

# Consideraciones metodológicas para estudiantes de política legislativa mexicana: sesgo por selección en votaciones nominales<sup>1</sup>

Francisco Cantú, Scott Desposato y Eric Magar<sup>2</sup>

**Resumen:** El estudio de la conducta legislativa en la Cámara de Diputados durante el periodo de 1998-2006 presenta un problema potencialmente serio: no todos los votos han sido publicados en la Gaceta Parlamentaria. Analizamos la naturaleza de estos datos en aras de explorar la representatividad de la muestra de votos disponibles y las posibles repercusiones de este problema en los análisis legislativos existentes. Para esto, aprovechamos la aparición de una página de internet que registra la totalidad de votos de la Cámara de Diputados desde el año 2006 en paralelo con el sistema existente desde 1998. Mediante la exploración de los mecanismos que generan la omisión de votos y la comparación distintas estimaciones del comportamiento legislativo, concluimos que los votos no publicados merman la precisión de estimadores de uso común pero no introducen ningún tipo de sesgo. A la par, hacemos pública una base de datos para el estudio del Congreso Mexicano.

*Palabras clave:* votaciones legislativas, Congreso Mexicano, problemas de datos incompletos.

## Methodological Considerations for Students of Mexican Legislative Politics: Selection Bias in Roll-Call Votes

**Abstract:** This paper examines the nature of the data available for studying legislative behavior in Mexico. In particular, we evaluate a potentially serious problem: only a subset of roll-call votes have been released for the critical transition period of 1998–2006. We test whether this subset is a representative sample of all votes, and thus suitable for study, or whether it is biased in a way that misleads scholarship. Our research strategy takes advantage of a partial overlap between two roll call vote reporting sources by the Chamber of Deputies: the site with partial vote

---

<sup>1</sup> Eric Magar agradece a Mariana Meza Hernández y Lucía Motolinía, por su ayuda en la investigación, y a la Asociación Mexicana de Cultura, A.C. por patrocinar parte de esta investigación. Los errores y omisiones son responsabilidad de los autores.

<sup>2</sup> Francisco Cantú, es profesor asistente en el Departamento de Ciencia Política de la Universidad de Houston. Dirección: Philip Guthrie Hoffman Hall 429, Department of Political Science, The University of Houston, Houston, TX, 77204-3011. Correo electrónico: [fcantu10@uh.edu](mailto:fcantu10@uh.edu)

Scott Desposato es profesor asociado en el Departamento de Ciencia Política de la Universidad de California, San Diego. Dirección: Department of Political Science. University of California, San Diego, 9500 Gilman Drive, MC 0521, La Jolla, CA 92093-0521. Correo electrónico: [swd@ucsd.edu](mailto:swd@ucsd.edu)

Eric Magar es profesor en el Departamento de Ciencia Política del Instituto Tecnológico Autónomo de México (ITAM). Dirección: Río Hondo # 1, Col. Tizapán San Ángel, Del. Álvaro Obregón, C.P. 01000 México, D.F. Tel: 5628-4000 Ext. 3702. Correo electrónico: [emagar@itam.mx](mailto:emagar@itam.mx)

Artículo recibido el 19 de diciembre de 2012 y aceptado para su publicación el 23 de abril de 2013

Traducción: Danielle Stren y Eric Magar

disclosure, created in 1998 and still in place today; and the site with universal vote disclosure since 2006 only. An examination of the data generation and publication mechanisms, comparing different estimations of legislative behavior, reveals that omitted votes reduce the precision of estimates but do not introduce bias. Scholarship of the lower chamber can therefore proceed with data that we make public with the publication of the paper.

*Key words: Roll-call votes, Mexican Congress, incomplete data problems.*

El partido hegemónico relegó al Congreso mexicano al rol de “apéndice del

ejecutivo” durante la mayor parte del siglo veinte. El papel secundario de la rama legislativa se vio reforzado por una falta general de recursos para e interés de los ciudadanos y académicos por seguir el comportamiento de los legisladores. Por décadas, las votaciones nominales no fueron publicadas sistemáticamente en el *Diario de los Debates* – el equivalente mexicano del Hansard en los sistemas Westminster o el *Congressional Record* en los EE.UU. --- sólo la cuenta total de votos al final de la votaciones. Esto opacó el proceso legislativo en beneficio del sistema de control del Partido Revolucionario Institucional (PRI).

La transición gradual a la democracia restableció la autoridad constitucional que el Congreso tiene con respecto al ejecutivo. Con una mayoría máxima topada en tres quintos de la cámara baja y una pérdida de competitividad electoral, el PRI necesitó de la oposición para las enmiendas constitucionales del ambicioso programa de reformas de Salinas de Gortari. Fue hasta 1997, cuando el PRI perdió la mayoría de los diputados, que los académicos empezaron a formular preguntas acerca de la legislatura y su papel en el proceso de creación de políticas (p.ej., [Cantú y Desposato 2012](#); [Morgenstern 2003](#); [Morgenstern y Nacif 2002](#); [Rosas y Langston 2011](#); [Weldon 1997, 2005](#)). La reaparición de datos para el estudio del Congreso mexicano pareciera atestiguar su resurgimiento. Recursos tales como las proposiciones e iniciativas de ley presentadas por los legisladores al pleno, los dictámenes de las comisiones y las votaciones nominales son publicados en la página web de los Diputados desde noviembre de 1998. De la noche a la mañana empezó a fluir la clase de información común para el análisis del comportamiento legislativo en el mundo y que proporciona al ciudadano un instrumento para hacer responsables a sus representates. No obstante, un problema adoleció al caudal de nueva información. Las votaciones nominales publicadas representan sólo una porción, grande pero no completa, de todos los votos

celebrados en el recinto parlamentario. De hecho, en promedio, desde la aparición de la Gaceta Parlamentaria y hasta la 61ª legislatura (2009-2012) inclusive, casi el 20 por ciento de los votos no fueron publicados. La cifra alcanzó su máximo, 31 por ciento de votaciones no publicadas, en la 58ª legislatura (2000-2003), y ha menguado en cada legislatura subsecuente. Esta discrepancia y aparentemente selectiva publicación de las votaciones nominales plantea una serie de desafíos. Si el subconjunto de los votos disponibles difiere sistemáticamente del universo de votos, el grueso de la investigación existente sobre el comportamiento legislativo del país está distorsionada, sesgando nuestra comprensión de la política legislativa de México en el periodo. Más aún, si tales votos no fueran una muestra representativa de los llevados a cabo en el pleno, los votantes tendrían en sus manos una herramienta deficiente para dar seguimiento a lo que sus representantes hacen en el Congreso. Los periodos de transición revisten importancia especial para los estudiantes de la democratización en lo general, y de la política mexicana en lo particular. ¿Están quizás siendo retenidos los votos más importantes? ¿Ofrecen acaso las votaciones nominales publicadas una perspectiva capaz de despistar, por incompleta, el examen de la naturaleza de la política legislativa durante la transición? El analista necesita saber si puede o no confiar en estos datos y utilizarlos para hacer inferencias acerca de este importante periodo.

Este trabajo examina la naturaleza de los datos disponibles para los estudios legislativos durante una coyuntura crítica para la democracia mexicana del periodo comprendido entre 1998 y 2006. Nuestra estrategia de investigación saca provecho de dos características de las votaciones nominales reportadas por la Cámara de Diputados. Desde septiembre de 2006, cuando quedó instalada la 60ª legislatura, el

Servicio de Información para la Estadística Parlamentaria (INFOPAL) ha publicado la totalidad de votaciones nominales, y ha coexistido a un lado de la Gaceta Parlamentaria que publica un subconjunto solamente. Esto nos ha dado la oportunidad de contrastar el subconjunto de votos elegido por la Gaceta contra el universo a lo largo de las 60ª y 61ª legislaturas (2006-2012). Segundo, aunque no contamos con todos los votos nominales para el periodo previo (1997-2006), sí pudimos recopilar el resultado agregado de cada votación – cuántos diputados votaron a favor y en contra – para todas las mociones sometidas a consideración del pleno, publicadas en la Gaceta o no. Esta fuente de datos, que obtuvimos directamente del Diario de los Debates, nos permite comparar las características claves, como la cohesión partidista y la tasa de abstención, de los votos publicados y los no publicados. Recurrimos a métodos estadísticos varios para responder si la muestra de votos publicados en 1998-2006 es representativa de todos los votos. Nuestro análisis sugiere que la muestra es estadísticamente indistinguible de la población más amplia de votaciones nominales. De cumplirse el supuesto de que los criterios de selección de votaciones no publicadas por la Gaceta se mantuvieron sin cambio antes y después de 2006 – un supuesto fundamental de nuestro análisis y que, también demostramos, se sostiene – podemos afirmar que es posible estudiar el periodo 1998-2006 con los datos incompletos sin temor de sesgar las conclusiones. Las implicaciones de nuestro análisis descartan una omisión selectiva de votaciones nominales, explicando el problema como el resultado de una falta de recursos y no de opacidad del sistema de rendición de cuentas.

