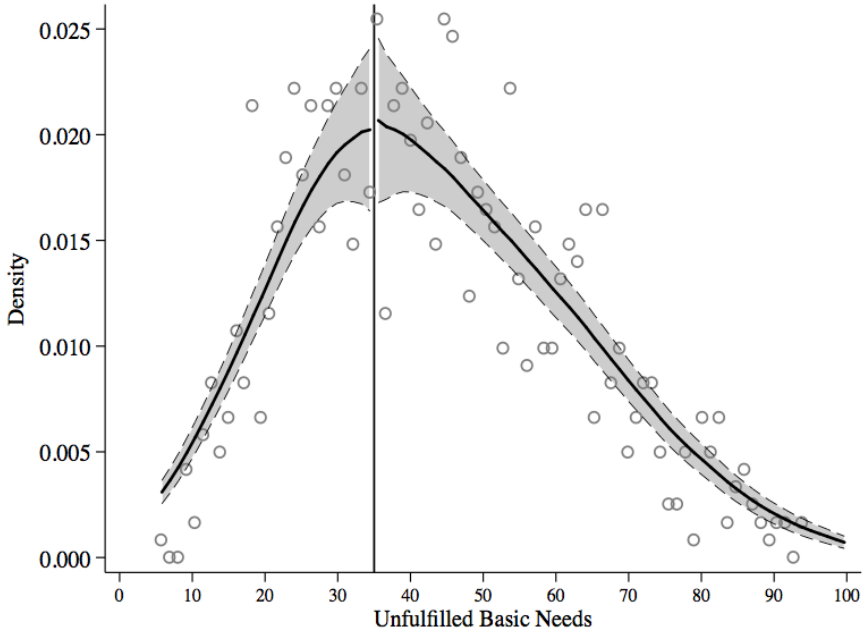


Figure 2.8. McCrary test: Sorting around the 0.35 level of UBN

To make sure that any discontinuous variation in our outcome variables at the *UBN* = 35% threshold could only be attributable to the discontinuous change in the amount of royalties transferred to municipalities that takes place at the threshold, we finally assess the assumption that all predetermined covariates vary smoothly at the cutoff. We check this assumption by estimating equation (1) using as dependent variables a variety of characteristics of the municipalities at the baseline, which include fiscal per capita revenue, property tax, population, municipal category and transfers from the national government. We find no

statistically significant differences at the threshold in the levels of these municipality characteristics at the *UBN* cutoff (Table 2). Municipalities that are more frequently in contact with national entities in charge of allocating royalties are more prone to receive other kind of transfers or royalties. In column 5, we show that differences in transfers from the central government are not statistically significant around the cutoff. Since municipalities only receive royalties from the Regional Compensation Found or direct royalties, in columns 6, 7 and 8 of Table 2.2, we show that differences in direct royalties, log of direct royalties or even the probability of receiving royalties, are not statistically significant around the cutoff at the 5% level. Appendix A presents the continuity of formal employment both in levels and logs, at the baseline (years 2011 and 2012). Although formal employment in levels does not register any significant change, it does in logs for some economic sectors. We will include a control for baseline self-sector formal log employment in our estimations below.

Table 2.2. Direct royalties and continuity over potential employment determinants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable							
	Tax Revenues	*Property Tax*	*Population*	*Municipal Category*	*National Transfers*	*Direct Royalties*	*Log of Direct Royalties*	*Pr(Direct Royalties>0)*
UBN	-75.0	-7.9	5,153	-0.033	-54.6	-1,358,057,600	-5.01	-0.39*
	(169.19)	(13.26)	(14,978)	(0.03)	(126.45)	(910,442,624)	(3.34)	(0.20)
Bandwidth	11.967	11.666	6.793	8.626	11.592	11.644	9.548	8.643
Observations	1,063	1,063	1,064	1,064	1,063	1,064	1,064	1,064
Observations in bandwidth	499	488	279	346	486	489	386	346

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). The estimates are obtained using a triangular kernel and a quadratic polynomial. Results are robust to using an Epanechnikov or uniform kernel, or a linear polynomial. Tax revenues, property tax, national transfers are figures of year 2010 taken from the central municipalities government fiscal statistics, the population is taken from Population Census 2005, and the direct royalties are figures of year 2012. We use data of all municipalities in categories 4, 5 and 6.

Table 2.3. RDD nonparametric estimates of the effects of royalties on formal employment

	(1)		(2)		(3)		(4)		(5)		(6)	
Second Stage: Dep. Var.: Ln(Labor)												
Biased-Corrected Coefficient of Ln(Royalties)	0.826	**	0.818	**	0.895	**	0.784	*	0.877	**	1.004	**
Std. Err.	0.403		0.416		0.397		0.457		0.431		0.416	
P-Value	0.040		0.049		0.024		0.086		0.042		0.016	
First Stage: Dep. Var.: Ln(Royalties)												
Biased-Corrected Coefficient of 1[UBN>35%]	-0.545	***	-0.570	***	-0.618	***	-0.532	***	-0.561	***	-0.590	***
Std. Err.	0.115		0.124		0.129		0.132		0.132		0.132	
P-Value	2.E-06		4.E-06		2.E-06		6.E-05		2.E-05		8.E-06	
Bandwidth Left+Right	13.32		10.90		9.01		18.05		18.60		16.63	
Number of Observations	3,015		3,015		3,015		3,015		3,015		3,015	
Effective Number of Obs. in Bandwidth	794		668		538		1049		1100		980	
Polinomial Order	Linear						Quadratic					
Kernel	Tri		Epa		Uni		Tri		Epa		Uni	

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We control for tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Finally, we control for the log of 2012 formal employment. We use data of all municipalities in categories 4, 5 and 6, for years 2013, 2014 and 2015. Standard errors are robust and clustered at the municipality level.

In Table 2.3 we report our main result: increases in royalties leads to a substantial and statistically significant rise of formal employment. The table presents non-parametric estimates following Calonico et al. (2014a), and using our data for 2013, 2014, and 2015 (similar results are obtained if instead we only consider years 2014 and 2015). We report bias corrected coefficient, robust standard errors clustered at by municipality, and optimal bandwidths for our first and second stages. Columns 1-3 include a linear polynomial while columns 4-6 a quadratic polynomial. To show robustness to the importance of observations around the cutoff, estimates in columns 1 and 4 weight observations using a triangular kernel, columns 2 and 5 an epanechnikov kernel, and columns 3 and 6 allow all observations to have the same weight. We control for 2012 tax revenues, property tax, and national transfers, 2005 population, 2012 direct royalties, and for the log of 2012 formal employment (controlling for the log of 2011 formal employment leads to similar results). Estimates without control (not presented) are always positive and significant at the 5% level.

The non-parametric estimates are statistically significant across all specifications. In line with the shape of the discontinuity illustrated in Figure 5, the first stage estimation suggests that municipalities around the cutoff whose proportion of population with UBN exceeded 35% leads to receive 53-62% less royalties in reference to municipalities where less than 35% of population had UBN. On the other hand, the point estimate for the second stage shows that a one percentage point increase in the level of royalties explain an increase of 0,78-1,0% on formal employment. A one percentage point increase in royalties for municipalities belonging to categories 4, 5 or 6, for the period 2013-2015, would be equivalent to an increase in royalties of about COP$ 8,3 million, while say, a 0.9 percentage point increase in formal employment would be about 7,3 additional formal employees, that is about COP$ 1,14 million per formal employee, or 185% of a monthly minimum wage in year 2014.

The effect seems reasonable once one considers the weight of royalties in the funding of royalty-promoted projects by municipalities, partly funded by the RCF: 25% is funded with the own resources of municipalities, 11% is funded with direct royalties, 4% is funded with resources from the Regional Development Fund (RDF), and finally, 60% is funded with resources of the Regional Compensation Fund (RCF), out of which, we are assessing the effects of its 40% that is mostly targeted based on the UBN (See column 17, row 3, in Appendix B). That is, an average project of municipalities, co-funded with resources of the RCF, would invest about $4,74 million of 2014 per year, nearly $4,61 million of 2013, that is, 7,7 monthly minimum wages of year 2014. This figure is about 41% the average savings by firms per formal employment generated due to the payroll taxes reduction introduced by the tax reform of 2012, which Morales and Medina (2017) found to be of $11,2 million of 2013 per year. Our figure is smaller since we are only considering the smallest municipalities, which offer much smaller formal wages than the average firms considered in their analysis.

Our results do not depend on our choice of a specific bandwidth. Panel A and B of Figure 2.9 shows the estimated coefficient (and the 95% confidence interval) for the first and second stage when a bandwidth between 50% and 150% as large as the optimal bandwidth is used, and we do not include controls. Thus, the first value (50%) includes municipalities at a half distance of the optimal bandwidth estimated in Table 2, which means that it includes all municipalities with UBN between 31,54 (35-6,925/2) and 38,46 (35+6,925/2). For each bandwidth, equations (1) and (2) are estimated based on the first column of Table 3. The

results confirm that there is indeed a strong discontinuity in the log of royalties received by municipalities that have a lower rate of UBN (but close) at 35% and those with higher UBN index (but close) to 35%. The graph shows that the discontinuity is statistically significant no matter the size of the bandwidth. Also, the magnitude of the effect is very stable is this range.

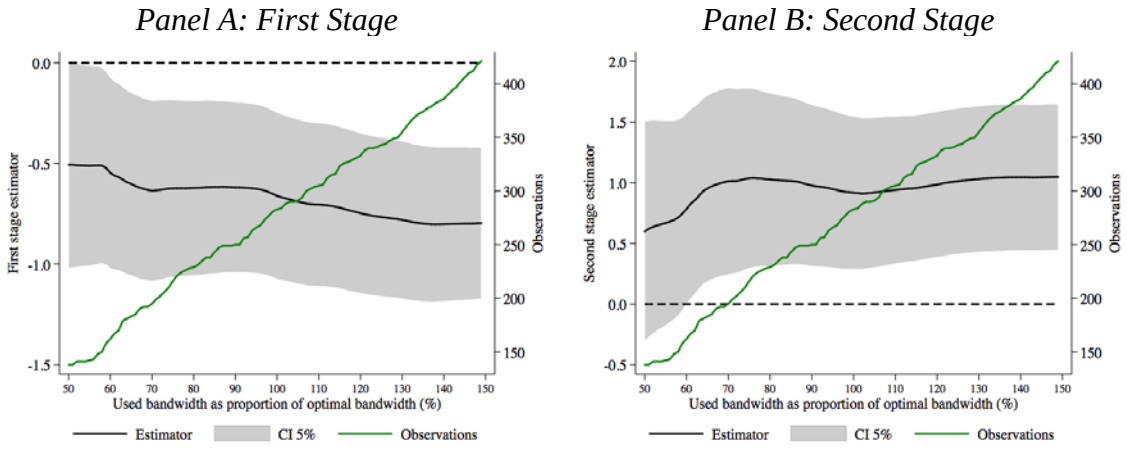


Figure 2.9. Robustness to bandwidth selection

Our results are largely held regardless of the inclusion of controls, or whether our estimates are parametric or not. In Appendix C1 we replicate Table 2.3 without control variables, while Appendix C2 presents parametric estimates with and without control variables. The inclusion of covariates reduce the magnitude of our nonparametric estimates, although in a statistically insignificant magnitude, and their sign keeps being highly statistically significant. Similar results are found with our parametric estimates, although in that case, the inclusion of controls reduces more the magnitude of the estimated effect.

The first active year of the reform was 2012. However, because municipalities were not fully engaging to the new system, less than 5% of them received transferences trough the Regional Compensation Found. However, this proportion increased to 93% in 2013 and it was greater than 95% in the next couple of years. Table 2.4 shows that the effect for years 2013, 2014 and 2015 are of similar magnitude, although they are not always statistically significant. Column 4 shows the effect when we include the three years in a single regression (the same result presented in column (4) of Table 3). Appendix D includes the results when instead of a piecewise continuous quadratic polynomial in z we control for a linear polynomial in z, finding similar, although more statistically significant results.

Table 2.4. Prevalence of the effect

	2013	2014	2015	All
Year	(1)	(2)	(3)	(4)
Second Stage: Dep. Var.: Ln(Labor)				
Biased-Corrected Coefficient of Ln(Royalties)	0.759	0.632	0.996 **	0.784 *
Std. Err.	0.518	0.470	0.487	0.457
P-Value	0.143	0.178	0.041	0.086
First Stage: Dep. Var.: Ln(Royalties)				
Biased-Corrected Coefficient of 1[UBN>35%]	-0.489 ***	-0.571 ***	-0.524 ***	-0.532 ***
Std. Err.	0.123	0.141	0.141	0.132
P-Value	0.000	0.000	0.000	0.000
Bandwidth Left	11.2	8.6	9.5	12.6
Bandwidth Right	11.2	8.6	9.5	12.6
Bandwidth Left+Right	22.4	17.3	19.0	25.1
Number of Observations	996	1009	1010	3015
Effective Number of Obs. in Bandwidth	445	339	375	1049

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We control for a piecewise continuous ***quadratic*** polynomial in *z*, tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Finally, we control for the log of 2012 formal employment. Standard errors are robust and clustered at the municipality level in column (4). We use data of all municipalities in categories 4, 5 and 6.

Table 2.5. Heterogeneous effects by sector. Quadratic Polynomial.

