

■ Nota técnica sobre el Índice de Democracia Electoral

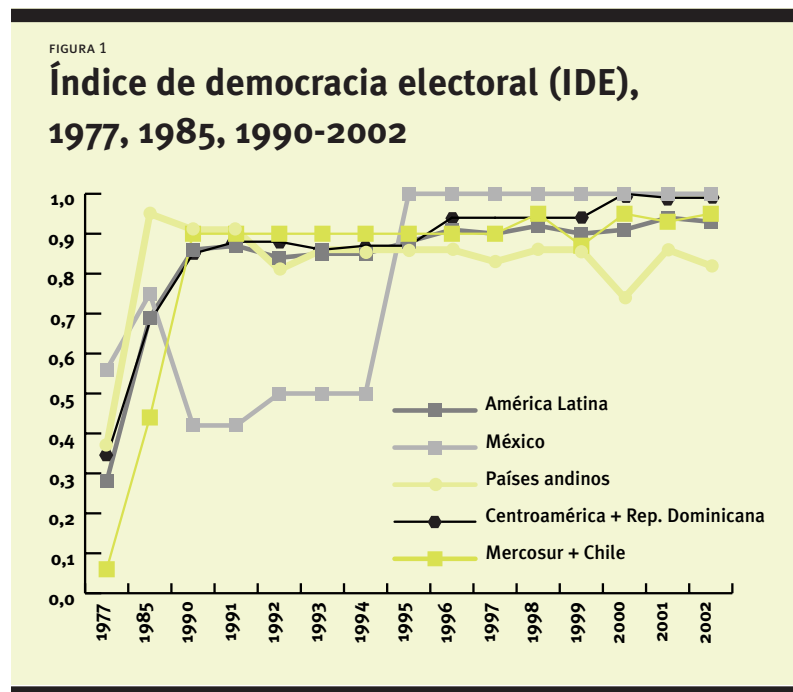
Esta nota describe los pasos dados y las pruebas estadísticas realizadas para la construcción del Índice de Democracia Electoral (IDE), una medida compuesta sobre los derechos políticos relacionados con la elección de los gobiernos. Se presentan además aclaraciones sobre su interpretación y utilización.

Construcción del IDE

La elección de los componentes

El primer paso para la construcción del IDE, y probablemente el más importante, consistió en la elección de sus cuatro componentes: sufragio, elecciones limpias, elecciones libres y cargos públicos electos. Ellos fueron seleccionados considerando los elementos centrales tradicionalmente invocados por los teóricos sobre la democracia para la definición de un régimen democrático, y abarcan una serie de temas que en general han sido considerados centrales, incluso necesarios, para cualquier evaluación sobre el carácter democrático de un régimen político (figura 1).

En segundo lugar, estos elementos se refieren a derechos de la ciudadanía, cuya vigencia es responsabilidad del Estado, y que pueden ser interpretados claramente en términos de la teoría de la democracia vigente. De este modo, se evitan problemas asociados con elementos tales como la concurrencia de los votantes a las urnas o la desproporcionalidad electoral, que reflejan tanto las acciones estatales como las de los ciudadanos. Ello permite garantizar que el índice pueda ser interpretado claramente como una medida del grado en que el Estado garantiza los derechos de la ciudadanía referidos al régimen político, de modo diferenciado de la acción de los ciudadanos. De la misma manera, se evitan problemas relacionados con medidas de significación poco claras respecto del grado de democracia de un régimen, tales como



Nota: El gráfico se basa en datos presentados en el Compendio Estadístico del Informe.

la diferencia entre reglamentaciones electorales proporcionales y mayoritarias, o entre sistemas presidencialistas o parlamentaristas. Estos aspectos no carecen ciertamente de importancia, pero no están tan claramente conectados con el grado de democracia de un régimen como los cuatro elegidos.

En tercer lugar, estos elementos posibilitan contar con datos válidos y confiables del último año calendario. En este sentido, se ha puesto énfasis en la medición de componentes estrictamente observables, evitando el uso de encuestas sobre percepciones. Así, algunos componentes que podrían haber sido incluidos fueron dejados de lado por razones “prácticas”.

