

Impacto de la política pública de combate a la informalidad en la ciudad de México^{1,2}

Víctor González Ruelas
Universidad del Istmo
Ixtepec, Oaxaca, México
vgruelas@hotmail.com

Juan Carlos Méndez Ferrer³
Universidad Autónoma de San Luis Potosí
San Luis Potosí, San Luis Potosí, México
jcmferrer@uaslp.mx

Resumen

El artículo muestra diferentes metodologías para evaluar el impacto de los programas de combate a la informalidad, que componen la Política de Promoción de Ocupación Remunerada, implementada en la Ciudad de México durante el 2010. La medición se realizó con el enfoque de emparejamiento no paramétrico llamado *propensity score matching* junto con el método de diferencias en diferencias y la técnica de control reflexivo utilizando grupos de comparación no equivalentes, los cuales arrojaron que la eficacia en el combate al trabajo informal disminuye a medida que transcurre el tiempo, debido a la falta de coordinación y seguimiento de los instrumentos señalados; situaciones que de corregirse podrían mejorar los efectos de la intervención pública.

Palabras clave: Capacitación, dicotomía, informalidad, política pública y evaluación.
Clasificación JEL: J08, J46, J48.

¹El presente documento forma parte de una investigación más amplia titulada “Evaluación de Impacto de la Política de Promoción de Ocupación Remunerada que Combate la Informalidad en la Ciudad de México”.

²Los autores agradecen la colaboración del Dr. Álvaro Castañeda Arredondo para la elaboración de este documento.

³Facultad de Economía.

Abstract

The paper shows different methods to assess the impact of programs to combat informality that make Promotion Policy Gainful Occupation, implemented in Mexico City during the 2010. The measurement was performed with the nonparametric matching approach called propensity score matching method together with the differences in differences and reflexive control techniques using non-equivalent comparison groups. These methods showed that the effectiveness in combating informal employment decreases as time passes, due to the lack of coordination and monitoring instruments indicated; to correct there situations could enhance the effects of government intervention.

Keywords: Training, dichotomy, informality, public policy and evaluation.

JEL classification: J08, J46, J48.

1. Introducción

Desde los primeros estudios de la informalidad se ha observado que América Latina ha sido una de las regiones con elevados índices de ocupación informal⁴, situación que ha propiciado múltiples efectos negativos en sus economías (ilegalidad, corrupción, violencia, lavado de dinero, baja recaudación fiscal, tráfico de drogas y armas, etc.) motivando continuos debates con respecto a la intervención pública en prácticamente todas las países de la zona (Monnet, 2006). En México, las actividades informales provocan bajo crecimiento económico, degradación del tejido social y violación del estado de derecho debido al inadecuado funcionamiento de las instituciones, al exceso de regulaciones y a la inobservancia de la ley por parte de gobernantes, autoridades, políticos, empresarios y ciudadanos⁵. Toro Vázquez (1992) y Guerrero (2002) señalan que el fenómeno se agudiza durante las crisis económicas, debido al despido de un gran número de trabajadores que al no tener opciones se ocupan en la informalidad (ver Figura 1). En la Ciudad de México la ocupación informal ha provocado diversos problemas, entre los que destacan: corrupción, extorsión, explotación, inseguridad pública, delincuencia, venta de productos de dudosa procedencia, evasión de disposiciones legales para ejercer oficios y/o actividades formales, etc. En el área geográfica que comprende el Distrito Federal, la Tasa de Ocupación en el

⁴Los países que integran la región tienen una economía informal que se encuentra dentro del rango de 24.6 a 55% de su PIB (Chapa, 2007).

⁵www.elsemanario.com.mx, divulgación: 22/01/2009, 05/11/2009, 12/11/2010 y 13/06/2011.

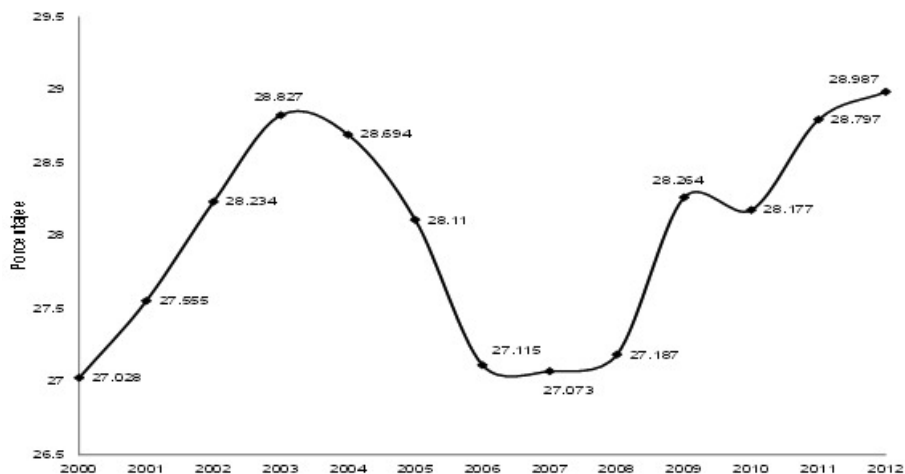


Figura 1: Tasa de Ocupación en el Sector Informal (TOSI) con respecto a la Población Económicamente Activa (PEA) en México, Porcentaje (2000 – 2012). **Fuente:** elaboración propia con base en la Encuesta Nacional de Empleo (ENE) para el periodo 2000 – 2004, y el resto de la Encuesta Nacional de Ocupación y Empleo (ENOE).

Sector informal (TOSI) se ha incrementado a partir del 2007 (ver Figura 2), y aunque es menor a la media nacional la brecha se ha ido cerrando en los últimos años; ello implica la necesidad de implementar medidas para abatirla.

Para enfrentar el incremento del trabajo informal, el gobierno del Distrito Federal incorporó en el Programa General de Desarrollo (2007-2012) una gran cantidad de objetivos que coinciden con las propuestas teóricas para abatir el fenómeno de la informalidad. Algunas de las medidas presentadas en el citado estatuto fueron: Capacitación para el Impulso de la Economía Social (CAPACITES), Calidad y Modernización (CYMO) y Seguro de Desempleo (SDD), las cuales integran la política pública de ocupación remunerada de la Ciudad de México, implementada en el 2010.

Para este fin se sometieron a evaluación los programas de la política de ocupación remunerada. El análisis se desarrolló, a través de la definición de dos grupos: 1) tratamiento (personas que obtuvieron el beneficio) y 2) control (individuos que no recibieron el apoyo de los programas); así mismo, se contemplan tres momentos, 6, 12 y 24 meses, a partir de la implementación.

La presente investigación se divide en seis partes: 1) introducción, 2) teorías de la ocupación informal, 3) características de los programas, 4) metodología de

evaluación, 5) resultados de las estimaciones econométricas, 6) conclusiones y referencias.

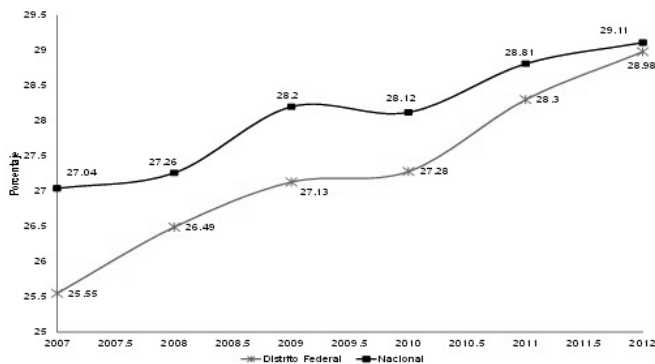


Figura 2: Tasa de Ocupación en el Sector Informal y Población Económicamente Activa en el Distrito Federal y República Mexicana (2000 – 2012). Fuente: elaboración propia con base en la Encuesta Nacional de Empleo (ENE) para el periodo 2000 – 2004, y el resto de la Encuesta Nacional de Ocupación y Empleo (ENOE).

