

# ABCD' ECONOMÍA VBCD, ECONOMÍA V

**TOMO III**

**Alejandro Mosiño**  
**COORDINADOR**

UNIVERSIDAD DE  
GUANAJUATO





# ABCD'ECONOMÍA

## Tomo III

Alejandro Mosiño  
coordinador

UNIVERSIDAD DE  
GUANAJUATO



*ABCD'ECONOMÍA, Tomo III*, Primera edición, 2020

D.R. © Del texto: los autores

D.R. © De la edición:

Universidad de Guanajuato, Campus Guanajuato

División de Ciencias Económico Administrativas

Departamento de Economía y Finanzas

Fraccionamiento 1, Col. El Establo, C.P. 36250, Guanajuato, Gto., México

D.R. © De la edición: Secularte A.C.

<p>Esta obra está conformada por los trabajos sometidos y aceptados de acuerdo con la convocatoria emitida por el Departamento de Economía y Finanzas de la Universidad de Guanajuato. Los trabajos se sometieron al sistema de dictaminación de doble ciego por pares académicos con reconocimiento nacional en sus áreas. El resultado de los dictámenes fue positivo.</p>
--

Maquetación: Luis Villalobos

ISBN: 978-607-441-786-9

El presente ejemplar es de distribución y descarga en acceso abierto conforme a lo permitido por la licencia Creative Commons.

Hecho en México - *Made in Mexico*

# Índice

Presentación		7
Limpieza y análisis de bases de datos: un ejemplo introductorio	<i>Alejandro Mosiño</i>	11
Introducción a la metodología de la investigación. Un caso práctico	<i>Antonio Baez Morales</i>	23
Macroeconomía dinámica aplicada: cambio estructural en México	<i>Édgar Cruz</i> <i>Fernando García Barragán</i>	49
Migración internacional	<i>Miriam Saldaña Hernández</i> <i>Coralía Quintero Rojas</i>	85
Redes sociales	<i>José Antonio Carrillo Viramontes</i> <i>Lari Arthur Viñanto</i> <i>Alejandro Tatsuo Moreno Okuno</i>	101

# Índice

Presentación		7
Limpieza y análisis de bases de datos: un ejemplo introductorio	<i>Alejandro Mosiño</i>	11
Introducción a la metodología de la investigación. Un caso práctico	<i>Antonio Baez Morales</i>	23
Macroeconomía dinámica aplicada: cambio estructural en México	<i>Édgar Cruz</i> <i>Fernando García Barragán</i>	49
Migración internacional	<i>Miriam Saldaña Hernández</i> <i>Coralía Quintero Rojas</i>	85
Redes sociales	<i>José Antonio Carrillo Viramontes</i> <i>Lari Arthur Viñanto</i> <i>Alejandro Tatsuo Moreno Okuno</i>	101

## Lista de autores

**Departamento de Economía y Finanzas**

**Universidad de Guanajuato**

Alejandro Mosiño  
alejandro.mosino@gmail.com

Antonio Baez Morales  
antonio.baez@ugto.mx

B. Édgar Cruz González  
be.cruz@ugto.mx

Fernando García Barragán  
garcia.barragan@gmail.com

Miriam Saldaña Hernández  
miriam.saldana@ugto.mx

Coralía A. Quintero Rojas  
coralia\_azucena@yahoo.com

José Antonio Carrillo Viramontes  
ja.carrillo@ugto.mx

Lari A. Viianto  
la.viianto@ugto.mx

Alejandro Tatsuo Moreno Okuno  
atatsuo@hotmail.com

## Presentación

**E**l Tomo III de la serie *ABCD'Economía* tiene como principal objetivo el de complementar las herramientas presentadas en los tomos precedentes. En el primer tomo presentamos las que consideramos son las cuatro grandes áreas en Economía: Métodos Cuantitativos, Microeconomía, Macroeconomía y Econometría, de las cuales se derivan todas las demás áreas de interés para los economistas, así como un sinnúmero de aplicaciones. Algunas de estas fueron presentadas en el segundo tomo, el cual tenía como objetivo despertar la curiosidad del lector y su entusiasmo por analizar los fenómenos económicos a su alrededor. Además, vimos que en Economía no todo son matemáticas, si no que el estado del arte de esta ciencia es consecuencia de la interacción entre varias corrientes de pensamiento.

En este tomo tenemos, además de los objetivos planteados en los tomos I y II, un reto adicional: despertar el interés de nuestros lectores, y en particular aquellos que cursan los últimos semestres de la carrera o maestría en Economía, por tomar el camino de la investigación.

En el Capítulo 1 tratamos de saldar una deuda pendiente con nuestros estudiantes. Si bien en nuestros cursos intentamos entrenarlos para hacer análisis empíricos utilizando datos y softwares especializados, casi nunca les mostramos cómo es el proceso de búsqueda y recolección de información, así como el trabajo que implica el limpiar las bases de datos para que estas queden listas para introducirse a la computadora para su análisis posterior. Para ver la importancia de esto, basta con tirarse un clavado a las encuestas del Instituto Nacional de Estadística e Informática (INEGI), las cuales están divididas en varios módulos y/o cuestionarios. ¿Cómo aprovechamos toda esta información? ¿Cómo descartamos información que resulte irrelevante?

El Capítulo 2 constituye una primera aproximación a la realización de un trabajo de investigación en Economía. En este se explica paso a paso cuál es la metodología que podría seguir un economista para analizar y

reportar un tema de interés. En específico, el autor toma como ejemplo una investigación publicada en el libro *Problemas del Desarrollo en México II* para mostrarnos, primero, cuáles son sus partes. Luego, el autor nos muestra cómo obtiene y cómo se analiza la información. Así mismo, aprenderemos cómo una teoría abstracta se traduce en un análisis empírico. Finalmente, veremos un ejemplo de cómo realizar la presentación de resultados. Es importante aclarar que, si bien el autor hace referencia al paquete estadístico y de análisis de datos Stata, todo el procedimiento puede realizarse utilizando otros paquetes computacionales, ya sea comerciales o de distribución libre.

En el Capítulo 3 tenemos un ejemplo de Macroeconomía aplicada. En específico, los autores analizan el concepto de cambio estructural, el cual hace referencia a la relocalización de la actividad económica a través de los sectores productivos: agrícola, industrial y de servicios. Esta relocalización se observa en la mayoría de los países y tiene importantes implicaciones para el desarrollo y el crecimiento. Primero, los autores nos proporcionan algunos hechos estilizados sobre el cambio estructural. Luego nos presentan algunas de las teorías que nos ayudan a explicar el fenómeno de cambio estructural. Si bien la teoría es abstracta, los autores nos muestran cómo puede explicarse el fenómeno de cambio estructural en México utilizando datos reales. A diferencia de los métodos más conocidos para la realización de análisis empíricos en economía (la econometría, específicamente), los autores llevan a cabo una simulación computacional de un modelo mediante el uso del paquete computacional Dynare.

El capítulo 4 nos muestra un ejemplo de cómo obtener lo más de los datos disponibles y de cómo llevar a cabo un análisis utilizando no más que gráficas y estadística descriptiva. Para esto, las autoras abordan el fenómeno de la migración internacional. Específicamente, hacen una revisión global de las tendencias migratorias y ponen en contexto la migración internacional de México y sus principales causas. A partir de los datos se obtiene, por ejemplo, que, contrario a la creencia popular, en la actualidad los migrantes que entran a los Estados Unidos de manera ilegal son un porcentaje relativamente bajo. Además, la gran mayoría de los migrantes han reportado tener los documentos legales que les permiten trabajar en ese país.

Finalmente, el capítulo 5 nos presenta uno de los temas que más interés han atraído entre los economistas y científicos de otras áreas. Este es el del análisis de las redes sociales. Contrario a lo que podríamos pensar, las redes sociales son más que aquellas vinculadas con las plataformas

de internet, tales como Facebook, Instagram, YouTube, entre otras. Más bien, una red social intenta representar los diversos tipos de relaciones e interacciones que surgen entre los individuos que forman parte de estas, las cuales suelen tener un origen de carácter emocional. Primero, los autores nos enseñan cómo se representan las redes sociales. Luego, veremos en qué consiste la formación estratégica de redes, la cual se encarga de analizar y definir las formas particulares que toman las redes sociales. Finalmente, los autores nos muestran algunos ejemplos y aplicaciones.

Como siempre, no resta más que agradecerle a nuestros lectores por el tiempo que puedan dedicar a la lectura de este tomo y a los precedentes de la serie. Deseamos que estos estén a la altura de sus expectativas y que, sobre todo, logren esclarecer sus dudas con respecto a la labor diaria de los economistas, así como despertar su interés por realizar investigación autónoma.

Alejandro Mosiño



# Limpieza y análisis de bases de datos: un ejemplo introductorio

Alejandro Mosiño

## 1 Introducción

En contadas ocasiones la información que utilizamos para llevar a cabo cualquier tipo de análisis económico (o de cualquier índole) se encuentra disponible en forma limpia y lista para introducirse a la computadora. Usualmente mucho es el trabajo previo. De hecho, sobran los econométristas que opinan que el trabajo más pesado se encuentra en la búsqueda de la información relevante, así como en la limpieza y organización de las bases de datos. Adicionalmente, la mayor parte de los resultados de las encuestas son presentados en partes y el vaciado de las respuestas no siempre nos permite hacer un análisis preciso de la información. Por ejemplo, la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) del INEGI consta de 14 módulos y la Encuesta Nacional de Salud y Nutrición (ENSANUT) del INSP consta de 14 cuestionarios. En estos casos, podría resultar indispensable que unamos o combinemos dos o más bases de datos, así como hacer un análisis exploratorio previo de la base de datos resultante para identificar, por ejemplo, respuestas incoherentes o eliminar variables o datos que no sirvan para los propósitos del investigador.

En este capítulo, daremos un ejemplo utilizando una base de datos (ficticia) que contiene información obtenida de varios hospitales en Estados Unidos sobre sus pacientes con cáncer. El objetivo es: 1) realizar un análisis exploratorio de la información y 2) limpiar la base de datos tal que sea posible hacer análisis estadísticos más profundos. Cabe mencionar que este tipo de procedimientos pueden llevarse a cabo utilizando una gran gama de paquetes computacionales. Estos van desde los más básicos como MS Excel o Google Sheets, hasta los más sofisticados como Matlab, R o Python. Para los econométristas, este tipo de procedimientos

pueden realizarse en paquetes comerciales como Stata y Eviews o bien en paquetes de distribución libre como Gretl. Por regla general, no se recomienda utilizar MS Excel o Google Sheets más que para almacenar la información.

## 2 Limpieza de la base de datos y valores ausentes

La mayor parte de las bases de datos que se encuentran en Internet son distribuidos en formato xls o xlsx de Excel, CSV, Stata o similares. En cualquiera de los casos, los paquetes computacionales de análisis de datos pueden (virtualmente) importar cualquier tipo de archivo. Sin embargo, conviene siempre asegurarse que los datos se descargan en un formato legible.

La base de datos inicial que aquí consideramos consta de 24 variables y 120 observaciones que simulan información sobre pacientes con cáncer. Una pequeña muestra de los datos se presenta en la Tabla 1. Como primer paso, conviene hacer un análisis descriptivo simple. Esto nos permitirá no solo conocer la distribución de la muestra, si no que también nos permite detectar valores sospechosos. Los estadísticos descriptivos más importantes de nuestra base de datos se muestran en la Tabla 2. Entre los indicadores más relevantes se encuentran el número de observaciones (N), la media (Mean) la desviación estándar (St. Dev.), los valores mínimos y máximos (Min y Max, respectivamente) y el primero y tercer percentiles de la muestra (Pctl25 y Pctl75, respectivamente). En nuestra base de datos podemos ver, por ejemplo, que la edad promedio de los pacientes (age) es de 53.59 años y que el número promedio de tumores detectados por paciente (ntumors) es de 3.15.

Una simple inspección a la Tabla 2 nos permite detectar algunos valores sospechosos. Por ejemplo, el paciente con más edad en la base de datos (columna Max) tiene ¡357 años! Naturalmente esto es un error que puede deberse a varias causas, pero la más obvia es un descuido en la captura de la información sobre ese paciente.

La base de datos que analizamos tiene otro tipo de información que, por su propia naturaleza, no puede presentarse en la Tabla 2. Esta es, principalmente, información cualitativa. Por ejemplo, el hospital en el cual se atendieron los pacientes, el género del paciente, el estadio del cáncer, entre otros. Para ejemplificar, en la Tabla 3 presentamos el resumen estadístico de la variable género (sex). Como podemos ver ahí,

Tabla 1: Base de datos

	hospital	hospid	docid	dis_date	sex	age	test1	test2	pain	tumorsize	co2	wound
1	UCLA	1.00	1-1	6-Sep-09	male	64.97	3.70	8.09	4.00	67.98	1.53	4.00
2	UCLA	1.00	1-1	7-Jan-11	female	53.92	2.63	0.80	2.00	64.70	1.68	3.00
3	UCLA	1.00	1-1	4-Sep-10	male	41.37	-99.00	2.13	3.00	86.44	1.45	3.00
4	UCLA	1.00	1-1	25-Jun-10	male	46.80	3.89	1.35	3.00	53.40	1.57	4.00
5	UCLA	1.00	1-1	1-Jul-09	male	51.93	1.42	2.19	4.00	51.66	1.42	5.00
6	UCLA	1.00	1-1	6-Mar-09	female	53.83	2.29	8.61	3.00	78.92	1.71	4.00
7	UCSF	2.00	2-201	17-Jul-10	male	49.11	2.62	5.99	6.00	58.36	1.57	6.00
8	UCSF	2.00	2-201	18-Feb-11	female	54.06	9.56	2.91	8.00	74.90	1.75	5.00
9	UCSF	2.00	2-216	25-Mar-10	male	52.07	11.07	6.88	7.00	69.65	1.45	7.00
10	UCSF	2.00	2-216	11-Apr-10	female	51.96	6.70	9.37	9.00	78.52	1.70	6.00
11	UCSF	2.00	2-216	11-Dec-09	female	51.09	12.42	1.36	6.00	74.02	1.64	6.00
12	UCSF	2.00	2-177	9-Jun-10	male	58.74	7.62	9.31	9.00	63.45	1.48	9.00

Tabla 2: Estadística descriptiva de una base datos

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
hospid	120	1.483	0.502	1	1	2	2
age	120	53.590	28.679	34.192	47.748	55.012	357.890
test1	120	-1.989	24.392	-99.000	1.560	5.707	12.416
test2	120	-1.226	24.639	-99.000	2.249	6.166	17.228
pain	120	5.325	1.567	1	4	6	9
tumorsize	120	70.083	11.655	49.099	62.179	77.478	109.010
co2	120	-0.890	15.616	-98.000	1.495	1.686	1.942
wound	120	5.592	1.590	2	5	7	9
mobility	120	6.033	2.004	2	5	7	9
ntumors	120	3.150	2.779	0	1	5	9
remission	120	0.350	0.479	0	0	1	1
lungcapacity	120	-18.328	39.384	-99.000	0.514	0.872	0.998
married	120	0.633	0.484	0	0	1	1
lengthofstay	120	5.308	1.143	3	4	6	8
rbc	120	4.970	0.254	4.360	4.825	5.150	5.535
bmi	120	29.383	6.745	18.450	24.505	34.305	58.000

tenemos 120 observaciones, 0 datos ausentes y ¡3 valores diferentes! Estos valores son: hombre, mujer y ¡12.2! Tenemos 74 mujeres (61.7 % de la muestra), 45 hombres (37.5 % de la muestra) y 1 “12.2” (0.8 % de la muestra). Obviamente la observación cuyo género es “12.2” constituye un error que debemos corregir.

Tabla 3: Estadística descriptiva para una variable cualitativa

<b>sex</b>			
N	120		
Missing	0		
Distinct	3		
Value	12.2	female	male
Frequency	1	74	45
Proportion	0.008	0.617	0.375

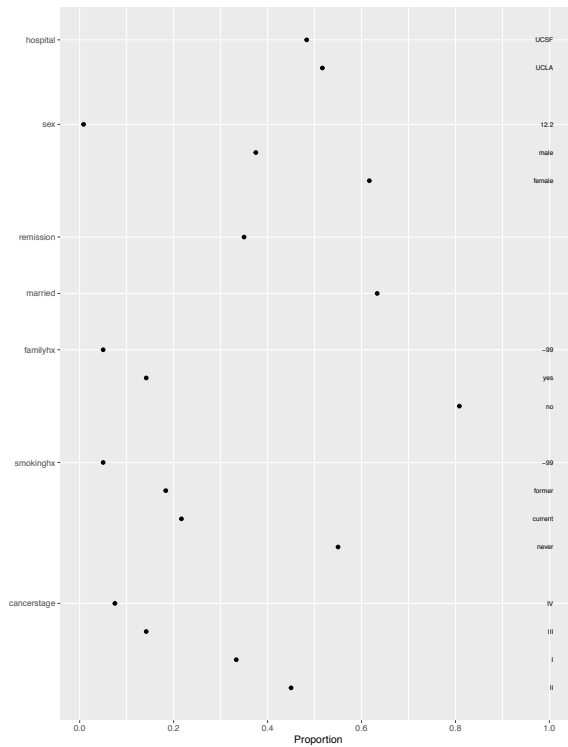


Figura 1: Estadística descriptiva. El caso de las variables cualitativas.

Existen otras maneras de analizar las bases de datos y de buscar incoherencias. Algunos paquetes computacionales nos permiten visualizar todo el resumen estadístico a partir de gráficas en lugar de tablas. Uno de estos paquetes es R, el cual nos permite graficar la estadística descriptiva como en las Figuras 1 y 2. Como podemos ver, la Figura 1 es una gráfica de puntos para las variables categóricas (discretas o cualitativas). En el eje vertical izquierdo la gráfica nos indica la variable y en el eje vertical derecho las categorías. Por ejemplo, vemos que la variable de género (sex) tiene, como hemos dicho, tres categorías, una de las cuales constituye una incoherencia.

La Figura 2 contiene un conjunto de histogramas para las variables continuas (o cuantitativas). Cada histograma tiene su propia escala. En este sentido, cuando un histograma está demasiado alargado pero, a la vez, tiene grupos muy compactos, podríamos estar en presencia de datos extremos. Por ejemplo, el histograma para la variable de edad (age) es muy compacto del lado izquierdo, pero aun así está alargado debido al paciente que tiene, supuestamente, 357 años.

En general, las gráficas tienden a revelar más información sobre los valores sospechosos, a la vez que son más simples de analizar que las tablas numéricas. Sin embargo, ambas revelan la misma información. En nuestro caso, además de las incoherencias mencionadas antes, las Figuras 1 y 2 revelan que:

- Algunas variables, como test1, test2, lungcapacity, entre otras tienen valores de -99. Usualmente, en las encuestas este valor corresponde “no sabe” o “ausente”.
- Algunas variables, como co2 y lungcapacity, tienen valores de -98. Usualmente, en las encuestas este valor corresponde a “no quiso contestar”.
- La variable wbc (conteo de glóbulos blancos en la sangre) debería ser una variable numérica. Sin embargo, debido a que tiene valores “Not Assesed” (no calculado), los paquetes computacionales tienden a confundirla con una variable cualitativa.

Por supuesto, los valores de -98 y -99 no nos sirven, al igual que todas las incoherencias encontradas. Entonces, para que nuestra base de datos sea utilizable, podemos reemplazar todos estos valores sospechosos por valores ausentes. Usualmente, esto corresponde a un “NA” en la mayor parte de los paquetes estadísticos (o celdas vacías en MS Excel). Claro que lo ideal sería reemplazar los valores sospechosos con sus valores verdaderos, pero esto casi siempre es imposible porque se trata de información a la que, por lo general, no tenemos acceso.

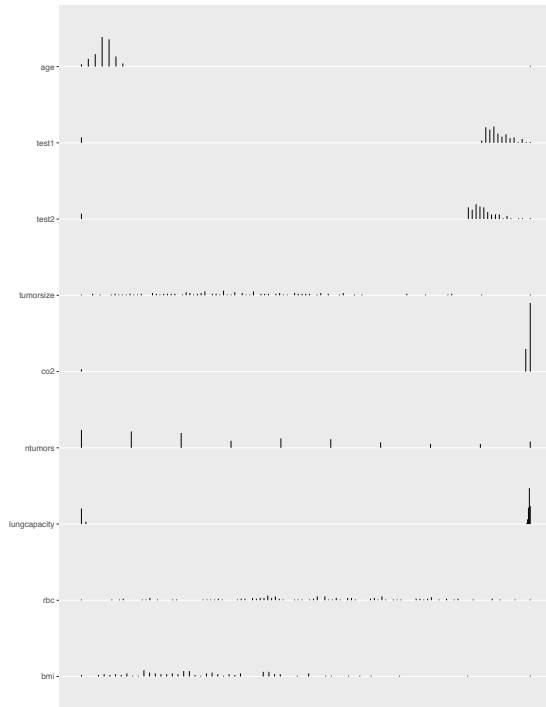


Figura 2: Estadística descriptiva. El caso de las variables cuantitativas.

Reemplazamos entonces los valores -98 y -99 por NA, al igual que el género de la persona “12.2”, la edad de las personas menores a 18 años y mayores a 120,<sup>1</sup> y los “Not Assesed” de toda la base de datos.

¿Cómo se ven las Figuras 1 y 2 luego de estas modificaciones? El resultado se muestra en las Figuras 3 y 4. Como podemos constatar, en el caso de las variables discretas no parecen existir más incoherencias. En el caso de las variable continuas, los histogramas lucen menos alargados y sus datos mejor distribuidos. Además, como hemos reemplazado los valores incoherentes con NA, la gráfica también incluye un mapa de calor para indicarnos el número de valores ausentes para cada una de las variables.

<sup>1</sup> La base, aunque ficticia, analiza los casos de cáncer en personas adultas. Por otra parte, aunque es muy complicado encontrar personas de 120 años, biológicamente no es imposible.

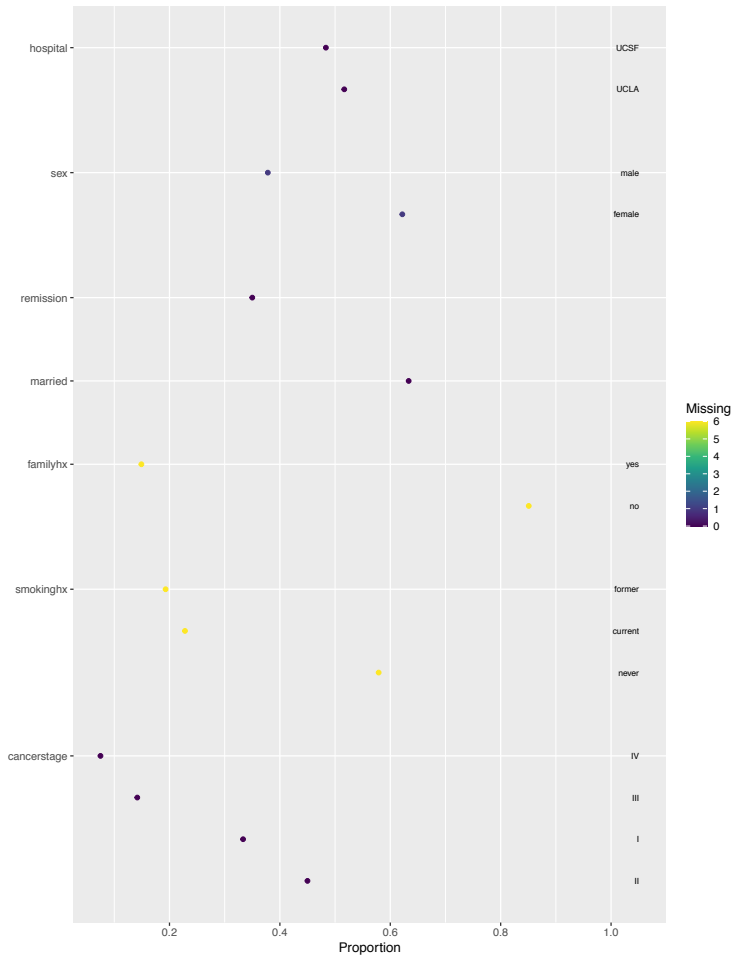


Figura 3: Estadística descriptiva. El caso de las variables cualitativas. Base de datos limpia.

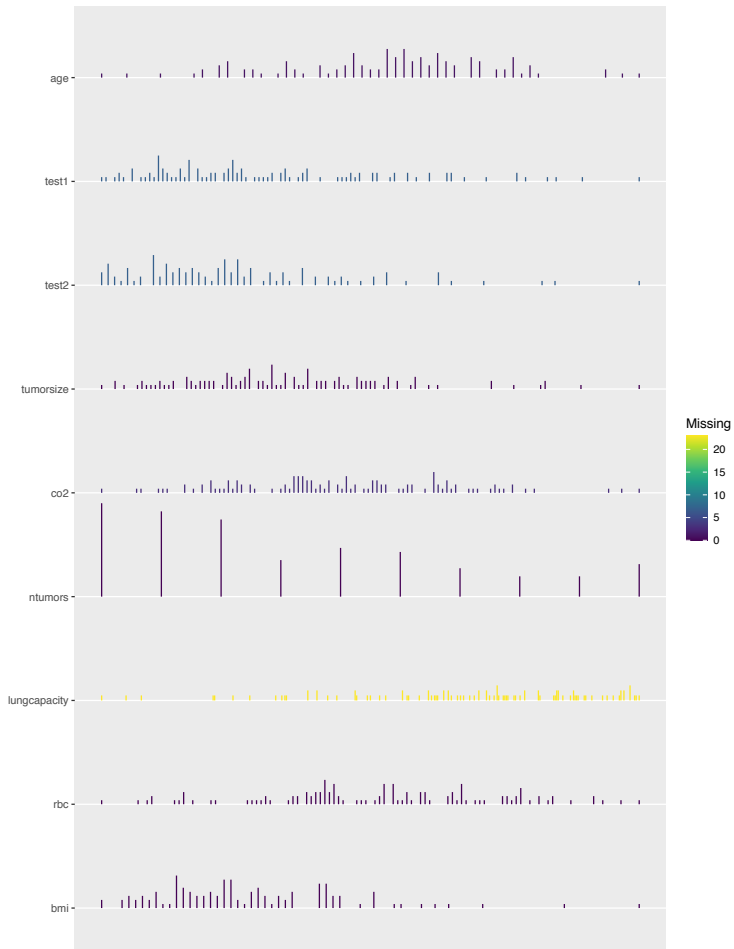


Figura 4: Estadística descriptiva. El caso de las variables cuantitativas. Base de datos limpia.

### 3 Unir y combinar bases de datos

Como hemos dicho, en ocasiones las bases de datos están divididas en varias partes. Cada parte puede contener las mismas variables para nuevos individuos, o cada parte puede contener nuevas variables para los mismos individuos. En el primer caso, lo que tenemos que hacer es unir las diferentes partes de las bases de datos, y en el segundo lo que tenemos que hacer es combinarlas.

#### 3.1 Unir bases de datos

Considera una segunda parte de la base de datos que hemos considerado hasta ahora. Una muestra de la información la presentamos en la Tabla 4. Esta nueva base contiene prácticamente las mismas variables que la base de datos original, pero contiene nuevos pacientes (de un tercer hospital). Otras diferencias entre las dos bases de datos resultan de las manipulaciones que hemos hecho arriba a la base de datos original.

Tabla 4: Complemento a base de datos

	hospital	hospid	docid	dis_date	sex	age	test1	test2	pain	tumorsize	co2	wound
1	Cedars-Sinai	3.00	3-227	1-Oct-09	male	54.20	3.87	1.37	6.00	69.77	1.53	5.00
2	Cedars-Sinai	3.00	3-227	18-Feb-10	female	55.07	7.76	4.50	8.00	68.03	1.69	4.00
3	Cedars-Sinai	3.00	3-227	30-Jun-09	female	58.06	5.73	-99.00	7.00	65.06	1.56	4.00
4	Cedars-Sinai	3.00	3-227	15-Nov-09	female	51.52	8.58	4.61	9.00	71.44	-98.00	5.00
5	Cedars-Sinai	3.00	3-227	17-Feb-10	female	54.63	6.70	2.64	4.00	69.52	1.68	5.00
6	Cedars-Sinai	3.00	3-227	22-Dec-09	male	49.09	2.87	5.46	5.00	89.74	1.89	6.00

El trabajo de unir dos o más bases de datos es el más sencillo de los que aquí analizaremos. Para ver por qué, basta con imaginar cuál sería el procedimiento en el paquete de manejo de datos más sencillo del que disponemos (o al menos al que estamos más habituados): MS Excel. En principio, lo único que tendríamos que hacer es copiar la segunda base de datos y pegarla a la base original a partir de la primera fila de celdas vacía.<sup>2</sup> El resultado sería muy similar al de la Tabla 5.

Naturalmente, en el proceso de unir dos bases de datos podríamos encontrar algunas complicaciones. La más común es la que las bases de datos no tengan exactamente las mismas columnas; en otras palabras, que una de las bases de datos que vamos a unir tenga columnas que la otra no. También los datos en las bases podrían ser de diferentes tipos aun perteneciendo a la misma variable. Por ejemplo, en una de las bases los

<sup>2</sup> Por supuesto, aquí estamos obviando el hecho de que la segunda base de datos también tiene que limpiarse siguiendo el mismo procedimiento que realizamos en la sección anterior.

Tabla 5: Unión de bases de datos

source	hospital	hospid	docid	dis_date	sex	age	test1	test2	pain	tumorsize	co2	
1	1	UCLA	1.00	1-1	6-Sep-09	male	64.97	3.70	8.09	4.00	67.98	1.53
2	1	UCLA	1.00	1-1	7-Jan-11	female	53.92	2.63	0.80	2.00	64.70	1.68
3	1	UCLA	1.00	1-1	4-Sep-10	male	41.37		2.13	3.00	86.44	1.45
4	1	UCLA	1.00	1-1	25-Jun-10	male	46.80	3.89	1.35	3.00	53.40	1.57
5	1	UCLA	1.00	1-1	1-Jul-09	male	51.93	1.42	2.19	4.00	51.66	1.42
6	1	UCSF	2.00	2-201	18-Feb-11	female	54.06	9.56	2.91	8.00	74.90	1.75
7	1	UCSF	2.00	2-216	25-Mar-10	male	52.07	11.07	6.88	7.00	69.65	1.45
8	1	UCSF	2.00	2-216	11-Apr-10	female	51.96	6.70	9.37	9.00	78.52	1.70
9	1	UCSF	2.00	2-216	11-Dec-09	female	51.09	12.42	1.36	6.00	74.02	1.64
10	1	UCSF	2.00	2-177	9-Jun-10	male	58.74	7.62	9.31	9.00	63.45	1.48
11	2	Cedars-Sinai	3.00	3-394	2-Apr-10	male	54.67	0.33	5.88	6.00	70.62	1.75
12	2	Cedars-Sinai	3.00	3-394	6-May-09	male	47.19	2.88	6.45	6.00	64.92	1.71
13	2	Cedars-Sinai	3.00	3-227	17-Feb-10	female	54.63	6.70	2.64	4.00	69.52	1.68
14	2	Cedars-Sinai	3.00	3-297	25-Mar-10	male	70.27	2.31	2.90	4.00	63.43	1.49
15	2	Cedars-Sinai	3.00	3-351	20-Oct-10	female	52.40	6.05	1.86	6.00	79.02	1.78
16	2	Cedars-Sinai	3.00	3-362	18-Jul-09	female	61.07	1.03	1.40	3.00	70.63	1.53

datos de una columna podrían ser del tipo texto, mientras que en la otra ser del tipo numérico. En presencia de complicaciones, el procedimiento en MS Excel es tedioso, pero no tan complicado. Afortunadamente, los paquetes de cómputo más sofisticados pueden automatizar el proceso. Por ejemplo, para la primer complicación, la computadora podría crear automáticamente las columnas faltantes en cada una de las bases de datos y rellenar los espacios con NA.

### 3.2 Combinar bases de datos

Ahora supongamos que una tercera base de datos contiene nuevas variables, pero (virtualmente) los mismos individuos. Nota que, a diferencia del caso anterior, ahora debemos pegar columnas en lugar de filas. En este caso, las bases de datos pueden combinarse *si existe alguna o algunas variables que nos permitan identificar a los individuos*. Este procedimiento no es tan simple como el anterior, al menos no en paquetes como MS Excel. En un mundo ideal, los individuos en cada una de las bases están colocados exactamente en la misma posición. Por ejemplo, el individuo 1 en la primer base es el individuo 1 en la segunda, el individuo 2 es el segundo en ambas bases y así sucesivamente. También, en un mundo ideal, no faltan (o sobran) individuos en alguna de las bases. Lamentablemente, en la práctica, es muy difícil encontrar estas condiciones.

Consideremos los datos de la Tabla 6, la cual presenta una parte de la información sobre los médicos que han atendido a pacientes con cáncer en cualquiera de los hospitales considerados. Esta información incluye, además del identificador de médico, sus años de experiencia, el tipo de

escuela de medicina de procedencia y el número de demandas que ha recibido durante su carrera.