Por último, se analizó un conjunto de factores que podrían haber sido incluidos y no lo fueron, en gran parte debido a la dificultad de desarrollar mediciones apropiadas y a tiempo para este primer informe. Ello incluye factores asociados al ejercicio del dere-

cho al voto, tales como el proceso de obtención de documentos de identidad, la inscripción o registro para votar y la votación misma, y las condiciones para la competencia libre, afectada por factores tales como el financiamiento de partidos y campañas, el uso de recursos públicos, el acceso a los medios de comunicación y la libertad de prensa. Otras cuestiones importantes involucran a las prácticas electorales en el nivel subnacional y a la estabilidad del régimen. Se deja como tema para futuras discusiones, la conveniencia de la construcción de nuevos índices.

La medición de los componentes

El segundo paso en la construcción del IDE –la medición de sus cuatro componentes– requirió tomar dos decisiones clave. La primera tiene que ver con las reglas del proceso de codificación y la segunda, con el proceso de codificación en sí.

En cuanto a las reglas del proceso de codificación, las escalas –tres ordinales de cinco puntos y una ordinal de tres puntos– se construyeron determinando primero los puntos finales teóricamente significativos, y luego identificando distintos valores de escala lo más alejados conceptualmente como fuera posible, comenzando con el punto medio. Los valores de la escala fueron elegidos para reflejar distinciones relevantes de la bibliografía, evitando pequeñas variaciones entre casos, aunque fueran verificables. En los casos que no correspondían con precisión a ninguno de los puntos de las escalas ordinales, se introdujo el uso de signos más y menos como modo de registrar valores intermedios. Las escalas se construyeron también de manera que cada punto se correspondiera con situaciones y acontecimientos relativamente concretos, y que las decisiones de su codificación pudieran ser tomadas estrictamente sobre la base de observables. No se han incluido en el índice datos basados en encuestas sobre percepciones.

Más aún, como una forma de asegurar la replicabilidad del ejercicio de codificación, y frente a posibles arbitrariedades, se subrayó la importancia de documentar las bases de las decisiones de codificación a través de la referencia a fuentes de información disponibles públicamente. No se han requerido puntajes

para cada caso en cada año. Antes bien, sólo fueron requeridos los puntajes para tres de los componentes –sufragio, elecciones limpias y elecciones libres– para los años en los que se celebraron elecciones. Las condiciones para las elecciones están afectadas por acontecimientos y decisiones tomadas entre las elecciones, y la codificación ha tomado información entre los períodos electivos. Pero la significación de estos acontecimientos y decisiones para el proceso por el cual los actores obtienen el acceso a los cargos gubernamentales, que es el interés central del ejercicio de medición, se cristaliza en el acontecimiento electoral en sí mismo. Así, aunque los puntajes fueron asignados a algunos componentes sólo durante los años de elecciones, estos puntajes han sido entendidos como síntesis de procesos más amplios.

El segundo conjunto de decisiones se refiere al proceso de codificación en sí. En este sentido, se han utilizado dos procedimientos complementarios para codificar los casos. Un codificador particular realizó una codificación basándose en una investigación extensa y en consultas con numerosos expertos, a lo largo de un período de muchos meses. Los puntajes asignados fueron presentados y discutidos en profundidad en varios encuentros, incluido uno con un grupo de participantes invitados que trabajaban en diversos contextos (política, academia, organizaciones internacionales), provenientes de diferentes países de las Américas (Argentina, Brasil, Canadá, Colombia, Ecuador, Estados Unidos, México y Uruguay). Estas discusiones condujeron a la identificación de desacuerdos que llevaron a sucesivas investigaciones y más discusiones grupales. Finalmente, luego de este proceso iterativo, emergió un alto grado de consenso en lo que respecta a la codificación de las cuatro dimensiones del IDE.