## 1 La *Gaceta* y la omisión de votos

La *Gaceta Parlamentaria* fue creada cuando la cámara de Diputados instaló un nuevo sistema de votación electrónica al principio del segundo año de la 57ª Legislatura. Se trata de un acervo, accesible por internet, de documentos e información pertinentes

para las actividades enumeradas en el orden del día de cada sesión plenaria. Empezó a publicarse el 10 de noviembre de 1998. Antes de ello, cuando el PRI controlaba sistemáticamente la mayoría en la cámara, los diputados iniciaban las sesiones sin mucho más que el orden del día. Quien deseara consultar el dictamen de una comisión, los anexos o cualquier información pertinente a una moción específica, tenía que solicitarla al personal de la cámara. Dichas solicitudes, si eran atendidas, solían llegar tarde, incitando la ira de los legisladores de la minoría. Como consecuencia, los diputados a menudo votaban con los ojos esencialmente “vendados” en el pleno, con escasa información de la sustancia de la legislación.<sup>3</sup>

Además de proporcionar acceso a documentos claves para el trabajo de los legisladores, la Gaceta también se convirtió en una fuente invaluable para estudiar al Congreso, ofreciendo información detallada del proceso legislativo, como las votaciones nominales, por primera vez en décadas (p.ej., [Casar 2002](#); [Nacif 2002](#); [Rosas y Langston 2011](#)). A pesar de ser amigable para el usuario, la Gaceta no es explícitamente un archivo y de esto se desprenden problemas para la investigación. La Gaceta intenta, antes que cualquier otra cosa, tener accesible la información relevante para los trámites del orden del día. Esto no siempre es posible. Los coordinadores de

---

3 Una sesión ordinaria de la 56ª Legislatura, en el ocaso de la hegemonía del PRI, ofrece un ejemplo elocuente. El diario de los debates sugiere que el PAN y el PRI habrían acordado enmendar el artículo 105 de la constitución sin la anuencia del PRD. A la 1:55pm., el diputado del PAN Francisco José Peniche y Bolio hizo la propuesta para dispensar de la segunda lectura y proceder de inmediato a la votación. El diputado del PRD José Jesús Ortega Martínez fue entonces reconocido para expresar su oposición a la suspensión del procedimiento ordinario. “[L]o que va a esperar el pueblo es nuevamente que los diputados estemos legislando sobre las rodillas” dijo, “y legislar es un decir, porque sólo se nos presentan las iniciativas y prácticamente se nos llama a que votemos a favor y a que votemos en contra (...) [E]l día de hoy a las 12:00 o a la 1:00 de la tarde (...) la gran mayoría de todos nosotros estamos conociendo el contenido de la ley que hoy vamos a discutir y supuestamente vamos a aprobar (...) ¿Cuál es nuestra responsabilidad, compañeras y compañeros y particularmente mis compañeros de Acción Nacional, que son los que hacen la propuesta de que ahorremos trámites?” (*Diario de Debates*, segunda sesión ordinaria del primer año de la 56ª Legislatura, 24 de Abril, 1995).

las fracciones partidistas negocian el orden del día en la Junta de Coordinación Política y, a menudo, terminan de hacerlo minutos antes de que dé inicio la sesión plenaria. Las iniciativas que son anunciadas con anticipación (o que pueden ser anticipadas por el personal de la cámara) cuentan con toda la documentación en la *Gaceta* del día. Pero aquellos votos que son programados tarde o después del comienzo de la sesión plenaria, así como votos procedimentales, suelen no aparecer en la *Gaceta* del todo. Y, por ende, la plantilla utilizada para recibir la votación nominal tampoco está lista. La votación queda excluida de la *Gaceta*.<sup>4</sup>

Para identificar la totalidad de votaciones celebradas en el pleno y compararlas con las publicadas, descargamos versiones digitalizadas del Diario de los Debates (Diario, abreviado) correspondientes a las sesiones regulares y extraordinarias de la cámara de la 57<sup>a</sup> a la 59<sup>a</sup> Legislaturas. Analizamos el texto con un procedimiento automatizado que busca el anuncio formal de cada votación nominal. Este procedimiento y su exactitud son discutidos con detalle en el apéndice. Para las 60<sup>a</sup> y 61<sup>a</sup> Legislaturas, en cambio, el listado completo de las votaciones está disponible en el INFOPAL. De ahora en adelante adoptaremos la etiqueta “votaciones *Gaceta*” para referirnos a aquéllas que sí fueron publicadas en la *Gaceta Parlamentaria*, “votaciones omitidas” para aquéllas excluidas de la *Gaceta* y “todas las votaciones” para el universo de las votaciones nominales, sin importar si fueron o no publicadas.

#### [TABLA 1]

La tabla 1 ofrece una primera comparación de 3,032 votaciones del universo y del subconjunto de 2,439 publicadas en la *Gaceta* en el periodo de 1998-2012—en total, la *Gaceta* omitió casi de 20 por ciento de las votaciones nominales. Las votaciones están agrupadas en seis categorías mutuamente excluyentes. Las mociones en el

---

4 Contactamos al staff legislativo para recuperar los votos no publicados. Después de varios intentos, nos dijo un director que esos votos habían sido desechados después de las sesiones y que simplemente no existían.

Congreso mexicano son tramitadas con una agenda progresiva (vea [Heller y Weldon 2003](#)). Se las vota primero en lo general contra el status quo. Si una mayoría se manifiesta en favor, la moción es entonces sometida a una o más enmiendas en votaciones en lo particular. La categoría de votaciones en lo general se conforma en su casi totalidad de dictámenes de comisión, aunque un puñado fueron propuestas directas al pleno cuando las reglas se suspenden por procedimientos extraordinarios.

La información de la tabla muestra efectos de selección substanciales en los tipos de proyectos que son publicados por la *Gaceta*. Primero, los votos en lo general están modestamente sobrerrepresentados en la *Gaceta*: alcanzan 68.5% de las votaciones publicadas contra el 56.8% de la totalidad. Los dictámenes de comisión suelen conocerse de antemano y, por ende,, suelen también aparecer en la *Gaceta*: las votaciones en lo general representan solo el 8.4 por ciento de las omitidas por la *Gaceta*, principalmente dictámenes colocados al último en el orden del día, pero también puntos de acuerdo mandados directamente por la Junta de Coordinación al pleno (p.ej. el presupuesto de la cámara) y dictámenes negativos para rechazar una o más propuestas simultáneamente.

Los votos en lo particular están a penas sobrerrepresentados en los publicados. Las votaciones en esta categoría inciden sobre iniciativas que ya sotrearon el voto en lo general. Las enmiendas introducidas en comisión se conocen a priori, dándole tiempo al personal de preparar su inclusión en la *Gaceta*. Pero las enmiendas también pueden ser introducidas cuando la Junta negocia la orden del día y, si esto pasa al último minuto, pueden ser omitidas por la *Gaceta* del día. Aquellas raras enmiendas introducidas directamente en el pleno, cuando las reglas son suspendidas por una mayoría calificada, también tienden a quedar fuera de la *Gaceta*. Las enmiendas representan casi la cuarta parte de las votaciones publicadas y la mitad de esta



proporción entre las no reportadas

De los votos procedimentales, menos de la mitad fueron publicados en la *Gaceta* durante este periodo, dejándolos muy sobrerrepresentados en la porción omitida por la *Gaceta* con respecto del universo. La categoría incluye votos que estructuran el proceso legislativo—enmiendas a las reglas permanentes del Congreso, mociones para suspender las reglas y adoptar procedimientos extraordinarios, mociones para aceptar enmiendas no contempladas de el dictamen de la comisión, etc.—y otros asuntos procedimentales de menos fondo. Como éstos tienden a plantearse espontáneamente en el transcurso de la sesión plenaria, son típicamente omitidos por la *Gaceta*.

Las tres categorías restantes son prácticamente excluidas de entre los votos publicados. La categoría “toma de posición” son votaciones sobre si incluir o no inmediatamente un punto de acuerdo en el orden del día. Aunque algunos decretos de la cámara de importancia (p.ej. enmiendas al reglamento de la cámara o nombramientos a las comisiones permanentes) son puntos de acuerdo, aquéllos incluidos en esta categoría son ejercicios de *position-taking* (Mayhew 1974), tales como solicitar que la Secretaría de Hacienda conceda créditos libres de intereses a las víctimas de huracanes o lamentarse por un comentario hecho por el presidente que pudo haber herido los sentimientos del régimen cubano. Mientras que el punto de acuerdo en sí mismo se decide por votación económica, la decisión de modificar el orden del día para tomar dicha votación requiere de una mayoría calificada y, por lo tanto, de una votación nominal. Casi una quinta parte de los votos sin reportar son de este tipo. Los “proyectos privados” y “viajes presidenciales” autorizan a ciudadanos para desempeñar labores para o aceptar honores de gobiernos extranjeros y al ejecutivo para salir del país, respectivamente. Aunque algunas controversias que

pueden definir las líneas de conflicto legislativo pueden aparecer aquí, la *Gaceta* suele ignorarlos por su escasa importancia.