2. Teoría de la ocupación informal

El término de “informalidad” no es posible definirlo en términos estándar; contrariamente, hay diversos puntos de vista que se han establecido en la literatura. Así que se tiene un problema epistemológico debido a las múltiples interpretaciones que genera su conceptualización.

Roubaud (1995) plantea que la informalidad no forma parte de la economía de mercado. La Organización Internacional del Trabajo (OIT) a través de su Programa Regional de Empleo para América Latina y el Caribe (PREALC) la definió como: “conjunto de actividades desarrolladas por trabajadores y/o empresas no formalizadas que utilizan procedimientos tradicionales”. Para Judisman (1993) el sector informal está integrado por: a) trabajadores domésticos; b) empleados, asalariados y trabajadores a destajo; c) trabajadores por cuenta propia y aquellos que no reciben pago. Roubaud (1995) puntualiza que el sector informal urbano es el conjunto de las unidades económicas que no están registradas ante las autoridades, y las personas físicas que no declaran impuestos.

Por su parte, Kathleen Staudt (1998) afirma que la informalidad está integrada por los trabajadores ocasionales y quienes se dedican a las tareas de auto-

construcción de viviendas. Fernando Mires (1998) argumenta que la informalidad está relacionada con los sectores atrasados y empobrecidos de la economía. Rodríguez Oreggia, Martín Lima y Alberto Villalpando (Rodríguez-Oreggia, Lima y Villalpando, 2006) plantean que el sector informal está conformado por los trabajadores que carecen de prestaciones de seguridad social, por lo que no gozan de los derechos y prestaciones que por ley les corresponden.

En el 2011, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) señaló que las actividades informales son aquellas que no se registran ante las autoridades, por lo que evaden impuestos y contribuciones a la seguridad social. En 1993 la OIT señaló que el sector informal está conformado por las empresas que no están constituidas legalmente y, desde entonces, el Instituto Nacional de Estadística y Geografía (INEGI) adopta dicha definición.

En la década de los sesenta, el Centro de Investigación y Acción Social para el Desarrollo Social en América Latina (DESAL, 1965 y 1969) diferenció entre el sector tradicional (marginal y rudimentario) y el moderno (urbano e industrial), originando una postura dual, puesto que los problemas estructurales provocaron que las personas emigraran del campo a la ciudad; pero al no estar capacitados para desarrollar actividades especializadas, se convirtieron en un ejército de reserva de mano de obra.

Posteriormente, la teoría Desalinista enfatizó que en las economías capitalistas dependientes, algunas ocupaciones se van atrasando en relación con el progreso y en consecuencia, tienden a desaparecer, generándose así actividades productivas irregulares que se realizan en unidades pequeñas, de manera improvisada, con predominio de la mano de obra familiar, con bajos recursos financieros y evadiendo toda disposición gubernamental (Quijano, 1971).

La corriente estructuralista destaca que el fenómeno se origina por el desequilibrio entre la oferta y la demanda en el mercado laboral; principalmente, porque la primera resulta ser mayor a la segunda, lo cual se debe a la migración de la gente del campo a la ciudad e integración de la mujer en el mercado laboral⁶.

Para la escuela de la extralegalidad, la informalidad se utiliza como válvula de escape a la excesiva regulación del Estado; lo que ha propiciado un “sistema legal de excepción” que tolera y, en la mayoría de los casos, regula la realiza-

⁶Para los estructuralistas, los componentes que constituyen al sector informal son: a) baja productividad, b) escasa capacidad de acumulación, c) pequeñas unidades económicas, f) relaciones laborales desprotegidas con bajas remuneraciones, g) no se requiere fuerza de trabajo calificada, h) bajos estándares de competencia y calidad, entre otros. (Lewis, 1954; Hart, 1970 y 1973; OIT, 1972; Singer, 1980; Piore, 1975 y 1980; PREALC, 1981 y 1985; Tokman, 1978 y 1982).

ción de actividades informales, afectando a todas aquellas que cumplen con los derechos y obligaciones de ley. (De Soto, 1987)

La escuela institucionalista explica que la existencia del sector informal se debe a que los agentes económicos evitan cumplir con el conjunto de reglas que articulan y organizan las interacciones económicas, sociales y políticas entre los individuos y los grupos sociales.

Sobre la base del contexto presentado se concluye que casi todos los enfoques revisados vinculan al sector informal con aspectos laborales; en este sentido, los trabajadores informales son asociados con individuos de escasa preparación, lo que les dificulta obtener un trabajo en la economía formal, recurriendo al autoempleo en unidades económicas irregulares.

3. La política pública en el combate a la informalidad

En este apartado se determinan las principales características de los programas que componen la política de ocupación remunerada de combate a la informalidad en la Ciudad de México, demostrándose que una política de mercado laboral es un instrumento que la combate.

Al respecto, el Programa General de Desarrollo del Distrito Federal 2007 – 2012, en sus ejes 2 y 4, asume la Recomendación 195 de la OIT para combatir las actividades informales mediante la *política de ocupación remunerada*, señalando como objetivo primordial el impulsar la incorporación de los trabajadores informales hacia las actividades formales mediante su capacitación. Por ello, lo anterior lleva a concluir que las autoridades están conscientes de los efectos negativos de la informalidad y, por eso, las estipulaciones plasmadas en el citado instrumento se vuelven obligaciones en el marco de la planeación oficial de los compromisos y estrategias que sigue la administración local.

En dicho contexto, la política está conformada por: a) Programa de Becas a la Capacitación para el Trabajo (BÉCATE), b) Programa de Capacitación para el Impulso de la Economía Social (CAPACITES), c) Programa de Capacitación y Modernización (CyMO), d) Programa de Atención Integral a Jóvenes Desempleados (COMUNA), e) Programa Seguro de Desempleo. (SDD), f) Programa de Reordenamiento del Trabajo no Asalariado y g) Programa Emergente de Apoyo a Trabajadoras y Trabajadores Migrantes Desempleados en el Distrito Federal.

Debido a que algunos programas exigen intervención federal y puesto que sólo

interesa analizar la forma en que la capacitación combate la informalidad en la Ciudad de México, acorde con la recomendación 195 de la OIT, la investigación toma únicamente para su análisis los programas CAPACITES, CyMO, y SDD, con sus subprogramas y modalidades.

El CAPACITES tiene como objetivo desarrollar y/o fortalecer los conocimientos, habilidades, destrezas y aptitudes laborales de los desempleados, subempleados, egresados de escuelas a nivel medio superior e internos de los Centros de Readaptación del Distrito Federal. Este programa incluye la entrega de apoyos en especie, así como una compensación monetaria de forma temporal. CAPACITES está compuesto por los siguientes subprogramas: 1) Capacitación para el Trabajo (SCAPAT); 2) Fomento al Autoempleo (SFA), y 3) Compensación a la Ocupación Temporal (SCOT). El primero está conformado por tres modalidades: a) Capacitación para el Autoempleo (CA); b) Vales de Capacitación (VC), y c) Consolidación de Iniciativas de Empleo (CIE). La modalidad CA está enfocada en alentar los conocimientos y destrezas técnicas y administrativas de individuos desempleados y subempleados, quienes reciben capacitación a través de cursos intensivos (cuya duración es de uno a tres meses); durante la preparación se otorga un apoyo económico por la cantidad equivalente a un salario mínimo mensual.