La generación de una base de datos rectangular con escalas normalizadas

El tercer paso en la construcción del IDE fue la transformación de los puntajes en las escalas de los componentes en una base de datos rectangular, es decir, una base de datos que incluye puntajes numéricos para todos los casos en todas las variables y todos los años, con escalas normalizadas.

Este paso implicó una serie de procedimientos. Primero, se enfrentaron aspectos bastante mecánicos. Los signos más y menos fueron convertidos en números, agregando y sustrayendo 0,33 del puntaje base (por ejemplo, un 3+ fue convertido en un 3,33). Los guiones (-) utilizados para indicar que la asignación de un puntaje no era aplicable, debido a que el gobierno no provenía de una elección, se convirtieron en ceros (0). Además, los puntajes de dos de los componentes a los que se les asignaron puntajes sólo para los años en los que se celebró una elección –sufragio y elecciones libres– fueron extendidos a los años intermedios, simplemente trasladando el puntaje de un año dado a los años subsiguientes, hasta que se asignara un nuevo puntaje (ya sea porque se hubiera celebrado una elección luego de un período en el que existía un gobierno no electo, o porque se celebrara una nueva elección, o porque el proceso electoral se hubiera interrumpido). La justificación para este procedimiento es que la forma en la que un gobierno se origina continúa siendo una característica que afecta su naturaleza más allá del momento de su instalación.

En el caso de las elecciones limpias, se siguió un proceso ligeramente más complejo. En la codificación de este ítem se utilizó una escala de tres puntos para facilitar la interpretación. Sin embargo, el 1 de esta escala no representa en realidad un punto medio; es mucho más cercano al 2. Así, cada 1 fue convertido en un 3 y cada 2 en un 4. Además, debido a que este elemento distingue los valores asignados a las elecciones presidenciales y parlamentarias, los puntajes no han sido simplemente trasladados de elección a elección. Los puntajes son un promedio de los puntajes de las elecciones presidenciales y parlamentarias.

Un segundo aspecto considerado fue la asignación de un solo puntaje por país y por año. Esta práctica habitual obedece a razones de parsimonia, y está bien justificada. Después de todo, el objetivo de generar un índice es ofrecer una síntesis de la situación de un país. Pero ello implica varios problemas, debido a que la situación de un país cambia en el curso de un año, y se utiliza un solo puntaje para caracterizar todo el período

anual. En algunos casos la solución es relativamente simple, cuando un acontecimiento clave –como la celebración de una elección– tenía lugar hacia fines de un año, el cambio de estatus debido a ese acontecimiento fue registrado en el año siguiente. Por ejemplo, cuando en 1985 se celebraron en Guatemala las elecciones que terminaron con un período de gobiernos dominados por los militares, esas elecciones tuvieron lugar a fines de año, y el cambio de gobierno se realizó en enero de 1986. Por tanto, aún cuando los puntajes para los elementos componentes fueron registrados en 1985, al computar el IDE esos puntajes fueron ingresados en 1986. Cuando los acontecimientos sucedieron en la primera mitad del año, fueron computados para tal año. Por ejemplo, las elecciones de 1994 en El Salvador se celebraron en marzo, y el cambio de gobierno en junio, por lo tanto el cambio fue registrado en 1994.

En otros casos la solución fue más complicada, ya que cuando un acontecimiento ocurrió en la segunda mitad del año fue registrado en ese año. Éste es el caso, por ejemplo, de las elecciones fraudulentas celebradas en mayo en la República Dominicana, y la nueva presidencia que asumió en agosto. También fueron problemáticos los casos en los que ocurrió más de un acontecimiento crítico en el mismo año. Por ejemplo, en el año 2000 Perú celebró dos elecciones fuertemente cuestionadas, en abril y mayo, que llevaron en julio a la asunción de Fujimori a la presidencia y luego a su renuncia en noviembre. En este caso, las elecciones problemáticas fueron registradas en el año 2000 y la rectificación de la situación en el año 2001.

