

P
A
R
T
E



Anexo estadístico y metodológico

Anexo Metodológico

Introducción

Uno de los principales objetivos del *Informe Estado de la Nación* es proveer información oportuna, que permita conocer el avance del país en el logro de sus aspiraciones de desarrollo humano sostenible. En su preparación cada año interviene una amplia red de instituciones e investigadores, quienes colaboran con el suministro de datos actualizados y la aplicación de técnicas y mediciones novedosas, que facilitan una comprensión más objetiva de la realidad nacional. Con esta incorporación de distintos instrumentos metodológicos se pretende dar una sólida base técnica a los hallazgos presentados en cada capítulo.

En este Anexo Metodológico se exponen los procedimientos seguidos para el abordaje de algunos temas incluidos en esta edición. Cuatro de ellos corresponden al capítulo “Equidad e integración social”, a saber: i) prospección del empleo con la matriz insumo producto (MIP), ii) índice adelantado de demanda laboral (IDL), iii) aporte de las distintas fuentes de ingresos a la desigualdad total, y iv) actualización de la estimación de clase social para las personas ocupadas.

En el capítulo “Oportunidades, estabilidad y solvencia económicas” la principal novedad metodológica es la estimación de una serie de modelos para el estudio del crecimiento económico, los ingresos, el empleo, el sector externo y el impacto de diferentes escenarios de reforma fiscal en los hogares.

A su vez, el capítulo “Fortalecimiento de la democracia”, aporta una nueva

mirada para el análisis sobre los liderazgos de los partidos políticos, así como una revisión comparativa de las ediciones del *Plan Nacional de Desarrollo* presentadas por los últimos tres gobiernos (períodos 2006-2010, 2010-2014 y 2014-2018) y una novedosa metodología para evaluar la gestión del Poder Ejecutivo.

Finalmente, para el capítulo 6, “Ordenamiento territorial: implicaciones para el desarrollo humano”, se explican los métodos utilizados para profundizar en dos temas: por un lado, la relación de la renta del suelo con los costos de transporte, las pendientes y el uso actual del suelo, y por otro, los factores determinantes de la densidad urbana según diversos índices.

Aportes metodológicos en materia de equidad e integración social

Prospección del empleo con la MIP

Para este Informe se diseñó un modelo para la prospección del empleo en Costa Rica, con un horizonte de mediano plazo y con base en las actividades económicas identificadas en la matriz insumo producto (MIP) del 2011 (Vargas, 2015). El modelo además permite cuantificar el efecto potencial que tendría sobre los distintos sectores de actividad, una serie de cambios supuestos en los diversos componentes de la demanda final.

La MIP es un instrumento desarrollado en el marco del nuevo Sistema de Cuentas Nacionales del Banco Central de Costa Rica (BCCR), que muestra las interrelaciones que se dan entre la oferta

y la demanda en la economía y provee una visión detallada y completa del vínculo entre vendedores y compradores de los diferentes sectores económicos. Como se dijo, esta herramienta, en su versión 2011, fue la principal fuente de información para el ejercicio de prospección realizado.

Se efectuó un reagrupamiento de las 77 actividades incluidas en la MIP 2011, hasta generar una matriz de 35 x 35. Esta especificación de sectores permite obtener la denominada “matriz inversa de Leontief”, con la cual es factible: i) examinar interrelaciones entre actividades económicas, ii) conocer el efecto multiplicador de las distintas ramas de actividad sobre la economía y iii) clasificar las actividades con base en sus efectos multiplicadores (Leontief, 1936).

Para identificar y clasificar las actividades se empleó la metodología de Rasmussen (1956), que calcula los encadenamientos “hacia atrás” y “hacia adelante” de una industria con el resto de la economía. También utiliza criterios de poder de dispersión (encadenamiento hacia atrás) y sensibilidad de dispersión (encadenamiento hacia adelante) de los impulsos generados por cambios en la demanda agregada.

El modelo básico de insumo producto¹ parte de la identidad que expresa el valor bruto de producción [X] como la suma de la demanda intermedia [DI] y la demanda final [Y], de la siguiente manera:

$$(1) \quad [X] = [DI] + [Y]$$

donde [X] es un vector de tamaño n

$\times 1$, n es el número de actividades de la economía y cada elemento X_i es la producción de la actividad i . El supuesto básico del modelo de insumo producto incluye la existencia de una relación lineal entre la demanda intermedia y el valor bruto de producción, que se puede expresar como:

$$(2) \quad [DI] = [A][X]$$

donde $[A]$ es la matriz de coeficientes técnicos o matriz tecnológica, cuyos elementos se obtienen dividiendo el consumo intermedio del producto i , utilizado por la industria j , entre la producción de la industria j . Esta matriz contiene coeficientes que muestran la cantidad de insumos directos que requiere cada industria para producir una unidad de su producto.

Sustituyendo (2) en (1), se obtiene:

$$(3) \quad [X] = [A][X] + [Y]$$

Dado que interesa encontrar la producción necesaria que satisface tanto las necesidades intermedias como la demanda final, se puede obtener el vector $[X]$ así:

$$(4) \quad \begin{aligned} [X] - [A][X] &= [Y] \\ [I - A][X] &= [Y] \\ [X] &= [I - A]^{-1} [Y] \end{aligned}$$

donde $[I]$ es una matriz identidad y $[I - A]^{-1}$ es la matriz inversa de Leontief, cuyos elementos, llamados coeficientes de requerimientos totales, muestran el impacto de un cambio exógeno de la demanda final sobre todas las actividades de la economía. Así, la producción de un bien requiere varios insumos intermedios, los que a su vez requieren otros insumos, todo lo cual genera una cadena de efectos de interacción en el proceso productivo.

Mientras la matriz de coeficientes técnicos $[A]$ describe la estructura de la economía en términos estrictamente estáticos, la matriz inversa de Leontief permite examinar las interrelaciones entre actividades económicas y considera todo un conjunto de reacciones y movimientos en la cadena productiva. De acuerdo

con Schuschny (2005), la matriz inversa de Leontief se puede escribir como una sumatoria de matrices, de la siguiente manera:

$$(5) \quad [I - A]^{-1} = [I] + [A] + [A]^2 + [A]^3 + \dots + [A]^n + \dots = \sum [A]^k$$

Esto muestra cómo la matriz inversa incluye efectos directos e indirectos de la demanda final sobre el proceso de producción. El primer término, es decir, la matriz identidad, se refiere a la producción necesaria para satisfacer directamente la demanda final de un sector j . El segundo término, la matriz $[A]$, da cuenta de la producción adicional del sector j requerida para atender las necesidades de insumos de producción de los sectores que le proveen insumos. El tercer término da cuenta de la producción adicional para atender la producción necesaria de la primera ronda, y así sucesivamente. Esta matriz permite obtener los encadenamientos hacia atrás y hacia adelante, y constituye la base para realizar la clasificación de actividades de acuerdo con sus impactos multiplicadores.

Cada elemento de esta matriz es llamado coeficiente de requerimiento total, y representa la producción que debe realizar el sector i (ubicado en las filas) para satisfacer una unidad de demanda final del sector j (ubicado en las columnas). Dependiendo de su posición en la matriz, cada coeficiente constituye un efecto, ya sea directo o indirecto, sobre el resto de la economía. Los efectos directos se localizan en la diagonal principal de la matriz y son mayores o iguales a 1. Esto significa que si la demanda final de una industria incrementa en una unidad, es necesario que su propia producción aumente en el valor del elemento correspondiente a la diagonal principal. A su vez, los efectos indirectos son los que se hallan fuera de la diagonal principal. Es decir, ante un cambio unitario en la demanda de la industria j -ésima, el efecto indirecto contabiliza los cambios en la producción en todas las industrias diferentes a la j -ésima. Esto estaría representado por los elementos de la columna j , excepto el elemento ubicado en la diagonal principal.

Modelo de prospección del empleo

El modelo planteado considera un conjunto de interrelaciones entre las diversas actividades de la economía. Específicamente, supone que es posible identificar con precisión todas esas relaciones (¿quién le compra a quién en la cadena de producción?), así como su grado de profundidad (¿cuánto le compran a cada quién?).