La VC consiste en proporcionar un vale – personal e intransferible – a los desempleados mayores de 16 años canjeable por un curso de capacitación ante una institución registrada; además se cubren los costos de inscripción, colegiatura, materiales, pago de la certificación laboral y ayuda para transporte; los técnicos y operarios reciben un apoyo de 1.5 veces el salario mínimo vigente, y dos veces para profesionistas.

Finalmente, el CIE tiene el propósito de impulsar los conocimientos y habilidades de los desempleados o subempleados por medio de la capacitación y ayuda en especie que permitan la consolidación y desarrollo de un proyecto empresarial, siendo su segmento de interés las jefas de familia, adolescentes embarazadas, personas con discapacidad, población callejera e internos quienes estén próximos a ser liberados, preliberados o externados de los Centros de Readaptación Social del Distrito Federal.

Los apoyos que ofrece el SFA están encaminados a promover actividades productivas por cuenta propia, iniciativas empresariales registradas o de nueva creación. Los solicitantes (desempleados o subempleados) alcanzan financiamientos parciales para la adquisición de maquinaria, mobiliarios, equipos, herramientas y accesorios otorgados en condición de comodato, y demostrando el uso continuo y adecuado, se obtiene su posesión, siempre y cuando cumplan con los siguientes requisitos: comprobar los conocimientos y experiencia de cuando menos seis

meses en el giro en el que se desea incursionar, contar con una edad mínima de 18 años, ganar menos de seis salarios mínimos, no estar en posibilidades para acceder a un crédito financiero, demostrar una residencia de cuando menos dos años en el Distrito Federal y no tener deudas.

El SCOT está dirigido a la población desempleada mayor de 18 años que labora en trabajos temporales en instituciones públicas o privadas por un periodo de hasta tres meses; los apoyos otorgados no generan relación laboral o contractual con las instituciones propietarias del proyecto. Los beneficiarios reciben materiales, insumos y herramientas, además de una compensación económica o en especie; esto con el fin de fortalecer sus habilidades y destrezas de manera teórico-práctica.

El CyMO está orientado a mejorar la productividad de las empresas (micro, pequeñas y medianas empresas) y sus trabajadores con cursos de capacitación en las áreas de administración, mercadotecnia, puntos de venta y atención a clientes. La intención es asegurar que las unidades económicas se mantengan en el mercado y conserven su plantilla laboral, lo cual evitará que las empresas cierren y su personal ingrese a la informalidad como válvula de escape al desempleo.

Por último, el SDD otorga una prestación equivalente a treinta días de salario mínimo general vigente en el Distrito Federal por un periodo de seis meses a todos los que hayan perdido su empleo en el último año y que, previamente, hayan laborado al menos durante seis meses para una persona física o moral con domicilio fiscal en la Ciudad de México; periodo en el que deberán presentar una bitácora que demuestre que están buscando empleo en el sector formal.

4. Metodología

La estimación de los alcances de la Política de Ocupación Remunerada del 2010 se desarrolló con la técnica de evaluación de impacto, en virtud de que ésta tiene como propósito determinar si un programa produjo los efectos deseados en sus beneficiarios y si son atribuibles a la intervención pública. Dicho procedimiento determina los efectos de los programas públicos, dado un “escenario en el que no los hubo”. Esto se logra comparando la situación de los beneficiarios posterior a la implementación de la política con un entorno en el que no recibieron apoyos. Empero, es imposible observar a los mismos individuos en contextos opuestos, por lo que es necesario construir escenarios que simulen la situación de los beneficiarios bajo el contexto de no participación en los programas.

Lo anterior se consiguió estimando el indicador o variable dicotómica “tipo

de ocupación: formal e informal” de los beneficiarios de los programas de capacitación⁷. A partir de esta construcción fue factible evaluar las relaciones de causa efecto entre los objetivos del programa y sus resultados⁸.

La técnica que se utilizó para generar los datos fue la *cuasi experimental*, debido a que se respetaron las limitaciones teóricas de los modelos experimentales; no obstante, la creación de grupos de control análogos a los de tratamiento exigió el uso combinado del método cuasi experimental con el proceso de emparejamiento o *matching*⁹. Con el fin de evitar los efectos negativos en la muestra y en los resultados de la evaluación, atribuibles al sesgo de selección por el uso del diseño cuasi experimental, se utilizó el método semi-paramétrico.

La estimación de la evaluación se realizó con el modelo probit, en virtud, de que se utilizan variables explicativas cualitativas o dicotómicas¹⁰ debido a que adquieren valores de 1 ó 0 (el cero indica ausencia del atributo y el uno presencia o posesión del atributo); por tal razón, los resultados de los modelos arrojaron la probabilidad de ocurrencia del acontecimiento objeto de estudio ($Y_i=1$).

4.1. Selección de la muestra

La selección de la muestra se eligió a partir del registro proporcionado por la Secretaría del Trabajo y Fomento al Empleo (STyFE) a través del Servicio Nacional de Empleo del Distrito Federal, por lo que todos los individuos en esta base de datos tuvieron la misma probabilidad de ser seleccionados.

El tamaño de la muestra se definió a través de la aplicación de la fórmula de Cochran. Tomándose un nivel de confianza del 95 por ciento, probabilidades de ocurrencia y no ocurrencia del 50 por ciento, respectivamente, y un nivel de precisión del 5 por ciento; el tamaño de la muestra para el grupo de tratamiento y control por programa¹¹ se obtienen los resultado del Cuadro 1.

⁷Bajo este diseño, el indicador impacto se estima en su versión de diferencias en diferencias, tomando en consideración la diferencia en la variable de impacto entre beneficiarios y controles en la situación post programa (t1) y en la situación pre-programa (t0).

⁸La evaluación se orientó a estimar la probabilidad de que los beneficiarios encuentren empleo a partir de la intervención pública en el corto y mediano plazo; es decir, el interés versa sobre los beneficios que no desaparecen si se deja de participar en alguno de los programas de capacitación.

⁹El proceso *matching* compara la situación de cada uno de los participantes con la de aquellos individuos que sean similares en un conjunto de atributos, o al menos, ponderan más la comparación con individuos que tengan atributos similares que la de individuos cuyos atributos sean muy diferentes.

¹⁰Las variables dicotomas son también conocidas como dummy, artificiales, indicadoras, binarias, categóricas o cualitativas.

¹¹Del registro de la STyFE se eligió aleatoriamente al grupo de tratamiento y el de control

PPOR	GT	GC	Enc.
Capacitación al autoempleo	530	779	1,309
Vales de capacitación	80	196	276
Consolidación de iniciativas de empleo	233	345	578
Subprograma de Fomento al autoempleo (SFA)	327	470	797
Subprograma Compensación a la Ocupación Temporal (SCOT)	170	170	340
Programa de Capacitación y Modernización (CyMO)	331	466	797
Programa Seguro de Desempleo (SDD)	480	731	1,211
Siete instrumentos	2,151	3,242	5,393

Cuadro 1: Muestras de los grupos de beneficiarios y control por programa. PPOR significa “Política pública de ocupación remunerada”, GT denota el “grupo de tratamiento”, GC el “grupo de control” y Enc. el número de “encuestados”. **Fuente:** elaboración propia con base en registros de la STyFE.

