

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE ECONOMÍA Y RELACIONES INTERNACIONALES

PROGRAMA DE DOCTORADO EN CIENCIAS ECONÓMICAS



**“EL FENÓMENO DE LA INFORMALIDAD LABORAL EN MÉXICO:
UNA PERSPECTIVA MACROECONÓMICA DUAL E INSTITUCIONAL”**

**TESIS QUE COMO REQUISITO PARA OBTENER EL GRADO DE
DOCTORA EN CIENCIAS ECONÓMICAS**

PRESENTA:

MAYRA YESENIA NAVA RUBIO

DIRECTOR DE TESIS:

DR. ROGELIO VARELA LLAMAS

Tijuana, Baja California, diciembre de 2016.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE ECONOMÍA Y RELACIONES INTERNACIONALES

PROGRAMA DE DOCTORADO EN CIENCIAS ECONÓMICAS



**“EL FENÓMENO DE LA INFORMALIDAD LABORAL EN MÉXICO:
UNA PERSPECTIVA MACROECONÓMICA DUAL E INSTITUCIONAL”**

**TESIS QUE COMO REQUISITO PARA OBTENER EL GRADO DE
DOCTORA EN CIENCIAS ECONÓMICAS**

PRESENTA:

MAYRA YESENIA NAVA RUBIO

COMITÉ TUTORIAL:

DR. ROGELIO VARELA LLAMAS

DR. JUAN MANUEL OCEGUEDA HERNÁNDEZ

DR. RAMÓN AMADEO CASTILLO PONCE

DR. SALVADOR GONZÁLEZ ANDRADE

DR. ROBERTO IVÁN FUENTES CONTRERAS

Tijuana, Baja California, diciembre de 2016.

DEDICATORIA

*A Dios, por darme la fuerza para seguir superándome cada día y por bendecirme
concediéndome a una excelente familia.*

*A mis padres, por ser mi guía en todo momento y por alentarme a concluir esta etapa,
entregándome lo mejor de cada uno y su amor infinito.*

*A mis hermanos, Jesús y Pedro Agustín, quienes han sido mis compañeros de vida y me han
hecho sentir especial al demostrarme todo su cariño, apoyo y respeto.*

AGRADECIMIENTOS

La culminación de este trabajo de tesis, representa el esfuerzo conjunto de algunas instituciones y personas a quienes dedico este espacio para expresarles mis más sinceros agradecimientos por todo el apoyo otorgado de distintas formas durante el proceso de mis estudios en el programa de Doctorado en Ciencias Económicas.

En primera instancia debo agradecer a la Universidad Autónoma de Baja California (UABC) por el esfuerzo que realiza para ofrecer programas de posgrado de gran calidad educativa como los que se encuentran en la Facultad de Economía y Relaciones Internacionales (FEYRI), a la cual agradezco a su vez la oportunidad concedida para continuar con mi formación académica a este nivel.

También debo reconocer la importancia del apoyo económico ofrecido por parte del Consejo Nacional de Ciencia y Tecnología (CONACYT) para la realización de mis estudios y el desarrollo de la presente investigación.

Por otra parte, deseo agradecer infinitamente al Dr. Rogelio Varela Llamas por todo su apoyo como director de esta tesis, por sus enseñanzas que son de gran valor para mí, por su dedicación, por su paciencia, por su esfuerzo, por sus comentarios, por sus observaciones, por su disponibilidad, por demostrar interés en esta investigación, por ser un gran director de tesis y una excelente persona a quien tuve el privilegio de conocer desde mis estudios de licenciatura. Por todo esto y más, gracias.

También agradezco al Dr. Juan Manuel Ocegueda Hernández por sus valiosas aportaciones en cada coloquio y por sus pertinentes observaciones y recomendaciones que hicieron de este un mejor trabajo. Asimismo expreso mi profundo agradecimiento al Dr. Ramón Amadeo Castillo Ponce por su gran compromiso, ayuda e importantes y valiosas contribuciones para la realización de esta investigación. A ambos, gracias por todo.

Por último, agradezco a todos aquellos profesores que no menciono pero que contribuyeron con sus enseñanzas a mi formación académica que hoy se ve realizada en un trabajo de investigación doctoral. Asimismo agradezco al personal administrativo por su apoyo.

Gracias!!!

ÍNDICE

	Pág.
CAPÍTULO I: PROTOCOLO DE LA TESIS	1
I.1. Introducción	1
I.2. Antecedentes	5
I.3. Relevancia del Estudio de la Informalidad	8
I.4. La Informalidad Laboral en México: Un Análisis Preliminar	11
I.5. Análisis Empírico de la Informalidad: ¿Qué indican los hechos sobre el fenómeno?.....	14
I.6. Metodología de Análisis	23
CAPÍTULO II: DISCUSIÓN TEÓRICA. ENFOQUES QUE EXPLICAN LA INFORMALIDAD LABORAL	25
II.1. El Enfoque Dual de la Informalidad Laboral	25
II.1.1. El Enfoque Estructuralista de la Informalidad Laboral	31
II.2. El Enfoque Institucionalista de la Informalidad Laboral.....	33
CAPÍTULO III: METODOLOGÍAS PARA EL ANÁLISIS DE LA INFORMALIDAD LABORAL. MODELOS CON DATOS EN PANEL.....	42
III.1. Modelo de Regresión Agrupada.....	44
III.2. Modelo de Efectos Fijos.....	45
III.2.1. Efectos Fijos con Variables Dicotómicas.....	46
III.2.2. Efectos Fijos en Primeras Diferencias	47
III.3. Modelo de Efectos Aleatorios	48
III.4. Modelo con Variables Instrumentales.....	50

III.5. Elección del Modelo Apropriado con Datos en Panel.....	51
CAPÍTULO IV: FUENTES DE INFORMACIÓN Y ESPECIFICACIÓN DEL MODELO DE INFORMALIDAD LABORAL	59
IV.1. Fuentes de Información.....	59
IV.2. Especificación del Modelo con Datos en Panel	60
IV.2.1 Descripción y Justificación de las Variables incluidas en las Especificaciones del Modelo	61
IV.3. Análisis Gráfico de las Variables Incluidas en el Modelo	70
IV.3.1. Análisis Gráfico de las Variables Duales (D_{it}).....	71
IV.3.1.1. PIB	71
IV.3.1.2. Escolaridad.....	71
IV.3.1.3. Desempleo.....	72
IV.3.1.4. Sector Moderno.....	73
IV.3.2. Análisis Gráfico de las Variables Institucionales (I_{it}).....	73
IV.3.2.1. Costo Laboral Unitario.....	73
IV.3.2.2. Eficiencia en Recaudación	74
IV.3.2.3. Inversión Pública.....	75
IV. 3.2.4. Cobertura del Seguro Popular	76
IV.3.3. Análisis Gráfico de las Variables Instrumentales	77
IV. 3.3.1. Número de Escuelas.....	77
IV. 3.3.2. Tasa de Deserción en Secundaria.....	78
IV.3.4. Análisis Gráfico de las Variables Dependientes TILI Y TILD	78

IV.3.4.1. Tasa de Informalidad Laboral Independiente	78
IV.3.4.2. Tasa de Informalidad Laboral Dependiente.....	79
CAPÍTULO V: PRESENTACIÓN, ANÁLISIS E INTERPRETACIÓN DE RESULTADOS ...	81
V.1. Regresión Agrupada vs Efectos Fijos vs Efectos Aleatorios	81
V.2. Variables Instrumentales	90
V.3. El “Mejor” Modelo.....	93
V.4. Análisis e Interpretación de Resultados	96
CAPÍTULO VI: CONCLUSIONES	104
REFERENCIAS	113
ANEXOS.....	122
A. Clasificación de la Ocupación Informal en México.	122
B. Resultados de las Estimaciones	122
B.1. Estimaciones con TIL.....	122
B.1.1. Resultados y Pruebas de la Regresión con todas las Variables Explicativas	122
B.1.2. Resultados y Pruebas de la Regresión sin IP	124
B.1.3. Resultados y Pruebas de la Regresión sin IP y PIB.....	126
B.1.4. Resultados y Pruebas de la Regresión sin IP y NDES. El “Mejor” Modelo	128
B.1.5. Pruebas de Endogeneidad y Exogeneidad	131
B.1.6. Efectos Marginales	133
B.1.7. Primeras Diferencias.....	133
B.1.8. Pruebas de Raíz Unitaria	134

B.2. Estimaciones con TILI.....	134
B.2.1. Resultados y Pruebas de la Regresión sin IP y NDES. El “Mejor” Modelo	134
B.3. Estimaciones con TILD.....	136
B.3.1. Resultados y Pruebas de la Regresión sin IP y NDES. El “Mejor” Modelo	136
C. Correlaciones	138

ÍNDICE DE GRÁFICAS

	Pág.
Gráfica 1: Tasa de Informalidad Laboral (TIL)	12
Gráfica 2: Tasa de Informalidad Laboral (TIL) por Género	12
Gráfica 3: Tasa de Informalidad Laboral (TIL). Tres entidades con las mayores tasas y tres entidades con las menores tasas.	13
Gráfica 4: Mercado Dual de Trabajo.....	28
Gráfica 5: Producto Interno Bruto (PIB) y Tasa de Informalidad Laboral (TIL)	71
Gráfica 6: Escolaridad de la PEA (ESCO) y Tasa de Informalidad Laboral (TIL)	72
Gráfica 7: Número de Desempleados (NDES) y Tasa de Informalidad Laboral (TIL)	72
Gráfica 8: Sector Moderno (SMODERNO) y Tasa de Informalidad Laboral (TIL)	73
Gráfica 9: Costo Laboral Unitario (CLU) y Tasa de Informalidad Laboral (TIL)	74
Gráfica 10: Eficiencia en Recaudación (ER) y Tasa de Informalidad Laboral (TIL).....	75
Gráfica 11: Inversión Pública (IP) y Tasa de Informalidad Laboral (TIL).....	76
Gráfica 12: Cobertura del Seguro Popular (COBERTURA) y Tasa de Informalidad Laboral (TIL)	76
Gráfica 13: Número de Escuelas (NESCUELAS) y Escolaridad de la PEA (ESCO)	77
Gráfica 14: Tasa de Deserción en Secundaria (TDS) y Escolaridad de la PEA (ESCO)	78
Gráfica 15: Tasa de Informalidad Laboral Independiente (TILI)	79
Gráfica 16: Tasa de Informalidad Laboral Dependiente (TILD)	79

ÍNDICE DE TABLAS

	Pág.
Tabla 1: Especificaciones del Modelo de Informalidad Laboral	60
Tabla 2: Variables Explicativas Duales (D_{it}) e Institucionales (I_{it}).....	69
Tabla 3: Resultados de la Regresión Agrupada	83
Tabla 4: Resultados de la Regresión con Efectos Fijos	84
Tabla 5: Resultados de la Regresión con Efectos Aleatorios.....	85
Tabla 6: Prueba F. Regresión Agrupada vs Efectos Fijos.....	85
Tabla 7: Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios	86
Tabla 8: Prueba de Sobreidentificación. Efectos Fijos vs Efectos Aleatorios.	87
Tabla 9: Resultados de la Regresión con Efectos Fijos. Sin IP.....	88
Tabla 10: Resultados de la Regresión con Efectos Fijos. Sin IP y PIB	89
Tabla 11: Resultados de la Regresión con Efectos Fijos. Sin IP y NDES	90
Tabla 12: Matriz de Correlaciones	90
Tabla 13: Resultados Prueba de Endogeneidad No. 1	91
Tabla 14: Resultados Prueba de Endogeneidad No. 2	92
Tabla 15: Resultados Prueba de Exogeneidad	92
Tabla 16: Efectos Marginales.....	94
Tabla 17: Resultados de la Regresión en Primeras Diferencias.....	95
Tabla 18: Resultados de las Pruebas de Raíz Unitaria para los Residuos del “Mejor” Modelo ...	96
Tabla 19: Elasticidades. No.1.....	97

Tabla 20: Elasticidades. No.2..... 101

Tabla 21: Elasticidades. No.3..... 102

RESUMEN

El fenómeno de la informalidad laboral en México se analiza desde una perspectiva macroeconómica, considerando dos enfoques teóricos que son opuestos: el dual y el institucional. La principal diferencia entre ambas explicaciones estriba en que el primero le atribuye un carácter involuntario a la informalidad, mientras que el segundo se asocia a un proceso de decisión que es voluntario. El interés por abordar el objeto de estudio surge de observar elevados niveles de informalidad en el mercado laboral a escala nacional y por entidad federativa. En particular interesa determinar qué factores influyen en la explicación de la informalidad laboral y en qué magnitud lo hacen. Se estima un modelo econométrico con datos en panel de efectos fijos para el periodo 2005-2014, para el total de la informalidad y para aquella que se considera independiente y dependiente. Se encuentra que las variables duales ejercen una mayor influencia sobre los niveles de informalidad laboral, en particular, resalta el papel preponderante de la educación formal y por el lado del enfoque institucional la cobertura del seguro popular.

Palabras clave: informalidad laboral, enfoque dual, enfoque institucional, panel de datos.

**El Fenómeno de la Informalidad Laboral en México:
Una Perspectiva Macroeconómica Dual e Institucional**

CAPÍTULO I: PROTOCOLO DE LA TESIS

I.1. Introducción

La dinámica del empleo alrededor del mundo es de particular interés desde el punto de vista del análisis económico, debido a que representa la principal fuente de ingresos para los trabajadores y sus familias. Una característica del empleo es que puede ser abordado considerando distintos aspectos de acuerdo a su tipología. Por ello, las investigaciones que existen en torno a este eje temático resultan importantes, ya que pueden aportar un diagnóstico más completo de sus determinantes por área o región geográfica.

El presente trabajo se enfoca específicamente en el estudio del fenómeno de la informalidad laboral. En el plano teórico, son dos los paradigmas que proveen una explicación respecto al fenómeno; la visión tradicional y la institucionalista. La primera está integrada por los enfoques asociados a la teoría dualista y a la estructuralista, donde ésta última se caracteriza por tener sus bases en la primera. En general, el enfoque tradicional le confiere la existencia de la informalidad a la economía y a la estructura del mercado de trabajo, señalando que tiene un carácter involuntario. Por su parte, el enfoque institucionalista señala que la causa de la informalidad son las instituciones del Estado, mismas que a su vez pueden influir a favor de la elección de un empleo informal por parte del individuo, por lo cual le concede un carácter voluntario a la informalidad. Bajo este paradigma se considera que los individuos son quienes deciden ser informales escapando al cumplimiento de las obligaciones que implica el sector formal tales como el pago de impuestos, o en el caso de los empleadores la provisión de prestaciones laborales que implican un costo para ellos.

En el plano estadístico (para conformar la base de datos) la informalidad puede ser desglosada a grandes rasgos en dos partes. La primera considerando la perspectiva metodológica de medición de la unidad económica y la segunda referente a las condiciones de los individuos en sus empleos. La informalidad laboral que es de interés y que se estudia en este trabajo pertenece a

ambas perspectivas y es analizada de manera conjunta y separada. La primera perspectiva corresponde al tipo de informalidad que capta a los individuos que se encuentran al mando de unidades económicas que poseen la característica de ser clasificadas como informales, es decir, tanto trabajadores por cuenta propia como empleadores y aquellos que incurren en la agricultura de subsistencia. La segunda permite analizar a los trabajadores subordinados remunerados y no remunerados que se desempeñan en condiciones de informalidad.

El desarrollo de esta investigación tiene como objetivo general contrastar empíricamente los enfoques teóricos dual e institucional, y a partir de ello encontrar una explicación más razonable del fenómeno de la informalidad en México. Para cumplir con el objetivo establecido, se plantean dos objetivos específicos: uno es estimar tres versiones de un modelo econométrico basado en datos en panel que es conformado por variables duales e institucionales y otro es realizar la selección del mejor modelo para cada versión, mediante el uso de distintas metodologías y la aplicación de algunas pruebas econométricas formales.

Las tres versiones del modelo giran en torno a la perspectiva dualista e institucionalista buscando confrontar ambas posturas. Por lo cual, se analizan los factores que determinan la informalidad laboral en México durante el periodo anual que va del año 2005 al año 2014, realizando una comparación entre estos. La información utilizada corresponde a 24¹ de las 32 entidades que conforman la geografía nacional. En el marco del enfoque dualista se identifican tres factores explicativos; el Producto Interno Bruto (PIB) que es utilizado para capturar el impacto de la situación de la economía del país sobre el aumento o disminución de la informalidad laboral; también se contempla la educación de la población como una variable proxy de capital humano bajo el supuesto de que al incrementarse puede influir en la reducción de los niveles de informalidad debido a que los individuos con mayores habilidades son requeridos en el lado formal de la economía. Para la construcción de esta variable, se consideran los años promedio de escolaridad de la Población Económicamente Activa (PEA). Adicionalmente, se incluye un tercer factor que es la proporción que representa el empleo en el sector manufacturero en el empleo total. Esta variable se estima que es una aproximación a la población que labora en el sector moderno de la economía.

¹ La principal razón para considerar solamente 24 entidades, es la falta de información. Las entidades excluidas del análisis son: Baja California, Baja California Sur, Durango, Guerrero, Oaxaca, Querétaro, Tlaxcala y Veracruz. En el capítulo cuatro se provee una justificación más detallada al respecto.

En lo que respecta a los factores relacionados con el enfoque institucionalista, son tres las variables que los representan; la primera se relaciona con el costo laboral en el sector moderno de la economía (sector manufacturero) que es calculado mediante la fórmula del costo laboral unitario. La segunda variable se encuentra asociada a la eficiencia del gobierno, específicamente la eficiencia en materia de recaudación de impuestos, medida como el cociente de los egresos en remuneraciones al personal de la Secretaría de Administración y Finanzas entre los ingresos obtenidos vía impuestos. La tercera variable es la cobertura del seguro popular, que busca evidenciar la elección del individuo por pertenecer a la informalidad basándose en la posible sustitución de la atención médica que proporciona el mercado formal de trabajo.

La información proviene de diversas fuentes oficiales como el Instituto Nacional de Estadística y Geografía (INEGI), el Banco de Información Económica (BIE), la Encuesta Mensual de la Industria Manufacturera (EMIM), la Encuesta Industrial Mensual (EIM), la Encuesta Nacional de Ocupación y Empleo (ENOE), de los registros administrativos de las Finanzas Públicas Estatales y Municipales, del Tercer Informe de Gobierno 2014-2015, del Consejo Nacional de la Población (CONAPO), la Secretaría de Educación Pública (SEP) y del Instituto Nacional para la Evaluación de la Educación (INEE).

Sin embargo, la diferencia entre las versiones radica en que en el primer ejercicio se estima el modelo tomando como variable dependiente el total de la informalidad laboral y como variables explicativas los factores ya mencionados para cada uno de los enfoques teóricos. En el segundo ejercicio la variable dependiente sólo considera el trabajo informal independiente o la informalidad desde la perspectiva metodológica de la unidad económica. En el tercer ejercicio, la variable dependiente captura la informalidad desde la perspectiva de las condiciones laborales, o bien, el trabajo informal dependiente. Cabe mencionar que la variable a explicar de los dos últimos ejercicios es tomada en el marco de la metodología utilizada por el Instituto Nacional de Estadística y Geografía (INEGI) en colaboración con la Organización Internacional del Trabajo (OIT), para la medición del fenómeno informal en el país a partir de las bases de datos de la ENOE (véase sección A en los anexos).

Las razones que justifican la estimación del modelo con datos en panel desagregando la informalidad en independiente y dependiente, son teóricas y empíricas por lo cual se destaca la importancia de la realización de un análisis de este tipo. A nivel teórico, se ha argumentado que

existe una gran diferencia dentro de la informalidad y que ésta se compone tanto por individuos que se encuentran involuntariamente dentro del sector informal como por individuos que se encuentran de manera voluntaria, lo que da origen a las teorías que se confrontan en esta investigación. La justificación empírica descansa en que los individuos informales evidentemente no son iguales, y que las causas que inciden en que estos pertenezcan a uno u otro tipo de informalidad pueden ser distintas e influir en diferente medida. En otras palabras, la informalidad es un fenómeno mixto que merece ser estudiado de manera desglosada de acuerdo a su tipología.

Las preguntas de investigación que se derivan del objeto de estudio son las siguientes, ¿qué enfoque teórico explica el fenómeno de la informalidad laboral para el caso de México? y ¿qué acciones de política pública pueden emprenderse para disminuir la informalidad observada?. La hipótesis vinculada a estas preguntas, es que el fenómeno de la informalidad laboral en México, puede ser explicado considerando ambos enfoques teóricos (dualista e institucionalista) debido a que una parte de la informalidad tiene su origen en una debilidad estructural de la economía, pero también otra parte se podría explicar por factores de orden institucional. Esto debe cumplirse si se analiza la informalidad total o de manera desglosada en independiente y dependiente. Sin embargo, se destaca que son los factores de orden dual los que ejercen mayor influencia en los niveles de informalidad observados en el país debido a que aun cuando los individuos pudieran desear pertenecer al lado formal de la economía no pueden hacerlo porque las condiciones son ajenas a estos, en cambio por el lado institucional los individuos pueden optar o no por ser informales.

De manera puntual, la presencia de la informalidad laboral en México se debe a la capacidad limitada del sector formal para emplear a todos los individuos que buscan un empleo o en términos generales a que la situación de la economía no siempre es favorable. En este sentido, si el reto es abatir las elevadas cifras de informalidad en el país, entonces las acciones de política deben enfrentar el problema considerando principalmente incidir en variables de orden dual a nivel agregado, tales como el Producto Interno Bruto (PIB), la capacidad de empleo en el sector formal y la educación. Por ejemplo, se deben emprender aquellas políticas que dinamicen la economía y que como consecuencia incrementen la capacidad de empleo en el sector formal, así como aquellas que fomenten la educación de la población en general y la formación de los individuos para el trabajo, mejorando así la empleabilidad en la formalidad.

La principal razón que justifica el análisis de la informalidad laboral en México, es que refleja un desequilibrio del mercado laboral debido a que algo ocurre al interior y fuera del mismo, que provoca que algunos individuos se empleen voluntaria o involuntariamente en la informalidad, teniendo como consecuencia que la calidad del empleo disminuya para esa porción de la población. La informalidad se puede caracterizar por condiciones laborales no apropiadas para el trabajador como es la falta de protección laboral y la falta de acceso a la atención médica. Además es un sector relacionado con empleos de baja productividad y con personal ocupado que posee bajos niveles de escolaridad. De aquí la relevancia de su análisis.

El documento se encuentra estructurado en seis capítulos; el primero corresponde al protocolo donde se destaca el tema de interés, el objetivo, la hipótesis, y la relevancia del tema de investigación. Se describen los antecedentes de la informalidad laboral en México y se analiza su evolución a través de datos estadísticos que ayudan a dimensionar el problema de estudio; asimismo se presentan algunos resultados empíricos en torno a la temática de interés y se comenta brevemente la metodología y las variables utilizadas. En el segundo capítulo se procede con la revisión de los enfoques teóricos que facilitan la explicación de la informalidad desde las dos grandes visiones (*dual e institucional*) y se introduce a un enfoque teórico mixto. Las diversas metodologías de análisis se presentan de manera detallada en el tercer capítulo, así como las distintas pruebas formales que son de utilidad para elegir el mejor modelo que permita obtener resultados confiables. En el capítulo cuarto, se describen las fuentes de información, se especifican los modelos a estimar, se definen las variables involucradas y se realiza un análisis gráfico de éstas. En el quinto capítulo se presentan, analizan e interpretan los resultados derivados de las estimaciones de los modelos de panel de datos. Las conclusiones se reportan en el capítulo seis. Y al final del documento se ubican las referencias y los anexos.

I.2. Antecedentes

El concepto de informalidad ha ido evolucionando al paso del tiempo y han surgido nuevas metodologías para su medición y definición. En retrospectiva, los orígenes sobre el estudio de la informalidad se remontan a los trabajos del antropólogo Keith Hart, quien estudió el caso de Ghana a principios de los años setenta, acuñando el término de sector informal. En dicho estudio, Hart se refiere al sector informal como “un conjunto de actividades económicas que se encuentran fuera de la fuerza laboral organizada” y que son objeto de una examinación detallada

en su trabajo. En este contexto histórico planteó la idea de que el empleo informal actúa como un amortiguador para quienes se encuentran sin empleo (Hart, 1973).

La historia institucional del sector informal en el ámbito de la Organización Internacional del Trabajo (OIT) la retrata Bangasser (2000) en tres fases, la primera corresponde a la década de 1970, a la cual el autor denomina los años de incubación y que consiste en el desarrollo y arraigo del concepto del sector informal sobre todo a partir de la misión de Kenya en 1972. La segunda fase se desarrolla en la década de 1980, años en donde el concepto se propaga entre diferentes actores, quienes inician estudios sobre dicho sector y diversos programas. Finalmente se encuentran los años de la oficialización que se circunscriben a la década de 1990 en donde el concepto adquiere reconocimiento internacional.

La OIT hace uso del término de economía informal para incluir al sector informal así como al empleo informal; de este último menciona que puede presentar características como la ausencia de protección. Por ejemplo, cuando no se le retribuye al individuo por su trabajo, cuando se le obliga a hacer sobretiempo o trabajar turnos extraordinarios, cuando es despedido sin aviso ni compensación, o labora en condiciones inseguras y sin beneficios como las pensiones, el reposo por enfermedad o el seguro de salud.

Debido a que el término definido por la OIT comprende dos grandes aspectos de la informalidad (el sector informal en el sentido de las unidades económicas y el empleo informal desde las condiciones laborales) es posible decir que es el concepto que mejor guía la comprensión de la informalidad en el mercado de trabajo y que captura el fenómeno en un amplio sentido al considerar todo tipo de informalidad existente.

La economía informal se define como el conjunto de actividades económicas que realizan los trabajadores y las empresas que, tanto en la legislación como en la práctica, están insuficientemente contempladas por los sistemas formales o en el extremo no lo están. Dichas actividades no están amparadas por la ley, lo que significa que se desempeñan al margen de ella; o no están contempladas en la práctica. Si bien estos individuos y empresas operan dentro del ámbito de la ley, ésta no se aplica o no se cumple; o la misma no fomenta su cumplimiento al ser inadecuada, compleja o al imponer costos excesivos (OIT, 2012). Para Webb, Tihanyi, Ireland y Sirmon (2009), la economía informal se conforma por actividades que reconocen y aprovechan

las oportunidades generadas fuera de los límites institucionales formales, pero que se encuentran dentro de los límites institucionales informales.

En un trabajo de la OIT (2013) se sostiene que el sector informal se compone de pequeñas empresas o de empresas por cuenta propia, con poca o ninguna organización formal o capital, y con empleo informal. Donde además dicho sector posee la característica de ser una de las fuentes principales de empleo en muchos países del mundo. Respecto al empleo informal, éste incluye a las personas ocupadas que, por ley o en la práctica, no se encuentran sometidas a la legislación laboral nacional, al impuesto sobre la renta o que no tienen derecho a la protección social ni a los beneficios laborales. Asimismo, el empleo informal puede presentarse tanto en el sector formal como en el sector informal de la economía².

Desde un enfoque de medición se pueden mencionar básicamente dos perspectivas de la informalidad. La primera aborda el fenómeno desde el punto de vista de la unidad económica y se conoce como enfoque de sector informal. Se centra en las características de las unidades económicas no agropecuarias que no están constituidas como empresas y que no cumplen con los registros mínimos que la legislación demanda de proveedores de bienes y servicios. La segunda perspectiva se concentra en analizar las condiciones laborales de los individuos, es decir, las formas de trabajo que carecen de garantías y acceso a la seguridad social (INEGI, 2014). Para esta investigación ambas perspectivas son consideradas, pues el objetivo es explicar el fenómeno bajo un amplio espectro que no subestime los niveles de informalidad observados en el país.

Otros estudios sostienen que la informalidad es un trabajo que se realiza sin contar con el amparo del marco legal o institucional, no importando cuál sea el tipo de unidad económica (Negrete y Eternod, 2011). Algunos otros subrayan que el término puede tener significado diferente para distintas personas pero la característica común es que significa cosas “malas” tales como:

² La informalidad no es exclusiva del sector informal, también se presenta en el sector formal de la economía. Un ejemplo de ello es cuando los individuos desempeñan sus labores como empleados pero sin contar con un contrato formal que defina la relación patrón-obrero. En otras palabras, algunos individuos trabajan bajo condiciones de informalidad en el sector formal. Tal es la importancia de este hecho que en las cifras de la ENOE se reporta la ocupación informal para México separándola en: ocupación informal en el sector informal y ocupación informal fuera del sector informal (que puede asumirse que se refiere al sector formal). Bajo esta clasificación, en promedio para el año 2014 la informalidad en el sector informal representó el 47% del total de la informalidad, en tanto que la informalidad fuera del sector informal fue el restante 53%. Es interesante como para México la informalidad fuera del sector informal posee una gran representatividad del total, es decir, es mayor el número de individuos informales fuera del sector informal que dentro de éste. Por lo cual la situación se convierte en un fenómeno de tal relevancia que merece ser analizado.

trabajadores sin protección, regulación excesiva, baja productividad, competencia desleal, evasión de la autoridad de la ley, pago incompleto o falta de pago de impuestos y un trabajo desde las sombras (Maloney y Saavedra, 2008).

Por su parte Guha-Khasnobis, Kanbur y Ostrom (2006) parten de asumir que no existe consenso en la definición de lo que es formal e informal, sin embargo resaltan dos aspectos, el primero es la noción de lo informal como fuera del alcance de los diferentes niveles y mecanismos del gobierno oficial, como contraparte lo formal es alcanzable por estos mecanismos. El segundo aspecto se refiere a que lo informal se identifica a menudo con la falta de estructura y lo formal con lo estructurado, por lo cual el término sector no organizado también puede ser utilizado para hacer referencia al sector informal.

I.3. Relevancia del Estudio de la Informalidad

La informalidad genera efectos positivos y negativos sobre la economía en su conjunto y es precisamente esta característica la que permite que su estudio haya adquirido recientemente gran relevancia alrededor del mundo, pero sobre todo en las economías en desarrollo que es donde el fenómeno alcanza una mayor dimensión. Entre los efectos positivos se puede mencionar principalmente que es una forma de ocupación que permite abatir los niveles de desempleo de un país, propiciando que los hogares accedan a un ingreso sin el cual su bienestar se vería totalmente reducido. Es por esta razón que se convierte en una salida a la falta de empleo formal o a la incapacidad del sector formal para generar los suficientes empleos que demanda la sociedad en edad de trabajar.

Sin embargo a pesar de los aspectos positivos que puede generar, también es motivo de interés para algunos autores concentrarse en los efectos negativos que pueden impactar directamente sobre el bienestar de los individuos a través de las condiciones laborales, el impacto sobre la estructura del mercado de trabajo y en general, el efecto sobre la economía. Esta preocupación surge de la tendencia creciente en los distintos países.

Por el lado de los efectos negativos Aguilera y Velázquez (1994), argumentan que puede traer consigo implicaciones negativas para el crecimiento económico de un país en el mediano y largo plazo por ejemplo, al impactar sobre la productividad debido a que el sector informal es menos productivo que el formal. Sostienen que la informalidad puede afectar a la economía por medio

del sistema impositivo ya que al ser reducido el número de contribuyentes cautivos, estos deben de alguna manera absorber el costo impositivo convirtiéndose en un círculo vicioso, en donde unos pagan impuestos y otros los evaden.

A partir de lo anterior, es posible decir que el hecho de que algunos individuos evadan impuestos (como es el caso de los informales) tiene una penalización en cierto sentido sobre aquellos que si cumplen con su pago de impuestos, para éstos tal situación puede representar una falta de aplicación de la ley tributaria por igual (es decir una debilidad institucional) y convertirse en un incentivo para evadir impuestos o bien para optar por la informalidad y con ello reducir los costos que implica la formalidad. Por ello Aguilera y Velázquez (1994) enfatizan en el círculo vicioso que se genera a partir de la aplicación del sistema impositivo.

En el contexto de un modelo de equilibrio general, Leal (2014) menciona que existen tres distorsiones inducidas por la informalidad. La primera se refiere a distorsiones en la demanda óptima de trabajo de las empresas informales, esto es, para las empresas existe un incentivo a permanecer pequeñas con el fin de seguir evadiendo la detección de su actividad informal. Por lo cual se puede decir que es posible asociar la informalidad a empresas pequeñas y a su baja productividad, que se traduce en bajo crecimiento de las mismas. La segunda distorsión que menciona el autor se refleja en las elecciones de ocupación debido a que para un grupo de individuos es más atractivo convertirse en empresarios en lugar de empleados. Por último se encuentra la mala asignación de recursos, es decir, a las empresas con baja productividad se les asignan más recursos de los que deberían. En general, Leal (2014) destaca que reducir la informalidad traería consigo ganancias importantes.

Otro efecto negativo es señalado en Elgin y Uras (2013) donde se plantea que surge de la relación entre el desarrollo financiero y el tamaño de la economía informal, destacando que éste último afecta al primero. El sector informal perjudica al desarrollo financiero a través del aumento de la represión financiera, debido a la evasión fiscal existente. Por esta razón, proponen que suprimir el tamaño del sector informal beneficia el nivel de desarrollo financiero en la economía.

Para Rodríguez (2007) la informalidad genera ciertos costos sociales, entre ellos el hecho de que los trabajadores y las compañías quedan atrapados en operaciones de baja productividad en las que los trabajadores no cuentan con prestaciones sociales y se evaden la mayoría de las

regulaciones. Sin embargo, propone que la informalidad puede generarse de una forma más rápida a causa de una regulación y fiscalización muy fuerte. Desde un punto de vista social la parte preocupante es que si un gran número de individuos opera en la informalidad, entonces una gran parte de la población no accede a beneficios sociales, tales como la atención médica. Este constituye un efecto negativo que repercute de manera directa sobre el bienestar de la población.

Centrándose en los individuos, la OCDE (2014) afirma que los trabajadores en el sector informal son sujeto de mayor inseguridad y además los empleadores en este sector dedican menos recursos a sus trabajadores. Por esta razón el organismo propone que el rendimiento de un mercado laboral no debe evaluarse sólo con base en el número de puestos de trabajo disponibles o en el número de personas que participan en la fuerza laboral. Es decir, para poder evaluar la situación real del mercado deben considerarse cuestiones más profundas, por ejemplo, las relacionadas con las condiciones de esos empleos (informalidad y subempleo).

Finalmente, lo que preocupa son las condiciones laborales de los individuos que realizan alguna actividad económica en el mercado de trabajo, sobre todo porque existe una relación entre baja calidad del empleo y la informalidad laboral, aunque debe hacerse la precisión de que pueden presentarse situaciones en las que los individuos encuentren mejores condiciones laborales en la informalidad, por ejemplo la flexibilidad en el horario de trabajo que permite combinar la crianza de los hijos y el cuidado del hogar con la actividad laboral (sobre todo para aquellos individuos que se encuentran al frente del hogar, en términos del cuidado de este). Asimismo es posible que los ingresos que percibe un individuo en un empleo informal sean mejores que los ingresos que puede recibir en un empleo formal, aunque cabe aclarar que es de esperarse que esta situación no se presente en la mayoría de los casos, ya que si así fuera la informalidad sería aún mayor, basándose solamente en esta comparación.

En este sentido surge el siguiente cuestionamiento, *¿por qué sigue siendo relevante el estudio de la informalidad laboral?* aun cuando es posible encontrar en ciertos trabajos empíricos que algunos individuos pueden estar mejor en un empleo informal que en uno formal. La posible respuesta, es que la informalidad es de interés y preocupación por sus características negativas que se asocian con la baja calidad del empleo. Algunos autores como Varela (2015), han logrado demostrar que no se está mejor en la informalidad que en la formalidad, al analizar las diferencias de ingreso entre trabajadores con empleo formal e informal en los estados de la Frontera Norte de

México. De esta manera preocupa que la informalidad laboral vaya en aumento debido a que se incrementan también los problemas de calidad del empleo de un país, lo cual se traduce en un problema de bienestar social y de baja productividad laboral.

En el trabajo de Rodríguez (2007) se resaltan los beneficios de reducir la informalidad, por ejemplo, contar con una mayor población con prestaciones sociales, desarrollo social del país por medio de los impuestos, el crecimiento económico, el fortalecimiento de los derechos de propiedad, la reducción de ciertos costos, acceso a infraestructura de servicios y créditos y tomar ventajas de la expansión de mercados. Por estas razones debe ponerse especial atención en las acciones de política tomadas para reducir la informalidad, ya que los beneficios son variados.

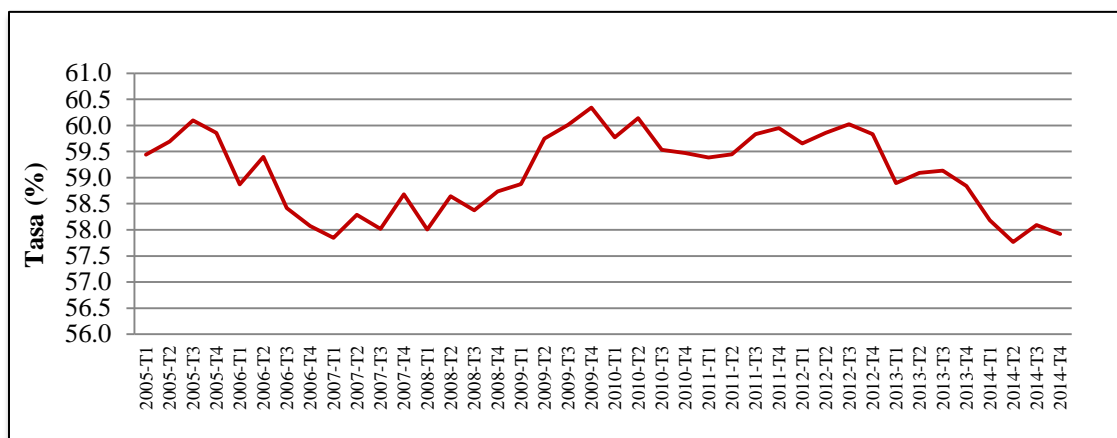
I.4. La Informalidad Laboral en México: Un Análisis Preliminar

En esta sección se describe el fenómeno de la informalidad capturado desde un enfoque estadístico con el objetivo de mostrar su comportamiento al paso del tiempo a escala nacional, lo que permite contextualizar dicho fenómeno en un escenario económico caracterizado por acontecimientos como la crisis de finales de 2008. Se muestran cifras y gráficas correspondientes al periodo que va del primer trimestre del año 2005 al cuarto trimestre del año 2014, mismo que coincide con el periodo de análisis para la muestra de entidades.

En la gráfica número 1 se muestra el comportamiento de la tasa de informalidad laboral total a nivel nacional, que se encuentra como TIL1 en la base de datos de la ENOE y hace referencia a la informalidad laboral considerando el sector agropecuario. Es una tasa calculada contra la población ocupada total. La evolución observada es preocupante debido a que para el periodo, en promedio el 58.8% de la población ocupada es informal, esto es, un elevado porcentaje de la población ocupada se encuentra laborando en un empleo de baja calidad en términos generales.

También se destaca el comportamiento ascendente de la tasa de informalidad a partir de 2007 alcanzando el nivel máximo en el cuarto trimestre de 2009 con un valor de 60.4% que puede asociarse con la difícil situación que se vivió a partir del comienzo de la crisis económica a finales de 2008. Resulta interesante la forma en la que después de este evento, la tendencia a la alza se revierte, sobre todo si se analiza el cambio observado del cuarto trimestre de 2012 al primer trimestre de 2013, al pasar la TIL de 59.8% a 58.9%. Al final, la TIL logra alcanzar niveles muy similares a los registrados en 2007 que son de los más bajos en el periodo de estudio.