Tabla 6: Una tercera base de datos

	docid	experience	school	lawsuits	medicaid
1	1-1	25.00	average	3.00	0.61
2	1-11	10.00	top	0.00	0.60
3	1-21	21.00	average	3.00	0.48
4	1-22	22.00	top	3.00	0.48
5	1-33	16.00	top	0.00	0.58
6	1-48	23.00	average	3.00	0.22

Nuestro objetivo es incluir la información de la Tabla 6 en la base de datos de la Tabla 5. La base de datos completa podría ser útil para analizar, por ejemplo, la probabilidad de éxito en la atención de un paciente, o la probabilidad de que un médico reciba una nueva demanda. Una de las principales dificultades en el procedimiento se encuentra, por ejemplo, en que el médico con identificador 1-1 ha atendido a más de un paciente (al menos los primeros 5 en la Tabla 5, pero obviamente hay más cuando consideramos los datos completos). Además, es posible que el mismo médico haya atendido a otros pacientes en alguno de los otros hospitales. Lo mismo el médico 1-11, 1-21, etcétera. ¿Qué tan fácil es asociar la información de cada médico a cada uno de sus pacientes? Imagina que tuviéramos que hacer esto “manualmente” copiando y pegando celdas en MS Excel. Nuevamente, y para nuestra fortuna, los paquetes estadísticos y econométricos cuentan con rutinas relativamente simples de automatizar este proceso.

Finalmente, existen algunas otras complicaciones a tomar en cuenta antes de combinar dos bases de datos. Por ejemplo, ¿qué pasa si tenemos uno o varios médicos en la Tabla 6 que no aparecen en la Tabla 5? ¿Qué haríamos si tuviéramos uno o varios pacientes en la Tabla 5 cuyo médico no aparece en la Tabla 6? Para los paquetes computacionales esto no es un problema, pero el analista tiene que decidir qué hacer en cada caso. Una solución podría ser eliminar a los pacientes sin información sobre su médico. Otra opción es rellenar la información ausente con “NA”. Para cada caso existe alguna solución simple a nivel computacional. Una parte de la base de datos resultante se muestra en la Tabla 7.

Ahora sí, con la base de datos completa, podemos proceder a hacer cualquier análisis estadístico o econométrico a partir de ella.

Tabla 7: Base de datos final

	source	hospital	hospid	docid	dis_date	sex	age	experience	school	lawsuits	medicaid
1	1	UCLA	1.00	1-1	6-Sep-09	male	64.97	25.00	average	3.00	0.61
2	1	UCLA	1.00	1-1	7-Jan-11	female	53.92	25.00	average	3.00	0.61
3	1	UCLA	1.00	1-1	4-Sep-10	male	41.37	25.00	average	3.00	0.61
4	1	UCLA	1.00	1-1	25-Jun-10	male	46.80	25.00	average	3.00	0.61
5	1	UCLA	1.00	1-1	1-Jul-09	male	51.93	25.00	average	3.00	0.61
6	1	UCLA	1.00	1-1	6-Mar-09	female	53.83	25.00	average	3.00	0.61

## 4 Conclusión

En este capítulo hemos visto una parte del procedimiento necesario para limpiar y analizar una base de datos. En específico, hemos visto que la mayor parte de las bases contienen datos incoherentes y errores que debemos tomar en cuenta antes de comenzar con cualquier estudio. Además, la información contenida en las encuestas casi siempre se encuentra en partes, por lo que es muy común que tengamos que hacer un esfuerzo adicional para unir y/o combinar las bases tal que estas puedan introducirse a la computadora para su análisis posterior.

El procedimiento que hemos visto aquí ha sido muy genérico, por lo que los pasos en concreto van a depender del paquete computacional elegido por el analista. Existen muchas opciones, desde los basados en hojas de cálculo hasta los que requieren de un lenguaje de programación más sofisticado. En definitiva, es muy recomendable que los responsables de manejar bases de datos elijan la segunda opción. Si bien el aprendizaje de un lenguaje nuevo requiere de tiempo, el beneficio es, a la larga, bastante alto.

# Introducción a la metodología de la investigación. Un caso práctico

Antonio Baez Morales

## 1 Introducción

**E**n el presente capítulo se explica el desarrollo que se tuvo al elaborar el texto “Relación entre las motivaciones para abrir una microempresa y su eficiencia. El caso de México”, de Baez-Morales & Negrete-García (2019), que fue publicado en el libro *Problemas del desarrollo en México II*; esto con el fin de que el lector, y en particular los estudiantes de economía, tengan un referente más en la elaboración de una investigación, de manera especial en la serie de pasos prácticos que se llevaron a cabo para la elaboración del capítulo en mención.

Para tener una mayor contextualización del tema del capítulo libro, y de cualquier tema a investigar, lo primero que se debe hacer es la revisión de la literatura, para lo cual se deben leer artículos relacionados al tema de investigación. En la primera sección de este libro se hace una breve revisión de la conformación de un artículo científico. En el capítulo del libro analizado, por ejemplo, se profundiza en la importancia de los micronegocios en los países en desarrollo, con especial interés en México.

Dentro del capítulo analizado, el tema central gira en torno a la *eficiencia* en los micronegocios y cómo se ven afectados por las *motivaciones* que tienen los dueños cuando deciden abrir este tipo de unidades económicas. Es por ello que en el presente texto se explica brevemente la obtención de dichas variables y el manejo de las bases de datos en el capítulo analizado. Una vez que se tiene la idea general de la base de datos, se presenta la metodología usada por los autores, así como un breve repaso del paquete estadístico usado. Por último, se presentan algunos pasos prácticos en la presentación de los resultados.

## 2 Breve guía de la composición de un artículo científico

Cuando estamos interesados en investigar un tema en particular tenemos que revisar los artículos relacionados con el tema que nos servirán para la elaboración de la sección de *revisión de la literatura* de nuestro capítulo o artículo que estemos trabajando. Para tener el contexto del tema que estamos estudiando se hace una búsqueda exhaustiva de tesis, artículos, capítulos de libro, libros, etcétera, que tengan relación con nuestra investigación. A partir de la identificación del material necesario procedemos a identificar los aspectos más importantes y conocidos del tema, así como también los puntos desconocidos.

Ya que es una tarea primordial revisar los diferentes artículos científicos, el objetivo de esta sección es explicar cómo se conforman para tener una mejor comprensión. Para lograr este objetivo se toma a manera de ejemplo el artículo “Microenterprise origins and efficiency in Mexico”, de Otero et al. (2013), donde se estudia el papel de las motivaciones en la eficiencia de los micronegocios.

Al tener el artículo en mención en las manos podemos darnos cuenta de que en la primera página se nos proporciona un resumen (*abstract*, en inglés) de la investigación, donde se nos menciona el objetivo del estudio, de dónde se obtuvo la información, la metodología empleada, así como los principales resultados de la investigación. A continuación se reproduce el resumen del artículo de Otero et al. (2013):

This study focuses on the technical efficiency of microenterprises. We use data from Mexico’s National Survey of Microenterprises to estimate a stochastic frontier production model, and we find that the reason the owner started the business is significant in explaining variations in the levels of technical efficiency. Those who inherited the business or started it because of family tradition initially ran more efficient operations. These owners may have benefited from better knowledge and contacts. Over time, however, those who started their businesses to increase their income or, because they could not find a job, became more efficient. We argue that these owners benefited from greater motivation to be successful. Keywords: mi-

croenterprises; technical efficiency; Enamin; Mexico.

Como podemos observar, los autores se enfocan en la eficiencia técnica de los micronegocios en México. Hacen uso de la Encuesta Nacional de Micronegocios (Enamin) para la obtención de la base de datos. La metodología usada es la estimación del modelo de producción de fronteras estocásticas. Dentro de los principales resultados del estudio está que las motivaciones importan para explicar las variaciones en los niveles de la eficiencia técnica. De esta manera, podemos ver que esta parte nos da la información relevante de la investigación.

Continuando con la revisión del artículo, podemos encontrar la sección de Introducción, que es la que empieza propiamente el artículo. Su función principal es darnos una visión general o panorámica respecto a la totalidad del artículo. Es decir, es la sección que nos orienta de manera más amplia hacia dónde se quiere dirigir la investigación. De estas dos partes del artículo, el resumen y la *Introducción*, ya podemos tener una idea general si la investigación es de nuestro interés y, de ser así, continuar con la lectura.

En el artículo de Otero et al., inmediatamente después de la *Introducción* podemos encontrar la sección de *Revisión de la literatura*. En esta sección los autores exponen los puntos más relevantes hasta el momento de la elaboración del artículo sobre el tema que les interesa. De manera similar, el estudiante hará su propia sección de *Revisión de literatura* del tema que esté trabajando.

La sección que le procede es la de *Metodología*, donde Otero et al. plantea el modelo que usaron para perseguir los objetivos planteados en el trabajo. En este caso particular tenemos lo siguiente:

To study the role that motivation to start a business could have on technical efficiency, a stochastic frontier production function with inefficiency effects proposed by Battese and Coelli (1995) was used. Specifically, the translog production frontier for firm  $i$  is given by:

$$\ln Q_i = \beta_0 + \beta_1 \ln K_i + \beta_2 \ln L_i + \beta_3 \ln K_i^2 + \beta_4 \ln L_i^2 + \beta_5 \ln K_i \ln L_i + \delta X_i + v_i - u_i$$

The inefficiency effect ( $u_i$ ) in (1) above can be made a function of business motivations as well as other factors:

$$u_i = Z_i + w_i$$

Como podemos observar, lo autores explican el tipo de modelo usado, así como la referencia de dicho modelo. En este caso, explican las variables usadas y la forma en que están medidas cada una de ellas. Esta información es relevante, ya que nos proporciona la forma en que los

autores toman y miden las diferentes variables y, en su caso, nos proporciona la primera pista de qué información podemos necesitar en nuestra propia investigación.

Una vez que se nos ha explicado la sección de la *Metodología* le sigue la sección de *Resultados*. Esta última sección es precisamente dónde podemos observar con mayor detalle los resultados encontrados por los autores. En el caso del artículo mencionado, primero empiezan con una descripción de la Enamin. Por ejemplo, se nos dice a quiénes entrevistan en este tipo de encuestas y qué tipo de información podemos encontrar. Se menciona el número de micronegocios que fueron tomados en cuenta para el análisis, los sectores donde se desenvuelven y las motivaciones al abrir los negocios.

Se prosigue a revisar las estadísticas descriptivas de las variables utilizadas, para lo cual presentan el Cuadro 1, donde se puede ver que desagregan las variables, primero para toda la muestra de los micronegocios para proseguir con la de los sectores de servicios, manufactura y comercio. En estas estadísticas se nos proporcionan la media y la desviación estándar. La importancia del cuadro reside en que nos proporciona el primer panorama de las variables de importancia para el estudio. Por ejemplo, para la variable *años que lleva operando el negocio*, el promedio en toda la muestra fue de 8.3 años, mientras que para el sector de comercio fue de 6.5 años.

En un siguiente cuadro (ver Figura 1) nos presentan los resultados de estimar el modelo de fronteras estocásticas y las ineficiencias de los micronegocios. En el caso que nos presentan, son cuatro las especificaciones hechas.

Como se puede notar, los autores nos presentan sus resultados y explican el sentido de ellos, es decir, los efectos de las variables independientes sobre la ineficiencia. Por ejemplo, la variable de escolaridad tiene una relación negativa con la ineficiencia en todos los sectores. De manera similar, hacen la descripción de todas las variables del modelo.

Después de la sección de *Resultados*, la sección que prosigue es la de *Conclusiones*, que de manera general nos da información de los puntos clave de la investigación que llevaron a cabo los autores.

Figura 1: Ejemplo de estimaciones

Table 2. Stochastic Frontier Results with Years in Business Interaction Effects

Frontier	Full Sample		Services		Manufacturing		Commerce	
	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Constant	6.047***	0.091	6.244***	0.123	4.858***	0.210	6.322***	0.208
Ln Capital	0.318***	0.016	0.266***	0.024	0.422***	0.042	0.325***	0.028
Ln Labor	-0.291***	0.093	-0.074	0.127	0.468***	0.202	-0.864***	0.174
Ln Capital Squared	-0.011***	0.001	-0.008***	0.001	-0.014**	0.002	-0.013***	0.001
Ln Labor Squared	0.186***	0.044	0.186***	0.063	-0.052***	0.087	0.295***	0.093
Ln Capital x Labor	0.053***	0.009	0.126**	0.013	0.023	0.022	0.092***	0.017
Northern States (1=Yes; 0=No)	0.086*	0.045	0.114*	0.058	0.124	0.118	0.012	0.083
Central States (1=Yes; 0=No)	0.004	0.040	0.017	0.053	0.127	0.110	-0.069	0.071
Southern States(1=Yes; 0=No)	0.010	0.044	0.079	0.057	0.049	0.114	-0.088	0.079
Border States (1=Yes; 0=No)	0.356***	0.043	0.383***	0.055	0.566***	0.119	0.241***	0.077
Manufacturing (1=Yes; 0=No)	-0.331***	0.029						
Commerce (1=Yes; 0=No)	-0.213***	0.022						
<b>Technical Inefficient Effects</b>								
Constant	0.505	0.510	-0.417***	1.139	-5.094***	1.973	2.008***	0.327
Years of Schooling of Owner	-0.246***	0.062	-0.239**	0.077	-0.124***	0.037	-0.102***	0.017
Years in Business	-0.170***	0.049	-0.228***	0.092	0.002	0.091	-0.078	0.060
Registered with SHCP (1=Yes;0=No)	-1.191***	0.306	-1.713*	0.602	-8.026***	2.576	-0.201***	0.076
Outside Financing (1=Yes;0=No)	-0.219***	0.082	-0.243***	0.129	0.628***	0.239	-0.070	0.087
Family Tradition (1=Yes;0=No)	-3.420***	0.946	-3.438**	1.290	-4.674**	2.198	-0.724*	0.386
Increase Income (1=Yes;0=No)	-0.744***	0.282	-2.306***	1.001	0.944	0.638	-0.100	0.290
No Other Job (1=Yes;0=No)	-1.937***	0.590	-3.305	1.251	-2.831***	1.703	-0.353	0.312
Years in Business X Family Tradition	0.205***	0.062	0.086	0.074	0.198	0.137	0.082	0.062
Years in Business X Increase Income	0.063**	0.031	0.112*	0.062	-0.103	0.080	0.047	0.057
Years in Business X No Other Job	0.138***	0.049	0.159**	0.075	0.142	0.130	0.068	0.059
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	3.790***	0.836	5.568***	1.751	6.126***	1.807	1.200***	0.097
$\gamma = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2)$	0.845***	0.033	0.918***	0.025	0.900***	0.031	0.355***	0.095
Log Likelihood Function	-13,532		-6,769		-2,035		-4,561	
n	9,719		5,123		1,444		3,152	

Notes: \*, \*\*, and \*\*\* denote significance at 10%, 5%, and 1% levels, respectively.

### 3 Exploración de la base de datos

Una vez que se tiene un panorama de la composición de los artículos científicos, y suponiendo que se ha hecho la *revisión de la literatura* del tema que se está investigando, el siguiente paso será encontrar información sobre las variables que podrían ayudar al estudio. En el caso del capítulo de Baez-Morales & Negrete García (2019) se hace una exploración de las distintas bases de datos oficiales que hay en el país; en particular se indaga en las del Instituto Nacional de Estadística y Geografía (INEGI). Dentro de ellas se encuentran diferentes encuestas hechas a los negocios. Para el estudio, se toma la Encuesta Nacional de Micronegocios

(Enamin). Esta encuesta está disponible para los años de 1992, 1994, 1996, 1998, 2002, 2008, 2010 y 2012.

El estudio del capítulo del libro analizado se centra en la edición de la Enamin 2012. Vale la pena mencionar que cuando se tiene diferentes encuestas se debe tener especial atención sobre la comparabilidad de la información a través de los años, ya que, por ejemplo, pueden ser representativas a diferentes niveles de cantidad de población o que el cuestionario haya cambiado o ampliado las preguntas de una edición a otra.

Una vez que identificamos la edición de la encuesta a utilizar se procede a examinar el cuestionario. La Enamin proporciona un cuestionario básico donde se pueden identificar las preguntas contestadas por los dueños del negocios. El cuestionario contiene las secciones de condición de la actividad, antecedentes laborales, fundación del negocio, registros de contabilidad, información del local, la determinación de los precios, tiempo dedicado al negocio, afiliación, personal ocupado, equipamiento y gastos, pertenencia como asalariado, sector de actividad económica, ingresos derivados de la industria manufacturera, gastos derivados de la actividad manufacturera, ingresos derivados de negocios comerciales, gastos derivados de la compra de mercancía, ingresos derivados de la prestación de servicios o construcción, gastos derivados de la prestación de servicios o construcción, financiamiento, capacitación, programas de apoyo y expectativas.

Toda la información de las encuestas a los micronegocios puede encontrarse en la sección de microdatos de la página oficial de la Enamin. Hay que tener en cuenta que el sistema arrojará esta información en formato *DBF*; es útil saber que este tipo de formato puede ejecutarse en el programa *Excel*. Las diferentes secciones de la encuesta vienen en varios archivos; por ejemplo, el archivo *Ing\_Gast.dbf* solo contiene las preguntas relacionadas con la sección de ingreso y gastos de los micronegocios, por lo que de necesitar información de varias secciones se tendrá que trabajar para unir estos archivos. Una vez que tenemos los archivos, al abrirlos nos aparecerá el formato como en la Figura 2:

Hay que tener en cuenta que no podemos identificar a un micronegocio en particular, ni su localización exacta. Solo podemos acceder a la entidad federativa a la que pertenece. Por otra parte, en el caso de la Enamin 2012, se nos proporciona un archivo con el nombre de *enamin12\_des\_archivos.pdf* que describe cada una de las variables y esto nos ayudará a identificar cada columna del archivo. A manera de ejemplo, la columna *P55\_REN* nos hace referencia al número de renglón del trabajador. En

Figura 2: Ejemplo de base de datos

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	PER	T_LOC	ENT	CON	UPM	V_SEL	N_HOG	H_MUD	N_REN	P55_REN	P55_0	P55_1	P55_2	P55_3	P55_4	F
2	412	1	17	0143	00488	05	1	0	02	01	1	-	1	33	04	0
3	412	1	17	0148	00826	02	1	0	01	01	1	-	2	45	02	0
4	412	1	20	0004	00348	02	1	0	02	01	1	-	2	57	06	0
5	412	1	20	0004	00348	02	1	0	02	02	1	-	1	28	07	0
6	412	1	17	0126	00594	02	1	0	01	01	1	-	1	22	07	0
7	412	2	20	0012	01009	03	1	0	02	01	1	-	2	11	02	0
8	412	1	17	0130	00699	01	1	0	02	01	1	-	1	53	03	0
9	412	1	20	0014	00083	01	1	0	01	01	1	-	2	27	04	0
10	412	1	20	0014	00083	01	1	0	01	02	1	-	1	25	07	0

el caso de la columna *P55\_2*, nos da la información del sexo del dueño del micronegocio, y es aquí donde debemos tener cuidado acerca de qué representa cada número. En este caso, el número 1 es el hombre y el número 2 representa a la mujer. Una muestra de cómo es presentada la información se puede ver en la Figura 3.

Figura 3: Ejemplo de descripción de las variables

<i>P55_REN</i>	Carácter	2	01 a 15 ó 99	Número de renglón del trabajador
			01-15	Número del renglón
			99	No especificado
<i>P55_0</i>	Carácter	1	1, 2 o 3	Atención para entrevistador
<i>P55_1</i>	Carácter	50	Texto	Nombre del trabajador
<i>P55_2</i>	Carácter	1	1 ó 2	Sexo
			1	Hombre
			2	Mujer

En el caso particular de la investigación de Baez-Morales & Negrete García (2019), había interés en las variables de *ingresos*, *capital*, *trabajo* y las *motivaciones* por las cuales las personas deciden emprender un micronegocio. Si bien, se utilizaron más variables de control, como la zona geográfica, en este capítulo nos centraremos solo en las variables mencionadas.

Para el caso de la variable *ingreso*, la Enamin 2012 hace la pregunta expresa en cuanto a los montos, como se observa en la Figura 4.

En el caso de las variables *capital*, se tiene la pregunta expresa, como puede verse en la Figura 5. Para la variable *trabajo* (número de trabajadores), en el estudio se incluyó al dueño como parte de los trabajadores por la naturaleza de este tipo de negocios, como se ve en la Figura 6.

Figura 4: Ejemplo de pregunta sobre los ingresos

76. ¿Cuál es el monto de los ingresos que obtuvo el mes pasado por los siguientes conceptos?  
(Lee las opciones, circula y anota la cantidad indicada)

1 Por total de servicios (Incluye materiales y servicios de reparación y mantenimiento)	\$ _____
2 Por venta de mercancías	\$ _____
3 Otros ingresos	\$ _____
4 No generó ingresos el mes pasado	
5 Se negó a contestar	
0 Exclusivo capturista	

Figura 5: Ejemplo de pregunta sobre el capital

57. Si usted tuviera que vender las herramientas, utensilios, maquinaria, mobiliario, equipo, terreno, local y vehículos de su propiedad que utiliza en su negocio o actividad, ¿en cuánto los vendería?  
(Escucha y circula la opción indicada y en su caso anota la cantidad en pesos)

1	\$ _____
2	No tiene activos
3	Los activos con los que cuenta no tienen valor comercial
4	Son prestados
5	Se negó a contestar
9	No sabe

Figura 6: Ejemplo de pregunta sobre los trabajadores

8. De las personas que ocupa o le ayudan, ¿cuántas son...  
(Lee las opciones, circula las indicadas y anota el número de personas)

1	trabajadores que reciben un pago?	□ □
2	socios?	□ □
3	trabajadores sin pago?	□ □
4	No sabe qué tipo de trabajadores son	□ □
	Total de trabajadores	□ □
0	Exclusivo capturista	

Como se puede notar, la cantidad de *ingresos* generados por los micro-negocios podemos obtenerla de la pregunta 75. Para tener una estimación anual se multiplicó el ingreso mensual por los doce meses del año. En el caso de la variable *capital*, podemos obtenerla de la pregunta 57;

Figura 7: Ejemplo de pregunta sobre las motivaciones

**29. ¿Cuál es el motivo principal por el que se inició en este negocio o actividad?**

*(Escucha, anota la respuesta y después clasificala en la opción indicada)*

- 
- 01 Por tradición familiar o lo heredó
  - 02 Para complementar el ingreso familiar
  - 03 Para mejorar el ingreso
  - 04 Tenía dinero y encontró una buena oportunidad
  - 05 Para ejercer su oficio, carrera o profesión
  - 06 Fue la única manera que tuvo para obtener un ingreso
  - 07 No tenía la experiencia requerida para un empleo
  - 08 No tenía la escolaridad o capacitación requerida para un empleo
  - 09 Estaba sobrecapacitado para un empleo
  - 10 Los empleos que encontró estaban mal pagados
  - 11 Requería un horario flexible
  - 12 No había oportunidades de empleo
  - 13 Otra razón

en este caso, tomamos el valor que le están dando a todos sus activos. Por último, la variable *trabajo* se obtuvo de la pregunta 8, que toma en cuenta los trabajadores que colaboran en el negocio, ya sea con pago o sin él.

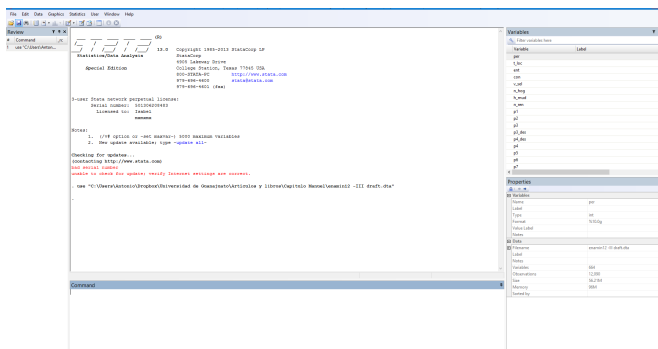
En el caso de las variables de interés para el estudio, en particular, las *motivaciones*, se tomó la pregunta 29, donde se pide explícitamente el motivo por el cual se inició el micronegocio, como se puede ver en la Figura 7. Al tener diferentes opciones como respuestas, los autores deciden agrupar los motivos en dos grandes grupos. El primero agrupó a las motivaciones asociadas al *emprededurismo*, es decir, por iniciativa de las personas y no por presiones económicas del momento. A este grupo se le llamó *motivopro*, que incluye los motivos uno, cuatro, cinco y nueve.

Hasta aquí se ha visto la forma como se identificaron las variables de interés en el cuestionario de la Enamin 2012 y de dónde se obtuvo la base de datos. Sin embargo, los estudiantes pueden encontrarse con otro problema en la práctica: cómo manipular la base de datos, y en particular, cómo hacerlo desde el uso de un paquete de software estadístico. En el caso del capítulo de libro en mención, se usó el programa *Stata*; este

programa maneja archivos de datos *DTA*, por lo que hay que hacer una conversión de datos del formato *Excel* a *Stata*.

A continuación se hace una breve exploración del uso del software para la importación de la base de datos, pero antes debemos familiarizarnos con el programa *Stata*. En la Figura 8 podemos ver la pantalla de inicio de *Stata*. En la parte superior derecha encontraremos las pestañas de *file*, *edit data*, *graphics*, *statistics*, *user*, *window*, *help*. Tenemos dos ventanas centrales: en la ventana central superior podemos ver la información del software, donde se desplegarán los resultados de las operaciones o comandos que le hayamos ordenado a *Stata*. En la ventana central inferior con el encabezado *comandos* es donde le daremos las órdenes o comandos al programa.

Figura 8: Pantalla principal del programa Stata



En la ventana de la izquierda que tiene como encabezado *review*, podemos visualizar todas las órdenes o comandos que hayamos hecho; es similar a un historial de las órdenes hechas durante la sesión. En la ventana superior de la derecha con el encabezado *variables* es donde podemos visualizar las variables que tiene la base de datos, así como la etiqueta explicativa de cada una de ellas. Por último, la ventana inferior izquierda con el encabezado *properties* nos da la información de las variables usadas tales como el nombre, etiqueta, tipo de variables, formato de las variables, valor de la etiqueta. En esta misma ventana podemos visualizar también nombre del archivo, etiqueta, notas, el número de variables, el número de observaciones, el tamaño del archivo, la memoria usada.

Figura 9: Ventana para importar datos en Stata

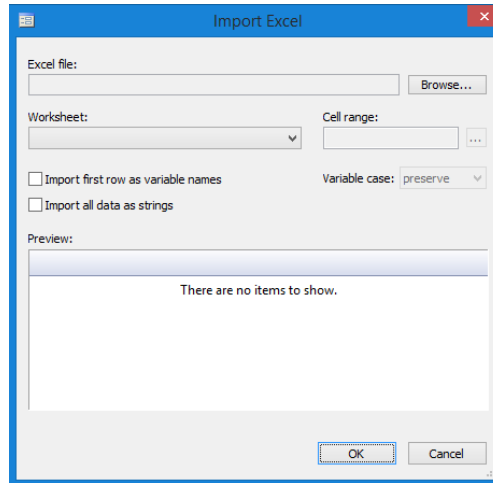


Figura 10: Visualización de datos en Stata

p58_4_3	p58_4_4	p58_4_c	p58_5_1	p58_5_2	p58_5_3
BIMESTRAL	10800	6	1	250	MENSUAL
BIMESTRAL	1200	6	2	0	
BIMESTRAL	5400	6	1	200	SEMANAL
BIMESTRAL	1200	6	2	0	

Vale la pena mencionar que una vez que conocemos la función de cada una de las ventanas, la de *properties* es muy útil para saber si la variable seleccionada en la ventana de *variables* es una variable *cadena* (*string*) o una variable numérica. Las variables *string* se refieren a cadenas de caracteres tales como “Hombre” o “Mujer”, “Sí” o “No”, “E. Gómez” o “A. Velasco”. En el caso de las variables numéricas estaríamos hablando de números como 0,1, 2, 5.7, y así sucesivamente. Esto es de importancia, ya que al hacer operaciones con las bases de datos vamos a necesitar que las variables estén en formato numérico.

Una vez que estamos familiarizados con la ventana de presentación de *Stata* podemos hacer la conversión de las bases de datos, importándolas. En el caso del programa *Stata* podemos ir a *File >>Import >>Excel*

*spreadsheet*. Una vez que entramos en esta opción nos arroja una ventana (ver Figura 9) donde podemos explorar en nuestro directorio para encontrar nuestro archivo de interés e importarlo.

Cuando tengamos nuestra información cargada podremos ver nuestra base de datos. En las pestañas que se encuentran en la parte superior izquierda podemos visualizar las variables en su conjunto al apretar en el botón *Data Editor (Browse)*, en donde las variables que sean *string* saldrán en color rojo y las variables numéricas en color negro, como se muestra en la Figura 10.

## 4 Acerca de la metodología

### 4.1 Modelos de fronteras estocásticas

Una vez que tenemos nuestra base de datos y estamos familiarizados con el programa estadístico, los autores exponen la metodología, en particular la de modelos de fronteras estocásticas. Este tipo de modelos parten de una frontera eficiente de producción  $f(x)$  que define la máxima cantidad del producto que una unidad económica —en el caso del capítulo de libro analizado, los micronegocios— pueden producir a partir de un conjunto de insumos  $x$ . La (in)eficiencia técnica de las empresas correspondería a las diferencias entre un máximo que se daría teóricamente y lo que pasa en la realidad, es decir, lo que produce los micronegocios con los insumos.

Para tener un mejor entendimiento podemos hacer una breve explicación de la naturaleza de las fronteras estocásticas. Supondremos que un productor tiene una función de producción  $f(z_i, \beta)$ . Suponiendo que no hay error o ineficiencia, el  $i$ -ésimo micronegocio produciría:

$$q_i = f(z_i, \beta)$$

El análisis de fronteras estocásticas asume que cada firma produce potencialmente menos de lo que podría debido a un grado de ineficiencia. Tenemos que:

$$q_i = f(z_i, \beta)\xi_i$$

Donde  $\xi_i$  es el nivel de eficiencia del micronegocio;  $\xi_i$  debe estar en un intervalo de  $(0, 1]$ . Si  $\xi_i = 1$ , el micronegocio está logrando la producción óptima con la tecnología que conlleva la función de producción  $f(z_i, \beta)$ . Cuando  $\xi_i < 1$ , el micronegocio no estaría aprovechando al

máximo los insumos  $z_i$  dada la tecnología que conlleva la función de producción  $f(z_i, \beta)$ . Ya que se asume que la producción es estrictamente positiva (es decir que,  $q_i > 0$ ), el grado de eficiencia técnica es asumido estrictamente positivo, lo que es  $\xi_i > 0$ .

También se puede asumir que la producción esté sujeta a shocks aleatorios, implicando que:

$$q_i = f(z_i; \beta)\xi_i \exp(\nu_i)$$

Tomando los logaritmos naturales de ambos lados, tenemos:

$$\ln(q_i) = \ln(f(z_i; \beta)) + \ln(\xi_i) + \nu_i$$

Asumiendo que hay  $k$  insumos y la función de producción es lineal en logaritmos, definiendo  $u_i = -\ln(\xi_i)$  tenemos:

$$\ln(q_i) = \beta_0 + \sum_{j=1}^k \beta_j \ln(z_{ij}) + \nu_i - u_i$$

De tal manera que  $u_i$  es sustraída de  $\ln(q_i)$ , restringiendo  $u_i > 0$  implica que  $0 < \xi_i < 1$ .

Para la ejecución de este tipo de modelos en el programa *Stata* se proporciona el comando *frontier*. De hecho, el comando va con la forma:

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \nu_i - u_i$$

En el contexto de los micronegocios y en el capítulo del libro estudiado, el modelo que expresado de la siguiente manera:

$$Y_i = \beta_0 + \beta_1 K_i + \beta_2 L_i + \delta X_i + \varepsilon_i$$

$$\varepsilon_i = \nu_i + \mu_i$$

$$E_i = \theta' Z_i + w_i$$

En la primera ecuación podemos ver la variable  $Y_i$  que representa la producción en términos logarítmicos de la microempresa  $i$ . La variable  $K_i$  es el logaritmo del capital físico. Y la variable  $L_i$  es el trabajo, expresado mediante el logaritmo del número de trabajadores. La variable  $X_i$  es un vector de variables de control que incluye la región geográfica y los sectores de la economía.

Hay que recordar que ya se ha visto de donde se ha sacado la información de la *producción* que en el capítulo analizado son los *ingresos* de los micronegocios. El *capital* que es el valor de todos sus activos y el *trabajo* que son el dueño y los trabajadores que se tengan.