En tercer lugar, se normalizaron las escalas de los componentes, es decir, fueron trasladadas a una métrica común, a través de una normalización lineal simple del intervalo unidad:

$$\text{valor normalizado} = \frac{\text{valor de escala original}}{\text{máximo valor posible en escala original}}$$

Prácticamente cualquier opción que se hubiera tomado habría sido algo arbitraria, dado que no existen unidades de medida para la libertad electoral ampliamente acepta-

das y comparables a unidades como kilogramos o dólares. Sin embargo, la elección del procedimiento de normalización tal como fue aplicado a las escalas ordinales de cinco puntos –con la modificación introducida en el elemento elecciones limpias, al transformar los puntajes de las escalas componentes en una base de datos rectangular, todas las escalas utilizadas para medir los cuatro componentes son escalas ordinales de cinco puntos– es transparente y justificable.

Por un lado, todas las escalas tienen puntos finales con significación teórica, y puede suponerse que caen en el intervalo unidad, en el que 0 indica ausencia total de la propiedad y 1 indica presencia total de la propiedad. El valor inferior de la escala ordinal corresponde a la negación de la propiedad en cuestión, mientras que el valor superior corresponde a su presencia completa. Un caso con valor de sufragio 0 no presenta sufragio en absoluto, mientras que un caso con valor de sufragio 1 luego de la normalización tiene sufragio adulto completo, el estándar teóricamente establecido. Así, el problema de la distancia se refiere únicamente a los puntos comprendidos entre el punto inicial y el punto final.

Por otro lado, la mayor parte de las escalas fueron construidas de modo que cada punto de la escala fuera interpretable teóricamente y que los distintos valores de la escala estuvieran lo más alejados conceptualmente entre sí como fuera posible. Los valores de la escala fueron elegidos para reflejar distinciones identificadas como relevantes en la bibliografía, evitando variaciones menores entre los casos verificables. Por lo tanto, la probabilidad de introducción de un error importante es relativamente pequeña. Aunque podrían utilizarse otros métodos psicométricos más sofisticados, éstos son más complicados, menos accesibles, fuertemente dependientes de los datos, y frecuentemente no funcionan mucho mejor que este procedimiento simple.

La elección de reglas de agregación

El cuarto paso para la construcción del IDE –la elección de reglas de agregación para formalizar la relación entre los elementos componentes del índice– fue resuelto a tra-

vés del uso de una regla de agregación simple. La idea central utilizada para ello es la opinión bien establecida de que los cuatro elementos componentes del IDE son partes que constituyen un sistema en virtud de la forma en que están combinadas, y más aún, que estos cuatro componentes son tan fundamentales para la caracterización global de un régimen que su ausencia lo tornaría directamente no-democrático. Por ejemplo, como han argumentado largamente los teóricos sobre la democracia, el hecho de que los sistemas de tipo soviético tuvieran elecciones con sufragio completo carece de significación desde la perspectiva de la democracia, debido a que el electorado no tenía opción entre candidatos alternativos y debido a que esas elecciones no llevaron al acceso a cargos que ejerciesen efectivamente poder estatal.

Por lo tanto, los cuatro elementos componentes del IDE son postulados como condiciones individualmente necesarias, insustituibles y de igual peso. Esta concepción queda formalizada calculando el producto del valor de cada uno de los elementos componentes. En términos formales, el IDE se calcula siguiendo la siguiente ecuación:

$$\text{Índice de Democracia Electoral} = \text{Sufragio} \times \text{Elecciones Limpias} \times \text{Elecciones Libres} \times \text{Cargos Públicos Electos}$$

Esta ecuación recoge una idea clave de la teoría sobre la democracia: cuando un elemento componente se encuentra por completo ausente, el régimen debe ser considerado como no-democrático. En la práctica, esta operación asegura que un valor *cero* en cualquiera de los cuatro elementos componentes lleva a clasificar el caso como no-democracia.