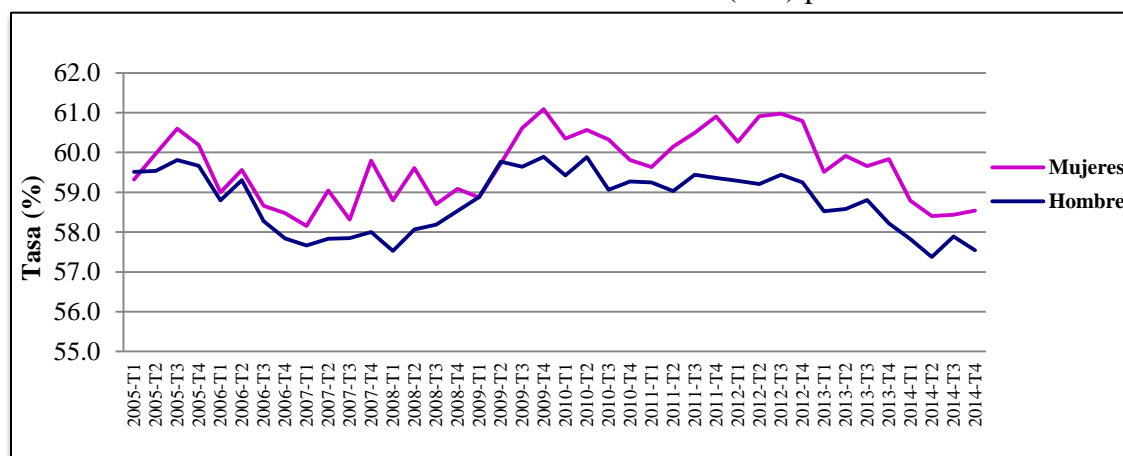
Gráfica 1: Tasa de Informalidad Laboral (TIL)



Fuente: Elaboración propia con datos de la ENOE.

En la gráfica número 2 se describe la TIL por género, la relevancia de analizar dichos datos se debe a la creciente participación de la mujer en el mercado laboral. El comportamiento de las series hace evidente que la informalidad es superior en las mujeres que en los hombres y el resultado coincide con el argumento de que posee ciertas características como la flexibilidad laboral que permite, en este caso a las mujeres combinar su actividad laboral con la crianza y cuidado de los hijos. Esto se comenta porque generalmente es la mujer quien se encarga del cuidado de los hijos en el hogar, aunque la participación de los hombres en este sentido es cada vez mayor. Por otro lado, el número de mujeres que son los jefes del hogar ha tendido a crecer. Debido a estas razones no es de extrañar que la tasa de informalidad de las mujeres a nivel nacional se encuentre por encima de la de los hombres en el periodo 2005 a 2014.

Gráfica 2: Tasa de Informalidad Laboral (TIL) por Género



Fuente: Elaboración propia con datos de la ENOE.

y Nuevo León su tasa de informalidad se encuentra por debajo del referente nacional lo que indica que los mercados laborales en estas economías son distintos a su contraparte al menos en términos de niveles de informalidad y por lo tanto, es de esperarse que las condiciones laborales en general sean mejores en dichas entidades. Las tres entidades con las tasas de informalidad laboral más elevadas poseen en promedio, el doble de población ocupada que es informal en comparación con las tres entidades con los menores valores de informalidad laboral. La heterogeneidad presente a nivel estatal, requiere de la utilización de un modelo econométrico que tenga en cuenta este hecho. El modelo seleccionado para llevar a cabo el estudio del fenómeno es uno con datos en panel, pues entre sus ventajas se encuentra que considera las diferencias entre los individuos de la muestra.

I.5. Análisis Empírico de la Informalidad: ¿Qué indican los hechos sobre el fenómeno?

En esta sección se describen los principales resultados empíricos de algunas investigaciones realizadas en torno a la temática de interés, éstas fueron llevadas a cabo por diversos autores, con distintos enfoques y metodologías para diversos países, entre ellos México. El interés por abordar los trabajos empíricos responde a la necesidad de enlazar las perspectivas teóricas del fenómeno de estudio con los hechos de la realidad económica-social.

Algunos trabajos como el de Tannuri-Pianto y Pianto (2002) consideran datos de la encuesta de hogares “Pesquisa Nacional por Amostra de Domicílios (PNAD)” para el año de 1999 para el caso de Brasil. Realizan un ejercicio en dos etapas estudiando a los individuos formales e informales en varios puntos de la distribución de ingresos (por cuantiles). En la primera etapa determinan la probabilidad de que una persona trabaje en el sector informal dado un conjunto de características. En la segunda etapa calculan una ecuación de salarios en la que incluyen características no observables de los individuos que afectan la decisión acerca del sector en el que trabajarán.

Los autores concluyen que las diferencias más relevantes de las funciones de ingresos son sus recompensas y penalizaciones por cuestiones como el analfabetismo, el género y la afiliación sindical. La penalización es muy elevada para los individuos analfabetas en el sector formal a través de todos los cuantiles de la distribución de ingresos, en contraste en el sector informal la penalización sólo es para los individuos que se encuentran en los cuantiles más altos. Los

individuos de ingresos bajos y medianos en el sector formal reciben altos retornos por un grado universitario en comparación con los individuos en el sector informal. Los trabajadores informales son quienes más se benefician de tener un grado de escolaridad de primaria y las mujeres son más penalizadas en los cuantiles más altos del sector formal que en los mismos cuantiles pero en el sector informal.

Asimismo, encuentran que ser miembro de un sindicato tiene efecto positivo y significativo sobre los ingresos de los formales a través de todos los cuantiles, en cambio sólo los trabajadores informales en los cuantiles más altos reciben un beneficio de la afiliación sindical. En cuanto a las características no observables, un resultado importante es que éstas impactan los ingresos de los individuos en el sector informal. Por todos estos resultados, los autores no pueden negar la hipótesis de segmentación del mercado laboral, lo que permite que su trabajo se encuadre bajo el enfoque dual que provee una de las explicaciones de la informalidad laboral.

El estudio llevado a cabo por Badaoui, Strobl y Walsh (2007), se centra en investigar la existencia de una penalización salarial al empleo informal tomando como unidad de análisis a los hombres no auto empleados, que trabajan en industrias no pertenecientes al sector público. Los datos pertenecen a la Encuesta de la Población Activa de África del Sur para los años 2001 a 2003; se concluye que cerca del 75% de la penalización salarial de los trabajadores informales se debe a diferencias en capital humano y a las características del empleo cuando se comparan los ingresos brutos. Cuando estas características son controladas, el logaritmo de los ingresos brutos continúa siendo 37% menor para los empleados en el sector informal. En un ejercicio adicional se controlan los factores no observables y la penalización salarial disminuye a 18%.

Finalmente encuentran que cuando se analizan sólo aquellos individuos para los que es posible calcular los retornos netos al empleo (después de impuestos) y asumiendo que los informales pueden evadir el pago de impuestos, las discrepancias desaparecen. Esta investigación se puede ubicar en los trabajos empíricos que intentan probar la validez del argumento de diferencia salarial en el mercado laboral correspondiente a la teoría dual de la informalidad.

En el estudio de La Porta y Shleifer (2008) se toman como referencia de análisis tres enfoques que relacionan la economía informal con el desarrollo económico. Estos representan la visión romántica, la perspectiva parásito y la visión dual. Para corroborar el cumplimiento de una u otra

visión se consideran datos para distintos países de tres encuestas del Banco Mundial, que son la Enterprise Survey, Informal survey y Micro survey, para los años 2003 a 2006. Los resultados obtenidos son acordes con la visión dual de la relación entre la informalidad y el desarrollo económico que sostiene que las empresas informales son altamente ineficientes, no presentan amenaza para las formales, y no contribuyen al crecimiento económico que es impulsado por las empresas formales eficientes. Entre los resultados resalta el hecho de que la elevada productividad proviene de las empresas formales y particularmente de las grandes. Asimismo, las empresas formales son dirigidas por gestores mejor educados, tienden a hacer mayor uso de capital, poseen diferentes clientes, comercializan sus productos y hacen uso de financiamiento externo en un mayor grado que las informales.

Marcouiller, Ruiz y Woodruff (1995) se preguntan qué tipo de personas trabajan en el sector informal y si los trabajadores informales perciben salarios más bajos que los trabajadores similares en el sector formal. La hipótesis que manejan es que quienes deciden trabajar pueden libremente elegir el sector en el que lo harán. Por lo tanto, proponen el análisis de datos, estimaciones de ecuaciones de salarios y de un modelo probit para El Salvador, Perú y México con datos de encuestas de hogares correspondientes a los años 1990, 1985 y 1990 respectivamente para cada país. Ello les permite encontrar ciertos patrones para los tres países, algunos de los cuales se enuncian a continuación:

- Los más jóvenes y los de mayor edad presentan mayor probabilidad de ser informales.
- Los trabajadores formales generalmente poseen mayor educación que los informales.
- El comercio al por menor es la actividad más grande de los trabajadores informales.
- En el sector formal e informal los retornos de la educación y de la experiencia son positivos. Este resultado es interesante porque no se ajusta a la concepción del sector informal como un sector secundario en la teoría dual.
- Para El Salvador y Perú existen primas salariales significativas asociadas con el trabajo en el sector formal. En cambio para México la prima se asocia con el trabajo informal.
- Para los hombres la probabilidad de empleo formal se incrementa cuando son jefes de hogar, con el matrimonio, y con el número de adultos inactivos en el hogar. Para las mujeres la probabilidad de empleo formal disminuye con el matrimonio y con el número de niños en el hogar, e incrementa con el número de adultos inactivos en el hogar.

Dado lo anterior Marcouiller et al. (1995) ponen en duda la idea de que el sector informal es siempre y en todas partes un refugio pobremente pagado pero de fácil entrada para quienes no tienen otras oportunidades de empleo. Principalmente dudan de ésta interpretación por los resultados de la prima salarial encontrada en el sector informal de México.

Hay otros trabajos que estudian la diferencia salarial entre el sector formal e informal donde se miden las ganancias y pérdidas en los salarios derivados de la formalización para el caso de México en el periodo 2000-2003 (Moreno, 2007). En dicho trabajo se reconoce que dos pueden ser las fuentes de la heterogeneidad; la primera es una tasa impositiva aplicada a un sector relativo al otro y la segunda es la autoselección de los agentes entre uno u otro sector. Las estimaciones realizadas permiten concluir que la fuente de esa diferencia salarial es la autoselección. Entre otros resultados interesantes, se encuentra evidencia a favor de que las personas con elevados niveles de educación obtienen una ganancia neta de cambiarse al sector formal, en cambio los trabajadores formales ganan menos si se cambian a la informalidad. Sin embargo para el autor, la evidencia también sugiere que las diferencias salariales entre los sectores se han reducido. Este tipo de análisis empírico se puede encuadrar dentro del enfoque dual.

Sin embargo, diversos autores señalan que basarse sólo en la comparación de los ingresos entre trabajadores formales e informales no prueba la segmentación en el mercado laboral o que los individuos se encuentren voluntaria o involuntariamente en la informalidad, pues aun cuando generalmente los ingresos de los individuos informales son menores a los ingresos de los formales existen otras razones por las cuales los individuos eligen la informalidad, por ejemplo la ya señalada flexibilidad laboral. En este sentido Packard (2007) propone que las diferencias salariales no logran captar el valor de las características no medibles de los diferentes tipos de trabajos, tales como las prestaciones no salariales, la compensación por el riesgo, la independencia, los pagos en especie o los costos implícitos de formación.

Por las razones anteriores, Packard (2007) examina la elección de sector en Chile y modela tres tipos de empleos: un sector formal de empleo asalariado bajo contrato, un sector informal que consiste de empleo asalariado sin contrato y por último el autoempleo. Sobre estos tipos de empleo realiza una serie de hipótesis que le permiten probar si pertenecen o no a la visión dualista del mercado laboral, es decir, puntualmente desea conocer si el sector informal y el

autoempleo son o no sectores residuales de la economía. Los resultados son los siguientes: para el autoempleo no existe evidencia de que este sea un sector residual de un mercado laboral dualista, es decir los individuos eligen el autoempleo, por el contrario el empleo informal si exhibe características de un sector residual tal como se plantea en la literatura perteneciente al enfoque dual.

En el trabajo de Friedman, Johnson, Kaufmann y Zoido-Lobaton (2000) se realiza un estudio de la informalidad desde una perspectiva institucional para 69 países, encontrando que elevadas tasas de impuestos se asocian con menor actividad informal, en cambio la mayor burocracia, mayor corrupción y un entorno legal débil se asocian con una gran economía informal. La razón para que altas tasas impositivas reduzcan la informalidad es que los ingresos por impuestos permiten la provisión de un mejor entorno legal. Se argumenta que los empresarios se trasladan a la informalidad no para evadir impuestos sino con el fin de reducir la carga burocrática y la corrupción.

En esta misma dimensión analítica, Heckman y Pagés (2000) investigan sobre el costo de regulación de la seguridad laboral argumentando que las regulaciones reducen el empleo agregado y promueven la desigualdad entre los trabajadores. Construyen un índice de seguridad laboral (JS) que refleja la magnitud de los costos de despido para un grupo de países principalmente de América Latina así como para algunas naciones industrializadas a finales de 1990. Se plantean como objetivo cuantificar el impacto de las regulaciones de seguridad laboral sobre las tasas de empleo y desempleo por medio de tres estimaciones: mínimos cuadrados ordinarios, efectos fijos y efectos aleatorios.

Los autores concluyen que la seguridad laboral tiene un efecto negativo sobre las tasas de empleo, sin embargo, no tiene el mismo efecto para todos los trabajadores, pues el impacto para los jóvenes es mayor. A pesar de que la evidencia no es concluyente en cuanto al impacto de la seguridad laboral sobre la composición del empleo formal-informal, los resultados indican que un incremento en la seguridad laboral está asociado con una disminución en el empleo formal y con un aumento del empleo informal. Por último, las provisiones de seguridad laboral son ineficientes debido a que reducen la demanda de trabajo y aumentan la desigualdad dado que incrementan el tamaño del sector informal. En este contexto es posible ubicar el trabajo de Heckman y Pagés dentro de una perspectiva de carácter institucional.

Existen otros trabajos que investigan los determinantes de la economía informal por regiones entre ellas África, Asia, Países Industrializados, América Latina y Países en Transición (Krakowski, 2005). Los resultados de dicho trabajo indican que es de importancia el factor regional en la explicación del tamaño de la economía informal. En lo que concierne a las variables explicativas del tamaño de la economía informal (variable dependiente), la efectividad del gobierno es significativa sólo para los países industrializados y aquellos en transición, además el signo del coeficiente indica que ésta variable reduce el tamaño de la economía informal. Por su parte, la regulación laboral solamente es significativa para los países industrializados y sugiere que aumenta el tamaño de la economía informal. Este documento también se puede incluir entre aquellos de orden institucional.

Hay otros autores como Günther y Launov (2006) que reconocen que el mercado laboral informal es muy heterogéneo debido a que existe un nivel superior en el que se encuentran los individuos mejor remunerados y uno inferior en el que se encuentran los trabajadores con bajos salarios. Por lo tanto en su estudio desean probar si la estructura heterogénea del mercado laboral informal es resultado de la segmentación laboral o no. Para esto utilizan datos de la encuesta de hogares de Costa de Marfil correspondiente al año de 1998.

Los resultados indican que efectivamente el mercado laboral informal consta de dos segmentos con una ecuación salarial distinta, el segmento del nivel superior es mayor en términos de ingresos, de retornos de la educación y de experiencia, por lo que los individuos de este segmento son voluntariamente informales. En cambio el segmento inferior es una estrategia de escape al desempleo involuntario, encontrándose evidencia a favor de la segmentación como causa de la heterogeneidad del mercado laboral informal. Aun cuando esta investigación se puede ubicar en el enfoque dual de la informalidad, también es posible notar el enfoque mixto de la misma, pues además los autores hacen referencia a la informalidad voluntaria o a la explicación institucional de la informalidad.

Dreher y Schneider (2006) emplean una base de datos correspondiente al periodo 1999-2002 para 120 países, con la finalidad de analizar la relación entre la corrupción y la economía informal. La evidencia sugiere que pasar a la informalidad es una alternativa a la corrupción en los países de elevados ingresos. Entre otros resultados, logran determinar que una mayor regulación conduce a más corrupción y a una economía informal de mayor dimensión. Es posible identificar el corte

institucional del análisis en este trabajo, ya que permite cuestionarse sobre los “beneficios” de una economía más regulada, sobretodo su posible efecto negativo sobre los niveles de informalidad.

Otro estudio de particular interés es el de Rei y Bhattacharya (2008) quienes determinan el impacto de las instituciones y la política sobre la economía informal para 111 países en desarrollo (ingresos bajos y medios). A través de un análisis de regresión, identifican que tanto la regulación del mercado de trabajo como los impuestos ayudan a disminuir la economía informal cuando se combinan con un gobierno eficaz. Asimismo un sistema bancario supervisado y controlado promueve la economía formal. Además la inversión extranjera directa también puede promover la economía informal, sin embargo deben revisarse los canales por los que lo hace. En general determinadas políticas e instituciones pueden ayudar a reducir la economía informal.

Siguiendo en la misma línea de análisis, Singh, Jain-Chandra y Mohommad (2012) estudian los determinantes de la economía informal con énfasis en el papel de las instituciones y el estado de derecho, bajo el supuesto de que grandes economías informales limitan la capacidad del estado para ofrecer gobernabilidad y fuertes instituciones. La pregunta clave es el grado en que las deficiencias en la calidad institucional pueden propiciar el ambiente para la actividad económica informal. La unidad de análisis es una muestra de 100 países donde se incluyen algunas naciones avanzadas, emergentes y en desarrollo. La medida de calidad institucional se compone del promedio de seis indicadores de percepción de calidad: voz y rendición de cuentas, estabilidad política y ausencia de violencia, efectividad gubernamental, calidad regulatoria, estado de derecho y control de la corrupción. Estos indicadores son tomados de la base de datos de gobernabilidad del Banco Mundial.

Los resultados principales son tres; el primero indica que mejores instituciones se asocian con una participación menor de la economía informal, en segundo lugar las instituciones son el determinante más importante del tamaño de la economía informal y finalmente los países con mayor corrupción tienden a poseer economías informales más grandes. Por lo tanto, este documento de investigación también corresponde a la visión institucional de la informalidad.

En Maloney (1999) se encuentra una alternativa a la perspectiva dual del mercado de trabajo argumentando que los trabajadores pueden elegir trabajar en el sector informal debido a dos

factores que disuaden la elección del trabajo asalariado formal; el primero son las leyes de protección laboral que con frecuencia cobran un impuesto implícito a los trabajadores y el segundo factor son las características deseables que presentan los trabajos informales tales como la flexibilidad o posibilidades de formación. El estudio de Maloney emplea dos fuentes, la Encuesta Nacional de Empleo Urbano (ENEU) y la Encuesta Nacional de Micronegocios (ENAMIN), con ello busca determinar si la estructura dual del mercado de trabajo se cumple, si es así, entonces deben observarse los siguientes patrones de movilidad entre los sectores:

- Tasas muy bajas de rotación del sector formal.
- Un enorme flujo unidireccional de trabajadores del sector informal al formal, donde permanecen hasta su jubilación.
- La probabilidad de moverse al trabajo asalariado formal es una función creciente de la experiencia, esto en relación con permanecer en el sector informal.

Los resultados contradicen todos los patrones, esto es, el sector formal muestra elevados niveles de movilidad con altas tasas de rotación, asimismo el flujo no es unidireccional debido a que existe búsqueda de oportunidades de empleo entre los sectores y la probabilidad de moverse al sector formal no crece conforme la experiencia. Por lo cual desde este punto de vista se encontró poca evidencia para México en favor de la visión dualista de la informalidad.

La excesiva informalidad en México provoca costos en términos de productividad (Busso, Fazio y Levy, 2012). Considerando los años de 1998 y 2008 los autores argumentan que la informalidad se debe a la regulación asimétrica entre trabajadores asalariados y no asalariados, esto es, la regulación laboral establece ciertos derechos para los asalariados así como obligaciones para las empresas (por ejemplo la inscripción al Instituto Mexicano del Seguro Social, el pago por despido, el salario mínimo, etc.) que los contratan. Por el contrario tales derechos y obligaciones no aplican para los no asalariados (donde entran los informales).

Por tanto, la vía por la que afecta la informalidad a la productividad es por medio de las distorsiones creadas en el precio de los factores dada la asimetría en la regulación laboral. El hecho que los autores consideren cuestiones de regulación permite encuadrar su trabajo bajo la perspectiva institucional de la informalidad laboral, asimismo corroboran los efectos negativos de este fenómeno sobre la economía.

Dentro de las posibles explicaciones de la informalidad también ha sido incluido el análisis del efecto de la competencia en el mercado sobre el incremento de ésta, tal como se señala en Rodríguez, Lima y Mulato (2013), donde por medio de un método en dos etapas y haciendo uso de microdatos de la Encuesta Nacional de Empleo de 1987 a 2004 y de encuestas industriales para México, estructuran un panel de datos que les permite calcular el efecto de la competencia por industria sobre la informalidad, que es medida mediante aquellos empleos no cubiertos por la seguridad social.

El resultado principal indica que una mayor competencia en productos del sector manufacturero incrementa la informalidad dado que aumenta la probabilidad de que un trabajador sea informal, de aquí que los autores sugieran que la desregulación sea acompañada de una reforma laboral bien delineada. Sin embargo, otro resultado que llama la atención y que coincide con uno de los argumentos del enfoque institucional de la informalidad, es que se demuestra que si hay otro miembro del hogar que cuente con seguridad social, entonces el trabajador tiene menor probabilidad de ser formal. O bien se puede decir que la cobertura en materia de seguridad social, del trabajador y algunos miembros de su familia, incentiva la informalidad.

El paso del sector informal al formal ha sido estudiado ampliamente, sin embargo resulta interesante conocer si es posible que los ingresos de los individuos mejoren ante tal cambio. En este sentido, Huesca y Padilla (2012) realizan un estudio en el que su objetivo es determinar si las oportunidades de obtener un trabajo formal en la región Frontera Norte de México durante 2009 a 2011 (un periodo de crisis) mejoran los ingresos de los informales haciendo énfasis en Chihuahua y el resto de estados de la frontera.

Por medio de la implementación de una técnica contrafactual para datos tomados de la Encuesta Nacional de Ocupación y Empleo (ENOE), Huesca y Padilla concluyen que los asalariados del estado de Chihuahua, si pueden mejorar sus salarios en comparación con el resto de los estados de la Frontera Norte. En cambio para los que se desempeñan por cuenta propia (informales) la situación sería diferente, pues sus ingresos se deteriorarían al pasar de la informalidad a la formalidad. En el caso de Chihuahua y para los demás estados de la Frontera son casi imperceptibles las ganancias. Por esta razón se concluye que no existen o son casi nulas las oportunidades en el sector formal para aquellos individuos que se quieran desempeñar por cuenta propia. Este último resultado da cuenta de lo que argumenta la visión dual de la informalidad

pues para ésta, es el contexto económico el que determina que los individuos se encuentren en uno u otro sector (formal e informal).

De la revisión de literatura empírica tanto para otros países como para México es posible notar que los trabajos considerados en este documento han ido en el sentido de verificar cual visión es la que explica la informalidad (por separado), específicamente si es la visión dual o verificar si se cumplen algunos de los supuestos que implica el enfoque institucional, desde una perspectiva microeconómica o macroeconómica.

Por el lado dual, los autores tienen en cuenta cuestiones como el cálculo de la probabilidad de que un individuo se encuentre en el lado informal de la economía, las diferencias salariales entre formales e informales, la manera en que las empresas informales no contribuyen a la mejora de la economía de un país y también permiten intuir el papel que juega la educación para disminuir los niveles de informalidad. En cuanto al enfoque institucional, los autores mencionados en esta sección consideran aspectos relacionados con la aplicación de la regulación por parte del Estado como es la recaudación de impuestos y la regulación del mercado laboral que puede incrementar los costos para los empleadores formales, también incluyen en sus análisis el desempeño del gobierno por ejemplo en forma de corrupción, efectividad y burocracia a la que deben enfrentarse los individuos. Por lo tanto, el hecho de que la presente investigación busque enfrentar la visión dual e institucional de forma simultánea considerando un nivel macroeconómico, resulta de particular interés.

I.6. Metodología de Análisis

La metodología seleccionada para cumplir con el objetivo del análisis de la informalidad laboral en México, son los modelos de regresión con datos en panel, debido a que la información de interés considera a 24 de las 32 entidades de la federación para el periodo que va desde el año 2005 al año 2014. Los distintos modelos estimados corresponden a las metodologías de regresión agrupada, efectos fijos, efectos aleatorios, variables instrumentales y primeras diferencias.

En términos muy generales los modelos planteados poseen como variable a explicar la informalidad laboral total (IL) que a su vez puede ser desglosada en informalidad laboral independiente (ILI) e informalidad laboral dependiente (ILD), cabe aclarar que todas estas son utilizadas en tasas por lo que su denominación cambia más adelante y se identifican como TIL,

TILI y TILD respectivamente. En lo que concierne a los factores explicativos, los modelos consideran dos vectores de variables; D_{it} es un vector conformado por variables explicativas relacionadas con el enfoque dual e I_{it} es un vector conformado por las variables explicativas asociadas a la perspectiva institucional. En las estimaciones seleccionadas para el análisis e interpretación de resultados, cada vector se compone de tres variables distintas.

En el capítulo tres de esta investigación se encuentra la descripción a detalle de la teoría econométrica sobre datos de panel, las diversas metodologías que pueden emplearse, así como las distintas pruebas formales aplicadas para seleccionar el mejor modelo. En el cuarto capítulo se provee la justificación de la consideración de solamente 24 de las 32 entidades de la federación. También se encuentra la descripción de cada una de las variables que conforman a los modelos estimados.

CAPÍTULO II: DISCUSIÓN TEÓRICA. ENFOQUES QUE EXPLICAN LA INFORMALIDAD LABORAL

En este capítulo se discuten los enfoques teóricos que ayudan a proveer los argumentos fundamentales en la explicación del fenómeno de la informalidad. Sus postulados e implicaciones de política económica han sido materia de debate por los especialistas tanto en el plano teórico como en el empírico. En la literatura existente es posible encontrar trabajos que abordan el tema de manera muy distinta. Sin embargo, muchos se caracterizan por adoptar alguno de los dos grandes enfoques ya mencionados, los cuales representan las posiciones dominantes en el campo teórico.

II.1. El Enfoque Dual de la Informalidad Laboral

La visión dual de la informalidad parte del reconocimiento de la existencia de dos sectores en la economía, uno al que se le suele denominar primario o moderno y otro que es el sector secundario o tradicional, mismos que presentan ciertas características distintivas mencionadas más adelante y que en síntesis se pueden asociar a la economía formal e informal respectivamente.

La teoría del mercado dual es parte de la literatura de los mercados laborales segmentados (MLS) que estudia la presencia de dos o más segmentos en el mercado laboral. En palabras de Leontaridi (1998) la dualidad en la literatura de MLS implica que los empleos pertenecen a un sector primario o a uno secundario del mercado laboral. Es común encontrar que en la literatura se haga referencia a estos sectores con distintas denominaciones, como las que menciona Fields (2004) refiriéndose a un sector llamado “capitalista, formal, moderno, industrial o urbano” y otro considerado “de subsistencia, informal, tradicional, agrícola o rural”. Sin embargo, indistintamente del término utilizado para cada segmento del mercado de trabajo la racionalidad teórica es la misma.

La teoría de los mercados laborales segmentados desafía a la teoría económica neoclásica y a la teoría del capital humano sobre la base de que los trabajadores y los puestos de trabajo no se corresponden de acuerdo con un mecanismo universal de mercado. En su lugar, la teoría sugiere que el empleo y la fuerza de trabajo son divididos en segmentos de mercado de trabajo que pueden ser, duales, tripartitas, cuádruples o jerárquicos (Bauder, 2001). Por su parte Cain (1976)

también reconoce que la teoría de los mercados laborales segmentados desafía a la teoría ortodoxa y que la teoría dual surge como alternativa ante la desilusión de la eficacia de la intensificación de las habilidades como forma de redistribución del ingreso, poniendo atención a los tipos de trabajos que se encuentran restringidos para los trabajadores desfavorecidos.

Para Dickens y Lang (1992) una característica o atributo de la teoría de la segmentación es que comúnmente es modelada en dos grandes segmentos, el primario y el secundario. Que de acuerdo con lo señalado en esta sección es la teoría dual. Refiriéndose a los empleos de salarios elevados, Dickens y Lang proponen que existe un grado de rigidez salarial aun cuando existe una gran oferta de trabajo que desea este tipo de empleos. Bulow y Summers (1985) proporcionan en cierto sentido una explicación sobre la rigidez salarial o bien sobre la existencia de diferencias salariales, la idea central es que para las empresas del sector primario puede resultar conveniente pagar un salario más elevado, pues es posible que con este obtengan un mayor esfuerzo de sus trabajadores.

La dualidad alcanza a los trabajadores, pues es una categoría que puede ser estudiada como un segmento de ocupados formales e informales. En este sentido, Reich, Gordon y Edwards (1973) sostienen que las diferencias entre grupos de trabajadores es una cuestión que ha permanecido a lo largo del tiempo, de tal forma que la fuerza laboral se encuentra fragmentada o segmentada. En este contexto la teoría de la segmentación propone que es incorrecto ver las fuentes de segmentación de los mercados laborales como exógenas al sistema económico, es decir, los mercados laborales segmentados son el resultado de un proceso de segmentación.

Por otra parte, algunos autores representativos del planteamiento dualista como Fields (2004), señalan al trabajo de Lewis (1954) como un punto de partida analítico. Sin embargo otros autores toman como referencia de dicho enfoque, el trabajo de Harris y Todaro (1970) pues éstos consideran la migración rural-urbana para plantear un modelo de dos sectores en el que la migración ocurre en respuesta a las diferencias urbanas-rurales en los ingresos esperados, donde la tasa de desempleo actúa como una fuerza equilibradora de tal migración.

Así pues, el mercado dual de trabajo se compone de dos sectores, uno denominado sector primario y otro secundario. El primero corresponde al mercado formal de trabajo, en el que se encuentran los empleos de elevada calidad mientras que el segundo corresponde al mercado

informal de trabajo, que se integra de aquellos empleos considerados de baja calidad. Este hecho provoca que los salarios sean distintos entre ambos sectores, en el moderno (primario) los salarios son altos y en el tradicional (secundario) son más bajos si se les compara con los primeros. Además, es preciso hacer la aclaración que en el ámbito de la presente investigación el término sector formal será tomado para referirse al sector primario y el término sector informal para hacer referencia al sector secundario indistintamente.

El trabajo pionero de Lewis (1954) plantea un modelo compuesto por dos sectores uno es el capitalista y el otro el de subsistencia. Al primero lo define como esa parte de la economía que utiliza capital reproducible y paga a los capitalistas por el uso del mismo. En cambio el sector de subsistencia es toda la parte de la economía que no utiliza el capital reproducible, además se caracteriza porque el producto por trabajador es menor en este sector que en el capitalista. Asimismo existe un diferencial salarial entre los dos sectores a favor del sector capitalista, Lewis calcula esta brecha en 30% o más.

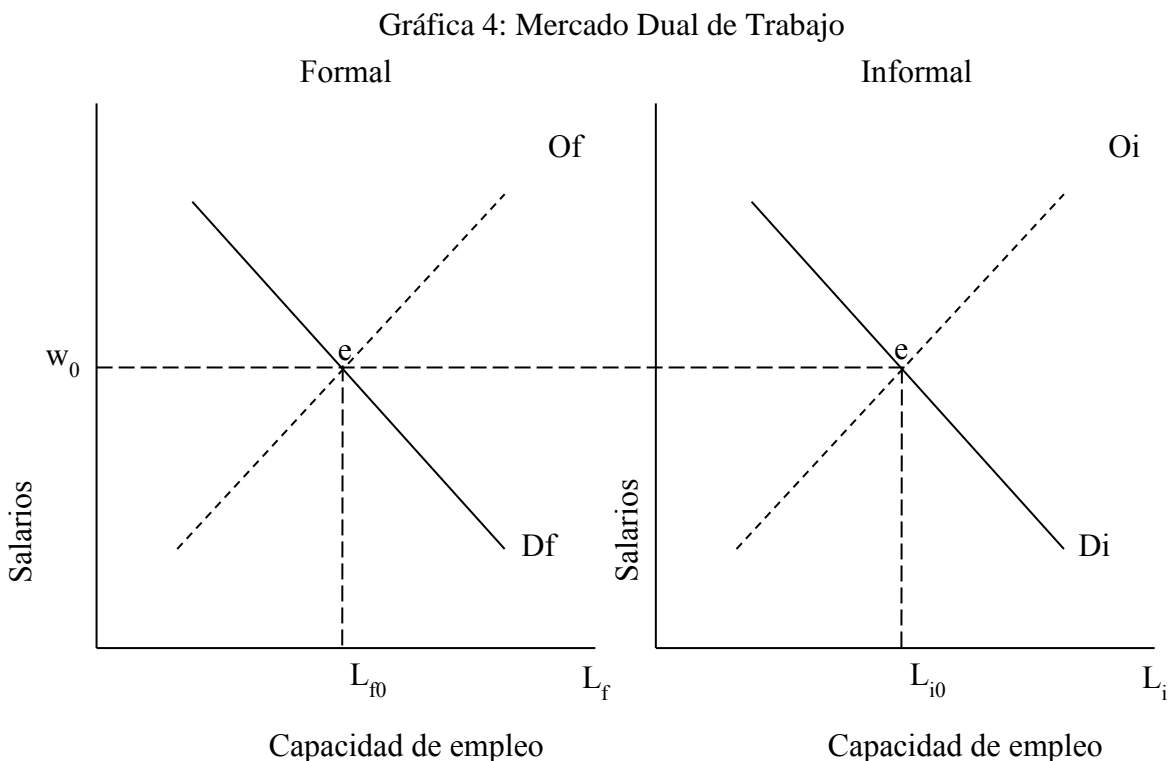
Leontaridi (1998) argumenta que el sector primario comprende todos los “buenos” empleos que se caracterizan por salarios negociados elevados, seguridad económica y rápida rotación que conducen a la promoción profesional. En contraste el sector secundario contiene a los “malos” empleos que suelen ser no calificados, que regularmente no ofrecen escalas profesionales, con niveles salariales que son bajos y determinados competitivamente. Estas líneas permiten identificar claramente la dualidad teórica pues los empleos son calificados en buenos y malos, así como pueden calificarse como formales e informales, esto acorde con la clasificación realizada en la presente investigación.

Los segmentos del mercado de trabajo se caracterizan por ser heterogéneos y es precisamente este hecho lo que le ha concedido gran relevancia en la literatura al enfoque dual. Representa una postura contraria a los planteamientos teóricos previos que afirmaban la existencia de una tendencia hacia la homogenización de la fuerza laboral. En este contexto, una idea adicional que desafía la teoría dual es la que señalan Rebitzer y Robinson (1991) al afirmar que es un enfoque que sostiene que la ley de un solo precio no prevalece en los mercados de trabajo. De acuerdo con dicha ley, bienes idénticos deberían venderse por el mismo precio en un mercado de libre competencia. Esto trasladado al contexto de los mercados de trabajo, implicaría que trabajadores con características productivas idénticas deberían ser pagados con el mismo salario, lo cual no

ocurre según los autores del enfoque dual. De aquí que se presenten diferencias salariales en los mercados primario y secundario.

Una vez identificados los componentes del mercado dual de trabajo y algunas de sus características generales, corresponde ahora explicar su funcionamiento. Riascos (2007) plantea que los mercados de trabajo formal e informal funcionan de manera similar y como un mercado convencional, para ambos sus componentes principales son la oferta y demanda de trabajo, sin embargo difieren en que en uno de ellos la oferta y la demanda es formal y en el otro es informal. Además, en el mercado formal no hay desempleo o es un evento atípico, debido a que el precio (salario real) es perfectamente flexible.

En la gráfica número 4 se muestra el mercado dual de trabajo bajo una situación de estado de equilibrio. Cuando el salario en términos reales aumenta, la oferta laboral en el mercado laboral formal (Of) crece y la demanda laboral (Df) disminuye ocasionando que las tasas de desocupación crezcan o que se presente un exceso de oferta laboral en el mercado formal que lleva a los individuos a trasladarse al mercado informal (reduciéndose las tasas de desocupación), lo que tiene como consecuencia que en este mercado los salarios se vean reducidos.



Fuente: Riascos (2007).

De lo anterior puede rescatarse que los individuos en el mercado secundario (informal) son excluidos del mercado primario (formal) o bien empujados hacia éste como consecuencia del propio funcionamiento del mercado dual. En este sentido, Jiménez (2013) resume adecuadamente lo que sucede en el mercado de trabajo, al sostener que para la teoría dualista, los trabajadores informales constituyen la porción excluida del mercado de trabajo formal en su intento por escapar del desempleo en tal mercado, dada la insuficiente creación de puestos laborales en el sector formal de la economía. Por lo tanto, la inserción en el mercado informal es involuntaria al ser excluidos del mercado formal. Por esta razón, en Perry et al. (2008) se le denomina la visión de exclusión debido a que los ciudadanos quedan fuera de las instituciones formales. Es decir los individuos son excluidos de los beneficios de pertenecer a la formalidad.

Autores representativos de la teoría del mercado dual tales como Dickens y Lang (1985) mencionan que los defensores de esta concepción coinciden en dos principios básicos. El primero se refiere a la caracterización de los empleos del mercado en primarios y secundarios y el segundo a que los empleos primarios se encuentran racionados ya que algunos individuos encuentran difícil emplearse en el sector primario. Asimismo la teoría del mercado dual sostiene que los individuos no siempre pueden elegir el sector de su preferencia, algunos trabajadores que prefieren emplearse en el sector primario no pueden encontrar trabajo en éste.

En Harrison y Sum (1979) se señala que de acuerdo con la teoría dual, los trabajos son organizados en dos segmentos, el centro y la periferia. El centro es dominado por un mercado de trabajo primario que se caracteriza por poseer un elevado poder de mercado y que es capaz de generar los suficientes beneficios que le permiten pagar salarios de no pobreza que provocan que los trabajadores valoren sus empleos y haya estabilidad del trabajo. Por el contrario, la periferia se compone de varios subsectores donde el más importante es el mercado de trabajo secundario, en el que a diferencia del primario los salarios son bajos, se presenta ausencia casi total de beneficios complementarios y las condiciones de trabajo son indeseables, como resultado hay inestabilidad en el trabajo.

El trabajo en el sector secundario es indeseable pues la característica esencial del dualismo es que el sector primario ofrece salarios y otros términos y condiciones de empleo relativamente atractivos (Fields, 2007). El otro sector ofrece salarios menos atractivos, de tal forma que el trabajo en el sector informal es pensado como un tipo de empleo de último recurso. Se puede

decir que las características tan opuestas entre ambos sectores generan que el sector secundario sea visto como una segunda opción forzada o una alternativa indeseable para los individuos, ante la imposibilidad de encontrar un empleo en el sector primario.

En Fields (2005) se retoma la idea del sector informal como un espacio de libre entrada, es decir, un sector en el que cualquiera que desee un empleo puede obtenerlo, o bien en el que las barreras a la entrada son pocas o inexistentes, por esta razón los procesos de determinación de salarios son distintos entre el sector formal e informal. De hecho, esta idea resulta bastante interesante pues logra ofrecer una explicación del porque para los individuos la informalidad es una opción, es decir, las pocas restricciones a la entrada la vuelven bastante atractiva y más aún en aquellos escenarios en los que el desempleo se incrementa o en los que la economía enfrenta dificultades.

Ahora bien, para Piore (1972), la hipótesis del mercado laboral dual (segmentación del mercado en primario y secundario) fue diseñada para explicar los problemas de aquellos individuos que se encuentran en desventaja en el mercado laboral, esto es, aquellos que operan en el sector secundario. Sin embargo, comenta que existen distinciones entre los empleos del sector primario que son tan importantes como la distinción de los empleos en primarios o secundarios.

Por lo tanto, la heterogeneidad planteada en la teoría del mercado dual de trabajo marca la pauta para el reconocimiento de que pueden existir diferencias incluso al interior de los sectores primario y secundario. Es por ello que Watcher (1974) comenta que con el fin de analizar el mercado primario, los dualistas se han concentrado en describir el mercado de trabajo interno lo que resulta en una teoría cuya hipótesis principal es que los salarios y los puestos de trabajo no se distribuyen entre los trabajadores de acuerdo a su productividad, sino más bien, de acuerdo a otros factores.

Una característica adicional de los puestos de trabajo en el sector secundario es que no poseen gran habilidad específica, para su realización se requiere poca o ninguna formación y cuando se le proporciona capacitación al trabajador, es general (Uys y Blaauw, 2006). Para Losby et al. (2002) en el sector secundario se encuentra el trabajo informal que incluye a las personas que operan su propia pequeña empresa de forma no regulada (comúnmente se les conoce como autempleados) o personas que trabajan para empleadores pero fuera de los libros (en el ámbito de la presente investigación, son los informales dependientes).