En la segunda ecuación nos muestra el componente del error  $\varepsilon_i$ , que es la suma o diferencia del disturbio normalmente distribuido  $v_i$ , el cual es una especificación del error, y del disturbio,  $\mu_i$ , el cual representa la eficiencia no negativa asumida como aleatoria y con una media normal, normal truncada y exponencial. Los términos  $v_i$  y  $\mu_i$  se asumen como independientes de cada uno y con respecto a las observaciones. Este es precisamente el modelo teórico explicado líneas arriba.

De manera práctica, una vez identificadas las preguntas asociadas a las variables de *ingreso*, *capital* y *trabajo* se hace la conversión a logaritmos como lo explica el modelo teórico. Un ejemplo de los comandos en el programa *Stata* para hacer esta transformación es la siguiente:

```
gen lny = log(y)
gen lnI = log(I)
gen lnk = log(k)
```

En el caso de las variables que identifiquen la *región geográfica*, la encuesta nos proporciona la entidad federativa en donde se localiza el micronegocio, por lo que podemos crear variables *dummy* para las diferentes regiones que existen en el país. Para el caso de los *sectores* también la encuesta nos proporciona en que sector se ubican; de esta manera, también se crean variables *dummy* para los sectores de la economía. Un ejemplo de la generación de estas últimas lo podemos ver a continuación:

```
gen manufac = (p66_t_1!=.)
gen commerc = (p71_t_1!=.)
gen servic = (p76_t_1!=.)
```

Continuando, una vez que se ha obtenido el modelo de fronteras estocásticas se estima la medida de eficiencia  $\mu_i$  y se renombra como  $E_i$  con el fin de distinguirlas. Tras estimar la eficiencia se corre una regresión lineal, donde la variable dependiente es la eficiencia y el vector  $Z_i$  agrupa las variables de control mencionadas en la sección anterior; es decir, las características de la empresa y las variables *dummies* de los motivos por los que se iniciaron los negocios. En otra regresión lineal se hace

el mismo procedimiento, pero en lugar de incorporar todos los motivos, como variables independientes, solo incluimos dos variables *dummies*, correspondientes a las dos clasificaciones *motivopro* y *motivonopro*.

En el caso del programa de *Stata* como ya se mencionó anteriormente, el comando *frontier* es el que nos ayuda a estimar el modelo de fronteras estocásticas. Un ejemplo de cómo ordenarle al programa puede quedar de la siguiente manera:

```
frontier lny lnk lnl north center south northeast northwest west east north-  
central southcentral southeast southwest manufacture commerce services  
predict Inefficiency, u  
predict Efficiency, te
```

Observemos que en las líneas anteriores el comando *frontier* se pone al inicio, ya que es la forma de ordenar al programa *Stata* que haga las operaciones sobre las variables que le especificamos. Seguido del comando, tenemos el logaritmo del ingreso (*lny*); poniéndolo en esta posición, es decir, seguido del comando *frontier*, le estamos indicando al programa que esta es la variable dependiente. Con la posición del resto de las variables le estaríamos indicando que son las variables independientes. El comando *predict* nos arrojaría la medida de eficiencia o ineficiencia de los micronegocios explicado en la teoría anteriormente.

Una vez que tenemos la medida de *eficiencia* de los micronegocios podemos ver las estadísticas básicas de esta variable. Podemos visualizar, por ejemplo, la eficiencia por número de *motivo* o *laeficiencia* por grupos, que en el caso del capítulo se proponen dos, ya explicados, como son *motivopro* y *motivonopro*. Un comando que nos sirve para visualizar este tipo de estadística es *sum*, que nos permite tener la siguiente información, como en la Figura 11.

## 4.2 Método de descomposición Oaxaca - Blinder

Una vez que tenemos la medida de eficiencia, en el capítulo de libro se propone hacer una descomposición, en específico la de Oaxaca-Blinder (1973). Esto con el fin de comparar a los dos grupos ya mencionados anteriormente (*motivopro* y *motivonopro*). Este tipo de composición es ampliamente usada para estudiar las diferencias en las medias entre grupos tales como las brechas salariales por sexo o raza.

Este tipo de metodología descompone las diferencias de las medias de los logaritmos, por ejemplo, de los salarios basados en modelos de regre-

Figura 11: Visualización de una variable con el comando *sum* en Stata.

```
. sum Efficiency
```

Variable	Obs	Mean	Std. Dev.	Min	Max
Efficiency	12053	.5074937	.1423551	.0314464	.86573

```
.
Command
sum Efficiency
```

sión de manera contrafactual. En el ejemplo de los salarios, esta metodología divide el diferencial salarial entre dos grupos en una parte que es “explicada” por las diferencias de grupo en las características de productividad tales como la educación o la experiencia laboral. La otra parte es un residual que no forma parte de tales diferencias en los determinantes del salario. Esta parte “inexplicada” es usualmente usada como una medida de discriminación, pero también sería parte de las diferencias de grupo de predictores inobservados.

De manera breve explicaremos en que consiste el método de Oaxaca-Blinder. Tenemos dos grupos, A y B, y una variable de resultado  $Y$ , así como una serie de predictores. Si pensamos en hombres y mujeres, el logaritmo de los salarios sería la variable resultado, y los indicadores de capital humano, tales como la educación o experiencia laboral, serían los predictores. La pregunta reside en cuánto es la diferencia media del resultado:

$$R = E(Y_A) - E(Y_B)$$

Donde  $E(Y)$  es el valor esperado de la variable resultado y estaría explicado por las diferencias de grupo en los predictores.

Basado en un modelo lineal:

$$Y_l = X_l' \beta_l + \epsilon_l, \quad E(\epsilon_l) = 0, \quad l \in \{A, B\}$$

Donde  $X$  es un vector que contiene los predictores y una constante,  $\beta$  contiene la pendiente de los parámetros y el intercepto, y  $\epsilon$  es el error, la diferencia media del resultado puede ser expresado como una diferencia en la predicción línea en la media del grupo específico de los regresores.

$$R = E(Y_A) - E(Y_B) = E(Y_A)' \beta_A - E(Y_B)' \beta_B$$

Donde

$$E(Y_i) = E(X_i' \beta_i + \epsilon_i) = E(X_i' \beta_i) + E(\epsilon_i) = E(X_i' \beta_i)$$

Con  $E(\beta_i) = \beta_i$  y  $E(\epsilon) = 0$  como supuesto.

Para identificar la contribución de las diferencias de grupo en los predictores respecto a la diferencia del resultado general, la ecuación anterior puede ser arreglada como sigue:

$$R = [E(X_A) - E(X_B)]' \beta_B + E(X_B)' (\beta_A - \beta_B) + [E(X_A) - E(X_B)]' (\beta_A - \beta_B)$$

Esta es la descomposición en tres partes, esto es, la diferencia en el resultado es dividida en tres partes:

$$R = E + C + I$$

En el caso del primer sumando

$$E = [E(X_A) - E(X_B)]' \beta_B$$

Esta parte captura las diferencias en los predictores entre los grupos. A esto se le conoce con el nombre de *efecto dotación* o *endowment effects*. Con este mismo nombre lo presenta el programa de *Stata*.

El segundo componente es:

$$C = E(X_B)' (\beta_A - \beta_B)$$

El cual mide la contribución de las diferencias en los coeficientes el cual incluye las diferencias en el intercepto.

El tercer sumando:

$$I = [E(X_A) - E(X_B)]' (\beta_A - \beta_B)$$

Es un término de interacción que toma en cuenta el hecho de las diferencias en las dotaciones y coeficientes existen simultáneamente entre los dos grupos.

En el caso práctico del capítulo analizado, se expresó de la siguiente manera:

$$\bar{E}_{pro,i} = \bar{Z}_{pro,i}\beta_{pro,i} + u_{pro,i}$$

$$\bar{E}_{nopro,i} = \bar{Z}_{nopro,i}\beta_{nopro,i} + u_{pro,i}$$

$$\bar{E}_{pro,i} - \bar{E}_{nopro,i} = (\bar{Z}_{pro,i} - \bar{Z}_{nopro,i})\beta_{pro,i} + \bar{Z}_{nopro,i}(\beta_{pro,i} - \beta_{nopro,i})$$

Como se explicó líneas arriba, la  $E$  es la eficiencia de los micronegocios, los cuales fueron divididos en dos grandes grupos, *motivepro* y *motivonopro*. El grupo de empresas con *motivopro* ( $E_{pro}$ ) y el grupo *motivonopro* ( $E_{nopro}$ ) están expresados en las ecuaciones anteriores. Los vectores  $Z_{pro}$  y  $Z_{nopro}$  representan las características de las microempresas o las dotaciones,  $\beta$  expresa los retornos de las dotaciones, y  $u_i$  es el termino de error.

Seguido de la ecuación de cada grupo, se hace la descomposición por medio de la diferencia de la eficiencia entre grupos. De tal manera que la primera parte se refiere a las diferencias de las *características* entre grupos que tienen los mismos retornos de las dotaciones de las microempresas con *motivos proactivos*. La segunda parte se refiere a la diferencia entre los *retornos de las dotaciones* cuando las dotaciones que tienen son las mismas.

Como se puede notar, se está haciendo el mismo procedimiento que en la explicación teórica del modelo, pero que tenía como variable resultado o dependiente a los *ingresos*. En nuestro ejemplo, la variable resultado es la *eficiencia* de los micronegocios. En el modelo teórico tenemos dos grupos A y B, que pueden ser hombres y mujeres, en nuestro ejemplo serían el grupo de empresas ligadas al *emprendedurismo* (*motivopro*), y las que no están ligadas a este grupo (*motivonopro*).

Para poder hacer los cálculos de acuerdo a esta metodología, el comando *oaxaca* nos permite hacer las operaciones en *Stata*, como se puede ver a continuación:

```
Oaxaca eficiencia edad edad2 años años2 educ_2 educ_3 educ_4 educ_5
sexo casado, by (motivepro) weight (1)
```

Como podemos notar, en primer lugar tenemos que poner el comando para indicarle al programa que vamos a utilizar el método de descomposición Oaxaca-Blinder (1973). En seguida tenemos que poner la variable resultado, que en nuestro caso es la *eficiencia*. Inmediatamente después ponemos todas las variables independientes que podrían explicar las diferencias de los grupos. En la parte final del comando tenemos

que especificar el grupo de referencia a comparar; en el caso analizado *fuemotivopro*, como se puede notar arriba.

## 5 Presentación de los resultados

Una vez que tenemos las bases de datos y hemos empleado la metodología propuesta, en el capítulo del libro se procede a la presentación de resultados. Se empieza esta sección con las presentaciones de algunas estadísticas básicas que versan sobre las variables de interés, que en este caso son las motivaciones por las que abren los micronegocios los dueños de los negocios. También se muestran los porcentajes que ocupa cada motivo del total de las respuestas. *Stata* nos permite obtener la información por medio del comando *tab* seguido de la variable de interés. Un ejemplo se muestra en la Figura 12.

Figura 12: Visualización de una variable con el comando *tab* en *Stata*.

```
. tab motive1
```

motive1	Freq.	Percent	Cum.
0	11,258	93.40	93.40
1	795	6.60	100.00
Total	12,053	100.00	

```
.  
Command  
tab motive1
```

Como podemos observar, la variable *motive1* es una *variable dummy* que toma el valor de 0 si es cualquier otro motivo, y toma el valor de 1 si es el *motivo número uno*. Entonces podemos ver que, de 12,053 observaciones, 795 dueños de micronegocios respondieron que el *motive1* era la razón por la cual decidieron abrir su negocio y representan el 6.60% del total de la muestra.

Esta información se presenta para el resto de las motivaciones e inmediatamente después en un gráfico, donde se muestra el número que concentra cada una de las motivaciones. Este tipo de información se pue-

de poner en la parte introductoria de esta sección, dándole al lector un preámbulo de los resultados.

Los autores presentan después las estadísticas descriptivas, centrándose en la variable de interés, que es la *eficiencia* por el número *demotivo*. Para obtener esta información se puede hacer uso del comando *sum*, ya visto anteriormente, pero ahora se agrega el condicional *if* (si) para que el programa *dStata* solo seleccione la información del nivel de *eficiencia* de acuerdo con el *motivo* especificado. Un ejemplo lo podemos ver en la Figura 13.

Figura 13: Visualización de una variable con el comando *sum* y la condicional *if* en *Stata*

```
. sum Efficiency if motive1==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
Efficiency	795	.5247312	.1440241	.0579595	.8201813

```
.  
Command  
sum Efficiency if motive1==1
```

Como se puede apreciar, la forma específica del comando *sum eif* (si) nos permite visualizar la *eficiencia*, el número de observaciones, la media, desviación estándar, el dato mínimo y máximo del motivo que le estemos especificando. Obteniendo la información por cada motivo se puede construir una tabla para poder comparar las eficiencias por motivos y dar al lector un panorama práctico que resume la información.

En el capítulo analizado se procede a mostrar dos regresiones lineales, las cuales se llaman *Modelo 1* y *Modelo 2*. En cada uno de los modelos la variable dependiente es la *eficiencia*, y las variables independientes, en el caso del *Modelo 1*, son cada uno de los motivos que nos proporciona la encuesta sumado a las variables de control como son la *edad del dueño del negocio*, esta misma variable al cuadrado, los *años que lleva activo el negocio*, esta misma variable, pero al cuadrado, los diferentes *niveles educativos*, el *sexo* y si se está *casado*. En el *Modelo 2* se presentan las mismas variables de control, solo que con la variable *dummy* que agrupa los motivos por *motivopro*, para ver si hay significancias estadísticas.

Se puede ordenar al programa *Stata* el hacer las regresiones del párrafo anterior de la siguiente manera:

```
reg Efficiency edad edad2 años años2 educ_2 educ_3 educ_4 educ_5 sexo
casado motive1 motive2 motive3 motive4 motive5 motive7 motive8 motive9
motive10 motive11 motive12 motive13
```

```
outreg2 using "C:\Users\Mis documentos", e (N, r2) replace ctitle (mo-
tives)
```

```
reg Efficiency edad edad2 años años2 educ_2 educ_3 educ_4 educ_5 sexo
casado motivepro motivenopro
```