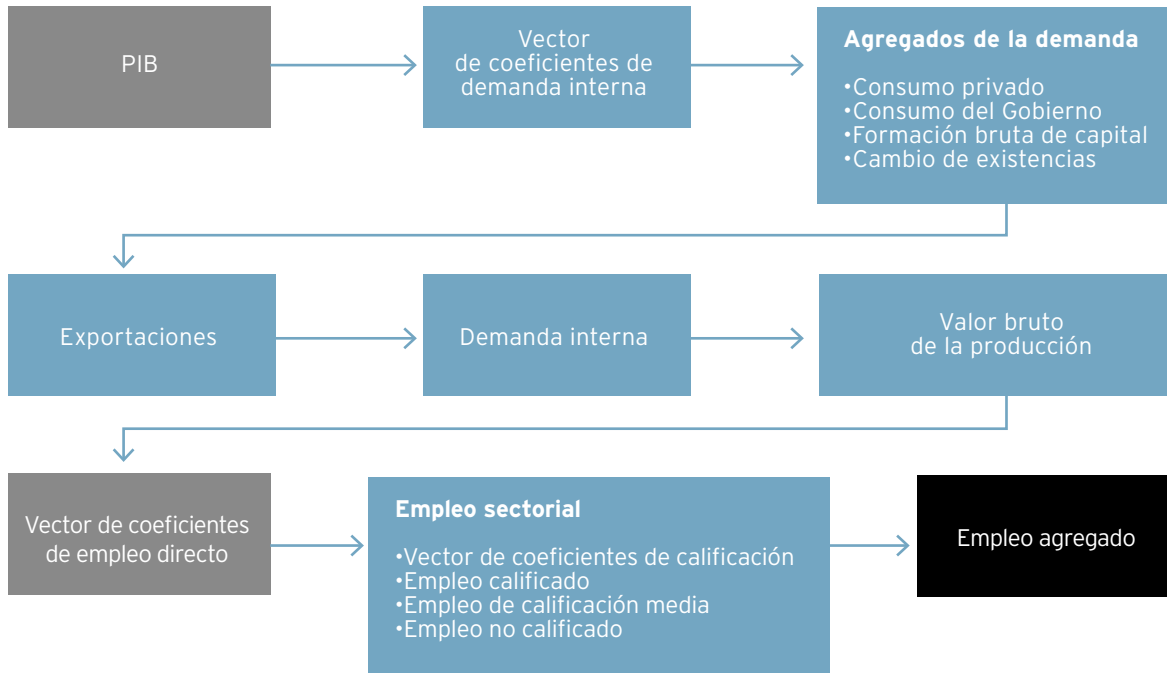
Los escenarios se plantean en términos de la evolución del PIB, la cual, dado un vector de coeficientes de demanda interna, brindan un conjunto de valores para los agregados de la demanda final, a saber: consumo privado, consumo del Gobierno, formación bruta de capital, aumento en existencias y exportaciones. El diagrama 7.1 ilustra los componentes e interrelaciones del modelo. Este conjunto de elementos, junto con la matriz inversa de Leontief, proporcionan el valor bruto de la producción (Leontief, 1986). A su vez, con los valores generados para cada componente es posible calcular el empleo directo (vector de coeficientes de empleo directo). Cada elemento muestra la cantidad de personas requeridas para producir un millón de colones del producto en un determinado sector. Es muy importante señalar que, para efectos de la simulación del modelo, este vector solo se refiere al empleo directo. Para cada uno de los años se estima el empleo directo, y el resultado de su sumatoria es el empleo agregado.

Las simulaciones realizadas no incluyen modificaciones en la productividad media del trabajo, lo que podría conllevar a estimaciones máximas de generación de empleo por unidad de crecimiento económico². Es factible obtener una estimación de la mano de obra según su calificación, a partir de los datos de empleo de la MIP, ya que para cada sector se tiene el desglose en tres grupos: calificado, calificación media y no calificado. Para efectos de la dinámica del modelo, dicho vector de coeficientes de calificación se supone constante a lo largo de las simulaciones. Los supuestos básicos del modelo son³:

- Cada mercancía es producida por una sola industria o sector económico, lo cual implica que no existen productos secundarios.

DIAGRAMA 7.1

Estructura e interrelaciones del modelo para la prospección del empleo



Fuente: Vargas, 2015.

- La función de producción es la postulada por Leontief, por lo que no es posible la sustitución de insumos (hay una cantidad y unas proporciones específicas de insumos que se requieren para producir cada bien, y no es posible la producción si uno de ellos no está disponible).
- Hay rendimientos constantes a escala.
- Existe una dotación limitada de factores (solo se modifica un factor específico: la mano de obra).
- En el corto plazo no ocurren cambios en la estructura productiva de cada sector (la productividad media del trabajo es constante).
- Los insumos comprados por cada sector económico solo dependen del nivel de producción de ese sector. Se asume que las variaciones en la demanda final que dan lugar a las variaciones en la producción de los diversos sectores y, a su vez, a la variación de la producción de los insumos requeridos por el aumento en la producción, ocurren simultáneamente.

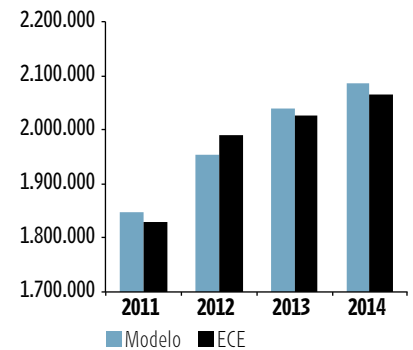
Algunas limitaciones del modelo son las siguientes:

- Dado que fue elaborado antes del cierre de las operaciones de manufactura de la empresa Intel, se toma en cuenta esa actividad, a pesar de que actualmente está restringida a la prestación de algunos servicios.
- No se consideran variaciones en la productividad en la mano de obra, debido a que, como se indicó, el vector de empleo directo opera como una constante. Lo mismo sucede con el porcentaje de mano de obra según calificación. Eventualmente, esto podría llevar a estimaciones máximas de generación de empleo por unidad de crecimiento económico⁴.

El objetivo del ejercicio de prospección fue obtener una estimación razonable del comportamiento esperado del empleo en los próximos años (2015-2021). Dado que la MIP utilizada para tal efecto corresponde a una “fotografía” de la estructura productiva en 2011, un primer ejercicio consistió en estimar el empleo generado en el período 2011-2014 según el modelo

GRÁFICO 7.1

Comparación de la estimación de personas ocupadas según el modelo de prospección y la Encuesta Continua de Empleo (ECE)



Fuente: Vargas, 2015.

y contrastarlo con lo efectivamente reportado por la Encuesta Continua de Empleo del INEC. Los resultados a nivel agregado mostraron una sobreestimación del total de ocupados que osciló entre 0,69% y 1,04%, la cual, aunque no es significativa, podría explicarse por una de las limitaciones antes señaladas:

suponer que la productividad media del trabajo permanece constante (gráfico 7.1).

Para proyectar los componentes de la demanda final se empleó la siguiente metodología. Primero se utilizaron las tasas reales de crecimiento de cada componente entre 2011 y 2014. Para el período 2015-2016 se emplearon los datos del Programa Macroeconómico del BCCR y a partir de 2017 las proyecciones se realizaron con promedios móviles y análisis de tendencias. Luego, para establecer el vínculo entre el crecimiento del PIB y los componentes de la demanda final, se analizó la relación existente entre ambos y se dividió cada agregado entre el nivel del PIB, considerando el 2011 como el nuevo año base. Los resultados de este ejercicio se ilustran en el cuadro 7.1.

Con estos escenarios, el crecimiento proyectado del PIB es el que muestra el gráfico 7.2. Los supuestos de crecimiento son arbitrarios o, si se prefiere, presuntivos. Naturalmente puede optarse por el pesimismo o el optimismo. En todo caso, aun en el escenario optimista, la evolución esperada de la economía es modesta, entre un 4,32% y un 4,78% en el período proyectado, con una tendencia a la desaceleración.

Encadenamientos entre sectores productivos según la MIP

La idea central del enfoque de encadenamientos es que no todas las actividades económicas tienen la misma capacidad para incidir sobre otras, ya que algunas se caracterizan por provocar efectos más intensos de arrastre o de empuje. Un rasgo esencial de la producción es la interdependencia entre las distintas actividades: variaciones de un sector genera cambios en cadena en otros

sectores que lo abastecen de insumos, así como en las ramas a las cuales ese sector provee materias primas. Hirschman (1988) distinguió entre encadenamiento hacia atrás (*backward linkage*) y hacia adelante (*forward linkage*). El primero ocurre cuando una actividad provoca el desarrollo de otras, al demandar bienes de consumo intermedio procedentes de ellas. El segundo tiene lugar cuando a partir del desarrollo de una actividad se obtienen productos que otras ramas utilizarán como insumos intermedios.

Para conocer cómo se distribuyen los impactos de un sector a través de toda la economía, lo que procede es comparar un sector de alto impacto, pero muy concentrado, con uno de menor impacto, pero muy difundido o disperso. Es en este contexto que Rasmussen (1963) introdujo dos conceptos importantes para el cálculo de los encadenamientos, conocidos como medidas de dispersión: el poder de dispersión y la sensibilidad de dispersión.

El índice de poder de dispersión (π_j) de un sector j es una medida del estímulo potencial que este provoca en el resto de la economía, debido a un incremento unitario de su demanda final neta de importaciones. Se calcula con base en el encadenamiento hacia atrás. Si π_j es mayor a 1 ($\pi_j > 1$), significa que los requisitos de insumos intermedios, generados por un aumento unitario de la demanda final del sector j , son mayores para este sector que para el promedio de la economía y, por lo tanto, se trata de un sector con un fuerte poder relativo de arrastre hacia atrás sobre el sistema productivo. Asimismo, es una actividad altamente interconectada (fuerte encadenamiento), por lo que un incremento en su demanda se irradia al resto de actividades, estimu-

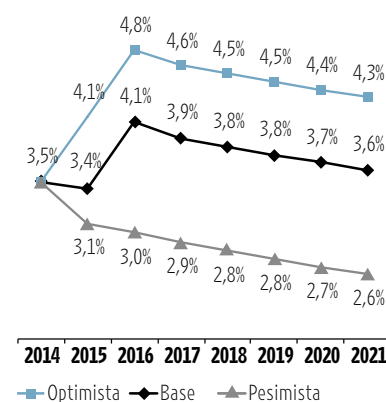
lando la producción y el crecimiento. Por el contrario, si π_j es menor a 1 ($\pi_j < 1$), el estímulo generado es inferior al promedio de la economía y su encadenamiento es débil, con un impacto poco significativo.