Hodson y Kaufman (1982) presentan un modelo teórico dual que descansa en cuatro elementos básicos y se enfoca al nivel de la economía nacional más que a un nivel local de los mercados laborales. El primer elemento es la economía dual, es decir, hay dos sectores de la economía, el centro y la periferia, que se distinguen por su poder de mercado donde las empresas del centro son monopólicas u oligopólicas en tanto que las de la periferia son competitivas. El segundo elemento son los mercados duales de trabajo correspondientes al centro y a la periferia, estos son el mercado primario y el secundario. Como tercer elemento plantean tres resultados para el trabajo: en primer lugar la probabilidad de movilidad entre sectores, en segundo término las condiciones de trabajo en cuanto a recompensas tangibles e intangibles y finalmente la conciencia de los trabajadores o su conocimiento de las condiciones e intereses colectivos. Por último, el cuarto elemento se refiere a las divisiones sociales del trabajo, por ejemplo la división de la fuerza de trabajo por raza, género, entre otros.

II.1.1. El Enfoque Estructuralista de la Informalidad Laboral

En este subapartado se describe brevemente una explicación sobre la informalidad laboral, denominada en la literatura como el enfoque estructuralista. Sin embargo, al poseer sus bases en el enfoque dual, presenta características similares y es casi indistinguible en cuanto a la explicación que ofrece la teoría dual. La perspectiva estructuralista es una teoría que le concede gran importancia a la demanda de trabajo en el mercado laboral formal, como el principal factor explicativo de la informalidad laboral, debido a que la cantidad de empleos formales no se corresponden con la cantidad de individuos que demandan un empleo de este tipo.

El enfoque estructuralista destaca que los aspectos macroeconómicos influyen de manera positiva o negativa en el incremento de la informalidad. Es decir, el comportamiento de la economía es relevante como causa de la informalidad. Al respecto, Roldan y Ospino (2009) comentan que la existencia de crisis o recesiones en algunos periodos es un hecho que afecta al sector formal pero también al informal. Por lo cual se puede decir que el comportamiento de la economía a nivel macro, afecta al mercado de trabajo incrementando o disminuyendo la informalidad.

El trabajo de Fields (1972) puede situarse dentro de la vertiente estructuralista debido a que parte del reconocimiento de la migración (un tema a nivel macroeconómico) como causa de la informalidad. Cabe señalar que este autor fue de los primeros en integrar el sector informal en la

literatura, argumentando que el crecimiento del sector urbano o moderno se debe a la migración que se da del sector rural al moderno y no al crecimiento biológico, en este proceso los migrantes más afortunados logran obtener un empleo que es permanente en el sector moderno donde dichos empleos son considerados como los mejores. Los individuos restantes se ven obligados a encontrar algún medio para obtener un ingreso, conformando lo que el autor denomina el sector turbio o lóbrego (*“murky sector”, en inglés*) y es de esta manera como introduce la idea del sector informal. En el sector lóbrego los ingresos son menores que en el sector moderno y comprende un amplio rango de actividades que pueden ser trabajos asalariados o autoempleos. Entre otras características, se encuentran la libre entrada, el hecho que los individuos son considerados como empleados, horarios de trabajo flexibles y finalmente el sector representa una nueva opción de empleo remunerado para cada miembro.

Otro punto de vista interesante en esta vertiente estructuralista de la informalidad y que es de gran interés, es el tema de conocer de qué forma pueden presentarse las relaciones entre los sectores formal e informal. En Tokman (1978) existe un interés por expresar como son tales relaciones, clasificándolas como benignas y de subordinación. Las primeras se dice que son benignas por que el sector informal proporciona empleo aun cuando puede carecer de relaciones con el resto de la economía (sector informal autónomo). Sin embargo reconoce también que el potencial de crecimiento del sector informal autónomo se ve fortalecido cuando las actividades que lo conforman están integradas al resto de la economía es decir las relaciones son complementarias.

Por su parte, las actividades informales son subordinadas cuando por ejemplo el progreso tecnológico se concentra en los sectores modernos de la economía (sector formal) o no se difunde de manera uniforme, o bien es posible encontrar que el sector informal tiene acceso limitado a los insumos modernos y a los mercados de productos (en comparación con el formal) lo que limita a su vez el crecimiento del sector. Finalmente Tokman (1978) concluye que el sector informal no debería ser percibido ni como completamente integrado ni como autónomo, sino más bien como uno que mantiene vínculos de importancia con el resto de la economía, exhibiendo un grado considerable de autonomía.

Por último, Chen (2012) señala que los estructuralistas ven a la economía formal e informal como intrínsecamente ligadas. Ven a las empresas informales y a los trabajadores informales

asalariados como subordinados a los intereses del desarrollo capitalista, proporcionando bienes y servicios que son baratos. También argumenta que los gobiernos deben regular las relaciones comerciales y de empleo para abordar esta relación desigual.

Se puede decir que la visión dual (y su producto estructuralista) atribuye la presencia de la informalidad laboral al comportamiento de la economía así como al mercado de trabajo, es decir, una situación económica difícil provoca que aumente la informalidad debido a que el sector formal expulsa individuos generándose un desequilibrio en el mercado de trabajo y quienes no logran emplearse en el sector formal (primario) lo hacen en el sector informal (secundario), con lo que es posible decir que algunos individuos se encuentran en la informalidad de manera involuntaria o debido a que son “empujados a ella”, donde los empleos poseen características que los hacen estar en desventaja frente a los empleos formales.

II.2. El Enfoque Institucionalista de la Informalidad Laboral

Una segunda explicación teórica respecto a la informalidad laboral acentúa el hecho de que ciertos individuos se encuentran en la informalidad debido a la decisión racional que estos realizan, así como también al hecho de que algunas de las instituciones del Estado pueden crear incentivos para que los individuos opten por la incursión en la informalidad, lo que constituye una perspectiva voluntaria del fenómeno. Este enfoque es denominado institucionalista.

En el marco de la teoría institucionalista se considera que las decisiones tomadas desde las instituciones del Estado generan incentivos que influyen en la elección racional de los agentes al escoger entre la ocupación formal o informal (Jiménez, 2012). Además quienes están en el sector informal lo hacen por decisión y no porque les toca hacerlo de forma involuntaria como en el caso de quienes se encuentran en el sector secundario en el ámbito de la explicación dual. Asimismo la explicación institucionalista de la informalidad resalta que son las reglas las que guían la conducta de los individuos.

De acuerdo con Uribe y Ortiz (2004), el enfoque institucionalista ignora las restricciones que puedan provenir de la estructura económica (a diferencia de la gran importancia que le concede el enfoque dual a este hecho) y se concentra en las decisiones que a nivel individual se realizan con respecto a la decisión de participar en la actividad económica, por lo tanto la informalidad es resultado de una decisión voluntaria. A esta explicación voluntaria de la informalidad Perry, et al.

(2008) se refieren como la visión de escape y puntualmente Maloney y Arias (2008) señalan que desde esta óptica buena parte del sector informal ofrece trabajos que son valorados por los trabajadores como igualmente buenos a aquellos que podrían obtener en el sector formal. Esto implica que muchos trabajadores informales están tan bien (en términos generales de bienestar) como lo estarían en otros trabajos formales apropiados para sus habilidades y que pueden moverse al sector formal pero es su decisión no hacerlo. Con lo que la explicación de una parte de la informalidad voluntaria queda justificada.

Bajo el enfoque institucionalista de la informalidad los individuos (y también las empresas) escapan al cumplimiento con las instituciones, un ejemplo es la evasión de impuestos. Para Pickhardt y Prinz (2012) la fuerza motivacional de una persona que se involucra en actividades de evasión fiscal o en actividades de la economía informal es la posibilidad de incrementar los ingresos, sin embargo esta fuerza parece ser independiente del nivel de ingresos que un individuo obtiene o puede obtener sin pertenecer a la economía informal. En esta última línea los autores permiten entrever que incluso los individuos que se encuentran del lado formal de la economía pueden tener razones para operar en la informalidad voluntariamente.

Otro ejemplo de la parte voluntaria de la informalidad se presenta cuando las empresas deciden evadir ciertas instituciones tales como los impuestos laborales que deben pagar por sus empleados. De acuerdo con Ulysea (2014) si una empresa decide ingresar en el sector formal, se enfrenta a elevados costos fijos y variables (ingresos e impuestos laborales), pero mediante la contratación de trabajadores fuera de los libros las empresas formales pueden evitar tales costos. En este caso si bien las empresas continúan siendo formales, la informalidad asciende debido al número de individuos que son contratados bajo condiciones de informalidad que exentan a los empleadores de cumplir con obligaciones que suponen un costo. De aquí que la informalidad se presente tanto dentro como fuera del sector informal o bien no sea exclusiva de las empresas constituidas como informales.

Maloney (2003) concuerda con la visión voluntaria de la informalidad sobre todo si se trata de los autoempleados informales al argumentar que este sector debe ser visto como un sector empresarial relativamente deseable ya que ofrece a los trabajadores pobres, trabajos decentes como los que podrían obtener en el sector formal, por lo cual no hay nada inferior en él. Esta idea contrasta con la explicación dual en la que el trabajo informal es indeseable.

Un hecho interesante es determinar hasta qué punto puede afirmarse que la informalidad es voluntaria, puesto que en la literatura se reconoce también que las instituciones pueden fomentarla de manera indirecta y es de esto de lo que toman ventaja los individuos, tal es el caso de la regulación laboral, la estructura de impuestos, la falta de credibilidad en las instituciones o la percepción de una incorrecta aplicación de la legislación, solo por mencionar algunos ejemplos.

Para Choi y Thum (2003) una de las razones principales de la existencia del sector informal es el intento de los emprendedores por encontrar un refugio de las distorsiones inducidas por el gobierno, como los impuestos excesivos, la regulación excesiva o la corrupción, ya que entrar a la economía informal les ahorra los costos derivados de las distorsiones. En este mismo contexto el BID (2003) dedica un capítulo al papel de las regulaciones e instituciones en el desempeño de los mercados laborales, en dicho apartado reconoce que una gran cantidad de empleadores, economistas y políticos afirman que la reglamentación laboral obstaculiza la capacidad de los mercados laborales para funcionar bien, es decir, la normativa puede dejar desempleados, asignación ineficiente del empleo entre sectores, empresas y plantas y llevar a la evasión de las leyes laborales. Como ejemplo se puede mencionar la regulación sobre el salario mínimo que puede crear por si misma el aumento en la informalidad.

Sarte (1999) propone que pueden existir dos comportamientos de la burocracia, una es eficiente y controla la entrada y salida de empresas en la economía formal pero también puede existir una burocracia depredadora que busca obtener ingresos de las empresas y este comportamiento es el que daña el bienestar y el crecimiento económico. Por lo tanto, un sector informal grande puede surgir como resultado de las empresas que buscan escapar de la burocracia asociada a la economía formal.

Por su parte Maloney (2004) afirma que si buena parte del sector informal es voluntario en el sentido de que los trabajadores prefieren su empleo actual a uno en el sector formal, implica que el empleo informal debe ser al menos de la misma calidad del empleo formal, esto medido a lo largo de un conjunto amplio de características del empleo. Además los trabajadores parecen encontrar sustitutos para la protección o los servicios ofrecidos por las instituciones formales. Por otro lado, Maloney aclara también que encontrarse en el sector informal es a menudo la decisión óptima de los trabajadores dadas sus preferencias, las limitaciones que enfrentan en términos de

su nivel de capital humano y el nivel de productividad laboral del sector formal en el país. En este mismo orden de ideas, Kucera y Roncolato (2008) mencionan que la noción de trabajo informal voluntario presupone la posibilidad de elección, esto es, que supuestamente los trabajadores informales voluntarios podrían ocupar un empleo formal si quisieran. Sin embargo los hechos demuestran que esto no sucede así para todos los individuos.

Desde una perspectiva de la empresa, Chen (2007) argumenta que el interés de las empresas formales por la producción flexible especializada, competencia global o por la disminución de los costos laborales, ejerce influencia en que estas prefieran las relaciones de empleo informal. Es decir, el punto importante es que algunas empresas formales eligen este tipo de relaciones de empleo informal como medio para evitar sus obligaciones formales como empleadores. En estos casos es la empresa formal no el trabajador informal quien decide operar informalmente y disfrutar de los beneficios de la informalidad. Por lo cual esta realidad que la autora observa le hace subrayar la necesidad de reexaminar la idea de que el empleo informal es voluntario desde la perspectiva de los trabajadores asalariados del sector informal y no sólo desde los autoempleados.

En Bivens y Gammage (2005) se muestra una síntesis sobre la teoría institucionalista que puede considerarse acertada pues señalan que para esta explicación, el empleo informal no es necesariamente un síntoma de la degradación de la calidad del empleo, pero si es una consecuencia del esfuerzo que realizan los empresarios por escapar de la pesada regulación y/o la corrupción oficial asociada al sector formal. En estas líneas es posible identificar la relación entre la informalidad, la regulación y la corrupción.

El sector informal surge cuando se imponen regulaciones e impuestos excesivos por parte del gobierno quien carece de la capacidad para hacer cumplir (Loayza, 1997). En este sentido, un sistema de regulación excesivo hace a la economía formal poco atractiva al imponer elevados costos de entrada a la legalidad así como elevados costos de permanecer legal. Pero también debe resaltarse que de igual manera los individuos enfrentan costos de encontrarse operando en la informalidad.

Los costos de la formalidad de acuerdo con Loayza se pueden dividir en los costos para acceder al sector formal y los costos de permanecer formal. Entre los primeros se encuentran aquellos

relacionados con el registro legal y los costos financieros de entrada. Los segundos tipos de costos los representan los impuestos, las regulaciones que controlan la utilización de los recursos, la forma de producción y la distribución de la producción y de las ganancias. Algunos ejemplos señalados por el autor son las regulaciones relacionadas con la protección del medio ambiente, la asignación de los insumos importados, la protección al consumidor y la calidad, la disponibilidad de capital financiero y el bienestar de los trabajadores y los requisitos burocráticos como la burocracia y el papeleo. Por otra parte, dentro de los costos de la informalidad se encuentran las sanciones cuando se detecta la actividad informal así como la incapacidad para obtener el máximo provecho de los bienes suministrados por el gobierno (Loayza, 1997).

Para Robles (2008) los impuestos en cualquier economía son el medio que posee el gobierno para obtener los ingresos necesarios para su operación y para la provisión de bienes y servicios públicos, sin embargo ciertas distorsiones pueden surgir cuando existen restricciones en la capacidad del gobierno para la recaudación de impuestos, una de las consecuencias de esta falta de capacidad es el surgimiento del sector informal. La capacidad de recaudación se puede decir que se ve reflejada en que la base de contribuyentes no aumente o bien en que un gran número de individuos se encuentren fuera del pago de impuestos, tal como ocurre cuando los empleos de estos pertenecen al sector informal de la economía u operan bajo condiciones de informalidad.

Si bien el aumento de la economía informal o la incurrencia en la informalidad puede deberse a los impuestos, puntualmente al incremento de estos, también puede explicarse por la percepción que pueden tener los individuos sobre los mecanismos de cobro o sobre el cobro mismo ya que si consideran que no son apropiados esto constituirá una razón que los incentivará a evadir impuestos y por lo tanto incurrir en la informalidad. A esta situación descrita, Schneider y Williams (2013) la conceptualizan como moralidad tributaria (*tax morale*) y lo que significa es que si la gente piensa que el sistema impositivo se ha vuelto menos justo y que sus “vecinos” están evadiendo impuestos, entonces aquellos serán más propensos a trabajar en la economía informal. De aquí que la percepción que tienen los individuos sobre las acciones del Estado sean tan importantes en la explicación del aumento o disminución de la informalidad.

Ligado a lo anterior, Schneider y Enste (2002) reconocen que la actividad económica informal tiende a ser más reducida en países donde las instituciones gubernamentales son fuertes y eficientes, por lo que no son las tasas impositivas más altas las que incrementan el volumen de la

economía informal sino la aplicación ineficaz y discrecional del sistema y las leyes tributarias por parte de los gobiernos. Los mismos autores señalan que una economía muy reglamentada, aunada a una administración débil y discrecional de la ley, constituye un terreno particularmente fértil para las actividades informales, y representa también las condiciones en las que prospera la corrupción. En este mismo orden de ideas Friedman et al. (2000) reconocen que cuando las empresas se enfrentan a una burocracia onerosa, elevados niveles de corrupción y un sistema legal débil, ocultan sus actividades “informalmente”, o en otras palabras se puede decir que las instituciones causan el incremento de la informalidad.

Un punto de vista distinto sobre el efecto de los impuestos es que pueden incidir en la disminución de la informalidad. Hibbs y Piculescu (2005) plantean que es posible un círculo virtuoso en el que coexistan amigablemente los elevados impuestos y una baja tasa de evasión de impuestos (con ello una baja tasa de informalidad). La explicación va en el sentido de que se contará con los ingresos fiscales necesarios para mantener el buen funcionamiento de las instituciones (de la calidad de éstas) que sustentan la elevada tolerancia a los impuestos. Es decir, los ingresos que se obtienen por medio de los impuestos permiten que las instituciones ofrecidas por el Estado sean de calidad, lo que eleva la tolerancia a la implementación de impuestos y que los individuos tengan incentivos a permanecer en la formalidad.

Por otro lado, los beneficios que dictan las instituciones que deben recibir los individuos empleados en el sector formal pueden contribuir al incremento de la informalidad, tal es el caso del seguro médico. El empleo del jefe de hogar, en algunos casos ofrece seguro médico para los familiares dependientes principalmente cónyuges e hijos. Como resultado, estos individuos pueden obtener relativamente poca rentabilidad de las contribuciones de seguridad social que deberán pagar si se emplean en el sector formal (OCDE, 2004).

Del argumento anterior de la OCDE se puede concluir que destacan dos efectos de la atención médica que brindan los empleos formales. El primer efecto es positivo y va en el sentido de que puede elevar los niveles de empleo formal pues es un beneficio con el que no cuentan los informales, es decir, puede convertirse en un incentivo a la búsqueda de empleo en el sector formal o bien un incentivo a la permanencia de los individuos en la formalidad. El segundo efecto es contrario y se relaciona con el hecho de que es posible que si algún miembro de la familia cuenta con atención médica, esto desincentive al resto de los miembros a optar por un empleo en

el sector formal pues ya estaría cubierto medicamente, lo que culminaría en un incremento de la informalidad.

Continuando con el tema de la seguridad social Antón, Hernández y Levy (2011) mencionan que algunos países poseen un modelo de seguridad social “dual” que se caracteriza por poseer una parte contributiva y una no contributiva, la primera se refiere a aquella que es obligatoria y proporcionada por el empleo de los trabajadores que reciben su pago en forma de salario, la segunda es voluntaria puesto que es proporcionada por el gobierno a los trabajadores que se encuentran empleados pero no reciben su pago específicamente como salario o bien poseen relaciones no contractuales y que en el ámbito de la presente investigación se puede decir es la cobertura social que reciben los informales. A esta configuración de seguridad social los autores la consideran como una mala política debido a que se traduce en un impuesto al trabajo asalariado (por la parte contributiva de la seguridad social) y en un subsidio al trabajo no asalariado (por la parte no contributiva de la seguridad social).

En este orden de ideas es necesario mencionar que si bien la cobertura social a la que acceden los informales no es de calidad equiparable con aquella que reciben los formales, si constituye un sustituto que puede influir en la decisión voluntaria de los individuos por formar parte de la informalidad pues en este caso se encontrarían cubiertos sin haber pagado un impuesto directo por ello, en comparación con los formales a los que se les deduce del salario de manera obligatoria. También se refuerza la idea de que las instituciones gubernamentales por si mismas crean incentivos para que los individuos incurran en la informalidad, si bien este no es el objetivo de las acciones del Estado si se generan efectos no deseados como puede ser la mayor informalidad.

Un tema adicional relacionado con la explicación de la informalidad es la forma en que el entorno institucional puede influir en el emprendimiento en la economía informal y con esto como consecuencia, en la creación de empresas que proporcionan empleos informales tanto para el creador como para otros individuos, lo que finaliza en el incremento de la informalidad en la economía. Al respecto Webb, Bruton, Tihanyi y Ireland (2012) reconocen que las características del entorno institucional influyen en el proceso del emprendimiento empresarial debido a que las diferencias en las definiciones de las instituciones formales e informales y el grado de burocracia en el ambiente institucional permiten a los emprendedores identificar las oportunidades en la

economía informal. De aquí que algunos individuos emprendedores exploten las oportunidades de operar fuera de la parte formal de la economía, con lo cual una vez más se ve reflejado que para determinados individuos el hecho de pertenecer a la informalidad es una decisión voluntaria, en este caso son los emprendedores quienes eligen voluntariamente operar bajo el esquema informal.

El objetivo de las instituciones creadas por el Estado es remediar algún problema presente en la economía o en la sociedad, sin embargo en algunas ocasiones éstas pueden generar otros problemas, tal es el caso que hasta ahora se ha presentado en el mercado de trabajo pues las instituciones de ciertas formas pueden explicar el surgimiento de la informalidad. Esto mismo ha sido reconocido por autores como Betcherman (2013), quien tiende a bien señalar que las sociedades casi siempre introducen las instituciones del mercado de trabajo para proteger a los trabajadores o para redistribuir el ingreso a ellos. Betcherman enfatiza el efecto de la seguridad laboral o la legislación de protección al empleo, pues posee efectos en los niveles de empleo y de desempleo. Asimismo reconoce que las instituciones pueden influir en la posición en el empleo ya sea por ejemplo, como formal o informal.

En este mismo sentido, la OCDE (2004) reconoce que instituciones tales como el seguro por desempleo pueden incrementar la informalidad, esto ocurre cuando por ejemplo las prestaciones por desempleo son pagadas bajo controles poco estrictos lo que promueve el trabajo no declarado. En cambio si el control es riguroso, el pago de prestaciones puede desincentivar este tipo de trabajo (informal).

Por su parte, Schneider y Klinglmair (2004) listan las siguientes tres causas como las principales que determinan la economía informal: a) los impuestos y las cargas de la cotización a la seguridad social, por que incrementan los costos del trabajo en la economía formal lo que genera mayores incentivos a evadirlos mediante la operación en la informalidad, b) el incremento de la intensidad de las regulaciones que reduce la libertad de los individuos en la economía formal y c) el deterioro de la calidad de los bienes y de la administración pública que incentiva la participación en la economía informal.

Los argumentos anteriores permiten decir que para la visión institucionalista la informalidad surge de la decisión que realizan los individuos ante la posibilidad de escapar a las regulaciones

del Estado, por lo cual se le ha concedido el carácter de voluntario a la explicación que provee esta teoría basada en la elección racional de los agentes en la economía.

II.3. ¿Enfoque Mixto de la Informalidad?

Cabe mencionar que aun cuando las perspectivas dual e institucionalista son opuestas, la mayoría de los autores sobretodo en el plano empírico sugieren que la explicación de la informalidad debe considerar ambas, pues tan importantes son los factores de un enfoque como los del otro y si se pretende abordar el fenómeno de la informalidad laboral es apropiado tener en cuenta las dos visiones ya que el resultado será el análisis integral del problema.

En este contexto, Fields (1990) reconoce que dos cosas muy diferentes están sucediendo en el sector informal, es decir dos partes lo conforman. Una parte consiste de empleo caracterizado por la libre entrada, salarios bajos y que es indeseable en relación al empleo en el sector formal. En cambio otra parte consiste de empleo con limitaciones a la entrada, salarios elevados y que es preferido al empleo en el sector formal. Desde el punto de vista del trabajador, estos dos grupos son distintos, donde los primeros se encuentran peor en su posición relativa al sector formal y los segundos se encuentran mejor. Por estas razones se propone la combinación de enfoques.

A manera de conclusión de este capítulo se puede decir que, las dos grandes visiones teóricas de la informalidad ofrecen explicaciones contrarias al fenómeno de estudio. Por la parte dual la existencia de la informalidad va en el sentido involuntario ya que las condiciones difíciles de la economía impactan el equilibrio del mercado de trabajo generándose un incremento del sector informal. Por el lado institucional, la explicación de la informalidad posee un carácter voluntario, es decir, los individuos y empresas son quienes eligen las relaciones laborales informales, teniendo como objetivo escapar a las regulaciones u obligaciones que implica el sector formal de la economía.

A partir de todo lo anterior se puede reforzar la idea de que la informalidad laboral es de preocupación debido a que su existencia y sobre todo a que su incremento son indicadores de que el mercado de trabajo presenta un desequilibrio o de que las instituciones existentes no son las mejores y terminan fomentando la informalidad, con lo que debe ponerse especial atención en las acciones de política gubernamental a implementar con el fin de reducir los niveles de informalidad en el país.

CAPÍTULO III: METODOLOGÍAS PARA EL ANÁLISIS DE LA INFORMALIDAD LABORAL. MODELOS CON DATOS EN PANEL

En el presente apartado se muestran algunas de las opciones metodológicas disponibles para el análisis del fenómeno de interés en la presente investigación. El punto de partida para la elección de la metodología apropiada, es determinado por las características que presenta la información del fenómeno a explicar, o bien, los datos.

En general, la información se puede encontrar en tres formas, distinguiéndose por su enfoque que puede ser hacia el tiempo, las unidades de análisis (entidades, individuos, etc.), o ambos. El primer caso corresponde a las series de tiempo y se definen como una secuencia ordenada de observaciones a través del tiempo (Wei, 2006). Las series de tiempo se presentan cuando se observan valores de una o más variables a través de un intervalo temporal para una unidad de análisis; en este caso el enfoque es evidentemente el tiempo y no las unidades bajo estudio. A su vez, cuando los datos están centrados en las unidades de análisis, se denominan datos de corte transversal, en este caso se observa más de una unidad en un punto del tiempo, o en palabras de Wang y Jain (2003) los datos de corte transversal son una fotografía instantánea de los datos en un punto dado en el tiempo, por ejemplo para algunas regiones. Finalmente, la información también puede combinar las unidades bajo análisis con el tiempo, constituyendo los datos en panel o datos longitudinales que son aquellos que siguen a una muestra de individuos a través de un periodo y por lo tanto proporcionan múltiples observaciones sobre cada individuo (Hsiao, 2014).

De las tres posibles estructuras de datos, la que se ajusta a la información con la que se cuenta es la que corresponde a los datos en panel debido a que la conformación de la base contempló a 24 entidades federativas de la República Mexicana a través de una frecuencia anual que va de 2005 a 2014. Por lo tanto, la muestra utilizada permite observar a las distintas entidades a través del tiempo.

La estructura de datos en panel presenta ciertas ventajas en comparación con las series de tiempo y con el corte transversal, dichas ventajas han influido en que cada vez sea más común encontrar investigaciones que utilizan los datos en panel. Mayorga y Muñoz (2000) destacan entre las ventajas de utilizar técnicas con datos en panel: a) una mejora en las estimaciones debido a que el

investigador dispone de un mayor número de observaciones que incrementan los grados de libertad y reducen la colinealidad entre las variables explicativas, b) capturan la heterogeneidad no observable ya sea entre unidades individuales de análisis así como en el tiempo, c) permiten estudiar de una mejor manera la dinámica de los procesos de ajuste y d) permiten elaborar y probar modelos más complejos, esto si se les compara con los análisis de series de tiempo y de corte transversal. En este mismo orden de ideas Greene (2012) argumenta que la ventaja fundamental de un panel de datos es que permite al investigador gran flexibilidad al modelar las diferencias del comportamiento entre los individuos. En la presente investigación la heterogeneidad surge de la inclusión de las diversas entidades en el análisis, pues cada una presenta características propias.

Una característica interesante de los datos en panel es que pueden ser balanceados o desbalanceados. El panel será balanceado si los datos están disponibles para todos los individuos en todos los años (Cameron y Trivedi, 2005). Para Baum (2006) un panel balanceado es aquel en el que cada unidad “ i ” es observada en cada periodo “ t ” en contraste con el panel desbalanceado en el que para algunos individuos los datos no están disponibles.

Partiendo de que el panel balanceado implica que t es igual para todas las unidades de análisis i y que en el panel desbalanceado t es diferente para algunas i , en lo que respecta a la informalidad laboral se decidió utilizar solamente la información de los estados para los que se cuenta con la información completa para todas las variables, es decir, se trabaja con un panel balanceado⁴ para 24 entidades que son observadas en el periodo 2005-2014. Es decir, se tiene que $t = 10$ e $i = 24$, de tal forma que el número total de observaciones es $NT = 240$.

Ahora bien, existen distintas técnicas metodológicas para trabajar con datos en panel dependiendo de los supuestos que se realicen sobre las variables incluidas en el modelo a estimar, por esta razón a continuación se describen algunas formas de modelar los datos en panel comenzando desde una versión básica o general, hasta algunas más complejas. El elemento clave entre las distintas metodologías es el hecho de que se asuma que existe o no un efecto no observado y que éste a su vez se encuentre o no correlacionado con alguna de las variables explicativas.

⁴ En el cuarto capítulo se encuentra la justificación para el uso del panel balanceado.

III.1. Modelo de Regresión Agrupada

El modelo de regresión agrupada es el modelo base para el análisis de datos en panel. En Wooldridge (2002) se señala que el panel de datos puede ser utilizado bajo ciertos supuestos para obtener estimadores consistentes en presencia de variables omitidas o también denominadas variables no observadas. Se desarrolla el modelo comenzando por asumir que tanto y como $\mathbf{x} \equiv (x_1, x_2, \dots, x_K)$ son variables aleatorias observadas y que c es una variable aleatoria no observada que captura las características de un individuo que están dadas y no cambian a través del tiempo. Entonces el interés recae en el vector $(y, x_1, x_2, \dots, x_K, c)$ y en los efectos parciales de las x_j variables explicativas en la función de regresión poblacional, es decir, $E(y \mid x_1, x_2, \dots, x_K, c)$ donde c se mantiene constante.

Ahora bien, considerando un modelo lineal con c introducida de forma aditiva con las x_j variables, se tiene que $E(y \mid \mathbf{x}, c) = \beta_0 + \mathbf{x} \boldsymbol{\beta} + c$. En cuanto a los estimadores, el vector $\boldsymbol{\beta}$ (de dimensión $K \times 1$) es de particular relevancia por lo cual es importante establecer si puede ser estimado consistentemente, para ello es necesario poner atención en la variable c . Si no existe correlación entre c y las x_j , entonces no existe un problema debido a que c solamente será otro factor no observado que afecta a y . En otros términos, lo que se buscaría para estimar consistentemente a $\boldsymbol{\beta}$ es que se cumpla que la covarianza entre las variables explicativas y el efecto no observado sea exactamente igual a cero, es decir que $\text{Cov}(x_j, c) = 0$ para alguna j , con ello se podrá afirmar que no existe ningún grado de variación conjunta entre estos elementos o bien no hay dependencia. En cambio si resultara que $\text{Cov}(x_j, c) \neq 0$ para alguna j , entonces incluir a c en el término de error constituye un problema de gran magnitud, es decir, la estimación inconsistente de $\boldsymbol{\beta}$. Cabe resaltar que para estimar una regresión agrupada por Mínimos Cuadrados Ordinarios (MCO) es necesario que se cumpla que $\text{Cov}(x_j, c) = 0$.

Adicionalmente a la estimación base de un panel de datos por medio de la regresión agrupada, Wooldridge (2002) presenta un ejemplo en el que aterriza la parte introductoria de los párrafos previos. El ejemplo es de un panel para dos periodos de tiempo (siguiendo esta lógica, el modelo puede extenderse a más periodos), supone que puede observar a y y \mathbf{x} en dos puntos del tiempo

distintos y que además la variable omitida c es contante en el tiempo, por lo cual la función de regresión poblacional es:

$$E(y_t | \mathbf{x}_t, c) = \beta_0 + \mathbf{x}_t \boldsymbol{\beta} + c, \quad t = 1, 2 \quad [1]$$

Donde $\mathbf{x}_t \boldsymbol{\beta} = \beta_1 x_{t1} + \dots + \beta_K x_{tK}$ y el subíndice indica la variable j en el tiempo t . Luego se reescribe la ecuación 1 en términos de error como:

$$y_t = \beta_0 + \mathbf{x}_t \boldsymbol{\beta} + c + u_t, \quad [2]$$

Donde se esperaría que $E(u_t | \mathbf{x}_t, c) = 0$ y para poder aplicar MCO es necesario que $E(x_t' c_t) = 0$, pero si c está correlacionada con alguna variable contenida en \mathbf{x}_t entonces debe buscarse alguna alternativa para estimar los datos en panel, por ejemplo a través del modelo de efectos fijos o del modelo de efectos aleatorios ya que MCO agrupados reportará estimadores sesgados e inconsistentes.

III.2. Modelo de Efectos Fijos

A continuación se presenta el modelo de efectos fijos, como una extensión del modelo de regresión agrupada, y la diferencia con el anterior es que su aplicación es posible cuando se asume que está presente una relación entre el componente de error y las variables independientes, o en términos del modelo anterior existe correlación entre c y \mathbf{x}_t . Si se encuentra presencia de correlación, entonces estimar $\boldsymbol{\beta}$ por medio de un modelo de regresión agrupada sería erróneo, pues proporciona estimadores sesgados e inconsistentes en estos casos.

El desarrollo del modelo se realiza siguiendo a Davidson y MacKinnon (1993) quienes parten de un modelo con datos en panel como:

$$y_{it} = x_{it}(\beta) + u_{it}, \quad t = 1, \dots, T, \quad i = 1, \dots, n \quad [1]$$

En la ecuación 1 t hace referencia a la dimensión temporal que va desde 1 hasta T periodos, e i es la dimensión de corte transversal considerando desde 1 hasta n unidades de análisis. Por lo tanto, el número total de observaciones es nT . Para tratar los datos en panel los autores proponen un enfoque que se basa en utilizar los modelos de componentes de error de los que forman parte los

modelos de efectos fijos y de efectos aleatorios. La idea de los modelos de componentes de error es modelar u_{it} (el error) como un componente formado de tres choques individuales que poseen la característica de ser independientes entre sí, es decir:

$$u_{it} = e_t + v_i + \varepsilon_{it}, \quad [2]$$

De la ecuación 2 resalta que e_t afecta a todas las observaciones para el periodo de tiempo t , que v_i afecta a todas las observaciones para la unidad de corte transversal i y que ε_{it} sólo afecta a la observación ti . En este contexto, existen dos formas generales de estimar un modelo de regresión cuyo término de error se compone como en la ecuación 2, éstas son por medio del modelo de efectos fijos y el modelo de efectos aleatorios. El modelo de efectos fijos puede ser estimado por MCO y el de efectos aleatorios requiere el uso de Mínimos Cuadrados Generalizados (MCG) o el método de Máxima Verosimilitud (MV). Una ventaja de efectos fijos es que no es necesario que los regresores e_t y v_i sean independientes.

Para simplificar el análisis, los autores asumen que $e_t = 0$ para todos los t , por lo cual el error queda compuesto solamente por los choques v_i y ε_{it} . Además suponen que la función de regresión para la observación ti es $\mathbf{X}_{it} \beta$ y señalan que la idea de efectos fijos es tratar a los v_i 's como parámetros desconocidos y estimarlos junto con β .

En este escenario, el modelo de componentes de error lo reescriben como:

$$y_{it} = \mathbf{X}_{it} \beta + v_i + \varepsilon_{it} \quad [3]$$

III.2.1. Efectos Fijos con Variables Dicotómicas

Para estimar el modelo de efectos fijos, partiendo de un modelo general como el planteado en la ecuación 3 de la sección anterior, se pueden seguir dos procedimientos; uno es agregar n variables dicotómicas y otro es trabajar con las variables en primeras diferencias. Davidson y MacKinnon (1993) optan por la primera opción al agregar n variables dicotómicas D_{it}^j a la ecuación 3, donde cada variable es igual a la unidad cuando $i = j$ y en cualquier otro caso es igual a cero. Además con el objetivo de evitar la trampa de la variable dicotómica, deberá

omitirse una de estas si \mathbf{X}_{it} incluye un término constante o el equivalente. La versión de efectos fijos en notación matricial utilizando la opción de las variables dicotómicas es tal como en la ecuación 4, donde \mathbf{v} es un vector n con el elemento v_i . Siempre que los ε_{it} 's sean independientes e idénticamente distribuidos el modelo [4] podrá ser estimado por MCO.

$$y = X\beta + Dv + \varepsilon \quad [4]$$

III.2.2. Efectos Fijos en Primeras Diferencias

La segunda opción para estimar el modelo de efectos fijos es trabajarlo en primeras diferencias. Para presentar esta alternativa se sigue lo planteado por Greene (2012). El autor comienza con el modelo de regresión base de un panel de datos que puede expresarse como en la ecuación 1 o en la 2:

$$y_{it} = \mathbf{x}'_{it}\beta + \mathbf{z}'_i\alpha + \varepsilon_{it} \quad [1]$$

$$y_{it} = \mathbf{x}'_{it}\beta + c_i + \varepsilon_{it} \quad [2]$$

En la ecuación 1 hay K regresores en \mathbf{x}_{it} , sin incluir un término constante. La heterogeneidad o efectos individuales son capturados en $\mathbf{z}'_i\alpha$ donde \mathbf{z}_i contiene un término constante y un conjunto de variables individuales o grupales específicas que son constantes en el tiempo y que pueden ser observadas o no observadas. De aquí que si \mathbf{z}_i se observa para todos los individuos, entonces el modelo en la ecuación 1 o 2 (en la ecuación 2 el término $\mathbf{z}'_i\alpha$ se sustituye por c_i) puede tratarse por MCO. Sin embargo, generalmente c_i es no observada (comprende a las variables omitidas) y se pretende que no se encuentre correlacionada con las variables incluidas en el modelo. Por esta razón, el objetivo en el modelo de primeras diferencias es eliminar c_i (efecto no observado) para eliminar la posible correlación con las variables observadas.

Por lo tanto, la ecuación 2 en primeras diferencias es:

$$\Delta y_{it} = \Delta c_i + (\Delta \mathbf{x}_{it})'\beta + \Delta \varepsilon_{it} \quad \circ$$

$$\Delta y_{it} = (\Delta \mathbf{x}_{it})'\beta + \varepsilon_{it} - \varepsilon_{i,t-1} \quad \circ$$

$$\Delta y_{it} = (\Delta \mathbf{x}_{it})' \boldsymbol{\beta} + u_{it} \quad [3]$$

Para Greene (2012) la ventaja de este enfoque en primeras diferencias es que elimina la heterogeneidad latente, pero por otra parte, la desventaja es que la diferenciación también remueve del modelo cualquier variable constante en el tiempo. De este modo la ecuación 3 se puede estimar utilizando Mínimos Cuadrados Generalizados (MCG) debido a que esta metodología trata a los errores como una perturbación de media móvil y esto es u_{it} .

III.3. Modelo de Efectos Aleatorios

De acuerdo con Cameron y Trivedi (2013) un modelo de efectos aleatorios es un modelo en el que se asume que el efecto individual (efecto no observado hasta lo que aquí se ha mostrado) se encuentra distribuido independientemente de los regresores, en cambio un modelo de efectos fijos trata al efecto individual como distribuido aleatoriamente y potencialmente correlacionado con las variables explicativas.

El modelo de efectos aleatorios es desarrollado con base a Wooldridge (2010) quien comienza con un modelo de efectos inobservables tal como el que se muestra a continuación:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \dots + \beta_k x_{itk} + a_i + u_{it} \quad [1]$$

En el planteamiento de la ecuación 1 se incluye el intercepto (β_0) y se identifica al efecto no observado con a_i . En este contexto, si se supone que a_i no está correlacionado con ninguna variable explicativa en todos los t periodos, entonces no es necesario realizar transformación alguna para eliminar a_i ya que en caso contrario el resultado son estimadores ineficientes, por esta razón en efectos fijos se parte de asumir que sí hay correlación entre a_i y las variables explicativas, con lo que el modelo es tratado con variables dicotómicas o en primeras diferencias bajo el objetivo de eliminar el efecto no observado a_i de acuerdo con lo presentado en la sección anterior.