Éste es un estándar “duro”, que puede ser visto como menos “perdonador” que otras reglas de agregación. La concepción de que los elementos componentes del IDE son condiciones individualmente necesarias es altamente exigente, por ello ha sido utilizada conjuntamente con un criterio conservador en la asignación de ceros a los elementos componentes. Esto es así tanto para las escalas construidas de modo tal que un cero sea utilizado sólo en casos extremos, en los cuales se encuentra totalmente ausente una propiedad

ampliamente considerada como vital para la existencia de la democracia, como en el sentido de que la evidencia necesaria para asignar un cero debe ser convincente. Por lo tanto, sólo cuando las normas democráticas han sido indiscutiblemente dejadas de lado el IDE calificará a un país como no-democrático.

Probando el IDE

Confiabilidad entre codificadores y estimación de error

Por razones de tiempo no se realizó una prueba formal de confiabilidad entre codificadores. Sin embargo, para saber si otros codificadores podrían haber tenido asignación de valores diferentes de los elementos componentes del IDE, se llevó a cabo un análisis de sensibilidad. Este análisis se basa en perturbaciones a las codificaciones asignadas, de acuerdo con un diseño experimental y el examen del índice global “replicado” resultante. Así, utiliza las matemáticas para crear codificadores “virtuales” sesgados de diversos modos; por ejemplo, sesgados en una valoración inferior de uno o más componentes.

Los resultados de esta prueba mostraron que el IDE es bastante estable—las correlaciones de rango con todas las otras “réplicas” fueron 0,99 o mayores— y los cambios en la media y la dispersión fueron bastante predecibles, mostrando un sesgo negativo o positivo esperado según el diseño experimental. Esta prueba también proporcionó algunos márgenes de error básicos del IDE sobre la base de las “réplicas”. Para valores del índice entre 0,25 y 0,75, un ancho de banda generoso es aproximadamente de $\pm 0,07$, y un ancho de banda razonablemente conservador es de $\pm 0,1$. De acuerdo con el estándar más conservador posible, los valores del IDE están dentro de $\pm 0,2$. Este ancho es razonablemente constante a lo largo del intervalo citado, pero los límites precisos dependen del valor del índice y en general son más estrechos cerca de los puntos finales. Se realizó un control matemático utilizando la inversión de la conocida y muy conservadora prueba de Kolmogorov-Smirnov para la función de distribución—basada en matemáticas completamente diferentes— y se obtuvieron resultados similares.

La solidez de las reglas de agregación

Se realizó una prueba de comparación de cuatro reglas posibles de agregación para combinar los elementos componentes del IDE: el producto de los cuatro componentes utilizados en el IDE, el valor mínimo de los cuatro componentes de la escala, la media geométrica de los cuatro componentes y la media aritmética de los cuatro componentes. Los resultados mostraron que, sin importar qué regla se utilizaba, las correlaciones de rango son siempre muy altas, lo que indica que se preserva el ordenamiento general de los casos. Pero existen diferencias entre los índices, siendo por un lado la media aritmética y la geométrica similares entre sí, y por otro lado, también similares entre sí el valor mínimo y el IDE. La diferencia más importante se encuentra entre las medias y las desviaciones estándar (DE). Las medias geométrica y aritmética son de 0,92 y 0,91, respectivamente, y DE de 0,20 y 0,21 respectivamente. En cambio, el valor mínimo y el IDE tienen medias de 0,84 y 0,82 respectivamente, y DE de 0,26 y 0,28 respectivamente. Esto sugiere que aplicando estas últimas reglas es mejor la dispersión de los casos para evitar el conglomerado de casos que hace difícil interpretar sus diferencias con claridad.

El carácter dimensional de los elementos componentes

La prueba de escalabilidad de los cuatro elementos componentes del IDE dio como resultado un alfa de Cronbach de 0,92; lo que sugiere que el IDE es una medida de un fenómeno unidimensional. Sin embargo, cuando se realizó una prueba en dos períodos (1960-1985 y 1990-2002), los alfas de Cronbach resultantes fueron de 0,95 y 0,23 respectivamente. Esto indica que, mientras que en el primer período los componentes fueron unidimensionales, esto ya no resulta válido en el período posterior a 1990. Este resultado es consistente con la teoría utilizada al seleccionar las reglas de agregación para el IDE. En efecto, es importante notar que los modelos de medición aditivos estándar descansan en la presunción de que la agregación opera en múltiples mediciones paralelas. En contraste, dado que los componentes del IDE son, por teoría, considerados insustitui-

bles, la decisión de agregarlos a un puntaje único no está invalidada por ninguna desviación potencial de la unidimensionalidad. De este modo, la prueba de escalabilidad proporciona aún más validación de la elección de las reglas de agregación propuestas en vez de la muy habitual regla de adición.