	All	Agriculture	Mining	Manufacturing	Electricity	Manufacturing & Electricity
	(0)	(1)	(2)	(3)	(4)	(3 & 4)
Second Stage: Dep. Var.: Ln(Labor)						
Biased-Corrected Coefficient of Ln(Royalties)	0.784 *	0.638	0.078	1.022	1.297	1.358 *
Std. Err.	0.457	0.956	0.815	0.889	0.924	0.807
P-Value	0.086	0.505	0.924	0.251	0.160	0.093
First Stage: Dep. Var.: Ln(Royalties)						
Biased-Corrected Coefficient of 1[UBN>35%]	-0.532 ***	-0.529 ***	-0.531 ***	-0.535 ***	-0.519 ***	-0.532 ***
Std. Err.	0.132	0.134	0.134	0.132	0.135	0.133
P-Value	6.E-05	7.E-05	7.E-05	5.E-05	1.E-04	0.000
Bandwidth Left+Right	18.05	16.58	16.84	18.41	15.81	18.06
Number of Observations	3,015	3,015	3,015	3,015	3,015	3,015
Effective Number of Obs. in Bandwidth	1049	977	998	1088	932	1052

	Construction	Commerce	Transport	Financial Intermediation	Real Estate	Social & Personal Services
	(5)	(6)	(7)	(8)	(9)	(10)
Second Stage: Dep. Var.: Ln(Labor)						
Biased-Corrected Coefficient of Ln(Royalties)	-0.445	0.512	0.674	1.561 **	1.133 **	0.465
Std. Err.	0.874	0.772	0.738	0.619	0.535	0.516
P-Value	0.611	0.507	0.361	0.012	0.034	0.367
First Stage: Dep. Var.: Ln(Royalties)						
Biased-Corrected Coefficient of 1[UBN>35%]	-0.533 ***	-0.532 ***	-0.532 ***	-0.530 ***	-0.534 ***	-0.532 ***
Std. Err.	0.132	0.133	0.133	0.133	0.132	0.133
P-Value	6.E-05	6.E-05	6.E-05	7.E-05	5.E-05	6.E-05
Bandwidth Left+Right	17.73	17.98	17.64	16.51	17.38	18.06
Number of Observations	3,015	3,015	3,015	3,015	3,015	3,015
Effective Number of Obs. in Bandwidth	1028	1040	1025	974	1010	1052

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We control for tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Finally, we control for the log of 2012 formal employment. Standard errors are robust and clustered at the municipality level. We use data of all municipalities in categories 4, 5 and 6.
Agriculture: agriculture, hunting, forestry and fishing; Mining: mining and quarrying; Electricity: electricity, gas, water and sewerage; Commerce: wholesale and retail trade and hotels; Transport: transport, storage and communications; financial intermediation; Social and Personal Services: services to companies, social and personal services.

Table 2.5 breaks down the dependent variable by economic sector according to the registered Classification of Economic Activities (CIIU). We group employment in 10 categories as follow (sector size according to formal employment in 2012 in municipal categories 4, 5 and 6 in parenthesis): agriculture, hunting, forestry and fishing (8,5%); mining and quarrying (3,7%); manufacturing (5,1%); electricity, gas, water and sewerage (0,8%); construction (4,4%); wholesale and retail trade and hotels (11.2%); transport, storage and communications (3,0%); financial intermediation (3,0%); real estate (26,3%); and services to companies, social and personal services (34,0%). Table 5 shows that the sectors that would have benefited the most of the royalties were financial intermediation (column 8), real state (column 9) and the aggregation of the manufacturing and agricultural sectors (column "3 & 4"). The results are robust to whether we use a quadratic or a linear polynomial in *z*, as it is shown in Appendix E, which in addition, also suggests there might as well be positive effects in agriculture.

Effect of Royalties on the Source of Funding of Municipal Projects

We now use data of an official report provided by the Department of National Planning (SUIFP), containing the census of the projects registered in its Bank of Investment Projects (BPIN by its acronym in Spanish) in years 2013, 2014 and 2015. The report contains all investment projects by year and entity (states, municipalities, private entities, public firms, etc.). For our empirical estimations we only consider the information for central governments of municipalities in categories 4, 5 and 6, as described in Appendix B, columns 13 to 18. These municipalities account for a total of 6,657 projects, 4,622 of which were at least partly funded with resources of the Regional Compensation Fund. For each municipality, we estimate the average share of the projects funded with resources of each of the sources included in Appendix B: (i) own resources, (ii) the 0,5% of the resources of the SGR, (iii) the Science, Technology and Innovation Fund of the SGR, (iv) the Regional Compensation Fund of the SGR, (v) the Regional Development Fund of the SGR, and finally, (vi) resources coming from direct royalties of the SGR. We then estimate the effect of the royalties transferred to municipalities from the RCF, on the share of the investment projects funded with resources of the RCF, and on the share of the projects funded with the own resources of the municipalities. We also get estimates of the effect of the royalties transferred to municipalities from the RCF, on the logarithm of the amount funded with resources of the RCF, and with the own resources of the municipalities.

Table 2.6. RDD nonparametric estimates of the effects of royalties on the source of funding of projects promoted by municipalities

	Dep Var: Source of Funding			
	A. logarithms*		B. Shares**	
	Own Resources	RCF	Own Resources	RCF
	(1)	(2)	(3)	(4)
Second Stage: Dep. Var.: Ln(Labor)				
Biased-Corrected Coefficient of Ln(Royalties)	-5.939	7.253 *	-0.202	0.185
Std. Err.	5.341	4.023	0.124	0.309
P-Value	0.266	0.071	0.104	0.549
First Stage: Dep. Var.: Ln(Royalties)				
Biased-Corrected Coefficient of 1[UBN>35%]	-0.532 ***	-0.533 ***	-0.321 **	-0.379 ***
Std. Err.	0.133	0.133	0.127	0.129
P-Value	0.000	0.000	0.011	0.003
Bandwidth Left+Right	17.3	17.1	12.8	15.6
Number of Observations	3,015	3,015	2,460	2,460
Effective Number of Obs. in Bandwidth	1,010	1,001	631	784

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We control for tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Standard errors are robust and clustered at the municipality level. We use data of all municipalities in categories 4, 5 and 6, for years 2013, 2014 and 2015.

* Logarithm of the amount of the investment project funded with resources of the Regional Compensation Fund of the SGR (RCF), or with own resources of the municipality. ** Share of the investment projects funded with resources of the Regional Compensation Fund of the SGR (RCF), or with own resources of the municipality.

Source: Report provided by the Department of National Planning (SUIFP) with the census of the projects registered in its Bank of Investment Projects (BPIN by its acronym in Spanish) for years 2013, 2014 and 2015.

The results are presented in Table 2.6. Panel A includes the effects on the logarithms of the investment funded with the resources of the RCF, and with their own resources, and panel B presents the effects on the shares of their respective participations in the total funding of the projects. Among the municipalities' own resources we consider their own taxes (property, industry and commerce, etc.), and other earnings that depend on the municipalities' own fiscal effort. In both cases we find a significant discontinuity in the transfers of royalties from the RCF at the UBN cutoff of 35%, with the poorer municipalities actually receiving smaller transfers, as previously found and explained. We find no significant effect of the royalties transferred of the RCF to the municipalities on their amounts of own resources invested (column 1), although we do find an increase in the amount of funds received of the RCF and invested in their projects (column 2). There is no effects on the shares these items contribute to their investment projects but a marginally significant (at the 10,4%, column 3) reduction in the share of their own resources that might be explained by the larger use of other sources, like those of the RCF.

Effect of Royalties on the Fiscal Effort of Municipalities

Table 2.7. Effect of Royalties on the Fiscal Effort of Municipalities

	Municipalities in Categories 4, 5 & 6			
	Nat. Transfers	All Taxes	Property	Ind & Comm
	(1)	(2)	(3)	(4)
Second Stage: Dep. Var.: Ln(Labor)				
Biased-Corrected Coefficient of Ln(Royalties)	-6,000,000	25,000,000	-710,000	-1,600,000
Std. Err.	5,400,000	64,000,000	660,000	2,500,000
P-Value	0.267	0.692	0.282	0.527
First Stage: Dep. Var.: Ln(Royalties)				
Biased-Corrected Coefficient of 1[UBN>35%]	-0.397 **	-0.160	-0.455 ***	-0.660 ***
Std. Err.	0.165	0.381	0.152	0.210
P-Value	0.016	0.675	0.003	0.002
Bandwidth Left+Right	10.9	20.7	12.1	15.0
Number of Observations	3,803	3,815	3,965	4,313
Effective Number of Obs. in Bandwidth	806	1,618	1,169	1,305

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We control for tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Standard errors are robust and clustered at the municipality level. We use data of all municipalities in categories 4, 5 and 6, for years 2013, 2014 and 2015.

We now assess the effects of the royalty transfers on the fiscal effort municipalities exert on their taxes earnings, and in particular, on their main tax earnings, namely property tax, and the tax to industries and commerce. Municipalities receiving additional resources based on royalty transfers from the RCF might either complement those new economic resources with their own based mostly on property taxes and on taxes to industrial and commercial firms, which would require them to increase their fiscal effort, or substitute their own resources with the new ones from the RCF, lowering their fiscal effort.

The first column of the table shows that royalties transferred from the RCF do not affect the amount of national transfers. This was the expected result since national transfers do not depend on the discretion of local policymakers, but on municipalities' characteristics such as their population, poverty levels, etc. Columns 2 to 4 of the table show that royalties from the RCF do not affect the aggregate of the municipalities' tax earnings, neither their property, nor their industrial and commercial taxes.

2.6 Conclusions

In this chapter, we investigate the effects on formal employment, of a subset of the royalties that are transferred to Colombian municipalities, coming from a fund that seeks to progressively redistribute among all municipalities (the Regional Compensation Fund), a share of the royalties generated from the exploitation of nonrenewable resources. This fund represents about 6% of all royalties transferred to subnational governments, which in turn, represent more than 1% of the GDP, more than their earnings due to their taxes on industry

and commercial firms, which have been registered to fluctuate between 0,9% and 1,0% of the GDP.

We focus on the assessment of the effects of the Regional Compensation Fund since it is distributed among municipalities according to a rule that allows us to credibly identify its effect on employment by means of a regression discontinuity methodology. The law orders to distribute royalties of the Regional Compensation Fund based on a mixture of the municipalities' population, and on whether the percentage of the municipalities' population with unsatisfied basic needs was above or below 35%.

We find positive and significant effects of these royalties on formal employment, mainly in the sectors of financial intermediation, real estate, the union of industry and commerce, and to a lesser extent, on agriculture. These results are robust to several specifications regarding the nonparametric estimation and the control variables included.

Furthermore, we assess whether royalties coming from the RCF affected the sources of funding of the subnational investment projects on which municipalities participate, and find that it increases the amount of resources from the RCF supporting those project without reducing the amount of other sources of funding, in particular, of the municipalities' own resources.

Finally, we assess whether royalties coming from the RCF affected the municipalities' fiscal effort, by reducing their total taxes earnings, or their property, or their industry and commercial tax earnings, and found no effect on none of them.

Even though our assessment only covered up to somewhat more than three years beyond the implementation of the law, this preliminary evidence seems very positive, since the generation of formal employment was significant, while it did not affect the targeting of the municipalities' own resources, neither their levels of fiscal effort.

2.7 Appendix

Appendix A. Continuity over baseline employment variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Dependent variable										
	Formal Employment 2011										
	All	Agriculture	Mining	Manufacturing	Electricity	Construction	Commerce	Transport	Financial Intermediation	Real State	Social and Personal Services
UBN	1,559.0	81.1	104.629*	29.2	-3.7	62.4	192.2	64.7	22.7	847.6	339.7
	(1,783.4)	(75.9)	(57.0)	(51.9)	(10.9)	(57.3)	(221.9)	(102.8)	(16.3)	(844.5)	(503.8)
Bandwidth	5.445	5.390	6.234	5.290	6.025	4.934	5.379	6.830	4.267	4.880	5.709
Observations	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064
Observations in bandwidth	228	226	252	224	251	200	226	281	169	199	237

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Dependent variable										
	Log of Formal Employment 2011										
	All	Agriculture	Mining	Manufacturing	Electricity	Construction	Commerce	Transport	Financial Intermediation	Real State	Social and Personal Services
UBN	-0.872*	-0.907	-0.677	-0.886*	-1.241**	-0.153	-1.085**	-1.127**	-0.672	-1.016**	-0.726*
	(0.447)	(0.566)	(0.481)	(0.531)	(0.621)	(0.604)	(0.552)	(0.559)	(0.49)	(0.476)	(0.438)
Bandwidth	10.864	9.062	11.544	10.801	7.033	10.243	13.565	9.904	8.864	9.714	9.745
Observations	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064
Observations in bandwidth	446	364	483	444	289	422	553	401	352	392	393

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Dependent variable										
	Formal Employment 2012										
	All	Agriculture	Mining	Manufacturing	Electricity	Construction	Commerce	Transport	Financial Intermediation	Real State	Social and Personal Services
UBN	1,665.2	122.1	117.96**	19.2	-3.9	78.4	222.2	68.1	20.3	910.6	296.3
	(1,868.4)	(95.4)	(56.4)	(56.2)	(12.3)	(71.6)	(241.4)	(113.6)	(15.5)	(902.2)	(481.1)
Bandwidth	5.417	5.507	6.066	5.498	6.070	4.709	5.490	7.173	4.485	4.883	5.605
Observations	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064
Observations in bandwidth	227	229	251	229	251	193	229	292	183	199	232

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Dependent variable										
	Log of Formal Employment 2012										
	All	Agriculture	Mining	Manufacturing	Electricity	Construction	Commerce	Transport	Financial Intermediation	Real State	Social and Personal Services
UBN	-0.932**	-0.654	-0.641	-1.039**	-0.894	-0.52	-0.925	-1.058**	-0.845*	-1.061**	-0.697*
	(0.426)	(0.61)	(0.43)	(0.53)	(0.63)	(0.571)	(0.59)	(0.538)	(0.47)	(0.443)	(0.414)
Bandwidth	12.047	8.292	11.171	11.196	7.680	12.641	11.618	10.444	8.866	9.435	10.786
Observations	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064	1,064
Observations in bandwidth	503	335	462	464	308	529	488	430	352	381	444

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). The estimates are obtained using a triangular kernel and a quadratic polynomial. Results are robust to using an Epanechnikov or uniform kernel, or a linear polynomial. We use data of all municipalities in categories 4, 5 and 6.