Entonces, la ecuación 1 se convierte en un modelo de efectos aleatorios cuando se supone que no existe correlación entre el efecto no observado a_i y alguna de las variables explicativas, lo que se representa como $\text{Cov}(x_{ijt}, a_i) = 0$ para $t=1,2,\dots,T$ y $j=1,2,\dots,k$. Ahora bien, el término de error

compuesto es v_{it} y se define como $v_{it} = a_i + u_{it}$, con lo que la ecuación 1 se reescribe de la siguiente manera:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \dots + \beta_k x_{itk} + v_{it} \quad [2]$$

En la ecuación 2, a_i se ubica en el error compuesto en cada periodo lo que provoca que las v_{it} se correlacionen serialmente en cada periodo de tiempo. Así, debido a que la correlación serial puede ser sustancial es posible utilizar MCG. Bajo los supuestos de efectos aleatorios se tiene:

$$\text{Corr}(v_{it}, v_{is}) = \frac{\sigma_a^2}{(\sigma_a^2 + \sigma_u^2)}, \quad t \neq s \quad [3]$$

Para eliminar la correlación serial Wooldridge (2010) define el parámetro $\lambda = 1 - [\sigma_u^2 / (\sigma_u^2 + T\sigma_a^2)]^{1/2}$ que se encuentra entre cero y uno y que permite obtener la ecuación 4 transformada, en la que la barra superior indica los promedios a lo largo del tiempo y datos cuasi deducidos en cada variable. En dicha ecuación los errores no presentan correlación serial, además permite incluir variables constantes en el tiempo y esta representa una ventaja de efectos aleatorios en comparación con efectos fijos.

$$y_{it} - \lambda \bar{y}_i = \beta_0(1 - \lambda) + \beta_1(x_{it1} - \lambda \bar{x}_{i1}) + \dots + \beta_k(x_{itk} - \lambda \bar{x}_{ik}) + (v_{it} - \lambda \bar{v}_i) \quad [4]$$

Finalmente, dado que λ no se conoce, puede estimarse como $\hat{\lambda} = 1 - \{1/[1 + T(\hat{\sigma}_a^2 / \hat{\sigma}_u^2)]\}^{1/2}$ donde $\hat{\sigma}_a^2$ es un estimador consistente de σ_a^2 y $\hat{\sigma}_u^2$ es un estimador consistente de σ_u^2 . Por lo cual el estimador de MCG que utiliza $\hat{\lambda}$ se denomina estimador de efectos aleatorios que es consistente y posee una distribución normal asintótica a medida que N aumenta con T fijo.

A manera de conclusión, las metodologías de efectos fijos y efectos aleatorios comparten la característica de permitir trabajar con datos en panel e incluir un efecto no observado (c_i o a_i , en los modelos anteriores) en su especificación matemática. Sin embargo, la principal diferencia entre ambas radica en que en contraste con efectos fijos, en efectos aleatorios se asume que el efecto no observado no está correlacionado con ninguna variable explicativa. Por lo que una metodología es alternativa de la otra y deberán utilizarse pruebas formales para elegir entre estas dos.

III.4. Modelo con Variables Instrumentales

Sumado a lo anterior, un panel de datos puede ser estimado haciendo uso de la metodología de variables instrumentales (VI). Precisamente se utiliza VI cuando se sospecha que alguna (o algunas) de las variables independientes o explicativas se encuentra correlacionada con el término de error, es decir la variable es endógena en el modelo y no exógena como sería lo apropiado. Para Johnston y DiNardo (1997) bajo los supuestos clásicos, los estimadores de MCO son los mejores lineales e insesgados y uno de los más relevantes es la independencia entre las variables explicativas y el término de error. Sin embargo, cuando este supuesto no se satisface los estimadores de MCO son sesgados e inconsistentes y por lo tanto no será apropiado utilizarlos bajo esta situación.

Ahora bien, la fuente de endogeneidad puede atribuirse a los errores en la medición de las variables tal como señalan Davidson y MacKinnon (1993) al argumentar que casi todas las variables económicas se miden con error y que las consecuencias estadísticas son severas porque generalmente se correlacionan con los términos de error. Entonces se puede afirmar que en presencia de endogeneidad debe recurrirse a una metodología alternativa como es VI en la que variables conocidas como instrumentos sustituyen a las variables originales permitiendo solucionar el problema, dando lugar a estimaciones consistentes. Las dos condiciones principales para elegir una variable como instrumento es que no se encuentre correlacionada con el término de error (es decir que sea una variable exógena) y que por el contrario si esté fuertemente correlacionada con la variable explicativa que se supone es endógena en el modelo. Pindyck y Rubinfeld (2001) desarrollan la estimación por VI en forma matricial partiendo de un modelo como a continuación:

$$Y = X\beta + \varepsilon \quad [1]$$

Para obtener estimaciones consistentes de β se utiliza una matriz de dimensión $N \times k$ instrumentos $Z = (Z_1, Z_2, \dots, Z_k)$, donde cada Z_i tiene un número N de observaciones. Además Z debe satisfacer las condiciones de que cada instrumento no se correlacione con el término de error, que si se correlacione con las variables en X (respectivamente) y ser linealmente independiente. Una vez que se cuenta con los instrumentos apropiados (aquellos que satisfacen las condiciones), la ecuación 1 se premultiplica por Z obteniendo:

$$\mathbf{Z}'\mathbf{Y} = \mathbf{Z}'\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}'\boldsymbol{\varepsilon} \quad [2]$$

$$\text{De lo cual se deriva que, } \hat{\boldsymbol{\beta}}^* = (\mathbf{Z}'\mathbf{X})^{-1} \mathbf{Z}'\mathbf{Y} \quad [3]$$

Donde $\hat{\boldsymbol{\beta}}^*$ será un estimador consistente de $\boldsymbol{\beta}$ y éste es precisamente el objetivo de la metodología de VI. Como nota final, los modelos de efectos fijos y de efectos aleatorios también es posible estimarlos mediante el uso de la metodología de variables instrumentales (VI).

III.5. Elección del Modelo Apropriado con Datos en Panel

Hasta este punto se presentaron las diferentes alternativas para la estimación de un modelo con datos en panel, ya sea por medio de la regresión agrupada, efectos fijos con variables dicotómicas o en primeras diferencias, efectos aleatorios y variables instrumentales. Sin embargo, una pregunta muy razonable que surge es ¿cómo elegir apropiadamente un modelo con datos en panel?, es decir, ¿en qué situación se aplica un modelo u otro?, para ello es importante hacer uso de algunas técnicas formales que ayuden en la elección.

En este contexto, Montero (2011) proporciona una guía bastante interesante y de utilidad en la elección de un modelo con datos en panel, partiendo de un modelo general como el de la ecuación 1, donde se supone que el error puede descomponerse en dos, una parte fija constante para cada individuo (v_i) y otra aleatoria (u_{it}) y que a su vez cumple los requisitos para estimarse por MCO.

$$y_{it} = \alpha + \beta X_{it} + v_i + u_{it} \quad [1]$$

Montero propone que para resolver qué modelo aplicar, se deben solucionar las siguientes cuestiones:

1. ¿La varianza de v_i es significativamente distinta de cero? Si es así, entonces existe un componente no observado de la varianza asociada a cada individuo y la estimación por MCO estará sesgada. En este caso no se puede aplicar el modelo de regresión agrupada.
2. Si la respuesta a la pregunta 1 es afirmativa, entonces la siguiente cuestión es ¿se tiene un panel en el que se incluyen todos los individuos del universo? Si la respuesta es sí, entonces aplicar un modelo de efectos fijos.

3. Si la respuesta a la pregunta 2 es negativa entonces la cuestión que surge es ¿las estimaciones de efectos fijos y de efectos aleatorios son significativamente distintas? Si es así es mejor escoger efectos fijos, si son iguales entonces se elegirá efectos aleatorios.

Dado lo anterior, lo lógico es preguntarse *¿cómo se pueden responder formalmente las cuestiones mencionadas?*. Para ello es de gran utilidad el uso de pruebas formales, entre las que se encuentran la prueba F restringida, la prueba del Multiplicador de Lagrange de Breusch-Pagan (también conocida como prueba LM) y la prueba de Hausman. Cada una tiene un objetivo diferente pero en conjunto constituyen una gran herramienta de elección sustentada para discernir entre los distintos métodos de estimación.

Comenzando con la prueba F restringida, es posible decir que ésta aplica para comparar entre regresiones no restringidas y restringidas. En este contexto, el modelo de regresión agrupada es un modelo restringido debido a que asume que existe un intercepto común para todos los individuos (por ejemplo β_0 en la ecuación 2 de la sección III.1). Por esta razón se utiliza la prueba F para comparar entre un modelo de regresión agrupada y uno de efectos fijos donde este último es considerado como el modelo de regresión no restringido. Aun cuando en la notación de la sección III.2 no se hace referencia explícita al intercepto del modelo de efectos fijos, en el caso que si se considerara éste sería denotado como β_{0i} donde claramente se vuelve evidente que el intercepto es diferente para cada individuo o unidad de análisis i . Siguiendo a Gujarati y Porter (2010) la prueba F restringida puede expresarse de la siguiente manera:

$$F = \frac{(\sum \hat{u}_R^2 - \sum \hat{u}_{NR}^2) / m}{\sum \hat{u}_{NR}^2 / (n - k)} \quad [2]$$

Donde:

$\sum \hat{u}_{NR}^2$: Suma de cuadrados de residuos de la regresión no restringida.

$\sum \hat{u}_R^2$: Suma de cuadrados de residuos de la regresión restringida.

m : número de restricciones lineales⁵.

⁵ En este caso aquella que surge de asumir un intercepto común para todos los individuos.

k : número de parámetros en la regresión no restringida.

n : número de observaciones.

Asimismo la ecuación 2 sigue la distribución F con m , $(n-k)$ grados de libertad y los subíndices NR y R representan no restringida y restringida, respectivamente. La prueba F también la presentan los autores en términos de R^2 como a continuación:

$$F = \frac{(R_{NR}^2 - R_R^2)/m}{(1 - R_{NR}^2)/(n - k)} \quad [3]$$

Donde R_{NR}^2 y R_R^2 son los valores R^2 obtenidos de las regresiones no restringida y restringida respectivamente. Adicionalmente debe observarse que se cumpla lo siguiente de las condiciones de desigualdad 4 y 5:

$$R_{NR}^2 \geq R_R^2 \quad [4]$$

$$\sum \hat{u}_{NR}^2 \leq \sum \hat{u}_R^2 \quad [5]$$

Finalmente, la hipótesis nula H_0 es que los coeficientes diferenciales son iguales a cero, o en otras palabras, la estimación del modelo de regresión agrupada es apropiada por que el intercepto es igual para todos los individuos. En cambio la hipótesis alternativa asume que los coeficientes diferenciales son distintos de cero o bien el intercepto es distinto para cada individuo, en este caso la estimación del modelo de efectos fijos es el apropiado. En resumen, si el valor F es estadísticamente significativo ($p\text{-value} < 0.05$) se rechaza la H_0 y entonces se deberá elegir la estimación por efectos fijos y no la estimación de regresión agrupada.

En cuanto a la prueba LM, Baltagi, Mátyás y Sevestre (2008) señalan que Breusch y Pagan (1979) propusieron el uso de la prueba del multiplicador de Lagrange para verificar la ausencia de efectos individuales. La idea es que si no hay tal efecto, entonces las perturbaciones del modelo son completamente idiosincráticas o en términos de lo presentado en la sección de los modelos, no existe un componente no observado en el error, por lo cual en este caso se puede aplicar el modelo de regresión agrupada por MCO. En esta situación la varianza de las perturbaciones debería ser muy similar a la de sus medias individuales (Baltagi et al., 2008).

La hipótesis nula que prueba el multiplicador de Lagrange es que la varianza del término de perturbación denotado por μ sea cero o bien $H_0 : \sigma_\mu^2 = 0$ (Baltagi, 2011). El estadístico LM se calcula como:

$$LM = (NT/2(T-1)) \left[\left(\frac{\sum_{i=1}^N e_i^2}{\sum_{i=1}^N \sum_{t=1}^T e_{it}^2} \right) - 1 \right]^2 \quad [6]$$

En la ecuación 6, e_{it} representa los residuales de MCO del modelo de regresión agrupada y e_i denota la suma a través de t . Bajo la hipótesis nula, el estadístico LM se distribuye como χ_i^2 . Si la hipótesis nula se rechaza entonces es preferible utilizar efectos aleatorios y si se acepta se usa la regresión agrupada.

Por su parte, Hausman (1978) propuso una prueba muy general para la especificación de un modelo econométrico y su implementación tiene el objetivo de probar efectos fijos y aleatorios (Nerlove, 2005). Es por ello que la prueba de Hausman es conveniente para decidir qué modelo elegir entre efectos fijos o aleatorios. La hipótesis nula es que los estimadores de efectos fijos y de efectos aleatorios no difieren, por lo que si se rechaza la hipótesis nula entonces será mejor utilizar efectos fijos y si por el contrario se acepta la hipótesis, se utilizarán efectos aleatorios. De una forma general, considérense dos estimadores $\hat{\theta}$ y $\tilde{\theta}$ (que pueden ser los de efectos fijos y aleatorios respectivamente), entonces la situación que se pretende probar mediante la prueba de Hausman según lo señalado por Cameron y Trivedi (2005) es:

$$H_0 : \text{plim}(\hat{\theta} - \tilde{\theta}) = 0,$$

$$H_1 : \text{plim}(\hat{\theta} - \tilde{\theta}) \neq 0 \quad [7]$$

El estadístico de prueba de Hausman es:

$$H = (\hat{\theta} - \tilde{\theta})' (N^{-1} \hat{V}_H)^{-1} (\hat{\theta} - \tilde{\theta}) \quad [8]$$

En la ecuación 8, \hat{V}_H denota la matriz de varianza en la distribución límite. El estadístico H se encuentra distribuido $\chi^2(q)$ asintóticamente bajo la H_0 . Por lo que H_0 se rechazará a un nivel α si se cumple que $H > \chi_\alpha^2(q)$.

Sin embargo, la prueba de Hausman presenta un inconveniente y es que en ocasiones el estadístico resulta con valor negativo y no es posible realizar tal prueba. Por esta razón la alternativa es realizar una prueba de sobreidentificación de restricciones (comando *xtoverid* en STATA) haciendo uso del estadístico Sargan-Hansen. Esta prueba es más robusta en comparación con la de Hausman y garantiza un estadístico no negativo y la posibilidad de elegir entre un modelo de efectos fijos o aleatorios, cuando se estima robusto a problemas de autocorrelación y heterocedasticidad.

Generalmente la prueba de sobreidentificación y como consecuencia el estadístico de Sargan-Hansen suelen asociarse comúnmente con los modelos de variables instrumentales para probar la validez de los instrumentos, pero por otra parte también puede utilizarse para elegir entre los modelos de efectos fijos y aleatorios. Baum (2007) señala que la rutina *xtoverid* de M. Schaffer y S. Stillmans puede utilizarse para llevar a cabo una prueba de Hausman de efectos fijos contra efectos aleatorios, o en otras palabras, para elegir entre estos.

El estimador de efectos fijos emplea las condiciones de ortogonalidad de que los regresores se encuentran no correlacionados con el error idiosincrático e_{it} , es decir, $E(X_{it} * e_{it}) = 0$. En cambio el estimador de efectos aleatorios utiliza las condiciones adicionales de ortogonalidad de que los regresores no están correlacionados con el error específico de grupo u_i o bien $E(X_{it} * u_i) = 0$. Por lo cual estas condiciones adicionales de ortogonalidad constituyen restricciones de sobreidentificación y entonces es posible emplear la prueba de sobreidentificación para elegir entre efectos fijos y aleatorios⁶. La hipótesis nula en este caso es la misma que en Hausman, por lo que el rechazo de la H_0 indica que se debe utilizar el modelo de efectos fijos en lugar del modelo de efectos aleatorios.

Por otra parte, si el modelo con el que se realiza el análisis del fenómeno de interés corresponde a variables instrumentales (VI), entonces el principal objetivo es realizar alguna prueba que permita corroborar que alguna o algunas de las variables explicativas son endógenas y que por tanto esto justifique el uso de la metodología de VI. A continuación se presentan dos formas de probar endogeneidad en un modelo de regresión.

⁶ Esta explicación y más detalles se encuentran en el archivo de ayuda del programa STATA para el comando *xtoverid*.

La primera prueba de endogeneidad se puede encontrar a manera de ejemplo en Wooldridge (2015). Sin embargo de una forma muy puntual se puede decir que esta prueba consiste en realizar una rutina de estimaciones siguiendo una serie de pasos que es posible describirlos de forma general como a continuación:

1. Estimar el modelo seleccionado como el mejor ya sea por efectos fijos o aleatorios. Sin incluir los instrumentos.
2. Estimar una regresión por efectos fijos o aleatorios en la que la variable dependiente sea la variable explicativa endógena y la variable independiente sea la variable instrumental.
3. Obtener los residuales de la estimación en el paso 2.
4. Estimar el modelo del paso 1 pero considerando como una variable explicativa adicional los residuales obtenidos de la regresión en el paso 2.
5. De la estimación obtenida en el paso 4 se pone especial atención a la significancia estadística del coeficiente asociado a los residuales de la estimación en el paso 2. Si éstos no son significativos entonces no hay fuerte evidencia de endogeneidad, es decir, no existe endogeneidad si el $p\text{-value} > 0.05$.

La segunda prueba de endogeneidad es una alternativa a la primera y se conoce en el programa STATA como la prueba de endogeneidad de regresores endógenos. El comando que sirve para obtenerla es `endog`⁷. Bajo la hipótesis nula el regresor (variable explicativa) especificado como endógeno puede ser tratado como exógeno, por lo cual si se cumple que $p\text{-value} > 0.05$ entonces no existe endogeneidad en el modelo y no es necesario el uso de variables instrumentales, en caso contrario la o las regresoras puestas a prueba serán endógenas y deberá buscarse un instrumento para estas.

En este contexto, una vez que se ha determinado que existe endogeneidad en el modelo de regresión por la correlación de una o más variables explicativas con el término de error, el siguiente paso es aplicar alguna prueba que permita conocer si las variables seleccionadas como instrumentos de las explicativas (que son endógenas en el modelo), poseen la característica de ser exógenas. Es decir, debe ser probada la validez de los instrumentos a considerar para evitar errores de especificación, para ello se puede hacer uso de la denominada prueba de Sargan.

⁷ Más detalles se encuentran en el archivo de ayuda del programa STATA para el comando `endog`.

La prueba de Sargan consiste en los siguientes pasos (Gujarati y Porter, 2010):

1. Se dividen las variables en dos grupos, aquellas que son independientes del término de error X_1, X_2, \dots, X_p y las endógenas o dependientes del término de error Z_1, Z_2, \dots, Z_q .
2. Los instrumentos de las variables endógenas son W_1, W_2, \dots, W_s , y debe cumplirse que $s > q$ porque en cualquier otro caso la prueba no es válida.
3. Se sustituyen las variables Z por las W y se estima la regresión original obteniendo los residuos (\hat{u}).
4. Se realiza una regresión de \hat{u} sobre una constante, las variables X y W , pero excluyendo todas las variables Z y se obtiene el R^2 de la regresión.
5. Finalmente, se calcula el estadístico SARG, definido como: $SARG = (n - k)R^2 \sim \chi^2_{s-q}$. Donde n es el número de observaciones y k el número de coeficientes en la ecuación de regresión original. La hipótesis nula de la prueba es que los instrumentos son exógenos. La prueba sigue la distribución χ^2 con $s - q$ grados de libertad, donde s corresponde al número de instrumentos y q al número de variables explicativas en la ecuación original. Si el valor calculado de χ^2 es estadísticamente significativo, se rechaza la validez de los instrumentos. Si por el contrario no es estadísticamente significativo, entonces se acepta como válido el instrumento elegido.
6. La hipótesis nula implica que son válidos todos los instrumentos utilizados. Si la χ^2 calculada excede el valor χ^2 crítico, se rechaza la hipótesis nula, y habrá evidencia de endogeneidad en los instrumentos, por lo tanto la estimación de variables instrumentales no será válida y en este caso deberán buscarse otros instrumentos.

Adicionalmente, todas las estimaciones ya sea regresión agrupada, efectos fijos, efectos aleatorios o variables instrumentales no deben presentar problemas tales como autocorrelación y heterocedasticidad. Por esta razón, las regresiones en la sección de presentación de resultados fueron estimadas robustas a autocorrelación y heterocedasticidad, haciendo uso del comando *cluster* del programa STATA.

A manera de cierre de este capítulo, se puede decir que las opciones metodológicas para trabajar un panel de datos son muy variadas, por lo que la tarea para el investigador es implementar aquella que le permita cumplir con el objetivo de su estudio. Sin embargo, la elección del modelo

apropiado deberá estar basada en la aplicación de diversas pruebas formales que a su vez sean de utilidad en la validación de los resultados obtenidos de la estimación y que le permitan realizar una serie de conclusiones pertinentes respecto al fenómeno de estudio.

CAPÍTULO IV: FUENTES DE INFORMACIÓN Y ESPECIFICACIÓN DEL MODELO DE INFORMALIDAD LABORAL

En este capítulo se describen las fuentes de información utilizadas para la obtención de los datos necesarios sobre las variables que conforman la especificación del modelo de informalidad laboral en sus tres versiones, mismas que permiten el análisis del fenómeno de manera global así como desglosada. Además se realiza la descripción de dichas variables que poseen la característica de representar a nivel empírico algunos aspectos de las dos grandes posturas teóricas sobre la informalidad laboral, que previamente se presentaron. Finalmente se realiza un análisis gráfico de las variables, con el objetivo de brindar el escenario de su evolución histórica.

IV.1. Fuentes de Información

Los datos que se utilizan para estimar los modelos fueron obtenidos de distintas fuentes de información, entre las que destaca el Banco de Información Económica (BIE) del Instituto Nacional de Estadística y Geografía (INEGI), la Encuesta Mensual de la Industria Manufacturera (EMIM), Encuesta Industrial Mensual (EIM), Encuesta Nacional de Ocupación y Empleo (ENOE), los registros administrativos de las Finanzas Públicas Estatales y Municipales y de otras como el Tercer Informe de Gobierno 2014-2015, Consejo Nacional de la Población (CONAPO), Secretaría de Educación Pública (SEP) y el Instituto Nacional para la Evaluación de la Educación (INEE) . Todas las fuentes permiten conformar una base con datos en panel para 24 de las 32 entidades federativas de México en un horizonte temporal que va desde el año 2005 al año 2014.

La razón que justifica la consideración de solamente 24 entidades, se basa en que para la construcción de la variable Eficiencia en Recaudación (ER) que es de orden institucional, la disponibilidad de los datos para las 8 entidades restantes es muy limitada y no es posible contar con información suficiente que permita pronosticar los datos faltantes y así calcular la ER apropiadamente.

Más aun, el análisis no sería suficientemente apropiado si se trabaja con un panel desbalanceado, esto desde el punto de vista de comparación ya que uno de los objetivos principales del documento es poder comparar las variables duales contra las institucionales pero también entre sí mismas. En este sentido, lo que se intentó al utilizar un panel de datos balanceado fue minimizar o evitar el posible sesgo que pudiera surgir de la falta de datos para un cuarto de la muestra en

este caso para ocho de las 32 entidades, por lo que al ser balanceado la muestra de las 24 entidades contiene información completa para cada una de ellas. Cabe mencionar que esto no significa que el uso de paneles desbalanceados sea inadecuado, por el contrario es de gran utilidad, sin embargo debido para el objetivo de la investigación lo más apropiado resultó un panel balanceado. Por otro lado, las entidades no consideradas son: Baja California, Baja California Sur, Durango, Guerrero, Oaxaca, Querétaro, Tlaxcala y Veracruz.

Por lo tanto, dadas las características de la información, la metodología apropiada que permite cumplir con el objetivo de la investigación, es la estimación de un modelo de regresión con datos en panel como los descritos en el capítulo tres, seleccionando entre las diversas opciones el considerado como el mejor de acuerdo con lo que señalen las distintas pruebas formales.

IV.2. Especificación del Modelo con Datos en Panel

La especificación del modelo se basa en lo descrito en las secciones anteriores. El modelo estimado captura las versiones distintas, donde la diferencia la representa la variable que se utiliza como dependiente. La primera versión considera el total de la informalidad laboral sin realizar ninguna distinción por tipo de informalidad (modelo global), la segunda sólo considera la informalidad laboral desde el enfoque de la unidad económica; es decir, únicamente a los trabajadores informales independientes. La tercera versión tiene en cuenta sólo a la informalidad laboral desde el enfoque de las condiciones laborales o bien a los trabajadores dependientes informales. Tanto la estimación global como las desglosadas se basan en información anual de las 24 entidades para el periodo 2005-2014.

Las tres versiones del modelo considerando las posibilidades de estimación por las metodologías de regresión agrupada, efectos fijos y efectos aleatorios, se muestran en la tabla 1 a continuación:

Tabla 1: Especificaciones del Modelo de Informalidad Laboral

Metodología		
Regresión Agrupada	Efectos Fijos	Efectos Aleatorios
$IL_{it} = \alpha + D_{it}\beta + I_{it}\delta + u_{it}$	$IL_{it} = \alpha_i + D_{it}\beta + I_{it}\delta + u_{it}$	$IL_{it} = \alpha_1 + D_{it}\beta + I_{it}\delta + v_{it}$
$ILI_{it} = \alpha + D_{it}\beta + I_{it}\delta + u_{it}$	$ILI_{it} = \alpha_i + D_{it}\beta + I_{it}\delta + u_{it}$	$ILI_{it} = \alpha_1 + D_{it}\beta + I_{it}\delta + v_{it}$
$ILD_{it} = \alpha + D_{it}\beta + I_{it}\delta + u_{it}$	$ILD_{it} = \alpha_i + D_{it}\beta + I_{it}\delta + u_{it}$	$ILD_{it} = \alpha_1 + D_{it}\beta + I_{it}\delta + v_{it}$

Fuente: Elaboración propia.

La versión del modelo que contempla el total de la informalidad IL_{it} se especifica en el primer renglón de la tabla número 1. En el segundo renglón se encuentran las especificaciones para la informalidad laboral independiente ILI_{it} , y en el tercer renglón las especificaciones para la informalidad laboral dependiente ILD_{it} . En cuanto a las columnas, la primera corresponde a la especificación considerando la metodología de regresión agrupada, la segunda a la metodología de efectos fijos y en la tercera a la de efectos aleatorios. Las variables dependientes IL_{it} , ILI_{it} e ILD_{it} son utilizadas en tasas. Por su parte, en todas las especificaciones D_{it} es un vector conformado por variables explicativas relacionadas con el enfoque dual e I_{it} es un vector conformado por las variables explicativas asociadas a la perspectiva institucional. Más adelante en la sección V.1. se presentan estas ecuaciones para el total de informalidad (como referencia, aunque lo mismo aplica para el desglose de la informalidad independiente y dependiente) en términos de las variables implementadas.

IV.2.1 Descripción y Justificación de las Variables incluidas en las Especificaciones del Modelo

Es de gran importancia describir cada una de las variables utilizadas para la especificación de las tres versiones del modelo, con la finalidad de esclarecer qué variables conforman uno u otro enfoque de la informalidad laboral, así como la diferencia entre las tres especificaciones. Por esta razón, a continuación se presentan a detalle las variables dependientes y las explicativas duales e institucionales. Comenzando con la variable dependiente que es la informalidad, se utiliza en tres formas distintas: total, independiente y dependiente. La primera que es IL_{it} se refiere a la tasa de informalidad laboral total, construida a partir de los datos provistos en la ENOE bajo el nombre de TIL1 calculada contra la población ocupada. La razón para utilizar esta variable y no la Tasa de Ocupación en el Sector Informal TOSI1, se debe a que se considera que la TIL1 expresa mejor el fenómeno ya que tiene en cuenta que se puede presentar dentro y fuera del sector informal de la economía. Recuérdese que también en el sector formal hay individuos que operan en condiciones de informalidad laboral, por lo que si se utilizara la TOSI1 se estaría subestimando el tamaño real de la informalidad. En este sentido, para efectos de la presente investigación se considera que la TIL1 es la medida más adecuada. Finalmente en el modelo a estimar se le denomina TIL_{it} .

La informalidad laboral puede ser analizada a través de diferentes niveles de desagregación, por ejemplo, por sector de actividad económica, sexo, edad, niveles de escolaridad y tipo de empleo, por mencionar algunos. En la presente investigación es relevante estudiar la informalidad desagregándola por tipo de empleo, en independiente y dependiente. En la segunda versión del modelo, la variable a explicar cambia a ILL_{it} que representa a la informalidad laboral independiente, que considera uno de los dos tipos en los que puede dividirse desde un enfoque por tipo de empleo. La informalidad independiente se compone por todos aquellos individuos que son trabajadores por cuenta propia y empleadores, es decir, aquellos que no están subordinados a un jefe o no dependen de uno. Asimismo, ILL_{it} estará representada por la tasa de informalidad laboral independiente $TILI_{it}$.

En contraste a lo anterior, ILD_{it} que es la variable dependiente de la tercera versión del modelo se conforma por la informalidad laboral dependiente que captura a aquellos individuos cuyo empleo informal puede ser subordinado y remunerado o subordinado y no remunerado; en este caso los trabajadores sí dependen de un jefe. Esta variable se representa por la tasa de informalidad laboral dependiente $TILD_{it}$.

Por otra parte, la razón que justifica el desglose de la informalidad laboral total es que el análisis por tipo de empleo informal permitirá considerar las diferencias que existen dentro de ella, pues no todos los empleos ni los individuos son iguales, es decir, se podrá capturar la estructura heterogénea del mercado laboral informal en México. A nivel teórico también es de gran importancia la desagregación debido a que será posible determinar si las causas que influyen en las variaciones de los niveles de uno u otro tipo de informalidad son diferentes, es decir, duales o institucionales o por el contrario si estas causas son similares. El resultado esperado es que de igual manera que en el caso global, la informalidad independiente y dependiente se explique en parte por ambos enfoques teóricos.

Por otro lado, se encuentran las variables explicativas contempladas en los vectores D_{it} e I_{it} , es decir, las duales e institucionales respectivamente. El vector de las duales se compone por cuatro variables⁸, una es el Producto Interno Bruto (PIB) que se incluye para mostrar cómo afecta la

⁸ Sin embargo en la estimación final la variable NDES no se considera.

actividad económica a la informalidad laboral; la relación esperada es negativa porque que si la situación económica es favorable, entonces la informalidad laboral disminuye debido a que el sector moderno se encontrará en mejores condiciones para emplear una cantidad mayor de individuos, reduciéndose el número de aquellos que no logran obtener un empleo formal. Por el contrario, una situación económica desfavorable provocará incrementos en la informalidad, puesto que las actividades informales tienden incrementarse en periodos de baja actividad económica o de crisis.

Una segunda variable enmarcada en el enfoque dualista es la Escolaridad de la Población Económicamente Activa (PEA) que se denota en el modelo como ESCO, la cual indica los años promedio de escolaridad de la PEA. En este caso, si se parte del supuesto razonable de que el sector informal o secundario contempla los empleos menos calificados, entonces entre menor sea la escolaridad de los individuos, el sector informal crecerá más. Recuérdese también, que se ha demostrado que los individuos con elevados niveles de educación obtienen una ganancia neta de cambiarse al sector formal (Moreno, 2007). Bajo este contexto, la relación entre informalidad y escolaridad se espera que sea negativa, a mayor escolaridad de la PEA menores niveles de informalidad y al contrario.

Otra variable de carácter dual que se utiliza para dimensionar el mercado de trabajo, es el número de desempleados (NDES), que pretende mostrar que si la situación del mercado de trabajo no es propicia para captar toda la fuerza laboral buscadora de empleo, entonces la informalidad aumentará. Un estado no propicio es aquel en el que el NDES crece y como consecuencia las condiciones de búsqueda de empleo se tornan más difíciles, ante un mayor número de individuos compitiendo por el empleo, lo que se ve reflejado en un crecimiento de la informalidad como salida voluntaria, transitoria o permanente al desempleo. La relación esperada entre el logaritmo natural del número de desempleados en la economía y la informalidad, es positiva⁹.

Como cuarta variable del vector dual se propone el tamaño del Sector Moderno o formal (SMODERNO) medido a través de la población ocupada en la industria manufacturera como proporción de la población ocupada total. La industria manufacturera es considerada como el

⁹ Al momento de realizar la estimación del modelo no es posible incluir PIB y NDES en la misma ecuación debido a que conlleva a resultados inadecuados por la correlación que existe entre estas dos variables y ello se refleja en los coeficientes estimados. Por ello en la sección de anexos se presenta la estimación que considera PIB y NDES simultáneamente, sólo como referencia.

sector moderno debido a que se compone de actividades que son reconocidas por hacer uso de diversos niveles de tecnología, incluso de alta tecnología (industria aeroespacial, farmacéutica, automotriz, por ejemplo) y esta tendencia ha ido en crecimiento. Además es bien reconocido que la innovación tecnológica es uno de los factores que influyen en el crecimiento a largo plazo para una economía y con ello la población en términos generales se ve beneficiada. Por lo tanto, entre esta variable propuesta y la informalidad se espera una relación negativa, la lógica detrás de ello es que la informalidad se moverá de acuerdo con la capacidad de empleo del sector formal o moderno. Si la población ocupada en el sector moderno disminuye, entonces la capacidad de empleo en este también lo hará, por lo que la informalidad se incrementará debido al aumento del desempleo. En otras palabras, aumentará el número de individuos desempleados que se trasladarán al sector informal puesto que la teoría dual reconoce que sí existe movilidad entre los sectores. Por el contrario, un incremento de la capacidad de empleo en el sector moderno implica una disminución de la informalidad laboral.

En lo que concierne a las cuatro variables institucionales que conforman el vector I_{it} , la primera a describir es el Costo Laboral Unitario (CLU) en el sector moderno (industria manufacturera). Debido a que el enfoque institucionalista resalta que tanto individuos como empresas buscan la informalidad como una medida para escapar a las obligaciones formales, entonces es de esperarse que por el lado de la demanda de trabajo las empresas en la búsqueda de la minimización de sus costos, consideren de gran importancia el comportamiento del costo laboral al momento de demandar trabajo formal. Si el costo laboral en el sector moderno aumenta, las empresas demandan menos trabajo o reducen su oferta de empleos formales, con lo que la informalidad se puede incrementar mediante dos vías: la primera es que el aumento del desempleo de los formales los lleve a elegir trabajar en el sector informal para obtener una fuente de ingresos. La segunda vía, es que las empresas pueden ver como alternativa incrementar su demanda por trabajo informal que implica menores costos al evadir obligaciones que no pueden ser evadidas si el trabajador es contratado como formal. En resumen, la relación esperada entre el costo laboral y la informalidad es positiva, pues a mayores costos laborales mayor informalidad y viceversa.

El cálculo del CLU se realizó siguiendo a McConnell, Brue y Macpherson (2003) quienes lo realizan de la siguiente forma:

$$\frac{\text{Costo Laboral Unitario}}{\text{Costo Laboral Unitario}} = \frac{\text{Costos Laborales Totales}}{\text{Cantidad de Producción}} \quad [1]$$

Donde el numerador de la fórmula 1 es el salario nominal medio multiplicado por el total de horas de trabajo utilizadas. Sustituyendo:

$$\frac{\text{Costo Laboral Unitario}}{\text{Costo Laboral Unitario}} = \frac{\text{Salario por Hora X Horas de Trabajo}}{\text{Cantidad de Producción}} \quad [2]$$

Y dividiendo tanto numerador como denominador por las horas de trabajo:

$$\frac{\text{Costo Laboral Unitario}}{\text{Costo Laboral Unitario}} = \frac{\text{Salario por Hora}}{\text{Cantidad de Producción / Horas de Trabajo}} \quad [3]$$

Con base a la información disponible para la industria manufacturera, los elementos de la fórmula 3 son calculados de la siguiente forma: el salario por hora es aproximado al dividir el total de remuneraciones por el total de horas del personal ocupado. Las remuneraciones son todos los pagos y aportaciones en dinero y especie antes de cualquier deducción, realizados por el establecimiento en el periodo de referencia, para retribuir el trabajo del personal remunerado dependiente de la razón social, en forma de salarios, sueldos, prestaciones sociales y utilidades repartidas al personal, ya sea que este pago se calcule sobre la base de una jornada de trabajo o por la cantidad de trabajo desarrollado (destajo) o mediante un salario base que se complementa con comisiones por ventas u otras actividades¹⁰. Debido a que las remuneraciones consideran otros conceptos además de los salarios tales como: sueldos, aguinaldo, prima vacacional, utilidades, etc, esta es una variable proxy a la que se refieren McConnell, Brue y Macpherson (2003). Sin embargo, es una variable de suma importancia porque captura el costo total para el empleador, que no sólo se ve reducido a los salarios sino que también incluye los sueldos y demás prestaciones ya señaladas. Por su parte, para la producción se utilizó el total del valor de la producción y finalmente las horas de trabajo son el total de horas del personal ocupado.

¹⁰ Definición disponible en el glosario de la EMIM, www.inegi.org.mx.

Asimismo, para el enfoque institucional es muy importante el papel del gobierno debido a que a través de sus acciones e instituciones puede influir en las variaciones de los niveles de informalidad. Por esta razón, se propone como segunda variable institucional la eficiencia en recaudación (ER) que mide la proporción de ingresos por impuestos que se gastan en remuneraciones al personal de la Secretaría de Administración y Finanzas. La fórmula utilizada es:

$$\text{Eficiencia en Recaudación} = \frac{\text{Egresos en Remuneraciones al Personal de la Secretaría de Administración y Finanzas}}{\text{Ingresos por Impuestos}} \times 100 \quad [4]$$

Para la construcción de la variable ER se utilizó como referencia lo planteado en el IMCO (2012), en el apartado correspondiente al gobierno eficiente y eficaz, donde se calcula la eficiencia en recaudación como el resultado de la nómina para finanzas entre la recaudación de impuestos. Se menciona también que la eficiencia en recaudación calculada de esta manera representa el valor de la recaudación fiscal en relación a su costo administrativo. En otras palabras, se mide qué tan eficiente es el gasto para recaudar los impuestos. Entonces, a menor valor de ER mayor es la eficiencia del gobierno en este sentido.

La relación entre la ER y la informalidad puede ser positiva o negativa dependiendo de la óptica que tome el individuo. En general los individuos intentan determinar qué hace el Estado con los impuestos que recauda, es decir, si esa recaudación es eficiente y si obtienen algún beneficio de ella. El sentido de la eficiencia recaudatoria que aquí se maneja implica que una mayor eficiencia se presenta cuando los valores de ER son menores y por el contrario una menor eficiencia implica valores de ER mayores.

La explicación de una relación positiva entre la ER y la informalidad va en el sentido de que si el gobierno es más eficiente, la informalidad disminuye debido a que una mayor cantidad de impuestos pueden ser destinados a fortalecer instituciones que beneficien a los individuos, por ejemplo programas de salud, educación, seguridad pública, inversión pública, etcétera. En este caso los individuos elegirán permanecer en la formalidad dado que perciben que sus contribuciones tributarias mejoran la calidad de las instituciones gubernamentales y el beneficio es directo para ellos.

Por el contrario, si la relación entre la ER y la informalidad es negativa, significa que el individuo considera que esa eficiencia lo perjudica porque con menor costo de recaudación se están extrayendo mejor los impuestos, es decir, el individuo va a buscar evadir aún más las obligaciones fiscales puesto que su costo de permanecer formal se eleva si el gobierno es más eficiente para obtener los impuestos, lo que culmina en el aumento de la informalidad. Asimismo esta relación implica que desde la perspectiva de los individuos, los recursos recaudados no son destinados en su beneficio, por lo que una mayor eficiencia no es motivo para que decidan permanecer formales o salir de la informalidad. En síntesis, puede decirse que una buena administración pública disminuye la informalidad y que por el contrario una administración pública deficiente ejerce influencia en el aumento de los niveles de informalidad en el país.

La tercera variable que forma parte del vector de institucionales es el gasto en Inversión Pública que realiza el gobierno (IP). Se espera que esta mantenga una relación inversa con la informalidad debido a dos posibilidades. La primera es que la IP puede convertirse en un incentivo a permanecer en la formalidad o moverse hacia ella si se parte del supuesto de que los individuos consideran que los recursos gubernamentales están siendo invertidos en cuestiones que los benefician directamente. La segunda posibilidad es que los recursos que invierte el gobierno se vean reflejados en acciones que tienden a fomentar la formalidad o disminuir la informalidad. En conclusión, es de esperarse que a mayor inversión pública los niveles de informalidad sean menores y que por el contrario cuando la inversión pública disminuya, la informalidad laboral aumente. Cabe mencionar que debido a que esta variable no resultó significativa en ninguna de las estimaciones se decidió no considerarla.

Por último se propone como variable institucional, la cobertura del seguro popular (COBERTURA). La razón para incluirla es el intento por capturar la decisión voluntaria del individuo por pertenecer a la informalidad considerando que este programa social es de gran influencia sobre tal decisión. Recuérdese que la discusión de Anton et al. (2011) sugiere que la cobertura social es como una forma de subsidio para aquellos trabajadores que no poseen relaciones contractuales (en el presente caso se puede asumir que son los informales). El seguro popular representa un sustituto¹¹ de la atención médica que otorga un empleo formal, por lo que

¹¹ El seguro popular es un sustituto, pero se cuestiona el hecho de que sea uno bueno y la razón es que la cobertura de enfermedades por parte de este es limitada en comparación con la cobertura de instituciones como el IMSS o el ISSTE, por ejemplo.

aun estando informal el individuo puede acceder a esta atención médica que es menos costosa que la formal y que la privada. Esta afirmación, se puede sostener al revisar el trabajo de Santana (2011) quien determina que el bajo costo del Seguro Popular ocasiona que algunos individuos prefieran permanecer en la informalidad en lugar de afiliarse a un esquema como el del IMSS.