En conjunto, desde una óptica meramente descriptiva, los tipos de votos que no reporta la *Gaceta* son marcadamente diferentes de aquellos que reporta. Y son muchos. ¿Acaso la investigación basada solamente en las votaciones publicadas (p.ej. Rosas y Langston 2011) está condenada a sacar conclusiones inevitablemente sesgadas? ¿Pueden los estudios legislativos aprender algo sobre el periodo 1998-2006 a pesar del mecanismo de selección empleado por la *Gaceta*? El resto del trabajo investiga en qué medida esto es problemático para el análisis de votaciones nominales de la Cámara de Diputados.

## 2 Análisis

La pregunta empírica central de nuestro trabajo es si las votaciones nominales publicadas por la *Gaceta* son o no una muestra representativa de la población más amplia de votaciones. De serlo, entonces académicos y analistas pueden estudiar el periodo crítico de transición del Congreso mexicano utilizando los recursos disponibles, sin temor a extraer conclusiones tendenciosas. Si en cambio no lo fuera, será necesaria una recopilación adicional de datos y/o ajustes a la muestra.<sup>5</sup>

En un marco más abstracto, cada voto puede ser considerado como un sorteo de una distribución categórica  $M$ -dimensional.  $M$  en este caso es el número de diputados y las categorías corresponden al número de acciones que puede llevar a cabo un legislador en una votación nominal—votar a favor, en contra, abstenerse, ausentarse. Las dimensiones no son independientes porque los legisladores negocian y son

---

<sup>5</sup> Por supuesto, aun si la muestra de la *Gaceta* fuera representativa, hay pérdida de información en virtud base de datos más pequeña. Esto conlleva conclusiones menos precisas, pero atinadas. Si estuviera en el futuro disponible el universo de las votaciones para ese periodo, ese censo será preferible que la muestra que consideramos aquí.

presionados por el partido y por otros actores. En este contexto, probar si los votos publicados por la *Gaceta* son una muestra aleatoria de todos los votos requeriría de una base de datos inmensa—mucho más grande que la proporcionada por los años de votos nominales que observamos.

Como una alternativa, nos enfocamos en cantidades de interés para los estudios legislativos clásicos: índices Rice de cohesión de los legisladores y los puntos ideales que las técnicas de escalamiento permiten recuperar (Cox 1987; Poole y Rosenthal 1997; Rosas y Langston 2011). Si bien existen otras cantidades de potencial interés, es extremadamente raro que quienes analizan votaciones nominales no le presten atención a estas dos cantidades clave. Nuestras pruebas, entonces, plantearán la misma pregunta a estos estadísticos: ¿difieren significativamente las cantidades derivadas a partir de las votaciones *Gaceta* de las cantidades derivadas de todos los votos? ¿Son los votos *Gaceta* representativos de todos los votos en términos de cohesión y perfiles de votación?

A continuación abordamos estas preguntas por medio del análisis de permutación no paramétrico. El procedimiento consiste en tomar muchas muestras aleatorias repetidas con remplazo y de tamaño  $n$  de las votaciones nominales del universo en cada legislatura. Para cada muestras aleatoria, se calculan cantidades de interés. Comparamos la distribución de las cantidades de interés de las permutaciones con el valor de las mismas medidas de las votaciones *Gaceta*. Si los valores de votaciones *Gaceta* son típicos de muestras aleatorias de la población más grande, entonces tendremos evidencia de que los votos publicados pueden utilizarse para el análisis. De lo contrario, habrá evidencia de que los mecanismos de selección deberán ser entendidos para contabilizarlos en los estudios legislativos de la cámara de diputados.

Procedimos en dos frentes, examinando por separado los datos de 1998-2006 y los

de 2006-2012. Para el periodo crítico 1998-2006 carecemos de casi una cuarta parte de las votaciones que no fueron publicadas por la Gaceta ni en ningún otro lugar. Sin embargo, aunque para los votos no publicados carecemos de información de cómo votó cada diputado, el Diario de los Debates sí enumera estadísticos resumen del desenlace de cada votación: el número de diputados que votaron a favor, en contra o que se abstuvieron. Consecuentemente, aunque calcular la cohesión de las fracciones parlamentarias o los puntos ideales de los diputados está fuera de nuestro alcance, sí resulta posible, por lo menos, preguntar si en sus descriptivos agregados las votaciones Gaceta tienen patrones similares de divisiones favor-contra y de tasas de participación y de abstención en comparación con todos los votos que detectamos en el Diario. Nuestro hallazgo es que, en su mayor parte, las votaciones Gaceta parecen ser una muestra representativa de todas las votaciones. La excepción más importante es que la muestra Gaceta contiene más votos unánimes que el Diario, especialmente en el año 2004. Esto es a primera vista lamentable ya que los votos unánimes esencialmente no contienen ninguna información para el escalamiento de puntos ideales. Sin embargo la segunda parte de nuestro análisis proporciona la seguridad que esto no es un problema. Además, el volumen de votos omitidos en ese último periodo, por debajo del 20 por ciento, fue relativamente menor.

Para el periodo 2006-2012 contamos con la información de todas y cada una de las votaciones nominales. No hubo más datos faltantes porque, desde el inicio de la 60ª Legislatura, todas las votaciones nominales son listadas en una nueva página web de la propia Cámara interesada en recopilar cada una de ellas (no como la Gaceta, cuya razón de ser es tener lista información necesaria para tramitar los asuntos de cada día). Al mismo tiempo, la Gaceta Parlamentaria continuó la práctica de publicar un subconjunto incompleto de los votos celebrados. Bajo el supuesto de que el proceso

de omisión de votaciones de la Gaceta no ha cambiado, explotamos la coexistencia de una fuente completa y otra parcial para investigar si ésta afecta la estimación de los puntos ideales de los diputados. De nuevo, nuestro hallazgo más importante es que no afecta. Las votaciones Gaceta parecen ser típicas de todas desde esta óptica también.

## 2.1 La comparación de datos agregados 1998–2006

Empezamos con un estadístico de resumen, examinando la cohesión anual media del pleno en las votaciones publicadas y las no publicadas. El tamaño de la muestra aleatoria  $n$  en este paso del análisis es el número de votaciones nominales publicadas en el año en la Gaceta. En otras palabras, el muestreo selecciona un gran número de muestras (10 mil) con tantos votos como las sí publicadas en la Gaceta. Definimos la

cohesión anual media como  $c = \frac{1}{m} \sum_{i=1}^m \frac{|y_i - n_i|}{(y_i + n_i)}$ , donde  $m$  es el número de votaciones nominales celebradas en el año (publicadas o no),  $y_i$  es el número de votos a favor de la moción y  $n_i$  los votos en contra. En otras palabras, obtenemos primero el índice de cohesión de [Rice \(1925\)](#) para cada moción y los promediamos para todas las votaciones del año. El índice de Rice está acotado por el intervalo  $[0,1]$  y proporciona una medida de conflicto o controversia entre los legisladores. Una moción adoptada o rechazada por unanimidad (todos los diputados que votan lo hacen en el mismo sentido) obtiene un valor de 1 en el índice; aquélla que divide la legislatura por mitades exactas, una a favor y la otra en contra, obtiene un valor de 0; y los casos intermedios varían entre estos extremos. La Tabla 2 muestra las estadísticas resumen de cohesión de todas las votaciones en las legislaturas 57<sup>a</sup>-59<sup>a</sup> (1998-2006) en contraste con las publicadas y las no publicadas en la Gaceta.

[TABLA 2]

La naturaleza de la medida de cohesión que aquí estamos calculando requiere elaborar. Comúnmente el término cohesión se utiliza para referirse a la conducta de un subgrupo de diputados – esto es, el índice de Rice se calcula por separado para diferentes grupos pertenecientes a una legislatura con el objeto de comparar la cohesión de los distintos partidos o de los representantes de intereses agrícolas o manufactureros, como en el trabajo original de Rice. En este caso, estamos calculando una puntuación de cohesión total para toda la legislatura, en otras palabras, que tan frecuentemente *todos* los legisladores votaron en el mismo sentido, sin importar el partido o pertenencia a otros grupos. Otra forma de pensar esto es que  $1 - c$  es la medida de qué tan cerca estuvo el voto total de la división en mitades perfectas ( $1 - c = 1$ ) o de la una votación para nada cerrada, con todos los legisladores votando igual ( $1 - c = 0$ ).<sup>6</sup>

De vuelta a los datos, el patrón general que se distingue en la Tabla 2 es que los votos omitidos tienen una medida de cohesión menor, son más controversiales en promedio que los votos Gaceta. La cohesión media de los votos publicados fue 0.82, contra 0.69 de los inéditos. Casi la mitad de esta diferencia es atribuible al diferencial de votaciones unánimes. Treinta por ciento de las votaciones inéditos fueron unánimes, comparados con 50 por ciento de las votaciones Gaceta. Muchos métodos, como el de puntos ideales que elaboramos más adelante, descartan los votos unánimes, con lo cual la diferencia de cohesión entre las votaciones publicadas y las no disminuye sustancialmente a 0.08 (0.64 para aquéllas, 0.56 para éstas, no mostrados).