Appendix B. Participation of the sources of funding in projects funded with the SGR

	Source of Funding	All Entities**						Only Municipalities***						Only Municipalities in Categories 4, 5 & 6****					
		COP$ Billion (10^{12})			Participation			COP$ Billion (10^{12})			Participation			COP$ Billion (10^{12})			Participation		
		All	With RCF	Without RCF	All	With RCF	Without RCF	All	With RCF	Without RCF	All	With RCF	Without RCF	All	With RCF	Without RCF	All	With RCF	Without RCF
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
1	Own Resources*	5.62	2.55	3.06	22%	22%	19%	2.95	1.64	1.30	28%	26%	26%	1.86	1.13	0.74	23%	25%	22%
2	0.5% of SGR**	0.20	0.09	0.11	0.8%	0.8%	0.7%	0.01	0.00	0.01	0%	0.00%	0%	0.01	0.00	0.01	0%	0.00%	0%
3	STIF, SGR	3.75	0.03	3.72	15%	0.23%	23%	0.27	0.01	0.26	2%	0.11%	5%	0.16	0.00	0.16	2%	0.02%	5%
4	RCF, SGR	8.05	8.05	0.00	32%	70%	0.00%	3.73	3.73	0.00	35%	59%	0.00%	2.76	2.76	0.00	35%	60%	0.00%
5	RDF, SGR	0.00	0.00	2.35	0%	0%	15%	0.00	0.31	0.35	0%	5%	7%	0.42	0.18	0.24	5%	4%	7%
6	DR, SGR	7.65	0.73	6.94	30%	6%	43%	3.71	0.68	3.03	35%	11%	61%	2.75	0.52	2.23	35%	11%	66%
Total		25.26	11.45	16.17	100%	100%	100%	10.66	6.36	4.94	100%	100%	100%	7.96	4.59	3.37	100%	100%	100%
		322%	146%	206%				173%	103%	80%				129%	75%	55%			
Number of Projects		10,970	6,240	4,730				8,155	5,529	2,626				6,657	4,622	2,035			
		150%	85%	65%				117%	79%	38%				96%	66%	29%			

Projects of 2012, 2013, 2014 and 2015. * Fiscal earnings of departments and municipalities. ** Departamentos, municipios, privadas, empresas públicas, otras (Cormagdalena - SGR, Corporaciones-SGR, Entidades Presupuesto Nacional - PGN). *** Only includes the Central Government. **** Only includes the Central Government for municipalities in categories 4, 5 & 6. STIF, SGR: Science, Technology and Innovation Fund; RCF, SGR: Regional Compensation Fund; RDF, SGR: Regional Development Fund; DR, SGR: Direct Royalties.

Appendix C1. RDD nonparametric estimates of the effects of royalties on formal employment. No control variables

	(1)		(2)		(3)		(4)		(5)		(6)	
Second Stage: Dep. Var.: Ln(Labor)												
Biased-Corrected Coefficient of Ln(Royalties)	1.163	***	1.158	***	1.259	***	1.169	***	1.052	***	1.168	***
Std. Err.	0.297		0.295		0.276		0.306		0.367		0.279	
P-Value	0.000		0.000		0.000		0.000		0.004		0.000	
First Stage: Dep. Var.: Ln(Royalties)												
Biased-Corrected Coefficient of 1[UBN>35%]	-0.851	***	-0.866	***	-0.877	***	-0.855	***	-0.829	***	-0.977	***
Std. Err.	0.185		0.188		0.180		0.204		0.229		0.209	
P-Value	4.E-06		4.E-06		1.E-06		3.E-05		3.E-04		3.E-06	
Bandwidth Left	6.41		6.12		6.07		9.79		7.70		9.15	
Bandwidth Right	6.41		6.12		6.07		9.79		7.70		9.15	
Bandwidth Left+Right	12.82		12.24		12.13		19.59		15.41		18.30	
Number of Observations	3,015		3,015		3,015		3,015		3,015		3,015	
Effective Number of Obs. in Bandwidth	1036		1004		1004		1584		1232		1484	
Polinomial Order	Linear						Quadratic					
Kernel	Tri		Epa		Uni		Tri		Epa		Uni	

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We use data for years 2013, 2014 and 2015. Standard errors are robust and clustered at the municipality level.

Appendix C2. RDD Parametric estimates of the effects of royalties on formal employment

	No Controls (1)		Controls* (2)	
Coefficient of Ln(Royalties)	1.232	***	0.326	*
	0.445		0.179	
	0.006		0.070	
Coefficient of 1[UBN>35%]	-0.788	***	-0.405	***
	0.182		0.096	
	0.000		0.000	
Number of Observations	1,194		1,194	

We use data for years 2013, 2014 and 2015, in a bandwidth of *z* between -9 and 9. Standard errors are robust and clustered at the municipality level. * We control for tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Finally, we control for the log of 2012 formal employment.

Appendix D. Prevalence of the effect controlling for a linear polynomial in *z*.

	2013	2014	2015	All
Year	(1)	(2)	(3)	(4)
Second Stage: Dep. Var.: Ln(Labor)				
Biased-Corrected Coefficient of Ln(Royalties)	0.807	0.698 *	0.965 **	0.826 **
Std. Err.	0.490	0.369	0.442	0.403
P-Value	0.1	0.059	0.029	0.04
First Stage: Dep. Var.: Ln(Royalties)				
Biased-Corrected Coefficient of 1[UBN>35%]	-0.503 ***	-0.573 ***	-0.524 ***	-0.545 ***
Std. Err.	0.118	0.113	0.125	0.115
P-Value	0.000	0.000	0.000	0.000
Bandwidth Left	6.4	8.4	6.3	10.4
Bandwidth Right	6.4	8.4	6.3	10.4
Bandwidth Left+Right	12.8	16.8	12.5	20.8
Number of Observations	996	1009	1010	3015
Effective Number of Obs. in Bandwidth	248	334	248	794

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We control for a piecewise continuous ***linear*** polynomial in *z*, tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Finally, we control for the log of 2012 formal employment. Standard errors are robust and clustered at the municipality level in column (4).

Appendix E. Heterogeneous Effects by Sector. Linear Polynomial.

	All	Agriculture	Mining	Manufacturing	Electricity	Construction
	(0)	(1)	(2)	(3)	(4)	(5)
Second Stage: Dep. Var.: Ln(Labor)						
Biased-Corrected Coefficient of Ln(Royalties)	0.826 **	1.400 *	0.366	1.146	0.584	-0.092
Std. Err.	0.403	0.792	0.732	0.787	0.739	0.758
P-Value	0.040	0.077	0.617	0.145	0.429	0.903
First Stage: Dep. Var.: Ln(Royalties)						
Biased-Corrected Coefficient of 1[UBN>35%]	-0.545 ***	-0.547 ***	-0.552 ***	-0.548 ***	-0.549 ***	-0.521 ***
Std. Err.	0.115	0.115	0.118	0.116	0.118	0.111
P-Value	2.E-06	2.E-06	3.E-06	2.E-06	4.E-06	3.E-06
Bandwidth Left+Right	13.32					
Number of Observations	3,015	3,015	3,015	3,015	3,015	3,015
Effective Number of Obs. in Bandwidth	794	806	758	773	701	893

	Commerce	Transport	Financial Intermediation	Real State	Social & Personal Services
	(6)	(7)	(8)	(9)	(10)
Second Stage: Dep. Var.: Ln(Labor)					
Biased-Corrected Coefficient of Ln(Royalties)	0.718	0.727	1.226 **	0.946 **	0.574
Std. Err.	0.671	0.658	0.523	0.473	0.456
P-Value	0.285	0.269	0.019	0.046	0.208
First Stage: Dep. Var.: Ln(Royalties)					
Biased-Corrected Coefficient of 1[UBN>35%]	-0.481 ***	-0.552 ***	-0.543 ***	-0.531 ***	-0.539 ***
Std. Err.	0.102	0.118	0.115	0.112	0.115
P-Value	2.E-06	3.E-06	3.E-06	2.E-06	3.E-06
Bandwidth Left+Right					
Number of Observations	3,015	3,015	3,015	3,015	3,015
Effective Number of Obs. in Bandwidth	1028	731	728	833	839

Optimal bandwidth, bias corrected, robust nonparametric estimates based on Calonico, Cattaneo and Titiunik (2014a). We control for tax revenues, property tax, and national transfers (figures of year 2010, taken from the central municipalities government fiscal statistics), and the population (from Population Census 2005). Finally, we control for the log of 2012 formal employment. Standard errors are robust and clustered at the municipality level.

Agriculture: agriculture, hunting, forestry and fishing; Mining: mining and quarrying; Electricity: electricity, gas, water and sewerage; Commerce: wholesale and retail trade and hotels; Transport: transport, storage and communications; financial intermediation; Social and Personal Services: services to companies, social and personal services.

Chapter 3 - Labor Market Opportunities for the Demobilized in Colombia[25]

Abstract

This paper calculates the probability of employment for demobilized on formal and informal activities, considering the high rick of recidivism that represents for this people their participation in the labor market through an informal way. This study proposes strategies that maximizes the success of the peace process signed between the Colombian government and the Colombian Revolutionary Armed Forces FARC. The study estimates, first, a profile of the demobilized, using information from the Agency for Reintegration and Normalization, and subsequently, uses Mincer equations to perform the respective analysis of the labor market.

JEL CODES: C21; C30; C90; J23

KEYWORDS: Demobilized, informality, Mincer equations.

[25] This paper has been published as: *"Las ecuaciones de Mincer y el mercado laboral para los desmovilizados en Colombia"*. Revista ESPACIOS. ISSN 0798 1015. Vol. 41 (Nº 21) Año 2020.

3.1 Introduction

In Colombia, the internal armed conflict has presented serious negative impacts on economic matters. In the 1980s, productivity losses associated with the increase in crime within the country amount to 1% of GDP (Cárdenas, 2002). In aggregate, the costs associated with crime are estimated at 2.9% of GDP (Badel et al., 1998). The intensity of the conflict generates a decrease in the economic growth rate of more than 2 percentage points (Vargas, 2003, as cited by Sánchez and Díaz 2005).

These economic costs are reflected in different sectors and phenomena. One of them, the accumulation of capital, directly and indirectly affects public and private institutions, infrastructure and other physical capital, and increases transaction costs (Echeverri et al., 2001). Another phenomenon has to do with the labor market. Certainly, the armed conflict generates more informality, destroys the capacities and business infrastructure, which are necessary for the generation of formal jobs.

Some of these labor market problems do not end once the groups outside the law are reincorporated to civilian life. The incentives generated in the labor market, are designed so that these people engage in labor in informal activities, or recidivate in activities outside the law or, ultimately, access formal jobs.

The total of demobilized combatants in Colombia, according to the Agency for Reincorporation and Normalization, amounts to about 60 thousand people, of which about 20% are in the department of Antioquia. Of the total of reinserted people, just over 60% are in an age range from 26 to 40 years. They are people with very low levels of education, and more than half of them have children and with spouse or partner.

The Colombian Agency for reintegration has revealed that 24% of those reintegrated combatants remain in illegal activities, while 51% manage to engage in labor, but in informal jobs, with only 21% of people doing so in formal activities. Consequently, for those reinserted in Colombia, the probability of recidivism is much higher than the probability of finding formal employment. In contrast, if you are working, you are highly likely to do so in informal activities.

Thus, this chapter seeks to calculate the probabilities that a demobilized person will be incorporated into formal and informal economic activities, considering the high risk of incidence that informality represents, seeking to propose strategies that maximize the success of the peace process signed with the FARC. To achieve this objective, the study estimates, first, a profile of the demobilized, using information from the Agency for Reintegration and Normalization, and subsequently, Mincer equations are estimated to perform the respective analysis of the labor market using data from the Large Integrated Household Survey for 2017 (January - October).

The article is integrated as follows. After this introduction, a conceptual framework is presented that includes an analysis of the literature on the subject, as well as the characterization of the demobilized population. The next chapter develops the methodology,

the data is described and the model is outlined. Consequently, the following section offers the main results and, finally, the conclusions and policy recommendations are presented.

Research Questions

According to the determinants of labor productivity, identified by the economic literature, what are the factors that explain the labor demand of the demobilized in Colombia? Is it possible, through different actions of the State and public policies, to improve the possibilities of engaging in the formal labor market for the demobilized in Colombia? Currently, does the Colombian labor market differentiate or discriminate the labor supply in such a way that for a demobilized the probability of finding a formal job is lower than for another person who does not possess this quality? And finally, what level of education should we offer to the demobilized and reinserted in order to minimize the likelihood that they will reoccur in the armed conflict that the country has suffered for so many years?

3.2 Conceptual Framework

The study of the labor market of the demobilized in Colombia requires the conceptual analysis, both of elements of the discriminated labor markets, as well as of the main political, social and economic aspects present in the recent history of the country. Consequently, this chapter consists of two elements, namely:

A. Some Theoretical References

Discrimination in labor markets has been an important object of study, within the economic sciences, in recent years. The analyzes coincide in pointing out that in order to talk about discrimination, the populations under study and comparison are required to be identical in all dimensions that can affect productivity, in the light of employers' considerations (Bertrand and Mullainathan, 2004) and (Altonji and Blank, 1999).

In the existing economic literature, some of the main difficulties that arise when considering discrimination analysis are found in the fact that, in many occasions, the institutional surveys through which the labor market data are compiled, do not fully collect the elements that determine competitiveness, as well as all the characteristics that employers observe when hiring, promoting or establishing salaries (Bertrand and Mullainathan, 2004).

Labor productivity is a key indicator to explain the performance of the workforce (Delery and Shaw, 2001) and (Datta et al., 2005) and tends to be defined as the total product in relation to the workforce (Samuelson and Nordhaus, 1989). The determinants of labor productivity have been an important object of study in economic matters. Consequently, literature has been able to identify them fully, some of them being represented by: human capital, especially aspects related to education; (Schultz, 1961), (Barron et al., 1987) and (Black and Lynch, 1996). The fundamental premise has to do with the fact that higher levels of training or education explain higher labor productivity.

On the other hand, other factors that also determine or explain labor productivity have to do with the economic sector in which the firm operates (Maliranta and Rouvineen, 2003), (Sapprasert, 2010) and (Alderete and Gutiérrez, 2012). Sex is another variable that has to be used as an explanatory factor of labor productivity. The underlying idea is that maternity is a problem for companies and it is often said that maternity protection is excessive. Since the direct costs of maternity leave are borne by the State through the health system, the issue is raised especially in relation to possible productivity problems generated by replacements with people who need to adapt to the job position. But, at the same time, the idea that this eventual decrease in productivity could be compensated by the greater efficiency of women appears in the debate (Todaro et al., 2002).