Ahora bien, puesto que los datos utilizados son de panel longitudinal, la información recolectada para la técnica de control reflexivo se levantó en cinco momentos: el primero para beneficiarios sin proyecto, el segundo para beneficiarios con proyecto (registro), el tercero se aplicó a los seis meses, el cuarto al año y el último a los dos años posteriores a la implementación de los programas. Las bases de datos creadas para desarrollar la evaluación en comento, son: línea base, registro, evaluación a los 6 meses, 1 año y dos años de haberse implementado la política de ocupación remunerada.

Por último, la evaluación se realizó mediante los métodos de *matching* y diferencias en diferencias del diseño de comparación no equivalentes, que exigen para el desarrollo de la estimación tanto un grupo de tratamiento (que se obtuvo con el método de control reflexivo) como uno de control (se eligió con características similares a los de tratamiento por lo que denotan sesgo de selección).

5. Resultados

Esta sección se conforma de dos apartados: el primero presenta los resultados bajo el modelo de control reflexivo que se caracteriza por no discriminar entre los grupos de tratamiento y control; es decir, todos los individuos de la muestra son iguales en el sentido de haber participado en los programas de capacitación;

aglomeró a los solicitantes restantes, buscándose una semejanza con los primeros.

el segundo bloque muestra las estimaciones cuando se hace la distinción entre grupos, utilizando el método de diferencias en diferencias y el ajuste correspondiente con la técnica *propensity score matching*.

5.1. Resultados del modelo control reflexivo

La información que se empleó para la estimación del modelo control reflexivo parte de la muestra de los individuos quienes constituyen el grupo de tratamiento; dejando a un lado – por el momento – al grupo de control, no importando la distinción entre los dos. El modelo que se estimó mediante el procedimiento Logit fue:

$$Y_i = \alpha_0 + \alpha_p P_i + \mu \quad (1)$$

donde

- $Y_i = 1$ indica el promedio de los beneficiarios que realizan actividades informales (PRAI) y es igual a cero para indicar el promedio de los beneficiarios que realizan actividades formales.
- $P_i = 1$ indica que fue posterior a la implementación del programa y cero en caso contrario.
- α_0 es el intercepto de la regresión.
- α_p señala en cuánto se modifica Y debido a la intervención pública. Se espera que $\alpha_p < 0$.
- μ es el término de error no observado.

El propósito que se busca con la estimación es comparar si hay cambios significativos entre los individuos (beneficiarios) que se inscribieron en los programas de capacitación antes y después de la puesta en marcha de los mismos. En consecuencia, el parámetro del segundo término del lado derecho de la Ecuación 1, permitirá distinguir si hay diferencias entre los dos momentos que se citan. Los resultados se observan en el Cuadro 2.

Del Cuadro 2 se desprende que a medio año de implementados los cursos y asesorías de formación, todos los programas reducen la informalidad. Incluso, se elimina en 99.69 por ciento entre los individuos que se inscribieron y fueron beneficiarios del CyMO. En el caso de VC y el SCOT comparten resultados similares, puesto que, disminuyen el número de informales en un 80 por ciento. Contrariamente, las modalidades CIE y SFA presentan los valores más pequeños, ya que la disminución se ubica en un 22 por ciento, aproximadamente.

PPOR	1/2 año		1 año		2 años	
	α_p	ZP	α_p	ZP	α_p	ZP
Capacitación al autoempleo	-0.7396**	3.60	-0.6716**	3.38	-0.1962	0.99
Vales de capacitación	-0.80**	4.00	-0.6625**	3.32	-0.1875	0.87
Consolidación de iniciativas de empleo	-0.2188**	1.06	-0.1502**	0.75	-0.0901	0.49
Fomento al autoempleo	-0.2232**	1.12	-0.1865**	0.86	-0.1070**	0.50
Compensación a la ocupación temporal	-0.7941**	3.99	-0.6764**	3.38	-0.3176**	2.53
Programa de Capacitación y Modernización	-0.9969**	4.99	-0.9425**	4.69	-0.8972**	4.46
Programa Seguro de Desempleo	-0.7083**	3.50	-0.5229**	2.61	-0.1979	0.99

Cuadro 2: Estimación del PRAI con el modelo control reflexivo. ** Significativos al 10 por ciento. PPOR significa Política pública de ocupación remunerada, y ZP significa “Z en Probit”. Fuente: elaboración propia.

Al analizarse las estimaciones a través del tiempo, se detecta la tendencia a la baja en el parámetro de impacto α_p . En consecuencia, pierden eficacia los programas en cuanto al combate de las actividades informales. Y todavía a los dos años, el CyMO inhibe el número de informales en 89 por ciento. No obstante, los resultados de la primera tabla deben de tomarse con cautela porque se obtienen de la información proveniente del grupo de tratamiento antes y después de la implementación de los programas públicos, por lo que, se generan problemas de sesgo.

Para mejorar las estimaciones, se agregan variables exógenas que podrían estar relacionadas teóricamente con Y y P en la Ecuación 1 (género, edad, estado civil, escolaridad, experiencia, dependientes económicos, nivel de ingresos, prestaciones sociales, flexibilidad en el trabajo, estabilidad, capacitación y jefe de hogar); se corrió la regresión considerando el paquete de variables citadas; se revisó la significancia estadística de cada uno de los coeficientes, quedando como argumentos las primeras seis para volver a computar el modelo que se indica en la Ecuación 2, y cuyos resultados se muestran en el Cuadro 3.

$$Y_i = \beta_0 + \beta_p P_i + \beta_z Z_i + \varepsilon \quad (2)$$

Tanto Y_i como P_i se definieron en la Ecuación 1, y donde además

- Z_i : Conjunto de variables exógenas. Se consideran *Edad*: número de años,

Escolaridad: número de años (variable exógena y ordinal), *Experiencia*: número de años (variable exógena o de ajuste independiente atribuible al trabajo posterior a la capacitación), *Dependientes económicos*: número de integrantes que conforman el hogar del entrevistado, *Estado civil*: 1 si es casado ó 0 en otro caso y *Género*: 1 si es mujer ó 0 si es hombre.

- β_0 : intercepto de la regresión.
- β_p : señala cuánto se modifica Y debido a la intervención pública, esperando que $\beta_p < 0$.
- ε : término de error no observado.

PPOR	1/2 año		1 año		2 años	
	α_p	ZP	α_p	ZP	α_p	ZP
Capacitación al auto- empleo	-0.6174**	3.06	-0.5494**	2.71	-0.0740	0.41
Vales de capacitación	-0.6598**	3.25	-0.5223**	2.60	-0.0473	0.23
Consolidación de ini- ciativas de empleo	-0.1463	0.71	-0.0777	0.41	-0.0176	0.08
Fomento al autoem- pleo	-0.1671	0.83	-0.1304	0.66	-0.0509	0.25
Compensación a la ocupación temporal	-0.6903**	3.49	-0.5726**	2.91	-0.2138	1.08
Programa de Capaci- tación y Moderniza- ción	-0.9528**	4.75	-0.8984**	4.48	-0.8531**	4.25
Programa Seguro de Desempleo	-0.5892**	2.89	-0.4038**	2.00	-0.0788	0.41

Cuadro 3: Estimación del PRAI con el modelo control reflexivo aislando el efecto de los factores exógenos (definidos en Z_i). ** Significativos al 10 por ciento. PPOR significa Política pública de ocupación remunerada, y ZP significa “Z en Probit”. **Fuente:** elaboración propia.

Los resultados del Cuadro 3 muestran que hay una sobreestimación del parámetro de impacto sobre el combate a la informalidad β_p en relación con las estimaciones del Cuadro 2. Se siguen obteniendo el signo correcto como se esperaba teóricamente, es decir, una relación inversa entre participar o no en los programas de capacitación y la variable dependiente sobre la condición de informalidad.