Los estudios de de la política legislativa también suelen analizar las abstenciones en y las ausencias del pleno, a menudo codificándolas como si se tratase

---

<sup>6</sup> Agradecemos a un dictaminador anónimo la sugerencia de esta aclaración.

de votos en contra (p.ej. Jones y Hwang 2005). Consideramos dicha codificación con escepticismo (vea Rosas y Shomer 2008), pero tratándose de una práctica común, sopesamos las diferencias potenciales entre las votaciones publicadas y las no publicadas con respecto a esto. Y ambas muestras presentan tasas de participación virtualmente idénticas. El número de abstenciones promedio en las votaciones Gaceta es de 6.73, contra 6.79 en los votos inéditos. Las votaciones Gaceta registraron 104 diputados ausentes del pleno en promedio, las inéditas 118, modestamente más alto. En otras palabras, las comparaciones simples de cohesión y de tasas de abstención y ausencia revelan diferencias pequeñas entre los votos. Las abstenciones, que son probablemente estratégicas y más costosas para el legislador, tienen tasas casi idénticas. Las ausencias, que pueden ser reflejo de otras prioridades y obligaciones de los legisladores, la diferencia parece consistente con el anuncio de la votación. Las votaciones Gaceta son de mayor perfil, pre-agendadas, y por ende menos probables de llegar al pleno cuando muchos legisladores están ausentes. publicados o no.

### [TABLA 3]

La Tabla 3 desglosa las tasas de cohesión anuales de las votaciones según las haya publicado la Gaceta o no entre 1998 y 2006. Hay diferencias modestas entre las publicadas y las inéditas, pero los patrones no son consistentes ni en sentido ni en magnitud. La mayor parte de las diferencias son pequeñas, aunque la brecha de 2004 es mayor, en 0.24. Esto significa que la típica moción publicada en ese año tuvo una coalición mayoritaria alrededor de 10 por ciento más grande que la típica moción inédita. Una vez más, la mayor parte de esta diferencia es debida a un número excepcionalmente elevado de votos unánimes ese año: 56 por ciento de las votaciones Gaceta fueron unánimes en 2004, contra sólo 21 por ciento de los votos

inéditos, una diferencia mucho mayor que la observada en cualquier otro año.

Aunque modestas, ¿son significativas estas diferencias? La Figura 1 muestra los resultados del análisis de permutaciones que revelan que las votaciones Gaceta parecen ser típicas de todas las votaciones nominales. En cada histograma de la figura, la línea vertical continua indica la cohesión media anual de todos los votos, mientras la línea punteada indica la misma cantidad para los votos Gaceta únicamente. Los histogramas muestran la distribución de cohesión anual media obtenida permutando muestras aleatorias de todos los votos. Se tomaron diez mil muestras del mismo tamaño que el conjunto de votaciones que sí fueron publicadas en la Gaceta. De este modo, se consigue emular el proceso de generación de votos no publicados para verificar si los que omite la Gaceta parecen ser una muestra aleatoria. Cuando la línea punteada está hacia el medio de la distribución, significa que la cohesión de los votos Gaceta es típica de una muestra aleatoria de todos los votos. Cuando la línea punteada está en la cola del histograma, significa que los votos Gaceta son significativamente diferentes del conjunto de todos los votos.

#### [FIGURA 1]

En todos los años excepto 2004, la línea punteada se ubica más cerca del centro de la distribución simulada que de un extremo. En otras palabras, aunque presenta diferencias puntuales visibles en 2002, 2003, 2005 y 2006, el promedio de cohesión Gaceta en esos años se ve como la del típico voto elegido aleatoriamente de la población de todos los votos. No hay evidencia estadística que permita rechazar la hipótesis de que los votos Gaceta de hecho son una muestra aleatoria de la población completa. En 2004, en cambio, la cohesión media anual de la muestra Gaceta se ubicó en el percentil 99º de los valores permutados. Sin ser el valor máximo, es más grande que el promedio de cohesión anual de casi todas las muestras permutadas. Esta particularidad pareciera resultar del elevado número de votos unánimes de ese año. Cuando restringimos el muestreo de modo que excluya votaciones nominales unánimes, también obtuvimos un resultado nulo para el año



2004, como en los demás.

El análisis de permutaciones sostiene la afirmación de que los votos publicados por la Gaceta Parlamentaria no son fundamentalmente diferentes de una muestra representativa de todos los votos celebrados, pero requiere de varias advertencias y observaciones adicionales. Primero, señalamos de nueva cuenta que la escasez de datos para este primer periodo 1998–2006 limita nuestro análisis de cohesión general de votos. Los datos disponibles sólo permiten comparar la medida de controversia agregada de los votos. No podemos descartar que los patrones de votación al interior de cada partido u otros subgrupos difieran entre lo publicado y lo no publicado. Segundo, el hecho que haya relativamente más votos unánimes en la Gaceta significa pérdida de votos controversiales cuando se trabaja únicamente con el subconjunto Gaceta. Tercero, aunque las diferencias en abstenciones y ausencias parecen soslayables entre lo publicado y lo no, es posible que se trate de distintos tipos de diputados en una y otra muestras. Esto podría tener efectos según qué modelo de votación faltante se utilice (Rosas y Shomer 2008). Finalmente, repetimos el análisis entero codificando todas abstención y ausencia como si fuera un voto en contra de la moción. Los resultados permanecieron esencialmente iguales, sin evidencia de diferencias sistemáticas entre la Gaceta Parlamentaria y la totalidad de votaciones nominales. Con esto en mente, pasamos ahora a buscar diferencias mediante un análisis espacial de las 60<sup>a</sup> y 61<sup>a</sup> Legislaturas.

## 2.2 Los puntos ideales de los diputados 2006–12

La segunda parte de nuestro análisis datos del periodo en el que ya está disponible la totalidad de votaciones nominales a través de la nueva página web del Infopal. Como en legislaturas anteriores, la *Gaceta* continuó publicando un subconjunto incompleto de votos. Mientras que para el primer periodo contamos sólo con el agregado de votos

a favor y en contra, y sólo pudimos comparar la cohesión general de la legislatura, para este segundo periodo tenemos el voto individual de cada legislador. Por lo tanto podemos explorar si el mecanismo de selección de la *Gaceta* tiene algún impacto en la estimación de puntos ideales (vea Poole y Rosenthal 1997 para una introducción a esta metodología). Con este fin, estimamos los puntos ideales de los diputados dos veces: una utilizando la totalidad de votaciones nominales celebradas, otra sólo con los que publicó la *Gaceta*. En tanto el proceso de selección de la *Gaceta* arroje una muestra no representativa del total de votaciones, las diferencias se manifestarán en una comparación directa de los puntos ideales estimados con cada grupo de votos.

Se especificó un modelo IRT (*item response theory model*, vea Gelman y Hill 2007) bidimensional, estimando los puntos ideales de los diputados y otros parámetros de interés con Monte Carlo y Cadenas de Markov (MCMC) ([Clinton, Jackman y Rivers 2004](#); [Poole 2005](#):116). Como es estándar, las votaciones nominales unánimes y las mínimamente impugnadas (aquellas con una minoría de 2.5 por ciento o menos, excluyendo las abstenciones y ausencias) fueron eliminadas del conjunto de datos debido a la poca información que ofrecían o porque no aportaban nada relevante; la tabla 4 ofrece un resumen. Por la misma razón, los diputados que ya sea nunca juraron o que votaron en favor o en contra de menos del 10 por ciento de todos las mociones consideradas en la legislatura fueron eliminados del análisis. La primera dimensión explica casi el 95 por ciento de los votos individuales y describe una segmentación gobierno contra oposición en el Congreso. En la segunda dimensión predomina aún la presión continua para la liberalización del sistema político (Robles Peiro 2009).