Por su parte, el índice de sensibilidad de dispersión (τ_i) mide el estímulo potencial de un crecimiento unitario de toda la economía sobre la demanda final neta de importaciones del sector i . Se calcula con base en el encadenamiento hacia adelante y mide qué tan sensible es un sector ante cambios generales de la demanda. La sensibilidad de dispersión será mayor a 1 ($\tau_i > 1$), si dicho encadenamiento es mayor que el del promedio de la economía. Los resultados obtenidos se muestran en los gráficos 7.3.

El enfoque de encadenamientos constituye el insumo básico para el cómputo de los índices de Rasmussen, los cuales permiten detectar actividades fuertemente

GRÁFICO 7.2

Crecimiento proyectado del PIB, según escenario



Fuente: Vargas, 2015.

CUADRO 7.1

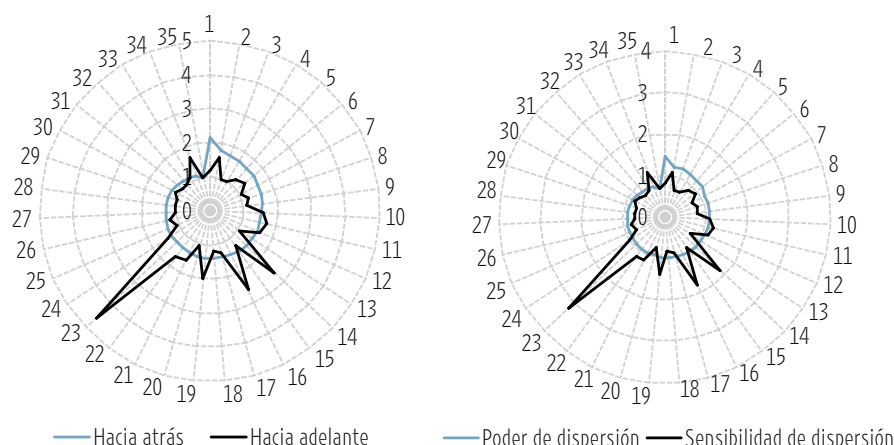
Crecimiento supuesto de los componentes de la demanda final. 2012-2016

Año	Consumo de los hogares	Consumo del Gobierno	Formación bruta de capital fijo	Variación de existencias	Exportaciones	Demanda final
2012	4,2	1,0	8,0	5,3	9,3	6,4
2013	3,3	2,8	12,3	-116,0	3,6	3,0
2014	3,9	3,5	4,5	457,2	-1,7	0,9
2015	5,0	3,2	7,7	-83,0	-7,4	2,2
2016	4,5	2,8	4,6	26,5	4,3	4,3

Fuente: Vargas, 2015, con base en el Programa Macroeconómico 2015-2016 e información del BCCR.

GRÁFICOS 7.3

Encadenamientos y medidas de dispersión^{a/} para los sectores de producción^{b/}



a/ En la escala de los índices de dispersión, un valor mayor a 1 indica que el poder o la sensibilidad del sector es mayor que el promedio de la economía.

b/ Para el detalle de los sectores de producción agrupados, véase Vargas, 2015, en el sitio www.estadonacion.or.cr

Fuente: Vargas, 2015.

CUADRO 7.2

Tipología sectorial según Rasmussen

	Poder de dispersión	
	$\pi_j > 1$	$\pi_j < 1$
Sensibilidad de dispersión	$T_i > 1$	Estratégico
	$T_i < 1$	Independiente
		Clave
		Impulsor

Fuente: Vargas, 2015, con base en Schuschny, 2005.

interrelacionadas, además de seleccionar sectores relevantes en el proceso de crecimiento económico. El método empleado consiste en identificar aquel sector en el que el poder y la sensibilidad de dispersión son mayores a 1. A este se le llama sector clave. El cuadro 7.2 muestra la clasificación sectorial de acuerdo con esta metodología.

Los sectores clave, con altos encadenamientos hacia atrás y hacia adelante, al ser fuertes demandantes y oferentes de insumos intermedios, son canales de paso obligado de los flujos intersectoriales. En el ejercicio realizado para este Informe, cinco sectores quedaron clasificados como clave. Estos proveen el 14,7% del empleo.

Los sectores estratégicos tienen baja demanda de insumos, pero abastecen sustantivamente de insumos a otros sectores. Esta tipología corresponde a siete sectores,

que representan el 53,6% del empleo.

Los sectores impulsores, o de fuerte arrastre, tiene altos encadenamientos hacia atrás, pero bajos hacia adelante. Este grupo está compuesto por once sectores, que absorben el 11,4% de la población ocupada.

Otros once sectores fueron clasificados como independientes; concentran el 20,4% del empleo, consumen una cantidad poco significativa de insumos intermedios y destinan su producción a abastecer, principalmente, la demanda final.

Índice adelantado de demanda laboral (IDL)

Con el objetivo de monitorear las tendencias de corto plazo en la generación de puestos de trabajo y complementar la información de la Encuesta Continua de Empleo (ECE) del INEC, para este

PARA MÁS INFORMACIÓN SOBRE **PROSPECCIÓN DEL EMPLEO**

véase Vargas, 2015, en www.estadonacion.or.cr

Informe se estimó un índice de demanda laboral (IDL), siguiendo la metodología y adaptando a Costa Rica otros indicadores similares construidos en Estados Unidos, Argentina, Uruguay y Bolivia (Barroso, 2013).

Para llevar adelante ese ejercicio, se construyó una base de datos con toda la información de las ofertas de empleo publicados en el periódico *La Nación* los días domingo y lunes de cada semana⁵, que ese medio de comunicación suministró al PEN, en formato de planas completas en archivos PDF. Los avisos se contabilizan (suman) mensualmente. El IDL se mide a partir de la comparación de los puestos de trabajo ofrecidos por cada sector económico en un mes, con respecto a la oferta que se presentó en un momento determinado (mes base).

El índice actual tiene como base el mes de mayo de 2011, pues a esa fecha la ECE tiene varios trimestres, y además cuenta con una captura completa de la demanda laboral por sectores económicos. El objetivo del indicador es medir la evolución y la tendencia de corto plazo en la oferta de puestos de trabajo del mercado, así como la condición económica en general, tomando en cuenta la estrecha relación que existe entre esas variables. Formalmente el índice tiene la siguiente estructura:

$$IDL \text{ por sector} = \frac{\sum \text{Puestos ofrecidos en el mes para el sector } i}{\sum \text{Puestos ofrecidos en el año base para el sector } i}$$

$$IDL \text{ mensual} = IDL \text{ por sector } \times g_i$$

donde g_i es una ponderación que relaciona la importancia del sector económico en la generación del empleo nacional. Los ponderadores fueron obtenidos en la base de datos de la ECE para el primer trimestre de 2011.

El marco muestral de la base de datos consta de 3.029 planas publicadas entre

enero de 2010 y marzo de 2015, es decir, cubre 63 meses. Se aplicó una técnica de muestreo sistemático por mes, cuyo resultado fue la selección de 345 planas, las cuales fueron incluidas en su totalidad en el análisis. Se construyó un factor de ponderación para estimar el total de avisos divulgados en ese lapso. Finalmente, se determinó que en el diario *La Nación* se publicaron 78.315 ofertas de puestos de trabajo los días domingos y lunes durante el período considerado.

Aporte de las distintas fuentes de ingresos a la desigualdad total

Para la presente edición del capítulo "Equidad e integración social", se planteó como tema de interés en el análisis de la desigualdad, determinar qué parte de la desigualdad total se puede atribuir a cada tipo de ingreso, según su procedencia. Para ello se distingue su naturaleza (rentas del capital, trabajo, transferencias, etc.) o su perceptor (sustentador principal, cónyuge, hijos, otros miembros, etc.). Con base en la metodología de Trejos y Oviedo (2012), Trejos (2015) descompuso el ingreso familiar per cápita por fuentes de ingreso e identificó el aporte de cada una de ellas a la desigualdad medida por el coeficiente de Gini.

Para la partición del índice de Gini por fuente se utiliza la relación que involucra la covarianza entre el ingreso y la posición que ocupan las observaciones en la curva de la distribución. Más específicamente, el coeficiente de Gini (G) se define como:

$$G = \frac{2 \text{cov}[y, F(y)]}{\mu}$$

donde $F(y)$ es la distribución acumulativa del ingreso y μ es el ingreso promedio.

Si Y_1, Y_2, \dots, Y_k son los "K" componentes del ingreso familiar per cápita (IFP), entonces:

$$\text{IFP} = \sum_{k=1}^K y_k$$

A partir de esa relación, y utilizando las propiedades de la covarianza, Lerman y Yitzhaki (1985) derivan una descomposición exacta del Gini en la siguiente forma funcional⁶:

$$G = \sum_{k=1}^K R_k G_k S_k$$

donde R_k representa el coeficiente de correlación de Gini entre el componente del ingreso k y el ingreso total IFP, G_k es el índice de Gini de la fuente de ingreso k , estimado para la totalidad de las personas y no solo entre los perceptores de esa fuente, y S_k es la proporción del ingreso k en el ingreso total (Y_k/IFP).