El costo por el servicio del seguro popular se maneja como una aportación familiar determinada con base al nivel de ingreso. El estudio socioeconómico que se realiza para determinar el costo por familia, no es tan riguroso, por lo que los individuos pueden reportar menores ingresos y con ello pagar una cuota mínima. En este sentido, si cada vez es mayor el número de individuos que pueden acceder al seguro popular entonces estos decidirán optar por la informalidad, debido a que recibirán atención médica que disminuye el costo de permanecer informal, con lo cual los niveles de informalidad se elevarán en el país. En Levy (2008) se pueden encontrar argumentos que refuerzan la idea de que los programas de protección social (como es el caso del Seguro Popular) inducen a los trabajadores a tomar un empleo informal¹². De esta manera, la relación esperada entre la COBERTURA y la informalidad es positiva, y la fórmula para calcularla es:

$$\text{Cobertura Seguro Popular} = \frac{\text{Número de Afiliados al Seguro Popular}}{\text{Población Total}} \times 100 \quad [5]$$

Por otro lado, es preciso señalar que las variables PIB, ESCO, NDES, CLU e IP son incluidas en términos de logaritmo natural, en tanto que las variables TIL, TILI, TILD, SMODERNO, ER Y COBERTURA se encuentran en términos porcentuales. Por lo tanto, los coeficientes obtenidos de las estimaciones son semielasticidades para el PIB, ESCO, NDES, CLU e IP y son elasticidades para SMODERNO, ER Y COBERTURA. Para hacer comparables los resultados, las semielasticidades se convirtieron a elasticidades¹³ y solamente se realizó la conversión para

¹² Levy (2008) se concentra en los efectos de la interfase entre la seguridad social y los programas de protección social sobre la productividad de los trabajadores (su disminución), distinguiendo entre los trabajadores pobres y los no pobres, argumentando que el efecto se concentra sobre los primeros ya que los programas inducen a que más trabajadores pobres elijan un empleo informal en comparación con los trabajadores no pobres.

¹³ La metodología considerada para convertir las semielasticidades a elasticidades se basa en lo propuesto en Gujarati y Porter (2010), como a continuación:

$$\text{La elasticidad, implica que: } \beta = \frac{\text{cambio relativo en Y}}{\text{cambio relativo en X}} = \frac{Y_t - Y_{t-1} / Y_{t-1}}{X_t - X_{t-1} / X_{t-1}} = \frac{\Delta Y / Y}{\Delta X / X} = \frac{\Delta Y}{\Delta X} \frac{X}{Y} \quad [1]$$

las variables que considera el modelo elegido como el mejor, es decir, para las variables PIB, ESCO, SMODERNO, CLU, ER y COBERTURA, obteniendo todos los coeficientes a interpretar en elasticidades. En la tabla número 2 se presenta un resumen de las variables explicativas y de los signos esperados de los coeficientes asociados a cada una de ellas con relación a la informalidad desde ambas posturas teóricas.

Tabla 2: Variables Explicativas Duales (D_{it}) e Institucionales (I_{it})

D_{it}		I_{it}	
Variable	Signo Esperado	Variable	Signo Esperado
PIB	-	CLU	+
ESCO	-	ER	+ o -
NDES	+	IP	-
SMODERNO	-	COBERTURA	+

Fuente: Elaboración propia.

Además de las seis variables de la tabla número 2, se consideró el uso de dos variables instrumentales en el contexto de la estimación por la metodología de VI debido a que una gran variedad de autores señalan que cuando se incluye la variable escolaridad en alguna especificación, es muy probable que esta no sea exógena, es decir se encuentre correlacionada con el término de error. Precisamente este problema se resuelve por el método de VI. Sin embargo, debe comprobarse que en tal especificación esté presente la endogeneidad mediante algunas pruebas formales y a partir de ello justificar el uso de VI.

A priori se reconoce que generalmente la educación resulta endógena cuando se estiman ecuaciones salariales (o de retornos de la educación), pues la educación depende a su vez de otras habilidades no observadas de los individuos, dichas habilidades también influyen en el salario y quedan capturadas en el término de error de la ecuación y entonces es así como surge la endogeneidad que permite el uso de VI. Sin embargo, la ecuación que aquí se plantea no

$$\text{Por su parte, la semielasticidad sugiere que: } \beta = \frac{\text{cambio absoluto en Y}}{\text{cambio relativo en X}} = \frac{Y_t - Y_{t-1}}{X_t - X_{t-1} / X_{t-1}} = \frac{\Delta Y}{\Delta X / X} \quad [2]$$

Entonces para convertir una semielasticidad a elasticidad, se sugiere multiplicarla por $\left(\frac{1}{Y}\right)$, donde Y es el valor

$$\text{medio de Y o bien } \bar{Y}. \text{ De tal manera, se tiene: } \left(\frac{\Delta Y}{\Delta X / X}\right)\left(\frac{1}{Y}\right) = \left(\frac{\Delta Y X}{\Delta X}\right)\left(\frac{1}{Y}\right) = \frac{\Delta Y X}{\Delta X Y} \quad [3]$$

De esta forma, la parte final de la ecuación 1 es igual a la parte final de la ecuación 3, lo que demuestra que es posible pasar de una semielasticidad como la de la ecuación 2, a una elasticidad como la de la ecuación 1, mediante el procedimiento en la ecuación 3.

corresponde al tipo que se señala por lo que debe comprobarse formalmente si se encuentra presente la endogeneidad o no.

Con el objetivo de elegir el mejor modelo se realiza también la rutina de estimación por la metodología de VI para resolver el posible problema de endogeneidad en caso de que se encuentre presente. A nivel empírico es común que se tomen como instrumentos de la escolaridad variables como la educación de los padres de los individuos, la distancia a la escuela más cercana, los costos de matrícula y la educación rezagada, por mencionar algunos¹⁴.

A nivel entidad no es posible contar con variables similares a las mencionadas, por lo que se decide tomar como instrumentos el número de escuelas por estado (NESCUELAS) y la tasa de deserción a nivel secundaria (TDS). Se asume que el número de escuelas por entidad explica la escolaridad porque es una forma de medir el acceso a la educación, es decir, a mayor número de escuelas mayor será el número de individuos que pueden acceder a la educación desde esta perspectiva, con lo que los niveles de escolaridad de la población pueden verse sustancialmente incrementados. En cuanto a la tasa de deserción a nivel secundaria, es de gran relevancia en la explicación de ESCO debido a que el promedio de años de escolaridad a nivel nacional es de alrededor de 9 años que justamente abarca hasta el nivel secundaria, entonces si la deserción a nivel secundaria aumenta ello impactará los niveles de escolaridad no permitiendo que estos sobrepasen los 9 años en promedio. A mayor deserción menores serán los años promedio de escolaridad de la población.

IV.3. Análisis Gráfico de las Variables Incluidas en el Modelo

Previo a la presentación de los resultados de las estimaciones econométricas es conveniente realizar un análisis gráfico, de las variables duales e institucionales incluidas en las especificaciones del modelo. Esto permite ofrecer un panorama de su desempeño en el tiempo y su relación con la informalidad, tomando como referentes los datos nacionales debido a que hacerlo por entidad complica el análisis gráfico. Sin embargo, resulta que para algunos pares de variables no son tan evidentes las relaciones de interés, y ello puede deberse a que el periodo no es tan extenso y a la presencia de una crisis económica.

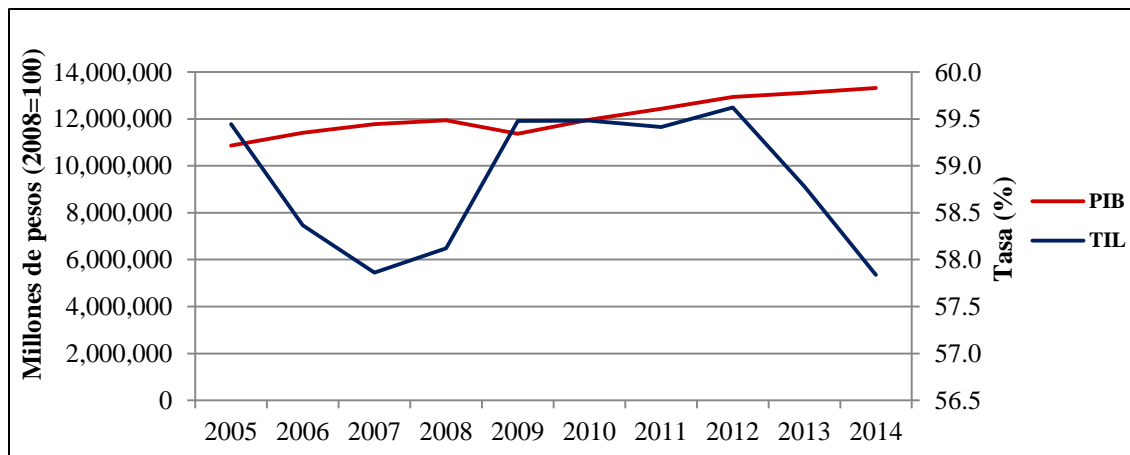
¹⁴ Algunos documentos en los que se utilizan variables instrumentales para la educación son por ejemplo el de Bhatti, Bourdon y Aslam (2013) y el de Card (1993).

IV.3.1. Análisis Gráfico de las Variables Duales (D_{it})

IV.3.1.1. PIB

En la gráfica número 5 se ilustran las variables PIB y TIL cuya relación es inversa. Cuando el PIB se incrementa en el periodo 2005-2008 la TIL tiende a disminuir, alcanzando uno de los dos valores más bajos para todo el periodo. En cambio, cuando el PIB disminuye como ocurrió en 2009, la TIL pasa de 58.1% en 2008 a 59.5% en 2009, es decir, se incrementó en 1.4% en un escenario adverso para la economía. Se puede argumentar que un comportamiento favorable de la economía tiende a disminuir la informalidad porque las condiciones son propicias para que el sector formal requiera de un mayor número de mano de obra. Por el contrario, el comportamiento desfavorable provoca el incremento como resultado de la expulsión de individuos del sector formal, quienes escapan al desempleo empleándose en una actividad informal.

Gráfica 5: Producto Interno Bruto (PIB) y Tasa de Informalidad Laboral (TIL)



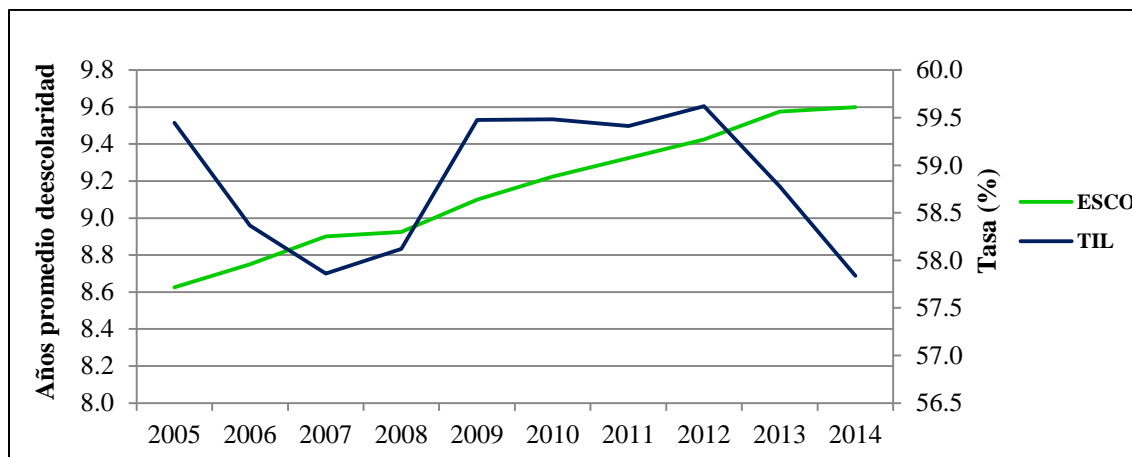
Fuente: Elaboración propia con datos de la ENOE y del BIE-INEGI.

IV.3.1.2. Escolaridad

Las variables ESCO y TIL se muestran en la gráfica número 6, sin embargo la relación negativa que se pretende evidenciar no es tan clara puesto que la escolaridad de la PEA es ascendente y en ningún momento descende. Por otra parte, si se toma en cuenta el periodo 2005-2007, es evidente que a mayor escolaridad menor es la informalidad laboral, lo mismo sucede en el periodo 2012-2014 donde la TIL muestra una tendencia a la baja y ESCO a incrementarse. Es decir, TIL y ESCO se mueven en sentido opuesto en algunos periodos. Los hechos confirman que siempre es benéfico el incremento de los niveles de escolaridad en el país, en este caso implica

que sea menor el número de individuos que terminan empleándose en la informalidad. Lo que permitiría que mayor parte de la PEA acceda a empleos con mejores condiciones laborales.

Gráfica 6: Escolaridad de la PEA (ESCO) y Tasa de Informalidad Laboral (TIL)

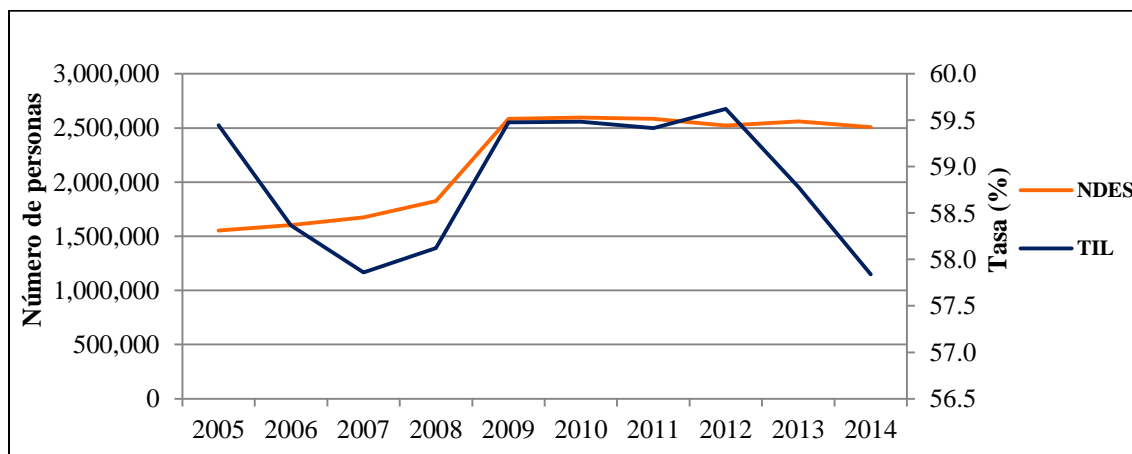


Fuente: Elaboración propia con datos de la ENOE.

IV.3.1.3. Desempleo

El desempleo se representa a través de la variable NDES. El comportamiento de la TIL y del NDES se muestra en la gráfica número 7. Se observa que se mueven en el mismo sentido, siendo evidente la relación positiva. Un hecho interesante es que en el periodo 2009–2011 cuando NDES muestra variaciones casi imperceptibles, la TIL también, es decir, la curva de ambas variables es casi plana. Entonces, puede afirmarse que la situación que enfrenta el mercado laboral sí afecta directamente a la informalidad. Cuando el desempleo crece hay un mayor número de individuos que buscan empleo y aquellos que no logran emplearse en la formalidad, lo hacen en lo informal.

Gráfica 7: Número de Desempleados (NDES) y Tasa de Informalidad Laboral (TIL)

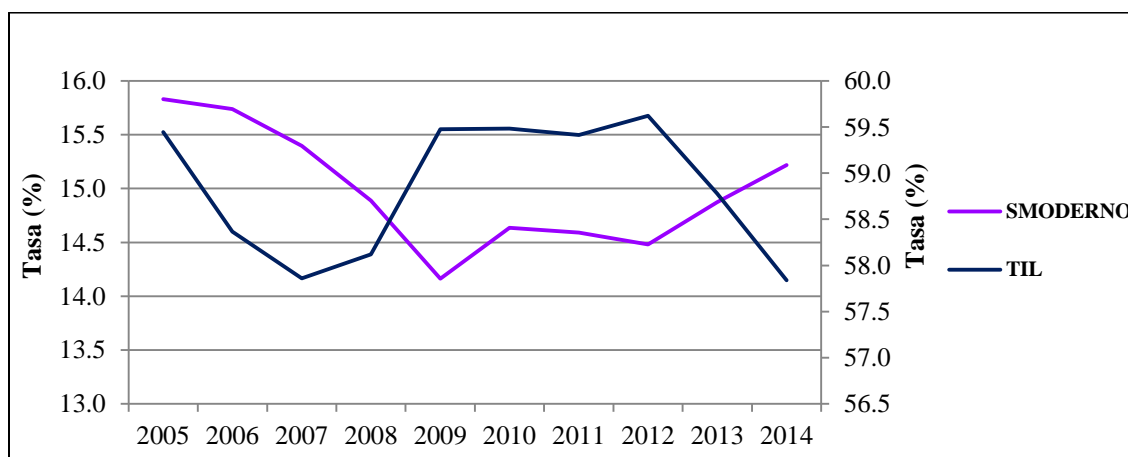


Fuente: Elaboración propia con datos de la ENOE.

IV.3.1.4. Sector Moderno

El tamaño del sector moderno y la TIL se describen en la gráfica número 8. La relación entre las variables no es lo suficientemente clara en todo el periodo, pero sí permite apreciar que cuando el sector moderno crece como proporción de la población total, la TIL disminuye y cuando SMODERNO disminuye, la TIL aumenta. El primer caso se puede constatar claramente en el periodo 2012-2014 donde SMODERNO aumenta y la TIL disminuye. El segundo caso se presenta, por ejemplo, en el periodo 2007-2009 donde la TIL exhibe un crecimiento considerable al pasar de 57.9% a 59.5%. Por su parte, la variable SMODERNO se reduce de 15.4% en 2007 a 14.2% en 2009. Por lo tanto, estas son algunas regularidades empíricas que corroboran la relación negativa esperada entre SMODERNO y TIL a nivel teórico para el enfoque dual. Se determina que cuando la capacidad de empleo del SMODERNO se contrae, entonces la informalidad crece y viceversa.

Gráfica 8: Sector Moderno (SMODERNO) y Tasa de Informalidad Laboral (TIL)



Fuente: Elaboración propia con datos de la ENOE.

IV.3.2. Análisis Gráfico de las Variables Institucionales (I_{it})

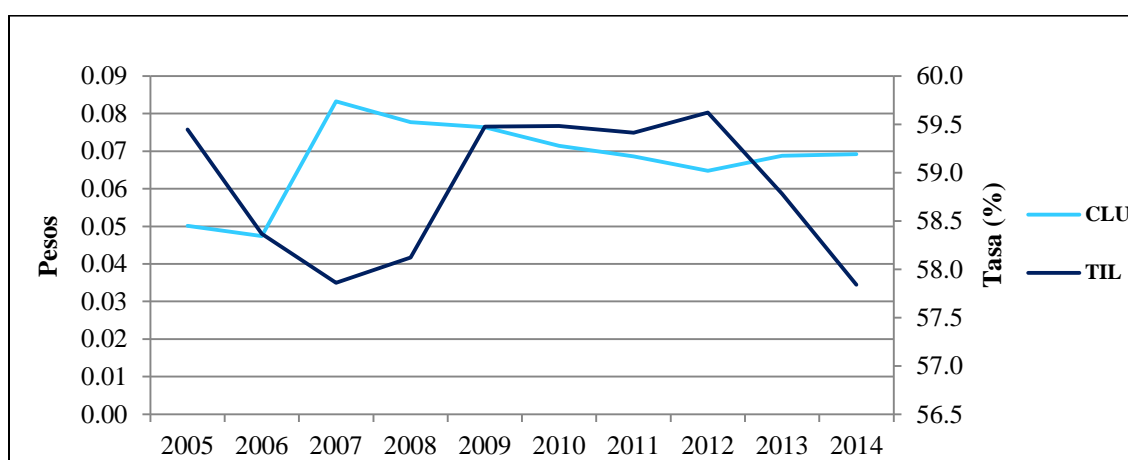
IV.3.2.1. Costo Laboral Unitario

Una medida propuesta para el costo laboral es el Costo Laboral Unitario (CLU) en la industria manufacturera, que debe relacionarse positivamente con la TIL. Un mayor costo implica que las tasas de informalidad serán mayores y viceversa. Las variables CLU y TIL se exhiben en la gráfica número 9, donde la relación directa no es tan evidente en todo el periodo. Sin embargo, si

el análisis se concentra por ejemplo en el periodo que va de 2009 a 2011, es posible observar que el CLU disminuyó y que la TIL también lo hizo.

El hecho de que en algunos periodos la relación entre el CLU y la TIL sea diferente a lo esperado, puede ser explicado argumentando que esto constituye una evidencia muy sencilla de que el efecto que posee el costo laboral sobre la TIL es muy pequeño, o bien, la TIL responde débilmente a las variaciones del CLU. También puede deberse a que además del CLU, otros factores determinan los niveles de informalidad observados en el país.

Gráfica 9: Costo Laboral Unitario (CLU) y Tasa de Informalidad Laboral (TIL)



Fuente: Elaboración propia con datos de la EIM, EMIM y ENOE.

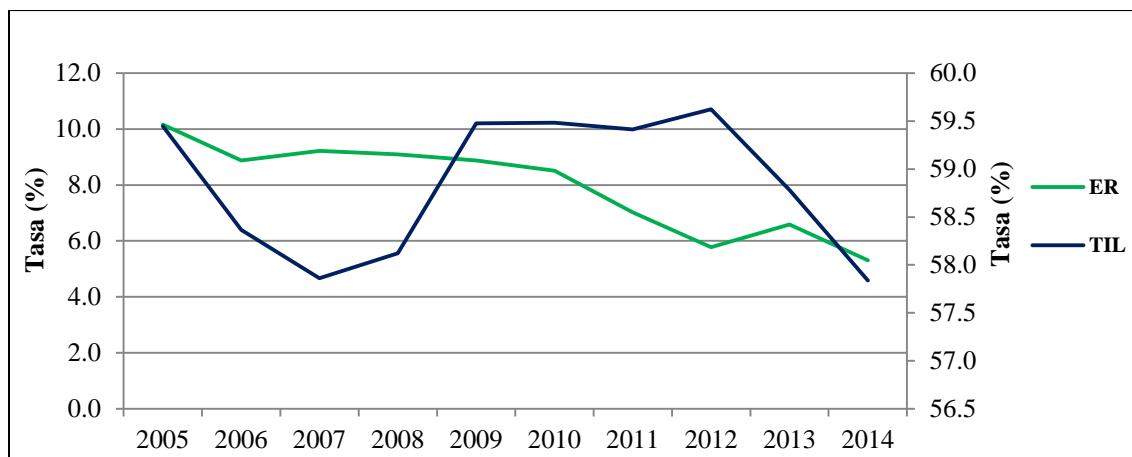
IV.3.2.2. Eficiencia en Recaudación

La relación entre la ER y la TIL es quizá la más compleja de detectar a simple vista para todo el periodo de análisis, a priori se asume que la relación puede ser positiva o negativa, sin embargo si se toman periodos cortos se corrobora que a mayor ER (valores más pequeños de esta variable) la informalidad disminuye. Por ejemplo, en la gráfica número 10 en el periodo de 2005 a 2006 la ER mejoró al pasar de 10.2% a 9.2% (recuérdese que un menor valor de la ER significa que el gobierno es más eficiente) y en este mismo sentido la TIL pasó de 59.4% a 58.4%. De 2010 a 2011 la ER pasa de un valor de 8.5% a 7.0% provocando que la TIL se mueva de 59.5% a 59.4%. Lo mismo ocurre de 2013 a 2014 pues la ER pasa de 6.6% a 5.3% y la TIL disminuye de 58.8% a 57.8%.

Como contraparte de lo anterior, la relación negativa también se encuentra presente, por ejemplo de 2006 a 2007 la eficiencia disminuyó al pasar de 8.9% a 9.2%, mientras que la TIL varía de

58.4% a 57.9%. Algo similar sucede de 2012 a 2013 al empeorar la ER de 5.8% a 6.6% y al disminuir la TIL de 59.6% a 58.8%. Los resultados del análisis gráfico coinciden con lo esperado, es decir, la eficiencia en la recaudación puede incrementar o disminuir la informalidad en el país dependiendo de cómo sea interpretada esa eficiencia por parte de los individuos que pagan impuestos. También destaca la tendencia hacia la baja en todo el periodo de la curva correspondiente a ER, lo que significa que la eficiencia gubernamental en términos de la recaudación de impuestos ha ido mejorando año tras año, salvo algunos casos como en 2007 y 2013 donde se presentan incrementos, pero el repunte más notorio se suscita en 2013.

Gráfica 10: Eficiencia en Recaudación (ER) y Tasa de Informalidad Laboral (TIL)

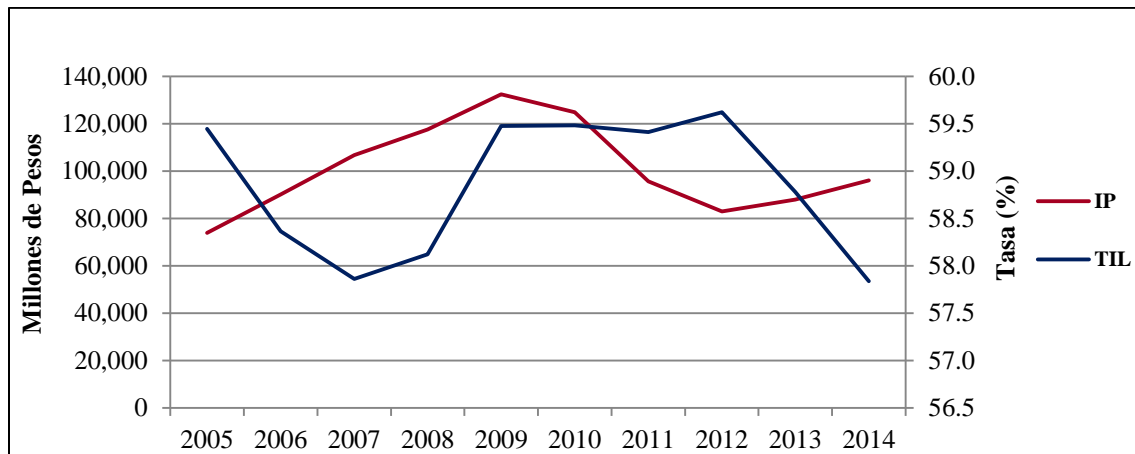


Fuente: Elaboración propia con datos de las Finanzas Públicas Estatales y Municipales. INEGI. ENOE.

IV.3.2.3. Inversión Pública

La relación entre la IP y la TIL se muestra en la gráfica número 11 y exhibe una relación negativa que indica que conforme se incrementa la inversión pública la informalidad disminuye y viceversa. Obsérvese que el periodo 2005-2008, donde la IP se incrementa considerablemente y la informalidad cae de 59.4% a 58.1%, corresponde al periodo previo a la crisis para el cual resulta muy interesante que se caracterice por incrementos sostenidos de la IP. Lo que sugiere que en tiempos de relativa estabilidad económica, el incremento en la IP ayuda a combatir los niveles de informalidad. En cambio, la IP desciende en el periodo 2009-2012 y la TIL pasa de 58.1% a 59.6%. Es decir, durante y después de la crisis donde el gobierno no apostó por la IP, la informalidad creció. Entonces, la IP trae consigo beneficios sobre la situación del mercado laboral en el país, al contribuir en la disminución de los niveles de informalidad.

Gráfica 11: Inversión Pública (IP) y Tasa de Informalidad Laboral (TIL)

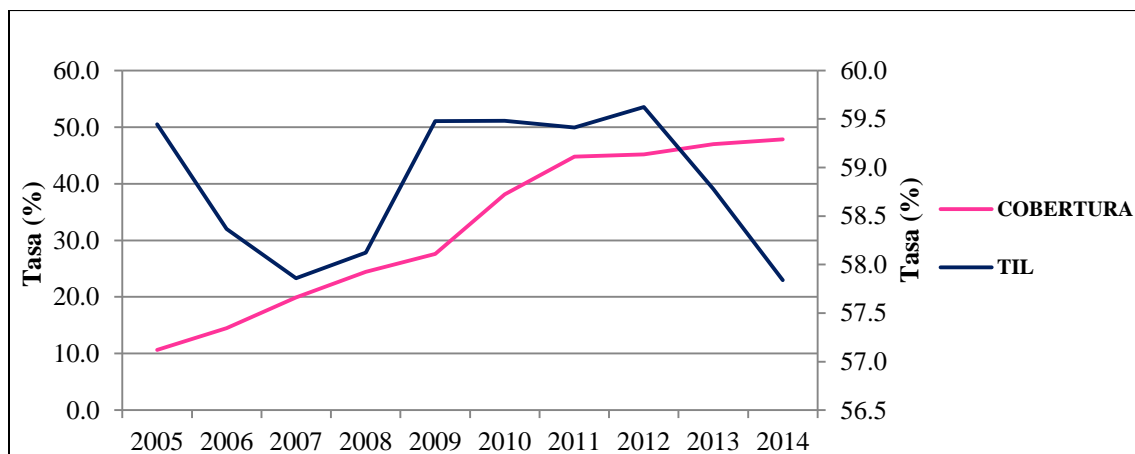


Fuente: Elaboración propia con datos de las Finanzas Públicas Estatales y Municipales. INEGI. ENOE.

IV. 3.2.4. Cobertura del Seguro Popular

La evolución de la COBERTURA y de la TIL se muestra en la gráfica número 12. El enfoque institucional sugiere que la relación debe ser positiva, y en términos generales así es, pues ambas variables a partir de 2007 y hasta 2012 exhiben un comportamiento ascendente. Cabe resaltar que la cobertura del seguro popular siempre tiende a incrementarse ya que es el objetivo de los programas sociales, además se extiende a la población en general y con pocas restricciones. De aquí que si un individuo no cuenta con atención médica de un empleo formal, él y su familia pueden estar cubiertos al afiliarse al seguro popular.

Gráfica 12: Cobertura del Seguro Popular (COBERTURA) y Tasa de Informalidad Laboral (TIL)



Fuente: Elaboración propia con datos del Tercer Informe de Gobierno 2014-2015 y estimaciones y proyecciones del CONAPO. ENOE.

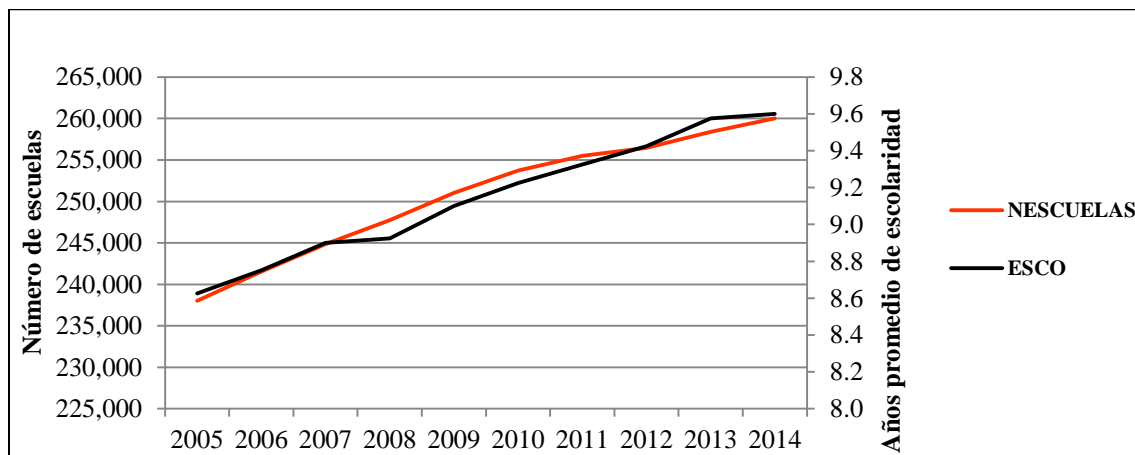
IV.3.3. Análisis Gráfico de las Variables Instrumentales

Además del análisis de las variables duales e institucionales, es pertinente observar el comportamiento de las variables que sirven como instrumentos de la variable dual ESCO, que puede ser endógena en el modelo. A continuación se muestra el desempeño de los instrumentos NESCUELAS y TDS por separado en el periodo 2005-2014. Las gráficas incluyen a ESCO con el objetivo de determinar si la relación entre esta y sus instrumentos se ajusta a lo planteado.

IV. 3.3.1. Número de Escuelas

El número de escuelas es un factor muy importante en la explicación de la escolaridad de la población, debido a que proporciona una idea de la situación del acceso a la educación. En la gráfica número 13 se muestra la evolución temporal de la variable instrumental NESCUELAS y de la posiblemente “endógena” ESCO, de la que se puede deducir que comparten una relación positiva en la que a medida que se incrementa el número de escuelas también lo hace la escolaridad. Esta relación es creciente y constante, pues tanto NESCUELAS como ESCO son variables que siempre tienden a crecer de un periodo a otro, aun cuando el cambio sea pequeño.

Gráfica 13: Número de Escuelas (NESCUELAS) y Escolaridad de la PEA (ESCO)



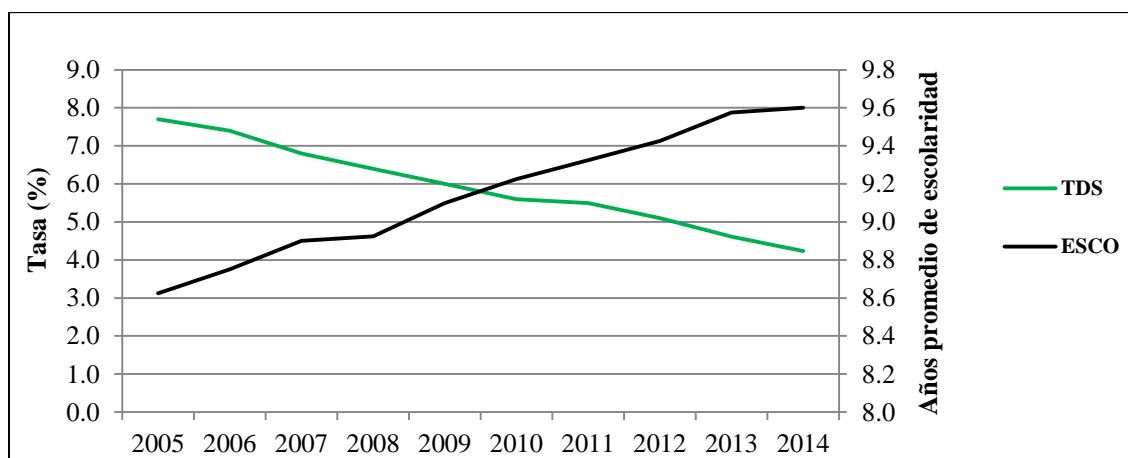
Fuente: Elaboración propia con datos de SEP y ENOE.

Puede concluirse, que entre mayor es el acceso a la educación mayores son los niveles de escolaridad de la población. En este sentido, las acciones gubernamentales que fomentan la construcción de escuelas son adecuadas ya que impactan positivamente sobre los años promedio de escolaridad y además tienen el poder de incidir en la disminución de la informalidad por esta vía.

IV. 3.3.2. Tasa de Deserción en Secundaria

La relación entre TDS y ESCO se plasma en la gráfica número 14 y precisamente coincide con lo planteado en la sección que describe las relaciones entre las variables, es decir, a menor tasa de deserción en la secundaria mayor será la escolaridad de la población y viceversa. Nótese que en todo el periodo la TDS siempre disminuye, lo que es un buen indicador de que la educación en el país avanza e impacta el número de años promedio de escolaridad (variable ESCO). La relación entre ambas variables tiene un sentido lógico, pues la tasa de deserción considerada corresponde a la del nivel secundaria y esto coincide con el número de años promedio de escolaridad que ha logrado alcanzar la población, es decir poco más de 9 años, que justamente abarca el nivel de secundaria.

Gráfica 14: Tasa de Deserción en Secundaria (TDS) y Escolaridad de la PEA (ESCO)



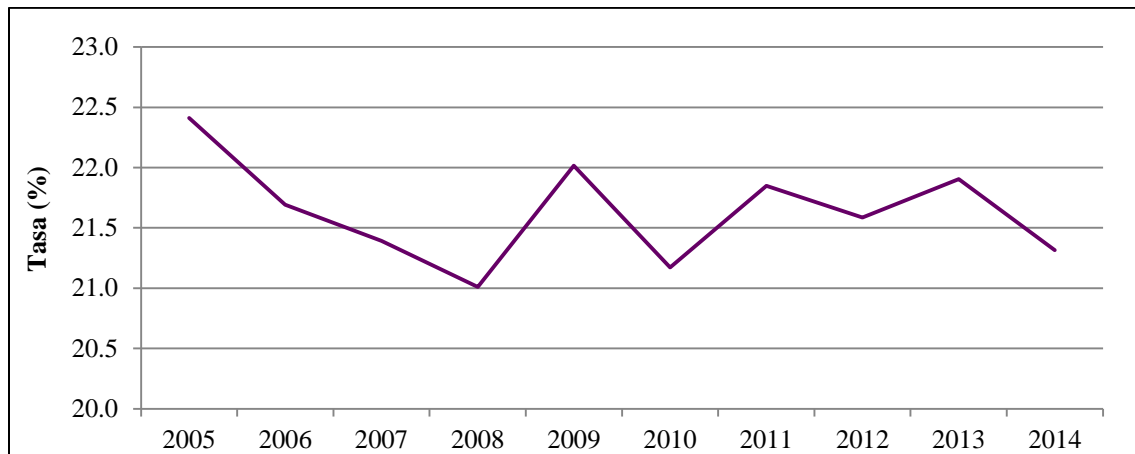
Fuente: Elaboración propia con datos del INEE y ENOE.

IV.3.4. Análisis Gráfico de las Variables Dependientes TILI Y TILD

IV.3.4.1. Tasa de Informalidad Laboral Independiente

El comportamiento de la TILI se presenta en la gráfica número 15. En promedio, la informalidad independiente es de 21.6% en todo el periodo. De 2005 a 2008 muestra un comportamiento descendente muy significativo al pasar de 22.4% a 21.0%. Sin embargo, en 2009 el valor de la TILI se incrementa a 22.0% y muestra variaciones positivas y negativas, pero sin regresar al valor más bajo que es el reportado en 2008. La crisis económica iniciada a finales de este año puede ser una explicación para el incremento en los niveles de informalidad de este tipo, también del tipo dependiente (gráfica número 16) y como consecuencia, del total (gráfica número 1).

Gráfica 15: Tasa de Informalidad Laboral Independiente (TILI)

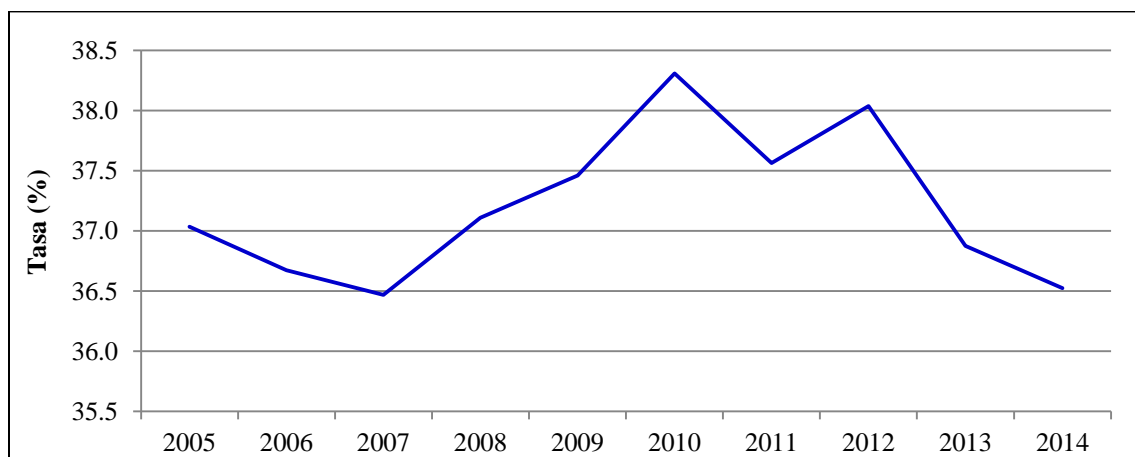


Fuente: Elaboración propia con datos de la ENOE.