La identificación del modelo se hizo asignando probabilidades anteriores semi-informativas a los puntos ideales de los miembros de los tres grandes partidos en coordenadas polares: el PRI al norte, el PAN al sureste y el PRD al suroeste.<sup>7</sup> Tres

<sup>7</sup>Definimos las distribuciones anteriores de los puntos ideales del siguiente modo ( $x,y$

cadenas se reiteraron diez mil veces cada una. Los parámetros habían alcanzado una etapa estable después de la primera mitad de todas esas repeticiones; se tomó cada 25ª simulación de la segunda mitad como muestra posterior de los parámetros, que reportamos.

[TABLA 4]

[FIGURA 2]

Los puntos ideales estimados, que aparecen en la Figura 2, son consistentes con estimaciones previas de la Cámara de Diputados de México (Rosas y Langston 2011). La mayoría de los diputados de los principales partidos manifiestan una ligera dispersión en torno a los vértices de un triángulo isósceles con reminiscencia de las probabilidades anteriores asignadas. La excepción es el PRD, alrededor de la mitad del cual se desplazó, junto con el PT, más al oeste en la Legislatura 60ª. Se trata de la facción de López Obrador, que tras no reconocer el triunfo de Felipe Calderón en la elección presidencial de 2006 comenzó una ruidosa campaña anti-PAN mientras el resto del partido negociaba con los demás partidos. La facción de López Obrador perdió control de la dirección del partido poco después, dejándolo con un disminuido, pero más radical contingente en la 61ª Legislatura. En esa legislatura, instalada después de las elecciones intermedias de 2009, el PRI ganó la pluralidad de la cámara, y se alió con dos partidos menores (el PVEM, su aliado pre-electoral común desde 2001 y el PANAL) en pos de una mayoría contra el presidente. Alcanzaron un grado de cohesión notable.

[FIGURA 3]

La Figura 3 compara, una dimensión a la vez, los puntos ideales estimados con la totalidad de votaciones nominales y aquéllos estimados con las votaciones de la Gaceta únicamente. Con una excepción, los legisladores de todos los partidos

---

son coordenadas en un plano cartesiano). Los diputados del PAN, PRI y PRD recibieron distribuciones anteriores normales bivariadas semi-informativas:

$\begin{pmatrix} x \\ y \end{pmatrix}_{\text{PAN}} \sim \mathcal{N}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{4}\right)$ ;  $\begin{pmatrix} x \\ y \end{pmatrix}_{\text{PRI}} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \frac{1}{4}\right)$ ; and  $\begin{pmatrix} x \\ y \end{pmatrix}_{\text{PRD}} \sim \mathcal{N}\left(\begin{pmatrix} -1 \\ -1 \end{pmatrix}, \frac{1}{4}\right)$ . Los demás diputados recibieron anteriores no-informativas.  $\begin{pmatrix} x \\ y \end{pmatrix}_{\text{rest}} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, 10\right)$ .

aparecen estrechamente alineados con la diagonal de 45° que indica la coincidencia exacta de coordenadas. La excepción es la segunda dimensión de la 60ª Legislatura, donde un grupo grande del PRD aparece en una posición más radical con los votos Gaceta únicamente que con la totalidad de votos. Esto sorprende porque la Gaceta excluye un buen número de mociones de posicionamiento y de enmienda, y su exclusión debería ejercer una fuerza centrípeta, no centrífuga.<sup>8</sup> Aún en este caso, la coincidencia general entre los puntos ideales estimados permite sostener que las diferencias entre los votos Gaceta y la totalidad son pequeñas.

### 2.3 Permutaciones de puntos ideales

El siguiente paso para evaluar los efectos de las votaciones nominales incompletas reproduce el análisis de permutaciones utilizado anteriormente pero con los puntos ideales en vez de los estadísticos de cohesión agregados. Como la estimación MCMC demanda muchos recursos de la computadora, usar un elevadísimo número de muestras de gran tamaño resultó impráctico. De entre todas las votaciones nominales celebradas en cada legislatura, tomamos 99 muestras aleatorias (en vez de 10 mil) de  $n = 50$  votos escalables. Hubiera sido preferible tomar muchas más muestras y que cada una tuviese el mismo tamaño que los votos escalables de la *Gaceta* (o sea  $n = 160$  para la 60ª y  $n = 234$  para la 61ª Legislaturas, respectivamente), pero el tiempo de cómputo continúa siendo un problema grave.<sup>9</sup> Para realizar la comparación de los puntos ideales estimados con las muestras aleatorias contra los obtenidos con la *Gaceta*, el punto ideal de cada legislador escalado con votos Gaceta únicamente fue restado, coordenada por coordenada, de cada uno de los puntos obtenidos con las 99

<sup>8</sup> Como se observó en el análisis de la cohesión del pleno, el mecanismo de selección de la Gaceta resulta en una proporción mayor de votos no informativos – votos unánimes o cuasi unánimes. Es posible que las fuerzas moderadoras del PRD ejercieran mayor influencia en votaciones no publicadas por la Gaceta, y que sea justamente su exclusión la que mueva este resultado.

<sup>9</sup> Aún así, se requirieron cerca de 24 horas para completar la rutina de estimación para cada legislatura usando tres procesadores multinúcleo en paralelo. Con tamaño de muestras entre tres y cinco veces más grandes, el tiempo de computación aumentaría una cuatro veces en promedio.

muestras aleatorias. Si la *Gaceta* es una muestra aleatoria más, los puntos ideales escalados con ella oscilarán estocásticamente en torno de los de las 99 muestras, sin cambios sistemáticos entre legisladores. Las diferencias estarán centradas en el cero, con raras desviaciones positivas o negativas anormales. Las Figuras 4 y 5 clasifican a los diputados de acuerdo a puntuaciones *Gaceta* ascendentes, y el eje vertical reporta la diferencia de coordenadas de las muestras aleatorias. Las líneas verticales gruesas y delgadas señalan los intervalos de 50 y 95 por ciento de la diferencia.

[FIGURA 4]

[FIGURA 5]

Para el caso de la 60ª Legislatura, la Figura 4 muestra que solo un puñado de los intervalos de 95 por ciento de los diputados excluye el cero, el punto donde la diferencia entre las muestras y la *Gaceta* no se manifiesta. Incluso son muchos los intervalos de 50 por ciento que incluyen el cero. Las discrepancias se concentran en la primera dimensión y en el extremo izquierdo de la distribución – los diputados afines a López Obrador son más vulnerables al muestreo de la *Gaceta*. El muestreo aleatorio devuelve puntos ideales estimados más volátiles para unos partidos que otros, notablemente el Partido Verde (PVEM) en ambas dimensiones, y el PRD en la segunda dimensión. Para el caso de la 61ª Legislatura (Figura 5), la variación de los puntajes ideales es significativamente más pequeña que la legislatura previa. Similar a la 60ª, el grupo de los diputados lopezobradoristas tiene puntuaciones menos radicales en la primera dimensión cuando las estimaciones utilizan una muestra aleatoria de todos los votos, en vez de la muestra *Gaceta*. Hay un efecto similar entre los miembros más moderados del PRI en la segunda dimensión. Pueden extraerse algunas generalizaciones de este análisis. Primero, los diputados del PAN y del PRI son generalmente menos sensibles al muestreo que el PRD y diputados de partidos pequeños. Segundo, los diputados extremistas del flanco izquierdo resienten el muestreo de la *Gaceta*, pero esto no parece suficiente para cambiar su posicionamiento ordinal con respecto a los miembros más moderados de su propio partido. Se debe tener cautela al estimar puntos ideales para este grupo de diputados, pero no encontramos diferencias importantes en sus perfiles de votación usando los

votos Gaceta y el conjunto completo.

### 3 Conclusiones

La transición democrática de México ha venido de la mano de un aumento en el estudio académico de las instituciones políticas, incluyendo el Congreso. Desde 1997, ningún partido ha obtenido la mayoría de escaños en la Cámara de Diputados y el dominio del poder legislativo por los Presidentes mexicanos ha terminado. El resultado es un Congreso fortalecido, cuya capacidad y institucional y profesionalismo están en pleno desarrollo. Pero si ha crecido dramáticamente el interés en la política legislativa y el Congreso, la naturaleza de los datos disponibles para su estudio merece un examen serio. En particular, como la publicación de las votaciones nominales entre 1998 y 2006 fue incompleta, surge la posibilidad de que las investigaciones de la conducta legislativa en el periodo hayan extraído conclusiones erróneas, deformes por el sesgo de selección. Es más, se plantea la interrogante de si la falta de un número cuantioso de información en los registros públicos de la Cámara fue por omisión o por comisión. Este estudio indirectamente ambas interrogantes. No encontramos evidencia ni de que las conclusiones hasta ahora alcanzadas estén irremediabilmente sesgadas ni de que la falta de información sea una política de opacidad intencional por el personal de la cámara.