```
outreg2 using "C:\Users\Mis documentos", e (N, r2) word append ctitle
(proactives)
```

Como se puede notar, se le está indicado al programa *Stata* que haga la regresión lineal por medio del comando *reg*, en seguida se pone la variable dependiente, que en caso de estudio es la *eficiencia*, seguido de las variables independientes que han sido ya explicadas. En los comandos arriba el primero se refiere al *Modelo 1*. En un segundo orden, en *Stata* aparece el comando *outreg2* que simplemente nos va a permitir exportar los resultados en forma de artículo científico. En un tercer orden se muestra la regresión para el *Modelo 2*. Y en un cuarto orden, de nuevo la exportación de los resultados. Una muestra de ellos la podemos ver a continuación en el Cuadro 1.

Cuadro 1: Muestra de los resultados: regresión lineal

Variables	Modelo 1	Modelo 2
Edad	0.0025*** (0.0005)	0.0025*** (0.0005)
Edad2	-3.91e-05*** (5.99e-06)	-3.96e-05*** (6.00e-06)
Edad_negocio	0.0013*** (0.0002)	0.0014*** (0.0003)
Edad_negocio2	-2.49e-05*** (6.68e-06)	-2.58e-05*** (6.67e-06)
Educ_primaria	0.0205*** (0.0068)	0.0211*** (0.0068)
Educ_secundaria	0.0253*** (0.0071)	0.0264*** (0.0071)
Educ_preparatoria	0.0326*** (0.0074)	0.0340*** (0.0074)
Educ_superior	0.0471***	0.0488***

	(0.0071)	(0.0071)
Mujer	-0.0470***	-0.0521***
	(0.0028)	(0.0026)
Casado	-0.0053*	-0.00696**
	(0.0029)	(0.0029)
Motivo_1	0.0094	
	(0.0064)	
Motivo_2	-0.0131***	
	(0.0049)	
Motivo_3	0.0088*	
	(0.0047)	
Motivo_4	0.0135**	
	(0.0065)	
Motivo_5	0.0095*	
	(0.0052)	
Motivo_7	-0.0058	
	(0.0296)	
Motivo_8	0.0165	
	(0.0236)	
Motivo_9	0.0739*	
	(0.0417)	
Motivo_10	0.0110	
	(0.0123)	
Motivo_11	0.0097	
	(0.0100)	
Motivo_12	0.00075	
	(0.0068)	
Motivo_13	-0.0016	
	(0.0055)	
Motivopro		0.0114**
		(0.00468)
Motivonopro		0.00158
		(0.00427)
Constante	0.510***	0.515***
	(0.0150)	(0.0150)
N	12053	12053

---

Una vez que se han presentado las tablas se procede hacer la descripción y análisis de los resultados. En el caso de la tabla anterior, las variables de interés para el estudio son por una parte la *eficiencia* como variable a explicar, y las variables de *motivaciones* como las variables de interés en explicar la *eficiencia*. De acuerdo con los objetivos planteados en el trabajo, se centró en la relación que tiene cada una de las *motivaciones*, es decir, en el efecto que tienen sobre la *eficiencia* de los micronegocios. Como podemos ver en la tabla de arriba, en el *Modelo 1* las motivaciones del 2 al 4 tienen un efecto significativo estadísticamente. A manera de ejemplo, el *motivo 2* “para complementar el ingreso familiar” tiene un efecto negativo sobre la eficiencia de los micronegocios; para hacer una interpretación más detallada debemos tener cuidado en la manera en que están medidas las variables para poder expresarla.

Por último, los autores presentan la tabla donde se nos proporcionan los resultados del método de descomposición de Oaxaca-Blinder (1973), como se muestra en el Cuadro 2.

Cuadro 2: Muestra de los resultados: Oaxaca-Blinder

Variables	Eficiencia	Dotaciones	Retornos
Motivopro	44.89*** (0.139)		
Motivonopro	43.95*** (0.231)		
Diferencia	0.945*** (0.270)		
Dotaciones	0.865*** (0.264)		
Retornos	0.0805 (0.0500)		
Edad2		0.875*** (0.262)	0.141 (0.0952)
Edad_negocio		-0.662*** (0.0450)	0.276* (0.146)
Edad_negocio2		0.616*** (0.0496)	-0.516*** (0.0814)
Educ_primaria		0.0977*** (0.0146)	0.189*** (0.0690)
Educ_secundaria		0.0404*** (0.0101)	0.206*** (0.0719)

Educ_preparatoria		-0.00693	0.0798
		(0.00556)	(0.0534)
Educ_superior		-0.165***	0.257**
		(0.0218)	(0.106)
Mujer		0.0436***	0.221*
		(0.00915)	(0.131)
Casado		0.0268***	-0.204***
		(0.00742)	(0.0738)
Constante			-0.569
			(0.363)
Observaciones	12,053	12,053	12,053
N_1	8853	8853	8853
N_2	3200	3200	3200

Tomando en cuanto lo expuesto en la sección de metodología, podemos observar que *Stata* nos arroja la tabla, donde podemos notar la media de la *eficiencia* por grupo, que en el caso analizado es *motivopro* y *motivenopro*. En seguida podemos ver la diferencia que existe entre ellos. La parte de *endowment effect* la podemos notar en la parte de la tabla que se llama *dotaciones*. La parte de las diferencias en los coeficientes la podemos ver en la tabla por medio de la notación *retornos*. En la parte de *dotaciones* nos estaría indicando cuánto de las diferencias entre grupos se debe a las que hay entre sus características, como el nivel de educación o los años del negocio en el mercado. La segunda parte cuantifica los cambios en la eficiencia del grupo de referencia (*motivepro*) cuando aplicamos los coeficientes del grupo *motivonopro* a las características del grupo *motivopro*; esta segunda parte también es interpretada como la forma en que se usan las dotaciones. Y como puede observarse, cada una de las partes viene desglosada por las variables que la componen, así como su efecto y significancia estadística.

Por último, en el trabajo se presentan las conclusiones y se sintetizan los puntos clave de la investigación, así como recomendaciones en nuevas áreas de estudio para futuras investigaciones.

## 6 Conclusiones

En este texto se ha revisado de manera práctica la elaboración del capítulo de libro “Relación entre las motivaciones para abrir una micro-empresa y su eficiencia. El caso de México”, de Baez-Morales y Negrete

García (2019). Se ha empezado exponiendo la composición de un artículo científico con el fin de saber identificar las partes que lo componen. Por otra parte, se abundó en la metodología utilizada en el capítulo analizado, pasando por la búsqueda de bases de datos, paquete estadístico, comandos y la intuición detrás de cada paso. Por último, se hizo una presentación de los resultados y una breve interpretación de estos.

Este artículo también pretendió mostrar de manera práctica a los estudiantes de economía la elaboración de un capítulo de libro basado en un trabajo de investigación que ha sido publicado. Sin embargo, se tiene que aclarar que se han omitido muchos pasos en cada una de las secciones que darían para muchas más hojas de explicación. No obstante, se espera que este capítulo sea solo una referencia o apoyo en los primeros pasos dentro del trabajo de investigación de los estudiantes.

## Referencias

- [1] Baez-Morales y Negrete García (2019) “Relación entre las motivaciones para abrir una microempresa y su eficiencia. El caso de México”. Problemas del desarrollo en México II. Grañén Porrúa.
- [2] Otero, María, y Elisabeth Rhyne. (1994) “The World of Micro-Enterprise Finance: Building Healthy Financial Institutions for the Poor” Kumarian Press, Bloomfield.
- [3] Oaxaca, Ronald. (1973) “Male-Female Wage Differentials in Urban Labor Markets” *International Economic Review* . 14:693.



# Macroeconomía Dinámica Aplicada: Cambio Estructural en México

Édgar Cruz

Fernando García Barragán

## 1 Introducción

Durante proceso del desarrollo y crecimiento económico, los países experimentan importantes transformaciones económicas. Por ejemplo, la tasa de fertilidad, definida como el número de hijos promedio disminuyen en la medida que la riqueza por persona se incrementa en el país, mientras que la tasa de participación laboral de las mujeres y los niveles de escolaridad de la población se incrementan. Otro proceso de transformación económica, que es común entre países, es el cambio estructural.

El concepto de cambio estructural hace referencia a la relocalización de la actividad económica a través de los sectores productivos. En particular, la relocalización de la actividad económica entre los tres sectores agregados que componen la economía, a saber: el sector agrícola, industrial y de los servicios. La relocalización de la actividad económica implica que la distribución de los factores productivos como el capital y el trabajo entre sectores cambia durante el proceso de desarrollo. Interesantemente, esta continua relocalización de los factores productivos es un hecho que se observa en la mayoría de los países y tiene importantes implicaciones para otras dimensiones del desarrollo y el crecimiento.

El objetivo de este capítulo es introducir al lector en la literatura económica sobre el cambio estructural presentando algunos hechos estilizados sobre el cambio estructural a través de dos dimensiones, empleo y valor agregado; las teorías que explican estos hechos; y la presentación y solución de un modelo de equilibrio general que engloba las principales explicaciones sobre el cambio estructural. Con el propósito de reforzar

el aprendizaje a través de métodos aplicados, este capítulo presenta la solución analítica y numérica (simulación computacional) del modelo propuesto mediante el uso del paquete computacional Dynare.

## 2 Hechos sobre el cambio estructural

En la literatura existen dos medidas sobre el cambio en la composición sectorial. La primera medida está basada en la contribución que hace cada sector al Producto Interno Bruto (PIB). En particular, la medida de cambio estructural es los cambios en el peso relativo (share) que tiene el valor de la producción de los sectores agrícola, industrial y de los servicios en el PIB.

El gráfico en la Figura 1, primera fila, muestra la relación entre el peso relativo de la agrícola, industrial y de servicios en el PIB y el logaritmo del valor del PIB per cápita en dólares y a precios constantes para varios países durante el periodo 1970-2018. Para realizar el gráfico los datos de los países para todos los años son combinados y controlamos las diferencias en los niveles de los shares entre países efectuando una regresión entre la composición sectorial (shares) y un polinomio de orden bajo del logaritmo del PIB per cápita y country fixed effects.<sup>1</sup> La inspección visual del gráfico muestra que existe un claro patrón en entre el nivel de desarrollo de un país y la composición sectorial de la economía. En niveles bajos de desarrollo, observamos que la composición sectorial del PIB es primordialmente definida por el sector agrícola. Por ejemplo, en países donde el PIB es mayor a de 560 dólares (6 en logaritmo) observamos el peso del sector agrícola en el PIB es alrededor de 60 % (por ejemplo, Afganistán). En contraste, cuando el nivel de desarrollo es elevado, observamos que la composición sectorial del PIB es definida primordialmente por el sector de servicios.

[Insertar Figura 1]

La segunda medida está basada en la contribución que hace cada sector al empleo total. En particular, la medida de cambio estructural es el peso relativo o share que tiene el empleo localizado en los sectores agrícola,

---

<sup>1</sup> Al controlar por características fijas de los países, los gráficos en la Figura 1 muestran el share o peso relativo de cada sector que corresponde al nivel de desarrollo considerando que existen otros elementos o variables, que son constantes en cada país, que pueden explicar las diferencias de los sectores entre países, por ejemplo, factores institucionales o geográficos.

industrial y de los servicios respecto al empleo total. El gráfico en la Figura 1, segunda fila, muestra la relación entre empleo sectorial y el nivel de desarrollo. En este caso, la composición del sectorial asociada a cada nivel de desarrollo. Con base en esta medida del cambio sectorial, observamos el mismo patrón que empleando datos sobre producción: el sector agrícola es el sector que emplea a la mayor parte de la fuerza laboral en las primeras etapas del desarrollo económico, mientras que el sector de servicios emplea la mayor parte del trabajo en niveles elevados de desarrollo.

[Insertar Figura 2]

Un hecho destacable del proceso de desarrollo esta determinado por la evolución del sector industrial. Los datos muestran claramente que los países siguen un proceso de industrialización (crecimiento del sector industrial) y de desindustrialización (disminución del sector) en la medida que el país se desarrolla. Este patrón de U invertida es una de las características mas relevantes del cambio estructural y objeto de análisis en la literatura como se muestra más adelante.

Es importante subrayar que, aunque los datos combinados muestran una tendencia clara en cuanto la relación entre el cambio en la composición sectorial y desarrollo económico, existen diferencias en la velocidad con la cual ocurren estas transformaciones entre regiones y países. Las gráficas en la Figura 2 muestran el tamaño del sector de servicios asociado a diferentes niveles desarrollo para distintas regiones del mundo.<sup>2</sup> Al igual que en el Figura 1, los datos muestran que el crecimiento del sector de servicios esta asociado a incremento del nivel de desarrollo. Sin embargo, los gráficos muestran que existen diferencias en el tamaño del sector de servicios entre regiones para niveles de desarrollo similares.

La similitud observada en los patrones de cambio estructural sugiere que existen factores comunes en el proceso de desarrollo entre países que conducen la relocalización de la actividad productiva y empleo entre sectores. La moderna teoría del crecimiento analiza los cambios en la estructura de la demanda y factores tecnológicos para explicar la evolución de la composición sectorial durante el desarrollo y, más importante, las implicaciones de estos cambios sobre el crecimiento económico.

<sup>2</sup> Las regiones geográficas que empleamos son las regiones definidas por el banco mundial son África, Asia oriental y el Pacífico, Europa y Asia central, Norte América Latina y el Caribe, Oriente Medio y Norte de África, Asia meridional

### 3 Literatura sobre el cambio estructural

En la moderna teoría del crecimiento económico, los cambios en la estructura de la demanda y las diferencias tecnológicas entre sectores son los principales mecanismos que conducen el cambio de la composición sectorial.

#### 3.1 Cambios en la estructura de la demanda o *Income effect*

El efecto de los cambios en la estructura de la demanda sobre la composición sectorial del PIB y el empleo es explicado con base en la ley de Engel. La ley de Engel es una observación empírica robusta sobre el cambio de los gastos dado variaciones en el ingreso. A grosso modo, esta ley establece que la proporción que representa el gasto en alimentos en el ingreso total disminuye en la medida que se incrementa el ingreso.<sup>3</sup> En otras palabras, la elasticidad ingreso de la demanda de los alimentos es menor a cero.

Este mecanismo es empleado en la teoría sobre el cambio estructural para explicar los cambios en la estructura de la demanda agregada y sus efectos sobre la composición sectorial. La intuición es la siguiente. En la medida que las elasticidades ingresos de la demanda difieran entre los sectores que componen la economía, la demanda por sus productos crecen de forma no equilibrada respecto al crecimiento del ingreso. En el contexto de una economía conformada por los sectores agrícola, industrial y de servicios, la elasticidad ingreso del sector de servicios es mayor que la unidad mientras que la elasticidad ingreso del sector agrícola es menor a la unidad. Por tanto, el crecimiento de la demanda por servicios es mayor que la demanda por productos industriales y agrícolas cuando aumenta el ingreso. Como consecuencia de este mecanismo, una cantidad mayor de recursos productivos (capital y trabajo) son localizados en dicho sector que en el resto de los sectores lo que incrementa la contribución del sector de servicios al PIB.

Este mecanismo del cambio estructural basada en los cambios en la estructura de la demanda se le conoce como *income effect*, dado que asocia los cambios en la composición sectorial a incrementos del ingreso.

<sup>3</sup> Es importante señalar que esto no implica necesariamente que el gasto en alimentos permanece constante ante los incrementos del ingreso. Por lo contrario, sugiere que los consumidores incrementan su gasto en alimentos, pero a un ritmo menor que el incremento experimentado por el ingreso.

### 3.2 Diferencias tecnológicas entre sectores o price effect

La explicación tecnológica del cambio estructural centra su análisis en papel que tiene las diferencias sectoriales en (i) el crecimiento de la productividad total de los factores (TFP por sus siglas en inglés) y ii) diferencias en la intensidad de uso del capital respecto al input trabajo sobre la composición sectorial del PIB y el empleo a través de sus efectos sobre los precios relativos sectoriales.

La intuición básica del efecto de las diferencias en el crecimiento de la TFP sectoriales sobre la composición sectorial de la economía tiene su origen en el trabajo de Baumol (1967). Baumol formuló la siguiente hipótesis: los recursos económicos, principalmente el trabajo, se relocalizan de los sectores progresivos (aquellos sectores con tasas relativamente altas de crecimiento de la productividad) hacia los sectores no progresivos.<sup>4</sup> Esta idea fue formalizada en los modelos de crecimiento multisectoriales para explicar el cambio estructural. En el contexto de la tricotomía en la composición sectorial, el crecimiento del sector de servicios, tanto en su contribución al PIB como empleo localizado en este sector, es explicado por el bajo crecimiento de productividad comparado con el crecimiento de la productividad en los sectores industrial y agrícola.<sup>5</sup> Esto ocurre debido a los efectos de la productividad (un factor tecnológico) sobre el crecimiento de los precios relativos de la agricultura y servicios respecto al precio de la industria.

El precio relativo es definido como el cociente entre el precio índice de un sector respecto al precio índice de otro sector que es usado como numerario. Si el crecimiento del precio del sector  $j$  relativo al precio

<sup>4</sup> En el trabajo de Baumol, el traslado de recursos económico hacia los sectores con menores tasas de crecimiento de productividad implica necesariamente una reducción de la tasa de crecimiento de la economía agregada. Esto es así debido al crecimiento del sector no progresivo al recibir recursos de forma continua. Consecuentemente, estos sectores contribuyen en mayor grado a la productividad agregada y su crecimiento al ser más relativamente más grandes que otros sectores en la economía. Este resultado se conoce como Baumol's cost disease.

<sup>5</sup> Ngai y Pissarides (2008) derivan las condiciones necesarias para observar el crecimiento económico y cambio estructural representado en las Figuras 1 y 2. En particular, estos autores muestran que el crecimiento de la productividad en el sector agrícola es relevante para explicar el crecimiento del sector de la industria y principalmente el de servicios en un modelo de crecimiento económico. Es decir, el sector progresivo en la economía, en términos de Baumol, es el sector agrícola. Las mediciones sobre productividad sectorial apoyan esta implicación del modelo.

del sector  $i$  es positivo (negativo), entonces el crecimiento de la productividad del sector  $j$  es necesariamente mayor (menor) que en el sector  $i$ . En una economía compuesta por tres sectores cuyos bienes son complementarios, el mayor crecimiento de la productividad agrícola relativo al sector industrial y de servicios implica que el precio del bien agrícola disminuya en el tiempo. Esto implica que la demanda por la producción de otros sectores se incremente (dada la complementariedad entre bienes) y, por tanto, más recursos (trabajo y capital) deben ser destinados al sector menos productivo (servicios) para poder satisfacer la demanda creciente. Este es el efecto de los precios relativos sobre el cambio estructural o *price effect*, donde los cambios de los precios son conducidos por las diferencias en productividad.

Cabe señalar que los cambios en los precios relativos sectoriales no resultado solo de las diferencias en el crecimiento de la productividad entre sectores. Otros factores tecnológicos pueden explicar los cambios observados en los precios relativos. Por ejemplo, las diferencias sectoriales en la intensidad de uso del capital relativo al trabajo también pueden explicar las variaciones en los precios relativos y, consecuentemente, en los cambios de la composición sectorial.

La intuición de este mecanismo es simple. Los sectores que componen la economía difieren en gran medida en la proporción que emplean capital y trabajo. Sectores que utilizan más unidades de capital por una unidad de trabajo se define como capital-intensivos. En la medida que observamos crecimiento económico, la oferta de capital relativo al trabajo en la economía aumenta. Como consecuencia de la mayor oferta relativa de capital, el precio del capital disminuye. Esta reducción del precio del capital favorece a aquellos sectores que son capital-intensivo induciendo (i) la sustitución de trabajo por capital (emplean más del factor productivo con menor precio) con el resultado de (ii) una reducción de costos de producción y una mayor productividad. En este contexto, los sectores capital-intensivos experimentan mayor productividad que otros sectores y un menor precio relativo sectorial. Una vez más, el *price effect* o las diferencias de precios inducen una relocalización del empleo hacia los sectores no capital-intensivos y la relocalización de capital hacia sectores intensivos. Estos cambios pueden explicar el cambio sectorial observados en el PIB y el empleo.<sup>6</sup>

<sup>6</sup> Acemoglu and Guerrieri (2008) desarrollan un modelo de equilibrio general para analizar el efecto de las diferencias en el uso del capital entre sectores.

### 3.3 Importancia del Income effect y price effect

En resumen, el income effect y price effect son dos explicaciones complementarias al mismo fenómeno: el cambio estructural durante el proceso de crecimiento económico. La diferencia entre esta explicación radica en los fundamentos microeconómicos. Mientras que el income effect se fundamenta en los cambios de la demanda debido al crecimiento del ingreso, el price effect tiene su fundamento en diferencias tecnológicas entre sectores.

La evidencia empírica muestra que la importancia de uno u otro efecto para explicar el cambio estructural depende del nivel de desarrollo. En particular, los datos de largo plazo sobre la economía estadounidense muestran que el income effect es más importante para explicar la composición sectorial de la economía desde mediados del siglo XIX hasta a inicios de la segunda guerra mundial, mientras que el price effect es más importante para explicar el cambio sectorial en la era postguerra. Es decir, los cambios en la composición del gasto que realizaron los hogares americanos antes de la segunda guerra mundial indujeron un importante crecimiento del sector industria y de servicios, mientras que el progreso tecnológico experimentado a partir de 1945 promovió el crecimiento del sector de servicios.<sup>7</sup>

En la moderna teoría del crecimiento económico, las implicaciones se analizan mediante modelos de equilibrio general dinámico (DGE por sus siglas en inglés). La siguiente sección muestra un modelo simple de equilibrio general que contiene los dos principales mecanismos citados en la literatura que inducen el cambio estructural con el propósito de mostrar la aplicación del modelaje DGE para describir pautas de desarrollo y analizar algunas de las implicaciones del cambio estructural sobre el crecimiento económico.

## 4 Un modelo de equilibrio general

Para mostrar los mecanismos del cambio estructural sobre el crecimiento económico, empleamos un modelo simple de equilibrio general dinámico. La estructura del modelo es simple. La economía está compuesta

---

<sup>7</sup> Dennis and Iscan (2009) muestran que el crecimiento de la productividad fue el principal promotor de las transformaciones sectorial de la economía americana después de la segunda guerra mundial, mientras que el efecto ingreso es el factor relevante para explicar los cambios en la composición sectorial desde el siglo XIX hasta 1945.

solo por consumidores y empresas. Los consumidores son caracterizados por obtener utilidad del consumo de tres tipos de productos agregados: agrícola (alimentos), industrial (TVs, smartphones, etc) y de servicios (educación, salud, etc). Cada tipo de producto es producido por empresas en los sectores agrícola, industrial y de servicios. Para simplificar el análisis y solución del modelo, asumiremos que la tecnología empleada por las empresas en cada sector es similar en el uso de intensivo del capital, pero difieren en el crecimiento de la productividad. Es decir, en este modelo los precios relativos sectoriales dependerán solo de los cambios en la productividad.

En las siguientes secciones mostramos los pasos a seguir para resolver el modelo. Primero analizamos el problema de los consumidores, después resolvemos el problema de las empresas y posteriormente caracterizamos el equilibrio de la economía.

## 4.1 Los consumidores

### 4.1.1 Las preferencias del consumidor

Como es usual en la literatura, asumimos que existen  $N$  consumidores que viven infinitamente y son idénticos en preferencias y dotación de recursos. En esta economía los consumidores obtienen utilidad del consumo de tres bienes a saber: el bien agrícola, el bien manufacturado y de servicios. En particular, asumimos que las preferencias del consumidor representativo son representadas por la siguiente función de utilidad,

$$u_t = \sum_{n=0}^{\infty} \beta^n \ln C_{t+n}, \quad (1)$$

donde  $\beta$  es la tasa de descuento subjetiva y  $C_{t+n}$  es el consumo compuesto/agregado de los tres bienes agrícola,

$$C_{t+n} = \left[ \eta_a (c_{at+n} - \bar{c})^{\frac{\varepsilon-1}{\varepsilon}} + \eta_m c_{mt+n}^{\frac{\varepsilon-1}{\varepsilon}} + \eta_s c_{st+n}^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}, \quad (2)$$

donde  $c_{at}$  es la cantidad consumida de bienes agrícolas,  $c_{mt}$  es la cantidad consumida bienes manufacturados y  $c_{st}$  es la cantidad consumida de servicios.<sup>8</sup> La importancia relativa de cada uno de los bienes en la

<sup>8</sup> Para evitar confusión con la nomenclatura en los subíndices de las variables, el lector debe tener en mente que las letras  $a$ ,  $m$  y  $s$  son los índices que denotan el sector productivo que genera los bienes, mientras que  $t$  y  $n$  son índices que denotan tiempo. En particular,  $t$  es el periodo inicial y  $n$  es el incremento del tiempo.

utilidad del consumidor es medida por el parámetro  $\eta$ , que asumiremos  $\eta_a + \eta_m + \eta_s = 1$ . La elasticidad de sustitución entre los tres bienes es definida por  $\epsilon > 0$ , y  $\bar{c}$  es un requerimiento constante de consumo mínimo o de subsistencia del bien agrícola. Es necesario subrayar que este tipo de preferencias (función de utilidad) tiene implicaciones sobre las características de las funciones de demanda de los bienes y servicios a través de la maximización de la utilidad sujeta a su restricción presupuestaria.

La primera implicación esta relacionada con la existencia de un consumo mínimo. Si existe un requerimiento consumo mínimo del bien agrícola, la demanda por bienes agrícolas crece menos que proporcional al crecimiento del ingreso. La intuición de este resultado es la siguiente. Los consumidores solo obtienen utilidad si consumen el bien agrícola por arriba del consumo mínimo. Si el consumo es inferior al mínimo, los consumidores mueren. Por tanto, los consumidores demandarán siempre una cantidad positiva del bien agrícola independientemente del nivel de ingreso. Una vez que el consumidor demanda supera el nivel de consumo mínimo, el consumo de ese bien se incrementa en menor proporción que el ingreso. Por tanto, el consumo mínimo en la función de utilidad implica que observaremos una disminución de la demanda del bien agrícola y un incremento en la demanda de bienes manufacturados y de servicios en la medida que aumenta el ingreso.

La segunda implicación está relacionada con la elasticidad de sustitución entre bienes. Bajo la especificación propuesta de la función de utilidad, permitimos que el consumidor pueda sustituir el consumo de un bien por otro, donde el parámetro  $\epsilon$  miden el grado que los consumidores pueden sustituir el consumo. Si los bienes son complementarios,  $0 < \epsilon < 1$ , entonces cuando el precio relativo de uno de los bienes, por ejemplo, el bien agrícola, la demanda por el resto de los bienes se incrementa. Por lo contrario, si los bienes son sustitutos,  $\epsilon > 1$ , entonces la demanda por el resto de los bienes se reduce cuando el precio relativo del bien agrícola disminuye.

#### **4.1.2 La restricción presupuestaria del consumidor**

Cada consumidor en la economía tiene una dotación de tiempo, igual a unidad de tiempo que destina inelásticamente a trabajar.<sup>9</sup> Además, cada consumidor tiene un activo que alquila a las empresas. Por tanto,

---

<sup>9</sup> Excluimos del analisis el papel del tiempo de ocio sobre el cambio estructural.

los consumidores obtienen su ingreso a partir del trabajo y capital y usa este ingreso en consumo e inversión. La restricción presupuestaria es

$$w_t + r_t k_t = E_t + \Delta k_{t+1}, \quad (3)$$

donde  $w_t$  es el salario;  $r_t$  es la tasa de alquiler del capital;  $k_t$  es el capital;  $\Delta k_{t+1} = k_{t+1} - k_t$  es la variación del stock de capital (inversión); y  $E = c_g + p_a c_a + p_s c_s$  es el gasto total en consumo. Es importante señalar que el gasto total es medido en unidades del bien manufacturado. Esto es así debido que el precio numerario es el bien manufacturado y consecuentemente,  $p_a$  and  $p_s$  son los precios relativos de los bienes agrícolas y de servicios.<sup>10</sup>

#### 4.1.3 Solución al problema del consumidor

Los consumidores deciden cuanto consumir del bien agrícola, de manufacturas y de los servicios con el propósito de maximizar la utilidad (1) sujeto a la restricción (3). A partir de la solución de este problema dinámico de maximización, podemos calcular los gastos que realizan los consumidores en bienes agrícolas, manufacturados y de servicios como una función de la evolución del gasto agregado.<sup>11</sup> Es conveniente escribir estas soluciones como la fracción que representa cada consumo en el gasto total que realiza un consumidor. En este sentido, la fracción de gasto en bienes agrícolas esta determinado por

$$\frac{p_{at}c_{at}}{E_t} = \frac{x_{at}}{\chi_t} \left(1 - \frac{\bar{C}_t}{E_t}\right) + \frac{\bar{C}_t}{E_t}, \quad (4)$$

la fracción de gasto en servicios

$$\frac{p_{st}c_{st}}{E_t} = \frac{x_{st}}{\chi_t} \left(1 - \frac{\bar{C}_t}{E_t}\right), \quad (5)$$

y la fracción del gasto en bienes manufacturados es determinado por

$$\frac{c_{mt}}{E_t} = \frac{1}{\chi_t} \left(1 - \frac{\bar{C}_t}{E_t}\right). \quad (6)$$

<sup>10</sup> En un economía sin dinero, un precio relativo indica el número de bienes que se debe renunciar a cambio de otro bien. En este caso,  $p_{at}$  y  $p_{st}$  indican el número de bienes manufacturados (el bien de referencia o numerario) que el consumidor debe renunciar (pagar) para obtener una unidad de los bienes agrícolas y de servicios.

<sup>11</sup> En el Apéndice de este capítulo se muestra con detalle como obtener estas ecuaciones.

donde  $\bar{C}_t = p_{at}\bar{c}_{at}$  y hemos definimos a  $\chi_t$  como un precio índice

$$\chi_t = 1 + x_{at} + x_{st}, \quad (7)$$

compuesto por  $x_{at} = \bar{\eta}_a p_{at}^{1-\varepsilon}$  y  $x_{st} = \bar{\eta}_s p_{st}^{1-\varepsilon}$ . Las ecuaciones (4), (5) y (6) muestran directamente que la *composición sectorial* del gasto agregado depende de ingreso (income effect) y de los precios relativos. Para analizar los efectos del ingreso y los precios relativos requerimos conocer la evolución temporal de  $E_t$ ,  $p_{at}$  y  $p_{st}$ .<sup>12</sup> La evolución del gasto total esta determinada por la ecuación de Euler. Con base en la condiciones de primer orden, la ecuación de Euler es

$$\frac{E_{t+1} - \bar{C}_{t+1}}{E_t - \bar{C}_t} = \beta(1 + r_{t+1}). \quad (8)$$

Esta ecuación dinámica dictamina el crecimiento del gasto como una función de la tasa de interes y el peso relativo del consumo de subsistencia. Consecuentemente, para completar el analisis, es necesario estudiar como se determina el precio del alquiler del capital y los precios relativos. En la siguiente sección, analizamos el problema de la empresa y la fijación de los precios.

## 4.2 Las empresas

La empresas en cada sector producen usando una tecnologia con rendimientos constantes a escala Cobb-Douglas a:

$$Y_{it} = A_{it} (s_{it}K_t)^\alpha (u_{it}L_t)^{1-\alpha}, \quad i = a, m, s \quad (9)$$

donde  $Y_{it}$  es la cantidad producida en el sector  $i$ ,  $\alpha \in (0, 1)$  es la elasticidad-output respecto al capital y mide la intensidad del uso del capital en la produccción,<sup>13</sup>  $s_{it}$  es la fracción del stock agregado de capital  $K_t$  destinado en el sector  $i$ , y  $u_{it}$  es la fracción del empleo total,  $L_t$ , que es destinado o localizado en el sector  $i$ . Finalmente,  $A_{it}$  mide la productividad total de los factores (TFP) en el sector  $i$  y asumimos que la TFP en cada sector es crece a una tasa constante,  $\gamma_i > 0$  Obviamente, las fracciones de

<sup>12</sup> A partir del problema del problema del consumidor, podemos obtener la regla de crecimiento del gasto agregado, mientras que la evolución de los precios relativos se obtiene a partir de la solución del problema de la empresa (siguiente sección). En el apéndice mostramos con detalle el procedimiento para obtener la ecuación de Euler.

<sup>13</sup> Como señalamos al inicio de este capítulo, asumimos que las empresas emplean con la misma intensidad el capital en su producción. Por tanto,  $\alpha$  es igual en todos los sectores.

capital y empleo deben satisfacer las siguientes restricciones

$$s_a + s_m + s_s = 1,$$

y

$$u_{at} + u_{mt} + u_{st} = 1.$$

Dado que cada individuo (consumidor) oferta inélasticamente la unidad de tiempo, la oferta total de trabajo en esta economía es igual al número de consumidores,  $N$ . Por lo tanto, en esta economía, el empleo total (demanda de empleo por parte de las empresas) satisface  $L = N$ . Es conveniente escribir la función de producción (9) en términos per capita

$$y_{it} = A_{it} (s_{it} k_t)^\alpha u_{it}^{1-\alpha}, \quad i = a, m, s, \quad (10)$$

donde  $y_{it} = Y_{it}/N$  and  $k_t = K_t/N$ . Bajo el supuesto de competencia perfecta y movilidad perfecta del capital y trabajo, cada factor de producción es pagado acorde con su productividad marginal y esas productividades marginales se igualan entre sectores. Esto implica que el precio del alquiler del capital es

$$r_t = \alpha p_{it} A_{it} (s_{it} k_t)^{\alpha-1} u_{it}^{1-\alpha}, \quad (11)$$

y el salario es

$$w_t = (1 - \alpha) p_{it} A_{it} (s_{it} k_t)^\alpha u_{it}^{-\alpha}, \quad (12)$$

donde  $p_{it}$  es el precio relativo asumimos que la tasa de depreciación del capital es igual a 1. Dado estos precios (11) y (12), obtenemos que la elección óptima de capital es

$$s_i^* = u_i, \quad (13)$$

es decir, las empresas eligen óptimamente la fracción del capital que igual a la fracción de empleo en cada sector. La igualación del precio de alquiler, el salario y la elección óptima de capital implica que el precio relativo del bien agrícola es

$$p_{at} = \frac{A_{mt}}{A_{at}}, \quad (14)$$

mientras que el precio relativo de los servicios es

$$p_{st} = \frac{A_{mt}}{A_{st}}. \quad (15)$$

Es importante señalar que los supuesto de misma intensidad de uso del capital en todos los sectores y competencia perfecta implican que los precios relativos solo dependen de las productividades (TFP) relativas. Más importante, acorde con la evidencia empírica, el precio relativo de los bienes agrícolas ha disminuido en el tiempo, mientras que el precio de los servicios se ha incrementado de forma sostenida. Bajo los supuesto del modelo aquí propuesto, esto implica que el crecimiento de la TFP agrícola es mayor que la TFP del sector de manufacturas y, consecuentemente, el crecimiento de la TFP del sector de manufacturas es mayor que la TFP del sector de servicios. En las siguientes secciones mantendremos este ranking de las tasas de crecimiento de las TFPs para caracterizar el equilibrio del modelo.

Finalmente, podemos calcular la asignación óptima de trabajo entre sectores usando la condiciones de vaciado de mercado, es decir que la oferta iguala a la demanda. En particular, asumimos que la producción de los bienes agrícolas y servicios son consumidos completamente y que solo el bien manufacturado es un bien de consumo y de inversión. Sustituyendo el gasto en bienes agrícolas (4) y servicios (5), la producción (10), la asignación óptima de capital (13), y los precios relativos (14) y (15) en  $y_{at}$  y  $y_{st}$ , y asumiendo el vaciado del mercado agrícola ( $y_{at} = c_{at}$ ) y de servicios ( $y_{st} = c_{st}$ ) despejamos  $u_{at}$  y  $u_{st}$ . Es decir, la asignación eficiente de trabajo en el sector agrícola es

$$u_{at} = \frac{x_{at}}{\chi_t} \left( \frac{E_t - \bar{C}_t}{A_{mt}k_t^\alpha} \right) + \frac{\bar{c}}{A_{mt}k_t^\alpha}, \quad (16)$$

mientras que la asignación eficiente en el sector de servicios es

$$u_{st} = \frac{x_{st}}{\chi_t} \left( \frac{E_t - \bar{C}_t}{A_{mt}k_t^\alpha} \right). \quad (17)$$