Sigue destacando el CyMO en el combate a la informalidad. A medio año de su implementación, el número de informales se redujo en un 95 por ciento, y a los dos años, su eficacia baja a un 85 por ciento; además, se identifica una significancia estadística del 10 por ciento. Contrariamente, la modalidad CIE y el SFA son programas que no tienen efecto alguno sobre la ocupación informal en los tres momentos analizados; situación similar ocurre para el resto de los programas, menos para el CyMO, a los dos años de evaluación.¹²

5.2. Resultados del método de diferencias en diferencias del grupo de comparación no equivalente

En primer lugar, se verificó si hay diferencias entre el grupo de tratamiento y de control mediante pruebas de diferencias de proporciones (estadístico Z) para la variable dependiente (PRAI) y las seis que se utilizaron como factores exógenos (género, edad, estado civil, escolaridad, experiencia y dependientes económicos) entre cada uno de los programas de capacitación, verificándose la presencia de divergencias, excepto para el CyMO, pero que sin duda alguna, se tienen argumentos para demostrar que los resultados que se obtuvieron con el diseño de control reflexivo tienen el problema de sesgo causado por las diferencias encontradas. Contando – ahora, en segundo lugar – con sustentos para la aplicación de la técnica *propensity score matching* a través de la técnica de *nearest neighbor matching* para el emparejamiento de grupos, y posteriormente – tercer lugar –, acudir a la evaluación de diferencias en diferencias con el siguiente modelo econométrico¹³:

$$Y_i = \alpha_0 + \alpha_p P_i + \alpha_T T_i + \alpha_{pT} T_i P_i + \mu \quad (3)$$

donde

- Tanto Y_i como P_i se definieron en la Ecuación 1.

¹²Se reitera que deben tomarse con precaución los resultados presentados en esta sección (Cuadros 2 y 3); puesto que no se está utilizando el método de diferencias en diferencias bajo una comparación no equivalente que se considera el medio más riguroso entre los esquemas cuasi-experimentales. Su prestigio lo ha ganado por las siguientes razones: a) trata de solucionar los problemas de validez interna del control reflexivo, ajustando por la técnica *propensity score matching*; b) toma en cuenta la información que se genera en los grupos de tratamiento y control, y c) se agrega una línea base a la evaluación (situación sin y con proyecto para los dos conjuntos).

¹³En promedio, se perdieron el 10 por ciento de las observaciones para las cuales no existen contrapartes suficientemente comparables.

- $T_i = 1$ si el individuo pertenece al grupo de tratamiento ó 0 si pertenece al grupo de control.
- α_0 es el intercepto de la regresión.
- α_p señala en cuánto se modifica Y debido a la situación antes y después de la capacitación.
- α_T indica en cuánto se modifica Y si el individuo pertenece al grupo de tratamiento o no.
- α_{pt} es el coeficiente de impacto de los programas (ver Cuadro 4), esperando que $\alpha_{pt} < 0$.
- μ es el término de error no observado.

Situación	GT	GC	Diferencia
Sin proyecto	$\alpha_0 + \alpha_T$	α_0	α_T
Con proyecto	$\alpha_0 + \alpha_P + \alpha_T + \alpha_{PT}$	$\alpha_0 + \alpha_P$	$\alpha_T + \alpha_{PT}$
Diferencia	$\alpha_P + \alpha_{PT}$	α_P	α_{PT}

Cuadro 4: Matriz de impacto en el diseño de diferencias en diferencias. “GT” significa *grupo de tratamiento* y “GC”, *grupo de control*. **Fuente:** Shoeni (2004).

En el Cuadro 5 se muestran los resultados de la estimación econométrica (Ecuación 3) mediante el modelo Logit, considerando la corrección por emparejamiento (*propensity score matching* a través de *nearest neighbor matching*); posteriormente, se aplican las diferencias en diferencias y se agregan al cálculo de la regresión las variables exógenas (género, edad, estado civil, experiencia y dependientes económicos).

A medio año de haber implementado los programas de capacitación, no hay la suficiente evidencia para afirmar que el SFA tiene relación con la variable dependiente (informalidad) a pesar del signo correcto en el parámetro. El resto de los programas tienen un impacto sobre el fenómeno en estudio, destacando el CyMO que reduce el número de informales en un 87 por ciento, aproximadamente; siguiéndole el programa SDD con el 60 por ciento. En contraste, la modalidad CIE tiene una eficacia del 12 por ciento.

Al revisar el alcance de los programas, se va identificando una pérdida en el impacto de la política pública; en términos de la significancia estadística, únicamente, el CyMO resulta un verdadero estímulo para la reducción de los trabajadores informales; en consecuencia, su evaluación arroja un saldo favorable. En

PPOR	1/2 año		1 año		2 años	
	α_{PT}	ZP	α_{PT}	ZP	α_{PT}	ZP
Capacitación al autoempleo	-0.4867**	2.48	-0.3977**	1.99	0.0900	0.49
Vales de capacitación	-0.5653**	2.83	-0.4067**	2.00	0.0753	0.41
Consolidación de iniciativas de empleo	-0.1254**	0.61	-0.0396	0.17	0.0380	0.16
Fomento al autoempleo	-0.1314	0.66	-0.0692	0.32	0.0151	0.08
Compensación a la ocupación temporal	-0.5389**	2.66	-0.3893**	1.98	-0.0243	0.13
Programa de Capacitación y Modernización	-0.8716**	4.65	-0.8000**	4.00	-0.7445**	3.67
Programa Seguro de Desempleo	-0.6042**	3.00	-0.3972**	1.99	-0.0633	0.32

Cuadro 5: Estimación econométrica del impacto de la política con diferencias. * Factores exógenos: género; edad; estado civil; escolaridad; experiencia, y dependientes económicos. ** Significativos al 10 por ciento. PPOR significa “Política pública de ocupación remunerada”, y ZP significa “Z en Probit”. **Fuente:** elaboración propia.

los demás se visualiza que no hay evidencia para afirmar que combaten las actividades no formales; inclusive, se obtienen parámetros no negativos (señalando una relación directa).

6. Conclusiones

Se concluye que para un alcance de corto plazo (un semestre después de la implementación de la capacitación), se aceptan las hipótesis alternativas para capacitación al autoempleo, vales de capacitación, consolidación de iniciativas de empleo, compensación a la ocupación temporal, programa de Capacitación y Modernización y programa Seguro de Desempleo. En suma, se verificaron que seis de las siete opciones sí tienen incidencia en la disminución del número de informales.

A un año de la implementación, solamente para cinco de las siete alternativas se cuentan con las evidencias de que sí combaten la informalidad. Por lo tanto, las hipótesis alternativas que no se rechazan son para capacitación al autoempleo, vales de capacitación, compensación a la ocupación temporal, programa de Capacitación y Modernización y programa Seguro de Desempleo.

Finalmente, en el tercer momento (dos años de implementación), únicamente el programa de Capacitación y Modernización tiene eficacia en bajar el número

de trabajadores que pudiesen incorporarse a la informalidad. Para el resto de los programas no se rechazan las hipótesis nulas para los dos años.

Resulta sobresaliente destacar que a través del tiempo se pierde la eficacia en las políticas públicas; por lo que su impacto tiende a la baja y, en general, se emite un juicio de evaluación no favorable. Esto representa áreas de oportunidad por cubrir, como por ejemplo establecer un control y seguimiento más allá de los doce meses, diseñar las propuestas de capacitación con base en las necesidades del mercado laboral, aprovechar las economías de alcance de cada uno en beneficio de los otros, es decir, fomentar una vinculación de recursos, entre otras.