Esta ecuación permite descomponer la desigualdad en el ingreso per cápita del hogar en tres partes: la primera da cuenta de la participación relativa de cada fuente en el ingreso total (S_k), la segunda relaciona la desigualdad del ingreso con el nivel de inequidad observado en la fuente k (G_k), y la tercera muestra la relación entre la fuente de ingreso k y el ingreso total IFP (R_k).

El coeficiente de correlación de Gini de la fuente de ingreso k se define como:

$$R_k = \frac{\text{cov}[y_k, F(y)]}{\text{cov}[y_k, F(y_k)]}$$

donde $F(Y_k)$ representa la función de distribución acumulativa del ingreso de la fuente k . Los valores de R_k se ubican en el intervalo $[-1, 1]$, R_k será igual a 1 cuando la fuente k es función creciente del ingreso, en tanto que si $R_k = -1$ significa que la importancia de la fuente analizada decrece con el ingreso total. Cuando la fuente de ingreso es constante, R_k es cero, de modo que esa fuente no aporta a la desigualdad, esto es, Y_k e IFP son independientes. No obstante, un aumento en el peso relativo de esa fuente si reducirá la desigualdad global. En el caso de que Y_k y IFP tengan una distribución de probabilidad normal, los valores de R_k coinciden con el coeficiente de correlación de Pearson.

Lerman y Yitzhaki (1985) desarrollaron la forma de medir el impacto de los cambios en cualquier fuente de ingresos en la desigualdad total. Si e_k representa el porcentaje de cambio en el ingreso de la k -ésima fuente de ingreso, entonces la variación marginal absoluta en el coeficiente de Gini se obtiene a partir de la siguiente expresión:

$$\frac{\partial G}{\partial e_k} = S_k(R_k G_k - G)$$

Dividiendo la expresión anterior entre el índice de Gini (G) se obtiene el cambio marginal relativo en el índice de Gini producto del cambio relativo en la fuente de ingreso. Esta se puede reescribir así:

$$\frac{\partial G / \partial e_k}{G} = \frac{S_k G_k R_k}{G} - S_k$$

Esa ecuación muestra que el cambio porcentual en el coeficiente de Gini a partir de una modificación en el valor de "e" en la fuente de ingresos "k", es igual a la contribución relativa de esa corriente a la desigualdad total, menos su participación en el ingreso total. La suma de los cambios marginales relativos es cero y, en caso de que todos los componentes del ingreso sean escalados por un factor "e", la desigualdad total permanece inalterada. De acuerdo con Stark, et al., (1986), cuando en esa ecuación la correlación de Gini entre la fuente "k" y el ingreso total (R_k) es negativa o cero, un incremento marginal del ingreso reducirá la desigualdad y, en caso contrario, su impacto dependerá del signo que asuma la expresión $R_k G_k - G$. Una condición necesaria para que la inequidad aumente es que la desigualdad en la fuente de ingresos "k" sea mayor a la desigualdad total: $G_k > G$ ($R_k \leq 1$).

De acuerdo con Yitzhaki (1990) y Wodon y Yitzhaki (2002), se puede estimar la elasticidad-Gini del ingreso k ($EGIk$) como la expresión siguiente:

$$EGIk = \frac{G_k R_k}{G}$$

lo que permite reescribir la ecuación así:

$$\frac{(\partial G / \partial e_k)}{G} = \frac{S_k G_k R_k}{G} - S_k = S_k (EGIk - 1)$$

De esta manera, un aumento porcentual en el ingreso de una fuente k con una $EGIk$ menor a 1 reduce la desigualdad, en tanto que las $EGIk$ superiores a la unidad la incrementan. Cuanto más pequeño sea el valor de la $EGIk$, mayor será su impacto redistributivo. Si la fuente de ingreso disminuye porcentualmente, el efecto en la desigualdad será el inverso. Cuando la $EGIk$ es unitaria, aumentos o reducciones porcentuales en ese ingreso no modifican la desigualdad. Por ello,

para ver los cambios de la desigualdad en el tiempo y el aporte de cada fuente, es necesario considerar tanto el valor de la elasticidad Gini del ingreso, como la dirección del cambio en el ingreso específico y en relación con las demás fuentes.



PARA MÁS INFORMACIÓN SOBRE FUENTES DE DESIGUALDAD

véase Trejos, 2015, en www.estadonacion.or.cr

Actualización de la estimación de clase social para las personas ocupadas

En su Decimoquinto Informe, el PEN propuso una nueva estructura para el estudio de las clases sociales, como herramienta complementaria para el análisis del desarrollo humano y la desigualdad en Costa Rica. Se entiende por clase social el conjunto de personas que reúnen ciertas características que las hacen compartir una posición en la vida económica de un país.

La aplicación de esta propuesta analítica utiliza como fuente de información principal las bases de datos de las encuestas de hogares aplicadas por el INEC durante el período 1987-2014, es decir, la Encuesta de Hogares de Propósitos Múltiples (1987-2009) y la Encuesta

Nacional de Hogares (2010-2014). La solidez estadística de estas encuestas las convierte en una herramienta altamente confiable, cuyo diseño muestral además permite hacer inferencias de los resultados a todo el país. Si bien no fueron concebidas para medir estructuras sociales, cuentan con una amplia batería de datos que ofrece una descripción detallada de la población, en especial de la económicamente activa.

En 2012 el INEC actualizó la Clasificación de Ocupaciones de Costa Rica (COCR-2011) y su respectivo manual, de modo que la clasificación efectuada antes de ese año, en el manual COCR-2000, quedó desactualizada y no es aplicable a las encuestas posteriores. Ante esta situación, para este Informe se decidió retomar la propuesta del PEN y reconstruir la clasificación de clases sociales con base en el nuevo manual.

Esta metodología se apega estrictamente a los criterios planteados en el Decimoquinto Informe, que clasifican a cada persona ocupada⁷ en función de diversos criterios teóricos⁸. La categoría ocupacional o posición en el empleo se usa como un primer discriminante en la distribución de la población, que para este efecto se subdivide en “asalariados o no remunerados”, “cuenta propia”, “patrones o empleadores” y “servicio doméstico asalariado”. El siguiente paso es diferenciar estas categorías según el tamaño de la

empresa o establecimiento para el cual laboran. En este sentido, el tamaño de la empresa (grandes o pequeños patronos) se entiende como una forma de aproximar el tema de las relaciones laborales y de producción.

Un tercer criterio tiene que ver con la pertenencia a cada sector de la economía. Por último, la valoración del nivel educativo y el tipo de ocupación que tienen las personas se torna indispensable, pues determina el tipo de inserción de cada individuo en la economía nacional. Este último paso es muy relevante, pues existen segmentos en los que la inserción laboral no está determinada por la posesión de bienes de producción, sino por la escolaridad, los títulos y conocimientos adquiridos, entre otros, que socialmente son aceptados como credenciales y, por tanto, como diferenciadores de pertenencia. La propuesta de clases sociales utilizada en el análisis se resume en el cuadro 7.3.



PARA MÁS INFORMACIÓN SOBRE METODOLOGÍA DE CLASES SOCIALES

véase el *Decimoquinto Informe Estado de la Nación* (2009), en www.estadonacion.or.cr

CUADRO 7.3

Clasificación de clases sociales

Clase social	Composición
Clase alta	Grandes empresarios, miembros de poderes de la República, dirigentes políticos y empresariales.
Medianos empresarios y expertos	Medianos empresarios, profesionales asalariados o independientes y jefes de departamento con bachillerato.
Clases intermedias	Dirigentes de organizaciones sociales, representantes de gobiernos locales, con y sin bachillerato, técnicos y profesionales medios, jefes de departamento sin bachillerato, dirigentes de organizaciones sociales sin bachillerato, empleados administrativos y de atención al público, empleados del comercio y capataces mayores agrícolas.
Pequeños propietarios	Pequeños productores agrícolas y dueños de pequeños establecimientos de servicios.
Obreros agrícolas	Trabajadores agrícolas en unidades pequeñas y grandes.
Obreros industriales	Trabajadores en establecimientos industriales grandes y pequeños.
Obreros en servicios	Trabajadores de servicios en establecimientos grandes y pequeños.
Otros trabajadores	Trabajadores en otros servicios y trabajadores en unidades domésticas.

Aportes metodológicos en materia de oportunidades, estabilidad y solvencia económicas

Crecimiento económico, ingresos, empleo y sector externo

Con el objetivo de conocer la relación entre el perfil sociodemográfico de los ocupados y la pertenencia a los sectores más dinámicos de la economía, Meneses y Anda (2015) estimaron una serie de modelos *logit* utilizando las bases de datos de las encuestas de hogares (INEC) desde el año 2003. Las autoras analizaron la pertenencia a la nueva economía y los servicios especializados (economía tipo 1), a partir de las diferencias entre variables sociodemográficas de las personas que trabajan en estos sectores y las que no. Los modelos permiten identificar las características que influyen en la pertenencia o no de los ocupados a cada actividad. Una vez evaluados los estadísticos de bondad de ajuste, así como los niveles de significancia de las variables predictoras, se determinó que las variables que mejor explican la inserción laboral en la economía tipo 1 son: el logaritmo natural de la edad del trabajador, residencia en la región Central, sexo, nivel de calificación, grupo de edad, logaritmo natural de los años de escolaridad, nacionalidad extranjera, dominio de un segundo idioma y tenencia de computador. Se debe aclarar que las transformaciones con logaritmo natural buscan normalizar la distribución de la variable, lo que garantiza que los errores se distribuyan con media cero y una varianza σ^2 . Si el coeficiente estimado para una característica es mayor a uno, se espera que los trabajadores que presentan ese rasgo tengan una mayor oportunidad o ventaja de pertenecer a la nueva economía y los servicios especializados. Por ejemplo, esa probabilidad es mayor a medida que aumentan los años de educación.