Other variables that also explain labor productivity have to do with age, health conditions (related to human capital) (Todaro et al., 2002) and marital status (Zarate et al., 2009). As a whole, these variables contribute to explaining the productivity of the workforce, as well as others, related to: the location conditions of the company, the economic sector in which the productive activity is developed, aspects related to the competitiveness of the country, among others.

In the same sense, it is possible to affirm that the results of the labor market are explained by a set of cognitive skills (experience, level of education, age; among others) and non-cognitive skills as well. Participation in illegal activities negatively affects labor market indicators, (Heckman et al., 2006). Indeed, in the presence of adverse conditions, for example, economies with high levels of informality or with high presence of illicit activities; labor market conditions tend to deteriorate.

In the particular case of Colombia, the labor market presents strong mismatches, basically due to: i) The conditions of the internal conflict (which has lasted for more than fifty years of fighting between the Colombian government and the revolutionary guerrillas, such as Revolutionary Armed Forces of Colombia –FARC-). ii) High levels of informality (close to 50% of total employment generated). And iii) the technological gap (present in important productive lines such as agriculture).

These factors combine and determine a situation of greater vulnerability for young people, women and people living in rural areas. Also, the age in interaction with gender makes men (youth, adolescents and even children) -more than their female counterparts- occupy the bulk of the agricultural workforce (including illegal crops) and the mass of combatants in all armies (Escobar y Meerten, 1997).

Subsequently, in the reintegration phase, in a particular way, it is very important to involve companies from different sectors so that they can offer employment to ex-combatants who have already received job training. For this, it is convenient for the government to provide tax incentives to these companies (Fisas, 2011).

That is, the most effective reintegration strategy involves the implementation of complex schemes, in which support for productive activities should be considered, either through self-employment or support for entering the formal labor market and, On the other hand, training activities, both for formal education and for work.

The rapid labor insertion, however precarious, provides ex-combatants (especially women) with new vital horizons that did not exist in the field: direct contact with the monetary economy and access to a new sociability allow them to break isolation, socialize, expand their relations with the outside and redefine their position in the family structure (Escobar y Meerten, 1997).

However, the country's labor market must be taken into consideration, when it comes to offering education to the demobilized, to not repeat what has occurred in Africa, where "the courses offered by development agencies are at a much higher level from what local labor markets require" (Valencia, 2007).

On the other hand, the phenomenon of recidivism in the demobilized population has been widely studied from different disciplines such as psychology and sociology, where problems such as threats to security, poverty and limited political participation have been identified as incentives for a former combatant to return to illegal activities (Bøas y Hatløy, 2008).

B. Characterization of the demobilized in Colombia

The country has consolidated vast experience in disarmament, demobilization and reintegration processes since the bipartisan conflicts of the 50s, through the amnesty and pardon processes with the M-19 and some fronts of the EPL, PRT, MAQL and the CRS[26] in the 80´s and 90´s, the demobilization of the *Autodefensas Unidas de Colombia* (AUC) in 2003, until the current peace process (Gutiérrez, 2016). The main objective in these reintegration processes is to voluntarily detach people from armed groups outside the law, granting them monetary and judicial benefits that guarantee them conditions of adaptability to society.

To this end, two fundamental institutions have been created in the country to advise on actions related to the structuring and development of peace policy, the formalization of dialogues and negotiated searches for peace agreements. On the one hand, in 1983 through the Decrees 240 and 2560 the government of Belisario Betancur created the position of the High Commissioner for Peace, transformed with Decree 1959 of 1994 into the Office of the High Commissioner for Peace (OACP); and, on the other, the Colombian Agency for the Reintegration of People and Groups Raised in Arms (ACR), created in 2006 as a High Presidential Council for Reintegration after the Reintegration of Civil Life Reintegration Program (PRCV), in operation between 2003-2006, were to be congested by the increase in individual and collective demobilizations, particularly with the massive process of AUC, transformed in 2011 and today known as the Agency for Reintegration and Standardization (ARN), in order to reform and prepare the demobilized individually, through psychosocial care, training and academic training, access to the national health system and monthly financial contributions (ACR, 2014).

[26] Movimiento 19 de abril (M-19), Ejército Popular de Liberación (EPL), milicias del Partido Revolucionario de los Trabajadores (PRT), Movimiento Armado Quintín Lame (MAQL) and Corriente de Renovación Socialista (CRS) that comes from Movimiento Izquierda Revolucionaria Patria Libre, and later merged with the Ejército de Liberación Nacional (ELN),

The latter is a specialized entity that has 29 care centers distributed throughout the country, and is responsible for coordinating, strengthening, managing and administering the reintegration program with a long-term vision. For this, with resolution 1356 of 2016, the Reintegration Route is designed, defined as the set of conditions, benefits, strategies, methodologies and actions defined by the ARN and agreed with the person in the process of reintegration, to promote the development of capacities, overcoming the situation of vulnerability and the autonomous exercise of citizenship in a scheme of activities in accordance with the real options of the individual in relation to the desired life project, without losing sight of the way in which social, economic, and legal benefits of the reintegration process are regulated (ARN, 2017). This route includes eight dimensions that can be seen in figure 3.1.

Figure 3.1. Dimensions of the Reintegration Route. Source: Agencia para la Reincorporación y la Normalización (ARN), 2017.

Table 3.1. Classification of participants in demobilization processes

In Process		**Out of the Process**			**Absent from the process**			**Completed**
Active	Inactive	Deceased	Loss of Benefits	Voluntary retirement	In investigation by abandonment of the reintegration process (6 months)	In investigation for causal supervening	Suspended	Completed
Total Population that Entered the Process								

Source: Own construction with information on ARN, 2017

However, there is a latent risk of abandonment of the process in any part of this route and it is due to the greater perceptible monetary benefits in large cities, in addition to other factors such as the low probability of arrest and the lower probability of recognition. It is very likely that those who do so end up inserting in criminal activities within the cities. This happens especially when the demobilized face difficulties in finding legal jobs, perhaps because of a

high unemployment rate or the lack of skills, knowledge or abilities as a result of abandonment of educational processes. In this sense, Table 3.1 shows how the demobilized classify according to the stage or part of the process they represent today.

It is for this reason that studying and understanding the needs of this population group together with the likelihood of recidivism takes on so much importance, mainly in a city like Medellin that houses a fifth of the country's demobilized (20.5%) and that contained 33.4% and 21.2% of the insurgents participating in the 2005 and 2006 processes, respectively (Figure 3.2).

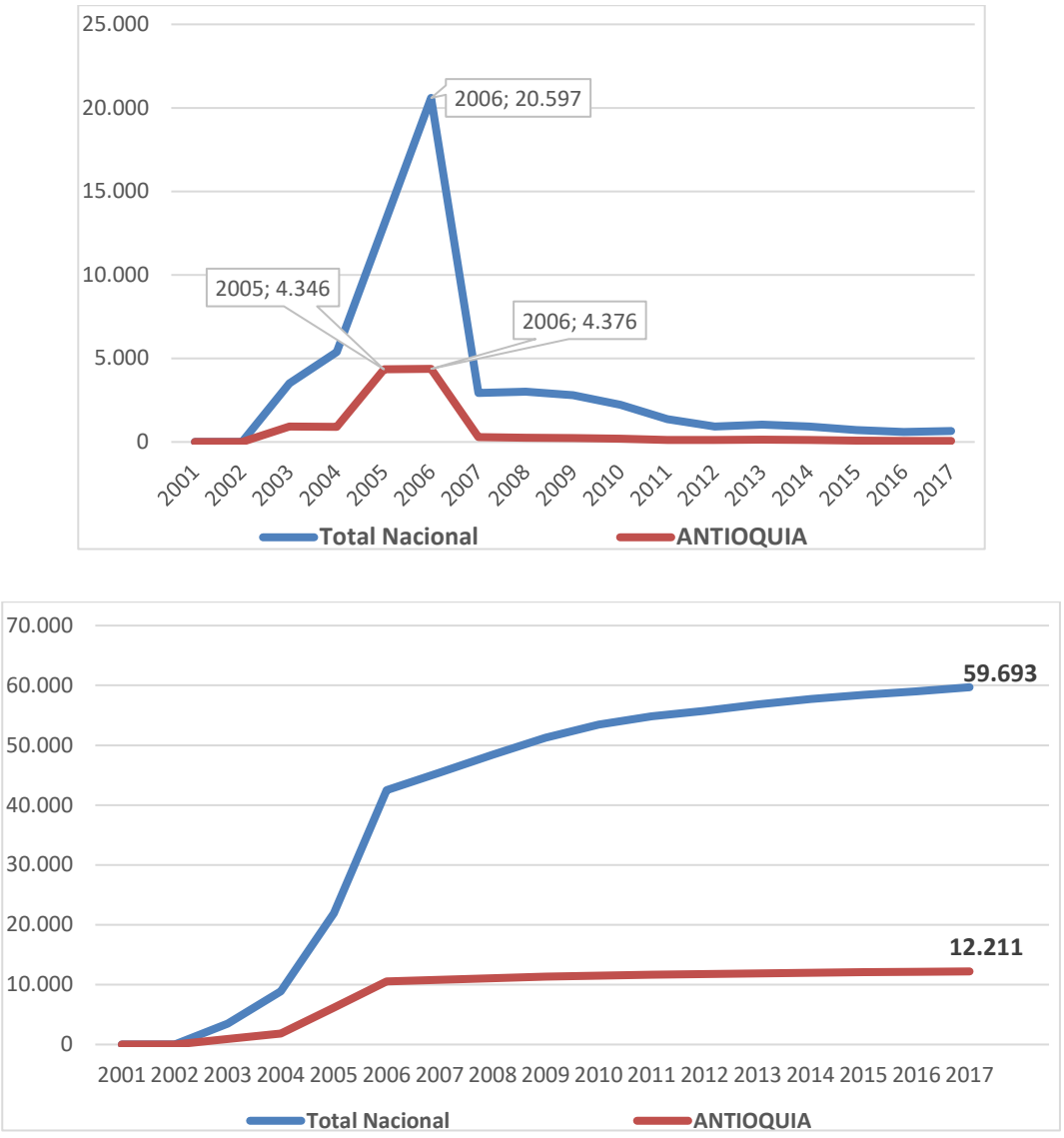


Figure 3.2. Demobilized Antioquia vs. Colombia. Source: Own construction with information on ARN, 2017

Figures 3.3, 3.4, and 3.5, show some of the most relevant variables related to the demobilized combatants and their commitment to stay in the process through the entire route drawn by the program. These variables are also the ones that mark how difficult it will be to access the labor market autonomously or prepare it for this purpose through the program. First, we find

family conditions, if you are married / have a spouse or not and if you have children, a situation that increases the haste of income generation. In both the Colombian and Antioquia case, the ex-combatants are divided very evenly according to this indicator. Second, there is the age group. Older people have more difficulty finding work and less time or availability to receive adequate training, according to the data 96% of the demobilized in Antioquia is over 26 years, moreover, 28% is over 41. Lastly, the educational level shows that grade 11 is the maximum level of schooling observed in the sample, 51.49% of the demobilized combatants in the country have secondary education or less and of these, 64.75% are located in Antioquia, where the same proportion is 49.73%. Finally, Table 3.2 gives a comparative summary of the main characteristics of the population that entered the demobilization process.

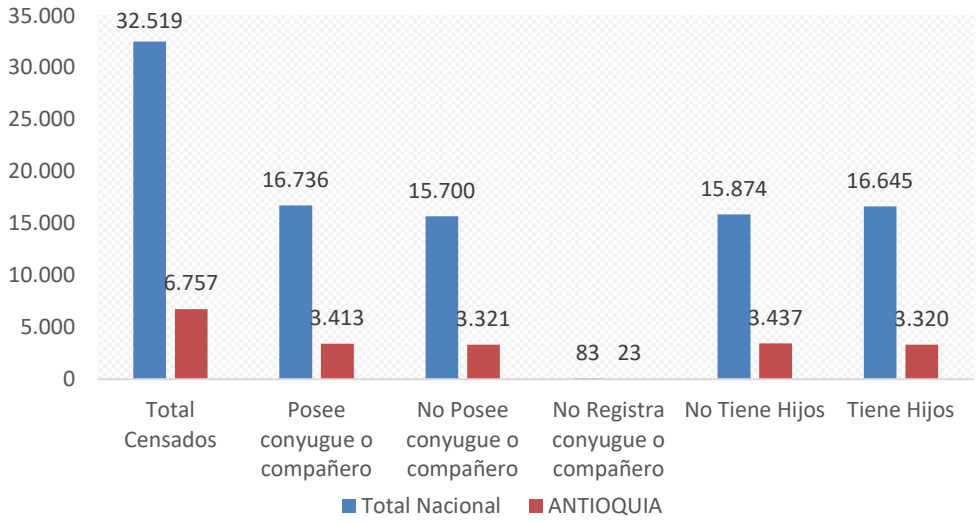


Figure 3.3. Family conditions of the demobilized, Antioquia vs. Colombia. Source: Own construction with information on ARN, 2017

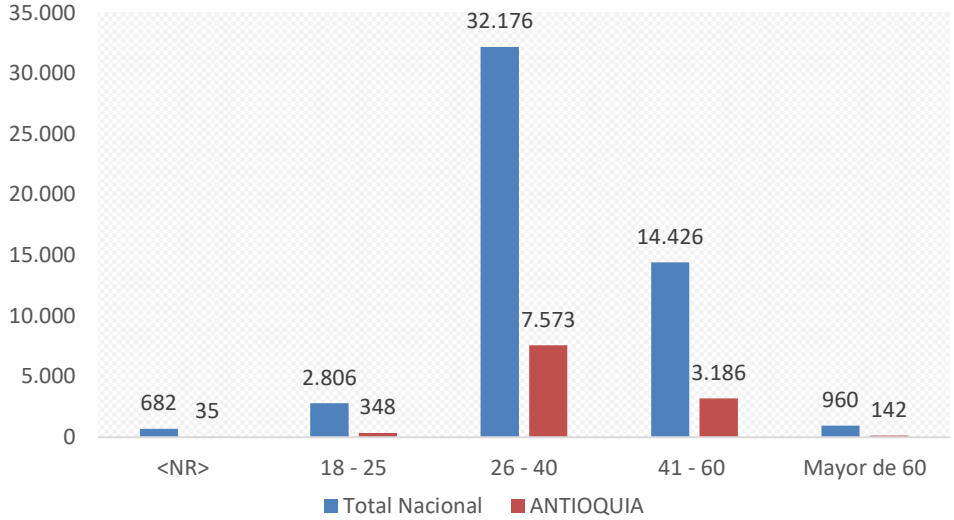


Figure 3.4. Age groups of the demobilized. Source: Own construction with information on ARN, 2017