Recepción: 24/09/2014. Aceptación: 19/11/2014.

Referencias

- [1] Chapa, Cantú, Daniel Flores y Jorge N. Valero (2007). “La Economía Informal”. México, Trillas.
- [2] De Soto, Hernando (1987). “El Otro Sendero. La Revolución Informal”. México, Diana.
- [3] DESAL (1965). “América Latina y Desarrollo Social”. Barcelona, Herder.
- [4] El Semanario, URL: www.elsemanario.com.mx, divulgación: 22/01/2009, 05/11/2009, 12/11/2010 y 13/06/2011
- [5] ENOE (Encuesta Nacional de Ocupación y Empleo) del Inegi, <http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/regulares/enoe/> se consultó en diversas ocasiones en los años 2011 y 2012.
- [6] Gobierno del Distrito Federal (2010). “Reglas de Operación del Programa de Capacitación y Modernización”. Publicadas en la Gaceta Oficial del Distrito Federal del 28 de enero de 2010.
- [7] Guerrero, J. Pablo (1995). “La Evaluación de las Políticas Públicas, Enfoques Teóricos y Realidades de Nueve Países Desarrollados”. Revista de Gestión y Política Pública, núm. 1, vol. IV, primer semestre, México, CIDE A.C.

- [8] Herrera, L. y Heijs J. (2003). “Difusión y Adicionalidad de las Ayudas Públicas a la Innovación: Una Estimación Basada en Propensity Score Matching”. España, Instituto de Análisis Industrial Financiero.
- [9] INEGI (Instituto Nacional de Estadística y Geografía). www.inegi.org.mx. Se consultó en diversas ocasiones en los años 2011, 2012 y 2013.
- [10] INEGI (Instituto Nacional de Geografía e Informática) (s.f.). www.inegi.gob.mx, consulta realizada el 05 de enero del 2010.
- [11] Judisman, Clara (1989). “Evolución del empleo y los mercados de trabajo en México”. Memorias de la III Reunión Nacional de Investigación Demográfica en México, México, SOMEDE.
- [12] Monnet, Jérôme (2005). “El prisma del ambulante. Conceptualización del ambulante, de los vendedores a los clientes: Un acercamiento a la metrópoli posfordista”. Memorias del Seminario “El ambulante en la ciudad de México: investigaciones recientes” PUEC-UNAM.
- [13] OIT (Organización Internacional del Trabajo) (1991). “Administración de la Seguridad Social”. Ginebra.
- [14] ---- (2000). “Informe sobre el Trabajo en el Mundo”. Ginebra.
- [15] ---- (s.f.). “Economía Informal en las Américas: Situación Actual, Prioridades de Políticas y Buenas Prácticas”.
- [16] ---- (2009). www.white.oit.org.pe/cimt/nn/documentos/economia_informal_.pdf, consulta realizada el 15 de noviembre del 2009.
- [17] ---- (s.f.). “Recomendación 195. Desarrollo de los Recursos Humanos: Educación, Formación y Aprendizaje Permanente”.
- [18] ---- (s.f.). www.oitcinterfor.org/sites/default/files/edit/docref/rec195.pdf, consulta realizada el 11 de febrero del 2010.
- [19] PREALC (Programa Regional de Empleo para América Latina y el Caribe) (1990). “Más Allá de la Regulación. El Sector Informal en América Latina”. Santiago de Chile, PREALC.
- [20] Quijano, Anibal (1971). “Polo Marginal y Mano de Obra Marginalizada”. Santiago de Chile, CEPAL.

- [21] Ramírez Rivera Faviola (s.f.). “Metodología para la Evaluación de Impacto. Propensity Score Matching”.
- [22] Rodriguez-Oreggia, Eduardo, M. Lima y A. Villalpando (2006). “Informalidad en México: Decisiones y premios salariales a lo largo de la curva de ingreso”, (agosto de 2006: <http://www.inegi.gob.mx>)
- [23] Roubaud, Francois (1995). La Economía Informal en México. México, D.F. Fondo de Cultura Económica.
- [24] Staudt, Kathleen (1998). Free trade? Informal Economics at the US-México Border, Temple University Press, Filadelfia.
- [25] Sianesi Barbara (2001). Implementing Propensity Score Matching Estimators with STATA. London, University College London.
- [26] Thoemmes F. (2012). Propensity Score Matching in SPSS. Germany, University of Tübingen.
- [27] Toro Vázquez, Juan José (1992). La Balanza de Pagos de México: su Importancia en el Análisis Económico, segunda edición, Ed. Universidad Autónoma de San Luis Potosí.

A. Anexo metodológico

La metodología para evaluar el impacto de la política de promoción de ocupación remunerada en el combate contra las actividades informales está sustentada en el argumento de qué hubiese sucedido con los resultados de los beneficiarios al no tomar la capacitación respectiva. De esta manera, se determina por diferencia si hay un contraste entre el valor de la variable dependiente – y de interés – cuando el individuo posee la cualidad de estudio versus su carencia – sino la tuviere –.

Así que, la descripción de la metodología a seguir se inicia con la definición de D_i como *dummy* para el i -ésimo individuo, tomando los valores de 1 si el i -ésimo individuo participa en la capacitación ó 0 en caso contrario. Además, se plantea que la variable dependiente Y_i significa que el i -ésimo individuo trabaja en una ocupación informal y recibió la capacitación, y sus resultados potenciales serán:

- Y_{0i} : valor de la variable respuesta cuando el i -ésimo individuo no se beneficia del programa de formación.

- Y_{1i} : valor de la variable respuesta cuando el i -ésimo individuo se beneficia del programa de formación.

Si el programa en evaluación arroja resultados positivos, se va a detectar con el cálculo de la diferencia $\Delta_i = Y_{1i} - Y_{0i}$ que se identifica como el “impacto” o “efecto de tratamiento”. Sin embargo, Imbens y Angrist (1991) indican que cualquier sujeto no puede encontrarse en las dos situaciones referidas (Y_{0i} e Y_{1i}) de manera simultánea; lo que sí puede suceder es que

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}. \quad (4)$$

El problema descrito en el párrafo anterior se supera con el seguimiento de muestras para dos grupos (a esta actividad se le conoce como métodos observacionales experimentales) en diferentes periodos de tiempo. Para el caso que es motivo de investigación, consistiría en manejar datos antes y después de la implementación de la política de promoción. El primer grupo está formado por individuos quienes estuvieron expuestos a los efectos de la política (grupo de tratamiento), y el segundo constituido por sujetos que fueron ajenos a los beneficios de los programas (grupo de control). Además, se necesita que los dos conjuntos tengan características semejantes que permitan la comparación.

Sin embargo, si el método experimental se realiza, se tiene la desventaja de que las características de los individuos para los dos grupos tenderán a cambiar a través del tiempo en esa tarea de seguimiento. Por lo tanto, Heckman, Ichimura, Smith y Todd (1998) señalan que bajo ciertas condiciones, el método de diferencias en diferencias será la opción que logra cubrir la dificultad en comento.

La variable dependiente cambia si se considera la distinción entre grupo de tratamiento y de control (Abadie, 2005; Cansino y Sánchez, 2008), quedando así:

- $Y_{1i}(t)$: Valores de la variable respuesta para el individuo i en un momento posterior a la finalización del programa t cuando i es beneficiario del mismo.
- $Y_{0i}(t)$: Valores de la variable respuesta para el individuo i en un momento posterior a la finalización del programa t cuando i no ha participado en el mismo (i es un individuo asignado al grupo de control).
- $Y_{1i}(t-1)$: Valores de la variable respuesta para el individuo i en el momento previo al inicio del programa $t - 1$ cuando i es beneficiario del mismo.