IV.3.4.2. Tasa de Informalidad Laboral Dependiente

El comportamiento de la informalidad dependiente se reporta en la gráfica número 16. El valor promedio para esta variable en el periodo es de 37.2%, que en comparación con la TILI es superior en 15.6 puntos porcentuales. Es evidente que la TIL se compone mayormente por la informalidad del tipo dependiente o bien un mayor número de individuos son informales subordinados a un patrón.

Gráfica 16: Tasa de Informalidad Laboral Dependiente (TILD)



Fuente: Elaboración propia con datos de la ENOE.

Es interesante notar que al principio del periodo, es decir de 2005 a 2007 la tendencia de la TILD sea negativa al igual que para la TILI. Sin embargo, la diferencia se encuentra en que la TILD

alcanza su valor más bajo en 2007 al ser 36.4%; en cambio la TILI lo alcanza en 2008 y es igual a 21.0%. El valor máximo para la TILD es de 38.3% en 2010, en tanto que para TILI es de 22.4% en 2007. Tanto la TILI como la TILD casi logran regresar a su valor más bajo en 2014.

CAPÍTULO V: PRESENTACIÓN, ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

En este capítulo se presentan los resultados obtenidos a través de las diversas estimaciones realizadas con base a la metodología de panel de datos. La selección del mejor modelo se sustenta aplicando pruebas formales a cada uno de ellos, eligiendo aquel cuyas pruebas así lo indican. Una vez que se elige el mejor modelo, se procede a analizar los resultados en el marco de la hipótesis del trabajo y las preguntas de investigación previamente formuladas, teniendo en consideración las posturas teóricas que ofrecen la explicación del fenómeno de interés así como el objetivo general de la investigación. Por último se interpretan los resultados.

V.1. Regresión Agrupada vs Efectos Fijos vs Efectos Aleatorios

En un primer ejercicio de selección se consideraron ocho variables como explicativas: Producto Interno Bruto (PIB), Escolaridad (ESCO), Número de Desempleados (NDES), Sector Moderno (SMODERNO), Costo Laboral Unitario (CLU), Eficiencia en Recaudación (ER), Inversión Pública (IP) y Cobertura del Seguro Popular (COBERTURA). Como variable dependiente se utiliza la Tasa de Informalidad Laboral (TIL). Con esta información fueron estimadas tres regresiones con datos en panel: 1) una regresión agrupada, 2) con efectos fijos y 3) con efectos aleatorios.

La ecuación 1 corresponde a la regresión agrupada. Donde α es un intercepto común para toda $i = 1, 2, 3, \dots, 24$ entidades y $t = 2005, 2006, \dots, 2014$. Por su parte, β_1 a β_8 son coeficientes asociados a cada una de las variables explicativas del modelo y reflejan factores relacionados tanto con el enfoque dual (PIB, ESCO, NDES y SMODERNO) como con el institucional (CLU, ER, IP y COBERTURA). El término de error se representa por u_{it} .

$$TIL_{it} = \alpha + \beta_1 PIB_{it} + \beta_2 ESCO_{it} + \beta_3 NDES_{it} + \beta_4 SMODERNO_{it} + \beta_5 CLU_{it} + \beta_6 ER_{it} + \beta_7 IP_{it} + \beta_8 COBERTURA_{it} + u_{it} \quad [1]$$

Por otro lado, en la metodología de efectos fijos el intercepto α es diferente para cada entidad y esto es representado agregando un subíndice i , con lo que el intercepto es α_i . Debe recordarse que la metodología de efectos fijos y de efectos aleatorios forman parte de los modelos de componente de error, donde el error incluye el efecto no observado c_i además de u_{it} y es

definido como $v_{it} = c_i + u_{it}$. Sin embargo la estimación por efectos fijos implica la eliminación de c_i . Bajo estas consideraciones, la ecuación 1 se transforma como en la ecuación 2.

$$TIL_{it} = \alpha_i + \beta_1 PIB_{it} + \beta_2 ESCO_{it} + \beta_3 NDES_{it} + \beta_4 SMODERNO_{it} + \beta_5 CLU_{it} + \beta_6 ER_{it} + \beta_7 IP_{it} + \beta_8 COBERTURA_{it} + u_{it} \quad [2]$$

La regresión por efectos aleatorios difiere de la agrupada y de la de efectos fijos tanto en el intercepto como en el error. En efectos aleatorios el intercepto es α_1 y la característica principal de este es que es un valor medio, por lo tanto se dice que las unidades de análisis tienen una media común para el intercepto. El término de error es v_{it} y se debe a que en efectos aleatorios se asume que c_i no se encuentra correlacionado con ninguna de las variables explicativas, por lo que no es necesario aplicar alguna técnica para removerlo de la ecuación. El modelo de efectos aleatorios es:

$$TIL_{it} = \alpha_1 + \beta_1 PIB_{it} + \beta_2 ESCO_{it} + \beta_3 NDES_{it} + \beta_4 SMODERNO_{it} + \beta_5 CLU_{it} + \beta_6 ER_{it} + \beta_7 IP_{it} + \beta_8 COBERTURA_{it} + v_{it} \quad [3]$$

A continuación en las tablas número 3, 4 y 5 se muestran los resultados de la estimación de una regresión agrupada, considerando efectos fijos y efectos aleatorios, respectivamente. Cabe hacer la aclaración que indistintamente todas las estimaciones son robustas ante la presencia de problemas de autocorrelación y heterocedasticidad en virtud de que se estiman instrumentando la opción *cluster*¹⁵, que los corrige.

Los resultados de la regresión agrupada se muestran en la tabla número 3 y se observan dos cosas importantes, en primer lugar, tres de los coeficientes no son significativos al 5% o al 10% y corresponden al PIB, al CLU y a la IP; en segundo, la relación esperada entre PIB y TIL no se cumple y lo mismo sucede con CLU. Se puede concluir que la regresión agrupada estimada por MCO no es el mejor modelo, puesto que algunos resultados no coinciden con lo que postula la teoría, lo que constituye una razón importante para no validarlo. Ello sugiere, que se deben

¹⁵ Es un tratamiento para los errores estándar asociado a una variación del comando robust que es desarrollado por Eicker F., Huber P. y White H. La implementación de este comando se realiza de acuerdo con los procedimientos de estimación encontrados en Wooldridge (2015).

explorar otras formas de especificación, más allá de las que indiquen las pruebas estadísticas de selección del modelo.

Tabla 3: Resultados de la Regresión Agrupada

Dependiente: TIL			
Variable	Coefficiente	Error Estándar	Estadístico t
Intercepto	220.6786*	19.6198	11.25
PIB	0.1037	1.3155	0.08
ESCO	-83.7633*	9.6952	-8.64
NDES	2.5613*	0.9815	2.61
SMODERNO	-0.6297*	0.1913	-3.29
CLU	-1.3720	1.4856	-0.92
ER	-0.0476*	0.0228	-2.09
IP	-0.2898	0.5822	-0.50
COBERTURA	0.1817*	0.0393	4.62

$R^2 = 0.8203$

* Coeficiente significativo al 5%.

**Significativo al 10%.

Fuente: Resultados de estimación.

En razón de lo anterior, en la tabla número 4 se ilustran los resultados del modelo con efectos fijos¹⁶, observándose que tanto las relaciones del modelo como la significancia estadística mejoran en algunas variables. Por ejemplo, el coeficiente asociado al PIB arroja un signo correcto (negativo) además de ser estadísticamente significativo al 10%. De igual manera sucede con el CLU ya que su signo también es el apropiado (positivo) y además resulta ser una variable significativa al 10%.

Sin embargo, en esta estimación la IP continúa presentando un signo contrario a lo esperado (es positivo cuando debería ser negativo) y no es estadísticamente significativa. Adicionalmente, surge otro problema, en este caso con el coeficiente de la variable COBERTURA ya que se torna estadísticamente no significativo aunque continua reportando un signo correcto. Por estas razones, en términos generales se puede afirmar que el panel estimado por efectos fijos

¹⁶ Nótese que la estimaciones por efectos fijos o por efectos aleatorios proveen tres valores para el estadístico R^2 . Entre estos el que debe considerarse en el caso de efectos fijos es el R^2 Whitin y en el de efectos aleatorios es el R^2 Overall.

proporciona mejores resultados en comparación con el panel estimado para datos agrupados con un mismo intercepto.

Tabla 4: Resultados de la Regresión con Efectos Fijos

Dependiente: TIL			
Variable	Coefficiente	Error Estándar	Estadístico t
Intercepto	125.3560*	23.5394	5.33
PIB	-3.1138**	1.5521	-2.01
ESCO	-29.9375*	7.8400	-3.82
NDES	3.9177*	0.6639	5.90
SMODERNO	-0.3233*	0.1104	-2.93
CLU	0.4209**	0.2285	1.84
ER	-0.0098*	0.0040	-2.41
IP	0.0295	0.1579	0.19
COBERTURA	0.0275	0.0177	1.55

R² within = 0.4477 R² between = 0.6626 R² overall = 0.6465
 * Coeficiente significativo al 5% .
 **Significativo al 10%.

Fuente: Resultados de estimación.

Por último, los resultados del modelo con efectos aleatorios¹⁷, se reportan en la tabla número 5. Al respecto, se puede notar una mejoría en comparación con las dos regresiones anterior debido a que el coeficiente de cada una de las variables muestra el signo adecuado conforme lo postula la teoría y/o evidencia empírica. Sin embargo las variables CLU e IP son estadísticamente no significativas.

En este sentido, si sólo se consideraran los resultados de las estimaciones en cuanto al signo de los coeficientes y su significancia estadística; de las tres se elegiría la que corresponde al modelo de efectos aleatorios, pero es necesario utilizar algunas pruebas econométricas formales que permitan orientar la selección del modelo más apropiado. En los siguientes párrafos se presentan tales pruebas.

¹⁷ Nótese que en el modelo de efectos aleatorios, el estadístico t cambia por el estadístico z. Sin embargo la interpretación es la misma. Esta diferencia se debe a que el método para estimar por efectos aleatorios es el de Mínimos Cuadrados Generalizados (MCG).

Tabla 5: Resultados de la Regresión con Efectos Aleatorios

Dependiente: TIL			
Variable	Coefficiente	Error Estándar	Estadístico z
Intercepto	154.6195*	17.2812	8.95
PIB	-3.2359*	0.8722	-3.71
ESCO	-41.7879*	6.6037	-6.33
NDES	3.8467*	0.6477	5.94
SMODERNO	-0.3836*	0.1035	-3.70
CLU	0.3896	0.2884	1.35
ER	-0.0116*	0.0043	-2.66
IP	-0.0182	0.1652	-0.11
COBERTURA	0.0607*	0.0149	4.08

R² within = 0.4347 R² between = 0.7445 R² overall = 0.7317
 * Coeficiente significativo al 5%.
 **Significativo al 10% .

Fuente: Resultados de estimación.

En primer lugar, para elegir entre un modelo de regresión agrupada y el modelo con efectos fijos, se instrumenta una prueba como la reportada en la tabla número 6. Se observa que el valor de este estadístico es relativamente elevado y su p-value, prácticamente de cero. La hipótesis nula ($H_0: \alpha_i = 0$) de la prueba F establece que cada intercepto α_i es igual a cero. La aceptación de H_0 implica que el mejor modelo es el de regresión agrupada debido a que este supone un intercepto común para todas las entidades. El rechazo de H_0 sugiere que deberá estimarse el panel con efectos fijos, ya que cada intercepto es distinto para cada entidad federativa. Entonces, si el p-value asociado a la prueba F es menor que 0.05 se rechaza H_0 y se elige efectos fijos en lugar del modelo de regresión agrupada. En este caso los resultados indican que se debe elegir el modelo de efectos fijos.

Tabla 6: Prueba F. Regresión Agrupada vs Efectos Fijos

Prueba F
F (8, 23) = 25.34
Prob > F = 0.0000

Fuente: Resultados de estimación.

También es posible probar formalmente si debe escogerse un modelo de regresión agrupada o uno de efectos aleatorios. Para ello se recurre a la prueba del Multiplicador de Lagrange (LM) en la que $H_0: \sigma_u^2 = 0$. Esta hipótesis plantea que la varianza de las perturbaciones es igual a cero, o bien que no existe el efecto no observado en el término de error tal como se asume en efectos aleatorios. Su aceptación indica que se utilice la regresión agrupada, en tanto que su rechazo sugiere utilizar efectos aleatorios. Si el p-value asociado a este estadístico es menor a 0.05 se rechaza H_0 y si su valor es mayor a 0.05, se elige el modelo de regresión agrupada estimado con MCO. En la tabla número 7 se presentan los resultados de la prueba LM que indican que se debe elegir el modelo de efectos aleatorios, debido a que el p-value < 0.05 permite rechazar H_0 .

Tabla 7: Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios

Resultados Estimados:		
	Var	S.D.
TIL	115.5133	10.7477
e	1.2876	1.1347
u	24.3223	4.9317
Prueba: Var (u) =0	Chibar2(01) =	799.05
	Prob > Chibar2 =	0.0000

Fuente: Resultados de estimación.

Una vez efectuadas las pruebas correspondientes y determinado que la regresión con intercepto común no es la más apropiada, la siguiente cuestión a resolver es *¿cuál modelo es el mejor, el de efectos fijos o el de efectos aleatorios?*. Para responder esta interrogante se debe aplicar la prueba de Hausman, pero el inconveniente con ella es que puede reportar valores negativos para el estadístico y no se puede determinar qué modelo elegir. Es común que esta situación se presente cuando las estimaciones son robustas a autocorrelación y heterocedasticidad¹⁸. Sin embargo, es de crucial importancia que las regresiones sean robustas ante estos problemas. Por lo cual, una alternativa a la prueba tradicional de Hausman consiste en utilizar la prueba de sobreidentificación (ya explicada) que se basa en la prueba de Hausman pero es robusta y asegura la obtención de un valor positivo para el estadístico y por lo tanto, permite elegir entre un modelo con efectos fijos y uno de efectos aleatorios.

¹⁸ Estas afirmaciones se realizan con base a la experiencia obtenida al realizar las estimaciones en STATA, pues el programa impide aplicar la prueba tradicional de Hausman cuando se utiliza el comando cluster para estimar el modelo robusto a autocorrelación y heterocedasticidad. Asimismo, STATA señala y emite un aviso cuando el resultado del estadístico es negativo.

La H_0 de la prueba de sobreidentificación es que los estimadores de efectos fijos y de efectos aleatorios son iguales o bien $H_0 : \text{plim}(\hat{\theta} - \tilde{\theta}) = 0$ donde $\hat{\theta}$ es el estimador de efectos fijos y $\tilde{\theta}$ el de aleatorios (de acuerdo con lo presentado en la sección de pruebas formales). Si se acepta H_0 entonces se deberá utilizar el modelo de efectos aleatorios, en caso contrario el de efectos fijos. Los resultados de la prueba se exhiben en la tabla número 8 de la que se concluye que es mejor utilizar efectos fijos al obtener un p-value < 0.05 , que indica que se debe rechazar la H_0 .

Tabla 8: Prueba de Sobreidentificación. Efectos Fijos vs Efectos Aleatorios.

Estadístico de Sargan-Hansen	P-value
137.213	0.0000

Fuente: Resultados de estimación.

Una vez efectuadas las pruebas de rigor, se determina que el mejor modelo es el de efectos fijos cuando se consideran las 8 variables explicativas. Pero es posible que el modelo se pueda mejorar mediante dos vías: la primera es excluir de las especificaciones 1, 2 y 3 a la variable IP, la razón para hacerlo, es que en todas las estimaciones resulta no significativa y en algunos casos con el signo opuesto a lo esperado. La segunda vía es excluir al PIB y/o al NDES. La razón que justifica este proceder es que teóricamente el desempleo y el PIB tienden a tener una correlación elevada (en el presente caso se calcula de 0.7398) lo que en el análisis empírico puede generar problemas de correlación serial y con ello que los resultados no sean confiables debido a que existiría colinealidad entre las variables. En este sentido, se propone la estimación de tres especificaciones alternativas (incisos a, b y c), a los modelos 1, 2 y 3, aplicando también las pruebas formales. Cabe señalar que en todas las variaciones las pruebas indicaron que el modelo que debe elegirse es el de efectos fijos o la especificación 2. Por lo que sólo considerando esta metodología se presentan los resultados en las tablas número 9, 10 y 11. En la sección de anexos se encuentran los resultados para los métodos de regresión agrupada y efectos aleatorios, así como sus respectivas pruebas.

a) Especificación 2 excluyendo la variable IP.

$$\begin{aligned}
 TIL_{it} = & \alpha_i + \beta_1 PIB_{it} + \beta_2 ESCO_{it} + \beta_3 NDES_{it} + \beta_4 SMODERNO_{it} + \beta_5 CLU_{it} + \beta_6 ER_{it} + \\
 & \beta_7 COBERTURA_{it} + u_{it}
 \end{aligned}
 \tag{4}$$

Los resultados que se reportan en la tabla número 9 arrojan resultados razonablemente buenos, todos los coeficientes estimados de las variables poseen el signo correcto. Además, para algunos coeficientes el valor de los errores estándar disminuyó, incrementándose el valor del estadístico t y disminuyendo el p-value. Sin embargo, en esta estimación la variable COBERTURA es positiva (posee el signo adecuado a lo predicho por el enfoque institucional) pero no es estadísticamente significativa. Por lo cual es probable que este resultado sea distinto en los modelos b y c de las tablas número 10 y 11.

Tabla 9: Resultados de la Regresión con Efectos Fijos. Sin IP

Dependiente: TIL			
Variable	Coefficiente	Error Estándar	Estadístico t
Intercepto	125.9798*	23.3821	5.39
PIB	-3.1039**	1.5528	-2.00
ESCO	-29.9946*	7.9004	-3.80
NDES	3.9244*	0.6633	5.92
SMODERNO	-0.3262*	0.1063	-3.07
CLU	0.4206**	0.2310	1.82
ER	-0.0100*	0.0043	-2.32
COBERTURA	0.0272	0.0175	1.55

R² within = 0.4476 R² between = 0.6647 R² overall = 0.6486
 * Coeficiente significativo al 5% .
 **Significativo al 10%.

Fuente: Resultados de estimación.

b) Especificación 2 excluyendo las variables IP y PIB.

$$TIL_{it} = \alpha_i + \beta_1 ESCO_{it} + \beta_2 NDES_{it} + \beta_3 SMODERNO_{it} + \beta_4 CLU_{it} + \beta_5 ER_{it} + \beta_6 COBERTURA_{it} + u_{it} \quad [5]$$

Los resultados de la estimación de la ecuación 5 se encuentran en la tabla número 10 considerando la metodología de efectos fijos, debido a que las pruebas formales señalan que este es el mejor modelo cuando no se consideran las variables IP y PIB. En lo que concierne a las relaciones esperadas todas son correctas y en cuanto a la significancia, nuevamente el coeficiente de COBERTURA no es estadísticamente significativo.

La conclusión de estos resultados en comparación con los anteriores (tabla 9), es que dejar fuera al PIB provoca que la significancia estadística se vea modificada, pues para la variable ER la significancia es ahora al 10%, en comparación con el 5% reportado en la tabla número 9. Además, en el caso de la variable COBERTURA el valor del estadístico t se ve disminuido. Basándose en estos resultados se puede afirmar que el PIB es de gran importancia por lo que debería ser considerado.

Tabla 10: Resultados de la Regresión con Efectos Fijos. Sin IP y PIB

Dependiente: TIL			
Variable	Coefficiente	Error Estándar	Estadístico t
Intercepto	94.7848*	15.4887	6.12
ESCO	-33.1115*	8.2733	-4.00
NDES	3.8969*	0.6738	5.78
SMODERNO	-0.3585*	0.1018	-3.52
CLU	0.4540**	0.2319	1.96
ER	-0.0084**	0.0043	-1.92
COBERTURA	0.0209	0.0164	1.28

R² within = 0.4320 R² between = 0.4896 R² overall = 0.4819
 * Coeficiente significativo al 5% .
 **Significativo al 10%.

Fuente: Resultados de estimación.

c) Especificación 2 excluyendo IP y NDES.

$$TIL_{it} = \alpha_i + \beta_1 PIB_{it} + \beta_2 ESCO_{it} + \beta_3 SMODERNO_{it} + \beta_4 CLU_{it} + \beta_5 ER_{it} + \beta_6 COBERTURA_{it} + u_{it} \quad [6]$$

Los resultados de la tabla número 11 corresponden a la estimación de la especificación 6 por el método de efectos fijos. El comportamiento de todos los coeficientes estimados en cuanto al tipo de relación esperada y en cuanto a la significancia estadística, son hasta el momento los mejores en comparación con las especificaciones 2, 4 y 5. Adicionalmente, la significancia mejoró para las variables SMODERNO, CLU, ER y COBERTURA al disminuir los valores del p-value y aumentar los valores del estadístico t. Entonces las pruebas señalan que el mejor modelo es el de efectos fijos en cualquier especificación y basándose en el comportamiento de los coeficientes, el mejor modelo es el que no incluye a las variables IP y NDES.

Tabla 11: Resultados de la Regresión con Efectos Fijos. Sin IP y NDES

Dependiente: TIL			
Variable	Coefficiente	Error Estándar	Estadístico t
Intercepto	138.2694*	24.9998	5.53
PIB	-2.8446*	1.1668	-2.44
ESCO	-16.2709*	7.4487	-2.18
SMODERNO	-0.5110*	0.1327	-3.85
CLU	0.8257*	0.2842	2.90
ER	-0.0095**	0.0048	-1.99
COBERTURA	0.0403**	0.0212	1.91

R² within = 0.2614 R² between = 0.3998 R² overall = 0.3968
 * Coeficiente significativo al 5% .
 **Significativo al 10%.

Fuente: Resultados de estimación.

Considerando que la especificación 6 es el mejor modelo, se procede también a estimar los coeficientes de correlación simple entre las variables explicativas. El criterio definido para indicar ausencia de correlación es que aquellos valores que resulten ser menores o iguales que 0.5 en términos absolutos no revelan un problema. La tabla número 12 muestra la matriz de correlaciones para las variables PIB, ESCO, SMODERNO, CLU, ER y COBERTURA. Se concluye que todas las variables explicativas del modelo son linealmente independientes.

Tabla 12: Matriz de Correlaciones

	PIB	ESCO	SMODERNO	CLU	ER	COBERTURA
PIB	1					
ESCO	0.3498	1				
SMODERNO	0.3494	0.1480	1			
CLU	0.2482	0.3077	0.1537	1		
ER	-0.2427	-0.3402	-0.2920	-0.3782	1	
COBERTURA	-0.2262	-0.0780	-0.3802	-0.1536	0.0410	1

Fuente: Resultados de estimación.

V.2. Variables Instrumentales

Una vez seleccionado el mejor modelo, a continuación se procede a verificar la posible existencia de endogeneidad, la cual se presenta cuando alguna o algunas de las variables explicativas se correlacionan con el término de error. La metodología que permite lidiar con este problema es el

uso de variables instrumentales¹⁹ que deben cumplir dos condiciones; correlacionarse con la variable endógena pero no con el término de error. En el presente caso, se plantea que la posible fuente de endogeneidad reside en la variable ESCO y se proponen a las variables NESCUÉLAS y TDS como instrumentos.

Un primer paso es realizar una prueba de endogeneidad tal como la que se señala en la sección III.5 basándose en Wooldridge (2015). Se obtienen los residuos de una estimación que posee como dependiente a la variable ESCO y como explicativas a NESCUÉLAS y TDS, tales residuos se incluyen como una variable explicativa adicional al modelo de efectos fijos seleccionado como el mejor y se estima poniendo atención a su coeficiente. La tabla número 13 contiene los resultados de la estimación para probar endogeneidad y es posible afirmar que no hay evidencia de endogeneidad debido a que el coeficiente de la variable RESIDUOS no es estadísticamente significativo, por lo cual no es necesario el uso de la metodología de variables instrumentales.

Tabla 13: Resultados Prueba de Endogeneidad No. 1

Dependiente: TIL			
Variable	Coefficiente	Error Estándar	Estadístico t
Intercepto	134.8196*	24.8971	5.42
PIB	-3.5481*	1.2734	-2.79
ESCO	-10.7219	8.3512	-1.28
SMODERNO	-0.4931*	0.1393	-3.85
CLU	0.8088*	0.2696	2.90
ER	-0.0105*	0.0048	-1.99
COBERTURA	0.0332	0.0209	1.91
RESIDUOS	-8.0757	8.0318	-1.01

R² within = 0.2690 R² between = 0.3092 R² overall = 0.3084
 * Coeficiente significativo al 5% .
 **Significativo al 10%.

Fuente: Resultados de estimación.

Una manera adicional de probar endogeneidad es hacer uso del comando endog del programa STATA o bien la prueba de endogeneidad de regresores endógenos, o como Baum (2007) le denomina probar la exogeneidad, que es exactamente lo mismo. En la tabla número 14 aparece el

¹⁹ La explicación a detalle de la metodología de variables instrumentales se encuentra en el capítulo tres y la justificación del uso de los instrumentos NESCUÉLAS y TDS se encuentra en el capítulo cuatro.

resultado de esta prueba, misma que permite corroborar que la variable escolaridad no es endógena, puesto que el p-value es igual a 0.2724 aceptándose la hipótesis nula que dice que el regresor (variable explicativa) especificado como endógeno puede ser tratado como exógeno.

Tabla 14: Resultados Prueba de Endogeneidad No. 2

Regresor: ESCO	
Prueba de endogeneidad de los regresores endógenos	Chi-sq (1) p-value
1.2040	0.2724

Fuente: Resultados de estimación.

Hasta este punto, no existen razones para seleccionar un modelo de variables instrumentales. Sin embargo debe probarse si los instrumentos definidos para ESCO son los apropiados, es decir, qué tan válidos son estos. Para ello es necesario probar la exogeneidad estricta de los instrumentos, o en otras palabras, probar que los instrumentos no se encuentran correlacionados con el término de error.

La prueba de exogeneidad por excelencia es la de Sargan (descrita en el capítulo tres). En la que aceptar la hipótesis nula indica que los instrumentos son exógenos y válidos. Los resultados de la tabla número 15 implican que el uso de NESCUELAS y TDS como instrumentos de ESCO es correcto, pues son exógenos y válidos. El estadístico que se reporta en el programa de STATA es el de Hansen J pero hace referencia al estadístico de Sargan-Hansen, su p-value es igual a 0.2106 y por lo tanto mayor que 0.05 lo que permite aceptar la hipótesis nula de exogeneidad.

Tabla 15: Resultados Prueba de Exogeneidad

Estadístico Hansen J	
Valor del Estadístico	Chi-sq (1) p-value
1.5670	0.2106

Fuente: Resultados de estimación.

En conclusión, se descarta la posible presencia de endogeneidad en el modelo con datos en panel seleccionado como el mejor. Por lo que puede afirmarse que no es necesario estimar la especificación 6 utilizando la metodología de variables instrumentales. Asimismo se comprueba

que las variables NESCUELAS y TDS son instrumentos válidos y que fueron correctamente seleccionados ya que cumplen con el requisito de ser estrictamente exógenos.

V.3. El “Mejor” Modelo

Llegado a este punto, fueron diversas las estimaciones que se realizaron para abordar el fenómeno de interés. Los métodos empleados van desde el más simple como es la regresión agrupada, después efectos fijos, efectos aleatorios, hasta métodos más complejos como el de variables instrumentales. Asimismo, se aplicaron diversas pruebas a los resultados de los modelos lo que permitió concluir que el mejor modelo es el de efectos fijos que considera seis variables explicativas que son el PIB, ESCO, SMODERNO, CLU, ER y COBERTURA.

Algunas características interesantes sobre el mejor modelo es que se compone por tres variables que corresponden al enfoque teórico dual (PIB, ESCO y SMODERNO) y por tres del enfoque institucional (CLU, ER y COBERTURA), también resalta que cuatro de los seis coeficientes estimados son estadísticamente significativos al 5% y solamente dos al 10%. Asimismo, las relaciones esperadas entre cada una de las variables y la informalidad son exactamente iguales a lo esperado de acuerdo con la teoría descrita en la sección correspondiente a ello.

Adicionalmente pueden realizarse los siguientes dos procedimientos que corroboran que la estimación mencionada es la mejor, de acuerdo con lo propuesto en Wooldridge (2015) y descrito en la sección de pruebas formales. Estos procedimientos son los efectos marginales y las primeras diferencias.

1. Efectos Marginales: una vez que se ha seleccionado el mejor modelo, se busca obtener los efectos marginales de la regresión y realizar la comparación entre estos valores y los coeficientes estimados. Lo deseable es que sean iguales para poder decir que no existe problema de especificación.

Al compararse los valores de los efectos marginales que se encuentran en la tabla número 16 en la columna dy/dx con los valores de los coeficientes estimados de la primera columna de la tabla número 11 resulta que estos son exactamente iguales, lo que implica que la especificación 6 estimada a través de efectos fijos no posee error de especificación, corroborándose que es el mejor modelo.

Tabla 16: Efectos Marginales

Variable	Efecto marginal dy/dx
PIB	-2.8446
ESCO	-16.2709
SMODERNO	-0.5110
CLU	0.8257
ER	-0.0095
COBERTURA	0.0403

Fuente: Resultados de estimación.

2. Primeras Diferencias: al igual que el procedimiento anterior, se debe tener seleccionado el mejor modelo y con base a ello realizar la estimación pero considerando una regresión lineal con las variables en primeras diferencias. Lo que se espera es que las estimaciones de primeras diferencias y de efectos fijos (mejor modelo) sean iguales porque ambas metodologías eliminan el efecto no observado c_i (revise capítulo tres), en tal caso no habrá diferencia entre utilizar un modelo de regresión con las variables en primeras diferencias o uno con efectos fijos.

Ahora bien, puede suceder que los coeficientes de ambas estimaciones no resulten ser iguales por lo que la decisión para elegir entre uno u otro modelo, se deberá a la naturaleza de la correlación serial en el término de error u_{it} . En este sentido, con poca correlación serial se prefiere un modelo con efectos fijos, y cuando la correlación serial sea elevada se prefiere un modelo en primeras diferencias. En el plano empírico, el procedimiento es estimar el modelo en primeras diferencias y realizar la prueba de raíz unitaria a los residuos del modelo con efectos fijos. Si la regresión de efectos fijos es espuria en niveles, es decir, no es $I(0)$ entonces se debe elegir la estimación de primeras diferencias y por el contrario si los residuos u_{it} son $I(0)$ entonces se elige el modelo con efectos fijos.

La tabla número 17 contiene los resultados de la estimación de la especificación 6 considerando todas las variables en primeras diferencias. De esta tabla pueden notarse que los coeficientes asociados a las variables explicativas son distintos en comparación con los resultados de la tabla

número 11 por lo que sí existe diferencia entre utilizar efectos fijos de esta forma y utilizar las primeras diferencias, además en primeras diferencias el coeficiente de CLU posee el signo incorrecto y es estadísticamente significativo lo que hace dudar de la pertinencia de esta estimación. Formalmente, la elección se basa en lo que dicten las pruebas de raíz unitaria sobre los residuales del modelo de efectos fijos.

Tabla 17: Resultados de la Regresión en Primeras Diferencias

Dependiente: D.TIL			
Variable	Coefficiente	Error Estándar	Estadístico t
D.PIB	-6.5813	2.1558	-3.05
D.ESCO	-25.8893	5.1400	-5.04
D.SMODERNO	-0.4426	0.0990	-4.47
D.CLU	-0.4863	0.2251	-2.16
D.ER	-0.0055	0.0023	-2.39
D.COBERTURA	0.0702	0.0168	4.17

$R^2 = 0.2540$
 * Coeficiente significativo al 5% .
 **Significativo al 10%.

Fuente: Resultados de estimación.

En la tabla número 18 se muestran los resultados de las pruebas de raíz unitaria aplicadas a los residuos u_{it} del modelo de efectos fijos seleccionado como el mejor. Las pruebas son las de Levin, Lin & Chu, Dickey-Fuller Aumentada (ADF) y Phillips-Perron (PP), en estas la hipótesis nula es que la variable en cuestión posee raíz unitaria, para rechazarla es necesario obtener que el $p\text{-value} < 0.05$. Si la hipótesis nula se rechaza, entonces la variable no posee raíz unitaria y por lo tanto es estacionaria ya sea en niveles $I(0)$ o en primeras diferencias $I(1)$. Sin embargo en el caso de los residuos u_{it} se busca que sean $I(0)$.

Las pruebas de raíz unitaria sobre u_{it} indican que son $I(0)$ debido a que en los tres casos (Levin, Lin & Chu, ADF y PP) el $p\text{-value}$ es menor que 0.05. Esto permite concluir que es preferible la estimación de efectos fijos. Una vez más se corrobora que la especificación 6 es el mejor modelo al estimarlo por efectos fijos. Se puede argumentar también que los mejores resultados son los de la tabla número 11 y que estos son los que deben utilizarse para interpretar y realizar las conclusiones del próximo capítulo.

Tabla 18: Resultados de las Pruebas de Raíz Unitaria para los Residuos del “Mejor” Modelo

Prueba	Estadístico	Prob.
Levin, Lin & Chu	-10.1559	0.0000
ADF	165.1800	0.0000
PP	134.9560	0.0000

Fuente: Resultados de estimación.

V.4. Análisis e Interpretación de Resultados

Los resultados que se analizan en esta sección del documento corresponden a la estimación de la especificación 6 que es un modelo de efectos fijos robusto ante el problema de autocorrelación y heterocedasticidad. Las variables que la integran son PIB, ESCO, SMODERNO, CLU, ER y COBERTURA. Las primeras tres variables se vinculan al enfoque teórico dual y el resto a una perspectiva institucional. Los resultados para TIL se encuentran en la tabla número 11 de la sección anterior y las distintas pruebas estadísticas señalan que es la mejor estimación considerando el total de la informalidad, lo mismo sucede en el caso de la informalidad independiente y dependiente²⁰.

Sin embargo, para poder verificar el cumplimiento de la hipótesis planteada al principio del documento es necesario que los coeficientes (o pendientes) a interpretar sean comparables, para ello existen dos opciones: la primera es que todas las variables explicativas se encuentren en la misma unidad de medida y la segunda es que aun cuando las variables se encuentren en diferentes unidades, los coeficientes estimados sean transformados de tal manera que al final sean comparables.

En el presente caso se aplicó la segunda opción debido a que algunas variables se encuentran en logaritmo natural y otras en tasas, en tanto que la dependiente es una tasa. Esto ocasiona que algunos de los coeficientes reportados por el modelo sean la razón de un cambio absoluto dado un cambio relativo y que puedan ser interpretados como semielasticidades debido a que la variable explicativa se encuentra expresada en términos de logaritmo natural y la dependiente en términos de tasa porcentual. Otros coeficientes son una razón de cambios absolutos cuando la

²⁰ Las pruebas aplicadas a los modelos para la informalidad independiente y dependiente, señalan que la mejor estimación es aquella que se basa en la metodología de efectos fijos, esto al compararse con las estimaciones de regresión agrupada y efectos aleatorios. Las pruebas son la F, la del Multiplicador de Lagrange (LM) y la de Sobreidentificación, mismas que se encuentran en la sección B.2. y B.3. de los anexos.

variable explicativa así como la dependiente se encuentran expresadas en términos de tasa porcentual, o bien son elasticidades. Por lo tanto todos los coeficientes que son semielasticidades se transformaron a elasticidades con el objetivo de unificar el análisis.

En la tabla número 19 se muestran las elasticidades calculadas para cada una de las variables incluidas en el modelo que considera al total de la informalidad laboral, es decir, aquel que posee como variable dependiente a TIL y que fue seleccionado como el mejor entre las distintas alternativas (especificación 6 por la metodología de efectos fijos). Comenzando por las variables representativas del enfoque dual de la informalidad laboral, la primera elasticidad a interpretar es la del PIB, cuyo valor es de -0.04. El signo corrobora lo esperado a nivel teórico, es decir, la relación entre el PIB y la informalidad es inversa. En este sentido, si el PIB se incrementa en 1% en promedio, la TIL disminuirá en 0.04% y de manera contraria.

Este resultado da elementos suficientes para decir que las condiciones económicas del país sí ejercen influencia sobre las variaciones de los niveles de informalidad observados en las entidades federativas. Sin embargo, cabe precisar que el valor de la elasticidad es muy pequeño y por lo tanto la TIL no es muy sensible a variaciones en el PIB (al compararlo con el resultado para otras variables).

Tabla 19: Elasticidades. No.1

Modelo con TIL		
Variable	Coefficiente Estimado	Elasticidad
Dual		
PIB	-2.8447	-0.0493
ESCO	-16.2709	-0.2818
SMODERNO	-0.5111	-0.1324
Institucional		
CLU	0.8257	0.0143
ER	-0.0096	-0.0034
COBERTURA	0.0404	0.0238

Fuente: Elaboración propia con base a resultado de estimación.

La siguiente variable del enfoque teórico dual es la escolaridad (ESCO) cuyo valor calculado de la elasticidad es -0.28 e indica una relación negativa entre ESCO y la informalidad. Es decir, un

cambio negativo de 0.28% en la TIL es ocasionado por un aumento de 1% en promedio en la escolaridad de la Población Económicamente Activa (PEA). En términos generales el resultado de la elasticidad señala que a mayores niveles de educación, la informalidad tenderá a disminuir y por el contrario a menores niveles de educación se observarán mayores tasas de informalidad en el país. Nótese que la sensibilidad de la TIL es mayor ante la escolaridad que ante cualquier otra variable del enfoque dual o institucional. De aquí la relevancia de la educación de la población como un medio para mejorar las condiciones laborales.

Por su parte, la variable SMODERNO reporta una elasticidad con valor de -0.13. Así pues, un incremento de 1% en promedio del SMODERNO genera un impacto negativo de 0.13% en la TIL. Por el contrario, si la capacidad del SMODERNO disminuye en 1% la TIL aumentará en 0.13%. Este resultado constituye evidencia para resaltar la importancia que tiene el hecho de que el sector formal (medido por la variable SMODERNO) posea la capacidad de absorber la suficiente mano de obra disponible para trabajar, con ello se lograría que disminuya el número de individuos que son empujados hacia la informalidad debido a que no encuentran una vacante en el lado formal de la economía. Si bien, el valor de la elasticidad no es elevado, si es mayor que el del PIB.

De la interpretación anterior referida sólo a la perspectiva dual del fenómeno de informalidad laboral, se puede resaltar que la variable que mayor efecto tiene sobre la TIL es la ESCO pues el incremento de 1% en promedio en esta, ocasiona una disminución de 0.28% en la TIL. En segundo lugar se encuentra el SMODERNO que influye negativamente en 0.13% sobre la TIL. Y finalmente el PIB es el que menor influencia posee sobre la informalidad al ser de 0.04%. Por lo tanto debe ponerse especial atención en la educación si se desea influir de forma más significativa en los niveles de la TIL a nivel nacional.

Continuando con la perspectiva institucional del fenómeno de investigación, la primera elasticidad a interpretar es la que corresponde a la variable de costo laboral unitario (CLU) con un valor de 0.01. En primera instancia su signo indica que la relación entre CLU y la TIL es de acuerdo con lo esperado, es decir positiva, por lo que un incremento de 1% en promedio en el CLU propicia un incremento cercano a 0.01% en la TIL. En contraste, si el CLU disminuye en 1% la TIL lo hace en 0.01%. Como es posible observar, el efecto del costo laboral sobre la informalidad observada es pequeño, pero estadísticamente significativo.

Es bien sabido que el costo que representan los trabajadores para los empleadores influye en la cantidad que deciden emplear, por esta razón un incremento del costo laboral en el sector moderno o formal traerá como consecuencia que disminuya el empleo en dicho sector generando así que el sector informal se incremente como resultado de tres efectos; el primero es que un mayor número de individuos no encuentran empleo en el sector formal debido a que la capacidad de absorción de mano de obra formal disminuye, esto si se considera que el incremento en el costo laboral ocasiona que cada trabajador sea más caro y eso afecta las decisiones de contratación de personal para una empresa, por lo que aquellos que no encuentren empleo en el lado formal si lo hagan en el lado informal. El segundo efecto es que los empleadores optarán por contratar individuos bajo condiciones de informalidad debido a que representan menor costo en comparación con el empleo bajo contrato formal. El tercer efecto es que algunos de los que tenían empleo formal pueden ser despedidos y terminar en el sector informal de la economía. Considerando la perspectiva institucional se puede decir entonces que las acciones gubernamentales, por ejemplo a través de políticas salariales o en general políticas del mercado de trabajo que tengan impacto sobre el costo laboral, son de interés por sus efectos sobre la informalidad.