La significancia global del modelo se prueba mediante la razón de verosimilitud, que resultó estadísticamente significativa para todos los años de estimación del modelo. Esto prueba que existe una relación significativa entre las características sociodemográficas analizadas y la probabilidad de pertenecer a la nueva economía y los servicios especializados.



PARA MÁS INFORMACIÓN SOBRE MODELOS ECONÓMICOS

véase Meneses y Anda, 2015
www.estadonacion.or.cr

Modelos de regresión logística de la Encuesta Nacional de Empleo

Para identificar los factores sociales que potencian la permanencia o inclusión de la fuerza laboral en un determinado sector económico, se indagó acerca de las probabilidades de estar trabajando en diferentes actividades, a saber, agricultura, industria, nueva industria y servicios de apoyo. La investigación empleó las bases de datos de la Encuesta Nacional de Empleo (ECE) para todos los años del período 2011-2014. A partir de esta información se diseñaron modelos de regresión logística que consideran como variable dependiente la pertenencia a cada sector. Como variables predictoras se utilizaron sexo, edad, dominio de un segundo idioma, nacionalidad extranjera, calificación educativa (secundaria completa o más) y residir en la región Central. Los coeficientes de regresión se evaluaron a un 95% de confianza.

En los modelos se mantuvieron las variables que fueron significativas en la mayor parte del período, pues la evidencia empírica demostró que fueron estables a lo largo de la serie y por cuestiones de mero azar no se mostraron significativas en un momento dado. Por el contrario, cuando la variable definitivamente permaneció como no significativa durante el período de análisis, fue descartada. Este procedimiento permite dar seguimiento a los coeficientes y su signo, así como a todas las razones de ventaja en el tiempo, de forma tal que es posible observar constancia, retrocesos o avances en los factores investigados.



PARA MÁS INFORMACIÓN SOBRE MODELOS DE EMPLEO

véase Jiménez-Fontana y Segura, 2015, en
www.estadonacion.or.cr

Impacto de diferentes escenarios de reforma fiscal en los hogares costarricenses

Con el fin de construir una tipología de los hogares costarricenses según sus estructuras de gasto, se llevó a cabo un análisis basado en información de la Encuesta Nacional de Ingresos y Gastos de los Hogares (Enigh) de 2013. Con los datos que provee esta encuesta es posible reconocer segmentos que tienen estructuras de gasto diferenciadas. En primer lugar se realizó un análisis factorial de los gastos mensuales de consumo de productos y servicios. Mediante esta técnica fue posible la reducción de datos a través de la identificación de variables correlacionadas y, como resultado, se obtuvieron catorce dimensiones que explican el 52,8% de la variabilidad.

El paso siguiente fue construir una segmentación, usando el procedimiento de análisis de conglomerados *two step cluster* o en dos fases. Este método permite la selección automática del número más apropiado de conglomerados, ofrece medidas para la selección de los distintos modelos y brinda la posibilidad de generar modelos basados al mismo tiempo en variables categóricas y continuas. Para este ejercicio se incorporaron las catorce variables creadas en el análisis factorial, así como los gastos que no son de consumo (impuestos de renta y no de renta, contribuciones sociales y transferencias de dinero) y los de capital (conservación de la vivienda propia, transacciones financieras y no financieras), transformados mediante normalización. De este modo se llegó a la identificación de seis segmentos con estructuras de gasto diferenciadas, que se caracterizan principalmente por el sexo, la escolaridad y el rango de edad de la jefatura del hogar.

El análisis metodológico se basó en los trabajos de Meneses y Segura (2011) y Barreix et al. (2011). Dada la dificultad para estimar las deducciones por créditos fiscales en las cadenas de producción, las simulaciones no consideraron la aplicación de un impuesto al valor agregado. Para conocer el impacto de distintas modificaciones al impuesto de ventas sobre la recaudación, el ingreso de los hogares, la equidad y la pobreza, se realizaron nueve escenarios de reforma fiscal:



PARA MÁS INFORMACIÓN SOBRE ESCENARIOS DE REFORMA FISCAL véase Meneses y Anda, 2015, en www.estadonacion.or.cr

Aportes metodológicos en materia de fortalecimiento de la democracia

Análisis sobre los liderazgos partidarios

Para analizar la renovación y fluidez de las cúpulas partidarias se construyó una matriz que contiene los nombres de los máximos líderes de cada partido político, con base en la metodología que se detalla a continuación.

1. Identificación de las cúpulas partidarias. El principal criterio utilizado para determinar quiénes conforman estas instancias es el ejercicio de cargos de dirección partidaria o representación política, es decir, puestos que, por su importancia, se espera sean asignados a militantes que tienen un alto perfil de liderazgo y conocimiento dentro de cada agrupación. En todos los casos se identificó a las personas que ocuparon los puestos indicados en el cuadro 7.4.

Cuando una persona ocupó más de un puesto en forma simultánea, solo se tomó en cuenta el de mayor duración o, en su defecto, el de mayor rango. Por ejemplo, si un dirigente aparecía como presidente de un partido y además como candidato

- Eliminación de todas las exoneraciones del impuesto de ventas. Tasa de 15% para todos los bienes y servicios.
- Eliminación de las exoneraciones del impuesto de ventas para la canasta básica tributaria (CBT) y aplicación de una tasa de 15% a estos productos.
- Eliminación de todas las exoneraciones del impuesto de ventas. Tasa de 15% para todos los bienes y servicios, con una compensación para las personas de los tres deciles más bajos, equivalente al monto promedio que paga una persona del segundo decil por este impuesto.
- Impuesto de ventas del 15% para los alquileres mayores a un salario base.
- Impuesto de ventas del 5% para todos los servicios de salud cuyo monto sea superior al 35% de un salario base.
- Impuesto de ventas del 15% para los servicios privados de cirugía y hospitalarios.
- Impuesto de ventas del 5% para los servicios de educación privada cuyo monto sea superior al 35% de un salario base.
- Impuesto de ventas del 15% para el consumo de energía eléctrica superior a 250 kW.

- Impuesto de ventas del 15% para el consumo de agua residencial mayor a cuarenta metros cúbicos.

Para conocer el impacto de cada escenario en los hogares y en el fisco se analizaron varios indicadores. El efecto en el gasto se midió en relación con el ingreso del hogar. Se utilizó la variación entre el gasto actual y el escenario, y se calculó la proporción que representa ese cambio en el ingreso corriente neto del hogar, sin valor locativo. En el ámbito de la recaudación fiscal, se estimó el porcentaje de los ingresos obtenidos mediante el impuesto de ventas con respecto al PIB de 2013. La recaudación actual se definió como la diferencia entre el gasto original (mensualizado de los hogares) sin impuesto y el gasto original. Estas consideraciones fueron la base para estudiar los escenarios fiscales planteados. Las estimaciones asumen que los bienes y servicios no exonerados efectivamente pagaron impuestos. Además, el impacto en la recaudación se midió bajo el supuesto de que no existe evasión o elusión.

También se analizó el impacto de las reformas sobre la equidad, con el fin de determinar si los escenarios son regresivos, progresivos o neutros. La equidad se midió como el cambio en el gasto de los hogares por deciles de ingreso. Finalmente, se calculó el efecto en la pobreza, entendido como el cambio en la proporción de la población que se ubica por encima de la línea de pobreza.

CUADRO 7.4

Criterios para la identificación de los líderes que conforman las cúpulas partidarias

Nivel	Criterios de selección	Fecha de inicio en el puesto	Fecha de salida del puesto
Partidario	Candidato(a) a la Presidencia de la República	Día en que inició formalmente el proceso electoral	Día en que se realizó la elección para el puesto al cual se postuló
	Presidente(a) del partido político	Día en que la asamblea del partido lo(a) nombró en el cargo	Fin del período en el cargo
	Secretario(a) General del partido político	Día en que la asamblea del partido lo(a) eligió en el puesto	Fin del período en el puesto
Legislativo	Jefe(a) de fracción legislativa	Inicio de la legislatura en que ostentó el cargo, considerando los cambios en las jefaturas.	Fin del período en el cargo
	Candidatos a diputados en puestos elegibles según antecedente ^{a/}	Día en que inició formalmente el proceso electoral	Día en que se realizó la elección para los puestos a los cuales se postulaban

a/ Para cada período electoral, se consideró la cantidad de curules obtenidas en los comicios anteriores. En todos los casos, y en especial para la primera participación electoral de cada partido, se tomó en cuenta al menos los primeros lugares de cada provincia.

presidencial, para efectos del conteo de días de permanencia en la cúpula, solo se consideró el primero de esos cargos. La función como candidato presidencial se desempeña por menos tiempo –los cuatro meses que dura la campaña electoral– y además ello ocurre en el mismo lapso ya contabilizado. De esta forma se evitó la duplicación de días asignados a cada persona.

La matriz además registra otras variables: sexo, puesto en la cúpula, fecha de inicio, fecha de salida, provincia y lugar en la papeleta (las dos últimas para el caso de las candidaturas a diputados).