Es importante señalar que las asignaciones eficientes de trabajo en los dos sectores corresponde a la fracción de empleo que es asignado a cada sector. Es decir,  $u_{at}$  y  $u_{st}$  representan la fracción del empleo total que trabaja en la agricultura y en los servicios y, consecuentemente,  $u_{mt}^* = 1 - u_{at}^* - u_{st}^*$  es el empleo de manufacturas. Estas ecuaciones son la representación matemática de los datos en la Figura 1 sobre empleo. Al igual que el consumo, estas ecuaciones dependen del ingreso (gasto) y los precios relativos. En la siguiente sección se presenta las ecuaciones dinámicas de estas variables y las características del equilibrio de mercado de esta economía.

### 4.3 El equilibrio

Esta sección caracteriza el equilibrio. Es decir, determinamos bajo que condiciones la economía esta en una senda de crecimiento sostenido (balanced growth path) mientras que observamos los cambios en la composición sectorial de la economía. Para cumplir con este proposito, es necesario definir las ecuaciones dinámicas que rigen el crecimiento en esta economía.

En la sección del consumidor, derivamos la ecuación de Euler. Esta ecuación dicta la evolución del gasto agregado y por tanto de los consumos agrícolas, manufacturados y servicios. Es necesario subrayar que esta ecuación depende crucialmente del stock de capital a través de la tasa de interés (ver ecuaciones 8 y 11). Más importante aún, los cambios en el gasto afectan a su vez el stock de capital, pues incrementos (disminuciones) en el gasto reducen (aumentan) la inversión y por tanto el stock de capital en los periodos siguientes. Es decir, requerimos definir la ecuación que dicta la evolución del stock de capital en la economía considerando los efectos del gasto agregado.

La ecuación dinámica del capital se determina a partir de la restricción presupuestaria (3) sustituyendo la tasa de alquiler del capital (11), el salario (12), las asignaciones eficientes de capital (13) y trabajo en el sector de manufacturas,  $u_m^*$ . La ley de acumulación de capital es

$$\frac{k_{t+1}}{k_t} = A_{mt}k_t^{\alpha-1} - \frac{E_t}{k_t} + (1 - \delta). \quad (18)$$

Finalmente, para caracterizar completamente el sistema de ecuaciones dinámicas, es necesario definir la ecuación de crecimiento de la TFP de las manufacturas. Observa que la ecuación de Euler, depende el stock de capital. El stock de capital depende del nivel de gasto agregado,  $E_t$ , y además del nivel de productividad. Consecuentemente, el gasto agregado también depende del  $A_{mt}$ . Para simplificar el modelo, asumimos que el crecimiento de la productividad es constante,  $\gamma_m > 0$ . Es decir,

$$\frac{A_{mt+1}}{A_{mt}} = 1 + \gamma_m \quad (19)$$

En resumen, las ecuaciones (8), (18) y (19) conforman un *sistema dinámico* que dicta la evolución en el tiempo del resto de las variables. Para caracterizar el equilibrio, reescribiremos el sistema dinámico conformado por las ecuaciones (8), (18) y (19) en términos de variables

estacionarias. Es decir, variables que en el tiempo convergen a un valor fijo, estacionario. Definiremos el equilibrio de este nuevo sistema dinámico, es decir sus valores en el estado estacionario, y observaremos como evoluciona el sistema original y el resto de variables en ese punto fijo.

Con este proposito, definiremos tres variables estacionarias. La primera variables es  $z_t = k_t/A_m t^{\frac{1}{1-\alpha}}$ . Esta variable representa el stock de capital por unidad de eficiencia del trabajo. Usando esta definición y sustituyendo en (11), obtenemos la tasa de interés como

$$r_t = \alpha z_t^{\alpha-1} - \delta. \quad (20)$$

La segunda variables es  $e_t = E_t/k_t$ . Esta variable es el gasto en consumo por unidad de stock de capital. Usando las definiciones  $e_t$ ,  $z_t$  y la ecuación del crecimiento del capital (18), podemos calcular el crecimiento de la variable  $z_t$  como<sup>14</sup>

$$\frac{z_{t+1}}{z_t} = \frac{z_t^{\alpha-1} - e_t + (1 - \delta)}{\tilde{\gamma}_m}. \quad (21)$$

donde  $\tilde{\gamma}_m = (1 + \gamma_m)^{\frac{1}{1-\alpha}}$ . La tercera variables es  $\bar{v}_t = \bar{C}_t/k_t$ . Esta variable es el gasto en el consumo mínimo por unidad de capital per capita. Sustituyendo  $\bar{v}_t$ ,  $e_t$  y la tasa de interés en la ecuación de Euler (8), tenemos

$$\frac{e_{t+1}}{e_t} = \frac{\beta}{\tilde{\gamma}_m} \frac{z_t}{z_{t+1}} \frac{(e_t - \bar{v}_t) (\alpha z_t^{\alpha-1} + 1 - \delta)}{e_t} + \frac{\bar{v}_{t+1}}{e_t}. \quad (22)$$

Finalmente, la tasa de crecimiento del consumo mínimo por unidad de capital es

$$\frac{\bar{v}_{t+1}}{\bar{v}_t} = \frac{1 + \gamma_a}{1 + \gamma_m} [z_t^{\alpha-1} - e_t + (1 - \delta)]^{-1}. \quad (23)$$

Con base en las ecuaciones (21), (22) y (23) podemos calcular el equilibrio de esta economía.<sup>15</sup> Para ello, asumiremos que la economía esta en un estado estacionario.

<sup>14</sup> En el apéndice C mostramos los pasos intermedios para obtener las ecuaciones dinámicas.

<sup>15</sup> Formalmente, dada las condiciones iniciales de las variables  $z(0)$ ,  $\bar{v}(0)$ , de los niveles de productividad  $A_a(0)$ ,  $A_m(0)$  and  $A_s(0)$ , **un equilibrio dinámico** es una senda de las variables  $\{z_t, \bar{v}_t, q_t\}_{t=0}^{\infty}$ , de los precios relativos  $\{p_{at}, p_{st}\}_{t=0}^{\infty}$  y la fracción de empleo en los sectores agrícola y de servicios  $\{u_{at}, u_{st}\}_{t=0}^{\infty}$  que resuelve el sistema de ecuaciones en diferencias (21), (22) y (23) y satisfacen las ecuaciones (16), (17) and  $A_{it} = A_i(0)(1 + \gamma_m)^t$ ,  $i = a, m, s$ .

### 4.3.1 El estado estacionario

El estado estacionario es un equilibrio en el cual los valores de las variables  $z_t$ ,  $\bar{v}_t$ , y  $q_t$  permanecen constante en el tiempo. Es decir, en un estado estacionario ocurre que el capital por unidad de eficiencia de trabajo es  $z_{t+1} = z_t = z^*$ , el gasto por unidad de capital  $q_{t+1} = q_t = q^*$  y el consumo mínimo es  $\bar{v}_{t+1} = \bar{v}_t = \bar{v}^*$ . Imponiendo esta condición en las ecuaciones (21), (22) y (23) encontramos los valores de equilibrio.

Primero, imponemos la condición  $z_{t+1} = z_t$  en la ecuación del crecimiento de  $z_t$ , para obtener

$$\tilde{\gamma}_m = z_t^{\alpha-1} - e_t + (1 - \delta), \quad (24)$$

y sustituimos este resultado en la ecuación de crecimiento de  $\bar{v}_t$ , para obtener

$$\frac{\bar{v}_{t+1}}{\bar{v}_t} = \frac{1 + \gamma_m}{1 + \gamma_a} \frac{1}{\tilde{\gamma}_m},$$

y sustituyendo  $\tilde{\gamma}_m = (1 + \gamma_m)^{\frac{1}{1-\alpha}}$ , tenemos la tasa de crecimiento del consumo mínimo es igual a

$$\frac{\bar{v}_{t+1}}{\bar{v}_t} = \frac{1}{(1 + \gamma_a)(1 + \gamma_m)^{\frac{\alpha}{1-\alpha}}},$$

observa que  $\frac{\bar{v}_{t+1}}{\bar{v}_t} \in (0, 1)$  por lo que, en el estado estacionario observamos que  $\bar{v}_t$  decrece a una tasa constante. Esto implica que  $\bar{v}_t \rightarrow 0$  de forma asintótica. Por lo que su valor de equilibrio es

$$\bar{v}^* = 0. \quad (25)$$

Ahora, sustituimos el valor de equilibrio de  $\bar{v}^*$  en la tasa de crecimiento del gasto por unidad de capital, es decir la ecuación (22). Imponiendo que la economía se encuentra en el estado estacionario,  $e_{t+1} = e_t$ , podemos encontrar el valor de estado estacionario del capital por unidad de eficiencia es

$$z^* = \left[ \frac{\alpha\beta}{\tilde{\gamma}_m - \beta(1 - \delta)} \right]^{\frac{1}{1-\alpha}}. \quad (26)$$

Finalmente, sustituimos los valores de estado estacionario  $z^*$  y  $\bar{v}^*$  en la ecuación (24) para que encontrar el valor en el estado estacionario del gasto por unidad de capital es

$$e^* = \frac{(1 - \alpha\beta)\tilde{\gamma}_m - (1 - \alpha)\beta(1 - \delta)}{\alpha\beta}. \quad (27)$$

Como habíamos mencionado antes, el sistema dinámico conformado por  $z_t$ ,  $e_t$  y  $\bar{v}$  dictaminan la evolución de todas las variables de la economía y nos muestran el estado de la economía en el largo plazo, en particular la composición sectorial. Para mostrar esto, podemos escribir las ecuaciones que determinan la cantidad de trabajo en el sector agrícola y en los servicios como funciones de las variables estacionarias y analizar cual será la composición sectorial en el largo plazo. En el largo plazo, la economía converge a una composición sectorial donde el sector agrícola no tiene importancia tanto en su contribución al empleo y al PIB. Usando la ecuación (16) tenemos

$$u_{at} = \frac{x_{at}}{\chi_t} \left( \frac{e_t - \bar{v}_t}{z_t^{\alpha-1}} \right) + \frac{\bar{v}_t}{z_t^{\alpha-1}}, \quad (28)$$

y su valor en el largo plazo es

$$u_{at}^* = 0,$$

mientras que la cantidad de trabajo localizada en el sector servicios es

$$u_{st} = \frac{x_{st}}{\chi_t} \left( \frac{e_t - \bar{v}_t}{z_t^{\alpha-1}} \right), \quad (29)$$

y su valor en el largo plazo es

$$u_{st}^* = \frac{(1 - \alpha\beta) \tilde{\gamma}_m - (1 - \alpha)\beta(1 - \delta)}{\tilde{\gamma}_m - \beta(1 - \delta)}.$$

El modelo predice que la composición sectorial de la economía esta dominada por el sector de servicios y el sector de manufacturas mientras que el tamaño del sector agrícola, en términos relativos, es nulo. Los datos sobre la composición sectorial muestran una continua reducción del tamaño del sector agrícola en la medida que hay más crecimiento económico. El modelo parece capturar la tendencia de los datos. Sin embargo, en la realidad, la composición sectorial continua cambiando: no se encuentra en el estado estacionario. Por lo tanto, es necesario analizar la transición dinámica del modelo. Es decir, analizar la predicciones que realiza el modelo cuando la economía se encuentra en la trayectoria hacia el estado estacionario.

## 5 Simulación por ordenador del modelo

### 5.1 Intuiciones básicas

Esta sección se propone analizar las predicciones del modelo sobre las trayectorias del empleo sectorial hacia su estado estacionario (transición dinámica). Para realizar este análisis es necesario resolver el sistema dinámico sin asumir que el sistema está en el estado estacionario como lo hicimos en la sección previa. La solución de este tipo de sistema dinámico requiere el empleo de técnicas numéricas por lo que empleamos el software Dynare.

Con base en la solución, analizamos si el modelo es adecuado para analizar los patrones de composición sectorial de la economía mexicana a lo largo del tiempo. Más importante aún, usaremos el modelo para no solo describir la evolución de la economía mexicana si no que usaremos el modelo para analizar si el progreso tecnológico (price effect) o el aumento del capital (income effect) son factores importantes para explicar el desarrollo de la economía. Para ellos realizaremos un par de ejercicios o experimentos llamados contrafactuales. Un experimento o contrafactual nos muestra un resultado que pudiera haber ocurrido si se hubieran dado determinadas circunstancias. Este tipo de experimento nos permite averiguar el grado de importancia que tiene un mecanismo sobre un resultado observado. Más adelante aunaremos en este aspecto.

La solución numérica del sistema dinámico requiere que asignemos valores a todos los parámetros del modelo. La asignación de estos valores no es una tarea trivial. La asignación arbitraria de valores a los parámetros genera resultados arbitrarios del modelo. Si requerimos que el modelo explique una parte de la realidad, necesitamos domar/educar el modelo a través de una asignación restringida de los valores de los parámetros.

Este ejercicio de asignación restringida es lo que se conoce en la literatura como calibración de los parámetros del modelo. En el proceso de calibración asignamos valores a todos los parámetros que conforman el modelo tal que el modelo replique una o varias características específicas de la realidad económica. A estas características de la realidad le llamamos targets de calibración. Por ejemplo, un target de calibración que es usado frecuentemente es el valor que tiene la compensación a los trabajadores como peso en el PIB. Se observa que en la economía estadounidense que este peso ronda el valor del 70 % del PIB y este valor

ha permanecido relativamente estable en el tiempo. Esta es una característica de la economía americana que es deseable que un modelo pueda replicar. Por tanto, el valor de 0.70 de PIB es un valor que usamos para calcular el valor de los parámetros de tal forma que el valor de las variables que componen el modelo, y representan la compensación de los trabajadores en el PIB, repliquen el valor observado.

En las siguientes secciones mostramos la intuición que seguimos para la calibración de los parámetros del modelo usando datos de la economía mexicana, los resultados de la simulación, así como el resultado de los experimentos contrafactuales para analizar la importancia del price y income effect sobre el desarrollo de la economía mexicana.<sup>16</sup>

## 5.2 Calibración del modelo

Como se ha mencionado antes, la asignación de los valores a los parámetros es una tarea complicada. Una primera fuente de complicación surge debido a la identificación de targets para realizar la calibración. Por ejemplo, un modelo que contenga 200 parámetros es mucho más difícil de calibrar que un modelo compuesto por solo 10 parámetros dado que el primero requeriría, en principio, contar con 200 targets, mientras que el segundo requeriría 10 targets. Una segunda fuente de complicación emerge por la carencia de información económica sobre los targets de calibración. Usando el ejemplo anterior sobre el peso de la compensación de los trabajadores en el PIB. Si las agencias de estadística económica no reportan este valor, entonces simplemente no disponemos de información con cual podemos calibrar los parámetros. En este capítulo, sólo reportamos los targets y valores de los parámetros y omitiremos el proceso completo de calibración dado que este procedimiento está más allá del alcance de este capítulo. La Tabla 1 reporta los valores de los parámetros para los cuales el modelo replica los valores de los targets seleccionados. Cabe señalar que la calibración de los parámetros se realiza de forma simultánea.<sup>17</sup>

[Insertar Tabla 1]

---

<sup>16</sup> Los datos y código empleados en esta sección pueden ser obtenidos bajo petición escribiendo al correo [be.cruz@ugto.mx](mailto:be.cruz@ugto.mx).

<sup>17</sup> Es decir, resolvemos un sistema de ecuaciones, que representan los targets de calibración, donde los parámetros son las variables que satisfacen los valores observados de los targets.

### 5.3 Simulación del modelo

Para realizar la simulación del modelo, usamos el software Dynare. En el apéndice D, mostramos el código empleado para simular la economía en este paquete computacional. Este código consta de varios bloques. En el primer bloque definimos número de años que simularemos. Dado que simularemos la economía mexicana desde 1950 a 2012, el número de periodos a simular son 63. En el segundo y tercer bloques, declaramos las variables endógenas y las variables exógenas del modelo. En el cuarto bloque, declaramos los parámetros del modelo y asignamos sus valores calibrados. En el quinto bloque, introducimos las ecuaciones dinámicas del modelo. Es decir, las ecuaciones (21), (22) y (23). Además, incluimos en esta sección, las ecuaciones que describen la composición sectorial, es decir, las ecuaciones (28) y (29). En el sexto bloque, proporcionamos valores iniciales que toman las variables de transformadas del stock de capital y del consumo mínimo de subsistencia. En el séptimo bloque, introducimos las variables exógenas. En este caso, los valores de las variables para cada año. En los bloques octavo y noveno, introducimos los valores de estado estacionario de las variables transformadas y de la composición sectorial y los valores que deben tomar las variables exógenas al final de la simulación. Finalmente, en el décimo bloque, introducimos el código para que Dynare compute los valores de las variables del modelo para cada año.

[Insertar Figura 3]

Los resultados de la simulación están reportados en la Figura 3. La figura muestra que el modelo replica los cambios en la composición sectorial del empleo en el periodo 1950 a 2012. En el caso de la agricultura el modelo predice el 96 por ciento del cambio total observado, mientras que replica el 93 por ciento en el caso de los servicios y sobre estima el cambio total del sector de manufacturas en un 5 por ciento. Con base en este resultado podemos decir que el modelo es adecuado para describir el cambio sectorial de la economía mexicana en el periodo 1950-2012. Es decir, que los mecanismos o efectos de precio e ingreso son relevantes para explicar el proceso de desarrollo. ¿Qué tan relevantes son estos mecanismos?

### 5.4 Experimentos o contrafactuales

Como mencionamos al inicio de esta sección, un experimento contrafactual nos muestra un resultado que pudiera haber ocurrido si se hubieran dado determinadas circunstancias. En el modelo, el cambio estruc-

tural esta conducido por el cambio en los precios y el efecto del consumo mínimo sobre la asignación de gastos.

Un primer ejercicio contrafactual sería preguntarnos cual habría sido la composición sectorial que observaríamos, si los precios relativos no hubiesen cambiado durante el periodo. Es decir, si no existiera crecimiento de la productividad de los sectores. Podemos usar el modelo calibrado para realizar este experimento. En particular, supondremos que los precios no cambian y estos son iguales al valor observado de 1950, después realizamos la simulación y observamos la composición sectorial que se obtiene. La Figura 4 muestra el resultado de este primer experimento. Notablemente, el cambio estructural que observaríamos sin crecimiento de las productividades sectoriales sería prácticamente nulo. En este experimento, la variación o cambio total en la agricultura solo habría sido un 91 por ciento al realmente observado, mientras que el cambio en los servicios sería menor al observado en un 104 por ciento. En este experimento, hemos cancelado el efecto de los precios y por tanto el único mecanismo que induce el cambio sectorial es el efecto ingreso o income effect. Dado el resultado obtenido, podemos deducir que el efecto ingreso no significativo para explicar el cambio sectorial.

[Insertar Figura 4]

El segundo ejercicio contrafactual consiste en preguntarnos cuál habría sido la composición sectorial que observaríamos, si no existiera un consumo mínimo de subsistencia y por ende un efecto ingreso. Es decir, el único mecanismo que explicaría el cambio sectorial es el crecimiento de las productividades sectoriales. La Figura 5 muestra el resultado de este experimento. Notablemente, el modelo reproduce la tendencia en los datos cuando solo los precios relativos cambian en el tiempo y no permitimos el cambio en el ingreso modifique la demanda por la ley de Engel. Este resultado confirma el resultado del primer experimento: el factor relevante para explicar el cambio sectorial en México es el crecimiento diferenciado de las productividades sectoriales. Es decir, un factor tecnológico explicaría la rápida disminución del sector agrícola y el crecimiento de servicios. Es importante recordar que el modelo ha omitido otro factor tecnológico que es la diferencia del uso de capital entre sectores. Este mecanismo puede contribuir al cambio sectorial en México de forma importante.

[Insertar Figura 5]

Los resultados de este simple ejercicio nos sugieren que los factores tecnológicos, como el crecimiento de la productividad sectoriales o diferencias en el uso del capital son relevante para explicar el desarrollo de la economía mexicana. Con base en estos resultados, podemos conjeturar que los factores que al crecimiento de las productividades sectoriales y el uso del capital como son la acumulación de capital humano e investigación y desarrollo podrían afectar de forma sustancial el desarrollo de la composición sectorial de la economía mexicana.

## 6 Conclusiones

En este capítulo mostramos que el proceso de crecimiento económico esta caracterizado por la transformación o cambio de la composición sectorial. En general, observamos que el peso relativo del sector agrícola en el PIB y en el empleo total se reduce, mientras que el tamaño del sector de servicios se incrementa. Más interesante es la evolución del sector industrial. Este sector sigue un proceso de U invertida: en etapas tempranas del desarrollo observamos un proceso de industrialización mientras que en etapas avanzadas observamos un proceso de desintrialización. Este patrón de desarrollo es similar entre países y grupos de países por región.

La moderna teoría del crecimiento señala que estos cambios en la composición sectorial del PIB y el empleo son conducidos por los cambios en la demanda inducidos por el crecimiento del ingreso (income effect) y por factores tecnológicos, específicamente, incrementos en la productividad en cada sector. El análisis de la importancia de estos efectos para explicar los hechos estilizados se realiza mediante el empleo de modelos de equilibrio general dinámico. En este capítulo presentamos un modelo de equilibrio general que contiene los mecanismos de precio e ingreso para explicar el cambio sectorial. Resolvemos el modelo y empleamos el paquete computacional Dynare para analizar el cambio sectorial en la economía mexicana a la luz del modelo.

La simulación del modelo muestra que este es capaz de describir las variaciones en la composición sectorial en la economía mexicana para el periodo 1950-2012. Finalmente, usamos el modelo para analizar la importancia de los mecanismos de precio e ingreso para el desarrollo mexicano. Los resultados obtenidos sugieren que el factor tecnológico es el principal motor del cambio en la composición sectorial del empleo. Con

base en estos resultados conjeturamos que los factores que promuevan el crecimiento de la productividad a nivel sectorial tendrían un impacto importante en la composición sectorial del empleo.

## Referencias

- [1] Acemoglu, D. , Guerrieri, V. (2008). Capital deepening and non-balanced economic growth. *Journal of Political Economy*, 116, 467-498.
- [2] Alonso-Carrera, J., Caballé, J. and Raurich, X. (2015). Sectoral Composition and Macroeconomic Dynamics. *BE J. Macroecon. Advances*, 15, 1-42,
- [3] Boppart, T. (2014). Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences, *Econometrica*, 82, 2167-2196.
- [4] Buera, F. and Kaboski, J. (2012). Scale and the origins of structural change, *Journal of Economic Theory*, 147, 684-712.
- [5] Dennis, B. and Iscan, T. (2009). Engel versus Baumol: accounting for sectoral change using two centuries of US data. *Explorations in Economic History*, 46, 186-202.
- [6] Duernacker, G. and Herrendorf, B. (2015). On the allocation of time - A Quantitative Analysis of the U.S. and France, CESifo Working Paper Series No. 5475.
- [7] Foellmi, R. and Zweimüller, J. (2008). Structural change, Engel's consumption cycles and Kaldor's facts of economic growth. *Journal of Monetary Economics*, 55,1317-1328.
- [8] Gollin, D., Parente, S. and Rogerson, R. (2004). Homework, farmwork, and international differences in productivity, *Review of Economic Dynamics*, 7, 827-850.
- [9] Herrendorf, B. , Rogerson, R. and Valentinyi, Á. (2013). Two Perspectives on Preferences and Structural Transformation. *American Economic Review*, 103, 2752-2789.
- [10] Herrendorf, B. ., Rogerson, R. and Valentinyi, Á. (2014). Growth and Structural Transformation, *Handbook of Economic Growth*, in:

Handbook of Economic Growth, edition 1, volume 2, chapter 6, pages 855-941 Elsevier.

- [11] Kongsamunt, P., Rebelo, S., Xie, D. (2001). Beyond balanced growth. *Review of Economic Studies*, 68, 869-882.
- [12] Ngai, L. R., Pissarides, C. (2007). Structural change in a multisector model of growth. *American Economic Review*, 97, 429-443.
- [13] Ngai, L. R., Pissarides, C. (2008). Trends in hours and economic growth, *Review of Economic Dynamics*, 11, 239-256.
- [14] M.P. Timmer, G.J. de Vries, and K. de Vries (2014). Patterns of Structural Change in Developing Countries. GGDC research memorandum 149.

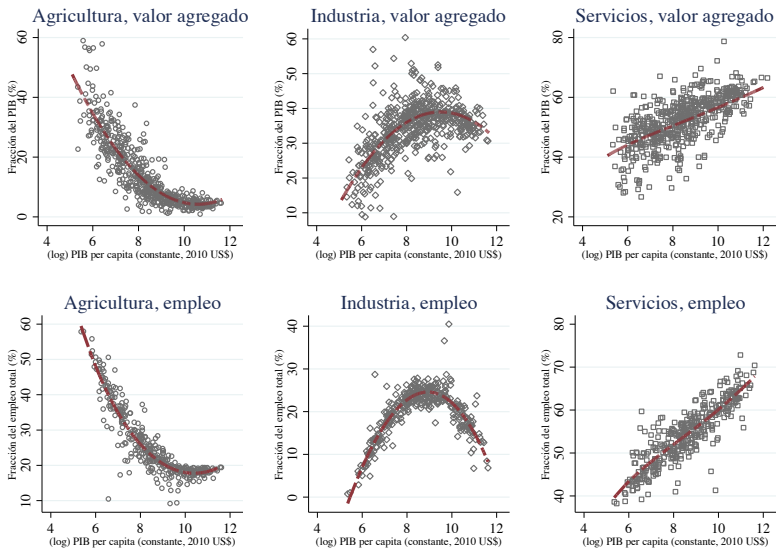
## A Tablas y Figuras

Cuadro 1: Calibración de los parámetros

Parámetro	Valor	Targets	Observado
$\alpha$	0.300	Compensación del trabajo (%PIB)	0.652
$\gamma_m$	0.010	Crecimiento del PIB per capita	0.012
$\gamma_s$	0.010	Crecimiento precio servicios	0.001
$\gamma_a$	0.04	Crecimiento precio agricultura	-0.030
$A_s$	0.4002	Nivel TFP inicial en agricultura	0.930
$A_s$	1.076	Nivel TFP inicial en servicios	2.430
$A_g$	1	Normalizado	-
$\tilde{\eta}_a$	6.390	Empleo en sector agrícola (1950)	0.581
$\tilde{\eta}_s$	7.285	Empleo en sector servicios (1950)	0.257
$\tilde{c}_a$	0.012	Empleo en sector industria (2012)	0.260
$\epsilon$	0.001	Relacion capital-ouput (promedio)	1.500

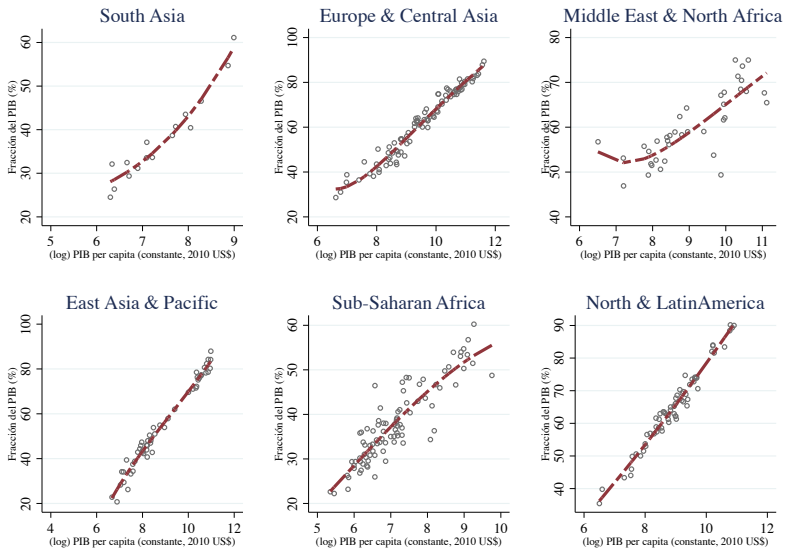
Los datos sobre la compensación del trabajo en el PIB y el cociente capita-producto que reportamos corresponde al promedio reportado en la base de datos Penn World Table version 9.0 para el periodo 1970-2010. La tasa de crecimiento per capita, precio relativos y el empleo en cada sector fue calculada usando datos 10 sector database de la universidad de Groningen para el periodo 1950 a 2012.

Figura 1: Patrones de cambio sectorial



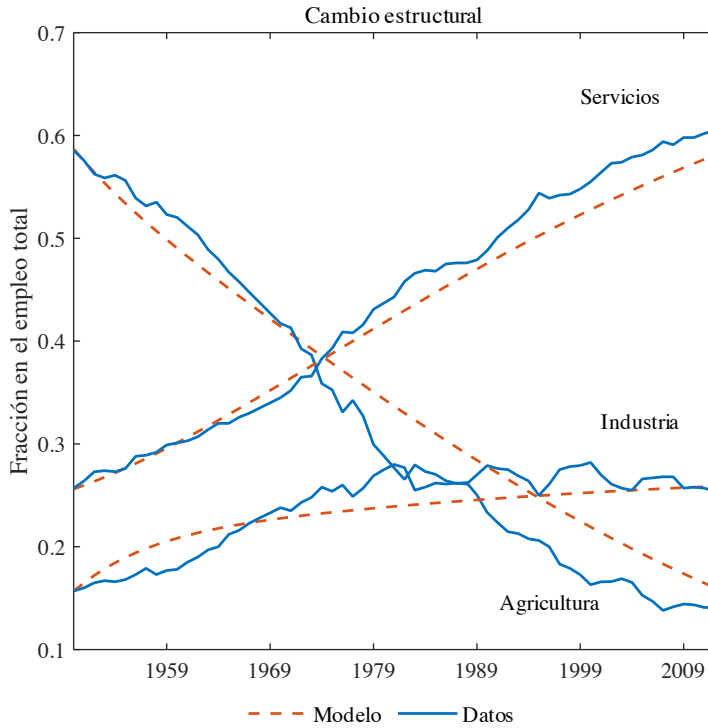
*Fuente:* Los datos empleo y valor añadido provienen de World Development Indicators. *Nota:* los datos muestran la relación entre PIB y composición sectorial del valor añadido (fila 1) y el empleo (fila 2) controlando por diferencias entre país que son constantes en el tiempo (country fixed effects). Cada gráfico muestra el valor esperado de empleo (o valor añadido) en cada sector dado su nivel de desarrollo (por nivel de PIB) una vez que consideramos diferencias constantes entre países, por ejemplo, geografía.

Figura 2: Patrones de cambio sectorial por región



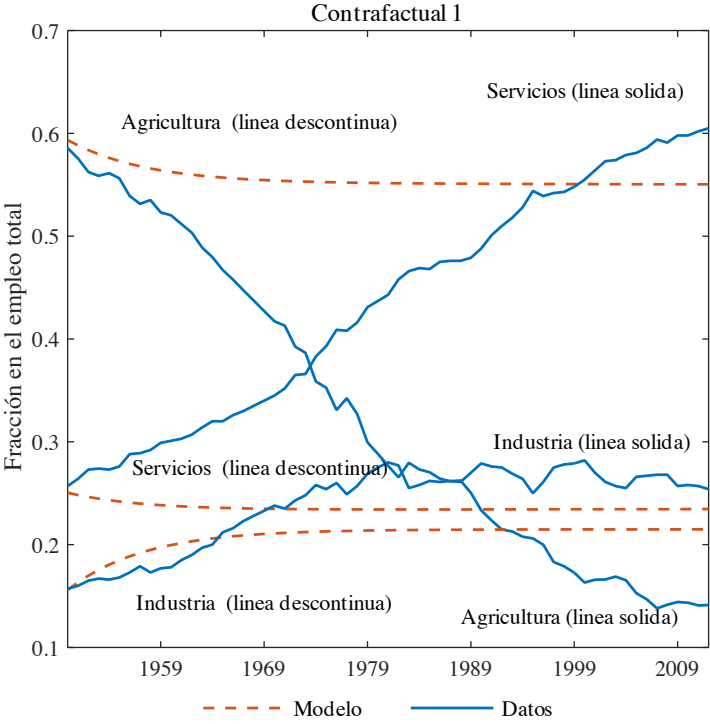
*Fuente:* Los datos empleo y valor añadido provienen de World Development Indicators. *Nota:* los datos muestran la relación entre PIB y composición sectorial del valor añadido (fila 1) y el empleo (fila 2) controlando por diferencias entre país que son constantes en el tiempo (country fixed effects). Cada gráfico muestra el valor esperado de empleo (o valor añadido) en cada sector dado su nivel de desarrollo (por nivel de PIB) una vez que consideramos diferencias constantes entre países, por ejemplo, geografía.

Figura 3: Simulación de la composición sectorial



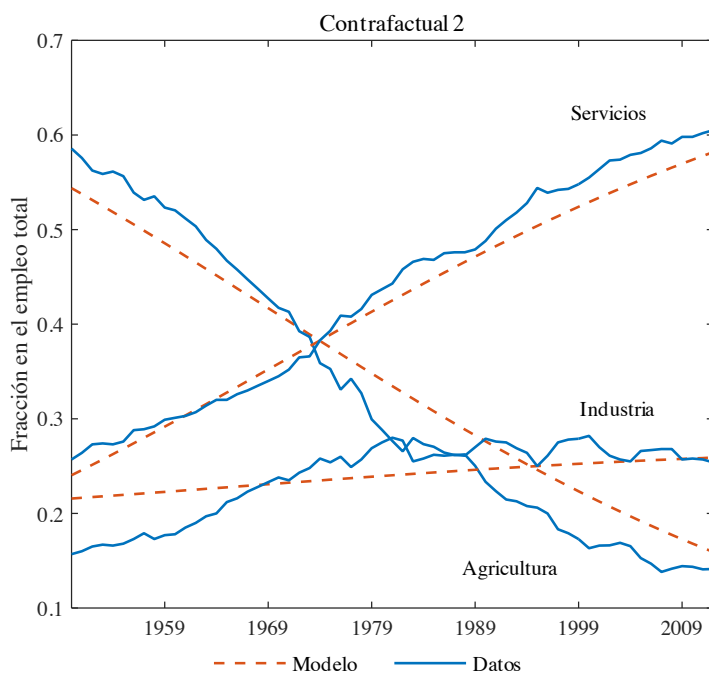
*Fuente:* Los datos empleo y valor añadido provienen de World Development Indicators. *Nota:* los datos muestran la relación entre PIB y composición sectorial del valor añadido (fila 1) y el empleo (fila 2) controlando por diferencias entre país que son constantes en el tiempo (country fixed effects). Cada gráfico muestra el valor esperado de empleo (o valor añadido) en cada sector dado su nivel de desarrollo (por nivel de PIB) una vez que consideramos diferencias constantes entre países, por ejemplo, geografía.

Figura 4: Simulación de la composición sectorial (sin price effect)



Fuente: Los datos empleo y valor añadido provienen de la base de datos 10-sector database de la universidad de Groningen .

Figura 5: Simulación de la composición sectorial (sin income effect)



Fuente: Los datos empleo y valor añadido provienen de la base de datos 10-sector database de la universidad de Groningen .

## B Solución al problema del consumidor

Los consumidores maximizan la función de utilidad sujeto a la restricción presupuestaria. El Lagrangiano asociado a este problema de maximización esta definido por

$$\mathcal{L} = \sum_{n=0}^{\infty} \{ \beta^n \ln C_{t+n} + \lambda_{t+n} (w_{t+n} + (1 + r_{t+n}) k_{t+n} - c_{mt+s} - p_{at+n} c_{at+n} - p_{st+n} c_{st+n}) \}$$

Las condiciones de primer orden respecto a  $c_{at+i}$ ,  $c_{mt+n}$ ,  $c_{st+n}$ , y  $k_{t+n}$  son, respectivamente,

$$\frac{\partial \mathcal{L}}{\partial c_{at}} : \beta^n C_{t+n}^{\frac{1-\varepsilon}{\varepsilon}} \eta_a (c_{at+n} - \bar{c})^{-\frac{1}{\varepsilon}} = \lambda_{t+n} p_{at+n}, \quad n \geq 0, \quad (30)$$

$$\frac{\partial \mathcal{L}}{\partial c_{mt}} : \beta^n C_{t+n}^{\frac{1-\varepsilon}{\varepsilon}} \eta_m (c_{mt+n})^{-\frac{1}{\varepsilon}} = \lambda_{t+n}, \quad n \geq 0, \quad (31)$$

$$\frac{\partial \mathcal{L}}{\partial c_{st}} : \beta^n C_{t+n}^{\frac{1-\varepsilon}{\varepsilon}} \eta_m (c_{st+n})^{-\frac{1}{\varepsilon}} = \lambda_{t+n} p_{st+n}, \quad n \geq 0, \quad (32)$$

$$\frac{\partial \mathcal{L}}{\partial k_t} : \lambda_{t+n} (1 + r_t) = \lambda_{t+n-1}, \quad n > 0. \quad (33)$$

Despues, resolviendo para  $n = 0$ , se procede a obtener los cantidades consumidas como funciones de los precios relativos, el salario y el gasto total de consumo. Para este fin, combinamos las ecuaciones (30), (31) y (32). Primero, sustituimos  $\lambda_t$  de la ecuación (31) en (30) para obtener el consumo del bien agrícola como una función del consumo manufacturado

$$c_{at} - \bar{c} = \left( \frac{\eta_a}{\eta_m} \right)^{\varepsilon} p_{at}^{-\varepsilon} c_{mt}, \quad (34)$$

donde hemos multiplicado la ecuación por el precio relativo agrícola. Siguiendo los pasos anteriores, sustituimos  $\lambda_t$  de la ecuación (31) en (32) para obtener el consumo de los servicios como una función del consumo manufacturado

$$c_{st} = \left( \frac{\eta_s}{\eta_m} \right)^{\varepsilon} p_{st}^{-\varepsilon} c_{mt}. \quad (35)$$

Despues, sustituimos (34) y (35) en la definición del consumo compuesto (2) para obtener

$$C_t = \left[ \eta_a \left( \frac{\eta_a}{\eta_m} \right)^{\varepsilon-1} p_{at}^{1-\varepsilon} c_{mt}^{\frac{\varepsilon-1}{\varepsilon}} + \eta_m c_{mt}^{\frac{\varepsilon-1}{\varepsilon}} + \eta_s \left( \frac{\eta_s}{\eta_m} \right)^{\varepsilon-1} p_{st}^{1-\varepsilon} c_{mt}^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}.$$

Esta ecuación puede simplificarse al obtener el factor común,  $\eta_m c_{mt}^{\frac{\varepsilon-1}{\varepsilon}}$ , y escribir el consumo compuesto como

$$C_t = \eta_m^{\frac{\varepsilon}{\varepsilon-1}} c_{mt} \chi_t^{\frac{\varepsilon}{\varepsilon-1}}. \quad (36)$$

donde definimos la ecuación auxiliar

$$\chi_t = 1 + \bar{\eta}_a p_{at}^{1-\varepsilon} + \bar{\eta}_s p_{st}^{1-\varepsilon}, \quad (37)$$

$\bar{\eta}_a = \left(\frac{\eta_a}{\eta_m}\right)^\varepsilon$  y  $\bar{\eta}_s = \left(\frac{\eta_s}{\eta_m}\right)^\varepsilon$ . Finalmente, sustituimos (36) en (31) en la condición de primer orden respecto al consumo de manufacturas

$$\frac{1}{c_{mt} \chi_t} = \lambda_t. \quad (38)$$

Observa que esta ecuación se mantiene para  $n = 0, 1, \dots, \infty$ . Por lo que podemos obtener,  $\lambda_{t+1}$  cuando  $n = 1$ , como

$$\frac{\beta}{c_{mt+1} \chi_{t+1}} = \lambda_{t+1}. \quad (39)$$

Calculamos la tasa de crecimiento, dividiendo  $\lambda_{t+1}$  por  $\lambda_t$ , es decir

$$\frac{\lambda_{t+1}}{\lambda_t} = \beta \frac{c_{mt} \chi_t}{c_{mt+1} \chi_{t+1}}. \quad (40)$$

Usando la condición de primer orden respecto al capital (33), cuando  $s=1$ , obtenemos que

$$\frac{\lambda_{t+1}}{\lambda_t} = \frac{1}{(1 + r_{t+1})}. \quad (41)$$

Finalmente, usando (40) y (41) obtenemos la ecuación de Euler.

$$\frac{c_{mt+1}}{c_{mt}} = \beta (1 + r_{t+1}) \frac{\chi_t}{\chi_{t+1}}. \quad (42)$$

Podemos escribir la ecuación de Euler en términos del gasto agregado  $E_t$ . Para ellos multiplicamos (34) y (35) por los precios relativos  $p_{at}$  y  $p_{st}$ , respectivamente, para obtener

$$p_{at} c_{at} = \left(\frac{\eta_a}{\eta_m}\right)^\varepsilon p_{at}^{1-\varepsilon} c_{mt} + p_{at} \bar{c}_{at},$$

y donde hemos multiplicado la ecuación por el precio relativo agrícola. Siguiendo los pasos anteriores, sustituimos  $\lambda_t$  de la ecuación (31) en (32) para obtener el consumo de los servicios como una función del consumo manufacturado

$$p_{st}c_{st} = \left(\frac{\eta_s}{\eta_m}\right)^\varepsilon p_{st}^{1-\varepsilon} c_{mt}.$$

Sustituimos en el gasto total en consumo  $E_t = c_{mt} + p_{at}c_{at} + p_{st}c_{st}$  para obtener

$$E_t = c_{mt} + \left(\frac{\eta_a}{\eta_m}\right)^\varepsilon p_{at}^{1-\varepsilon} c_{mt} + p_{at}\bar{c}_{at} + \left(\frac{\eta_s}{\eta_m}\right)^\varepsilon p_{st}^{1-\varepsilon} c_{mt}.$$

Usando la variable auxiliar (37), despejamos  $c_{mt}$  como función del gasto total en consumo.

$$c_{mt} = \frac{E_t - \bar{C}_t}{\chi_t}, \quad (43)$$

y  $\bar{C}_t = p_{at}\bar{c}_{at}$  es el gasto en el consumo de subsistencia. A partir de la ecuación anterior, podemos definir el gasto en el consumo de bienes agrícolas y el gasto en servicios como fracciones del gasto total. Sustituyendo (43) en (34)

$$c_{at} - \bar{c} = \left(\frac{\eta_a}{\eta_m}\right)^\varepsilon p_{at}^{1-\varepsilon} \frac{E_t - \bar{C}_t}{\chi_t},$$

después multiplicamos por el precio relativo

$$p_{at}c_{at} = \left(\frac{\eta_a}{\eta_m}\right)^\varepsilon p_{at}^{-\varepsilon} \frac{E_t - \bar{C}_t}{\chi_t} + p_{at}\bar{c},$$

y dividimos por el gasto total para obtener la ecuación (4) en el texto principal

$$\frac{p_{at}c_{at}}{E_t} = \frac{x_{at}}{\chi_t} \left(1 - \frac{\bar{C}_t}{E_t}\right) + \frac{\bar{C}_t}{E_t},$$

donde  $x_{at} = \bar{\eta}_a p_{at}^{1-\varepsilon}$ . Siguiendo los pasos anteriores, tenemos que la fracción del gasto total en servicios es

$$\frac{p_{st}c_{st}}{E_t} = \frac{x_{st}}{\chi_t} \left(1 - \frac{\bar{C}_t}{E_t}\right)$$

donde  $x_{st} = \bar{\eta}_s p_{st}^{1-\varepsilon}$ . Finalmente, sustituimos (43) en las ecuaciones (38)

$$\frac{1}{E_t - \bar{C}_t} = \lambda_t, \quad (44)$$

y en la ecuación (39)

$$\frac{\beta}{E_{t+1} - \bar{C}_{t+1}} = \lambda_{t+1}, \quad (45)$$

y usando (41) tenemos que la ecuación de Euler en términos del gasto total de consumo

$$\frac{E_{t+1} - \bar{C}_{t+1}}{E_t - \bar{C}_t} = \beta (1 + r_{t+1})$$

que puede ser reescrita como

$$\frac{E_{t+1}}{E_t} = \beta (1 + r_{t+1}) \left( 1 - \frac{\bar{C}_t}{E_t} \right) + \frac{\bar{C}_{t+1}}{E_t}.$$

## C Sistema de ecuaciones dinámicas

En este apéndice mostramos como obtener el sistema de ecuaciones dinámicas de las variables transformadas  $z_t$ ,  $e_t$  y  $\bar{v}_t$ . Primero, usando la definición de  $z_t$ , tenemos

$$z_t = \frac{k_t}{A_{mt}^{\frac{1}{1-\alpha}}}$$

luego dividimos por  $z_{t+1}$ , y ordenamos de la siguiente forma

$$\frac{z_{t+1}}{z_t} = \left( \frac{A_{mt}}{A_{m,t+1}} \right)^{\frac{1}{1-\alpha}} \frac{k_{t+1}}{k_t} \quad (46)$$

luego sustituimos la tasa de crecimiento de la productividad, ecuación (19), y la tasa de crecimiento del capital per capita, ecuación (18), tenemos

$$\frac{z_{t+1}}{z_t} = \left( \frac{1}{1 + \gamma_m} \right)^{\frac{1}{1-\alpha}} \left[ A_{mt} k_t^{\alpha-1} - \frac{E_t}{k_t} + (1 - \delta) \right],$$

y sustituimos  $e_t = E_t/k_t$  y  $z_t$ , para obtener la ecuación dinámica de  $z_t$  en el texto principal

$$\frac{z_{t+1}}{z_t} = \left( \frac{1}{1 + \gamma_m} \right)^{\frac{1}{1-\alpha}} [z_t^{\alpha-1} - e_t + (1 - \delta)]. \quad (47)$$

Después, obtenemos la ecuación de Euler basada en las variables transformadas. Primero usamos la definición  $e_t$  y  $\bar{v}_t$ , junto con la tasa de interés (20), en la ecuación de Euler (8), para obtener

$$\frac{k_{t+1}}{k_t} \frac{e_{t+1} - \bar{v}_{t+1}}{e_t - \bar{v}_t} = \beta (\alpha z_t^{\alpha-1} + 1 - \delta).$$

Después, podemos sustituir (46) y  $\tilde{\gamma}_m = (1 + \gamma_m)^{\frac{1}{1-\alpha}}$ , para obtener

$$\tilde{\gamma}_m \frac{z_{t+1}}{z_t} \frac{e_{t+1} - \bar{v}_{t+1}}{e_t - \bar{v}_t} = \beta (\alpha z_t^{\alpha-1} + 1 - \delta).$$

Despejando  $e_{t+1}$ , y dividiendo por  $e_t$ , tenemos la ecuación que reportamos en el texto principal.

$$\frac{e_{t+1}}{e_t} = \frac{\beta}{\tilde{\gamma}_m} \frac{z_t}{z_{t+1}} \frac{(e_t - \bar{v}_t) (\alpha z_t^{\alpha-1} + 1 - \delta)}{e_t} + \frac{\bar{v}_{t+1}}{e_t}. \quad (48)$$

Finalmente, la ecuación dinámica del consumo de subsistencia o mínimo esta determinado por el crecimiento del precio relativo de la agricultura y el crecimiento del capital per capita. Recordando que

$$\bar{v}_t \equiv \frac{\bar{C}_t}{k_t} = \frac{p_{at} \bar{c}_a}{k_t},$$

y dividiendo por  $\bar{v}_{t+1}$ , y ordenando términos tenemos

$$\frac{\bar{v}_{t+1}}{\bar{v}_t} = \frac{p_{at+1}}{p_{at}} \frac{k_t}{k_{t+1}}. \quad (49)$$

El precio relativo es el cociente entre la productividad del sector de manufacturas ( $A_m$ ) y el sector agrícola ( $A_a$ ). por lo que la tasa de crecimiento es igual al cociente de las tasas de crecimiento de las TFPs. Es decir,

$$\frac{p_{at+1}}{p_{at}} = \frac{1}{\frac{A_{at+1}}{A_{at}}} \frac{A_{mt+1}}{A_{mt}}.$$

Sustituyendo las tasas de crecimiento constantes,

$$\frac{p_{at+1}}{p_{at}} = \frac{1 + \gamma_m}{1 + \gamma_a}. \quad (50)$$

Usando las tasa de crecimiento del precio agrícola y la tasa de crecimiento per capita tenemos

$$\frac{\bar{v}_{t+1}}{\bar{v}_t} = \frac{1 + \gamma_m}{1 + \gamma_a} \frac{k_t}{k_{t+1}} = \frac{1 + \gamma_m}{1 + \gamma_a} \frac{z_t}{\tilde{\gamma}_m z_{t+1}},$$

simplificando al sustituir  $\frac{z_t}{z_{t+1}}$

$$\frac{\bar{v}_{t+1}}{\bar{v}_t} = \frac{1 + \gamma_m}{1 + \gamma_a} \frac{1}{z_t^{\alpha-1} - e_t + (1 - \delta)}.$$

## D Código en Dynare