Mediante un examen de la cohesión legislativa y los puntos ideales de los legisladores, este estudio exploró la posibilidad de considerar las votaciones nominales disponibles en línea a través de la Gaceta Parlamentaria como si se tratase de una muestra aleatoria de todas las votaciones celebradas en el periodo. La evidencia empírica muestra que, cuando se utilizan solamente los votos publicados, estadísticos comunes para el estudio del comportamiento legislativo no son significativamente diferentes de cuando se utiliza tanto la información publicada como la inédita. Una

implicación clave de este estudio es que pueden usarse los datos disponibles en línea para las 57ª, 58ª y 59ª Legislaturas sin preocuparse de la validez del análisis. En paralelo a la publicación de este artículo, hemos creado un acervo público que contiene las votaciones divulgadas por la Gaceta Parlamentaria entre 1998 y 2006 y por el Infopal de 2006 en adelante.<sup>10</sup>

Es importante reiterar una calificación hecha en el cuerpo del texto a nuestra conclusión. Hemos examinado y descartado el sesgo en las dos herramientas más comunes de los estudios legislativos en la actualidad – la cohesión en el pleno y los puntos ideales de los legisladores. Pero existen muchos otros estadísticos que no hemos analizado. Alentamos a la comunidad académica interesada en otras cantidades a verificar sus propiedades con los datos que hacemos públicos.

El problema de datos incompletos es común en nuevas democracias. El reto de los académicos, consultores, asesores y hasta servidores públicos consiste en evaluar si las inferencias hechas a partir de datos incompletos pueden estar sesgadas o si sólo contienen un mayor margen de error. Las herramientas que hemos propuesto en este estudio contribuyen a achicar las dudas al respecto de esta cuestión en los estudios legislativos mexicanos.

## Apéndice

Describimos aquí los pasos tomados para implementar una rutina automatizada capaz de rastrear, en el Diario de los Debates, las votaciones nominales celebradas en el pleno. Empezamos determinando, a partir de la página web de la cámara (<http://cronica.diputados.gob.mx/>), las fechas cuando la cámara estuvo en sesiones

---

<sup>10</sup>Los archivos digitalizados y con información sistematizada pueden encontrarse en las páginas de internet de los autores (véase, p.ej. <http://allman.rhon.itam.mx/~emagar/rollCall/>).

regulares o extraordinarias en las 57<sup>a</sup>, 58<sup>a</sup> y 59<sup>a</sup> Legislaturas. Una conexión al archivo que contiene el *Diario de los debates* de cada día (o la *versión estenográfica* para las raras ocasiones cuando la conexión al Diario falla) fue abierta para leer las líneas de código HTML y componer un objeto de R que llamamos `dd`. Procedimos después al análisis sintáctico de las líneas de texto en el objeto `dd`, con la ayuda de expresiones regulares ([Friedl 2006](#)), en busca del aviso de que una votación nominal se va a celebrar. Los votos individuales de los diputados no están disponibles en el Diario, pero sí la moción que se somete a consideración del pleno y el resultado de la votación con los votos a favor y en contra que obtuvo, agregados. Esto nos proporcionó un listado de las votaciones celebradas para después compararlas con las reportados en la *Gaceta*.

Leímos (esto es, humanamente) una muestra de Diarios de los Debates para detectar los patrones utilizados comúnmente cuando se anuncia una votación nominal. Después de que los diputados que fueron reconocidos para hablar en pro y en contra de la moción lo han hecho, el presidente lanza la pregunta de si se considera el tema suficientemente debatido, y el Secretario de la Cámara convoca a los diputados a votar con un anuncio protocolario que más o menos se lee así: “Ábrase el sistema electrónico de votación y háganse los avisos a que se refiere el artículo 144 del Reglamento de la Cámara de Diputados”. Antes de la enmienda al reglamento de la cámara de 2010, el artículo en cuestión era el 161. A menudo se presentan ligeras variaciones al protocolo, como cuando el Secretario se brinca alguna palabra o agregar algún adjetivo, así que la búsqueda del patrón fue calibrada, a través de prueba y error, en los Diarios que leímos humanamente. Tres patrones de búsqueda señalaron todas las votaciones, con una excepción que tratamos a continuación. Los tres patrones utilizados fueron los siguientes (`grep` es una función que ejecuta



expresiones regulares en R):

```
pat1 <- grep(dd, pattern="\b[ÁÁáá]bra\\w*\b la votaci[oó]n")
pat2 <- grep(dd, pattern="\b[ÁÁáá]bra\\w*\b el sistema [ed][le][e ]
[cv]
[to][rt][óoa][nc]i[coó][on]")
pat3 <- grep(dd, pattern="[Aa]rt[íi]culo 1[64]
[14].*de.*[Rr]eglamento")
```

Esta rutina detectó muchos eventos que no eran votaciones nominales, tales como mociones de quórum – los más comunes – o cualquier mención al artículo 141, 144, 161 o 164 de cualquier ley o reglamento. Estas fueron eliminadas de la lista, reteniendo solo las votaciones nominales.

Verificamos la exactitud del procedimiento al bajar también los Diarios de Debates entre el 1 de febrero y el 2 de diciembre de 2010, inclusive, para detectar las votaciones celebradas. Estas fueron sesiones que correspondían a la 61<sup>a</sup> Legislatura, para la cual el universo de 167 votaciones nominales es proporcionado por el Infopal. La rutina automatizada detectó todos los votos salvo uno – una precisión de más del 99%. El voto que faltó pertenece al Diario de Noviembre 15-16, cuando el presupuesto anual federal fue tramitado y aprobado en una sesión de 21 horas de duración que involucró discusiones eternas, recesos múltiples y muchas votaciones económicas de procedimiento para intentar de introducir enmiendas al dictamen de la comisión. En una instancia, la Secretaría convocó a una votación nominal sin hacer el anuncio protocolario correspondiente.

## Referencias

Cantú, Francisco and Scott Desposato. 2012. "The New Federalism of Mexico's Party System." *Journal of Politics in Latin America* (2):3–38.

- Casar, Mario Amparo. 2002. Executive-Legislative Relations: The Case of Mexico (1946-1997). In *Legislative Politics in Latin America*, ed. Scott Morgenstern and Benito Nacif. Cambridge University Press.
- Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *The American Political Science Review* 98(2):355–70.
- Cosío Villegas, Daniel. 1981. *El sistema político mexicano*. Mexico City: Joaquín Mortiz.
- Cox, Gary W. 1987. *The efficient secret*. Nueva York: Cambridge University Press.
- Domínguez, Jorge I. and James A. McCann. 1996. *Democratizing Mexico*. Baltimore: The John Hopkins University Press.
- Friedl, Jeffrey E.F. 2006. *Mastering regular expressions*. Third ed. Sebastopol, CA: O'Reilly.
- Gelman, Andrew y Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Nueva York: Cambridge University Press.
- González Casanova, Pablo. 1965. *La democracia en México*. Mexico City: Era.
- Heller, William B. and Jeffrey A. Weldon. 2003. Reglas de votación y la estabilidad en la Cámara de Diputados. In *El Congreso Mexicano después de la alternancia*, ed. Luisa Béjar Algazi and Rosa María Mirón Lince. Mexico City: Asociación Mexicana de Estudios Parlamentarios and Instituto de Investigaciones Legislativas del Senado de la República pp. 85–119.
- Jones, Mark P. y Wonjae Hwang. 2005. Party government in presidential democracies: extending the cartel and CPG theory beyond the U.S. Congress. *American Journal of Political Science* 49(2):267-82.
- Mayhew, David R. 1974. *Congress: The electoral connection*. New Haven: Yale University Press.
- Morgenstern, Scott. 2003. *Patterns of Legislative Politics: Roll-Call Voting in Latin America and the United States*. New York: Cambridge University Press.
- Morgenstern, Scott and Benito Nacif. 2002. *Legislative Politics in Latin America*. New York: Cambridge University Press.
- Nacif, Benito. 2002. Understanding Party Discipline in the Mexican Chamber of Deputies. In *Legislative Politics in Latin America*, ed. Scott Morgenstern and Benito Nacif. Cambridge University Press.
- Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting*. New York: Cambridge University Press.

- Poole, Keith T. y Howard Rosenthal. 1997. *Congress: A political-economic history of roll-call voting*. Oxford University Press.
- Rice, Stewart A. 1925. "The Behavior of Legislative Groups: A Method of Measurement." *Political Science Quarterly* 40(1):60–72.
- Robles Peiro, Gustavo Adolfo. 2009. *Dimensiones Espaciales de Votacion Legislativa en la Cámara de Diputados de la 60 Legislatura del Congreso de la Union, 2006- 2009*. Tesis de licenciatura, Departamento de Ciencia Política, ITAM.
- Rosas, Guillermo y Joy Langston. 2011. "Gubernatorial Effects on the Voting Behavior of National Legislators." *Journal of Politics* 73(2):477–493.
- Rosas, Guillermo y Yael Shomer. 2008. "Models of non-response in legislative politics." *Legislative Studies Quarterly* 33(4):573-601.
- Weldon, Jeffrey A. 1997. El presidente como legislador, 1917–1934. In *El Poder Legislativo en las décadas revolucionarias, 1908–1934*, ed. Pablo Atilio Piccato Rodríguez. *Enciclopedia Parlamentaria de México (Serie I, vol. I, tomo 3)* Mexico City: Instituto de Investigaciones Legislativas, Cámara de Diputados, LVI Legis- latura pp. 117–45.
- Weldon, Jeffrey A. 2005. "The Fall 2004 Term of the Mexican Congress (Mexican Congressional Report Series)." Center for Strategic and International Studies, Wash- ington DC.