- $Y_{0i}(t-1)$: Valores de la variable respuesta para el individuo i en el momento previo al inicio del programa $t-1$ cuando i no ha participado en el mismo (i es un individuo asignado al grupo de control).

La “nueva” medición del efecto causal debido a la aplicación de los programas de capacitación sobre el individuo i -ésimo será

$$Y_{1i}(t) - Y_{0i}(t). \quad (5)$$

Empero, un sujeto quien acudió y concluyó sus cursos de capacitación, no puede señalarse que simultáneamente, que él no asistió a los mismos. Es decir, no resulta lógico que los dos resultados de la Ecuación 4 se observaran en la realidad; lo que se hace es buscar al “gemelo” o contrafactual dentro del grupo de control (sujetos sin el beneficio del programa después que éste ha finalizado) para cada respectivo integrante – factual – del grupo de tratamiento (individuos con el beneficio del programa y que lo terminaron). De esta manera, el método de diferencias en diferencias se aplica bajo la condición de qué hubiera sucedido con el valor de la variable dependiente si los individuos no aprovecharon la protección de la implementación de los programas.

De aquí en adelante, se simplificarán las notaciones de los periodos con la eliminación del subíndice i , por lo que $t-1$ se indicará como el momento 0 (antes de la implementación), y t como el momento 1 (después de la implementación). Así que, las expresiones anteriores quedarán como $Y_1(1)$, $Y_0(1)$, $Y_1(0)$ y $Y_0(0)$.

En el momento 0, no se conoce quién se va a beneficiar o no por la puesta en marcha de los programas motivo de evaluación; por lo tanto, se tiene que $Y_1(0) = Y_0(0) = Y(0)$. Se hace notar que los posteriores resultados son independientes de que la decisión final del individuo sobre si desea o no participar en la capacitación en el futuro.

Card y Krueger (1994), Lmbens et al. (1998), y Angrist y Krueger (2000) proponen que si desea calcular el efecto promedio para cada uno de los programas motivo de evaluación después de su implementación (momento 1), se tiene que utilizar la siguiente expresión

$$E[Y_1(1) - Y_0(1) | D = 1]. \quad (6)$$

Imbens y Athey (2002) indican que si hay un seguimiento a través del tiempo del grupo de control sobre su comportamiento, servirá para conocer la evolución de los beneficiarios (grupo de tratamiento) en la circunstancia de que no lo hubieran seguido, de forma analítica:

$$E[Y_0(1) - Y_0(0) | D = 1] = E[Y_0(1) - Y_0(0) | D = 0]. \quad (7)$$

La Ecuación 7 indica que los efectos promedios del grupo de tratamiento en el tiempo, al compararse dos momentos, se puede calcular a través del comportamiento de los individuos que forman el bloque de control. Si lo anterior se cumple, entonces de la Ecuación 7 se deriva:

$$\alpha_{EMPS} = E[Y_1(1) - Y_0(1) | D = 1] = (E[Y(1) | D = 1] - E[Y(1) | D = 0]) - (E[Y(0) | D = 1] - E[Y(0) | D = 0]). \quad (8)$$

En el primer término se calcula la diferencia entre los resultados promedios observados después de la finalización de los programas tanto para el grupo de tratamiento como de control. Lo mismo se conjetura con el segundo término de la Ecuación 8 pero para el momento 0.

Con datos muestrales, la Ecuación 8 se convierte en:

$$\alpha_{EMPSDID} = \left(\frac{1}{n_1} \sum_{\substack{i=1 \\ D=1}}^{n_1} Y_i(1) - \frac{1}{n_0} \sum_{\substack{i=1 \\ D=0}}^{n_0} Y_i(1) \right) = \underbrace{\left(\frac{1}{n_1} \sum_{\substack{i=1 \\ D=1}}^{n_1} (Y_i(1) - Y_i(0)) \right)}_{1a. \text{ diferencia}} - \underbrace{\left(\frac{1}{n_0} \sum_{\substack{i=1 \\ D=0}}^{n_0} (Y_i(1) - Y_i(0)) \right)}_{2a. \text{ diferencia}}. \quad (9)$$

Diferencia en diferencia

donde

- n : tamaño de la muestra.
- n_1 : tamaño de la muestra del grupo de tratamiento (beneficiarios).
- n_0 : tamaño de la muestra del grupo de control.

y además $n = n_1 + n_0$. La Ecuación 9 permite estimar un parámetro α_{EMPS} a partir de una diferencia de diferencias.

Ashenfelter y Card (1985), y Abadie (2005) demuestran que el estimador de la Ecuación 9 se puede obtener con un modelo semi-paramétrico del siguiente tipo:

$$Y = \gamma + \lambda D + \omega T + \alpha(DT) + \mu \quad (10)$$

donde

- Y : Variable dependiente (recoge los resultados potenciales de los individuos).
- D : Variable explicativa dicotómica (binaria); asigna los valores de la siguiente forma: $D = 1$ (individuo participa en el programa), $D = 0$ (individuo no participa en el programa).
- T : Variable explicativa dicotómica (binaria); asigna los valores de la siguiente forma: $T = 1$ (momento posterior al programa), $T = 0$ (momento previo al programa).
- $D \cdot T$: Variable explicativa (resulta de multiplicar las variables D y T).
- λ, ω : Parámetros asociados a las variables explicativas D y T .
- γ : Parámetro que muestra los efectos fijos en el modelo.
- μ : Error aleatorio, de media cero; esto es $E[\varepsilon | D, T] = 0$.
- α : Parámetro que determina el efecto conjunto de las variables o interacción de las variables explicativas D y T .

Al modelo que se representa en la Ecuación 10 se le puede agregar un conjunto de variables independientes – adicionales – que se consideren afectan la variable dependiente, quedando como

$$Y = \gamma + \lambda D + \omega T + \alpha(DT) + \beta X + \mu. \quad (11)$$

Se identifica con β al parámetro que multiplica al conjunto de las características adicionales que pudiesen afectar a la variable dependiente.

Las complicaciones para el modelo – Ecuación 11 – son que las variables X 's cambien de valor a través del tiempo y que éstas estén correlacionadas con μ (el error aleatorio). Cansino y Sánchez (2008) proponen que la primera dificultad se puede superar con la incorporación de la variable binaria T en efecto multiplicativo con las variables X 's, teniéndose:

$$Y = \gamma + \lambda D + \omega T + \alpha(DT) + \beta_1(TX) + \beta_0[(1 - T)X] + \mu. \quad (12)$$

Se ha argumentado que el método experimental resulta costoso, poco ético y se dificulta el seguimiento de características de interés de la población objetivo en la investigación en diferentes periodos de tiempo (Heckman y Smith, 2005). Algunos autores han resuelto las complicaciones anteriores con el uso de los

métodos cuasi-experimentales, conformando dos grupos de estudio (tratamiento versus control), buscándose el mayor parecido entre los individuos en ambos conjuntos, y controlando las características no observables que pueden ocasionar sesgo en la comparación.