En lo que respecta al valor de la elasticidad asociada a la eficiencia del gobierno en términos de recaudación de impuestos ER es de -0.003 y por tanto bastante cercano a cero. La característica distintiva de ER es que puede reportar un signo positivo o negativo. En el presente caso el coeficiente resultó negativo e indica que un incremento de 1% en promedio en la ER (menor eficiencia) ocasiona una disminución de 0.003% en la TIL y de modo inverso. Este es sin duda el coeficiente con el valor más bajo que se reporta en el marco de las seis variables explicativas que representan a ambas posturas teóricas. Asimismo, es posible argumentar que la TIL no es sensible a variaciones en la ER. Por otra parte, el hecho de que resulte significativa es de utilidad para poder decir que al incrementarse la eficiencia recaudatoria (disminución del valor de la variable) que es medida como la proporción de ingresos por impuestos que se gastan en recaudar, la informalidad se incrementa como resultado de la búsqueda de los individuos por pagar menores impuestos o bien en su objetivo de evadir los costos de la formalidad.

Por último se encuentra la elasticidad calculada para la cobertura del seguro popular (COBERTURA) cuyo valor es de 0.02, comprobándose la relación positiva esperada con la TIL.

Un incremento de 1% en la COBERTURA en promedio genera un incremento de 0.02% en la TIL. En cambio, una disminución de 1% en la COBERTURA genera una disminución de 0.02% en la TIL. La mayor cobertura del seguro popular es contraproducente para los niveles de la TIL en el país pues la incrementan, y esto es debido a que el seguro popular es un sustituto de la atención médica que otorga el sector formal de la economía y a que los individuos en su esfuerzo por disminuir los costos de permanecer formal, optan por este tipo de cobertura médica que implica un menor costo. Recuérdese que la variable COBERTURA fue incluida para capturar la explicación netamente voluntaria del individuo por pertenecer a la informalidad y en este sentido al resultar significativa y además con el coeficiente más elevado dentro de los factores institucionales, se corrobora la explicación voluntaria del fenómeno.

Si se consideran solamente las variables institucionales, la que mayor impacto ocasiona sobre la TIL es la COBERTURA cuyo efecto es de 0.02%, seguida por el CLU que afecta a las variaciones de la TIL en 0.01% y por último la ER con un 0.003%. Debe destacarse el hecho que todos los valores son muy cercanos a cero, por lo que la sensibilidad de la TIL a los factores institucionales es muy baja en comparación con los factores de orden dual. Este hecho constituye un argumento para poder realizar la afirmación que se cumple la hipótesis planteada que menciona que el mayor efecto sobre la informalidad corresponde a los factores determinantes de orden dual.

Al principio de esta investigación se señala que son tres las versiones del modelo de informalidad que se estiman. Todo el análisis e interpretación anterior corresponde a la versión que considera el total de la informalidad laboral y que es la estimación principal. Por tanto, se procede a interpretar y analizar los resultados para las dos versiones restantes, una es aquella que considera a la tasa de informalidad laboral independiente (TILI) como variable dependiente y otra la que tiene a la tasa de informalidad laboral dependiente (TILD) como variable a explicar. Estos modelos son el 7 y el 8, respectivamente.

Al igual que en el caso general, los modelos fueron estimados utilizando la metodología de efectos fijos, robustos a problemas de autocorrelación y heterocedasticidad, también se realizó la transformación de los coeficientes a elasticidades para hacerlos comparables entre sí, y se especifican como a continuación:

$$TILI_{it} = \alpha_i + \beta_1 PIB_{it} + \beta_2 ESCO_{it} + \beta_3 SMODERNO_{it} + \beta_4 CLU_{it} + \beta_5 ER_{it} + \beta_6 COBERTURA_{it} + u_{it} \quad [7]$$

$$TILD_{it} = \alpha_i + \beta_1 PIB_{it} + \beta_2 ESCO_{it} + \beta_3 SMODERNO_{it} + \beta_4 CLU_{it} + \beta_5 ER_{it} + \beta_6 COBERTURA_{it} + u_{it} \quad [8]$$

Los resultados de las estimaciones se muestran en las tablas 20 y 21 para el modelo 7 y 8 respectivamente. Cabe aclarar que solamente se consideran las transformaciones para los coeficientes que resultaron estadísticamente significativos. La tabla número 20 contiene las elasticidades para el modelo con TILI como variable dependiente. Cuando se considera la informalidad independiente, los resultados demuestran que solamente dos variables son significativas en su explicación, una de ellas corresponde a la visión dual y la otra a la institucional. Esto permite confirmar la hipótesis planteada al inicio, es decir, lo que ocurre con el total de la informalidad también ocurre con la informalidad de manera desglosada, o bien los factores de ambos enfoques teóricos influyen en la explicación de los distintos tipos de informalidad al igual que del total de la informalidad.

Tabla 20: Elasticidades. No.2

Modelo con TILI		
Variable	Coefficiente Estimado	Elasticidad
Dual		
SMODERNO	-0.1872	-0.1345
Institucional		
ER	-0.0057	-0.0057

Fuente: Elaboración propia con base a resultado de estimación.

En cuanto a la interpretación, se encuentra evidencia que permite decir que una disminución de la TILI de 0.13% se presenta cuando el SMODERNO se incrementa en 1% en promedio y por el contrario. Además, en las tres versiones del modelo, la única variable que puede compararse es precisamente esta puesto que resulta significativa en todas y es interesante como el valor de la elasticidad es aproximadamente el mismo, con lo que se ratifica nuevamente la hipótesis mencionada en el párrafo anterior. Por otro lado, un incremento de 1% en promedio de la ER genera un decremento de 0.005% en la TILI, y por el contrario. El valor para esta variable es

mayor cuando se considera la informalidad independiente que cuando se considera la informalidad total (valor de -0.0034).

En la tabla número 21 se encuentran las elasticidades para la informalidad laboral dependiente, TILD. En general, los signos de todos los coeficientes y por lo tanto de las elasticidades calculadas, son los adecuados. Es posible notar un cambio respecto a las variables de la tabla número 20 y es que la ER que forma parte de las variables institucionales no es estadísticamente significativa y por lo tanto no aparece en la tabla, ello sugiere que al menos una de las variables institucionales no explica a la informalidad dependiente.

Tabla 21: Elasticidades. No.3

Modelo con TILD		
Variable	Coefficiente Estimado	Elasticidad
Dual		
PIB	-1.6560	-0.0448
ESCO	-15.0128	-0.4064
SMODERNO	-0.3236	-0.1310
Institucional		
CLU	0.6855	0.0186
COBERTURA	0.0518	0.0476

Fuente: Elaboración propia con base a resultado de estimación.

La interpretación y análisis de las elasticidades van en el mismo sentido que las anteriores pero ahora aplicado a la informalidad del tipo dependiente. Por el lado dual, si el PIB se incrementa en 1% en promedio entonces la TILD disminuye en 0.04% y viceversa, por lo que la situación de la economía del país explica las variaciones de la informalidad dependiente. También destaca que el valor de la elasticidad calculada es menor pero muy cercano que en el caso del modelo que considera a TIL, esto puede deberse a que la TIL se compone en mayor medida por la TILD.

Continuando con la explicación dual, un aumento de ESCO en 1% en promedio genera una disminución de la TILD en 0.40% y de forma contraria. Este resultado es muy interesante debido a que el valor calculado de la elasticidad es superior al presentado anteriormente para el modelo con TIL que es de 0.28% y lo que indica son dos cosas: en primer lugar que el mayor nivel de educación de los individuos contribuye en gran medida a la disminución de la informalidad del

tipo dependiente y en segundo, que los individuos con menores niveles de educación además de encontrarse en el sector informal lo hacen en un tipo de empleo en el que son subordinados a un patrón.

Por último, el incremento en 1% en promedio del SMODERNO ocasiona una disminución de 0.13% de la TILD y viceversa. En este caso llama la atención que el valor de la elasticidad sea casi igual, -0.1310 y -0.1324, al compararse con el valor del modelo general con TIL y lo que puede interpretarse a partir de este hecho es que el SMODERNO influye en la misma magnitud si se considera el total de la informalidad o si solamente es considerada la informalidad dependiente.

Dando paso a la explicación institucional, la primer variable es CLU cuyo aumento en 1% en promedio produce un incremento de 0.01% en la TILD y de forma inversa. Este resultado y el de la tabla número 18 son casi iguales, por lo que el costo que representan los trabajadores para los empleadores si es un factor que ejerce influencia (aun cuando es pequeña) sobre el número de trabajadores a contratar como formales.

La COBERTURA resulta ser una variable cuyo incremento en 1% en promedio provoca un aumento de 0.04% en la TILD y viceversa. Para esta variable sucede algo distinto y es que el valor de la elasticidad es mayor en esta versión del modelo que en la versión global, en el primer caso es de 0.04 y en el segundo de 0.02. Se puede afirmar que las variaciones en la COBERTURA ejercen mayor influencia sobre la informalidad dependiente en comparación con el total de la informalidad. Esto puede deberse a que los individuos aceptan un empleo informal en el que saben de antemano que la atención médica no será otorgada por el empleador, basándose en que tienen la alternativa de acceder a atención médica proporcionada por el seguro popular.

A manera de cierre de este subapartado, se puede destacar que los resultados son acorde a lo esperado. La TIL responde a las variaciones de los factores duales e institucionales, sin embargo el mayor efecto sobre esta le concierne a los duales. Lo mismo ocurre para la TILI y para la TILD. Además el modelo que considera a TILD es mejor en términos de número de variables significativas en comparación con el que contempla a TILI pero no mejor que el modelo con TIL.

CAPÍTULO VI: CONCLUSIONES

La presente investigación tuvo como objetivo identificar los determinantes de la informalidad laboral en México durante el periodo 2005-2014. Desde el contexto teórico se buscó confrontar las dos grandes posturas que proporcionan la explicación del fenómeno, estas son la visión dual y la institucional. La característica distintiva entre ambas posturas es que la primera atribuye las variaciones de la informalidad a factores determinantes que escapan a la elección del individuo, es decir, la informalidad es un fenómeno netamente involuntario que se ve influenciado por aspectos como la situación de la economía y la estructura del mercado de trabajo. Por su lado, la explicación institucional va en el sentido voluntario, es decir, el individuo es quien elige pertenecer a la informalidad basándose en los costos de la formalidad así como de la propia informalidad y en la observación de ciertas acciones realizadas por el gobierno que propician o desalientan tal elección.

Para cumplir con el objetivo y verificar la hipótesis se planteó la especificación de un modelo que permitiera incluir las dos posiciones teóricas. Los insumos del modelo provienen de diversas fuentes y corresponden a 24 de las 32 entidades federativas de México. Por lo tanto, se conformó un modelo con datos en panel que fue estimado a través de diversas metodologías a las cuales se les aplicaron distintas pruebas estadísticas con la finalidad de seleccionar el mejor modelo, que resultó ser el de efectos fijos robusto a problemas de autocorrelación y heterocedasticidad.

Las variables que constituyen el mejor modelo son seis, por el lado dual son el Producto Interno Bruto (PIB), Escolaridad de la PEA (ESCO) y Sector Moderno (SMODERNO). Por el lado institucional las variables que se utilizan son el Costo Laboral Unitario (CLU), Eficiencia en Recaudación (ER) y Cobertura del Seguro Popular (COBERTURA). Cabe resaltar que el modelo fue estimado en tres versiones en las que la diferencia radica en la variable dependiente. La primera versión corresponde a aquella que tiene como variable dependiente al total de la Tasa de Informalidad Laboral (TIL). La segunda versión toma como dependiente a la Tasa de Informalidad Independiente (TILI) y la tercera versión a la Tasa de Informalidad Dependiente (TILD). El objetivo de este desglose fue determinar si existen diferencias entre los tipos de informalidad en cuanto a los factores que proveen su explicación. La hipótesis que se planteó fue que al igual que para la informalidad total, los factores determinantes de la informalidad desagregada corresponden a ambas visiones teóricas confrontadas. Los resultados de la

estimación del mejor modelo para todas las versiones, permiten corroborar que las relaciones entre las variables consideradas y la informalidad laboral son acordes con la teoría.

A continuación se realizan una serie de conclusiones por cada uno de los determinantes de la informalidad laboral en México, comenzando por el modelo general. Primeramente, la situación de la actividad económica es capturada por el comportamiento del PIB, bajo este escenario se encuentra que las mejoras en la actividad económica ejercen influencia negativa sobre la informalidad mediante dos vías; la primera es que tales mejoras se reflejan en el aumento de la capacidad del sector moderno o formal para absorber parte de la población desempleada que se encuentra en situación de búsqueda de empleo y que tiene como opción emplearse en el sector informal, es decir, disminuye la informalidad al incrementar las oportunidades para los individuos de emplearse en el lado formal de la economía ya que este a su vez se encuentra en una situación favorable. La segunda vía es el incremento de las posibilidades de que algunos de los individuos que se encuentran laborando en la informalidad de manera involuntaria, logren obtener un empleo en la formalidad, pues debe recordarse que la teoría dual reconoce la existencia de movilidad entre los sectores. Por el contrario, una situación económica difícil provocará que la informalidad laboral se incremente debido a la mayor expulsión de individuos de la formalidad (crecimiento del desempleo en el sector formal), o al incremento de las dificultades para encontrar un empleo formal tanto para los desocupados como para los informales que buscan moverse hacia un empleo de este tipo.

En este sentido, en términos de acciones de política, es importante el fomento a las actividades productivas para mejorar las condiciones laborales de la población en general y con ello reducir los niveles de informalidad en el país. Particularmente en una economía cuyo crecimiento ha sido bastante discreto, como es el caso de la economía mexicana.

Por otra parte, de acuerdo con la teoría dual los empleos en el sector secundario o informal (para el presente documento) no requieren de mano de obra que cuente con elevados niveles de escolaridad, por lo que más bien el sector informal se compone de empleos cubiertos con mano de obra no calificada y en esta dirección de análisis los resultados obtenidos permiten confirmar que la escolaridad (ESCO) puede influir en que sea mayor o menor el número de individuos que terminan empleándose en el lado informal de la economía. En síntesis, si los niveles de escolaridad de la PEA son bajos es de esperarse que los niveles de informalidad sean elevados

debido a que el acceso a empleos formales que requieren de mayor calificación de la mano de obra, se ve limitado.

El resultado del factor ESCO permite argumentar que es de gran importancia que los programas sociales implementados por el gobierno en términos de educación se vean fortalecidos y que continúen siendo dirigidos hacia la población de todas las edades, un buen ejemplo de esto es la educación para los adultos que provee el Instituto Nacional para la Educación de los Adultos (INEA). Asimismo debe hacerse énfasis en que el objetivo de las acciones debe ser además de ofrecer una educación de calidad, el de incrementar los años promedio de escolaridad a nivel nacional, es decir, lograr pasar el promedio de nueve años que hasta ahora se ha registrado y que corresponde a una educación a nivel secundaria, con ello debería observarse una disminución de la informalidad.

En cuanto al tamaño del sector moderno (SMODERNO), este ejerce un efecto negativo sobre la informalidad, si se incrementa. La informalidad tenderá a disminuir si la capacidad de empleo en el sector moderno aumenta, porque se reduce el número de individuos que son empujados al sector secundario o residual de la economía (en términos de la teoría dual) que es el sector informal, es decir hay menores niveles de desempleo. En el caso opuesto, cuando el sector moderno disminuye, la informalidad aumenta como respuesta al mayor número de individuos que salen de la formalidad y para los que la única opción de salida de tal situación de desempleo es el empleo en el lado informal de la economía ya sea como subordinados, trabajadores por cuenta propia o empleadores informales.

Los efectos del SMODERNO permiten resaltar dos cosas importantes en términos de políticas, una es que todas aquellas acciones destinadas a dinamizar este sector permitirán cumplir con el objetivo de disminuir los niveles de informalidad e incrementar la calidad de empleo al que accede la población o al menos asegurar que los contratos sean formales y que los individuos reciban atención médica de mejor calidad. Lo segundo es que el enlace que realiza el gobierno entre empleadores y trabajadores a través de programas como las ferias de empleo, debe continuar siendo impulsado porque facilita que algunos de los individuos que se encuentran buscando empleo logren obtener uno que es formal, o bien aquellos que son informales tengan a su disposición opciones en el lado formal, lo que se vería reflejado en la reducción de la informalidad.

Considerando los factores institucionales que proveen la explicación del fenómeno, la variable asociada al costo laboral CLU permite concluir que aumentos en los costos laborales impactan de manera positiva a la informalidad, provocando que crezca. Esta variable logra capturar la decisión por el lado de los individuos empleadores o bien desde la demanda de trabajo, en cuanto al número de trabajadores que se deciden emplear y la forma en que se hace (bajo el esquema formal o bajo el informal). La obligación formal de pagar un salario y demás prestaciones se vuelve en contra de la informalidad cuando estos sufren incrementos ya que son parte del costo laboral que debe asumir el empleador. Entonces si el costo aumenta, se demandarán menos trabajadores en el lado formal de la economía y como consecuencia la informalidad tenderá a repuntar. Sin embargo, no es la única vía por la que la informalidad puede incrementarse como consecuencia del aumento del CLU, también puede presentarse la situación en la que el incremento del costo laboral propicie que las empresas formales empleen individuos pero bajo condiciones de informalidad, logrando así reducir sus obligaciones formales y por lo tanto su costo laboral.

Por otro parte, la inclusión de la eficiencia en recaudación (ER) como un factor explicativo tuvo la finalidad de determinar de qué forma influyen las acciones del gobierno sobre la informalidad, al menos en estos términos. Se parte de asumir que el individuo toma decisiones basándose en lo que percibe sobre determinadas instituciones, en este caso decide voluntariamente entre ser formal o informal. Los resultados señalan que el individuo percibe de forma negativa un incremento en la eficiencia, ello puede deberse a que asocia esa mejora con una mayor extracción de impuestos, entonces si su objetivo es disminuir los costos de la formalidad o evadir obligaciones, optará por la informalidad ya que la mayor eficiencia del gobierno se interpone con su objetivo. De esta manera, la forma que tiene el gobierno para desincentivar este comportamiento, es que el individuo vea reflejada su contribución tributaria en obras que lo beneficien, como en la educación, salud, seguridad pública, entre otras.

El impacto positivo de COBERTURA sobre la informalidad señala que los servicios médicos que ofrece el seguro popular son vistos como sustituto de la atención médica a la que accede el individuo cuando labora en el lado formal de la economía. En este caso, la atención médica otorgada por el seguro popular puede disminuir los costos de permanecer en el lado informal, por lo que a mayor cobertura del seguro popular mayores serán los niveles de informalidad

observados en el país. Sin embargo este resultado no debe ser interpretado como una prueba para decir que es un programa social del cual debe prescindirse, ya que sus beneficios en el sentido social han sido amplios y ha mostrado un gran impacto año tras año al poner al alcance de un mayor número de individuos los servicios médicos a los que no tendrían acceso.

Una vez que se reconoce que la mayor COBERTURA incrementa la informalidad, lo que se propone es revisar la configuración del seguro popular en términos de las condiciones en las que se otorga este beneficio a los individuos, es decir, debe ponerse especial atención sobre a quienes se les está concediendo y bajo qué términos. Asimismo, se vuelve evidente la parte voluntaria del fenómeno en la que el individuo observa las acciones gubernamentales y decide entre formar parte de la formalidad o de la informalidad basándose en los beneficios que puede obtener de una u otra posición en el mercado laboral, pero sobretodo las decisiones se encuentran basadas en los costos que puede evitar ya que se reconoce que es menos costosa la atención en el seguro popular que los servicios médicos del sector formal que son descontados de forma constante del salario que perciben.

Por otro lado, cuando se realiza la estimación considerando que la informalidad puede desglosarse en independiente y dependiente los resultados señalan que si existe diferencia entre ambos tipos de informalidad al menos en los factores que los explican, en tanto que las relaciones entre las variables se mantienen y concuerdan con lo esperado a nivel teórico y empírico. En el caso específico de la informalidad del tipo independiente, sus variaciones solamente pueden ser atribuidas a dos factores que son el SMODERNO y la ER, donde el primero tiene mayor impacto y corresponde a la perspectiva dual, en tanto que el segundo posee un impacto muy cercano a cero y forma parte de la visión institucional. Es decir, ambas posturas ejercen influencia en las variaciones de la informalidad independiente, y este hecho coincide con los resultados del modelo para el total de la informalidad.

La informalidad dependiente (a diferencia de la independiente) mostró resultados más parecidos a los de la informalidad total en términos del número de variables que son significativas en su explicación. De las seis variables consideradas, solamente la ER no contribuye en la explicación de las variaciones del fenómeno de investigación. La educación por su parte, es la variable con el mayor impacto e incluso este es mayor en comparación con el modelo que considera el total de la informalidad laboral. De aquí que se haga hincapié en su relevancia.

En términos generales los determinantes de la informalidad en México durante el periodo 2005-2014 son el PIB, ESCO, SMODERNO, CLU, ER y COBERTURA. Los factores corresponden a ambas posturas teóricas confrontadas en la presente investigación, o bien la informalidad laboral es un fenómeno tanto involuntario como voluntario. Por esta razón puede decirse que su explicación corresponde a un enfoque mixto (dual-institucional).

Los resultados encontrados en esta investigación confirman la hipótesis que sostiene que si bien los determinantes de la informalidad son en ambas direcciones teóricas, los duales son los que mayor efecto tienen sobre la informalidad ya que los valores de las elasticidades son superiores en comparación con los institucionales. Por ejemplo en el caso de la informalidad total, si se comparan los valores duales (en términos absolutos) de 0.04 para el PIB, 0.28 para ESCO y 0.13 para SMODERNO contra valores institucionales de 0.01 para CLU, 0.003 para ER y 0.02 para COBERTURA, resulta que los duales son mayores. Nótese que en general los valores de las elasticidades son bajos, sin embargo la significancia estadística que poseen les concede gran importancia en la explicación del fenómeno.

Dado todo lo anterior es posible afirmar que si el objetivo gubernamental es disminuir la informalidad laboral entonces las acciones deben dirigirse a fortalecer los niveles de educación de la población debido a que es el factor con mayor impacto sobre la tasa de informalidad laboral total. En segundo lugar, deben tomarse acciones encaminadas a incrementar la capacidad de absorción de empleo por parte del sector formal que a la vez se encuentra relacionado con las acciones dirigidas a dinamizar las actividades productivas y en general contribuir al buen comportamiento de la economía. Es decir, las acciones de mayor impacto deberían ir en el sentido teórico dual.

Los resultados permiten realizar recomendaciones específicas de política pública. El objetivo es fomentar algunos comportamientos que mejoren determinados aspectos del fenómeno y/o disuadir otros que agraven la situación. Para ello se puede hacer uso de políticas activas y/o pasivas. Las medidas que a continuación se sugieren, coinciden más con el sentido de las políticas activas ya que se caracterizan por lidiar con las causas del fenómeno y que en el ámbito de esta investigación se han tratado como los factores que determinan la informalidad laboral. Una primera acción es encauzar los esfuerzos para promover la movilidad de los individuos desde la informalidad hacia la formalidad, en dos sentidos; uno es fomentar los programas de

formalización dirigidos a los individuos que cuentan con un negocio informal, ya sea como empleadores o como trabajadores por cuenta propia. El otro sentido es enlazar la demanda de trabajo formal con los individuos que se encuentran en la informalidad subordinados a un patrón, es decir buscar que logren ubicarse en un empleo formal.

Adicionalmente, no debe olvidarse que si el objetivo de las políticas es disminuir la informalidad, también deben contemplarse los desempleados que pueden obtener un empleo en el lado formal o informal, entonces lo que se pretende es que los esfuerzos propicien que tal empleo sea formal. Por lo tanto, para hacer coincidir la oferta de trabajo con empleos formales se utiliza la intermediación laboral a través de las distintas opciones que forman parte del Programa de Apoyo al Empleo tales como las bolsas de trabajo, portales de empleo, ferias de empleo, servicio nacional de empleo por teléfono, observatorio laboral y talleres para buscadores de empleo. Principalmente estas herramientas se encuentran dirigidas a los desempleados y subempleados pero bien pueden extenderse a los informales ya que la finalidad es que los individuos que las utilizan encuentren un empleo formal o se autoempleen de manera formal. Con ello el bienestar de la población en general se verá incrementado, al cubrir gran parte de la oferta de trabajo.

Otra medida de política activa puede ser la de mejorar la formación de la oferta de trabajo para que los individuos logren obtener un empleo formal que requiere de habilidades y conocimientos específicos. Esto se logra mediante la capacitación ofrecida por parte del gobierno, por ejemplo en los centros de formación para el trabajo que consideran diferentes áreas. Sin embargo estos programas no son lo suficientemente conocidos por la población, por lo que habría que hacer una mayor difusión de la información. Algunas otras acciones ya no en el sentido de políticas del mercado de trabajo, como es revisar el acceso al seguro popular, también pueden ayudar a disminuir los niveles de informalidad, se menciona este por ser el más relevante dentro de la explicación institucional.

Finalmente pueden señalarse algunas limitantes del trabajo así como las posibles líneas de investigación que se desprenden de este documento. Dentro de las limitantes la primera a señalar es el periodo de análisis, debido a que solamente abarca 10 años y que por lo tanto es algo restrictivo si la intención es realizar un análisis de cointegración para el panel, sin embargo la fuente de la limitante son las estadísticas de la informalidad ya que no es posible obtener datos de este fenómeno para años previos. La segunda limitante se refiere al número de entidades

consideradas, es decir, sólo son incluidas 24 de las 32 entidades de la federación lo que puede verse como una restricción para que los resultados se generalicen a nivel país, pero se debe tener presente que la muestra incluye a la mayoría de las entidades por lo que es lo suficientemente representativa. La razón para excluir 8 entidades fue la falta de información para una de las variables institucionales y el hecho de incluirlas podría restarle comparabilidad a los resultados, sobretodo porque uno de los objetivos de esta investigación fue precisamente la comparabilidad. Como tercera limitante, es el hecho de que se consideró el sector manufacturero como el sector moderno a pesar de que algunas de las actividades dentro de este no poseen las características suficientes (como es el caso de la industria maquiladora) para considerarse como el sector moderno o primario, sin embargo recuérdese que se considera el total y que contribuye en gran medida a la actividad económica del país, al menos para el periodo 2005-2014 representa el 17% en promedio del total del PIB, por lo cual puede tomarse como referencia. También debe tenerse presente que dentro de la industria manufacturera se realizan actividades que hacen uso de alta tecnología como es la industria aeroespacial, farmacéutica y automotriz, por ejemplo. Además la información desglosada por actividad dentro del sector no se encuentra para todas las entidades ni en todos los años, lo que haría perder más información. Nótese que las tres limitantes giran en torno a la falta de información.

Finalmente, dentro de las posibles líneas de investigación se propone en primer lugar hacer uso de variables (ya sea que se obtengan directamente de alguna fuente o que se construyan) duales e institucionales que permitan incluir el total de entidades. Como segunda, realizar el análisis comparativo de los factores determinantes de la informalidad dentro y fuera del sector informal ya que esta última es ligeramente superior a la primera, por lo tanto puede existir diferencia entre los factores que determinan a cada una o bien en la magnitud en la que lo hacen. Como tercera, se propone investigar cómo son las condiciones de los empleos informales fuera del sector informal es decir, son mejores o peores que los que se encuentran dentro del sector informal. Por último, una parte de la perspectiva voluntaria de la informalidad puede ser estudiada a través de los resultados de la Encuesta Nacional de Micronegocios (ENAMIN) considerando solamente a las empresas informales, ya que se les pregunta a los individuos si aceptarían un trabajo asalariado con un ingreso similar a lo que actualmente ganan en su negocio o actividad, pero donde tuvieran acceso a los servicios del IMSS o bien si cambiarían su negocio o actividad actual, si les ofrecieran un trabajo con un ingreso similar al que tienen actualmente y además tuvieran un

derecho a pensión, asimismo se les pregunta cuál es el salario mensual por el que aceptarían dejar su negocio o actividad. Es decir, se puede determinar qué cantidad de individuos informales (al menos en los micronegocios) estarían dispuestos a moverse a la formalidad bajo el supuesto de que un trabajo formal es aquel que ofrece IMSS o pensión. También se comprobaría de cierto modo qué cantidad de individuos se encuentran de manera voluntaria en la informalidad al mando de un micronegocio (considerando que son aquellos que no desean moverse), o bien cuáles son las características de estos en comparación con aquellos que si estarían dispuestos a cambiarse a un empleo formal y que por lo tanto puede suponerse se encuentran de manera involuntaria en la informalidad. En resumen, sería posible realizar un análisis de informalidad laboral involuntaria-voluntaria a nivel de individuos en micronegocios.

REFERENCIAS

- Aguilera, N. y Velázquez, C. (1994). Los efectos de la informalidad. *Seguridad Social*, 1-7.
- Anton, A., Hernández, F. y Levy, S. (2011). The End of Informality in Mexico? Fiscal Reform for Universal Social Insurance. Inter-American Development Bank.
- Badaoui, E., Strobl, E. y Walsh, F. (November, 2007). Is There An Informal Employment Wage Penalty? Evidence from South Africa. *Discussion Paper Series. IZA. 3151*, 1-36.
- Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*. Fourth Edition. Chichester: John Wiley & Sons.
- Baltagi, B. H. (2011). *Econometrics*. Berlin: Springer
- Baltagi, B. H, Mátyás, L. y Sevestre, P. (2008). Error components models. Third Edition. En L. Mátyás y P. Sevestre (Eds), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*. Germany: Springer.
- Banco Interamericano de Desarrollo (BID). (2003). *Se buscan buenos empleos: Los mercados laborales en América Latina*. Washington, D.C.: BID.
- Bangasser, P. (2000). The ILO and the informal sector: an institutional history. *Employment Paper*. 2000/9. Geneva.
- Bauder, H. (2001). Culture in the labor market: segmentation theory and perspectives of place. *Progress in Human Geography*, 25(1), 37-52.
- Baum, C. (2006). *An introduction to modern econometrics using STATA*. United States of America: Stata Press.
- Baum, C. (2007). Instrumental Variables Estimation in Stata. *PDF Presentation*. Faculty Micro Resource Center. Boston College. March, 2007.
- Betcherman, G. (2013). Labor Market Institutions: A Review of the Literature. Background Paper for the World Development Report 2013. The World Bank.
- Bhatti, S., Bourdon, J. y Aslam, M. (2013). Economic Returns to Education in France: OLS and Instrumental Variable Estimations. *The Lahore Journal of Economics*, 18(2), 51-63.

Bivens, J. y Gammage, S. (2005). Will better workers lead to better jobs in the developing world? En T. Avirgan, L. J. Bivens y S. Gammage (Eds.), *Good Jobs, Bad Jobs, No Jobs*, Global Policy Network, Washington, DC.: Economic Policy Institute.

Bulow, J. y Summers, L. (July, 1985). A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination and Keynesian Unemployment. *NBER Working Paper*, 1666, 1-49.

Busso, M., Fazio, M. y Levy, S. (August, 2012). (In) Formal and (Un) Productive: The Productivity Costs of Excessive Informality in Mexico. *IDB Working Paper*, 341, 1-53.

Cain, G. (1976). The Challenge of Segmented Labor Market Theories to Orthodox Theory: A Survey. *Journal of Economic Literature*, 14(4), 1215-1257.

Cameron, A.C. y Trivedi, P. (2005). *Microeconometrics: methods and applications*. United States of America: Cambridge University Press.

Cameron, A.C. y Trivedi, P. (2013). Count Panel Data. Paper prepared for the Badi H. Baltagi ed., *Oxford Handbook of Panel Data Econometrics*, Oxford University Press.

Card, D. (1993). Using Geographic Variation in College Proximity to Estimate the Return to Schooling. *NBER Working Paper Series*, 4483, 1-29.

Chen, M. (2007). Rethinking the Informal Economy: Linkages with the Formal Economy and the Formal Regulatory Environment. *DESA Working Paper*, 46, 1-12.

Chen, M. (August, 2012). The Informal Economy: Definitions, Theories and Policies. *WIEGO Working Paper*, 1, 1-22.

Choi, J.P. y Thum, M. (2003). Corruption and the shadow economy. *Dresden Discussion Paper Series in Economics*, 2/03.

Davidson, R. y MacKinnon, J. (1993). *Estimation and inference in econometrics*. New York: Oxford University Press.

Dickens, W. y Lang, K. (1985). A Test of Dual Labor Market Theory. *The American Economic Review*, 75(4), 792-805.

Dickens, W. y Lang, K. (June, 1992). Labor Market Segmentation Theory: Reconsidering the Evidence. *NBER Working Paper*, 4087, 1-80.

Dreher, A. y Schneider, F. (January, 2006). Corruption and the Shadow Economy: An Empirical Analysis. IZA. *Discussion Paper Series, 1936*, 1-37.

Elgin, C. y Uras, B. (2013). Is informality a barrier to financial development?. *SERIEs*. Journal of the Spanish Economic Association, *4*, 309-331.

Fields, G. S. (1972). Rural-Urban Migration, Urban Employment and Underemployment, and Job Search Activity in LDCs [Electronic version]. Washington, DC: U.S. Agency for International Development.

Fields, G.S. (1990). Labour market modeling and the urban informal sector: Theory and evidence [Electronic version]. En D. Turnham, B. Salomé, & A. Schwarz (Eds.), *The informal sector revisited* (pp. 49-69). Paris: Organization for Economic Co-operation and Development.

Fields, G.S. (2004). Dualism in the labor market: A perspective on the Lewis Model After Half a Century. Retrieved [January 2015] from *Cornell University*, ILR school site: <http://digitalcommons.ilr.cornell.edu/articles/268/>

Fields, G.S. (April, 2005). A Guide to Multisector Labor Market Models. *Social Protection Discussion Paper Series, 0505*, 1-50.

Fields, G.S. (September, 2007). Labor Market Policy in Developing Countries: A Selective Review of the Literature and Needs for the Future. *Policy Research Working Paper*. The World Bank, *4362*, 1-76.

Friedman, E., Johnson, S., Kaufmann, D. y Zoido-Lobaton, P. (2000). Dodging the grabbing hand: the determinants of unofficial activity in 69 countries. *Journal of Public Economics*, *76*, 459-493.

Greene, W. (2008). *Econometric Analysis*. Fifth Edition. NJ: Prentice Hall.

Greene, W. (2012). *Econometric Analysis*. Séptima Edición (Internacional). London: Pearson Education.

Guha-Khasnobis, B., Kanbur, R. y Ostrom, E. (2006). Beyond Formality and Informality. In Guha-Khasnobis, B., Kanbur, R. y Ostrom, E. (eds.). *Linking the Formal and Informal Economy: Concepts and Policies*. Oxford: Oxford University Press.

Gujarati, D. y Porter, D. (2010). *Econometría*. Quinta Edición. México: McGraw-Hill.

Günther, I. y Launov, A. (September, 2006). Competitive and Segmented Informal Labor Markets. *Discussion Paper Series*. IZA, 2349, 1-29.

Harris, J. y Todaro, M. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. *The American Economic Review*, 60, 126-142.

Harrison, B. y Sum, A. (September, 1979). The Theory of “Dual” or Segmented Labor Markets. *Journal of Economic Issues*, XIII(3), 687-693.

Hart, K. (March, 1973). Informal Income Opportunities and Urban Employment in Ghana. *The Journal of Modern African Studies*, 11(1), 61-89.

Heckman, J. y Pagés, C. (August, 2000). The Cost of Job Security Regulation: Evidence from Latin American Labor Markets. *BID. Working Paper*, 430, 1-44.

Hibbs, D. A. y Piculescu, V. (2005). Institutions, Corruption and Tax Evasion in the Unofficial Economy. *Working Papers in Economics*. CEFOS and Department of Economics, Göteborg University.

Hodson, R. y Kaufman, R. (December, 1982). Economic Dualism: A Critical Review. *American Sociological Review*, 47(6), 727-739.

Hsiao, C. (2014). *Analysis of Panel Data*. Third Edition. United States of America: Cambridge University Press.

Huesca, L. y Padilla, M. (Noviembre, 2012). Empleo, escolaridad y sector informal en la Frontera Norte de México y Chihuahua: expectativas de ocupación en la crisis. *Ensayos de Economía*, XXXI(2), 57-86.

Instituto Mexicano para la Competitividad. (IMCO). 2012. Índice de Competitividad Estatal 2012. Recuperado de http://imco.org.mx/indice_de_competitividad_estatal_2012/archivos/libro/Indice_de_Competitividad_Estatal_2012.pdf

Instituto Nacional de Estadística y Geografía. (INEGI). 2014. La informalidad laboral. Encuesta Nacional de Ocupación y Empleo. Marco conceptual y metodológico. ISBN: 978-607-739-109-8. Recuperado de <http://www3.inegi.org.mx/sistemas/biblioteca/detalle.aspx?c=265&s=inegi&upc=702825060459&pf=prod&f=2&cl=0&tg=0>

- Jiménez, D. (Julio/Diciembre, 2012). La informalidad laboral en América Latina: ¿Explicación estructuralista o institucionalista?. *Cuadernos de Economía*, XXI(58), 113-143.
- Jiménez, M. (Marzo, 2013). La informalidad laboral en el sector formal. Un análisis preliminar. *Documentos de trabajo*, 10, 1-65.
- Johnston, J. y DiNardo, J. (1997). *Econometric Methods*. Fourth Edition. United States of America: McGraw-Hill.
- Krakowski, M. (2005). Determinants of the Informal Economy: The importance of Regional Factors. *HWWA Discussion Paper*, 313, 1-28.
- Kucera, D. y Roncolato, L. (2008). El trabajo informal. Dos asuntos clave para los programas políticos. *Revista Internacional del Trabajo*, 127(4), 357-387.
- La Porta, R. y Shleifer, A. (December, 2008). The Unofficial Economy and Economic Development. *NBER Working Paper Series*, 14520, 1-41.
- Leal, J. (November, 2014). The informal sector in contemporary models of the aggregate economy. *Banco de Mexico Working Papers*. 2014-24, 1-20.
- Leontaridi, M. (February, 1998). Segmented Labour Markets: Theory and Evidence. *Journal of Economic Surveys*, 12(1), 63-101.
- Levy, S. (2008). Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico. Washington, D.C.: Brookings Institution Press.
- Lewis, W. (May, 1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, 22(2), 139-191.
- Loayza, N. (February, 1997). The Economics of the Informal Sector. A Simple Model and Some Empirical Evidence from Latin America. *The World Bank. Policy Research Working Paper*. 1727, 1-52.
- Losby, J., Else, J., Kingslow, M., Edgcomb, E., Malm, E. y Kao, V. (December, 2002). Informal Economy Literature Review. Institute for Social and Economic Development and the Aspen Institute.
- Maloney, W. (1999). Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico. *The World Bank Economic Review*, 13(2), 275-302.

Maloney, W. (2003). Informal Self-Employment: Poverty Trap or Decent Alternative?. In Gary S. Fields and Guy Pfefferman, eds., *Pathways Out of Poverty*. Boston: Kluwer.

Maloney, W. (2004). Informality Revisited. *World Development*, 32(7), 1159-1173.

Maloney, W. y Arias O. (2008). El sector informal: La razón de ser del trabajador informal en *Informalidad: Escape y Exclusión*. Versión en Castellano. Banco mundial en coedición con Mayol Ediciones S.A.

Maloney, W. y Saavedra J. (2008). El sector informal: ¿Qué es, por qué nos importa y cómo lo medimos? en *Informalidad: Escape y Exclusión*. Versión en Castellano. Banco mundial en coedición con Mayol Ediciones S.A.

Marcouiller, D., Ruiz, V. y Woodruff, C. (Marzo, 1995). Formal Measures of the Informal Sector Wage Gap in Mexico, El Salvador, and Peru. Economics Department, Boston College, *Working Papers in Economics*, 1-29.

Mayorga, M. y Muñoz, E. (Septiembre, 2000). La técnica de datos de panel. Una guía para su uso e interpretación. *Documento de trabajo*. Banco Central de Costa Rica. Departamento de Investigaciones Económicas, 1-18.

McConnell, C., Brue, S. y Macpherson, D. (2003). *Economía Laboral*. España: McGraw-Hill.

Montero, R. (2011). Efectos fijos o aleatorios: test de especificación. *Documentos de Trabajo en Economía Aplicada*. Universidad de Granada. España.

Moreno, J. O. (Mayo, 2007). Los salarios del sector formal e informal en México: análisis de ganancias y pérdidas por formalización. *Ensayos*, XXVI(1), 1-44.