2. Selección de partidos políticos.

Se trabajó con información de siete agrupaciones de escala nacional a las que se viene dando seguimiento desde 2012, como parte de un proyecto de investigación del PEN y el Instituto de Formación y Estudios en Democracia (IFED), del TSE. Se trata de los siguientes partidos: Accesibilidad Sin Exclusión (PASE), Acción Ciudadana (PAC), Frente Amplio (FA), Liberación Nacional (PLN), Movimiento Libertario (ML), Renovación Costarricense (PRC) y Unidad Social Cristiana (PUSC). El período de análisis fue el comprendido entre el 1 de enero de 1982 y el 31 de mayo de 2015; este abarca la totalidad de años de existencia de las agrupaciones

consideradas, con excepción del PLN, que fue fundado en 1952. En el cuadro 7.5 se detalla el período específico que cubrió el estudio de cada cúpula partidaria.



PARA MÁS INFORMACIÓN SOBRE PARTIDOS POLÍTICOS

véase Gómez Campos y Sáenz, 2015, en www.estadonacion.or.cr

Comparación de los Planes Nacionales de Desarrollo 2006-2018

Se revisaron los Planes Nacionales de Desarrollo (PND) de las administraciones Arias Sánchez (2006-2010), Chinchilla Miranda (2011-2014) y Solís Rivera (2015-2018), con el objetivo de identificar similitudes y diferencias significativas en la visión de desarrollo planteada por cada gobierno.

Se construyó una base de datos en la que se recogieron todas las acciones estratégicas y programas institucionales incluidos en los PND, y en cada uno de ellos se identificó una serie de características: los temas abordados, las instituciones involucradas en el desarrollo de los programas o acciones, el tipo de financiamiento y el grado de similitud con los planteamientos de las otras administraciones.

Para conocer las similitudes se compararon, uno a uno, los objetivos específicos señalados para las acciones estratégicas y programas institucionales de cada PND. De esta forma fue posible establecer tres niveles de similitud: baja (coincidencia en menos de la mitad de los objetivos), alta (coincidencia en más de la mitad de los objetivos), y media (coincidencia en la mitad de los objetivos).

También se determinó si la acción estratégica o programa institucional era prioritario o no, tomando como base las prioridades expresadas por cada presidente en su discurso de toma de posesión (1 de mayo del primer año de gobierno), así como el contenido de los capítulos que definen la visión y los objetivos de desarrollo en los respectivos PND.



PARA MÁS INFORMACIÓN SOBRE PLANES NACIONALES DE DESARROLLO

véase Ramírez Cover, 2015, en www.estadonacion.or.cr

Análisis de la gestión del Poder Ejecutivo

En este Vigésimoprimer Informe se analiza la eficacia del Poder Ejecutivo en el control de la agenda legislativa y el manejo político en cuatro coyunturas

CUADRO 7.5

Período de análisis de las cúpulas partidarias, según partido político. 1982-2015

Partido	Fecha de fundación	Fecha del primer registro ^{a/}	Total de días abarcados ^{b/}	Tiempo total en años ^{c/}
PLN	15 de abril de 1952	1 de enero de 1982 ^{d/}	12.203	33,4
PUSC	17 de diciembre de 1983	18 de enero de 1984	11.456	31,4
ML	2 de diciembre de 1994	2 de diciembre de 1994	7.485	20,5
PRC	9 de diciembre de 1995	3 de junio de 1995 ^{e/}	7.302	20,0
PAC	3 de diciembre del 2000	20 de mayo de 2001	5.124	14,0
PASE	21 de agosto de 2004	1 de septiembre 2004	3.924	10,8
FA	16 de octubre de 2004	22 de mayo de 2005	3.661	10,0

a/ Se consideró la fecha de creación del primer comité ejecutivo que consta en el expediente oficial de cada partido ante el TSE, con excepción del PLN.

b/ Los días totales se obtienen al contabilizar los días transcurridos entre la fecha del primer registro y la fecha de corte, que para todos los partidos fue el 31 de mayo de 2015.

c/ El tiempo en años se obtiene de dividir los días totales entre 365.

d/ En el caso del PLN, se considera el comité ejecutivo vigente a partir del 1 de enero de 1982, que corresponde a la fecha de inicio del estudio.

e/ En el expediente oficial del PRC consta un comité ejecutivo provisional que llevó a cabo el proceso de inscripción oficial del partido.

conflictivas de alto perfil público. Con ese propósito se construyeron dos bases de datos, que se describen a continuación.

1. Tasa de éxito del Ejecutivo en sesiones extraordinarias. Se hizo una revisión exhaustiva de los proyectos convocados en las primeras legislaturas de las últimas siete administraciones (1990-2018). Para ello se recopilieron los decretos de convocatoria publicados en el diario oficial *La Gaceta* desde 1990 y, en el caso de la legislatura más reciente, se consultó el sitio web de la Asamblea Legislativa. Para determinar la tasa de éxito se calculó la relación entre las leyes aprobadas (utilizando una matriz suministrada por el Departamento de Servicios Parlamentarios de la Asamblea Legislativa) y el total de proyectos convocados en el período de estudio.

2. Apoyos y oposiciones hacia el Ejecutivo en conflictos de alto perfil público. Se identificaron las relaciones de apoyo y oposición al Poder Ejecutivo que mostraron diversos actores en cuatro conflictos de alto perfil público ocurridos durante el 2014. Esos episodios fueron, quizás, los de mayor tensión política del primer año de gobierno del presidente Solís Rivera. Dos de ellos implicaron serios enfrentamientos con algunos sectores sociales: la huelga de educadores y

el conflicto por el contrato con la empresa APM Terminals para la construcción de la nueva terminal de contenedores en el puerto de Moín, y los otros dos fueron parte de procesos institucionales: la aprobación del Presupuesto de la República para el 2015 y el levantamiento del veto presidencial a la reforma procesal laboral.

Se elaboró una matriz con información sistematizada de notas de prensa, artículos de opinión y campos pagados en tres periódicos nacionales: *La Nación*, el *Diario Extra* y el *Semanario Universidad*. Con base en ello fue posible identificar los actores que participaron y se manifestaron en cada uno de los conflictos, y sus posiciones específicas (cuadro 7.6).



PARA MÁS INFORMACIÓN SOBRE
**GESTIÓN Y DESEMPEÑO
DEL PODER EJECUTIVO**
véase Alpízar y Menocal, 2015, en
www.estadonacion.or.cr

Aportes metodológicos en materia de ordenamiento territorial

Relación de la renta del suelo con costos de transporte, las pendientes y el uso actual del suelo

Se desarrolló un modelo econométrico para estimar la relación entre la renta del suelo y variables como los costos de trans-

porte, las pendientes y el uso actual del suelo, específicamente en la Gran Área Metropolitana. Esto se hizo en dos pasos. Primero se determinaron los valores del suelo, a partir de los datos de hipotecas otorgadas por el Banco Nacional de Costa Rica, así como información de venta de terrenos y viviendas (en campo, internet y medios de comunicación escrita). Los valores por metro cuadrado fueron geolocalizados asignando las coordenadas de la propiedad, con datos tomados de la página web del Registro de Bienes Inmuebles. Los valores se pasaron de colones a dólares estadounidenses corrientes y luego a dólares reales de 2012, con la inflación de Estados Unidos. En total los registros geolocalizados por área metropolitana fueron: 2.315 para San José, 1.923 para Alajuela, 1.321 para Cartago y 6.213 para Heredia. El segundo paso fue analizar la relación entre las variables antes mencionadas (costos de transporte, pendientes y el uso actual del suelo). Para ello se tomó el caso del área metropolitana de Heredia, la región que concentra la mayor cantidad de observaciones (6.213), con el fin de obtener resultados más robustos (Sánchez, 2015).

La estimación del modelo de regresión implicó la transformación de mapas en celdas discretas, que luego se relacionaron mediante técnicas econométricas. Un registro queda compuesto por los valores de las distintas variables que se

CUADRO 7.6

Metodología para medir los niveles de apoyo y oposición al Poder Ejecutivo

Posición	Nivel	
	Fuerte	Débil
Oposición	Actores no tienen voluntad de negociar con el Poder Ejecutivo.	Actores están en contra de la acción realizada por el Ejecutivo, pero pueden cambiar de posición si se modifica la medida de acuerdo con sus intereses.
	Se utilizan medidas institucionales o no institucionales con el fin de lograr objetivos opuestos a los del Ejecutivo	No se amenaza con usar una vía de presión y solo se brinda una opinión en contra del Ejecutivo.
Apoyo	Actores intentan cooperar en la solución del conflicto, sin culpar al Ejecutivo de este.	Actores están a favor de la acción realizada por el Ejecutivo, pero pueden cambiar de posición si se modifica la medida en contra de sus intereses.
	Actores con potestad institucional de tomar decisiones, utilizan este recurso para apoyar al Ejecutivo.	Esta fue una categoría residual, es decir, en ella se ubicaron todas las muestras de apoyo que no cumplían con la característica de "apoyo fuerte".

ubican en el mismo lugar (misma coordenada geográfica). Estos modelos, que la literatura especializada denomina “estadísticos-espaciales”, típicamente han sido aplicados a problemas de cambio de uso del suelo.

Cabe destacar que el modelo de valor del suelo comparte características (y problemas metodológicos) con los modelos de cambio de uso del suelo. En particular, al ser las bases de datos producto de “discretizar” mapas (procesos espacialmente continuos), existe un importante riesgo de que los resultados sean sesgados por procesos de dependencia espacial (Irwin, 2010).