<CohesionPerms1.eps>

Figura 1: Las diferencias de cohesión en votaciones *Gaceta* y todas las votaciones no son significativas antes de 2004. Los histogramas muestran el rango que adopta la cohesión anual media en diez mil muestras aleatorias que emulan el subconjunto publicado en la *Gaceta* a partir del total de votaciones. Las líneas punteadas apuntan a la cohesión anual promedio en la *Gaceta*, las líneas continuas a la cohesión anual promedio en todas las votaciones.

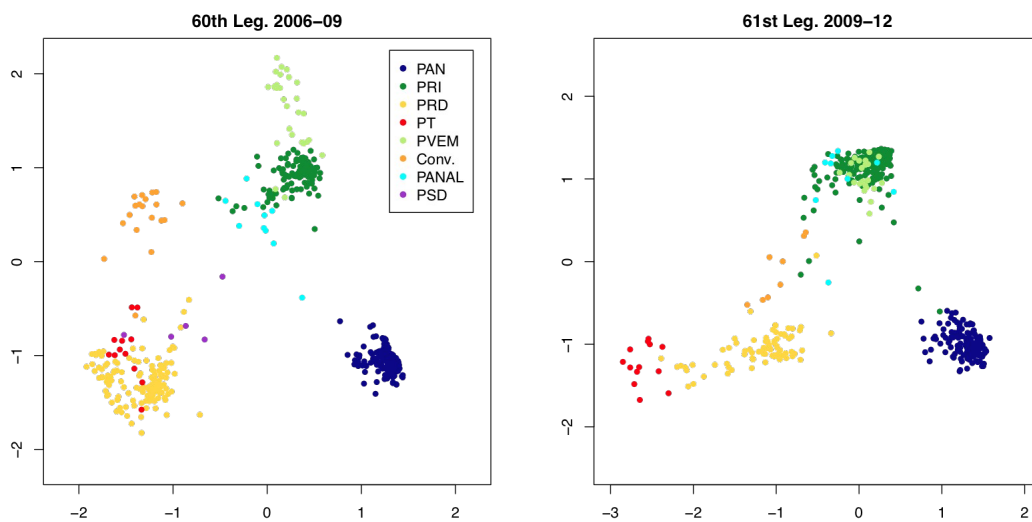
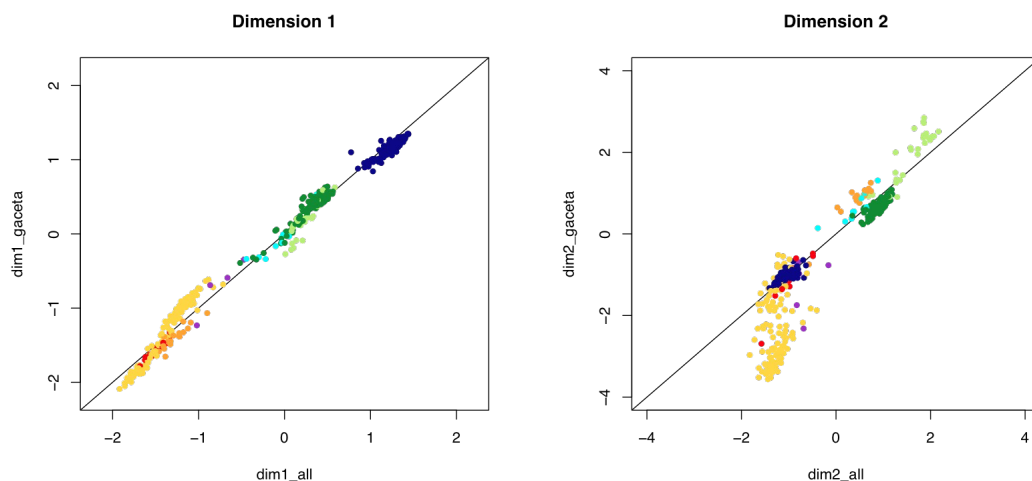


Figure 2: Ideal point estimates using all votes in two legislatures

Figura 2: Puntos ideales estimados con todos los votos en dos legislaturas. Nota: versiones a color de todos los diagramas del artículo pueden encontrarse en <http://allman.rhon.itam.mx/~emagar/graphs/pygPaper/>.

Part A. 60th Legislature, 2006–09



Part B. 61st Legislature, 2009–12

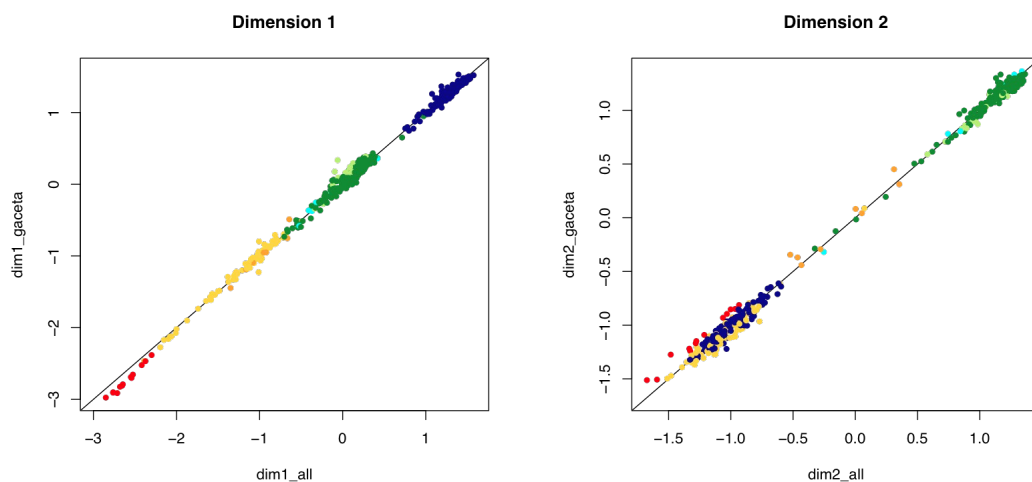
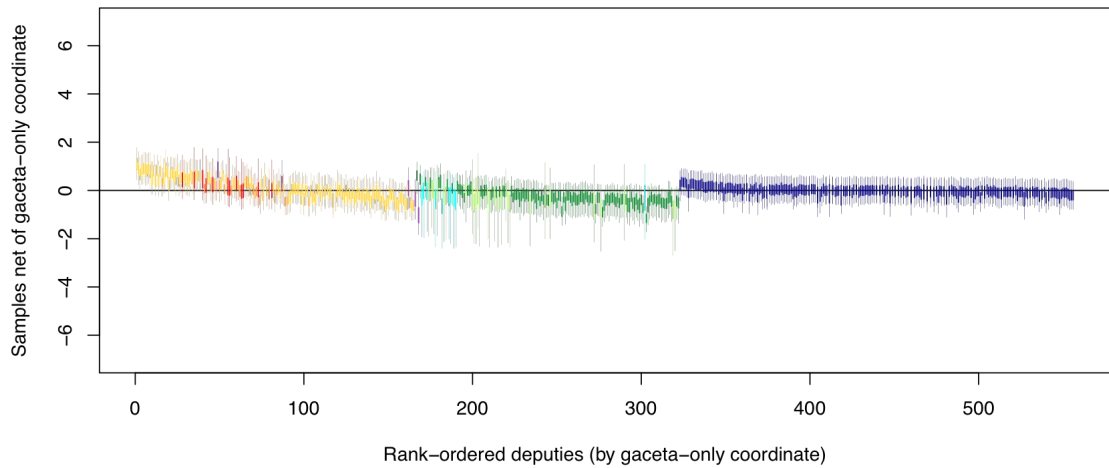


Figure 3: Dimension-by-dimension comparison of ideal points estimated with all votes (x-axis) and *Gaceta*-reported votes (y-axis)

Figura 3: Comparación coordinada-por-coordenada de los puntos ideales estimados con todos los votos (eje x) y con los votos publicados en la *Gaceta* (eje y). Nota: versiones a color de todos los diagramas del artículo pueden encontrarse en <http://allman.rhon.itam.mx/~emagar/graphs/pygPaper/>.