```

% MODELANDO LA ECONOMIA MEXICANA 1950-2012
%%*****
//
    @#define simulation_periods=63

%%*****
//          VARIABLES
// Endogenous variables

var z  ${z}$ (long_name='Capital stock per unit of efficient labor')
    e  ${e}$ (long_name='Expenditure per unit of capital ')
    v  ${v}$ (long_name='Minimum consumption expenditure per unit of capital')
    us ${us}$ (long_name='Labor share in services')
    ua ${ua}$ (long_name='Labor share in agriculture')
    ;

// Exogenous variables
varexo ps ${ps}$ (long_name='Relative price')
    pa  ${pa}$ (long_name='Relative price')
    Xs  ${Xs}$ (long_name='Auxiliar variable')
    Xa  ${Xa}$ (long_name='Auxiliar variable')
    ;

%%*****
//          PARAMETERS

parameters epsi  ${epsi}$ (long_name='Elasticity of substitution')
    cbar  ${cbar}$ (long_name='Minimum consumption ')
    beta  ${beta}$ (long_name='Discount factor in utility')
    etaS_bar  ${etaS_bar}$ (long_name='Auxiliar parameter.')
    etaA_bar  ${etaA_bar}$ (long_name='Auxiliar parameter.')
    alpha  ${alpha}$ (long_name='Capital share in production function')
    delta  ${delta}$ (long_name='Capital depreciation rate')
    gM  ${gM}$ (long_name='TFP growth rate in goods sector')
    gPS  ${gPS}$ (long_name='growth rate of relative price')
    gPA  ${gPA}$ (long_name='growth rate of relative price')
    gMA  ${gMA}$ (long_name='Auxiliar parameter')
    ps0  ${ps0}$ (long_name='Relative price at t=0')
    zss  ${zss}$ (long_name='z value at the steady state')
    ess  ${ess}$ (long_name='e value at the steady state')
    vss  ${vss}$ (long_name='v value at the steady state')
    uss  ${uss}$ (long_name='us value at the steady state')
    dist  ${dist}$ (long_name='distance to the steady state')
    z0  ${z0}$ (long_name='Initial value of z')
    v0  ${v0}$ (long_name='Initial value of v')
    ;

//          SET PARAMETERS

    epsi  = 0.0001;
    cbar  = 0.01259025;
    beta  = 0.9580;
    alpha = 0.3;
    delta = 0.050;
    gM    = 1.0120;
    gPS   = 1.0014;
    gPA   = 0.96786;
    ps0   = 0.93;
    pa0   = 2.48;
    gMA   = gM^(1/(1-alpha));
    zss   = ( alpha*beta/( gMA - beta*(1-delta) ) )^(1/(1-alpha)) ;
    ess   = ( (1-alpha*beta)*gMA - (1-alpha)*beta*(1-delta) )/(alpha*beta);
    vss   = 0;
    uss   = ess/(zss^(alpha-1)) ;

```

```

dist      = 1.42521;
z0        = dist*zss ;
v0        = (cbar/ (z0)) ;
etaS_bar  = 0.235*(ps0^(epsi-1))/( ( (ess-vss)/(zss^(alpha-1)) ) - (0.235 + 0.55) ) ;
etaA_bar  = 0.55*(pa0^(epsi-1))/( ( (ess-vss)/(zss^(alpha-1)) ) - (0.235 + 0.55) ) ;

%%*****
//          THE MODEL

model;
% Dynamic Equation of Total expenditure:
[name='Law of motion: expenditure per unit of capital']
e(+1) = (beta/gMA)*(z(-1)/z)*(e-v)*( alpha*(z(-1)^(alpha-1)) + 1 - delta ) + (v) ;

% Dynamic Equation of Minimum Consumption Expenditure:
[name='Law of motion: minimum consumption expenditure per unit of capital']
v = v(-1) / ( gM*( (z(-1)^(alpha-1)) - e + 1 - delta ) ) ;

% Dynamic Equation of Capital
[name='Law of motion stock of capital per unit of efficient labor']
z = z (-1)*( (z(-1)^(alpha-1)) - e + 1 - delta)/gMA ;

% Employment share in service sector
[name='Employment share in service sector']
us = Xs*( ( e - v )/(z^(alpha-1)) ) ;

% Employment share in agriculture sector
[name='Employment share in service sector']
ua = Xa*( ( e - v )/(z^(alpha-1)) ) + (v/z^(alpha-1)) ;

end;

%%*****
//          INITIAL VALUES

initial;
z = z0 ;
v = v0 ;

end;

%%*****
//          SHOCKS-BLOCK: define path of exogenous variables
// Definimos las trayectorias temporales de los precios como un de variables exogenas

values_ps=cumprod((gPS^1)*ones(@{simulation_periods},1))*ps0;
values_pa=cumprod((gPA^1)*ones(@{simulation_periods},1))*pa0;

// values_ps= ps0; // Precio relativo de los servicios
// values_pa=pa0; // Precio relativo de la agricultura

// construimos los vectores que contiene la evolución de Xa y Xs:

values_Xs = etaS_bar.*(values_ps.^(1-eps))/( 1 + etaA_bar.*(values_pa.^(1-eps)) ) + etaS_bar.*(values_ps.^(1-eps));
values_Xa = etaA_bar.*(values_pa.^(1-eps))/( 1 + etaA_bar.*(values_pa.^(1-eps)) ) + etaS_bar.*(values_ps.^(1-eps));

shocks;
var ps;
periods 1:@{simulation_periods};
values (values_ps);

var pa;
periods 1:@{simulation_periods};
values (values_pa);

var Xs;

```

```

periods 1:@{simulation_periods};
values (values_Xs);

var Xa;
periods 1:@{simulation_periods};
values (values_Xa);

end;
%%*****%

steady_state_model;
us=uss;
e =ess;
z=zss;
v=0;
end;

%%*****%
//          END VALUES

endval;

ps=ps0*(gPS)^(@{simulation_periods}+1);
Xs= etaS_bar*(ps^(1-eps))/( 1 + etaA_bar*(pa^(1-eps)) + etaS_bar*(ps^(1-eps)) );
Xa= etaA_bar*(pa^(1-eps))/( 1 + etaA_bar*(pa^(1-eps)) + etaS_bar*(ps^(1-eps)) );
end;

%%%*****%
//          RUN Simulation of the model

perfect_foresight_setup(periods=@{simulation_periods});
perfect_foresight_solver(maxit=100, stack_solve_algo=7, solve_algo=9, tolf=0.00001);

%%%*****%

%%%*****%
//          STORAGE SIMULATIONS

sim_ua = ua(2:end-1);
sim_us = us(2:end-1);
sim_um = 1 - sim_ua - sim_us;
sim_ps = values_ps(1:end);
sim_pa = values_pa(1:end);