Lo anterior, se formaliza al señalar que el individuo i toma la decisión de participar en un programa de capacitación en el momento q . Por lo que, $Y_{0i}(t)$ representa la ocupación informal del individuo i en el momento t en ausencia de capacitación ($P_i = 0$) y dependiente de otras características X_{it} . Posteriormente, $t > q$ periodos después del inicio de la capacitación, el individuo i puede seguir en la informalidad pero haberse beneficiado de la capacitación ($Y_{1i}(t)$) o no tenerla ($Y_{0i}(t)$) y siguen condicionados al conjunto de variables X_{it} . En términos de modelos, se tiene:

$$Y_{0i}(t) = \beta_{0t}X_{it} + u_{0it} \quad (13)$$

$$Y_{1i}(t) = \beta_{1t}X_{it} + u_{1it}. \quad (14)$$

La variable u en las dos ecuaciones representa el error aleatorio de las características no observables de los individuos. Regularmente se supone que $u_{0t} = u_{it}$ y $\beta_{0t} = \beta_{it}$.

Las dos ecuaciones (13) y (14) se pueden simplificar en una sola, al tomarse el promedio ponderado $Y(t)$ de las respuestas de los individuos que integran el grupo de tratamiento y control; además, suponiendo que $u_{0t} = u_{it}$ y $\beta_{0t} = \beta_{it}$, se tiene:

$$\begin{aligned} Y_{it} &= (1 - P_i)Y_{0it} + P_iY_{1it} \\ Y_{it} &= Y_{0it} - P_iY_{0it} + P_iY_{1it} \\ Y_{it} &= P_i(Y_{1it} - Y_{0it}) + Y_{0it} \\ Y_{it} &= P_i\alpha + Y_{0it} \\ Y_{it} &= P_i\alpha + \beta_tX_{it} + u_{it}. \end{aligned} \quad (15)$$

Recordar que el parámetro α es el estimador de diferencias en diferencias.

Suponiendo que $\beta_{0t} \neq \beta_{it}$ (hay variación a través del tiempo en X' s y $u_{0t} = u_{it}$ (se mantiene), entonces la Ecuación 15 se vuelve:

$$\begin{aligned} Y_{it} &= P_i(\beta_{1t} - \beta_{0t})X_{it} + \beta_{0t}X_{it} + u_{it} \\ Y_{it} &= (1 - P_i)(\beta_{0t}X_{it} + u_{0it}) + P_i(\beta_{1t}X_{it} + u_{1it}) \\ Y_{it} &= \beta_{0t}X_{it} + u_{0it} - P_i\beta_{0t}X_{it} - P_iu_{0it} + P_i\beta_{1t}X_{it} + P_iu_{1it} \\ Y_{it} &= P_i(\beta_{1t} - \beta_{0t})X_{it} + \beta_{0t}X_{it} + u_{0it} + P_i(u_{1it} - u_{0it}) \end{aligned}$$

y puesto que $u_{1it} = u_{0it}$

$$Y_{it} = P_i(\beta_{1t} - \beta_{0t})X_{it} + \beta_{0t}X_{it} + u_{it}. \quad (16)$$

Si $\beta_{0t} \neq \beta_{it}$ junto con $u_{0t} \neq u_{it}$, entonces,

$$Y_{it} = P_i(\beta_{1t} - \beta_{0t})X_{it} + \beta_{0t}X_{it} + u_{0it} + P_i(u_{1it} - u_{0it})$$

Otras expresiones a destacar sobre las características de los resultados son *treatment on the treated* (TT) y *average treatment effect* (ATE), que se indican a continuación:

$$TT = E[Y_{1it} - Y_{0it} \mid X_{it}, P_i = 1] = (\beta_{1t} - \beta_{0t})X_{it} + E[u_{1it} - u_{0it} \mid X_{it}, P_i = 1] \quad (17)$$

$$ATE = E[Y_{1it} - Y_{0it} \mid X_{it}] = (\beta_{1t} - \beta_{0t})X_{it} + E[u_{1it} - u_{0it} \mid X_{it}]. \quad (18)$$

Las ecuaciones (17) y (18) son iguales si se cumple

$$E[u_{1it} \mid X_{it}, P_i = 1] = E[u_{0it} \mid X_{it}, P_i = 1] = E[u_{0it} \mid X_{it}] = E[u_{1it} \mid X_{it}],$$

situación que sucede cuando los errores son iguales.

Se debe señalar que un aprieto en la metodología es el sesgo que estará presente cuando se genera la muestra del grupo de control, puesto que, no es seleccionada al azar. Una primera solución es la técnica del pareo; empero, resulta costosa por los requisitos a cubrir y su evaluación no es sencilla (Navarro *et al.*, 2006). Otra opción – lo más recomendable – es el empleo del enfoque sobre emparejamiento (*matching*) no paramétrico llamado *Propensity Score Matching* (PSM). Éste consiste en la obtención del efecto promedio entre el valor del resultado condicionado a un tratamiento versus su no exposición.

Desarrollando la idea sobre el PSM a detalle, la intención es seleccionar para cada individuo quien recibió la capacitación, un gemelo quien no la tomó pero que tiene la misma propensión a obtenerla; además, considerando un conjunto de variables adicionales y explicativas que pueden influir en la dependiente (ocupación en el sector informal). Las primeras se reducen a un escalar, el *propensity score* (PS) o $p(X)$, para que el emparejamiento sea más factible; formalmente, el planteamiento es:

$$p(x) \equiv P(D = 1 \mid X) = E(D \mid X). \quad (19)$$

Sin embargo, el hecho de que la asignación de la capacitación no es aleatoria, la variable dependiente cuando el individuo no participa en el programa de

capacitación no se trata de la estimación de una simple probabilidad sino que tienen que cumplirse los siguientes supuestos (Rosenbaum y Rubi, 1983; Becker e Ichino, 2002):

- Supuesto I: la distribución de probabilidad de las características de los individuos es independiente de la decisión sobre participar o no en los programas de capacitación.
- Supuesto II: el valor de la variable dependiente es independiente en cuanto a sí el individuo tomó o no la formación (Rubin, 1977).
- Supuesto III: los resultados probables de la variable dependiente sobre un individuo no están en función de las acciones que realicen los demás (Bryson, Dorsett y Purdon, 2002).

Una vez establecidos los supuestos que para esta investigación se cumplan, es decir, ya que no hay relación alguna entre los grupos de tratamiento y control, el efecto promedio del tratamiento puede ser estimado como

$$\begin{aligned}\tau &= E[Y_{1i} - Y_{0i} \mid D_i = 1] \\ &= E[E[Y_{1i} - Y_{0i} \mid D_i = 1, p(X_i)]] \\ &= E[E[Y_{1i} \mid D_i = 1, p(X_i)]] - E[E[Y_{0i} \mid D_i = 0, p(X_i) \mid D_i = 1]].\end{aligned}$$

El emparejamiento más cercano del *propensity score* es el *Nearest Neighbor Matching* (NNM) o vecino más cercano, y matemáticamente se obtiene como a continuación se expresa:

$$C_{(i)} = \underset{j}{\text{mín}} \|p_i - p_j\|$$

donde $C_{(i)}$ se refiere a los individuos de control emparejados con una unidad (i) de beneficiarios dado un *propensity score* (p_i).

El estimador del NNM es el siguiente:

$$\tau^{NNM} = \frac{1}{NT} \sum_{i \in T} Y_i^T - \frac{1}{NT} \sum_{j \in T} w_j Y_j^C,$$

donde

- Y_i^T : variable resultado del grupo de tratamiento.
- Y_j^C : variable resultado del grupo de control.
- N_i^C : indica el número de controles emparejados con observaciones $i \in T$.
- $w_{ij} = \frac{1}{N_i^C}$ sí $j \in C_{(i)}$ y $w_{ij} = 0$.