Por lo general este sesgo ha sido ignorado o mitigado con variables espacialmente explícitas. Por simplicidad, para reducir el sesgo por dependencia espacial se han introducido las posiciones X y Y de las coordenadas (empleadas por Müller y Zeller, 2002 y Hu y Lo, 2007 en el contexto de modelos estadísticos-espaciales de cambio de uso del suelo). Para los fines del estudio presentado en este Informe se planteó el siguiente modelo:

$$VS10 = \beta_0 + \beta_1 TVH + \beta_2 TVSJ + \beta_3 PEND + \beta_4 USOURB + \mu \quad [1]$$

donde:

- **VS10:** es el valor por metro cuadrado del suelo en dólares de 2012, según el tipo de cambio del día en que se registró la hipoteca.
- **TVH:** es el tiempo de viaje estimado en minutos al centro de Heredia (se suponen condiciones de flujo libre).
- **TVSJ:** es el tiempo de viaje estimado en minutos a la más cercana de las siguientes intersecciones: i) entre el límite de Heredia y la ruta nacional 1, ii) entre el límite de Heredia y la ruta nacional 3, iii) entre el límite de Heredia y la ruta nacional 5, y iv) entre el límite de Heredia y la ruta nacional 32. Estas son las cuatro vías más importantes que comunican la zona de estudio con San José; se suponen condiciones de flujo libre.
- **PEND:** es el porcentaje de pendiente estimado a partir del modelo de elevación digital (construido a partir de la cartografía nacional 1:10 000 para las zonas en que ésta existe y, en caso contrario, empleando la cartografía de Ceniga, escala 1: 25 000; se estimó

con celdas de cinco metros de lado para mantener consistencia con otras capas).

- **USOURB:** es una variable categórica igual a 1 si el uso del suelo era urbano en 2010 y 0 si no lo era. La extensión de las capas que se analizan está definida por la intersección entre la imagen satelital que se empleó para definir el uso del suelo y el límite sur de la ciudad de Heredia: el río Virilla.
- μ es un error aleatorio con distribución normal y media 0.

Se estimó un segundo modelo con las posiciones X y Y de cada celda (tomando como X = 0 y Y = 0 la celda superior izquierda del rectángulo en que se inscriben todas las celdas).

$$VS10 = \beta_0 + \beta_1 TVH + \beta_2 TVSJ + \beta_3 PEND + \beta_4 USOURB + \beta_5 PosX + \beta_6 PosY + \mu$$

Los resultados del modelo con los determinantes del valor del suelo se muestran en el cuadro 7.7.

CUADRO 7.7

Determinantes del valor del suelo^{a/}
(dólares estadounidenses de 2012 por metro cuadrado)

Variable	Modelo 1			Modelo 2		
	Coefficiente	Estadístico t	Probabilidad	Coefficiente	Estadístico t	Probabilidad
Tiempo de viaje a San José - condiciones de flujo libre (minutos)	-0,150	-1,439	<0,001	-0,110	-531,030	<0,001
Tiempo de viaje a Heredia - condiciones de flujo libre (minutos)	-0,030	-388,490	<0,001	-0,062	-247,470	<0,001
Uso del suelo (urbano = 1)	11,170	337,260	<0,001	111,653	336,563	<0,001
Pendiente (porcentaje)	-0,010	-14,240	<0,001	0,007	9,703	<0,001
Posición X				0,006	73,540	<0,001
Posición Y				0,009	125,236	<0,001
Intercepto	162,820	3,145	<0,001	1.328.995	478,437	<0,001
R ²	0,580			0,590		
R ² ajustado	0,580			0,590		
N	3.679.948			3.679.948		

a/ Ambos modelos son estadísticamente significativos.

Fuente: Sánchez, 2015, con datos de Pujol y Pérez, 2013.

Factores determinantes de la densidad urbana según diversos índices

Para conocer las variables que afectan el mercado de tierras, se evaluó la influencia de varios factores en la expansión de la urbanización y, por lo tanto, en la estructura urbana. Con ese propósito se crearon modelos econométricos que explican la cantidad, el crecimiento y la estructura de los patrones de área construida en la región. A partir de esos datos se pueden identificar factores que ocasionan cambios en los niveles y estructura del área construida (una mayor comprensión de las dinámicas asociadas al crecimiento urbano permite guiar eficientemente las intervenciones sobre los mercados de tierras, tanto en su objetivo como en sus instrumentos y aplicación) y generar información sobre las variables que deben controlarse para aislar de manera apropiada el fenómeno. En particular, en el desarrollo de modelos causales es fundamental para distinguir influencias que puedan confundir la medición del impacto cuya causalidad se investiga.

Tomando como variables dependientes la densidad poblacional, el índice de dispersión, el índice de contigüidad, el índice de compacidad, el área construida y la tasa de crecimiento, se hicieron cálculos de modelos econométricos de efectos aleatorios para identificar variables que podrían estar determinando la estructura urbana de la región. Las definiciones de las variables utilizadas son las siguientes:

- **Índice de dispersión:** para todas las celdas que se analizan (todas las celdas urbanas o todas las celdas correspondientes al crecimiento urbano), se estima el porcentaje de celdas no urbanas en una ventana de aproximadamente 1 km x 1 km centrada en la celda que se considera. Luego se promedia el valor de todas las celdas (área construida total o crecimiento del área construida) para toda la región o una parte de ella (por ejemplo un municipio). Si el valor es muy cercano a la unidad, los niveles de dispersión son muy altos, porque el desarrollo urbano representado por la celda está aislado (rodeado de muchas celdas no urbanas, es decir, alejado de

las áreas construidas consolidadas); a la inversa, si el valor es muy bajo, el nivel de dispersión también lo es.

- **Índice de compacidad:** se estima como el área construida consolidada principal (la mancha urbana contigua de mayor área de cada ciudad) dividida por el área desarrollable. Angel et al. (2005) definen el área desarrollable como el círculo de radio mínimo que contiene el área construida de la ciudad, excluyendo desarrollos muy dispersos (considerando solo áreas construidas que ocupen, contiguamente, 25 hectáreas). El índice se modificó en el marco de este estudio: se calcula como el área construida de cada municipio dividida por el área total edificable (con una pendiente menor al 30%) de ese municipio.
- **Índice de contigüidad:** se estima como el área de la mancha urbana continua más grande dividida por el área construida total. El índice se modificó en el marco de este estudio: se calcula el área construida contigua más grande dentro de los límites de un municipio, partida por el área construida total de ese municipio (se supone que cada municipio está aislado en sus bordes del resto de la ciudad, es decir, que las manchas urbanas se interrumpen en el límite cantonal aun en los casos en que éstas físicamente son continuas). Para la región en su conjunto, el índice de contigüidad se calcula como el área construida continua de mayor área (centrada en San José) dividida por el área construida total.
- **Modelos de precios hedónicos:** un modelo de precios hedónicos describe el precio de bienes que tienen varios atributos. Cuando un consumidor compra uno de estos bienes, en la práctica está adquiriendo una canasta que incluye todos sus atributos, es decir, obtiene utilidad de los atributos del bien, antes que del bien mismo. Así, es posible descomponer el precio del bien en función de sus atributos.
- **Propensity score matching (PSM):** el análisis PSM es una técnica no

paramétrica que permite comparar un grupo tratado con un grupo de control, seleccionando solo casos similares. Consiste en una extensión de las técnicas de experimentos, en la que el tratamiento se aplica de manera aleatoria a una serie de casos, a series de datos existentes donde el "tratamiento" no es aleatorio. Fue propuesta por Rosenbaum y Rubin en 1983 y ha sido ampliamente utilizada en diversos campos desde entonces.

El procedimiento para realizar un análisis PSM consiste en:

- Estimar una regresión de variable discreta (*logit*, *probit*) definiendo como variable dependiente el grupo tratado y predecir la probabilidad de pertenecer al grupo de control de cada dato. Esta probabilidad se convierte en el *propensity score*.
- Utilizar algún algoritmo de emparejamiento que seleccione datos con características (probabilidad) similares. La clave es que en cada "pareja" haya un dato en el grupo de control por cada uno o varios datos similares del grupo tratado. La similitud se evalúa con base en el *propensity score*.
- Realizar una prueba de diferencia de dos medias entre los grupos de datos seleccionados. La diferencia de medias es conocida como el parámetro ATT (*average treatment on the treated*) y representa el efecto neto del tratamiento.

En los análisis realizados, se estimó un modelo *logit* usando como variable dependiente LCU (variable dicotómica igual a 1 para las propiedades más allá del límite de crecimiento urbano). El argumento es el siguiente: si el límite de crecimiento responde a las mismas variables que el valor del suelo —una suposición razonable, dado que el valor del suelo es reflejo de las decisiones de localización de los agentes económicos en la ciudad y el límite de crecimiento urbano fue definido precisamente para modificar esas decisiones— entonces debe ser explicado, al menos parcialmente,

por los mismos factores que determinan ese valor del suelo.

Una vez estimado el modelo que predice el *propensity score*, se aplicó el algoritmo propuesto por Cochran y Rubin (1973) para asociar dos registros tratados con cada registro no tratado. Con base en este emparejamiento, se estimaron

diferencias entre las medias de todos los registros (que es un resultado sesgado y solo sirve como parámetro de comparación) y de los registros emparejados.