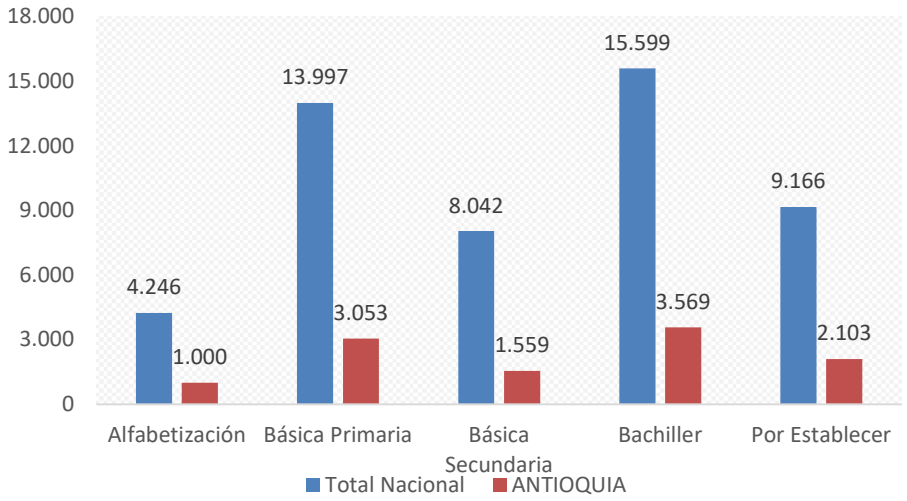


Figure 3.5. Educational level of the demobilized. Source: Own construction with information on ARN, 2017

Table 3.2.Comparative summary

		ANTIOQUIA		*COLOMBIA*	
Total Población que Ingresó al Proceso		11.284	22,1%	51.050	
OCUPACIÓN	Desocupados	642	16,6%	3.861	7,6%
	Ocupados en el sector Formal	2.432	26,9%	9.051	17,7%
	Ocupados en el sector Informal	3.691	19,3%	19.17	37,6%
	Población Económicamente Inactiva	1.630	22,6%	7.215	14,1%
EDUCACIÓN	Técnico Profesional	26	33,8%	77	0,2%
	Tecnológico	1.005	42,7%	2.353	4,6%
	Educación Superior	130	26,5%	490	1,0%
REINCIDENCIA	Nro. de Homicidios	920	30,7%	3.002	5,9%
	Nro. de Individuos con Registro de Riesgo	866	16,7%	5.187	10,2%
	Cálculo de Reincidencia Probada	1.600	32,4%	4.945	9,7%
PERSONAS ATENDIDAS	Total Atendidos	2.783	17,5%	15.87	31,1%
	Acompañamiento Psicosocial	2.574	16,8%	15.31	96,5%
	Gestión Educación	241	15,6%	1.540	9,7%
	Formación para el Trabajo	394	20,4%	1.927	12,1%

Source: Own construction with information on ARN, 2017

Table 3.3.Characteristics of the demobilized in Colombia

Ex Grupo	Género		Edad					Total general	%
	Femenino	Masculino	NR	18-25	25-40	41-60	> 60		
AUC	2.236	28.460	0,0%		54,2%	43,6%	2,1%	30.696	59,3%
ELN	856	2.729	0,2%	14,6%	65,9%	18,1%	1,1%	3.585	6,9%
EPL	28	78		15,1%	66,0%	18,9%		106	0,2%
ERG	46	83		0,8%	65,9%	30,2%	3,1%	129	0,2%
ERP	46	119			64,8%	34,5%	0,6%	165	0,3%
FARC	3.892	13.197	0,3%	8,7%	65,6%	23,9%	1,5%	17.089	33,0%
Sin dato	3	11		78,6%	21,4%			14	0,0%
Total general	**7.107**	**44.677**	**50**	**2.052**	**30.479**	**18.245**	**958**	**51.784**	100%
Participación	**13,7%**	**86,3%**	**0,1%**	**4,0%**	**58,9%**	**35,2%**	**1,8%**		

Ex Grupo	Nivel Educativo					Tiene Hijos		
	Alfabetización	Básica Primaria	Básica Secundaria	Bachiller	Por Establecer	Sí	No	No Censados
AUC	6,2%	25,5%	16,2%	35,9%	16,1%	32,9%	27,8%	39,2%
ELN	12,3%	28,9%	14,0%	26,1%	18,7%	32,2%	37,8%	29,7%
EPL	6,6%	37,7%	6,6%	26,4%	22,6%	28,3%	47,2%	24,5%
ERG	12,4%	30,2%	14,0%	29,5%	14,0%	31,8%	32,6%	35,7%
ERP	7,9%	30,3%	21,2%	26,7%	13,9%	33,9%	32,7%	33,3%
FARC	10,0%	33,4%	15,1%	27,1%	14,3%	33,6%	36,9%	29,2%
Sin dato	7,1%	64,3%	7,1%	7,1%	14,3%	28,6%	50,0%	21,4%
Total general	**4.103**	**14.730**	**8.134**	**16.703**	**8.114**	**17.134**	**16.361**	**18.220**
	7,9%	28,4%	15,7%	32,3%	15,7%	33,1%	31,6%	35,2%

Source: Own construction with information on ARN, 2017.

3.3 Methods

In accordance with the purpose of this work, three fundamental tasks are considered among the aspects used: (i) the characterization of the reinserted population according to the information described above, (ii) a scheme of incentives for the reinserted population is proposed considering the labor market they are going to face, and (iii) Mincer equations are estimated in order to find the determinants that have the greatest effect on the probability of obtaining formal employment, and on the expected wages of the population of interest in different scenarios. The first task was already described above, the other two are shown below.

A. Methodology

Initially, an incentive scheme for reintegrated combatants is proposed at the time of entering to participate in the labor market shown in Figure 2.6.

At the initial moment of a reinserted combatant seeking a job, he has a probability $\propto$ of getting a job and $(1-\propto)$ of not getting one. If he gets a job, he will have a probability β that his employment will be formal, and in that case, his wage would be W_F, which is a wage that includes all the respective social benefits, while the probability that the employment is informal will be $(1-\beta)$ and the respective wage would be W_I.

If the individual does not get a job, he will have two possible decisions: start the search again or reoffend as a criminal. In any case, the objective of this work is to look for the factors that most affect the probability that reintegrated combatants will get a job, and that such employment is in the formal sector, taking into account that informal jobs can increase incentives to keep committing crimes. That is, we are interested in maximizing the combination of probabilities. $\propto \beta$.

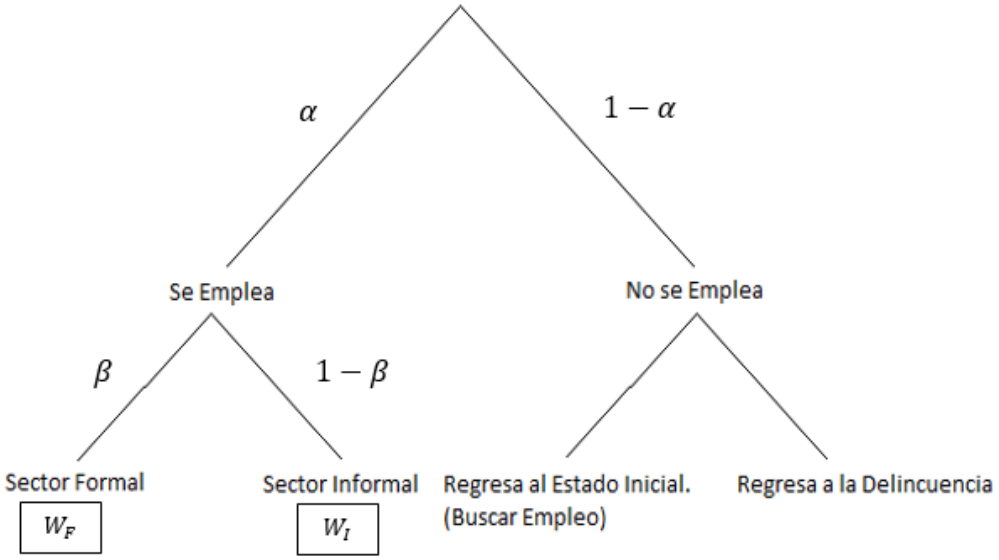


Figure 3.6. Scheme of Incentives for a Reinserted Combatant. Source: own elaboration.

In these terms, the expected income of the demobilized would be given by:

$$E[Income\ Formal\ Sector] = \alpha\beta W_F$$
$$E[Income\ Informal\ Sector] = \alpha(1-\beta)W_I$$

So the individual would choose the formal sector, as long as it is fulfilled that:

$$E[Income\ Formal\ Sector] > E[Income\ Informal\ Sector]$$

And this implies that:

$$\alpha\beta W_F > \alpha(1-\beta)W_I$$

Thus,

$$\beta W_F > (1-\beta)W_I$$

In this regard, it is necessary to mention that the probabilities α y β can be influenced through public policies that maximize the expected income of each individual by improving the characteristics of their resume.

Estimating Mincer equations, it is possible to estimate the probability of obtaining formal and informal employment for different profiles specified in terms of characteristics such as: Educational level, Age, Experience, Gender, Department and whether you reside in rural or urban areas.

B. Data

- ✔ Information from the Large Integrated Household Survey for 2017 (January - October).
- ✔ Information from the Agency for Reintegration and Normalization.

C. Model

The probabilities of obtaining formal or informal employment and the expected wages of each individual, according to their characteristics, are estimated using Mincer equations in two stages. The first stage estimates the probability of getting formal employment and the second stage shows the expected wage of each individual, taking into account the mentioned probabilities.

✔ First Stage: Probability of Employing in the Formal or Informal Sector:

$$Pr\,Pr\,(Formal\,Job_i = 1) = \alpha_0 + \beta_0 gender_i + \beta_1 headhousehold_i + \beta_2 education_i + \beta_3 experience_i + \beta_4 age_i + \beta_5 rural_i + \mu_d + \mu_i$$

$$Pr\,Pr\,(Informal\,Job_i = 1) = \alpha_0 + \beta_0 gender_i + \beta_1 headhousehold_i + \beta_2 educaction_i + \beta_3 experience_i + \beta_4 age_i + \beta_5 rural_i + \mu_d + \mu_i$$

The first stage is estimated using PROBIT regression models.

✔ In the second stage is estimate the equations of Mincer for wages $Ln(W_i)$:

$$Ln(W_i) = \alpha_0 + \beta_0 gender_i + \beta_1 headhousehold + \beta_2 education_i + \beta_3 experience_i + \beta_4 experience^2{}_i + \beta_5 age_i + \beta_6 formal_i + \beta_5 rural_i + \mu_d + \mu_i$$

The second stage is estimated using OLS regressions in logarithms.

3.4 Results

Table 3.4 shows the results in the first stage, where the probabilities of obtaining formal or informal employment are estimated. In general terms, it is common to find that the coefficients of a variable have opposite signs in the two models, taking into account that the aspects that potentiate the probability of obtaining formal employment are at the same time those that minimize the risk of falling into informality. For example, the coefficient for schooling is positive in the model that explains the probability of getting formal employment, but negative in the model for the probability of entering informality.

Table 3.5 shows the marginal effects on the average of the variables for the probability of obtaining formal or informal employment. These marginal effects show that, for example, men are 5.1% more likely to get formal employment compared to women. Additionally, on the average of the sample, an additional year in age decreases the probability of obtaining formal jobs, informality predominates in rural areas, and especially, a result that attracts a lot of attention is that for each year of schooling increases the probability of obtaining formal employment by 3.7% and reduces the probability of being employed informally in the same magnitude.

Table 3.4. Result Probability of Formal and Informal Occupation (Probit)

VARIABLES	Pr(Empleo Formal=1)	Pr(Empleo Informal=1)
Género (Hombre)	0.176***	-0.078***
	(0.001)	(0.001)
Jefe de Hogar	0.271***	-0.122***
	(0.001)	(0.001)
Edad	-0.002***	0.002***
	(0.000)	(0.000)
Años Escolaridad	**0.128*** **	**-0.119*** **
	(0.000)	**(0.000)**
Experiencia	0.001***	0.002***
	(0.000)	(0.000)
Zona (Rural)	-0.300***	0.362***
	(0.001)	(0.001)
Constant	-1.555***	0.894***
	(0.002)	(0.001)
Efectos Fijos Departamento	Sí	Sí
Observations	296,227	296,227
Pseudo R-squared	0.192	0.193
Standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Source: own elaboration.

Table 3.5. Marginal Effects Average Probability of Formal and Informal Occupation (Probit)

VARIABLES	dy/dx Empleo Formal	dy/dx Empleo Informal
Género (Hombre)	0.051***	-0.024***
	(0.000)	(0.000)
Jefe de Hogar	0.079***	-0.038***
	(0.000)	(0.000)
Edad	-0.001***	0.001***
	(0.000)	(0.000)
Años Escolaridad	**0.037*** **	**-0.037*** **
	(0.000)	**(0.000)**
Experiencia	0.000***	0.000***
	(0.000)	(0.000)
Zona (Rural)	-0.087***	0.113***
	(0.000)	(0.000)
Observations	296,227	296,227
Standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Source: own elaboration.