%%%*****%
//          GRAFICOS
load('real_mexico.mat') // carga los datos sobre empleo sectorial 1950-2012
% Agricultura
plot(sim_ua),
hold on
plot(agri),
% Servicios
plot(sim_us)
hold on
plot(ser),
% Industria
plot(sim_um)
hold on
plot(ind),

```

# Migración internacional

Miriam Saldaña Hernández

Coralia Quintero Rojas

## 1 Introducción

La migración es un fenómeno milenario y muy diverso, pues aun cuando la movilidad ha sido una característica intrínseca de todas las poblaciones, algunas personas se mueven más que otras y no todos los desplazamientos obedecen a las mismas causas. Por ejemplo, algunas personas se desplazan en búsqueda de mejores oportunidades económicas, como trabajo, estudios o alguna formación profesional; otros lo hacen por cuestiones personales, como reunirse con sus familias; mientras que otros lo hacen para escapar de situaciones violentas causadas por conflictos, terrorismo o violaciones de derechos humanos; o bien, parten de sus lugares de origen debido al deterioro de las condiciones de vida causado por el cambio climático, desastres naturales u otros factores ambientales. Esta gran diversidad está ampliamente vinculada al grado de desarrollo de los países y su distribución poblacional.

En suma, cada vez una mayor cantidad de personas vive en un país distinto de aquel donde nacieron. Mientras en 1980 los migrantes internacionales comprendían el 2.3 % de la población mundial, en el 2000 este porcentaje fue del 2.8 %. La tendencia siguió siendo creciente, de modo que en 2019 el número de migrantes alcanzó la cifra de 272 millones, 51 millones más que en 2010; es decir, el 3.5 % de la población mundial (OIM, 2020). Asimismo, puede decirse que las tendencias migratorias de las últimas décadas están muy relacionadas a las transformaciones mundiales más amplias en los ámbitos económico, social, político, tecnológico y ambiental.

Debido al impacto cada vez mayor de las migraciones en las sociedades de partida y de acogida, en este capítulo haremos una revisión global de las tendencias migratorias y de sus causas, además de poner en contexto,

la migración internacional de México. Así, en la sección 2 se presenta el concepto de migración y las tendencias actuales en el mundo. En la sección 3 abordamos la migración internacional de México y sus principales características.

## 2 La migración mundial en la actualidad

De acuerdo con la Organización Internacional para las Migraciones (OIM), la migración se define como el “movimiento de población hacia el territorio de otro Estado o dentro del mismo que abarca todo movimiento de personas sea cual fuere su tamaño, su composición o sus causas; incluye migración de refugiados, personas desplazadas, personas desarraigadas, migrantes económicos” (OIM, 2006).

### 2.1 Cambios observados de 2000 a 2020

El informe de la OIM del año 2000 destacaba que el “fenómeno milenario” de la migración se estaba acelerando como parte de las transformaciones más amplias de la globalización económica observada a partir de los años 1980. La cada vez mayor integración de las economías, a través del comercio y de la deslocalización de las diversas etapas de los procesos productivos, ha facilitado tanto el movimiento de bienes como de la fuerza laboral y el capital. En la Tabla 1 se resumen algunos indicadores clave de la migración para los años 2000 y 2020.

Tabla 1: Indicadores clave de la migración, 2000 y 2020

	2000	2020
Migrantes internacionales	150 millones	272 millones
Proporción de migrantes en la población mundial	2.8 %	3.5 %
Proporción de mujeres migrantes internacionales	47.5 %	47.9 %
Proporción de niños migrantes internacionales	16 %	13.9 %
Proporción de trabajadores migrantes	–	60.3 %
Proporción de refugiados	9.3 %	9.5 %

Fuente: OIM (2020). Los valores reportados son estimaciones.

Aun cuando la proporción de migrantes en la población mundial aumentó 0.7 puntos porcentuales, su composición se mantuvo más o menos estable. En particular, las proporciones de mujeres migrantes y de refugiados aumentaron muy ligeramente; por el contrario, la proporción de

niños migrantes bajó en 2.1 %. Finalmente, para 2020 se estima que el 60.3 % la migración corresponde a los trabajadores. No se cuenta con el ese dato para el año 2000, pero en 2017 esa proporción fue del 64 %. Esto indica que la migración laboral es la causa principal de la migración mundial, teniendo amplio impacto en diversos ámbitos de las economías de origen y destino de la fuerza laboral. Por ejemplo, en esas dos décadas las remesas crecieron un 446.8 %, pasando de un total estimado en 126,000 millones de dólares en 2000 a 689,000 millones en 2020.

Las remesas son transferencias de dinero o en especie que los migrantes hacen directamente a sus familias que viven en sus países de origen. El Banco Mundial compila datos mundiales sobre las remesas internacionales. De acuerdo con esos datos (Banco Mundial, 2019), en 2018 los cinco principales países de destino de las remesas fueron (en orden descendente) la India, China, México y Filipinas. Sin embargo, en términos del porcentaje del Producto Interno Bruto representado por las remesas, los cinco principales países de destino en ese año fueron Tonga (35.2 %), Kirguistán (33,6 %), Tayikistán (31 %), Haití (30,7 %) y Nepal (28 %). Por su parte, los lugares de procedencia de las remesas son, casi en exclusiva, los países de más altos ingresos, principalmente los Estados Unidos de América. Para los Estados miembros de las Naciones Unidas, este contexto pone de manifiesto la importancia de la migración internacional como motor del desarrollo (OIM, 2020, capítulo 11).

Asimismo, según Goldin, Cameron y Balajaran (2011), la migración también repercute positivamente en el crecimiento económico de las economías receptoras de distintas formas: los migrantes son en general más jóvenes que la población local, por lo que el PIB per cápita y el PIB global tienden a aumentar; la productividad por trabajador también aumenta con el mayor capital humano; y por último, la migración favorece la productividad total de los factores y la innovación. Todo esto se debe a que los migrantes internacionales no son solo trabajadores, también hay estudiantes, emprendedores, consumidores, contribuyentes fiscales, etcétera, por lo que contribuyen a aumentar la demanda, inversiones e innovación en los países de destino.

## **2.2 Composición de la fuerza laboral migrante internacional**

De acuerdo con estimaciones de la Organización Internacional del Trabajo (OIT, 2018), los trabajadores migrantes internacionales en edad de trabajar —es decir, de 15 años o más, representaron el 70 % de la pobla-

ción migrante internacional en 2017 (234 millones). Dentro de estos, la proporción de hombres fue del 54 % y la de mujeres del 46 %; es decir, hubo una brecha de género en la migración laboral del 8 %. Considerando la población migrante total, la brecha del género fue el doble: 58 % hombres frente a 42 % de mujeres migrantes internacionales. En términos generales, la brecha de género es mayor en los países de ingreso bajo y mediano bajo, que en los países de ingreso alto.

En cuanto a la dimensión geográfica de la migración, se tiene que casi el 61 % del total de trabajadores migrantes, reside en América del Norte; los Estados árabes; y Europa Septentrional, Meridional y Occidental. Por otra parte, la brecha de género entre los trabajadores migrantes es más acentuada en las regiones de Asia Meridional y los Estados árabes. Esta última región es además uno de los principales destinos de los trabajadores migrantes, quienes pueden llegar a dominar algunos sectores clave en esos países, tales como el de la construcción y el del trabajo doméstico. En particular, entre 2013 y 2017, el número de trabajadores migrantes aumentó más de un 5 % en los Estados árabes, dada una mayor demanda de trabajadores migrantes varones, principalmente en el sector de la construcción (OIT, 2018).

### **2.3 La feminización de la migración y la doble barrera en el mercado laboral**

Un fenómeno relativamente reciente es la llamada “feminización de la migración”, que se refiere no solo a la cada vez mayor proporción de mujeres migrantes, sino al cambio de patrones de migración. Esto se refiere a que ya no solo migran como parte de una familia, sino por su cuenta, mayoritariamente como trabajadoras, estudiantes o refugiadas. Sin embargo, enfrentan más dificultades de inclusión en el mercado laboral que los hombres migrantes y que las mujeres nacidas en el país. Esto se debe a que enfrentan dos tipos de barreras: ser mujeres y ser migrantes. En el caso de la Unión Europea (UE), por ejemplo, pese a tener una proporción de mujeres con educación superior similar a la de las mujeres nacidas en el país y a la de los hombres migrantes, las mujeres nacidas fuera de la UE tienen más probabilidades de estar sobrecalificadas para su trabajo y menos probabilidades de tener un empleo. Además, la brecha entre la proporción de mujeres empleadas no nacidas en la UE-28 y las mujeres nativas es 8 puntos porcentuales mayor que la brecha entre los hombres. Esto muestra claramente que las mujeres migrantes representan una fuente de habilidades sin explotar y subutilizada, y que

su potencial no está completamente explotado. Según un informe de la Comisión Europea (Comisión Europea, 2018).

Similarmente, evidencia reciente de la OCDE muestra que las mujeres refugiadas tardan más en establecerse en el mercado laboral en comparación con los hombres refugiados, y cuando están empleadas con frecuencia es en empleos de tiempo parcial. Asimismo, suelen presentar niveles más bajos de habilidades lingüísticas en comparación con los hombres en los primeros dos o tres años después de su llegada; esto se debe a que generalmente reciben menos apoyo de integración que los hombres, tanto en términos de capacitación lingüística como de medidas activas del mercado laboral. Así, a la cada vez mayor afluencia de refugiados a Europa se suma la urgencia de integrar a las mujeres migrantes. En particular, dada la evidencia de que el desarrollo de habilidades básicas en términos de logro educativo y capacitación en el idioma del país anfitrión tiene un alto rendimiento en términos de mejorar los resultados del mercado laboral (Liebig y Tronstad, 2018).

### **3 Migración internacional de México**

México ha sido históricamente el país con más emigración internacional de América Latina y el Caribe; y el segundo país de origen de migrantes del mundo, después de la India. Según datos de la OIM, en 2019 vivían en el extranjero alrededor de 12 millones de mexicanos. La mayoría de los emigrantes mexicanos viven en los Estados Unidos de América y su principal causa de migración fue la laboral. Cabe señalar que este flujo constituye el principal corredor migratorio de país a país de todo el mundo (OIM, 2020).

El flujo migratorio de México hacia Estados Unidos comenzó a mitad del siglo XX, cuando trabajadores mexicanos comenzaron a cubrir la escasez de mano de obra en Estados Unidos debido a la Segunda Guerra mundial. Así, surgió el Programa Braceros (Farm Labor Program para Estados Unidos), el cual, fue celebrado por más de 20 años entre estos dos países permitiendo la entrada de manera legal a trabajadores mexicanos quienes eran ocupados en su gran mayoría en el sector agrícola. Después de haberse concluido el Programa Braceros en 1964, a algunos trabajadores mexicanos les fue permitido permanecer legalmente en suelo estadounidense por el gobierno, mientras que otros trabajadores continuaron cruzando la frontera de manera ilegal en busca de trabajo. El Programa Braceros y la proximidad geográfica entre los países dio inicio a la creación de patrones migratorios entre México y Estados Uni-

dos, que hoy en día conforman el corredor de migrantes más grande del mundo.

De acuerdo con el Censo de Población y Vivienda 2010, de la población emigrante internacional el 89.4 % emigra hacia los Estados Unidos, un 7 % hacia otros países, y el 3.6 % no especificó el lugar de destino. Según datos de la Encuesta Nacional de la Dinámica Demográfica (ENADID), para el periodo de agosto de 2013 a septiembre de 2018, alrededor del 85 % de los mexicanos que emigraron internacionalmente lo hicieron hacia Estados Unidos. Tan solo para este periodo, el total de emigrantes del país ascendió a 760 mil personas y se estima que del total de emigrantes internacionales el 37.7 % regresa a México, mientras que el resto permanece en los Estados Unidos.

## **4 Composición de la migración internacional**

La Encuesta sobre la Migración en la Frontera Norte de México (EMIF Norte) provee información relevante sobre la dinámica demográfica con el país vecino. La EMIF Norte forma parte de las Encuestas sobre Migración en las Fronteras de México (EMIF), la cual se compone de la Encuesta sobre Migración en la Frontera Sur de México (EMIF Sur) y la EMIF Norte. Desde el año 2004, la EMIF Sur capta los desplazamientos migratorios de Honduras, Guatemala y El Salvador hacia México o Estados Unidos, mientras que la EMIF Norte capta desplazamientos migratorios hacia los Estados Unidos y desde los Estados Unidos hacia México. La EMIF Norte se lleva a cabo desde 1993, pero a partir de 1999 se realiza anualmente. La EMIF es una consulta obligada para aquellos que deseamos conocer más sobre los flujos migratorios y sus características en México.

Para contextualizar la migración internacional de México, nos interesa analizar la EMIF Norte. La EMIF Norte se compone de tres flujos migratorios, que a su vez se dividen en subgrupos y de ahí se obtiene la población objetivo. En uno de los flujos migratorios, el desplazamiento va de sur a norte, este flujo capta migrantes procedentes del sur de México y que tienen como destino la frontera norte de México o los Estados Unidos. Los otros dos flujos tienen un desplazamiento de norte a sur. Uno de estos flujos norte-sur, tiene como población objetivo los migrantes que son devueltos por las autoridades migratorias de los Estados Unidos. El segundo flujo norte-sur tiene tres poblaciones objetivo: 1) migrantes procedentes de la frontera norte, 2) migrantes procedentes de Estados Unidos vía terrestre y 3) migrantes procedentes de Estados

Unidos vía aérea. Cada población objetivo se convierte en una encuesta particular, para conocer sobre la composición de la migración internacional de México, nos enfocaremos en aquellos migrantes procedentes de Estados Unidos vía terrestre. Esta base de datos a comparación de la de los migrantes que provienen de Estados Unidos vía aérea, es recolectada anualmente desde 1999 mientras que la de vía aérea comenzó en 2004.

Utilizando un muestreo polietápico,<sup>1</sup> desde principios de 1990, la EMIF Norte ha tenido como principal objetivo determinar la magnitud del flujo migratorio, así como las características de los migrantes. En particular, brinda información en tres aspectos primordiales de los migrantes: características sociodemográficas, características laborales y características migratorias.

#### **4.1 Características sociodemográficas**

Como podemos observar en la Figura 1, a lo largo de estos últimos 19 años el porcentaje de mujeres migrantes ha ido incrementando. A finales del siglo XX y principios de siglo XXI el porcentaje de mujeres oscilaba entre 15 y 18 %, esta cifra fue en aumento y para el 2018 el porcentaje de mujeres migrantes alcanzó el 42 %. Este incremento en el porcentaje de mujeres migrantes no solo se debe a que son acompañantes de algún migrante, sino que cada vez más mujeres toman la iniciativa de migrar para ser el sustento de su hogar, sin embargo, son expuestas a ser violentadas durante el proceso migratorio (Inmujeres, 2005).

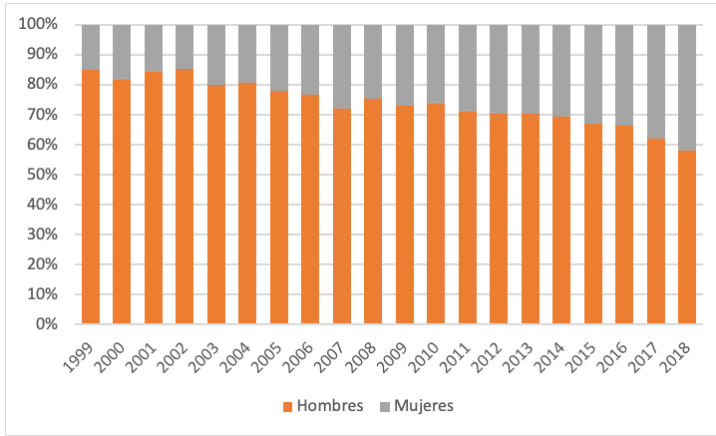
La composición de los migrantes por estatus marital ha permanecido constante a través del tiempo. Los migrantes en su gran mayoría son casados, para el periodo de la encuesta este grupo oscila entre 64-73 %, en el 2018 este grupo llegó a los 65 puntos porcentuales, el 5 % vivía en unión libre y el 8 % estaba separado o divorciado. El segundo grupo más predominante es el de solteros, para el 2018 representaron un 17 %. Aunado a la composición del estatus marital, la posición en el hogar refleja que en su gran mayoría, los migrantes son personas con dependientes económicos. Para el 2018, el 75 % de los migrantes reportó ser el jefe del hogar, cifra que ha permanecido más o menos constante desde 1999. Un indicador muy claro de que los migrantes son el principal sustento de su hogar.

Con respecto al nivel educativo de los migrantes, la Figura 2 muestra como anteriormente los migrantes tenían principalmente primaria o se-

---

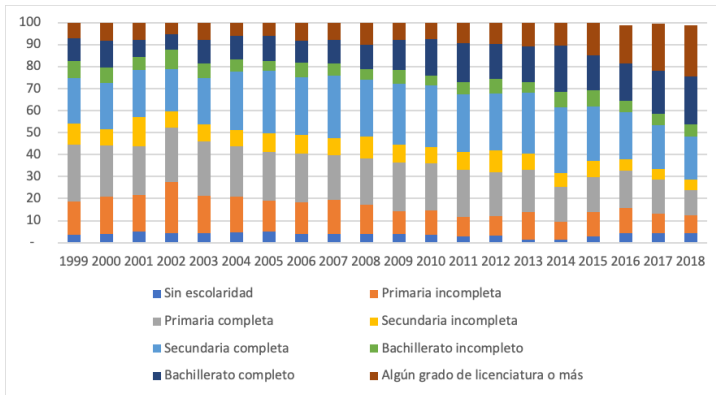
<sup>1</sup> En años recientes, se considera un marco muestral con dos ejes: tiempo y espacio; los cuales en conjunto definen el marco muestral de la encuesta.

Figura 1: Porcentaje de migrantes por género.



Fuente: Elaboración propia con datos de la EMIF Norte.

Figura 2: Nivel educativo de los migrantes.



Fuente: Elaboración propia con datos de la EMIF Norte.

cundaria terminada. Sin embargo, al paso de los años, el nivel educativo de los migrantes ha ido cambiando; en la actualidad hay un menor número de migrantes con primaria o secundaria y un mayor número de migrantes con bachillerato completo, e inclusive con algún grado de li-

cenciatura o más. En conjunto, estos dos últimos grupos suman el 55 % de los migrantes. Tiempo atrás, se pensaba que las personas tomaban la iniciativa de migrar porque no encontraban un trabajo bien remunerado o el salario no alcanzaba para cubrir sus necesidades básicas. Se asociaba esto con el nivel educativo del migrante, que como hemos visto eran en su mayoría personas con baja escolaridad. Hoy en día, personas con alto nivel de escolaridad están tomando la decisión de emigrar del país, lo que mundialmente se conoce como fuga de cerebros. Según una nota de *El Economista*, México encabeza el primer lugar en América Latina en la fuga de cerebros, seguido de Colombia y Cuba. También se menciona que uno de los canales que ha intensificado la fuga de cerebros es la educación en el extranjero, puesto que los mexicanos al momento de estudiar algún posgrado deciden hacerlo en el extranjero y otros más reciben ofertas de trabajo en otros países. Esto también se ve reflejado en el porcentaje de migrantes que hablan inglés. Este indicador comenzó a recolectarse a partir del año 2004, en el cual, el 54 % de los migrantes reportó que hablaba inglés, porcentaje que fue en aumento y para el 2018 el porcentaje de migrantes que reportó hablar inglés fue del 78 %.

Tabla 2: Grupos de edad de los migrantes.

	2000	2010	2018
Total %	100	100	100
15 a 19 años	2.80	2.99	0.84
20 a 24 años	9.55	8.66	4.81
25 a 29 años	14.35	12.60	7.86
30 a 34 años	15.36	12.15	8.40
35 a 39 años	15.67	16.54	10.78
40 a 44 años	12.99	12.72	10.73
45 a 49 años	10.30	11.04	14.14
50 a 54 años	7.97	7.79	12.79
55 a 59 años	3.88	6.03	9.83
60 a 64 años	3.27	3.49	8.36
65 años o más	3.86	5.95	11.33
No especificado	0.00	0.03	0.13

Fuente: Elaboración propia con datos de la EMIF Norte.

Otro aspecto sociodemográfico que es importante analizar es la edad de los migrantes. Para el año 2000, la EMIF Norte registró que prácticamente un 70 % de los migrantes tenía entre 25 y 49 años, para el 2010 este grupo comprendió un 65 %, mientras que para el 2018 bajó a un 52 %. Del 2000 al 2018 los grupos de 25 a 29 y 30 a 39 años se redujeron casi a la mitad, mientras que los grupos de migrantes entre 55-59, 60-64, y 65 años y más incrementaron casi tres veces su porcentaje. Esto sugiere que hay un cambio estructural en la composición por edad de los migrantes. La tabla 2 muestra el porcentaje por grupos de edad para los tres años seleccionados.

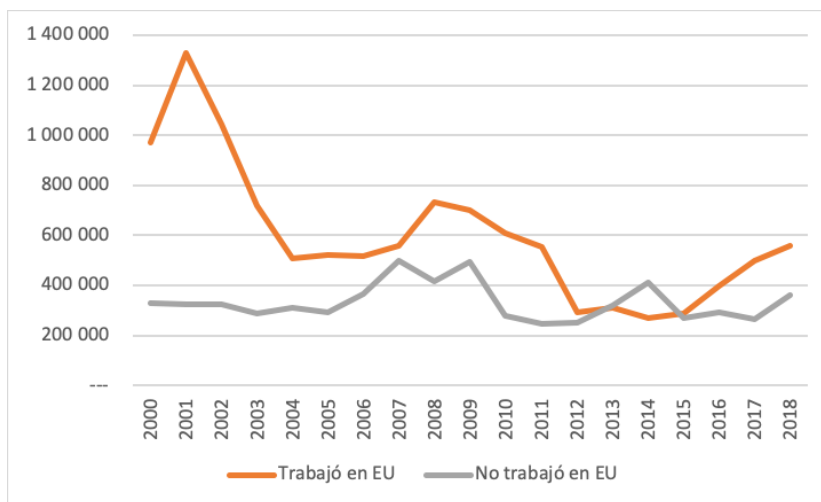
## 4.2 Características laborales

Respecto a las características laborales de los migrantes en Estados Unidos, la Figura 3 muestra el número de dos grupos de migrantes, los que laboraron y los que no laboraron en el país vecino del norte.<sup>2</sup> Podemos observar que el número de migrantes que no trabajó en los Estados Unidos se mantuvo relativamente constante oscilando alrededor de los 300,000 migrantes, los años en que se reportó un mayor número de migrantes en este rubro fue del 2007 al 2009. Este incremento podría atribuirse a la recesión económica que inició para ese periodo. En el caso de los migrantes que reportaron haber trabajado en los Estados Unidos, vemos cambios más drásticos. En el año 2001 fue donde se registró un mayor número de migrantes trabajadores, alrededor de 1.3 millones. Luego continuó una tendencia a la baja, la cual llegó a su punto más bajo en el 2014 con un poco más de 270 mil migrantes que habían trabajado. Hay que tener en perspectiva, que estos números se refieren a migrantes que trabajaron durante su último viaje a Estados Unidos, que por alguna razón están regresando a México y que pueden estar residiendo en cualquiera de los dos países. Otro aspecto que debemos tener presente es que estas cifras no corresponden al universo de migrantes que trabaja en Estados Unidos, sino que es una muestra. Conocer el número de migrantes mexicanos que trabajan en los Estados Unidos se torna una tarea difícil. Los datos reportados por el censo de Estados Unidos contemplan una parte de estos migrantes, sin embargo, dado que un gran número de migrantes se encuentra en Estados Unidos de forma ilegal, estas cifras están subestimadas. En el último censo de Estados Unidos (2010) se registró que había 31.7 millones de personas

---

<sup>2</sup> Aquellos que no laboraron puede ser porque permanecieron muy poco tiempo en Estados Unidos o porque buscaron y no encontraron trabajo.

Figura 3: Condición laboral de migrantes.

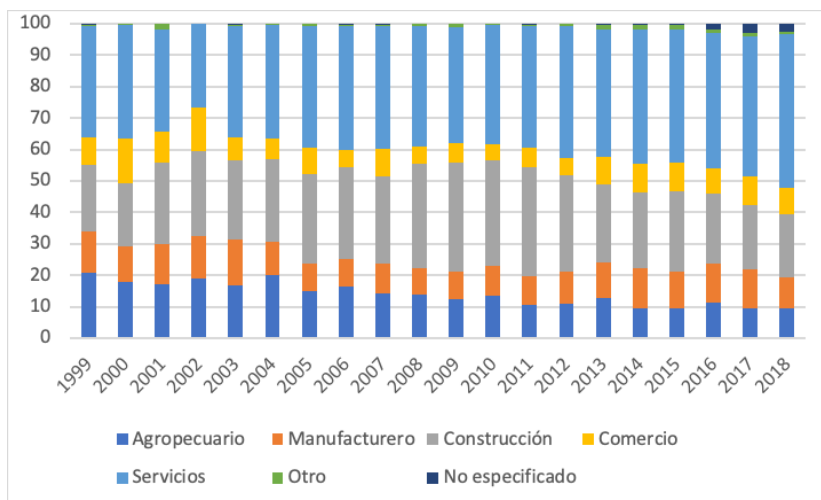


Fuente: Elaboración propia con datos de la EMIF Norte.

de origen mexicano, representando un 10.3 % del total de la población en dicho país. Para ese mismo año, en un artículo publicado por el Instituto de Política Migratoria (MPI, por sus siglas en inglés) reportó que el número de inmigrantes mexicanos era de 11.7 millones. En el 2018, la Current Population Survey (CPS, por sus siglas en inglés) estimó que 38.5 millones de personas que residen en Estados Unidos tienen origen mexicano y que 12.3 millones son migrantes mexicanos.

Respecto a los migrantes que trabajaron en Estados Unidos, podemos conocer el sector de ocupación, así como su posición en el trabajo y salario. La Figura 4 muestra la evolución de los sectores en que han laborado los migrantes. En 1999 el sector con un mayor número de trabajadores fue el de servicios con un 35 %, seguido de los sectores agropecuario y de construcción con 21 % cada uno. El sector servicios continuó siendo el sector preferido de lo migrantes a lo largo de los años, para el 2018 casi el 50 % de los trabajadores laboraron en este sector. Por el contrario, el sector agropecuario se redujo en doce puntos porcentuales para este mismo periodo. El sector de construcción permaneció relativamente constante, solo tuvo un auge en el periodo 2008-2012, pero para 2018 un 20 % laboró en este sector. Sectores como el manufacturero y comer-

Figura 4: Sector de ocupación.



Fuente: Elaboración propia con datos de la EMIF Norte.

cio se mantuvieron estables. La reducción del sector agropecuario y el incremento en el sector servicios, obedece en cierta medida a la globalización. Hoy en día, la fuerza laboral de los países desarrollados y de algunos países en desarrollo se concentra principalmente en el sector servicios mientras que el sector agropecuario emplea menos trabajo y mayor capital.

La gran mayoría de los migrantes reportó ser trabajador de sueldo fijo, para el 2018 el 74 % de los migrantes así lo reportó. La segunda posición en el trabajo fue trabajador a destajo o por obra con casi un 12 %, esta posición ha ido en decremento ya que para el 2000, el porcentaje de trabajadores a destajo era de 27 %. Ser trabajador por cuenta propia o patrón, son dos puestos que han tomado mayor relevancia al paso de los años, para 1999 representaban el 2.5 y 1.3 % respectivamente, mientras que para el 2018 fue de 8.7 y 4.3 %.

Finalmente, uno de los aspectos más importantes para los migrantes es el salario percibido por su trabajo. En el periodo analizado ha habido grandes cambios, inclusive, si se incluyen los datos de 1995, los cambios son más drásticos. La tabla 3 presenta el ingreso por rangos para años seleccionados. En 1995 se registró que casi un 50 % de los migrantes

que trabajaron recibían un ingreso mensual menor a mil dólares, y casi un 40 % percibía entre 1,000 a 1,999 dólares. Esto quiere decir que un 90 % de los migrantes percibía un salario mensual menor a dos mil dólares, cinco años después este porcentaje representaba un 70 %, es decir, un 20 % se distribuyó en rangos con un salario mayor. En el 2005 un 70 % reportó ganar entre 1,000 a 2,499 dólares mensuales, para el 2010 un 16 % reportó ganar más de 3,000 dólares y este porcentaje se duplicó para el 2018. Para este último año que se tiene registro, solo el 2 % reportó un salario menor a mil dólares mensuales, sin embargo, un gran porcentaje (33 %) no especificó su ingreso para este año por lo que debemos interpretar estos resultados con cautela.

Tabla 3: Ingreso mensual por rangos para años seleccionados.

	1995	2000	2005	2010	2018
Total	100 %	100 %	100 %	100 %	100 %
Menos de 1 000 dólares	48	19	6	4	2
1 000 a 1 499 dólares	28	31	29	19	4
1 500 a 1 999 dólares	11	20	21	25	7
2 000 a 2 499 dólares	4	11	17	21	14
2 500 a 2 999 dólares	3	3	5	10	8
3 000 o más dólares	5	10	13	16	32
No especificado	1	6	8	5	33

Fuente: Elaboración propia con datos de la EMIF Norte.

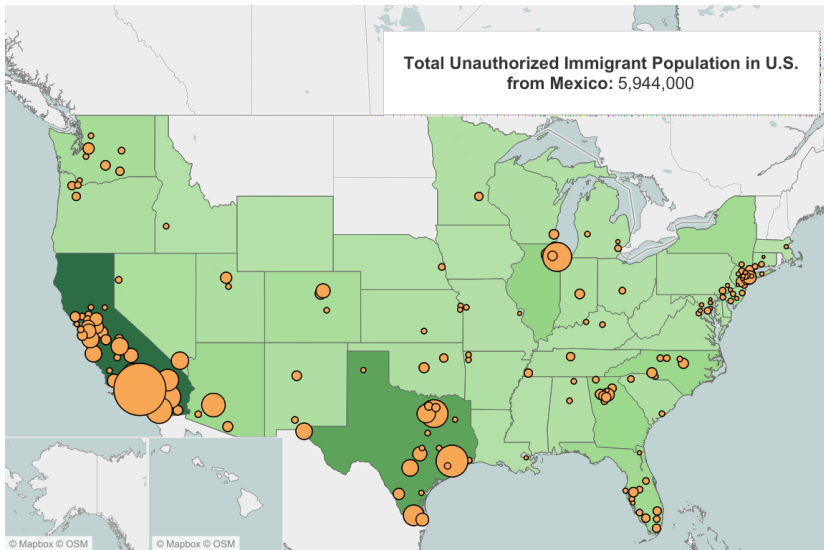
### 4.3 Características migratorias

Las ciudades que utilizan los migrantes para cruzar a los Estados Unidos son principalmente las ciudades más grandes a lo largo de la frontera norte de México. Las ciudades más populares que fueron reportadas en el 2018 fueron Tijuana (Baja California), Nuevo Laredo (Tamaulipas), y Ciudad Juárez (Chihuahua). Contrario a los que muchos piensan, en la actualidad los migrantes que entran a Estados Unidos de manera ilegal son un porcentaje relativamente bajo, para el 2018, el 17 % de los migrantes reportó haber cruzado la frontera sin tener los documentos legales para hacerlo. Un 50 % cruzó con documentos legales, y el resto no estaba en busca de trabajo o había cruzado desde pequeño a los Estados Unidos. Además, un 85 % de los migrantes reportó tener documentos

legales para trabajar en los Estados Unidos. Otro concepto erróneo que pudiéramos tener de la migración es el uso de pollero o coyote, en el año 2006 fue cuando se reportó el mayor porcentaje de migrantes que utilizaron este medio de cruce (22%). En el 2018, solo el 12.5% de los migrantes utilizaron un pollero para llegar al país vecino.

Los estados en donde la gran mayoría de los migrantes mexicanos se establecen se encuentran al suroeste de Estados Unidos: Arizona, California, Nuevo México y Texas. Estos cuatro estados concentran más del 80% de los migrantes registrados en la EMIF Norte; en especial, en California se concentra el 60%. La Figura 5 muestra un mapa con los estados y condados con una mayor concentración de mexicanos ilegales, siendo Los Ángeles, California el condado con un mayor número de migrantes, para el periodo 2012-2016 había 633 mexicanos viviendo de manera ilegal en dicho condado. Tan solo en el estado de California había poco más de dos millones de mexicanos indocumentados para ese periodo.

Figura 5: Migrantes mexicanos ilegales en Estados Unidos por estado y condado, 2012-2016.



Fuente: Migration Policy Institute. <https://www.migrationpolicy.org/programs/data-hub/charts/unauthorized-immigrant-populations-country-and-region-top-state-and-county>.

Una parte importante sobre la migración son las remesas enviadas a su lugar de origen. Como se mencionó anteriormente, México es uno de los principales países en el mundo que recibe remesas. En 2018 las remesas alcanzaron su punto máximo, con monto de casi 33.5 millones de dólares. Esta cifra representó a su vez el mayor grado de dependencia de las remesas en México, representando el 2.7% del PIB en México para dicho año (Conapo, 2019). El estudio de estos flujos monetarios (y en especie) requiere un capítulo aparte. En este capítulo nos limitamos a resaltar la magnitud de dicho flujo y la importancia que tiene para las familias receptoras, quienes utilizan este dinero para cubrir principalmente necesidades básicas.

## Referencias

- [1] Banco Mundial (2019). Migration and Remittances: Recent Developments and Outlook. Migration and Development Brief 31. Recuperado de [www.knomad.org/sites/default/files/2019-04/Migrationanddevelopmentbrief31.pdf](http://www.knomad.org/sites/default/files/2019-04/Migrationanddevelopmentbrief31.pdf)
- [2] Comisión Europea (2018). Integration of migrant women: A key challenge with limited policy resources. Recuperado de <https://ec.europa.eu/migrant-integration/feature/integration-of-migrant-women>
- [3] Consejo Nacional de Población, Conapo (2019). Anuario de Migración y Remesas México 2019. Séptima edición. México.
- [4] Current Population Survey, CPS (2018). Hispanic Population. Recuperado de <https://www.census.gov/programs-surveys/cps/data.html>
- [5] El Colegio de la Frontera Norte, Unidad de Política Migratoria, Registro e Identidad de Personas, Consejo Nacional de Población, Consejo Nacional para Prevenir la Discriminación, Secretaría del Trabajo y Previsión Social, Secretaría de Relaciones Exteriores, Secretaría de Bienestar (Bienestar), Encuesta sobre Migración en la Frontera Norte de México, [www.colef.mx/emif](http://www.colef.mx/emif)
- [6] El Economista (27, Diciembre, 2018). México encabeza fuga de cerebros en Latinoamérica. El Economista. Recuperado de <https://www.eleconomista.com.mx/arteseideas/Mexico-encabeza-fuga-de-cerebros-en-Latinoamerica-20181227-0075.html>

- [7] Encuesta Nacional de la Dinámica Demográfica, ENADID (2018). Tabulados principales: Migración Internacional. <https://www.inegi.org.mx/programas/enadid/2018/>
- [8] Goldin, I., Cameron, G. y M. Balarajan (2011). *Exceptional People: How Migration Shaped Our World and Will Define Our Future*. Princeton University Press, Princeton.
- [9] Instituto Nacional de las Mujeres, INMujeres (2005). *Mujeres migrantes y sus implicaciones desde la perspectiva de género*. Inmujeres, Primera edición. México.
- [10] Instituto Nacional de Estadística y Geografía, INEGI (2010). *Censo de Población y Vivienda 2010*.
- [11] Liebig, T. y K. Tronstad (2018), "Triple Disadvantage?: A first overview of the integration of refugee women", OECD Social, Employment and Migration Working Papers, No. 216, OECD Publishing, Paris, <https://doi.org/10.1787/3f3a9612-en>
- [12] Organización Internacional para las Migraciones, OIM (2006). *Glosario Sobre Migración. Derecho Internacional Sobre Migración No. 7*. Recuperado de [https://publications.iom.int/system/files/pdf/iml\\_7\\_sp.pdf](https://publications.iom.int/system/files/pdf/iml_7_sp.pdf)
- [13] Organización Internacional para las Migraciones, OIM (2020) *Informe sobre las migraciones en el mundo 2020*. Recuperado de [https://publications.iom.int/system/files/pdf/wmr\\_2020\\_es.pdf](https://publications.iom.int/system/files/pdf/wmr_2020_es.pdf)
- [14] Organización Internacional del Trabajo, OIT (2018). *ILO Global Estimates on International Migrant Workers – Results and Methodology*. Segunda edición. Recuperado de [www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms\\_652001.pdf](http://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_652001.pdf)
- [15] United States Census Bureau, (2010). *Censo decenal*. Recuperado de <https://data.census.gov/cedsci/table?q=mexican&tid=DECENNIALSF12010.PCT11>
- [16] Zong J., Batalova J. (2018). *Mexican Immigrants in the United States*, Migration Policy Institute. Recuperado de <https://www.migrationpolicy.org/article/mexican-immigrants-united-states>

# Redes sociales

José Antonio Carrillo Viramontes

Lari Arthur Viianto

Alejandro Tatsuo Moreno Okuno

## 1 Introducción

**H**oy en día, al escuchar “redes sociales” la gente suele vincular estas palabras con las plataformas de internet que nos facilitan una comunicación fluida e instantánea con personas de todo el mundo, pero en realidad estas son meras herramientas que permiten una comunicación diferente a la tradicional, de forma que uno se comunica a distancia, de manera instantánea, con un foro amplio. Sin embargo, cuando hablamos de red social nos referimos a algo más complejo, más profundo, y mucho más antiguo.

Los individuos nos organizamos de manera natural en sociedades, siendo estas de muy diversas maneras y tamaños, cada una con sus características propias y particularidades, pero con similitudes considerables. Estas sociedades, entre otras cosas, están basadas en los diversos tipos de relaciones e interacciones que surgen entre los individuos que forman parte de ellas. Las redes sociales son una representación de dichas relaciones y por lo tanto son una representación parcial de la estructura y el capital sociales subyacentes. Si bien, siendo parciales, no expresan toda la complejidad de la estructura, son elementos que permiten entender mejor muchos de los complejos mecanismos que vertebran la sociedad y permite el análisis de diversos efectos tanto, ya sean estos sociales, económicos, psicológicos, etcétera, con una perspectiva y una profundidad mayor a la de otros métodos de análisis.

Desde la economía clásica se considera al individuo de manera aislada, cuya interacción con el resto de la sociedad (otros individuos igualmente aislados) se produce simplemente a través de intercambios económicos

realizados en un mercado, considerado este como un entorno aséptico e impersonal, que garantiza el anonimato y la ausencia de cualquier interacción social significativa, más allá del mero intercambio de determinada cantidad a un determinado precio. Esta simplificación tiene mucho sentido para una basta particularidad de escenarios, donde las relaciones e interacciones quedan efectivamente relegadas a un segundo plano, y se realizan de manera anónima e impersonal, primándose el resultado beneficioso del intercambio de manera individual. Sin embargo, a la hora de describir el funcionamiento complejo de una sociedad, conviene incluir en el análisis las relaciones que las personas que forman la sociedad mantienen entre sí.

No somos individuos aislados, y mucho menos somos individuos impersonales, gran parte de nuestro bienestar (el elemento principal del estudio económico) se deriva de las relaciones que guardamos entre nosotros. No solo un bienestar desde un punto de vista general y ambiguo, como puede ser un bienestar emotivo, social, familiar, afectivo, sino que nuestro bienestar entendido desde el punto de vista económico clásico (ingreso, consumo, inversión, educación...) en realidad depende en gran manera de estas interrelaciones personales y el lugar que ocupamos en el entramado social que se expresa en las redes sociales.

Como seres sociales que somos, nos interrelacionamos entre nosotros de diversas maneras, por lo que existen diferentes redes en función del tipo de interrelación o interacción que deseemos expresar, y normalmente la funcionalidad de estas redes es diferente, por lo que se suelen representar un subconjunto de relaciones (o incluso un solo tipo de relación), en vez de representarlas todas de manera conjunta. Lo cual también es una simplificación de la complejidad que la red social tiene en realidad.

Lo que es indiscutible es que estas redes sociales existen, que nos relacionamos entre nosotros no de manera aleatoria, sino siguiendo patrones determinados, y con el claro deseo de establecer lazos más duraderos y buscar una continuidad en el tiempo de dichas interacciones.

La construcción de dichas redes sociales es un proceso natural, que realizamos en la mayor parte de casos de manera casi inconsciente, por lo que en muchos casos no existe una intención estratégica a la hora de generar estas interacciones que derivaran en la construcción de nuestra red social. Cabe decir también que la mayor razón para la construcción de dicha red es de carácter emocional y de necesidad psicosocial. Esto quiere decir que a pesar de que la red es una influencia importante en diversos aspectos de nuestras vidas, entre ellos muchos aspectos económicos, la finalidad de esta nunca fue la obtención de estos beneficios,

que ocurren de manera paralela, sino que la finalidad es esta cuestión emotiva y psicosocial. En el caso de que una persona construya sus redes sociales con finalidades económicas, o la búsqueda de otros beneficios materiales o inmateriales, normalmente es rechazado por la sociedad, en caso de ser descubierto, como una persona interesada, arribista, falsa, deshonesta... por lo que socialmente se considera una conducta no deseable y reprochable.

Lo anterior es de tremenda importancia para entender que a pesar de la relevancia de las redes para los diversos aspectos que se enumeran en este capítulo, de inicio consideraremos que en la mayoría (o la totalidad) de los casos, la construcción de dicha red no se realizó con la finalidad de estos, por lo que por definición, no existe una eficiencia en términos del efecto expresado, y que resulta inútil la construcción de una red eficiente en términos de alguno de estos beneficios, o incluso en términos de la combinación total de beneficios, si no incorporamos de manera expresa el elemento emotivo y psicosocial, algo que a estas alturas del desarrollo de las ciencias sociales resulta completamente imposible.

## **2 Representación de las redes sociales**

Para representar estas redes sociales, se requiere especificar primero qué tipo de relación o relaciones son las que van a ser representadas y cómo las definimos. Entre las relaciones más habituales se encuentran:

**Familiares:** Como seres humanos, no venimos solos al mundo, sino que nacemos en el seno de una familia (entendida desde un punto de vista amplio, no necesariamente consanguíneo). Este es nuestra red social básica, que representa nuestras primeras interacciones con el resto de la sociedad, y de las cuales heredaremos muchos rasgos que nos definen, como el lenguaje, la cultura inicial. De ella obtendremos mucho apoyo de manera multidimensional (afectivo, psicológico, histórico, económico, cultural... ). Si bien podemos definir la familia de manera legal, desde un aspecto social este concepto puede ser más complejo y no toda la familia legal es familia social, y personas que no son familiares desde un punto de vista legal, pueden ser familia desde un punto de vista de la influencia social que representan. Si bien nuestras relaciones dentro de la familia dependen parcialmente de nosotros mismos, la familia no la podemos escoger, y viene determinada por fuerzas externas.

**Amistades:** El siguiente elemento que construimos como vínculo con la sociedad son las amistades, personas externas a la familia con las que nos relacionamos y empezamos a compartir las primeras experiencias fuera

del seno y la protección familiar. Normalmente personas con las que compartimos algún espacio común, como el patio, el vecindario, el colegio. También tienen una fuerte influencia en la formación de nuestra identidad y gustos. Son una de las más importantes fuentes de información que se transmite a través de redes, sobre todo en escenarios de transmisión de ideas, conceptos, consumo (el boca a boca es muy estudiado en marketing) e incluso información laboral.

**Conocidos y compañeros:** De manera igual al anterior, pero con una carga emotiva inferior, normalmente se llegan a considerar de manera conjunta. En muchas ocasiones, la red de conocidos, al ser más amplia y menos cerrada, es el canal de difusión de información por excelencia.

**Colegas:** Personas con las que compartimos algún tipo de vínculo laboral, en la que se establece la cooperación mutua para la consecución de un objetivo marcado de mutuo beneficio, siendo el caso de los coautores uno de los casos más estudiados. El vínculo consiste en el trabajo compartido.

**Vecinos:** Personas con las que compartimos un espacio físico y que pueden desarrollar otras relaciones, como difusión de información, pequeños favores, prestamos, etcétera. Incluye los conceptos de compañeros de piso y similares.

**Parejas:** Personas con las que hemos establecido un vínculo afectivo o sexual, muy estudiado en los casos de enfermedades de transmisión sexual.

**Clientes y proveedores:** Utilizado sobre todo al analizar las relaciones entre empresas, para estudiar los sistemas de distribución y logística.

**Propiedad:** Análisis de las estructuras de propiedad entre empresas, como son holdings, carteles, la integración horizontal y vertical, y conceptos como el Keiretsu o el Chaebol.

**Deudores:** Relaciones económicas de préstamos, no necesariamente monetarios, que normalmente involucran algún tipo de familiaridad o conocimiento mutuo, pero no necesariamente.

**Otras:** Cualquier otra relación que se pueda definir y de la que se pueda tener constancia.

Como vemos, estas relaciones se pueden establecer entre personas, familias, empresas o incluso instituciones. Y si bien la familia no la podemos escoger, el resto de las relaciones dependen de nosotros mismos, pero también de los demás. El establecer alguna de estas relaciones requiere de la aceptación de ambas partes, pero el terminar alguna de estas se puede hacer de manera unilateral. Esto se conoce como *pairwise stabi-*

*tily*, para establecer un link este debe ser mutuamente beneficioso. Para deshacer un link, basta con que una de las partes no vea interés en este.

La diferencia básica entre el estudio de las redes sociales y otras redes es la característica de que los actores de una red social toman decisiones sobre la misma, construyendo la red de mutuo acuerdo. En el estudio de otros tipos de redes (internet, carreteras, redes neuronales. . .) existe un diseño establecido desde el exterior de esta, y tiene sentido hablar de la eficiencia de estas redes en cuanto a que existe un objetivo externo para el que fueron concebidas. Este no es el caso de las redes sociales, donde los participantes construyen de manera parcial su propio entorno.

A estos agentes capaces de decidir los llamaremos nodos, y son los actores principales de una red. Las relaciones que se establecen entre estos actores de mutuo acuerdo los llamaremos links. De manera básica estos son los dos únicos elementos necesarios para la construcción de una red social, actores (nodos) y relaciones (links).

Para enriquecer el modelo se pueden caracterizar tanto los nodos como los links, con características, habilidades, datos o cualquier otra información que se considere relevante.

Los links se establecen como relaciones binarias, o sea, relaciones entre dos, y solo dos, nodos. Esto limita en parte posibles relaciones donde una única relación incluya mas de dos personas, pero en la realidad hay pocos ejemplos prácticos de links no binarios.