Los resultados de los modelos con los factores determinantes del crecimiento urbano en la GAM se muestran en los cuadros 7.8, 7.9 y 7.10.

PARA MÁS INFORMACIÓN SOBRE
DETERMINANTES DE LA DENSIDAD URBANA

véase Sánchez, 2015, en
www.estadonacion.or.cr

CUADRO 7.8

Factores determinantes del crecimiento urbano en la Gran Área Metropolitana

Variable	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6	Modelo 7
	Densidad poblacional	Índice de dispersión	Índice de contigüidad	Índice de compacidad	Área construida nueva	Tasa de crecimiento interanual equivalente	Índice de dispersión
Intercepto	62,180 -4,190	-0,420 0,800	0,060 0,120	14,183 -1,529	-69,762 -2,173	-47,587 -2,110	-12,350 -2,734
Distancia a San José	-0,340 -3,311	0,160 -5,031	-0,090 -2,925	-0,196 -3,369	0,612 -3,072	0,240 -1,712	0,196 -6,975
Distancia a ciudad	-0,080 -1,652	-0,020 -1,980	0,030 -2,167	-0,018 0,600	-0,181 -1,492	-0,189 -2,238	-0,001 -0,075
Distancia a zona industrial	0,200 -2,198	0,050 -2,078	0,010 0,500	0,013 0,264	0,150 0,807	0,230 -1,772	0,051 -1,955
Trabajadores (normalizado)	-0,020 0,560	0,001 0,160	0,020 -1,687	-0,040 -1,679	0,088 0,764	-0,019 0,229	0,039 -2,415
Población estimada		-0,080 -5,320	0,040 -2,769	0,093 -2,921	0,573 -4,694	-0,240 -2,782	-0,045 -2,591
Índice de ingresos	-0,030 -1,216	-0,020 -3,470	0,020 -3,206	0,065 -3,765	0,155 -1,883	0,047 0,687	0,004 0,309
Pendiente promedio	-0,005 0,400	0,002 0,520	0,004 -1,021	0,002 0,232	0,051 -1,892	0,031 -1,605	-0,003 -0,848
Densidad de pozos (2 años) ^{a/}	-0,080 -2,908	0,010 -3,157	-0,010 -1,910	-0,016 0,996	0,218 -2,118	0,002 0,026	-0,003 -0,195
Porcentaje de área dentro de LCU	-0,200 0,920	-0,080 -1,164	0,480 -7,184	0,195 -1,606	0,267 0,647	0,121 0,419	-0,183 -3,139
Uso agrícola hortalizas (1 = sí)	0,250 -1,706	-0,030 0,710	-0,010 0,290	-0,028 0,343	-0,915 -2,994	-0,721 -3,326	0,002 0,046
Uso agrícola café (1 = sí)	0,040 0,550	0,060 -3,970	-0,050 -2,771	-0,090 -2,261	0,526 -3,510	0,565 -4,840	0,045 -2,053
Estadístico f	4,717 10 y 82 gl	40,800 11 y 81 gl	43,309 11 y 81 gl	16,175 11 y 81 gl	9,310 11 y 50 gl	7,808 11 y 50 gl	32,294 11 y 50 gl
R ² ajustado	0,322	0,738	0,855	0,599	0,542	0,510	0,707
Theta	0,455	0,697	0,598	0,451	0,319	0,136	0,266
N	93	93	93	93	62	62	62

a/ La densidad de pozos se estima con base en la cantidad de permisos de explotación aprobados por el Servicio Nacional de Aguas Subterráneas, Riego y Avenamiento (Senara) en el año en cuestión y el inmediatamente anterior (por ejemplo, para 1986 se consideran los permisos de 1985 y 1986). Con esa información se calcula la densidad de pozos para un mapa de celdas de 30 m x 30 m, en un radio de 10 km a partir del centro de cada celda. Sobre este mapa continuo, se obtiene el valor promedio de las celdas ubicadas dentro de los límites de cada municipio en cada período.

Fuente: Sánchez, 2015.

CUADRO 7.9

Resultados de propensity score matching: impacto del límite de crecimiento urbano sobre los valores del suelo

	Modelo 1	Modelo 2	Modelo 3	Modelo 4
Tratamiento	Fuera de límite de crecimiento urbano			
Variable respuesta	Valor del suelo (dólares por metro cuadrado, dólares reales de 2012)			
Cantidad de observaciones				
Con tratamiento	708	220	282	34
Sin tratamiento	379	120	183	18
Diferencia - todos los registros				
Diferencia de medias	-55,360	-67,080	-76,020	-34,400
Error estándar	1,980	5,170	6,170	4,500
Estadístico t (probabilidad)	-27,995 (<0,0001)	-12,963 (<0,0001)	-12,324 (<0,0001)	-7,645 (<0,0001)
Efecto promedio del tratamiento sobre los datos				
Diferencia de medias	-14,100	-45,090	-59,860	-22,190
Error estándar	3,650	16,150	19,130	6,940
Estadístico t (probabilidad)	-3,866 -0,001	-2,792 (<0,005)	-65,033 -0,002	-3,199 -0,001

Fuente: Sánchez, 2015.

CUADRO 7.10

Resultados de los modelos econométricos para determinar los valores del suelo

Variable	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Coefficiente	Estadístico t (probabilidad)	Coefficiente	Estadístico z (probabilidad)	Coefficiente	Estadístico z (probabilidad)	Coefficiente	Estadístico t (probabilidad)
Intercepto	86,645	60,134 (<0,001)	58,910	13,347 (<0,001)	121,950	18,640 (<0,001)	70,980	22,462 (<0,001)
Razón valor de construcción / valor de suelo [BtoLR]	-0,014	-5.521,000 (<0,001)	-0,011	-3,653 (<0,001)	-0,012	-3,710 (<0,001)	-0,001	-0,014 -0,989
Área de parcela [AREA]	-0,357	-16,224 (<0,001)	-0,323	-268,103 (<0,001)	-0,342	-25,727 (<0,001)	-0,276	-10,806 (<0,001)
Límite de crecimiento urbano (fuera = 1) [LCU]	-0,114	-3,850 (<0,001)	-0,019	-0,286 -0,775	-0,128	-2,279 -0,002	-0,169	-2,381 -0,018
Tiempo de viaje a San José - flujo libre [SJTT]	-0,186	-12,849 (<0,001)	-0,145	-5,671 (<0,001)	-0,313	-6,895 (<0,001)	-0,135	-3,657 (<0,001)
Tiempo de viaje a Heredia-Alajuela - flujo libre [HETT]	-0,153	-9,038 (<0,001)	-0,080	-30,708 -0,002	-0,222	-5,859 (<0,001)	-0,122	-2,785 -0,006
Pendiente [SLP]	-0,001	-0,562 -0,574	-0,001	-0,047 -0,962	-0,002	-0,333 -0,739	0,011	1,109 -0,268

Modelo 1 Ecuación [1] Método de estimación: mínimos cuadrados ordinarios corregidos por heterocedasticidad.
Modelo 2 Ecuación [2] Rezago espacial / Método de estimación: máxima verosimilitud.
Modelo 3 Ecuación [3] Rezago espacial y error espacial / Método de estimación: máxima verosimilitud.
Modelo 4 Ecuación [4] Sin evidencia de procesos espaciales ni de heterocedasticidad / Método de estimación: mínimos cuadrados ordinarios.

Fuente: Sánchez, 2015.

Este Anexo Metodológico fue preparado

por Natalia Morales, Pamela Jiménez, Karen

Chacón, Steffan Gómez y María Estelí Jarquín.

NOTAS

1 El lector interesado en las bases teóricas de la metodología, así como en las aplicaciones de insumo producto puede consultar: Miller y Blair (2009), Palomino y Pérez (2011) y Kozikowski (1988).

2 Si la productividad del trabajo aumenta, tal como ha sucedido en Costa Rica en los últimos años, serán necesarios menos trabajadores para obtener una cierta cantidad de producto.

3 Para mayor detalle véase Hernández (2012) y Palomino y Pérez (2011).

4 Mediante supuestos de cambios en la relación del producto por empleado con base en datos históricos, se podrían hacer simulaciones de empleo generado a nivel sectorial o, eventualmente, suponer que todos los sectores incrementan su productividad en cierto porcentaje. Esto se plantea como un desafío de investigación para futuras ediciones de este Informe.

5 Es preciso señalar dos limitaciones de la base de datos. Por un lado, utiliza una única fuente (avisos en el periódico *La Nación*), pese a que otros medios de circulación nacional también publican ofertas de puestos de trabajo. Por otro lado, se construyó en un momento en que el uso creciente de las tecnologías de información y comunicación ha provocado un cambio en las estrategias de las empresas para el reclutamiento de personal y una parte importante de las ofertas (que no es posible cuantificar) se divulga en sitios especializados en internet.

6 En Medina y Galván (2008) se detalla y explica la descomposición utilizada y se aplica a los países de América Latina. El análisis realizado en el presente Informe se basa en ese trabajo, que fue sintetizado por Trejos y Oviedo (2012).

7 Persona de 12 años o más que trabajó al menos una hora en el período de referencia o, si no lo hizo, tiene un empleo o trabajo que no realizó por motivos circunstanciales, planeados o inesperados. A partir del 2010 se calcula para las personas de 15 años o más.

8 Para más detalles, véase el capítulo 6 del *Decimoquinto Informe Estado de la Nación*.