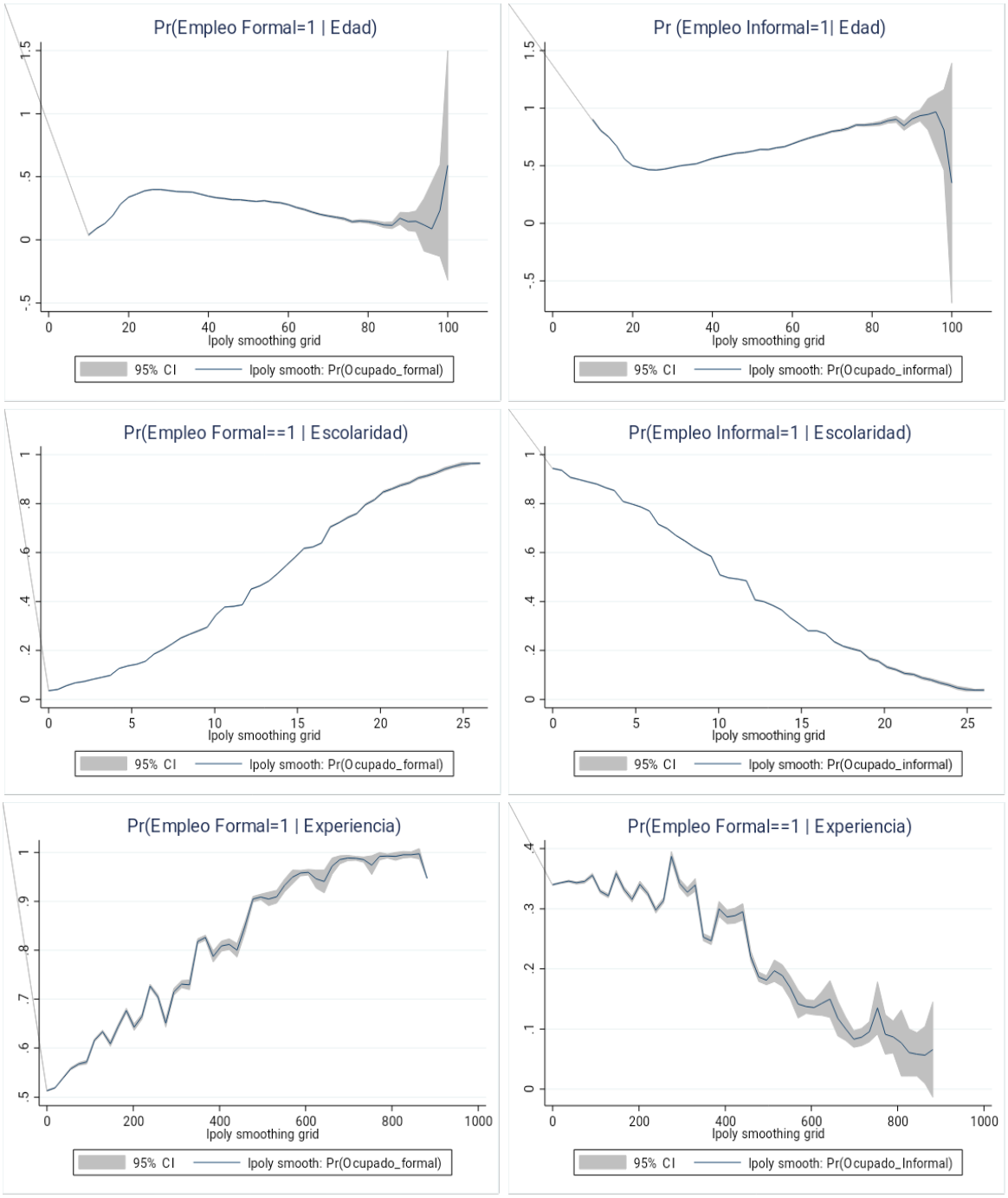


Figure 3.7. Probabilities of Getting Formal and Informal Employment (by Age, Education and Experience). Source: own elaboration.

On the other hand, Figure 3.7 shows the probabilities of obtaining formal and informal employment, taking into account the age of the individuals, the years of schooling and the months of experience. In general terms, it is observed that between the ages of 18 and 26, the probability of getting formal jobs increases, but from then on, it begins to decrease. With regard to years of education and experience, the results are very strong. Each year of education significantly increases the probability of being formally employed, and reduces

that of entering informality, and the same occurs with work experience. In this sense, the first clear path observed is that education and work experience are the variables that must be impacted with public policies to mitigate the risk of the demobilized combatants returning to crime.

Now, it is appropriate to show the results of the second stage of the model, to know those variables that maximize the salary of individuals, according to the sample that was observed in the characterization of the demobilized combatants. Thus, Table 3.6 shows the results for the wage equation estimates. In this case, the coefficients of the variables are in line with those of the estimates of the probabilistic models shown above. On average, one year of schooling increases a Colombian's salary by 6.7%, and the formality award implies wages above informal ones by 63%. It should be noted that in this modeling only employed persons who report some type of salary are taken into account.

Table 3.6. Wage Equation Result - Second Stage

VARIABLES	Ln (Ingreso Mensual)
Género (Hombre)	0.372***
	(0.006)
Jefe de Hogar	0.209***
	(0.006)
Edad	0.000*
	(0.000)
Años Escolaridad	0.067***
	(0.001)
Experiencia	0.002***
	(0.000)
Experiencia^2	-0.000***
	(0.000)
Formal	0.635***
	(0.006)
Area (Rural)	-0.103***
	(0.008)
Constant	12.064***
	(0.014)
Efectos Fijos Departamento	Sí
Observations	222,580
R-squared	0.410
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Source: own elaboration.

Rural areas represent lower wages, at around 10.3%. This is one of the main public policy challenges, taking into account that in the countryside, agriculture and agribusiness can represent one of the main sources of employment for the demobilized population, but in these

cases, the salary remuneration must be compensated, and also promote formality in these types of territories. Nevertheless, to have a clear picture of the demobilized employment situation, it is necessary to have a panorama that combines the results of the first and second stages.

Table 3.7. Probability of being employed, by profile

Perfil	Pr(Empleo Formal=1)		Salario Promedio Estimado		Brecha en Salario Esperado (Formal vs. Informal)	
	Sin Bachillerato	Con Bachillerato	Sin Bachillerato	Con Bachillerato	Sin Bachillerato	Con Bachillerato
Entre 18 y 25 Años Sin Experiencia	20.08%	40.42%	379,691	610,881	-68.7%	16.1%
Entre 26 y 40 Años Sin Experiencia	16.86%	44.10%	412,149	760,848	-69.5%	51.9%
Entre 40 y 50 Años Sin Experiencia	14.80%	45.37%	391,305	740,901	-74.5%	48.4%
Mayores de 50 Años Sin Experiencia	14.77%	50.88%	367,343	754,254	-77.6%	68.8%

Source: own elaboration.

Table 3.7 shows probabilities of being employed for different profiles of individuals by age. According to what we observe in the characterization of the demobilized in Colombia, we can conclude that most of these are between 26 and 40 years of age, have no schooling and much less work experience. Then, this group will be the one of main interest for the conclusions offered in this study, and for the public policy proposals.

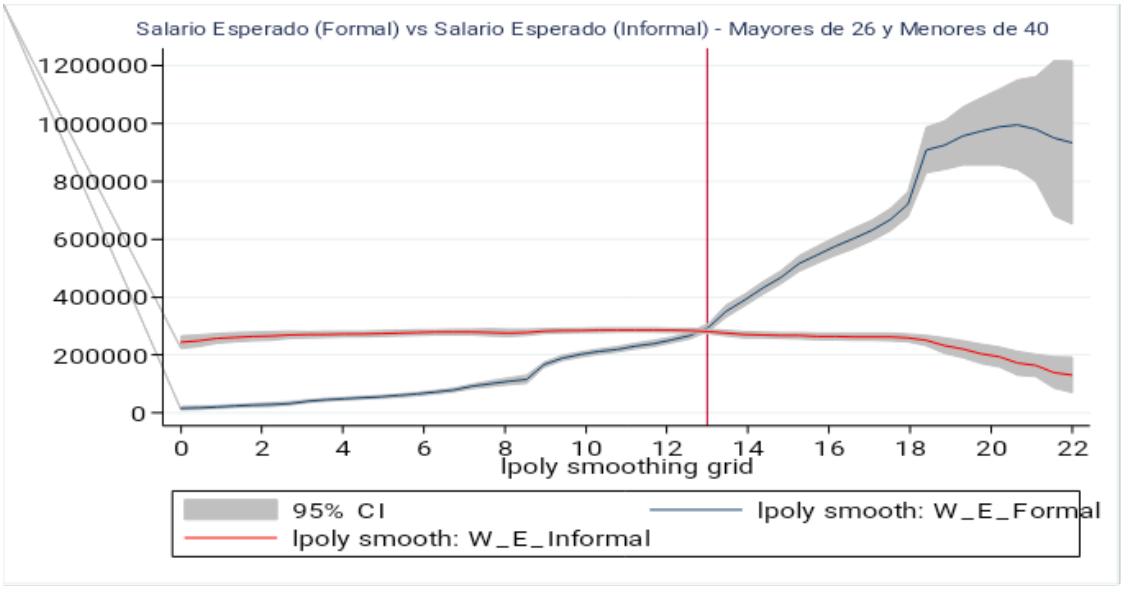


Figure 3.8. Expected Wage for Target Population in Formal or Informal Sector. Source: own elaboration.

For example, in this interest group, the probability of getting a formal job if you have a high-school diploma is 44.1%, while if you do not have a high-school diploma, it will be only 16.86%. In this case, the high-school diploma represents in this group of individuals a wage gap of 51.9% compared to the wage they would obtain if they were employed informally. In these terms, we already have elements to claim that providing high school studies the incentives to participate in the informal labor market are reversed.

By focusing the analysis on the aforementioned interest group, it is found that the incentives to belong in the formal or informal sector are reversed when each demobilized person has at least 13 years of education, as shown in Figure 3.8, below. Henceforth, the economic prize in the formal labor market more than compensates for the effort to study.

However, it is important to take into account the context in which these results are presented. For a population between 26 and 40 years, it is not feasible to offer 13 years of education, because the cost does not compensate for the benefit, neither for this population, nor for the country, taking into account that the youngest of the group (26 years) they would finish this cycle of studies with 39 years, and in this case, the labor market would discriminate them by age. Therefore, one has to be practical. These 13 years of education are equivalent to 11 years of complete high school plus two years of a technical or technological education.

In this sense, the ideal scenario would be to offer continuing education options in high school, which are intensive, recursive and accelerated. This will allow the demobilized combatants to graduate in about 1 or 2 years of high school and keep them busy to mitigate the risk of returning to illegality. After the completion of high school, there would be an offer of technical or technological education that can be accompanied by professional practices to complement their profiles with some academic experience. It is worth mentioning that these offers of higher education must be prioritized by economic sectors and geographical areas. In our results, we find that the areas where demobilized workers are most likely to succeed in the labor market are the departments of Antioquia, Bogotá, Caldas, Cundinamarca, Quindío, Risaralda, Valle del Cauca and Santander.

3.5 Conclusions and Development Proposal

In Colombia, and under the current circumstances, the labor market does not generate incentives for the incorporation of reinserted population. On the contrary, this population is more likely to relapse into crime, or to engage in informal activities, which easily brings them closer to illegality.

To improve the chances of engaging the reinserted population, it is necessary to guarantee training levels of at least 13 years. However, since this population is mostly between 26 and 40 years of age, it is not feasible to offer 13 years of education, since the cost does not compensate for the benefit, neither for this population, nor for the country, taking into account that the youngest of the group (26 years) would end this cycle of studies with 39 years, and in this case, the labor market would discriminate against them by age. Therefore, one must be practical. These 13 years of education are equivalent to 11 years of complete high school plus two years of a technical or technological education.

In this sense, the ideal scenario would be to offer continuing education options in high school, which are intensive, recursive and accelerated. This will allow the demobilized combatants to graduate in about 1 or 2 years of high school and keep them busy to mitigate the risk of returning to illegality. After the completion of high school, there would be an offer of technical or technological education that can be accompanied by professional practices to complement their profiles with some academic experience.

Additionally, it is important to mention that these higher education offers must be prioritized by economic sectors and geographical areas. In our results, we find that the areas where demobilized workers are most likely to succeed in the labor market are the departments of Antioquia, Bogotá, Caldas, Cundinamarca, Quindío, Risaralda, Valle del Cauca and Santander.

Finally, to meet the reinserted population, it is required, in the first instance, to level the years of education so as to facilitate their incorporation into the formal labor market. And, secondly, the creation of programs tending to favor the creation of companies (by way of enterprises) that allow this population to generate income is demanded.

All these proposals will help to improve the economic and social conditions of this population that has been part and victim of the armed conflict in the country, and will help to maintain the peace and sovereignty of the nation.

References

Chapter 1.