Los links pueden ser direccionales o no. Si son direccionales implican que la relación funciona en solo un sentido (por ejemplo, el flujo de información), si son no-direccionales, se asume que la relación funciona igual en ambos sentidos. Lo habitual en las ciencias sociales es asumir links no-direccionales, pero es muy común en redes de computadoras, redes logísticas, etcétera, que exista direccionalidad.

Los links podrían tener pesos, que reflejarían la importancia relativa de la relación respecto de otras, pero igualmente, por simplicidad, en las ciencias sociales es habitual no establecer pesos. Sin embargo, estos pesos en la práctica llegan a ser importantes.

## **2.1 Representación gráfica**

De manera gráfica los nodos se suelen representar como círculos, y los links como líneas que unen dos de estos círculos, con flechas si son direccionales, sin flechas para los no direccionales (Figuras 1 y 2).

Identificamos los nodos por números, y dada esta red, 1 guarda relación con 4 y 2, 2 con 1 y 5, 3 con 4 y 1, 4 con 1 y 3 y 5 con 2.

Figura 1: Red no direccional.

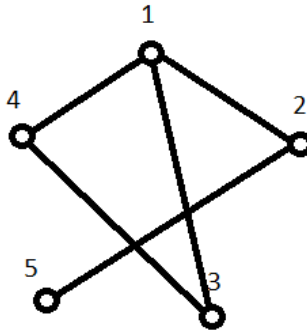
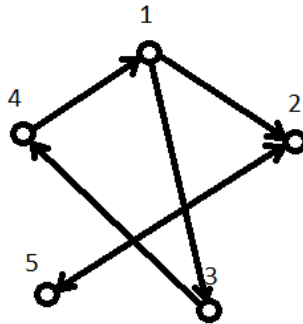


Figura 2: Red direccional.



De acuerdo con esta estructura de red, 1 guarda relación con 2 y 3, 2 con 5, 3 con 4, 4 con 1 y 5 con 2.

La representación gráfica es útil para redes pequeñas, mostrando una visualización clara de la estructura, pero es poco práctica para cálculos y análisis de la red, sobre todo si la red es extensa. Para el análisis de redes es mejor una representación matricial.

## 2.2 Representación matricial

En la representación matricial utilizamos una matriz cuadrada, donde las filas y las columnas corresponden con los nodos, la posición  $ij$  de la matriz indica una relación del nodo  $i$  con el nodo  $j$ . Si existe una relación, la posición  $ij$  toma el valor 1, si no existe, tomará el valor 0. En este punto sería posible establecer valores entre 0 y 1 para dar peso a las relaciones, pero complicaría todos los cálculos posteriores. La representación matricial de la Figura 1 se muestra en la Tabla 1.

Tabla 1: Representación matricial de la red de la Figura 1.

0	1	1	1	0
1	0	0	0	1
1	0	0	1	0
1	0	1	0	0
0	1	0	0	0

Si la red es no-direccional, la matriz será simétrica respecto de la diagonal, si es direccional puede no serlo. Nótese que cada nodo no se relaciona consigo mismo, y la diagonal es de ceros.

La representación matricial facilita el trabajo de análisis de redes, y su uso es recomendable para representar redes extensas.

Aun así, ver una matriz cuadrada de  $n \times n$ , siendo  $n$  muy grande, tampoco es muy informativo, pero partiendo de la matriz, es fácil calcular elementos sencillos para facilitar una síntesis de la información de la red. Algunos estadísticos básicos fáciles de calcular son:

- Número de nodos:  $n$ , número de actores en la red.
- Número de links:  $l$ , número de relaciones que se establecen en la red.
- Grado: Número de relaciones que tiene cada nodo (simplemente sumando filas o columnas).
- Distribución de grado: Cómo se distribuye el número de grado entre los nodos.
- Grado medio: Número medio de relaciones en la red.
- Densidad de red: Número de links que existen / Máximo número posible de links.

- Clustering: Qué tan común es que, si dos personas se relacionan con una tercera, estén también relacionados entre sí (¿Los amigos de mis amigos son mis amigos?) Es una medida de densidad local.
- Distancia: Qué tantos links existen entre dos nodos.
- Distancia media: En media, cuál es la distancia medida en links entre dos individuos cualquiera.
- Diámetro:Cuál es la distancia entre los dos nodos más alejados entre sí.

Son valores numéricos simples, de descripción estadística de la red, que son muy fáciles de calcular para una representación matricial, siempre que esté formada por 0 y 1. Son datos que analizan la red compleja y comprimen información valiosa para entender la red.

Esta red de relaciones y actores sirve para transmitir información, adopción de nuevas tecnologías, formación de hábitos de consumo, construcción de identidad, préstamos, obtención de recursos, palancas, influencias, entre otros.

A continuación, algunos ejemplos sobre las aplicaciones de las redes sociales.

### **3 Formación estratégica de redes**

La formación estratégica de redes es un tema relevante en la actualidad. La mayoría de nuestras decisiones sociales sobre con quién nos relacionamos o no, es decir, con quiénes formamos una red social, están determinadas por las personas involucradas y no por una regla arbitraria. De esta manera, la formación estratégica de redes se encarga de analizar y definir las formas particulares que toman las redes sociales, es decir, por qué y cómo las redes sociales han tomado ciertas formas.

El enfoque económico sugiere que, en la formación estratégica de redes, las personas o agentes representados por nodos, crearán relaciones (links) que son benéficas y eliminarán aquellas que no lo son. De acuerdo con Jackson (2018), existen dos aspectos cruciales para el modelamiento de redes desde un punto de vista estratégico. El primero, refiere a modelar de manera explícita los costos y beneficios que surgen a través de diferentes redes sociales. Esta modelación representa dos ventajas. Una de ellas es que permite conocer los incentivos que tienen los agentes

para formar links, y segundo esta modelación permite medir el bienestar social total. Esto permite tener predicciones acerca de cuáles son las mejores redes para toda la sociedad.

El segundo aspecto crucial es la predicción sobre los incentivos individuales que conllevan a la creación de redes estratégicas. En otras palabras, de qué manera incentivos individuales terminan por explicar el comportamiento agregado de una red social. Quizás usted se preguntará, ¿cómo la estructura de una red social puede tener efectos en el comportamiento individual y colectivo?

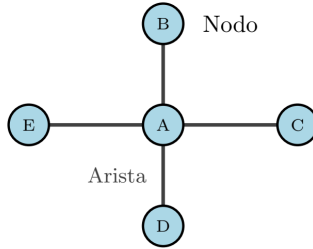
Existen algunos ejemplos que puedan responder esta pregunta. En el mercado laboral podemos encontrar a dos trabajadores con la misma habilidad, sin embargo, estos dos trabajadores pueden tener diferentes salarios. La razón quizás sería el número de conexiones personales (links) que podrían tener un trabajador. En este caso el trabajador con una red social más grande, es decir, mayores conexiones podrían recomendarle un mejor empleo y por lo tanto un mayor salario. Por lo tanto, estas ventajas de tener una red social más grande reflejarían la necesidad de individuos de formar conexiones con otros (Goyal, 2007).

En los párrafos anteriores hemos mencionado algunos conceptos importantes que son necesarios definirlos. Por ejemplo, red social. Si bien existen varias definiciones, podemos entender una red social como una estructura social compuesta por un conjunto finito de agentes (individuos, empresas, etcétera) y configurada en torno a una serie de relaciones entre ellos, que se puede representar en forma de uno o varios grafos. Los grafos se componen de nodos que representan agentes, y aristas (links o conexiones) que representan las relaciones entre ellos. Un tercer elemento de las redes es el límite (boundary) de la red (Aguirre, 2011).

Como podemos observar en la Figura 3, tenemos una red social la cual consiste en 5 nodos (A, B, C, D y E) y cuatro aristas o links entre los nodos. Los nodos pueden representar individuos, empresas, estados, países etcétera. Por otra parte, las aristas o links pueden representar relaciones entre ellos.

Para modelar matemáticamente la formación de redes necesitamos capturar los incentivos de los individuos. Usualmente la manera de capturar dichos incentivos es mediante una función de utilidad o pagos que cada agente recibe y que dependen de la red. Formalmente podemos representar los nodos de una red como  $N = \{1, \dots, n\}$ . Como se menciona anteriormente, la manera en que se modelan los incentivos es mediante una función de utilidad o pago representada por  $u_i : G(N) \rightarrow \mathbb{R}$ . Donde

Figura 3: Nodos y aristas.



$u_i(g)$  representa la utilidad o el beneficio neto del agente  $i$  cuando la red social en la que se encuentra es  $g$ .

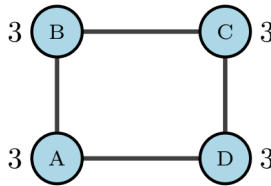
Existen aspectos importantes a destacar de esta función de utilidad  $u_i(g)$ . Uno de ellos es que el agente puede conocer la utilidad de añadir o perder un link o conexión en la red o actuar para incrementar su utilidad, es decir, añadir links o quitar links que disminuyan su utilidad (Jackson, 2008). Para poder analizar la formación de redes, es necesario partir del supuesto de que dos agentes forman un link dentro de la red solo si se tiene consentimiento mutuo. El concepto que cumple con este supuesto es el de *pairwise stability* o estabilidad de pares (Jackson & Wolinsky, 1996).

Formalmente el concepto de *pairwise stability* se representa de la siguiente forma:

1. para todo  $ij \in g$ ,  $u_i(g) \geq u_i(g - ij)$  y  $u_j(g) \geq u_j(g - ij)$ , y
2. para todo  $ij \notin g$ , si  $u_i(g + ij) > u_i(g)$ , entonces  $u_j(g + ij) < u_j(g)$ .

Intuitivamente, una red social es estable entre pares (*pairwise stability*) si ningún agente quiere eliminar un link y ningún par de agentes quiere añadir un link. La primera parte de la definición formal sugiere que un agente no tiene incentivos para eliminar un link si su utilidad es mayor de tener ese link. Además, implica que cada agente tiene el derecho unilateral de eliminar un link en el cual se encuentre participando. La segunda parte tiene dos implicaciones. La primera sugiere que si se añade un link por fuera de la red entonces un agente estaría mejor con ese link, pero otro tendría una pérdida. La segunda implicación sugiere

Figura 4: Red estable.



que la creación de un link dentro de la red  $g$  beneficia a ambos agentes con al menos uno de ellos con beneficio estricto, entonces la red  $g$  es no estable (Jackson 2008).

Veamos la Figura 4, que muestra una red estable entre pares.

Como podemos observar en la Figura 4, existen 4 nodos (A, B, C y D) y 4 aristas en la red social. El número al lado del nodo corresponde a el pago o utilidad de cada agente, es decir, cada agente tiene una utilidad de 3 debido a las dos aristas que cada nodo tiene con respecto a otro agente. En este caso, la red es *pairwise stable* o estable entre pares ya que ningún agente obtendría beneficio al eliminar o añadir un link.

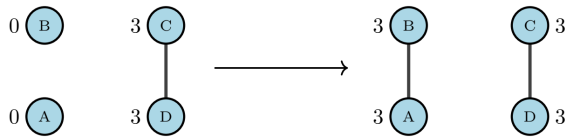
Como se menciona anteriormente, otro de los objetivos del análisis de formación estratégica de redes sociales es el de conocer las redes que benefician a la sociedad en general. Dado que conocemos las funciones de utilidad de cada agente en la red, se pueden rankear cada una de las posibles redes. Para alcanzar este objetivo se desarrolla el concepto de eficiencia.

En términos de eficiencia, la mejor red sería aquella que maximice la utilidad de todos sus agentes. Formalmente, una red  $g$  es eficiente en relación con las funciones de utilidades  $(u_1, \dots, u_n)$  si  $\sum_i u_i(g) \geq \sum_i u_i(g')$  para todo  $g' \in G(N)$ . Dada la definición, entonces existirá siempre al menos una red eficiente (Jackson 2008).

Veamos la Figura 5.

En la Figura 5 podemos observar dos redes sociales, en la de la izquierda, los agentes A y B tienen una utilidad de 0 ya que no tienen ningún link. Los agentes o nodos C y D tienen un link entre ellos por lo tanto su utilidad es de 3 para cada uno. En la red social del lado derecho se ha añadido un link entre los agentes A y B, por lo tanto, ahora toda la sociedad está mejor, es decir, ahora todos tiene una utilidad de 3. En

Figura 5: Red eficiente.



este ejemplo, la red social eficiente sería la del lado derecho. Quizás es bueno preguntarse ¿Y si agregamos otro link entre los agentes B y C? En este caso añadir un link entre los nodos B y C ciertamente beneficiaría a ambos, sin embargo, tanto A como D reducirían su utilidad.

## 4 Homofilia

Como anteriormente observamos, las redes sociales pueden ayudarnos a explicar varios fenómenos sociales y económicos. Uno de los conceptos importantes en las redes sociales es el de la homofilia. Este concepto refleja como los individuos tienden a establecer relaciones sociales con aquellos individuos que tienen características similares, ya sea de ideología, pensamiento, raza, estatus social, religión, etcétera. Las consecuencias de la homofilia son importantes, ya que explican la segregación o polarización de las redes sociales.

Consideremos el siguiente ejemplo de formación de redes con homofilia. Suponga que un amigo le presenta una persona a usted, de manera que usted forma una amistad con esa persona. Ahora considere el caso en el que usted conoce a esta persona en la escuela o trabajo y forma una amistad. En el primer escenario, la amistad surgió a través de la misma red (del amigo en común), sin embargo, en el segundo, el contexto en el que se desenvuelven los actores nos ayuda a entender un poco más la formación del link. En el segundo escenario, los actores comparten en común un entorno, quizás ciertas preferencias inclusive (Easley & Kleinberg, 2010).

Analicemos la situación previa de manera formal. Considere una red con tres nodos A, B, y C como en la Figura 6 del lado izquierdo. En este caso, los nodos A-B tienen un link o una amistad entre ellos. De la misma manera los nodos A-C. Los nodos B y C tienen como amigo en común el nodo A. La pregunta interesante sería ¿Qué características en común deben tener los nodos B y C con respecto al nodo A? De acuerdo con el

Figura 6: Homofilia.



principio de homofilia, lo mas probable es que tengan características comunes, por ejemplo, edad, ideología, gustos y preferencias. Este mismo principio sugiere la creación de un link entre los nodos B y C, como lo vemos en la red social del lado derecho. Sin embargo, dadas las características idiosincráticas de los nodos B y C, la amistad o link podría surgir inclusive sin que el nodo A interviniera.

Considerando el ejemplo anterior nos podemos preguntar ¿qué sucede si entonces la formación de links en una red obedece solamente a similitudes en las preferencias, ideologías o características raciales? En este caso la polarización o segregación social surge, como lo vemos en la Figura 7.

Figura 7: Polarización. Fuente: Easley & Kleinberg (2010).

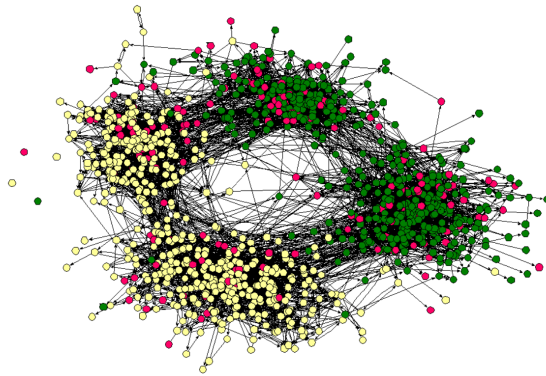
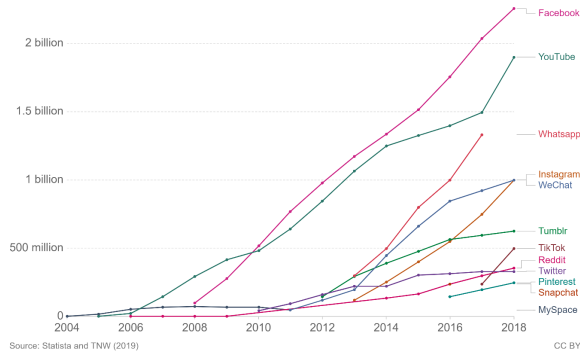


Figura 8: Uso de redes sociales. Fuente: Statista and TNW (2019).



## 5 Adopción tecnológica y procesos de contagio

El uso y difusión de innovaciones tecnológicas puede ser explicado desde un punto de vista de teoría de redes. La teoría de redes puede mostrarnos como la adopción de una nueva tecnología puede difundirse a través de las redes sociales. Veamos el ejemplo del uso de redes sociales tales como MySpace, Facebook etcétera.

En el 2003 la plataforma social MySpace hizo su aparición. Fue la primera red social cuya popularidad se maximizó en los años del 2005 al 2008. La plataforma fue creada por Brad Greenspan, fundador de la empresa eUniverse; una empresa dedicada al marketing por internet. Los primeros usuarios de la red social fueron los mismos empleados de manera que la difusión del uso de esta red social fue pequeña al inicio. Sin embargo, de acuerdo con el mismo Greenspan, la compañía llevó a cabo concursos para ver quién podría agregar más usuarios a la red social.

Como es común en las redes sociales, la transmisión o adopción de alguna nueva tecnología usualmente es recomendada por algún conocido. Esta transmisión y adopción de tecnología implica que la utilidad por adoptar nueva tecnología este en función de que otras personas también la adopten. En la Figura 8 vemos como se incrementó el uso de las redes sociales en los últimos 15 años.

La adopción de nueva tecnología y su proceso de difusión ha sido analizado en diferentes contextos. Uno de los ejemplos más comunes para

entender la adopción de nueva tecnología y su proceso de difusión en el sector salud es el analizado por Coleman et al (1959). Ellos analizaron el proceso de adopción de una nueva droga entrevistando a más de 260 médicos en cuatro ciudades durante un periodo de 15 meses. Al igual que en el caso de las plataformas sociales, la adopción y difusión de la nueva droga se llevo a cabo mayormente entre colegas médicos.

El alcance de la difusión y uso de la nueva medicina en la red social de los médicos se observa con el siguiente dato. Tan solo un tercio de aquellos médicos (que no tenían amigos médicos en común que si hayan utilizado la nueva droga) adoptaron el uso de la nueva medicina. Por otra parte, más de la mitad de los médicos (que mencionaron tener otros colegas médicos utilizando la nueva droga) utilizaron la nueva droga. Claramente la adopción de esta nueva tecnología medica tuvo una mayor difusión por la recomendación de colegas (Jackson, 2008).

## 6 Ejemplo del uso de redes

Ahora que ya conocemos los principales conceptos veamos el siguiente ejemplo aplicado. Vamos a utilizar un juego en redes que tiene aplicación en decisiones de consumo, comportamiento, formación de identidad, presión social y homofilia.

Un juego en redes es un juego estándar de teoría de juegos, pero que se juega con nuestra red social, con la particularidad de que solo podemos jugar una estrategia contra las de nuestros contactos. Esto es, cada nodo es un jugador que juega con sus contactos, a través de sus links, pero no puede diferenciar estrategias entre estos, por lo que tiene que tomar la estrategia que mejor se adapte al conjunto de sus contactos en vez de responder a la estrategia de un único jugador adversario.

Esta es una situación de posicionamiento público en la que tomamos acciones de manera pública, o una opción que refleja de cierta manera nuestra identidad, y que no podemos (o no debemos) variar según con quien nos encontremos. Algunos ejemplos son nuestras intenciones políticas, ser hinchas de un determinado equipo de deporte o el apoyo de determinadas causas o ideales, pero es igualmente extensible a opciones de consumo realizadas con grupos de amigos, como ir al cine, bares, tipos de música de los lugares que frecuentamos con nuestros amigos, o incluso cuestiones como el consumo de alcohol o fumar.

En este caso tenemos un juego de coordinación en el que los beneficios para los jugadores están en la diagonal del juego. Se consideran

tres estrategias o acciones (A, B y C) que se corresponden con tres tipos diferente de jugador (A, B y C).

Cada tipo de jugador tiene su acción preferida (la que le reporta mayor beneficio), que es idéntica a su tipo, una segunda mejor opción y una tercera, con beneficios descendientes.

El beneficio, cuando se coordina con una acción de su propio tipo, reporta un beneficio de dos (2) unidades; coordinar con la segunda opción reporta un beneficio de unidad y media (1.5); y la tercera opción genera solo una unidad (1).

Estas acciones se ranquean de la siguiente manera:

- Tipo A: A, B, C
- Tipo B: B, C, A
- Tipo C: C, A, B

De esta manera las tres acciones ocupan la primera, segunda y tercera opción para algún tipo de jugador.

Las matrices de pagos para cada tipo serían como en las Tablas 1 a 5. El lector puede buscar el resto de las combinaciones.

Tabla 2: Tipo A contra tipo A

	A	B	C
A	2, 2	0, 0	0, 0
B	0, 0	1.5, 1.5	0, 0
C	0, 0	0,0	1, 1

Tabla 3: Tipo B contra tipo B

	A	B	C
A	1, 1	0, 0	0, 0
B	0, 0	2, 2	0, 0
C	0, 0	0, 0	1.5, 1.5

Tabla 4: Tipo C contra tipo C

	A	B	C
A	1.5, 1.5	0, 0	0, 0
B	0, 0	1, 1	0, 0
C	0, 0	0, 0	2, 2

Tabla 5: Tipo A contra tipo B

	A	B	C
A	2, 1	0, 0	0, 0
B	0, 0	1.5, 2	0, 0
C	0, 0	0, 0	1, 1.5

Cada jugador obtiene beneficios solo cuando coordina su acción con sus contactos, pero desconoce el tipo de sus contactos, solo observa su acción.

Entonces, aun siendo de un determinado tipo, y teniendo preferencia por una determinada opción, si mi entorno juega acciones diferentes a la mía, mi beneficio se ve reducido a cero. Es entonces mejor cambiar de acción, hacia la segunda o tercera opción, en función de que acción reporta mayores beneficios.

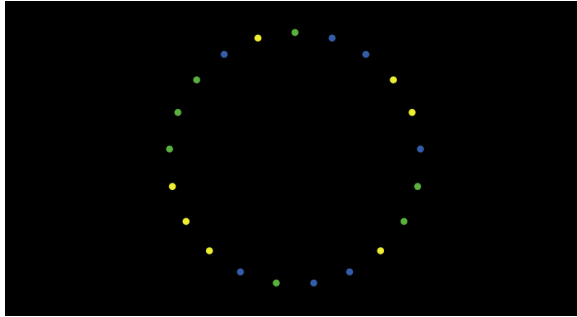
Aun siendo de un determinado tipo, si todos mis contactos toman la acción que menos beneficio genera, es preferible coordinar en la tercera opción que generar un beneficio de cero.

Esto implica una cierta forma de presión social, puesto que existe una opción que generaría para este individuo mayores beneficios, en caso de contar con contactos que tomen también esa misma opción (obsérvese que no necesariamente sean del mismo tipo, lo tipos no son observables, las acciones sí).

El interés en este caso está en ver cómo la estructura de red altera las acciones de los jugadores, que de esta manera pueden llegar a ocultar su tipo al tomar segundas y terceras opciones.

Para mantener la estructura simple y limitarnos al efecto de la estructura de red, utilizaremos un modelo donde cada tipo esta igualmente representado, y por lo tanto no existen minorías (problema serio en casos reales).

Figura 9: Nodos. Elaboración propia usando NetLogo.



Para no exigir demasiado nivel de computación, se ha optado por una sociedad muy pequeña (comparable a una clase o un ambiente laboral): 21 agentes o nodos, 7 de cada tipo.<sup>1</sup> Estos 21 agentes se relacionan de manera aleatoria y, por la misma simplicidad, hemos establecido que cada agente tendrá el mismo número de conocidos.

De manera gráfica, los nodos se establecen en un círculo, lo que simplifica la visualización de la red formada. Cada tipo se identifica con un color diferente. Esto se muestra en la Figura 9.

Si partimos de una red completa, todos tiene 20 contactos y todos se conocen entre sí como lo muestra la Figura 10.

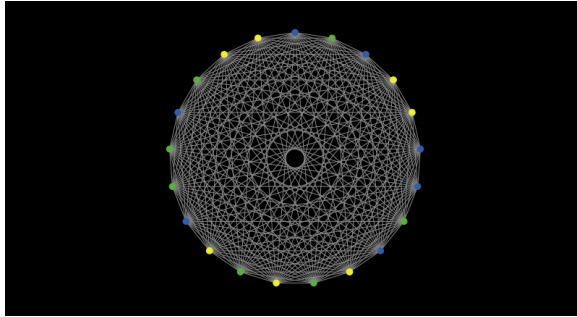
En el caso de la red completa, como veremos a continuación, es óptimo tomar la acción preferida por cada tipo, obteniéndose de esta manera el mayor beneficio total. Cada persona se relaciona con 6 personas de su propio tipo y 7 de cada otro tipo, por lo que el beneficio de mantener su opción preferida, revelando así su tipo, es mejor que cualquier otra opción.

Cada tipo observa 6 de su propio tipo (beneficio por esa acción  $6 \times 2 = 12$ ) y 7 de cada otro tipo (beneficio de  $7 \times 1.5 = 10.5$  para la segunda opción y  $7 \times 1 = 7$  para la tercera). Cada jugador genera así un bienestar de 12, y el bienestar total es de 252. Esta situación es estable, y ningún jugador quiere cambiar su acción.

---

<sup>1</sup> La generación de una red incrementa de manera exponencial su necesidad de capacidad de cómputo, de acuerdo con el número de nodos y links englobados en la red.

Figura 10: Red completa. Elaboración propia usando NetLogo.



Ahora, si reducimos el número de contactos que tienen los individuos en uno y la red no es completa, la gente empieza a coordinar en acciones diferentes.

La dinámica es la siguiente:

De inicio cada jugador escoge su acción preferida por default. Se escoge a uno de los jugadores al azar, este observa sus contactos (sus acciones, no su tipo) y cambia su acción si así consigue mayor beneficio. Realice el cambio o no, se escoge al siguiente jugador al azar, para comprobar si quiere cambiar su acción o no, y así hasta que todos los jugadores han tenido la oportunidad de cambiar su acción, si así lo desean. El proceso se repite hasta que en una ronda completa nadie cambia su acción, esto es, se alcanzó un equilibrio estable desde el punto de vista individual.

Al eliminar un solo contacto, elegir la acción preferida ya no es un equilibrio estable, y el conjunto de jugadores empieza a coordinar en alguna opción, de manera relativamente aleatoria, de forma que una de las tres acciones acaba siendo predominante, pero no necesariamente la única.

De esta forma la sociedad coordina de manera notoria en una de las opciones, pero en general, a pesar de ser la acción preferida por solo una tercera parte, el bienestar de la sociedad incrementa.

Supongamos un individuo que ha perdido por casualidad el contacto con uno de los jugadores de su tipo. Ahora tiene 5 de su propio tipo entre sus contactos ( $5 \times 2 = 10$ ), pero 7 de los que optan por su segunda mejor opción ( $7 \times 1.5 = 10.5$ ), por lo que cambiará su acción a su segunda mejor opción, pero eso empezará a generar una cadena de cambios, que hacen que una determinada opción predomine en la sociedad. Al final, si to-

dos coordinaran en la misma acción, tenemos 19 contactos, que generan incluso un beneficio mayor ( $19 \times 1.5 = 27$  para el que juega su segunda mejor opción, y  $19 \times 1 = 19$  para el que juega su tercera mejor opción, y obviamente  $19 \times 2 = 38$  los que tuvieron la suerte de que su opción fuera la elegida en este juego de coordinación). No solo incrementa el beneficio de la sociedad en su conjunto, se incrementa el beneficio de todos los jugadores.

El resultado en este caso es aleatorio, ya que dependerá de la dinámica aleatoria de los procesos de cambios, y no siempre termina con todos coordinando en la misma acción.

Simulando el escenario 100 veces se obtuvo una media de bienestar de 385.7, con una mediana de 389.75, un máximo de 455.5 y un mínimo de 314.5.

En media el, 35.619 % de los jugadores tomaron su opción preferida y un 32.19 % la segunda y tercera. Esto es muy cercano a  $1/3$ ,  $1/3$  y  $1/3$ , que correspondería al caso en el que todos coordinan en la misma acción, lo que ocurrió en un 55 % de las ocasiones.

Este es un primer resultado importante. Con eliminar un solo contacto, la sociedad cambia completamente su comportamiento agregado, de acciones separadas, hacia una acción colectiva.

Esto implica, para este caso en particular, que eliminando un contacto por persona, la sociedad coordina en un determinado estado, concentrándose en solo una opción (es menos diverso), pero con un mejor resultado en general.

Ahora bien, se ha generado desigualdad entre los jugadores (cosa que no ocurre en la red completa), lo cual puede generar a la larga problemas de otra índole, ya que a pesar de que todos salen beneficiados, unos se benefician más, y otros menos (¿por qué coordinamos en este equilibrio y no en el que yo prefiero?) Esto es un problema desde otro punto de vista de análisis, ajeno a este estudio en particular.

Entonces tenemos un problema de salto cuantitativo y cualitativo de la red completa (20 conocidos) a una red casi idéntica, pero con un contacto menos (19 conocidos).

Pero seguimos el análisis. Reduciendo aún más el número de contactos, para ver qué ocurre. La Tabla 6 ofrece los resultados en media para 100 simulaciones para cada caso. Podemos ver cómo el bienestar incrementa según lo hace el número de contactos de las personas, y cómo los que juegan su primera opción decrecen, mientras que aumentan los que toman su segunda y tercera opción como acciones.

Tabla 6: Resultados de bienestar y proporciones de opción como acción

	Bienestar	Primera	Segunda	Tercera
7	127	0.516	0.248	0.236
8	141	0.564	0.221	0.215
9	162	0.490	0.254	0.257
10	180	0.463	0.262	0.275
11	193	0.500	0.252	0.249
12	217	0.462	0.273	0.264
13	236	0.430	0.289	0.280
14	257	0.439	0.283	0.279
15	288	0.390	0.302	0.308
16	302	0.393	0.299	0.309
17	321	0.400	0.300	0.300
18	362	0.370	0.318	0.313
19	386	0.356	0.322	0.322
20	252	1.000	0.000	0.000

Vemos pues que la red completa genera un efecto de trunque, y con la red completa todos juegan su primera acción, pero limita sus beneficios al grupo de 7 que forman cada tipo en particular. Si el número de contactos es bajó, cerca de un 50 % de las personas aun toman su acción preferida, pero esto disminuye según se incrementa el número de contactos, siendo básicamente un 35 % para el caso de 19 contactos, donde la coordinación es mayor en alguna de las tres posibles acciones.

Si vemos un gráfico de dispersión con el total de 1400 resultados de las simulaciones, podemos observar la tendencia y la distribución de los valores de bienestar.

Podemos observar la tendencia al alza en el bienestar, con excepción de la red completa, y el incremento de varianza en los valores obtenidos. Según incrementa el número de contactos los valores de la simulación se vuelven más dispersos, incrementando la varianza. Si volvemos el valor de bienestar con el número de contactos, eliminando el caso de la red completa, obtenemos los resultados de la Tabla 7.

Figura 11: Gráfico de dispersión de los niveles de bienestar respecto al número de contactos.

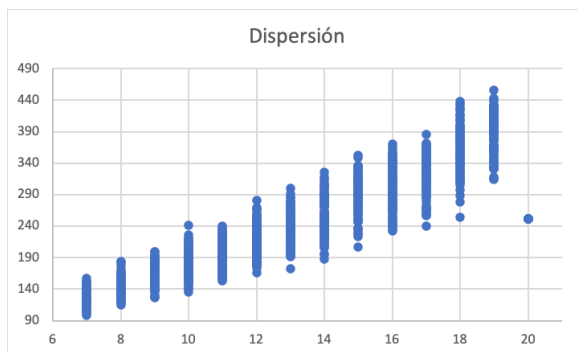


Tabla 7: Modelo de MCO usando 1300 observaciones.  
Variable dependiente: Bienestar

	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>Estadístico t</i>	<i>valor p</i>	
const	-33.4192	2.76033	-12.11	<0.0001	***
amigos	21.3489	0.204050	104.6	<0.0001	***

R-cuadrado	0.893994	R-cuadrado corregido	0.893913
F(1, 1298)	10946.61	Valor p (de F)	0.000000

Vemos en la variable amigos 7 que es altamente significativa, y que la regresión tiene un R cuadrado de 0.894, explicando casi la totalidad de la varianza observada en el bienestar.

Mientras no alcancemos la red completa, el incrementar el número de contactos mejora la coordinación en la sociedad, si bien genera menos diversidad en la misma. Esto implica la desaparición de dos de las tres posibles acciones que teníamos de inicio.

Esto es un ejemplo simple, pero sirve para ilustrar la presión social (si todos lo hacen, yo cambio mi acción), que es aplicable para muchas decisiones de consumo (como fumar, o dejar de hacerlo). También en un modelo más complejo, podríamos permitir la búsqueda dentro de la red, que en este modelo simple es estático y establecido al azar. Si los individuos pueden alterar su red, buscarán conectarse con gente que tome su acción preferida, rodeándose de los que son de su mismo tipo, y evitando a los otros dos tipos, generándose la homofilia, la tendencia de buscar personas similares a nosotros mismos en nuestros círculos de amigos (misma visión política, misma religión, mismo equipo deportivo...)

## **7 Decisiones, comportamiento, preferencias**

### **7.1 Aprendizaje y redes**

Nuestras redes sociales son muy importantes para nuestra formación de opiniones, para dar consejos y compartir información. Por ejemplo, usamos nuestras redes sociales para informarnos sobre posibilidades de trabajo, cuando buscamos casa o cuando buscamos donde estudiar. Algunas preguntas que nos interesa analizar cuando estudiamos la influencia de las redes sociales son las siguientes:

1. ¿Las creencias y las opiniones de los participantes llegan a un consenso o permanecen divididas?
2. ¿Cuáles son los individuos que tienen una mayor influencia en las opiniones de otros?
3. ¿Qué tan rápido actualizan sus creencias y opiniones los participantes en las redes?

Abajo hablaremos brevemente de dos tipos de modelos que intentan responder estas preguntas. El primer tipo de modelos es el de aprendizaje bayesiano, donde los individuos observan las acciones y los resultados experimentados por otros participantes y actualizan sus creencias con esta información. Sin embargo, los modelos de actualización bayesiana tienen la desventaja de ser muy difíciles de resolver.

El segundo tipo de modelos está basado en una manera mucho más sencilla de actualizar las creencias. Estos modelos asumen que los individuos intercambian información con otros individuos y actualizan sus creencias tomando un promedio ponderado de las creencias de estos individuos. Estos modelos son bastante manejables.

## 7.2 Aprendizaje bayesiano

En el aprendizaje bayesiano se asume que cada individuo observa una señal imperfecta (y diferente), por ejemplo, de la calidad de un producto. Los individuos observan las acciones de los otros jugadores y sus pagos, pero no la señal que observaron los otros individuos. Por ejemplo, cuando estamos considerando ir al cine, tenemos una idea de la calidad de la película. Cada persona puede observar si otras personas fueron al cine y si les gustó o no. Esto nos da una idea de que piensan los demás de la calidad de la película y podemos usar esta información para actualizar nuestras propias creencias. Una conclusión central del aprendizaje bayesiano es que, si los agentes pueden observar las acciones de las otras personas y sus beneficios de estas decisiones, y todas las personas tienen las mismas preferencias y se enfrentan a la misma incertidumbre, entonces eventualmente terminarán con pagos similares. La idea es que si un agente está obteniendo resultados significativamente peores que los demás, debe darse cuenta de esto y eventualmente cambiará sus acciones para copiar las decisiones de los demás y terminará tan bien como las otras personas. Esto implica que, en el límite, todos los agentes conectados en una red deberán de terminar con los mismos pagos.

## 7.3 Imitación y modelos de influencia social: el modelo de Degroot

En estos modelos, los individuos observan directamente las creencias de los otros individuos y usan el promedio ponderado de estas creencias para actualizar las suyas propias. En el modelo de Degroot (1974) se asume que los participantes no ajustan el peso con el que ponderan la opinión de los otros participantes en el tiempo.

Los líderes de opinión surgirán naturalmente en el modelo como individuos que son escuchados por otros individuos y que tienen influencia en la opinión de los demás. Esto es, la influencia de un individuo sobre las creencias finales en una red dependerá de cuanto peso otros individuos le pongan a su opinión.

Si en la red solo hay un grupo fuertemente conectado, las creencias de los miembros del grupo convergirán. Si los otros individuos tienen contactos directos con el grupo, deberán de llegar a la misma creencia. En el límite, estos individuos no tendrán ninguna influencia en las creencias finales.

En el caso en el que haya varios grupos en el que sus miembros están fuertemente conectados entre sí, cada grupo alcanzará sus propios

consensos, y las creencias finales de los individuos que no estén en ninguno de los grupos serán un promedio ponderado de las creencias de los grupos fuertemente conectados.

La importancia de las creencias de cada individuo en la creencia final de la red depende no solamente de su ponderación en la comunicación y en la actualización de las creencias de la red, sino de la ponderación de los individuos que los escuchan. Esto nos permite tener una medida de la influencia de cada individuo. Asimismo, podemos decir que los líderes de opinión son aquellos individuos que tienen una mayor ponderación en la comunicación y en el proceso de actualización de creencias.

## 7.4 Preferencias y decisiones

Muchas veces la utilidad que recibimos de un producto depende de si nuestros vecinos o amigos tienen ese producto. Por ejemplo, la utilidad de muchos programas de cómputo aumenta con el número de personas que tiene el mismo programa. O cuando buscamos una consola para videojuegos, nos interesa que nuestros amigos tengan la misma consola para compartir los juegos. A veces pasa lo contrario. Muchas veces cuando escogemos qué ropa comprar, buscamos que no la estén usando nuestros conocidos. En estas situaciones, las redes tienen una importancia fundamental en las decisiones de las personas.

Se dice que un producto presenta externalidades de redes si la utilidad de usar el producto aumenta con el número de personas que usan esos productos (Tirole, 1995). Algunos ejemplos son los programas de computadoras, los servicios de un banco, las redes sociales y las consolas para video juegos e incluso los teclados para computadoras. Hay dos razones por las que el número de personas que usan un producto impacta la utilidad de esos productos. Por el lado de la demanda, los consumidores se pueden beneficiar del número de consumidores de un producto ya que les permite compartir esos productos con más personas. En este caso, los consumidores deben de anticipar la tecnología que será usada por la mayoría de los consumidores, lo que introduce problemas de coordinación y puede generar varios equilibrios e ineficiencias. Por el lado de la oferta, el número de consumidores puede impactar la calidad y servicio de un producto (por ejemplo, el número de cajeros automáticos está en función del número de clientes de un banco).

David (1985) argumenta que el teclado actual de las computadoras QWERTY (llamado así por las primeras seis letras del teclado) es ineficiente y hay otros teclados que nos permiten escribir con mayor rapidez. David recuenta que el teclado QWERTY fue inventado en los Estados

Unidos a finales del siglo XIX. En la primera línea se pusieron las letras que se usan para escribir "Type Writer". Esto se hizo para que los vendedores de las primeras máquinas de escribir pudieran escribir estas palabras lo más rápido posible e impresionar a los compradores. David explica que lo más eficiente hubiera sido poner primero las teclas que se usan con mayor frecuencia al escribir. Sin embargo, el teclado QWERTY se estableció como el teclado dominante en el mundo. Ahora, por las externalidades de redes, todos quieren aprender a usarlo. Otros teclados que aparentemente son más eficientes y nos permiten escribir más rápido han fallado en establecerse dado el dominio del teclado QWERTY.

Muchas veces nos interesan las características sociales de un producto. En su artículo seminal, Becker (1991) intenta explicar por qué dos restaurantes aparentemente iguales, y que están muy cerca del otro, en Palo Alto, California, pero con popularidades muy diferentes, cobran precios similares. Mientras que el más popular tiene largas filas, el menos popular tiene mesas vacías la mayoría del tiempo. La pregunta es porque los dos restaurantes cobran precios similares si las teorías económicas estándar nos dirían que el restaurante más popular debería de cobrar un precio mayor (en sí el restaurante menos popular cobra precios ligeramente superiores al restaurante popular). El argumento de Becker es que lo que hace atractivo al restaurante popular es en sí su popularidad, lo que provoca que parte de la demanda tenga pendiente positiva. El restaurante popular no quiere subir los precios, porque una pequeña caída en sus clientes podría provocar que otros clientes no quieran ir. El punto interesante que introduce Becker en su artículo es que muchas veces queremos consumir un producto solamente porque otras personas los están consumiendo.

El comportamiento de rebaño de las hormigas puede ser instructivo para entender el de las personas. Kirman (1993) analiza cómo se comportan las hormigas cuando se les pone dos fuentes de comida idénticas. Las hormigas eligen una de las dos comidas y la mayoría de ellas van hacia ella. Sin embargo, ocasionalmente cambian de comportamiento y la mayoría va hacia la otra fuente de comida. Para modelar este comportamiento, Kirman asume que cada periodo una hormiga de la población es seleccionada. Con una probabilidad de  $1-\varepsilon$ , esta hormiga copia el comportamiento de otra hormiga elegida al azar. Con una probabilidad de  $\varepsilon$ , la hormiga no copia el comportamiento de otra hormiga, sino que escoge al azar a cuál de las dos comidas ir (con una probabilidad del 50% cada una). Cuando  $\varepsilon = 0$ , el comportamiento en este modelo es completamente de imitación. Cuando  $\varepsilon = 1$ , es un modelo en el que

los individuos escogen sus acciones al azar, sin tomar en cuenta las acciones de los otros individuos. Cuando  $\varepsilon$  es muy alta, pero sin ser 1, se da el comportamiento de las hormigas, en las cuales la mayoría de las hormigas se concentran en una fuente de comida, pero ocasionalmente cambian a la otra.

Calvó-Armengol y Jackson (2004) usan un modelo similar para modelar la decisión de padres e hijos con respecto a la educación. Si aquellos en una comunidad que rodean a un individuo influyen en la decisión del individuo, como en el caso de la educación superior, entonces tanto padres como hijos, que están relacionados con la misma comunidad, tomarán decisiones similares. Calvó-Armengol y Jackson muestran que un modelo sencillo de redes puede representar bastante bien la relación observada entre las decisiones de educación superior entre padres e hijos de varios países europeos.

Un modelo similar también nos podría explicar el fenómeno de la moda. En la moda, las personas cambian el diseño de los productos que compran a través del tiempo. Karni y Schmeidler (1990) muestran con su modelo de moda, como un grupo de personas, incluso con preferencias fijas, puede cambiar de elección a lo largo del tiempo. Los autores asumen que hay dos grupos de personas: uno tiene preferencias por consumir un producto que es consumido por su propio grupo; el otro prefiere consumir un producto si es consumido por los dos grupos. Ellos muestran que este modelo genera una solución en la que los individuos escogen cambiar periódicamente de los productos que consumen.

## Referencias

- [1] Aguirre, J. L. (2011). Introducción al análisis de redes sociales. Documentos de Trabajo del Centro Interdisciplinario para el Estudio de Políticas Públicas, 82, 1-59.
- [2] Becker, G. S. (1991). A Note on Restaurant Pricing and other Examples of Social Influence on Price, *The Journal of Political Economy*, 99 (5), 1109-1116.
- [3] Calvó-Armengol, A. y Jackson, M.O., (2004). Like Father, Like Son: Labor Market Networks and Social Mobility. mimeo: Caltech and Universitat Autònoma de Barcelona, <http://www.stanford.edu/jacksonm/mobility.pdf>, California Institute of Technology Working Paper # .

- [4] Coleman, J., Menzel, H., & Katz, E. (1959). Social processes in physicians' adoption of a new drug. *Journal of chronic diseases*, 9(1), 1-19.
- [5] David, P. A. (1985). Clio and the Economics of QWERTY, *The American Economic Review*, 75 (29), 332-337.
- [6] Degroot, M.H. (1974). Reaching a Consensus, *Journal of the American Statistical Association*, 69 (345), 118-121.
- [7] Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets* (Vol. 8). Cambridge: Cambridge university press.
- [8] Goyal, S. (2007). *Connections: an introduction to the network economy*. Princeton University Press.
- [9] Jackson, Matthew, O., (2008). *Social and Economic Networks*. Princeton University Press, New Jersey, U.S.A.
- [10] Jackson, M. O., & Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of economic theory*, 71(1), 44-74.
- [11] Karni, E. y D. Schmeidler, (1990). Fixed Preferences and Changing Tastes. *The American Economic Review*, 80 (2), 262-267.
- [12] Kirman, A.P. (1993). Ants, Rationality, and Recruitment, *The Quarterly Journal of Economics*, 108 (1), 137-156.
- [13] Tirole, J. (1995). *The Theory of Industrial Organization*, MIT Press.



En los tomos I y II de la serie *ABCD'Economía* se presentaron las que consideramos son las cuatro grandes áreas en Economía: Métodos Cuantitativos, Microeconomía, Macroeconomía y Econometría, de las cuales se derivan todas las demás áreas de interés para los economistas, así como un sinfín de aplicaciones. En este tomo tenemos como principal objetivo el de complementar estas herramientas. Pero, además, tenemos un reto adicional: despertar el interés de nuestros lectores por tomar el camino de la investigación mediante la presentación de nuevas aplicaciones y métodos.

UNIVERSIDAD DE  
GUANAJUATO



ISBN: 978-607-441-786-9



9 786074 417869