**2nd dimension**

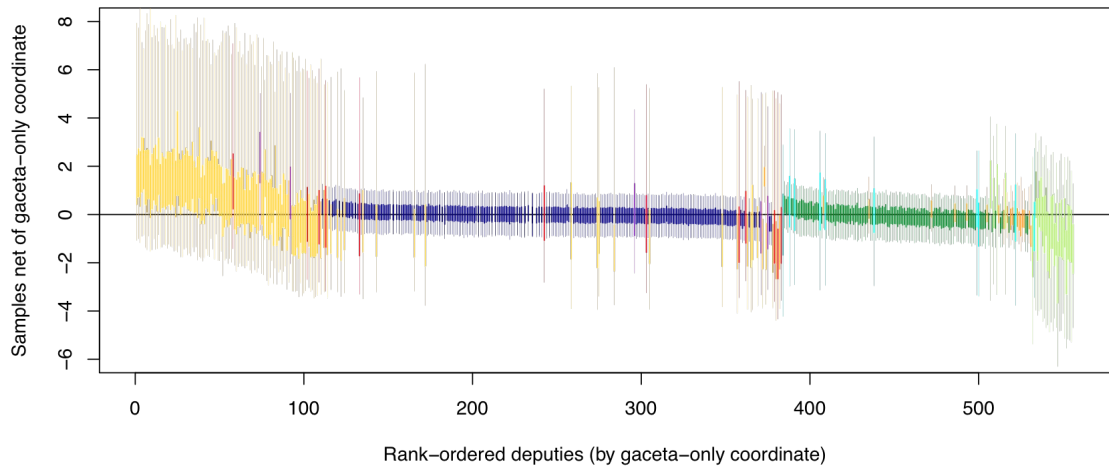


Figure 4: Samples of *Gaceta*-reported votes against theta, the point prediction of deputies's ideal scores using all votes, 60th Legislature

Figura 4: Puntos ideales estimados con muestras aleatorias de todos los votos y con los de la *Gaceta*, 60<sup>a</sup> Legislatura. El eje vertical indica el diferencial de las coordenadas obtenidas con las muestras y las obtenidas con la *Gaceta*. Nota: versiones a color de todos los diagramas del artículo pueden encontrarse en <http://allman.rhon.itam.mx/~emagar/graphs/pygPaper/>.

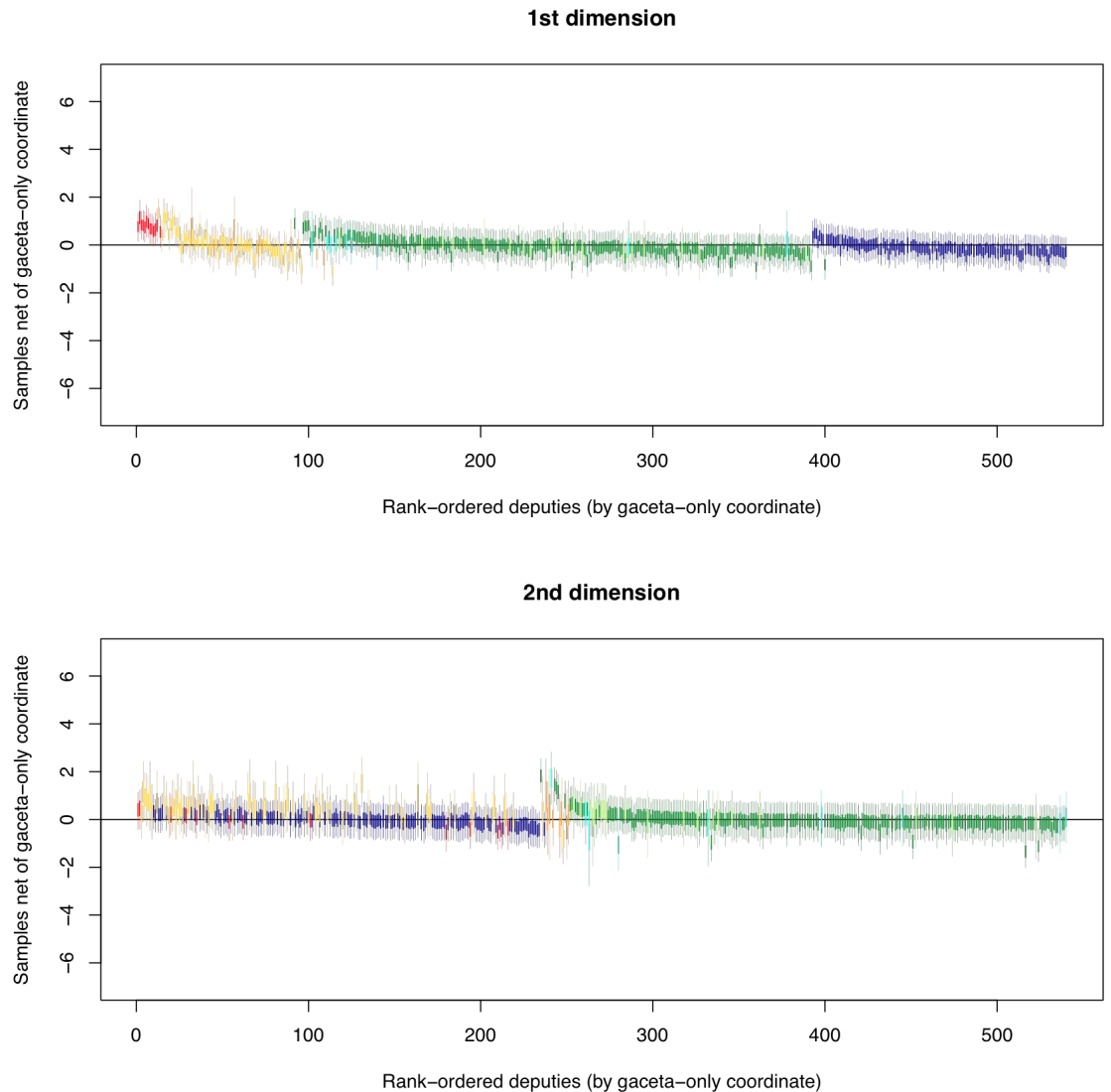


Figure 5: Samples of *Gaceta*-reported votes against theta, the point prediction of deputies's ideal scores using all votes, 61th Legislature

Figura 5: Puntos ideales estimados con muestras aleatorias de todos los votos y con los de la *Gaceta*, 61<sup>a</sup> Legislatura. El eje vertical indica el diferencial de las coordenadas obtenidas con las muestras y las obtenidas con la *Gaceta*. Nota: versiones a color de todos los diagramas del artículo pueden encontrarse en <http://allman.rhon.itam.mx/~emagar/graphs/pygPaper/>.



Tabla 1: Descripción de los votos *Gaceta* reportados y sin reportar, 57<sup>ava</sup>-61<sup>ava</sup>

legislaturas

Ver texto para criterios de categorización

Tipo de votación	Votaciones		Todas las votaciones
	reportadas en la <i>Gaceta</i>	no reportados en la <i>Gaceta</i>	
En lo general	68.5	8.4	56.8
En lo particular	24.4	12.0	22.0
Procedimental	5.3	26.3	9.4
Toma de posición	0.8	18.0	4.2
Proyecto privado	0.2	30.2	6.0
Viaje presidencial	0.8	5.1	1.6
Total	100%	100%	100%
(N)	(2,439)	(593)	(3,032)

**Tabla 2:** Comparación de la cohesión y la participación en Votaciones Publicadas e Inéditas, 57<sup>a</sup>-59<sup>a</sup> Legislaturas (1998-2006)

	Todas	<i>Gaceta</i>	Inéditas
Cohesión media	0.79	0.82	0.69
Media de votos	382	385	375
emitidos por votación			
Número de votaciones	1,702	1,285	417
nominales			
Porcentaje de	45	50	30
votaciones unánimes			

[Tabla 3](#): Cohesión media **anual por estado de publicaciónmuestra**.

Año	Todos los votos	Cohesión <i>Gaceta</i>	Inéditos	Diferencia
1998	0.64	0.62	0.67	-0.052
1999	0.62	0.61	0.66	-0.052
2000	0.71	0.77	0.60	0.171
2001	0.88	0.89	0.84	0.055
2002	0.84	0.87	0.78	0.086
2003	0.76	0.80	0.69	0.112
2004	0.79	0.84	0.60	0.239
2005	0.81	0.84	0.68	0.163
2006	0.92	0.94	0.73	0.215

**Tabla 4:** Votos Escalables y no escalables de los votos nominales, 2006–12

	60 <sup>a</sup>		61 <sup>a</sup>	
	Publicados en Gaceta	Todos los votos	Publicados en Gaceta	Todos los votos
Tipo de voto				
Unánimes	259	293	295	321
Casi sin oposición	65	82	138	150
Con oposición (escalables)	160	215	234	267
Participación media en votos escalables	370	365	348	341
Total	484	590	667	738