Estado de situación de encuestas

En todo el período revisado, se tienen 24 encuestas publicadas y fichas técnicas de 23 de ellas. Se revisaron las mismas y se vaciaron a una matriz inicial de análisis que se adjunta al documento. Las conclusiones iniciales del estado de situación son las siguientes:

* Aunque se tiene alguna información sobre las encuestas publicadas, solo en un caso (CELAG) hay datos ampliados sobre la metodología. Si se toman en cuenta los datos sin ello, es importante considerar que se está asumiendo que no hay errores metodológicos muestrales y que por tanto los datos son enteramente confiables. Por ejemplo, hay una encuesta que introduce criterios socioeconómicos, pero no es posible ver cómo operativizaron eso en el muestreo concretamente.
* Todas las encuestas tienen, teóricamente, el mismo nivel de confiabilidad de 95%, asumiendo que está bien aplicada la metodología, pero los márgenes de error varían.
* Es importante también, si se tuiviese acceso, contrastar que las preguntas sean compatibles y que no hayan preguntas que puedan interferir en la respuesta de los encuestados.
* El factor más importante para compatibilizar las diferentes encuestas en el tiempo es el alcance de la muestra. Hay 4 criterios centrales: ciudades del eje más El Alto, ciudades capitales más El Alto, ciudades intermedias y áreas rurales.
* Sin esta información en detalle, no es posible hacer esa valoración y en este momento solamente UNA de las encuestas cuenta con información específica sobre las localidades en las que se aplicó el formulario.
* La variable más importante para poder establecer una metodología de comparación es el alcance de la muestra. Mientras que el 100% de las encuestadoras consideran el llamado "eje" (La Paz, Cochabamba, Santa Cruz y El Alto), todo el resto del alcance puede ser muy variable y es necesario conocer los parámetros y las ciudades o comunidades que se incluyeron en la muestra. Se propone un instrumento como el siguiente ejemplo:



El instrumento nos permitirá tener una mirada general del alcance de las muestras utilizadas. Con ello, se puede hacer una valoración de las "piezas faltantes" en cada momento en el que tenemos una encuesta publicada:



Este instrumento nos permite identificar los meses en los que se cuenta con información más completa y nos dará pautas para la inferencia y la maetodología de "relleno" de los meses en los que no se tenga la información.

En ese caso, se considerará:

* El peso específico de las ciudades y comunidades faltantes ese mes respecto del padrón para determinar cuán grande sería el efecto de las mismas en el resultado.
* Cambios y variaciones en cada ciudad y localidad, para poder inferir resultados faltantes, por ejemplo en los casos en los que no haya datos pero la variación entre los datos que sí existan sea mínima.

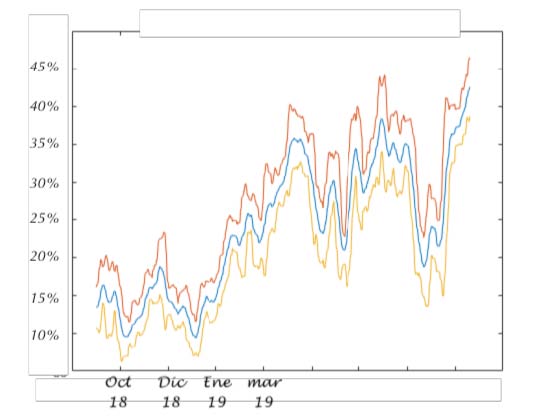
***Idea para rellenar vacíos de tos nacionales***

Generar inferencias en base a datos estadísticos de los cuales disponemos en lugar de tratar de extrapolar votación desagregada por partidos.

Es decir. Si asumimos que las encuestas históricamente han estado dentro de un margen de error aceptable, podemos usar datos de elecciones pasadas de los cuales se dispone, siguiendo uproceso similar al siguiente ejemplo:

1. En la mayor parte de las encuestas analizadas, solo se dispone de datos en el Eje Troncal + El Alto. Entonces se toman resultados de elecciones pasadas en EJE + El Alto (EAL), con el ﬁn de igualar la base de datos.
2. Una por una, en cada una de las elecciones pasadas, a estos datos de EJE+EAL se comparacon los datos a nivel NACIONAL, y se determina su variación, la cuál puede estardeterminada en puntos porcentuales de los, digamos, 5 partidos mas votados. Por ejemplo, supongamos que en la elección 2002, la votación de los primeros 5 fue así:
3. Se tabula una tabla análoga para cada elección nacional desde que se tienen datos, y se sacaun promedio de la diferencia absoluta.
4. Se puede estimar que el promedio de variación en esta ocasión estará dentro del promedio histórico. Se deduce entonces que si a los datos EJE+EAL se adicionan el resto de los datos nacionales, se obtendrá una banda de rango posible, digamos utilizando el resultado el ejemplo de +-3.83% si lo tomamos como el promedio total. Es decir, en el mejor de los casos al sumarle los datos nacionales, un partido en promedio obtendrá +3.83%, en el peor de los casos -3.83%
5. Se puede visualizar en forma de bandas:

|  |  |  |  |
| --- | --- | --- | --- |
| **Partido** | Nacional | 1. EJE+EAL | Diferencia (absoluta) |
| **MNR** | 22.46% | 24.13% | 1.67% |
| **MAS** | 20.94% | 26.98% | 6.04% |
| **NFR** | 20.91% | 16.42% | 4.49% |
| **MIR** | 16.32% | 12.06% | 4.26% |
| **MIP** | 6.09% | 8.78% | 2.69% |
| **Promedio para 2002** |  |  | 3.83% |



La línea azul sería la progresión de los resultado de las encuestas EJE+EAL. La línea amarilla sería el mejor de los casos si se le adiciona el resto de natos a nivel nacional, y la línea amarilla el peor de los casos.

**PROBLEMAS:**

- Como todo promedio, no contempla casos especíﬁcos, y por deﬁnición nada es el promedio, ósea que es una aproximación un poco cruda, pero con datos concretos, conocidos y bien calculados, sin supuestos ni extrapolaciones. -Al no tener datos históricos de ENCUESTAS y no conocer cómo en promedio las encuestas diﬁeren de los datos ﬁnales, estamos ignorando esa variable. Es decir, se asume que al rango que se presenta en el gráﬁco, habría que aumentarle un cierto porcentaje de variación que representaría la variación promedio histórica entre encuestas y resultados ﬁnales, pero esos datos no están disponibles.

-No se está tomando en cuenta, ni en las encuestas de base ni en la tabulación de batos, losresultados internacionales, que como se sabe pueden deﬁnir una elección