Interpretando y usando el IDE

El IDE es una escala de 0,00-1,00; en la cual 0,00 indica un régimen no democrático y cualquier número mayor a 0,00 un grado de democracia, siendo los puntajes más altos indicativos de un mayor grado de democracia. Para evitar confusiones, es importante notar que el índice no debe ser interpretado como una evaluación de las acciones del gobierno. Más bien es una medida del estado de un sistema, que se ve afectado por la acción o inacción de un gobierno, así como por otros agentes estatales y actores sociales. Además, se debe resaltar que el concepto que se mide es el de democracia electoral. Este concepto no es tan estrecho como algunos lo consideran. Por lo tanto, aunque se focaliza completamente en la celebración de elecciones inclusivas, libres y limpias, abarca más que “simples elecciones”. No sólo se ven afectadas las condiciones para celebrar tales elecciones por lo que sucede entre las elecciones. Además, lleva a considerar lo que ocurre con los mismos gobiernos entre las elecciones.

El IDE no es ciertamente una medida amplia de la democracia. Es más bien una medida de una concepción del régimen político democrático basada en los postulados más ampliamente compartidos en cuanto a los derechos políticos fundamentales. Esto es muy significativo. Por un lado, implica que cualquier defecto detectado por el IDE debe ser considerado como una restricción importante a los derechos políticos de los ciudadanos. Por otro lado, el hecho de que un país haya recibido un puntaje perfecto de 1,00 no debe interpretarse en el sentido de que no pueda mejorar tanto las dimensiones no incluidas en el índice como los estándares más exigentes de los elementos componentes del IDE.

El IDE puede ser utilizado para propósitos comparativos, tanto para comparar un país consigo mismo o con otros países. De estas dos formas, la comparación de un país consigo mismo en diferentes momentos es en general la más simple para interpretar. Después de todo, un país bien puede haber introducido mejoras notables y sin embargo quedar en inferioridad frente a otros países, si estos han hecho avances mayores. Sin embargo, es importante notar que cualquier comparación debe basarse en diferencias considerables y no menores. La razón de ello es que el IDE, como cualquier índice, tiene un cierto grado de error de medición y dentro de los límites de ese error se desaconseja realizar cualquier afirmación contundente sobre diferencias. De hecho, como fue estimado a través del análisis de sensibilidad, para los valores de IDE entre 0,25 y 0,75; un ancho de banda de error generoso es aproximadamente de $\pm 0,07$. Así, cualquier par de casos que difieran en menos de este valor –por ejemplo, un país con un IDE de 0,85 y otro con uno de 0,92– están demasiado cerca como para poder distinguirlos con validez. Por tanto, es metodológicamente injustificable ofrecer un ranking demasiado preciso de países, como es habitual en el contexto de otros índices, que simplemente transforman los puntajes del IDE en un ranking sin tomar en consideración los grados de incertidumbre asociados con los mismos. La identificación de casos de referencia que sean representaciones prototípicas de los rasgos asociados con una gama de puntajes puede ayudar a brindar mayor concreción al significado de cada número.

El IDE puede ser utilizado también como señal, en el sentido de que los puntajes específicos de cada país invitan al lector a volver a los cuadros de los elementos componentes para identificar precisamente qué aspecto o aspectos dan cuenta de ese puntaje. De este modo, el IDE puede ser usado como una herramienta analítica valiosa, en cuanto ofrece un puntaje resumido que permite a quienes lo usen identificar el carácter distintivo del régimen político de cada país, en términos de sus elementos componentes pero también en términos de la relación entre las partes constitutivas del régimen y su contribución al conjunto.