1. BARÓN, J. D. (2004). "¿Qué sucedió con las disparidades económicas regionales en Colombia entre 1980 y el 2000?", en A. Meisel (editor), Macroeconomía y regiones en Colombia, Banco de la República, Cartagena.
2. BARÓN, J. D. and MEISEL A. (2003). La Descentralización y las Disparidades Económicas Regionales en Colombia en la Década de 1990, Documentos de Trabajo sobre Economía Regional, CEER, Banco de la República, Cartagena, N.36.
3. BELSLEY D.A., KUH E. and WELSCH R.E. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Wiley Series in Probability and Statistics, Print ISBN: 9780471058564
4. BIRCHENALL, J., and G. MURCIA (1996). "Convergencia Regional: Una Revisión del Caso Colombiano", Desarrollo y Sociedad, 40, 273 – 308.
5. BOLLEN K. and JACKMAN R. (1990). Regression diagnostics: An expository treatment of outliers and influential cases. In: Fox J, Long J, editors. Modern Methods of Data Analysis. Newbury Park, CA: Sage. pp. 257–291.
6. BONET, J. and A. MEISEL (1999). "La convergencia regional en Colombia: una visión de largo plazo, 1926 – 1995", Coyuntura Económica, 29,1, 69 – 106.
7. BONET, J. and MEISEL A. (2006). "El legado colonial como determinante del ingreso per cápita departamental en Colombia", Documentos de trabajo sobre Economía Regional, Centro de Estudios Económicos Regionales (CEER), Banco de la República, Cartagena, N.75.
8. BRAKMAN, S., GARRETSEN H. and SCHRAMM M. (2004). "The spatial distribution of wages and Employment: Estimating the Helpman-Hanson model for Germany", Journal of Regional Science, 44, pp. 437-466.
9. BRAKMAN S., GARRETSEN, H. and VAN MARREWIJK, C. (2009). "Economic Geography within and between European nations: The role of Market Potential and density across space and time", Journal of Regional Science, 49(4), pp. 777–800.
10. BREINLICH H. (2006). "The spatial income structure in the European Union: what role for Economic Geography?", Journal of Economic Geography, 6(5), pp. 593–617.
11. BRUNA, F., LOPEZ-RODRIGUEZ, J. and FAÍÑA, A. (2015). "Market Potential, Spatial Dependences and Spillovers in European Regions", Regional Studies, 50(9), pp. 1551-1563.
12. CLARK, C., (1940). The Conditions of Economic Progress, Macmillan, London.
13. CLARK, C., F. WILSON and J. BRADLEY (1969). Industrial location and economic potential in Western Europe. Regional Studies 3(2), 197–212.
14. CROZET M. (2004). "Do Migrants Follow Market accesss? An Estimation of a New Economic Geography Model", Journal of Economic Geography, 4, pp. 439-458.
15. DANE (2013). Documento metodológico y resultados de la retropolación 1975 – 2005. Dirección de Síntesis y Cuentas Nacionales (DSCN).
16. FAÍÑA, A. and J. LÓPEZ-RODRÍGUEZ (2005). The economic geography of EU income: evolution since the eighties. In: Trívez, F.J., J. Mur, A. Angulo, M. Ben Kaabia, and B.

Catalán (eds.), Contributions in spatial econometrics, Ch. 9. Copy Center Digital, Zaragoza, Spain.

17. FAÍÑA, A. and J. LÓPEZ-RODRÍGUEZ (2006). European Union Enlargement, European Spatial Development Perspective and Regional Policy: Lessons from Population Potentials. Investigaciones Regionales 9, 3–21.
18. FINGLETON, B. (2006). The new economic geography versus urban economics: an evaluation using local wage rates in Great Britain. Oxford Economic Papers 58(3), 501–530
19. HANSON, G.H. (2005). "Market potential, increasing returns and geographic concentration", Journal of International Economics, 67(1), pp. 1–24.
20. HARRIS, C.D. (1954). "The Market as a Factor in the Localization of Industry in the United States", Annals of the Association of American Geographers, 44(4), pp. 315–348.
21. HEAD, K. y MAYER, T. (2000). "Non-Europe: The Magnitude and Causes of Market Fragmentation in the EU", Review of World Economics (Weltwirtschaftliches Archiv), 136, pp. 284-314.
22. HEAD, K. and T. MAYER (2006). Regional wage and employment responses to market potential in the EU. Regional Science and Urban Economics 36(5), 573–594.
23. HEAD, K. and T. MAYER (2011). Gravity, market potential and economic development. Journal of Economic Geography 11(2), 281–294.
24. KARAHASAN, B. C. and E. LÓPEZ-BAZO (2013). The Spatial Distribution of Human Capital: Can It Really Be Explained by Regional Differences in Market Access? International Re-gional Science Review 36(4), 451–480.
25. KEEBLE, D., OWENS, P.L. and THOMPSON, C. (1982): "Regional accessibility and economic potential in the European Community", Regional Studies, 16(6), pp. 419–432.
26. KRUGMAN, P. (1991). Increasing returns and economic geography. Journal of Political Economy 99(3), 483–499.
27. KRUGMAN, P. (1992). A Dynamic Spatial Model. NBER Working Paper 4219. National Bureau of Economic Research, Inc. On line: http://ideas.repec.org/p/nbr/nberwo/4219.html
28. LÓPEZ RODRÍGUEZ, J., M. A. MÁRQUEZ and A. FAÍÑA (2011). ¿Hasta qué punto la perificidad económica es responsable de las diferencias en el pib per cápita entre las provincias españolas? El Trimestre Económico LXXVIII-3(311), 583–611.
29. LÓPEZ-RODRÍGUEZ, J. and M. C. C. ACEVEDO-VILLALOBOS (2013). El acceso a los mercados y la disparidad del ingreso en los departamentos colombianos. El Trimestre Económico LXXX-4(320), 869–901.
30. LÓPEZ-RODRÍGUEZ, J. and A. FAÍÑA (2006). Does distance matter for determining regional in-come in the European Union? An approach through the market potential concept. Applied Economics Letters 13(6), 385–390.
31. LÓPEZ-RODRÍGUEZ, J., A. FAÍÑA and B. COSMIN-GABRIEL (2011). Economic Remoteness and Wage Disparities. In Romania. Tijdschrift voor economische en sociale geografie 102(5), 594–606.

32. LÓPEZ-RODRÍGUEZ, J., A. FAÍÑA and J. LÓPEZ-RODRÍGUEZ (2007). Human capital accumulation and geography: Empirical evidence from the European Union. Regional Studies 41(2), 217–234.
33. LUCAS, R. (1988). On the mechanisms of development planning. Journal of Monetary Economics, 22(1), 3-42.
34. MANKIW, G.,ROMER, D., and D.N. WEIL (1992). The Quarterly Journal of Economics, 107, 2, 407-437.
35. MEISEL, A. (1993). "Polarización o Convergencia? A Propósito de Cárdenas, Pontón y Trujillo", Coyuntura Económica, 23, 2, 153 – 160.
36. MORA, J., and B. SALAZAR (1994). "Fábula y Trama en el Relato de la Convergencia", Boletín Socioeconómico, 27, CIDSE – Universidad del Valle.
37. NIEBUHR, A. (2006). Market access and regional disparities. The Annals of Regional Science 40(2), 313–334.
38. NITSCH, V. (2000). "National borders and international trade: evidence from the European Union", Canadian Journal of Economics, 22, pp. 1091–1105.
39. OVERMAN, H.G., REDDING, S., and VENABLES, A.J. (2003). Economic Geography: survey of empirics. In CHOI, E.K. Harrigan, J. (eds), Hadbook of International Trade, Malden: Blackwell.
40. PSACHAROPOULOs, G., (1994). Returns to investment in education: a global update. World Development 22 (9),1325–1343
41. REDDING, S. J. (2011). Economic Geography: A Review of the Theoretical and Empirical Literature. In Bernhofen, D., R. Falvey, D. Greenaway, and U. Kreickemeier (eds.), Palgrave Handbook of International Trade, 1st edition, pp. 497-531. Palgrave Macmillan, London.
42. REDDING, S. J. and P. K. SCHOTT (2003). Distance, skill deepening and development: will peripheral countries ever get rich? Journal of Development Economics 72(2), 515–541.
43. REDDING, S. J. and A. VENABLES (2004a): Geography and export performance: external market access and internal supply capacity. In: Challenges to globalization: Analyzing the economics, University of Chicago Press, 95-130.
44. REDDING, S. J. and A. VENABLES (2004B): Economic geography and international inequality. Journal of International Economics, 62(1), 53-82.
45. ROCHA, R., and A. VIVAS (1998). "Crecimiento Regional en Colombia: Persiste la Desigualdad?", Revista de Economía del Rosario, 1, 1, 67-108.
46. ROMER, P.M. (1994). The origins of endogenous growth. The Journal of Economic Perspectives, 8(1), 3-22.
47. STEWART, J.Q. (1947). Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population," Geograbhical Review, XXXVII, 461-485
48. STEWART, J.Q. (1948). Demographic Gravitation: Evidence and Applications, Sociometry, XI (1948): 31-58
49. STEWART, J.Q. (1952). A Basis for Social Physics, Impact of Science on Society, 111 (1952): 110-133
50. TABELLINI, G. (2005). Culture and institutions: economic development in the regions of Europe. CESifo Working Paper Series No. 1492.

Chapter 2.

51. ANGRIST, J. and V. LAVY (1999). “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement” Quarterly Journal of Economics, 114(2), pp. 533-575.
52. BAHL, R. (2000). “Intergovernmental Transfers in Developing and Transition Countries: Principles and Practice” Municipal Finance Background Series, The World Bank, April.
53. BONET, J. and URREGO, J. (2014). “El Sistema General de Regalías: ¿Mejoró, Empeoró o Quedó Igual?” Documentos de Trabajo sobre Economía Regional No. 198, Enero.
54. CALONICO, S., M. D. CATTANEO, and R. TITIUNIK. (2014). “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs” Econometrica 82(6): 2295-2326.
55. CÁRDENAS, M. (2013) “Colombia: Reforma al Sistema General de Regalías” Ministerio de Hacienda y Crédito Público, July.
56. CORBI, R.; PAPAIOANNOU, E. and SURICO, P. (2014). “Federal Transfer Multipliers. Quasi-Experimental Evidence From Brazil” NBER WP No. 20751, December.
57. DNP (2013). “Sistema General de Regalías: Presentación General”.
58. DNP (2014). “Evaluación del Desempeño Integral de los Municipios y Distritos, Vigencia 2013.
59. Economía Urbana Ltda– Centro Nacional de Consultoría (2012). “Realizar una Evaluación de las Regalías Directas, una Evaluación de Resultados del Fondo Nacional de Regalías (FNR) y Elaborar un Estudio de Sostenibilidad Financiera a Mediano Plazo de las 80 Entidades Territoriales Mayores Receptoras de Regalías en el País” Programa para el Fortalecimiento de la Información Pública, Seguimiento y Evaluación para la Gestión por Resultados en Colombia, Crédito BIRF 7620 CO, Departamento Nacional de Planeación, Bogotá D. C., mayo 24.
60. ECHEVERRY, J.C.; ALONSO, G. and GARCÍA, A. (2011). “Por Qué es Necesaria la Creación de un Sistema General de Regalías” Notas Fiscales, Ministerio de Hacienda y Crédito Público No. 2, Enero.
61. FERRAZ, C. and FINAN, F. (2009). “Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance” NBER WP No. 14906, April.
62. HAHN, J.; TODD, P. and VAN DER KLAAUW, W. (2001). “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design” Econometrica, Econometric Society, vol. 69(1), pages 201-09, January.
63. MORALES, L.F. and MEDINA, C. (2017). “Assessing the Effect of Payroll Taxes on Formal Employment: The Case of the 2012 Tax Reform in Colombia” Economia Journal, The Latin American and Caribbean Economic Association - LACEA, Vol. 18, No. 1 (Fall), pages 75-124, November.
64. OATES, W.E. (1972). “Fiscal Federalism” (New York: Harcourt Brace Jovanovich).
65. OATES, W.E. (1993). “Fiscal Decentralization and Economic Development” National Tax Joumal, 46 (18): pp. 237-243.
66. SÁNCHEZ, F. (2006). “Descentralización y Progreso en el Acceso a los Servicios Sociales de Educación, Salud y Agua y Alcantarillado” Documento CEDE 15, marzo.

67. TANZI, V. (1995). “Fiscal Federalism and Decentralization: A Review of Some Efficiency and Macroeconomic Aspects” Annual World Bank Conference on Development Economics.
68. VAN DER KLAAUW, W. (2002). “Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach” International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 43(4), pages 1249-1287, November.

Chapter 3.

69. Agencia Colombiana para la Reintegración (2014). Evolución del Proceso de Reintegración Fortaleza Institucional basada en la experiencia y lecciones aprendidas 2014. Bogotá, Colombia.
70. ALDERETE, M. and GUTIÉRREZ, L. (2012). TIC y productividad en las industrias de servicios en Colombia. Lecturas de Economía, (77), 163-188.
71. ALTONJI, J. G., and BLANK, R. M. (1999). Race and gender in the labor market. Handbook of labor economics, 3, 3143-3259.
72. BADEL, M.E. and TRUJILLO, E. (1998). “Los Costos Económicos de la Criminalidad en Colombia: 1991-1996”. Archivos de Economía No 76.
73. BARRON, J. M.; BLACK, D. A. and LOEWENSTEIN, M. A. (1987). Employer size: The implications for search, training, capital investment, starting wages, and wage growth. Journal of Labor Economics, 5(1), 76-89.
74. BØÅS, M. and HATLØY, A. (2008). ‘Getting in, getting out’: militia membership and prospects for re-integration in post-war Liberia. The journal of modern African studies, 46(1), 33-55.
75. BERTRAND, M. and MULLAINATHAN, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. American economic review, 94(4), 991-1013.
76. BLACK, S. E. and LYNCH, L. M. (1996). Human-capital investments and productivity. The American economic review, 86(2), 263-267.
77. CÁRDENAS, M. (2002). “Economic Growth in Colombia: A Reversal of ‘Fortune’?”. Center for Internacional Development , Harvard University. Mimeo.
78. DATTA, D. K.; GUTHRIE, J. P. and WRIGHT, P. M. (2005). Human resource management and labor productivity: does industry matter?. Academy of management Journal, 48(1), 135-145.
79. DELERY, J. E. and SHAW, J. D. (2001). The strategic management of people in work organizations: Review, synthesis, and extension. In Research in personnel and human resources management (pp. 165-197). Emerald Group Publishing Limited.
80. ECHEVERRY, J. C.; SALAZAR, N. and NAVAS, V. (2001). “El Conflicto Colombiano en el Contexto Internacional”. En Martinez, A. (Editora), Economía, Crimen y Conflicto. Universidad Nacional: Bogotá; 77-128.
81. ESCOBAR, N. S. and MEERTENS, D. (1997). Desarraigo, género y desplazamiento interno en Colombia. Nueva sociedad, 148, 30-43.