Negrete, R. y Eternod, M. (2011). Informalidad y Género. XII Encuentro Internacional de Estadísticas de Género: empoderamiento, autonomía económica y políticas públicas. Instituto Nacional de Estadística y Geografía (INEGI).

Nerlove, M. (2005). *Essays in Panel Data Econometrics*. New York: Cambridge University Press.

Organización Internacional del Trabajo. OIT. (2012). Guía de recursos sobre economía informal. Recuperado de <http://ilo.org/public/spanish/support/lib/resource/subject/informal.htm>. Septiembre de 2014.

Organización Internacional del Trabajo. OIT. (2013). La medición de la informalidad: Manual estadístico sobre el sector informal y el empleo informal. Recuperado de http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_222986.pdf.

Septiembre de 2014.

Organización para la Cooperación y el Desarrollo Económicos. OCDE. (2004). Perspectivas del Empleo de la OCDE. Publicaciones de la OCDE.

Organización para la Cooperación y el Desarrollo Económicos. OCDE. (Septiembre, 2014). OCDE Perspectivas del Empleo de la OCDE 2014. ¿Cómo se sitúa México?. Publicaciones de la OCDE.

Packard, T. (May, 2007). Do Workers in Chile Choose Informal Employment? A Dynamic Analysis of Sector Choice. *World Bank Policy Research Working Paper*, 4232, 1-56.

Perry, G., Maloney, W., Arias, O., Fajnzylber, P., Mason, A. y Chanduvi, J. (2008). *Informalidad: Escape y Exclusión*. Versión en Castellano. Bogotá, Colombia: Banco mundial en coedición con Mayol Ediciones S.A..

Pickhardt, M. y Prinz, A. (2012). *Tax Evasion and the Shadow Economy*. United Kingdom: Edward Elgar.

Pindyck, R. y Rubinfeld, D. (2001). *Econometría: modelos y pronósticos*. Cuarta edición. Mexico: McGraw-Hill.

Piore, M. J. (Octubre, 1972). Notes for a Theory of Labor Market Stratification. *MIT Working Paper Department of Economics*, 95, 1-40.

Rebitzer, J. y Robinson, M. (January, 1991). Employer Size and Dual Labor Markets. *NBER Working Paper*, 3587, 1-17.

Rei, D. y Bhattacharya, M. (April, 2008). The impact of Institutions and Policy on Informal Economy in Developing Countries: An econometric exploration. Policy Integration and Statistics Department, International Labour Office. *Working Paper*, 84, 1-50.

Reich, M., Gordon, D. y Edwards, R. (May, 1973). Dual Labor Markets: A Theory of Labor Market Segmentation. *American Economic Review*, 63(2), 359-365.

Riascos, J. C. (2007). Análisis introductorio al mercado dual de trabajo. *Tendencias*, 8(2), 67-78.

- Robles, L. M. (2008). *The Economics of Tax Enforcement*. United States: ProQuest.
- Rodríguez, E. (2007). La dinámica comparativa del sector informal en México. Universidad Iberoamericana. *Serie Documentos de Investigación*, 19, 1-53.
- Rodríguez, E., Lima, J. y Mulato, D. (Septiembre/Diciembre, 2013). Competencia de mercado e informalidad en el sector manufacturero mexicano 1987-2004. *Economía, Sociedad y Territorio*, XIII(43), 721-750.
- Roldan, P. y Ospino, C. (2009). ¿Quiénes terminan en la informalidad?: Impacto de las características y el tiempo de búsqueda. *Revista de economía del caribe*, 4, 149-180.
- Santana, A. (Coord.). (2011). El Seguro Popular de Salud en México: Posibles efectos sobre la economía informal. *Avance-Análisis, Investigación y Estudios para el Desarrollo A.C.*
- Sarte, P. (1999). Informality and Rent-Seeking Bureaucracies in a Model of Long-Run Growth. *Working Paper Series*. Federal Reserve Bank of Richmond, 99-07, 1-57.
- Schneider, F. y Enste, D. (2002). *Ocultándose en las sombras. El crecimiento de la economía subterránea*. Temas de Economía. Fondo Monetario Internacional. Washington.
- Schneider, F. y Klinglmaier, R. (March, 2004). Shadow Economies around the World: What Do We Know?. *Discussion Paper*. IZA, 1043, 1-57.
- Schneider, F. y Williams, C. (2013). *The Shadow Economy*. Institute of Economic Affairs. London. Recuperado de http://works.bepress.com/colin_williams/23/
- Singh, A., Jain-Chandra, S. y Mohommad, A. (February, 2012). Inclusive Growth, Institutions, and the Underground Economy. *IMF Working Paper*, WP/12/47, 1-13.
- Tannuri-Pianto, M. y Pianto, D. (2002). Informal Employment in Brazil- A Choice at the Top and Segmentation at the Bottom: A Quantile Regression Approach. Paper presented at the XXIV Brazilian Econometrics Meeting, Rio de Janeiro, Brazil.
- Tokman, V. (1978). Las relaciones entre los sectores formal e informal. Una exploración sobre su naturaleza. *Revista de la CEPAL*, 5, 103-141.
- Ulyssea, G. (February, 2014). Firms, Informality and Welfare. *Job Market Paper*. University of Chicago and IPEA, 1-46.

Uribe, J. y Ortiz, C. (Octubre, 2004). Una propuesta de conceptualización y medición del sector informal. *Documentos de trabajo*. Universidad del Valle, Departamento de Economía, Centro de Investigaciones y Documentación Socioeconómica, 1-35.

Uys, M.D. y Blaauw, P.F. (2006). The dual labor market theory and the informal sector in South Africa. *Acta Commercii*, 6(1), 248-257.

Varela, R. (2015). Diferencias de ingreso en la Frontera Norte de México: Una perspectiva de empleo formal e informal. *Frontera Norte*, 27(53), 177-203.

Wang, G. y Jain, C. (2003). *Regression Analysis: Modeling and Forecasting*. United States of America: Graceway.

Watcher, M. (1974). Primary and Secondary Labor Markets: A Critique of the Dual Approach. *Brookings Papers on Economic Activity*, 3, 637-693.

Webb, J., Bruton, G., Tihanyi, L. y Ireland, R. D. (2012). Research on entrepreneurship in the informal economy: Framing a research agenda. *Journal of Business Venturing*, doi:10.1016/j.jbusvent.2012.05.003.

Webb, J., Tihanyi, L., Ireland, D. y Sirmon, D. (2009). You say illegal, I say legitimate: Entrepreneurship in the informal economy. *Academy of Management Review*, 34(3), 492-510.

Wei, William. (2006). *Time series analysis: univariate and multivariate methods*. Second Edition. United States of America: Pearson Addison Wesley.

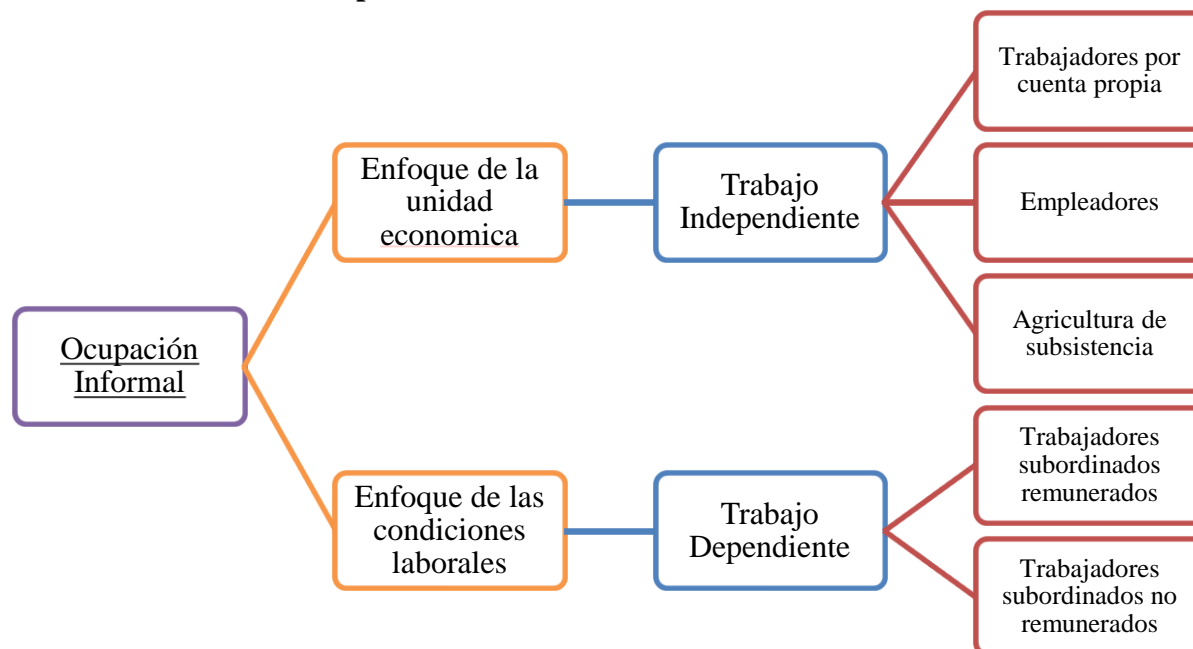
Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wooldridge, J.M. (2010). *Introducción a la econometría: un enfoque moderno*. Cuarta edición. México: CENGAGE Learning.

Wooldridge, J. (2015). Linear panel data models. Estimate. *PDF Presentation*. Department of Economics. Michigan State University. East Lansing, Michigan. May 29-31, 2015.

ANEXOS

A. Clasificación de la Ocupación Informal en México.



Fuente: Elaboración propia con datos de INEGI (2014).

B. Resultados de las Estimaciones

B.1. Estimaciones con TIL

B.1.1. Resultados y Pruebas de la Regresión con todas las Variables Explicativas

Regresión Agrupada

```
. reg til pib esco ndes smoderno clu er ip cobertura, cluster(entidad)
```

```
Linear regression                               Number of obs =      240
                                                F(  8,   23) =    26.38
                                                Prob > F      =    0.0000
                                                R-squared     =    0.8203
                                                Root MSE     =    4.6342
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	.1037176	1.315536	0.08	0.938	-2.617676	2.825111
esco	-83.76338	9.695277	-8.64	0.000	-103.8196	-63.70717
ndes	2.561326	.9815176	2.61	0.016	.530902	4.59175
smoderno	-.6297961	.1913587	-3.29	0.003	-1.025652	-.2339405
clu	-1.372	1.485637	-0.92	0.365	-4.445274	1.701275
er	-.0476822	.0228279	-2.09	0.048	-.0949052	-.0004592
ip	-.2898746	.5822781	-0.50	0.623	-1.494409	.9146594
cobertura	.1817399	.0393112	4.62	0.000	.1004185	.2630614
_cons	220.6786	19.61986	11.25	0.000	180.0918	261.2654

Fuente: Resultados de estimación.

Regresión con Efectos Fijos

```
. xtreg til pib esco ndes smoderno clu er ip cobertura, fe cluster(entidad)

Fixed-effects (within) regression      Number of obs   =      240
Group variable: entidad                Number of groups =      24

R-sq:  within = 0.4477                Obs per group:  min =      10
      between = 0.6626                    avg =      10.0
      overall = 0.6465                    max =      10

                                          F(8,23)        =      25.34
corr(u_i, Xb) = 0.6008                 Prob > F        =      0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	-3.113807	1.55215	-2.01	0.057	-6.324673	.097059
esco	-29.93756	7.840071	-3.82	0.001	-46.15598	-13.71913
ndes	3.917775	.6639923	5.90	0.000	2.544202	5.291348
smoderno	-.3233208	.1104091	-2.93	0.008	-.5517195	-.0949221
clu	.4209024	.2285542	1.84	0.078	-.051898	.8937029
er	-.0098891	.0040969	-2.41	0.024	-.0183642	-.0014141
ip	.0295421	.1579924	0.19	0.853	-.2972901	.3563743
cobertura	.0275024	.0177763	1.55	0.135	-.0092707	.0642755
_cons	125.356	23.53947	5.33	0.000	76.66088	174.0511
sigma_u	8.0356948					
sigma_e	1.1347556					
rho	.98044839	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Regresión con Efectos Aleatorios

```
. xtreg til pib esco ndes smoderno clu er ip cobertura, re cluster(entidad)

Random-effects GLS regression      Number of obs   =      240
Group variable: entidad                Number of groups =      24

R-sq:  within = 0.4347                Obs per group:  min =      10
      between = 0.7445                    avg =      10.0
      overall = 0.7317                    max =      10

                                          Wald chi2(8)    =      215.51
corr(u_i, X) = 0 (assumed)           Prob > chi2     =      0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
pib	-3.235941	.8722988	-3.71	0.000	-4.945616	-1.526267
esco	-41.78795	6.603718	-6.33	0.000	-54.731	-28.8449
ndes	3.846723	.6477416	5.94	0.000	2.577173	5.116273
smoderno	-.3836386	.1035731	-3.70	0.000	-.5866381	-.1806391
clu	.389657	.2884051	1.35	0.177	-.1756067	.9549207
er	-.0116469	.0043806	-2.66	0.008	-.0202327	-.003061
ip	-.0182453	.1652638	-0.11	0.912	-.3421565	.3056659
cobertura	.0607599	.0149029	4.08	0.000	.0315507	.089969
_cons	154.6195	17.28122	8.95	0.000	120.7489	188.4901
sigma_u	4.9317653					
sigma_e	1.1347556					
rho	.94971998	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Prueba F. Regresión Agrupada vs Efectos Fijos

Prueba F	
F (7, 23) =	25.34
Prob > F =	0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos

Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios

Breusch and Pagan Lagrangian multiplier test for random effects

$til[entidad,t] = Xb + u[entidad] + e[entidad,t]$

Estimated results:

	Var	sd = sqrt(Var)
til	115.5133	10.74771
e	1.28767	1.134756
u	24.32231	4.931765

Test: Var(u) = 0

chibar2(01) = 799.05

Prob > chibar2 = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Aleatorios

Prueba de Sobreidentificación. Efectos Fijos vs Efectos Aleatorios

Test of overidentifying restrictions: fixed vs random effects		
Cross-section time-series model: xtreg re robust cluster(entidad)		
Sargan-Hansen statistic = 137.213	Chi-sq(8)	P-value = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos.

B.1.2. Resultados y Pruebas de la Regresión sin IP

Regresión Agrupada

```
. reg til pib esco ndes smoderno clu er cobertura, cluster(entidad)
```

```
Linear regression                               Number of obs =      240
                                                F(   7,   23) =    20.85
                                                Prob > F       =    0.0000
                                                R-squared     =    0.8199
                                                Root MSE     =    4.6298
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	.1438717	1.298251	0.11	0.913	-2.541765	2.829509
esco	-83.34879	10.10196	-8.25	0.000	-104.2463	-62.4513
ndes	2.390303	.9663001	2.47	0.021	.3913587	4.389247
smoderno	-.6260104	.1915323	-3.27	0.003	-1.022225	-.2297957
clu	-1.392523	1.46135	-0.95	0.351	-4.415555	1.630509
er	-.0453783	.0214147	-2.12	0.045	-.089678	-.0010786
cobertura	.1864615	.0354698	5.26	0.000	.1130866	.2598364
_cons	214.5525	22.0451	9.73	0.000	168.9487	260.1563

Fuente: Resultados de estimación.

Regresión con Efectos Fijos

```
. xtreg til pib esco ndes smoderno clu er cobertura, fe cluster(entidad)

Fixed-effects (within) regression      Number of obs   =      240
Group variable: entidad                Number of groups =       24

R-sq:  within = 0.4476                  Obs per group:  min =       10
      between = 0.6647                      avg =      10.0
      overall = 0.6486                      max =       10

                                         F(7,23)        =      23.12
corr(u_i, Xb) = 0.6027                  Prob > F        =      0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
pib	-3.103993	1.552879	-2.00	0.058	-6.316369	.1083825
esco	-29.99468	7.900494	-3.80	0.001	-46.3381	-13.65127
ndes	3.924417	.6633666	5.92	0.000	2.552138	5.296695
smoderno	-.3262577	.1063401	-3.07	0.005	-.546239	-.1062764
clu	.4206298	.2310111	1.82	0.082	-.057253	.8985126
er	-.0100949	.0043437	-2.32	0.029	-.0190806	-.0011092
cobertura	.0272015	.017534	1.55	0.134	-.0090703	.0634733
_cons	125.9798	23.3821	5.39	0.000	77.61027	174.3494
sigma_u	8.0257198					
sigma_e	1.1321386					
rho	.98048925	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Regresión con Efectos Aleatorios

```
. xtreg til pib esco ndes smoderno clu er cobertura, re cluster(entidad)

Random-effects GLS regression      Number of obs   =      240
Group variable: entidad                Number of groups =       24

R-sq:  within = 0.4338                  Obs per group:  min =       10
      between = 0.7465                      avg =      10.0
      overall = 0.7338                      max =       10

                                         Wald chi2(7)    =     209.35
corr(u_i, X) = 0 (assumed)           Prob > chi2     =      0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
pib	-3.229608	.8639038	-3.74	0.000	-4.922828	-1.536387
esco	-42.32958	6.67171	-6.34	0.000	-55.40589	-29.25326
ndes	3.838422	.6504217	5.90	0.000	2.563619	5.113225
smoderno	-.3849118	.0997908	-3.86	0.000	-.5804981	-.1893254
clu	.3882174	.2896565	1.34	0.180	-.179499	.9559338
er	-.0115807	.0046107	-2.51	0.012	-.0206176	-.0025438
cobertura	.0625058	.0144349	4.33	0.000	.0342139	.0907976
_cons	155.3891	17.36889	8.95	0.000	121.3467	189.4315
sigma_u	4.7743854					
sigma_e	1.1321386					
rho	.94676394	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Prueba F. Regresión Agrupada vs Efectos Fijos

Prueba F	
F (7, 23) =	23.12
Prob > F =	0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos.

Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios

Breusch and Pagan Lagrangian multiplier test for random effects

$til[entidad,t] = Xb + u[entidad] + e[entidad,t]$

Estimated results:

	Var	sd = sqrt(Var)
til	115.5133	10.74771
e	1.281738	1.132139
u	22.79476	4.774385

Test: Var(u) = 0

chibar2(01) = 798.00

Prob > chibar2 = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Aleatorios.

Prueba de sobreidentificación. Efectos Fijos vs Efectos Aleatorios

Test of overidentifying restrictions: fixed vs random effects		
Cross-section time-series model: xtreg re robust cluster(entidad)		
Sargan-Hansen statistic = 95.298	Chi-sq(7)	P-value = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos.

B.1.3. Resultados y Pruebas de la Regresión sin IP y PIB

Regresión Agrupada

```
. reg til esco ndes smoderno clu er cobertura, cluster(entidad)
```

```
Linear regression                               Number of obs =      240
                                                F( 6,      23) =    24.62
                                                Prob > F      =    0.0000
                                                R-squared     =    0.8198
                                                Root MSE     =    4.6204
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
esco	-83.2793	10.05986	-8.28	0.000	-104.0897	-62.4689
ndes	2.486812	.9221707	2.70	0.013	.5791565	4.394467
smoderno	-.6271669	.1906677	-3.29	0.003	-1.021593	-.2327406
clu	-1.346518	1.333206	-1.01	0.323	-4.104464	1.411428
er	-.0451874	.0215123	-2.10	0.047	-.089689	-.0006858
cobertura	.1858282	.034711	5.35	0.000	.114023	.2576333
_cons	215.3257	19.02308	11.32	0.000	175.9734	254.6779

Fuente: Resultados de estimación.

Regresión con Efectos Fijos

. xtreg til esco ndes smoderno clu er cobertura, fe cluster(entidad)

```

Fixed-effects (within) regression          Number of obs   =       240
Group variable: entidad                   Number of groups =        24

R-sq:  within = 0.4320                    Obs per group:  min =        10
        between = 0.4896                    avg =           10.0
        overall = 0.4819                    max =           10

                                           F(6,23)         =       29.72
corr(u_i, Xb) = 0.4205                    Prob > F         =       0.0000

```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
esco	-33.11159	8.27337	-4.00	0.001	-50.22636	-15.99682
ndes	3.896956	.6738978	5.78	0.000	2.502892	5.29102
smoderno	-.3585192	.101886	-3.52	0.002	-.5692864	-.147752
clu	.4540131	.2319016	1.96	0.062	-.0257119	.9337382
er	-.0084472	.0043999	-1.92	0.067	-.017549	.0006547
cobertura	.02099	.0164206	1.28	0.214	-.0129786	.0549586
_cons	94.78487	15.4887	6.12	0.000	62.74404	126.8257
sigma_u	8.6075572					
sigma_e	1.1452739					
rho	.98260447	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Regresión con Efectos Aleatorios

. xtreg til esco ndes smoderno clu er cobertura, re cluster(entidad)

```

Random-effects GLS regression          Number of obs   =       240
Group variable: entidad                   Number of groups =        24

R-sq:  within = 0.4089                    Obs per group:  min =        10
        between = 0.7278                    avg =           10.0
        overall = 0.7147                    max =           10

                                           Wald chi2(6)     =      204.93
corr(u_i, X) = 0 (assumed)              Prob > chi2      =       0.0000

```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
esco	-46.86108	7.236809	-6.48	0.000	-61.04496	-32.67719
ndes	3.506699	.6496396	5.40	0.000	2.233429	4.779969
smoderno	-.447041	.0966994	-4.62	0.000	-.6365684	-.2575136
clu	.4311101	.3014887	1.43	0.153	-.1597969	1.022017
er	-.0100757	.0046603	-2.16	0.031	-.0192098	-.0009416
cobertura	.0635592	.0147118	4.32	0.000	.0347246	.0923937
_cons	129.2623	13.47324	9.59	0.000	102.8553	155.6694
sigma_u	4.6347902					
sigma_e	1.1452739					
rho	.94245355	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Prueba F. Regresión Agrupada vs Efectos Fijos

Prueba F	
F (6, 23) =	29.72
Prob > F =	0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos.

Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios

Breusch and Pagan Lagrangian multiplier test for random effects

til[entidad,t] = Xb + u[entidad] + e[entidad,t]

Estimated results:

	Var	sd = sqrt(Var)
til	115.5133	10.74771
e	1.311652	1.145274
u	21.48128	4.63479

Test: Var(u) = 0

chibar2(01) = 801.14

Prob > chibar2 = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Aleatorios.

Prueba de sobreidentificación. Efectos Fijos vs Efectos Aleatorios

Test of overidentifying restrictions: fixed vs random effects		
Cross-section time-series model: xtreg re robust cluster(entidad)		
Sargan-Hansen statistic = 80.956	Chi-sq(6)	P-value = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos.

B.1.4. Resultados y Pruebas de la Regresión sin IP y NDES. El “Mejor” Modelo

Regresión Agrupada

```
. reg til pib esco smoderno clu er cobertura, cluster(entidad)
```

```
Linear regression                               Number of obs =      240
                                                F( 6, 23) =      25.16
                                                Prob > F      = 0.0000
                                                R-squared     = 0.8058
                                                Root MSE     = 4.7966
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
pib	1.973266	1.239206	1.59	0.125	-.5902266	4.536758
esco	-81.32797	9.913494	-8.20	0.000	-101.8356	-60.82035
smoderno	-.5366684	.1848894	-2.90	0.008	-.9191413	-.1541955
clu	-2.314767	1.026839	-2.25	0.034	-4.438945	-.1905895
er	-.0520957	.0197898	-2.63	0.015	-.093034	-.0111574
cobertura	.1939375	.0359166	5.40	0.000	.1196383	.2682367
_cons	208.9859	21.75682	9.61	0.000	163.9785	253.9933

Fuente: Resultados de estimación.

Regresión con Efectos Fijos

```
. xtreg til pib esco smoderno clu er cobertura, fe cluster(entidad)

Fixed-effects (within) regression           Number of obs   =       240
Group variable: entidad                    Number of groups =        24

R-sq:  within = 0.2614                     Obs per group: min =       10
        between = 0.3998                     avg =           10.0
        overall = 0.3968                     max =           10

                                           F(6,23)        =       9.38
corr(u_i, Xb) = 0.1985                     Prob > F       =       0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	-2.844671	1.166858	-2.44	0.023	-5.258502	-.4308407
esco	-16.27091	7.448762	-2.18	0.039	-31.67985	-.8619774
smoderno	-.5110708	.1327729	-3.85	0.001	-.7857324	-.2364092
clu	.8257001	.2842894	2.90	0.008	.2376026	1.413798
er	-.0095933	.0048224	-1.99	0.059	-.0195692	.0003827
cobertura	.0403924	.0212002	1.91	0.069	-.0034635	.0842483
_cons	138.2694	24.99982	5.53	0.000	86.55336	189.9855
sigma_u	8.5878287					
sigma_e	1.3060121					
rho	.97739534	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Regresión con Efectos Aleatorios

```
. xtreg til pib esco smoderno clu er cobertura, re cluster(entidad)

Random-effects GLS regression           Number of obs   =       240
Group variable: entidad                    Number of groups =        24

R-sq:  within = 0.2376                     Obs per group: min =       10
        between = 0.6376                     avg =           10.0
        overall = 0.6280                     max =           10

                                           Wald chi2(6)   =      107.05
corr(u_i, X) = 0 (assumed)                 Prob > chi2    =       0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
pib	-1.279893	1.224256	-1.05	0.296	-3.679389	1.119604
esco	-34.23949	6.405811	-5.35	0.000	-46.79465	-21.68433
smoderno	-.5332433	.1346253	-3.96	0.000	-.7971041	-.2693826
clu	.7589446	.2841196	2.67	0.008	.2020804	1.315809
er	-.0108049	.005063	-2.13	0.033	-.0207283	-.0008816
cobertura	.0829844	.0159024	5.22	0.000	.0518162	.1141527
_cons	157.0011	19.40933	8.09	0.000	118.9595	195.0426
sigma_u	4.8177157					
sigma_e	1.3060121					
rho	.93154342	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Prueba F. Regresión Agrupada vs Efectos Fijos

Prueba F	
F (6, 23) =	9.38
Prob > F =	0.0000

Fuente: Resultados de estimación.
 *Conclusión: Elegir Efectos Fijos.

Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios

Breusch and Pagan Lagrangian multiplier test for random effects

$$til[entidad,t] = Xb + u[entidad] + e[entidad,t]$$

Estimated results:

	Var	sd = sqrt(Var)
til	115.5133	10.74771
e	1.705668	1.306012
u	23.21038	4.817716

Test: Var(u) = 0
 chibar2(01) = 731.04
 Prob > chibar2 = 0.0000

Fuente: Resultados de estimación.
 *Conclusión: Elegir Efectos Aleatorios.

Prueba de sobreidentificación. Efectos Fijos vs Efectos Aleatorios

Test of overidentifying restrictions: fixed vs random effects		
Cross-section time-series model: xtreg re robust cluster(entidad)		
Sargan-Hansen statistic = 56.820	Chi-sq(6)	P-value = 0.0000

Fuente: Resultados de estimación.
 *Conclusión: Elegir Efectos Fijos.

B.1.5. Pruebas de Endogeneidad y Exogeneidad

Prueba de Endogeneidad No. 1

```
. xtreg til pib esco smoderno clu er cobertura residuos, fe cluster(entidad)

Fixed-effects (within) regression      Number of obs   =       240
Group variable: entidad                Number of groups =        24

R-sq:  within = 0.2690                  Obs per group:  min =        10
      between = 0.3092                  avg =           10.0
      overall = 0.3084                  max =           10

                                         F(7,23)        =        9.11
corr(u_i, Xb) = 0.0923                  Prob > F        =       0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

til	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
pib	-3.548179	1.273468	-2.79	0.010	-6.182549	-.9138091
esco	-10.72194	8.351215	-1.28	0.212	-27.99775	6.55386
smoderno	-.4931597	.1393011	-3.54	0.002	-.7813261	-.2049934
clu	.8088807	.2696583	3.00	0.006	.2510501	1.366711
er	-.0105057	.0048187	-2.18	0.040	-.020474	-.0005374
cobertura	.0332154	.0209465	1.59	0.126	-.0101157	.0765465
residuos	-8.075708	8.031881	-1.01	0.325	-24.69092	8.539504
_cons	134.8196	24.89714	5.42	0.000	83.31591	186.3232
sigma_u	9.0648781					
sigma_e	1.3023829					
rho	.9797754	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

B.1.6. Efectos Marginales

```
. margins, dydx(pib esco smoderno clu er cobertura)

Average marginal effects          Number of obs =      240
Model VCE      : Robust

Expression      : Linear prediction, predict()
dy/dx w.r.t.   : pib esco smoderno clu er cobertura
```

	Delta-method				
	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]
pib	-2.844671	1.166858	-2.44	0.015	-5.131672 - .5576708
esco	-16.27091	7.448762	-2.18	0.029	-30.87022 -1.67161
smoderno	-.5110708	.1327729	-3.85	0.000	-.7713008 -.2508407
clu	.8257001	.2842894	2.90	0.004	.2685031 1.382897
er	-.0095933	.0048224	-1.99	0.047	-.019045 -.0001415
cobertura	.0403924	.0212002	1.91	0.057	-.0011592 .081944

Fuente: Resultados de estimación.

B.1.7. Primeras Diferencias

```
. reg D.(til pib esco smoderno clu er cobertura),nocons tsscons cluster(entidad)

Linear regression          Number of obs =      216
                          F( 5, 23) =      20.08
                          Prob > F      =      0.0000
                          R-squared      =      0.2540
                          Root MSE    =      1.1186

                          (Std. Err. adjusted for 24 clusters in entidad)
```

D.til	Robust				
	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
pib D1.	-6.581361	2.15585	-3.05	0.006	-11.04108 -2.121646
esco D1.	-25.8893	5.140019	-5.04	0.000	-36.52224 -15.25636
smoderno D1.	-.4426692	.0990052	-4.47	0.000	-.647477 -.2378613
clu D1.	-.4863242	.2251482	-2.16	0.041	-.9520787 -.0205697
er D1.	-.0055013	.0023001	-2.39	0.025	-.0102594 -.0007431
cobertura D1.	.0702297	.0168413	4.17	0.000	.0353907 .1050686

Fuente: Resultados de estimación.

B.1.8. Pruebas de Raíz Unitaria

Pruebas de Raíz Unitaria. Variable: Residuos

Panel unit root test: Summary
 Series: RESIDUOS
 Sample: 2005 2014
 Exogenous variables: None
 Automatic selection of maximum lags
 Automatic selection of lags based on SIC: 0 to 1
 Newey-West bandwidth selection using Bartlett kernel

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-10.1559	0.0000	24	205
Null: Unit root (assumes individual unit root process)				
ADF - Fisher Chi-square	165.180	0.0000	24	205
PP - Fisher Chi-square	134.956	0.0000	24	216

** Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.

Fuente: Resultados de estimación.

B.2. Estimaciones con TILI

B.2.1. Resultados y Pruebas de la Regresión sin IP y NDES. El “Mejor” Modelo

Regresión Agrupada

```
. reg tili pib esco smoderno clu er cobertura, cluster(entidad)
```

```
Linear regression                               Number of obs =      240
                                                F(   6,   23) =   14.39
                                                Prob > F       =   0.0000
                                                R-squared      =   0.7171
                                                Root MSE      =   2.6534
```

(Std. Err. adjusted for 24 clusters in entidad)

tili	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	1.164319	.8454615	1.38	0.182	-.5846514	2.913289
esco	-40.93643	5.885293	-6.96	0.000	-53.11109	-28.76178
smoderno	-.1920175	.1085639	-1.77	0.090	-.416599	.0325641
clu	.6967794	.6772978	1.03	0.314	-.7043178	2.097877
er	-.0128809	.0125958	-1.02	0.317	-.0389374	.0131755
cobertura	.0757772	.0216028	3.51	0.002	.0310885	.1204659
_cons	99.07259	12.30496	8.05	0.000	73.61784	124.5273

Fuente: Resultados de estimación.

Regresión con Efectos Fijos

```
. xtreg tili pib esco smoderno clu er cobertura, fe cluster(entidad)

Fixed-effects (within) regression              Number of obs   =       240
Group variable: entidad                       Number of groups =        24

R-sq:  within = 0.1176                        Obs per group:  min =        10
        between = 0.1108                       avg =          10.0
        overall = 0.1110                       max =          10

                                                F(6,23)         =        3.40
corr(u_i, Xb) = 0.0087                        Prob > F         =       0.0151
```

(Std. Err. adjusted for 24 clusters in entidad)

tili	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	-1.188248	1.020166	-1.16	0.256	-3.298622	.9221249
esco	-1.232836	4.353919	-0.28	0.780	-10.2396	7.773931
smoderno	-.1871863	.0752211	-2.49	0.021	-.342793	-.0315796
clu	.1398612	.1428071	0.98	0.338	-.1555577	.4352801
er	-.0056791	.0017168	-3.31	0.003	-.0092305	-.0021276
cobertura	-.0114739	.0139932	-0.82	0.421	-.040421	.0174732
_cons	42.19445	15.12575	2.79	0.010	10.90446	73.48444
sigma_u	4.6654856					
sigma_e	.84156785					
rho	.9684878	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Regresión con Efectos Aleatorios

```
. xtreg tili pib esco smoderno clu er cobertura, re cluster(entidad)

Random-effects GLS regression              Number of obs   =       240
Group variable: entidad                       Number of groups =        24

R-sq:  within = 0.0895                        Obs per group:  min =        10
        between = 0.5535                       avg =          10.0
        overall = 0.5343                       max =          10

                                                Wald chi2(6)     =       44.70
corr(u_i, X) = 0 (assumed)                  Prob > chi2      =       0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

tili	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
pib	-.2964962	.8544821	-0.35	0.729	-1.97125	1.378258
esco	-12.74807	3.975477	-3.21	0.001	-20.53986	-4.95628
smoderno	-.198233	.0632778	-3.13	0.002	-.3222552	-.0742108
clu	.179724	.1690897	1.06	0.288	-.1516858	.5111337
er	-.0062285	.0018696	-3.33	0.001	-.0098928	-.0025642
cobertura	.0162997	.0124647	1.31	0.191	-.0081306	.04073
_cons	55.75225	12.17341	4.58	0.000	31.8928	79.6117
sigma_u	2.8505833					
sigma_e	.84156785					
rho	.9198289	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Prueba F. Regresión Agrupada vs Efectos Fijos

Prueba F	
F (6, 23) =	3.40
Prob > F =	0.0151

Fuente: Resultados de estimación.
*Conclusión: Elegir Efectos Fijos.

Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios

Breusch and Pagan Lagrangian multiplier test for random effects

$$tili[entidad,t] = Xb + u[entidad] + e[entidad,t]$$

Estimated results:

	Var	sd = sqrt(Var)
tili	24.26096	4.925541
e	.7082365	.8415679
u	8.125825	2.850583

Test: Var(u) = 0

chibar2(01) = 769.39
Prob > chibar2 = 0.0000

Fuente: Resultados de estimación.
*Conclusión: Elegir Efectos Aleatorios.

Prueba de sobreidentificación. Efectos Fijos vs Efectos Aleatorios

Test of overidentifying restrictions: fixed vs random effects		
Cross-section time-series model: xtreg re robust cluster(entidad)		
Sargan-Hansen statistic = 31.316	Chi-sq(6)	P-value = 0.0000

Fuente: Resultados de estimación.
*Conclusión: Elegir Efectos Fijos.

B.3. Estimaciones con TILD

B.3.1. Resultados y Pruebas de la Regresión sin IP y NDES. El “Mejor” Modelo

Regresión Agrupada

```
. reg tild pib esco smoderno clu er cobertura, cluster(entidad)
```

```
Linear regression                               Number of obs =    240
                                                F(   6,   23) =   13.81
                                                Prob > F       =   0.0000
                                                R-squared     =   0.6863
                                                Root MSE     =   3.9438
```

(Std. Err. adjusted for 24 clusters in entidad)

tild	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	.8093599	.8842839	0.92	0.370	-1.019921	2.63864
esco	-40.39331	10.65564	-3.79	0.001	-62.43617	-18.35044
smoderno	-.3446846	.1264298	-2.73	0.012	-.6062247	-.0831446
clu	-3.01152	.6787955	-4.44	0.000	-4.415715	-1.607325
er	-.0392054	.0180758	-2.17	0.041	-.076598	-.0018128
cobertura	.1181634	.0266455	4.43	0.000	.0630431	.1732838
_cons	109.9122	19.71912	5.57	0.000	69.12006	150.7043

Fuente: Resultados de estimación.

Regresión con Efectos Fijos

```
. xtreg tild pib esco smoderno clu er cobertura, fe cluster(entidad)

Fixed-effects (within) regression              Number of obs   =       240
Group variable: entidad                       Number of groups =        24

R-sq:  within = 0.2449                        Obs per group: min =        10
        between = 0.4209                        avg =           10.0
        overall = 0.4164                        max =           10

                                                F(6,23)         =        6.98
corr(u_i, Xb) = 0.1653                        Prob > F        =       0.0003
```

(Std. Err. adjusted for 24 clusters in entidad)

tild	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
pib	-1.65596	.9481174	-1.75	0.094	-3.617291	.3053698
esco	-15.01282	5.38106	-2.79	0.010	-26.14439	-3.881248
smoderno	-.3235884	.0794547	-4.07	0.000	-.487953	-.1592238
clu	.6855054	.2666572	2.57	0.017	.1338829	1.237128
er	-.0039125	.0043685	-0.90	0.380	-.0129495	.0051246
cobertura	.0517869	.0179134	2.89	0.008	.0147302	.0888437
_cons	96.01052	17.79203	5.40	0.000	59.20491	132.8161
sigma_u	5.4038908					
sigma_e	.99850663					
rho	.96698521	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Regresión con Efectos Aleatorios

```
. xtreg tild pib esco smoderno clu er cobertura, re cluster(entidad)

Random-effects GLS regression              Number of obs   =       240
Group variable: entidad                       Number of groups =        24

R-sq:  within = 0.2360                        Obs per group: min =        10
        between = 0.5236                        avg =           10.0
        overall = 0.5158                        max =           10

                                                Wald chi2(6)    =       52.61
corr(u_i, X) = 0 (assumed)                  Prob > chi2     =       0.0000
```

(Std. Err. adjusted for 24 clusters in entidad)

tild	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
pib	-.9191317	.5890833	-1.56	0.119	-2.073714	.2354504
esco	-22.49832	4.916441	-4.58	0.000	-32.13437	-12.86227
smoderno	-.3352623	.0836491	-4.01	0.000	-.4992116	-.1713131
clu	.5894188	.2695989	2.19	0.029	.0610147	1.117823
er	-.0046122	.0045959	-1.00	0.316	-.01362	.0043956
cobertura	.0691667	.0157754	4.38	0.000	.0382475	.100086
_cons	102.6123	12.82366	8.00	0.000	77.4784	127.7462
sigma_u	3.7824795					
sigma_e	.99850663					
rho	.93485333	(fraction of variance due to u_i)				

Fuente: Resultados de estimación.

Prueba F. Regresión Agrupada vs Efectos Fijos

Prueba F	
F (6, 23) =	6.98
Prob > F =	0.0003

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos.

Prueba del Multiplicador de Lagrange (LM). Regresión Agrupada vs Efectos Aleatorios

Breusch and Pagan Lagrangian multiplier test for random effects

$$\tilde{tild}[entidad,t] = Xb + u[entidad] + e[entidad,t]$$

Estimated results:

	Var	sd = sqrt(Var)
tild	48.3364	6.952438
e	.9970155	.9985066
u	14.30715	3.782479

Test: Var(u) = 0

chibar2(01) = 734.78

Prob > chibar2 = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Aleatorios.

Prueba de sobreidentificación. Efectos Fijos vs Efectos Aleatorios

Test of overidentifying restrictions: fixed vs random effects		
Cross-section time-series model: xtreg re robust cluster(entidad)		
Sargan-Hansen statistic = 41.195	Chi-sq(6)	P-value = 0.0000

Fuente: Resultados de estimación.

*Conclusión: Elegir Efectos Fijos.

C. Correlaciones

Matriz de Correlaciones

```
. correlate pib esco smoderno clu er cobertura
(obs=240)
```

	pib	esco	smoderno	clu	er	cobert~a
pib	1.0000					
esco	0.3498	1.0000				
smoderno	0.3494	0.1480	1.0000			
clu	0.2482	0.3077	0.1537	1.0000		
er	-0.2427	-0.3402	-0.2920	-0.3782	1.0000	
cobertura	-0.2262	-0.0780	-0.3802	-0.1536	0.0410	1.0000

Fuente: Resultados de estimación.