82. FISAS, V. (2011). Introducción al Desarme, Desmovilización y Reintegración (DDR) de excombatientes. Quaderns de Construcció de Pau, 24, 1-20.
83. HECKMAN, J.J.; STIXRUD, J. and URZUA, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor economics, 24(3), 411-482.
84. RUEDA, R. P. (2017). La historia de las guerras. Debate.
85. SAMUELSON, P. A. and NORDHAUS, W. D. (1989). Economics (19th).
86. SOLOW, R.M. (1956). A Contribution to the Theory of Economic Growth. The Quarterly Journal of Economics, Vol. 70, 65-94.
87. SCHULTZ, T. W. (1961). Investment in human capital. The American economic review, 51(1), 1-17.
88. TODARO, R.; GODOY, L. and ABRAMO, L. (2002). Desempeño laboral de hombres y mujeres: opinan los empresarios. Cadernos Pagu, (17-18), 197-236.
89. TORRES, F. J. S. and DÍAZ, A. M. (2005). Los efectos del conflicto armado en el desarrollo social colombiano, 1990-2002. Universidad de los Andes, Facultad de Economía, CEDE.
90. VALENCIA, G.D. (2007). Reconstrucción analítica del proceso de desarme, desmovilización y reinserción con las Autodefensas Unidas de Colombia, 2002-2007. Perfil de coyuntura económica, (10).
91. VARGAS, J. (2003). «Conflicto interno y crecimiento económico en Colombia». Programa de las Naciones Unidas. Documento sobre conflicto y paz.
92. ZARATE, A.; CRESTTO, M.; MAIZ, A.; RAVEST, G.; PINO, M.I.; VALDIVIA, G.; ... and VILLARROEL, L. (2009). Influencia de la obesidad en los costos en salud y en el ausentismo laboral de causa médica en una cohorte de trabajadores. Revista médica de Chile, 137(3), 337-344.

Conclusions and Contributions of the Thesis

Conclusiones

Esta tesis doctoral se ha centrado en investigar tres temas de relevancia para Colombia. En el capítulo I investigamos el impacto del potencial de mercado sobre los diferenciales en los niveles de renta per cápita entre los distintos departamentos que componen la estructura territorial de Colombia. Nuestro análisis además de incorporar estimaciones de la ecuación nominal de salarios de la *Economía Geográfica* usando un panel de datos para el período 1990-2015 da un paso adicional estimando un panel espacial de esa misma ecuación para determinar el impacto que sobre los niveles de renta de un departamento tienen los departamentos vecinos.

Nuestros resultados a partir de una base de datos propia, que ha sido creada por el autor de la tesis al tener que enlazar datos procedentes de diferentes períodos temporales, nos permitió construir una base de datos homogénea (en términos de metodología y definición de variables) que incluye entre otras variables niveles de renta per cápita, dotaciones de capital físico, potencial de mercado y dotaciones de capital humano a nivel departamental. Nuestros análisis preliminares muestran que existe una relación positiva entre el potencial de mercado departamental y los niveles de renta per cápita correspondientes. La estimación empírica del modelo centro-periferia de Economía Geográfica que realizamos en el capítulo I tanto usando econometría de datos de panel como econometría espacial corrobora esas tesis preliminares. El potencial de mercado se convierte en una variable relevante a la hora de explicar los distintos niveles de desarrollo que observamos en los departamentos colombianos. Además los análisis de robustez llevados a cabo en nuestro análisis tanto en términos de endogeneidad del potencial de mercado (Variables Instrumentales, Valores históricos del potencial de mercado, potencial de mercado excluyendo la componente doméstica) como mediante la incorporación de potenciales elementos que pudieran estar influenciados por el potencial de mercado obedeciendo a dinámicas de acumulación como son las dotaciones de capital físico y capital humano nos muestran que el potencial de mercado se muestra como una variable relevante a la hora de explicar los niveles de renta en los departamentos colombianos (desde un punto de vista econométrico-coeficientes para el potencial de mercado positivos, económicamente significativos y estadísticamente significativos). Finalmente, la estimación de la ecuación nominal de salarios utilizando un panel espacial demuestra que alrededor del 50% del impacto del potencial de mercado en las disparidades de ingresos para cada departamento colombiano puede atribuirse a los departamentos vecinos.

El capítulo dos de la tesis evaluó los efectos que las regalías tienen sobre el empleo formal. En Colombia. Este capítulo permitió concluir que el Sistema General de Regalías SGR permitió redistribuir del monto de regalías y utilizar esa importante fuente de recursos, para impulsar el crecimiento regional y municipal, la equidad entre regiones dada la redistribución y disminuir los índices de pobreza vía empleo formal. Además, es claro que con el incremento en la participación de las regalías para los municipales que recibían menores proporciones antes de la reforma el empleo formal se dinamizó.

Este capítulo permitió concluir que focalizar las regalías es un problema de interés central en el país por la inmensa cuantía de las mismas (aproximadamente 1,6% del PIB), y porque son

una fuente importante de financiación para el desarrollo regional, la forma que su adecuada distribución es fundamental para favorecer el empleo formal y dinamiza el desarrollo de las regiones. El haber utilizado la ley 1530 de 2012 y el decreto 1073 de 2012, además de la regresión discontinua fundamenta en las necesidades básicas insatisfechas NBI permitió determinar el impacto en empleo formal al encontrar un efecto positivo de las regalías en el mismo, al tiempo que la reforma aumentó los recursos para entidades territoriales que no los recibían.

Dados los incipientes niveles de desarrollo de los entes territoriales, en especial los que se encuentran en sexta categoría los recursos provenientes de regalías se convirtieron en un vehículo de disponibilidad de recursos favorables a la dinámica territorial y en favor del empleo formal.

Finalmente, los resultados evidencian que hay efectos positivos y significativos de las regalías sobre el empleo formal, principalmente en los sectores de industria; electricidad, gas y agua; actividades inmobiliarias y empresariales; e intermediación financiera.

En el tercer capítulo la tesis aborda un tema de toda la relevancia en el contexto nacional e internacional, toda vez que, luego de un proceso de conflicto armado interno de más de 50 años en Colombia, se firma la paz con la guerrilla más antigua y más rica del mundo. Es allí, donde este capítulo permite comprender tanto algunos aspectos en común como otros asociados al mercado laboral que significarán esfuerzos significativos en materia de política pública. Los resultados de este capítulo de la tesis sugieren que en Colombia dadas las diferentes variables incluidas en el modelo el mercado laboral no genera estímulos para la incorporación de población reinsertada. Además de este resultado el estudio concluye que esta población presenta mayores probabilidades de reincidir en el delito y de engancharse en actividades informales. Esta situación posibilita que la población desmovilizada esté cerca de volver a la ilegalidad.

En el entendido de la conclusión anterior, un camino definitivamente debe ser buscar aumentar la probabilidad de la población reinsertada para engancharse al mercado laboral. Para dar cuenta de ello, el estudio concluye que se requiere garantizar niveles de formación en educación mínimo de 13 años. Una limitante que encuentra el estudio para poder lograr que tengan 13 años de formación es que la mayoría de esta población se encuentra entre los 26 y los 40 años. Un análisis costo – beneficio hace inviable esta posibilidad, incluso para los que hoy cuentan con 26 años pues al terminar el ciclo de formación tendrían alrededor de 40 años lo que para ese momento les limitaría posibilidades en el mercado laboral. Por lo anterior, el estudio encuentra viable diseñar una formación a la medida tanto en la formación básica como superior que sea compatible con la generación de ingresos. Otro camino posible es realizar formación a la medida tanto en sectores económicos como en términos geográficos.

El estudio además infiere que los departamentos en los que tienen mayor probabilidad de encontrar una oportunidad en el mercado laboral los desmovilizados son los departamentos de Antioquia, Bogotá, Caldas, Cundinamarca, Quindío, Risaralda, Valle del Cauca y Santander. Esto se compagina con el potencial de mercado y el PIB.

Finalmente, para atender a la población reinsertada, el estudio concluye que otro camino a explorar es programas tendientes que favorezcan la creación de empresas (a modo de emprendimientos) que permitan a esta población la generación de ingresos.

Conclusions

This doctoral thesis has focused on investigating three topics of relevance to Colombia. In Chapter I we investigate the impact of market potential on differentials in levels of per capita income between the different departments that make up the territorial structure of Colombia. Our analysis, in addition to incorporating estimates of the nominal wage equation of *Geographical Economics* using a data panel for the period 1990-2015, takes an additional step by estimating a spatial panel of that same equation to determine the impact that a department has on income levels of neighboring departments.

A novel database was created by having to link data from different time periods, allowing us to build a homogeneous dataset (in terms of methodology and definition of variables) which includes, among other variables, levels of per capita income, physical capital endowments, market potential and human capital endowments at the departmental level.

Our preliminary analyzes show that there is a positive relationship between the departmental market potential and the corresponding levels of income per capita. The empirical estimation of the center-periphery model of Geographical Economics that we carried out in Chapter I, using both panel data econometrics and spatial econometrics corroborates these preliminary theses. Market potential becomes a relevant variable when explaining the different levels of development that we observe in the Colombian departments.

Furthermore, We carry out robustness tests both in terms of endogeneity of the market potential (Instrumental Variables, Historical values of market potential, market potential excluding the domestic component) and by incorporating potential elements that could be influenced by market potential, obeying accumulation dynamics, such as the endowments of physical capital and human capital show us that market potential is shown as a relevant variable when explaining income levels in Colombian departments (from a econometric point of view - coefficients for market potential positive, economically significant and statistically significant).

Finally, estimating the nominal wage equation using a spatial panel, We show that around 50% of the impact of market potential on income disparities for each Colombian department can be attributed to neighboring departments.

Chapter two of the thesis evaluated the effects that royalties have on formal employment in Colombia. This chapter allowed to conclude that the General SGR Royalties System allowed to redistribute the amount of royalties and use that important source of revenue, to boost regional and municipal growth, increase equity between regions given the redistribution and decrease poverty rates via formal employment. In addition, we find that the increase in the participation of royalties for the municipalities that received lesser proportions before the reform, formal employment became more dynamic.

This chapter allowed to conclude that targeting royalties is a problem of central interest in the country due to the immense amount of these (approximately 1.6% of GDP), and because they are an important source of financing for regional development, the way that their adequate distribution it is essential to promote formal employment and stimulates the development of the regions. Having used Law 1530 of 2012 and Decree 1073 of 2012, in addition to the discontinuous regression based on unsatisfied basic needs -UBN- allowed determining the impact on formal employment by finding a positive effect of royalties on it, while the reform increased resources for territorial entities that did not receive them.
Given the incipient levels of development of territorial entities, especially those in the sixth category, revenue from royalties became a vehicle for the availability of resources favorable to territorial dynamics and in favor of formal employment.

Finally, the results show that there are positive and significant effects of royalties on formal employment, mainly in the industry sectors; electricity, gas and water; real estate and business activities; and financial intermediation.

In the third chapter, the thesis addresses an issue of relevance in the national and international context, since, after a process of internal armed conflict of more than 50 years in Colombia, peace is signed with the oldest and richest guerrilla of the world. This chapter allows us to understand some aspects in common as well as others associated with the labor market that will mean significant efforts in public policy. The results of this chapter of the thesis suggest that, in Colombia, given the different variables included in the model, the labor market does not generate stimuli for the incorporation of the reinserted population. In addition to this result, the study concludes that this population is more likely fall into recidivism and engage in informal activities. This situation enables the demobilized population to be close to returning to illegality.

In the understanding of the previous conclusion, one pathway should be to increase the probability of the reinserted population to engage in the labor market. To account for this, the study concludes that it is necessary to guarantee minimum education training levels of 13 years. A limitation that the study finds in achieving that they ex-combatants receive 13 years of training is that most of this population is between 26 and 40 years old. A cost-benefit analysis makes this possibility unfeasible, even for those who are now 26 years old because at the end of the training cycle they would be around 40 years old, which would limit their possibilities in the labor market at that time. Therefore, the study finds it feasible to design training tailored to both basic and higher education that is compatible with income generation. Another possible path is to carry out customized training both in economic sectors and in geographical terms.

The study also infers that the departments in which the demobilized ex-combatants are more likely to find an opportunity in the labor market are the departments of Antioquia, Bogotá, Caldas, Cundinamarca, Quindío, Risaralda, Valle del Cauca, and Santander. This is combined with market potential and GDP.

Finally, to attend to the reinserted population, the study concludes that another way to explore is tending programs that favor the creation of companies (as enterprises) that allow this population to generate income.

Contributions of the Thesis

Chapter I:

Seminar presentations:

1. "Income disparities across Colombian departamentos: A Geographical Economics explanation" Banco de la Republica de Colombia (Banrep), Medellín (Colombia) 2019.

2. "Income disparities across Colombian departamentos: A Geographical Economics explanation" Universidad de Antioquia, Medellín (Colombia) 2019.

Keynotes

1. "The role of Market Potential in Geographical Economics: Some empirical evidence" Guest lecture, Vest University of Timisoara (Romania) 2019

2. "The role of Market Potential in Geographical Economics: Some empirical evidence" Guest lecture, Universitate Babes-Bolyai (Romania) Fecha: 2019

Conference presentations:

1. "Market potential and income disparities across Colombian regions" XX Jornadas de Economía Internacional (AEEFI), Granada (Spain), 2019.

2. "Economía geográfica y disparidades en los niveles de renta en los departamentos colombianos: Análisis para el periodo 1990-2015", XX Reunión de economía mundial (REM), Almería (Spain, 2018).

3. "Geographical economics and income Disparities across Colombian Departamentos: Analysis for the period 1990-2015", 9th Global Business Conference, Dubrovnik (Croatia), 2018.

4. "Geographical economics and income Disparities across Colombian Departamentos: Analysis for the period 1990-2015", 15th EU-Real workshop, Palermo (Italia), 2018.

Chapter II:

Conference presentations:

1. The Latin American and Caribbean Economic Association or Asociación de Economía de América Latina y el Caribe (LACEA), Santa Cruz (Bolivia), 2018.

Seminar presentations:

1. Seminario de Economía Banco de la República. Medellín (Colombia), 2019.

Chapter III:

1. Published paper: Revista Espacios Vol. 41 año 2020

Seminar presentations:

1. Seminario de Economía No. 21 Banco de la República. Medellín (Colombia), 2018.

2. Foro Política de Reintegración Social y Económica de Desmovilizados en Colombia. Universidad de Antioquia, Medellín (Colombia), 2019.