MIXED-MODE SAMPLING FOR PUBLIC OPINION RESEARCH: MUCH MORE THAN A PLAN B

Sébastien Dallaire
Senior Vice President, Ipsos Canada
1. Political Polling at Ipsos
2. Single Mode, or Mixed-Mode?
3. Core Challenges of Mixed-Mode Polling
4. A Different Kind of Mixed-Mode: The Knowledge Panel®
5. Art & Science of Polling
IPSOS AND POLITICAL POLLING
Ipsos believes electoral polling is a very important activity because it is part of the electoral process.

– It brings relevant information on what is happening in an election campaign to the public.

– Well done, and well communicated, this is good for democracy.

For several reasons, it’s becoming much harder to do accurate electoral polls.

– The political environment is more volatile, which makes it harder to capture the rapid changes that happen during the electoral campaigns.

– Audience harder to reach and more divided politically, socially

– Constantly changing telecommunications landscape

Even with the best methods and teams, there is always a level of uncertainty and risk of errors.

– Nonetheless, we always strive to conduct our polling with the maximum level of rigor, based on our global expertise, which is why at Ipsos this is not exclusively a local activity.
Globally, the **Head of Election Research** is responsible for working with the countries on forward planning, and all electoral projects are monitored during execution.

- At Ipsos, the Global Head of Election Research is Oliviero Marchese.

Following a local election campaign, a **systematic post-election evaluation is conducted** locally, with involvement from the global Ipsos team.

- Furthermore, scientific programs are proposed to continuously improve methods and share lessons learned.

At Ipsos we **look at political polling as a continuous learning process** as opposed to being event-driven, with the aim of building knowledge and expertise across the company.
SINGLE MODE OR MIXED-MODE?
At Ipsos, our global breadth and depth of operational capabilities allows us to take a “mode agnostic” approach to research.

The research design drives the decision of the appropriate data collection method, not the reverse. Keys: question types, sample size, target population.


That is about starting from the research need, not from the data collection method which is most easily available.

A tailored approach to data collection is needed, taking in consideration:

- Specificity of each country
- Potential sample coverage errors
- Political environment that can lead to measurement error because of desirability effects.
In the end, the choice of mixed versus single mode is a complex one.

In pre-election polling, mixed-mode initially tried to deal with coverage error, as different profiles of voters had a tendency to respond to different data collection methods.

The approach also proved useful at mitigating measurement errors, like interviewer related effects.

Nevertheless, such benefits must be evaluated in each country, taking in consideration the specific geography, technological environment, culture and political circumstances.

Pro-active use of mixed-mode can help research to address necessary trade-offs, in the context of providing the highest quality while focusing against the research project’s main priorities:

- representativeness
- measurement validity
- consistency
- speed
- cost
A cell-oriented sample design.

A good design can correctly cover all the sample cells…yet different data-collection modes can bring different respondent profiles for the same set of sample cells.

Mixed-mode can compensate uncontrolled difference(s) in respondent profiles, by sourcing the sample cells in a more balanced way whenever it is possible.
Measurement error can vary across survey modes (interviewer vs self-administered).

Interviewer-administered surveys:
- less privacy and anonymity leads to acquiescence and/or social desirability.
- Both effects can severely impact election polling, e.g. turnout estimates, past vote recall and voting intention.

Whether the answer is **spoken, written, or typed** can also affect responses: timing, the level of thought, and amount of detail provided.

Experience can help understanding measurement bias, although a preliminary analysis of the political market is needed.
CASE IN POINT: ONLINE RESPONDENTS LESS POSITIVE ABOUT CANADA’S IMMIGRATION LEVELS

Q. In your opinion, do you feel there are too many, too few, or about the right number of immigrants coming to Canada?

(Anaided) Mar-18 OL

- Too few: 8%
- About right: 43%
- Too many: 40%

(Anaided) Mar-18 TEL

- Too few: 14%
- About right: 55%
- Too many: 27%

+16 pts

CORE CHALLENGES OF MIXED MODE POLLING
Representation and measurement issues are often intertwined. Comparing results for a set of cells can help to appreciate and possibly to mitigate mode-effect(s).
SPECIFIC CHALLENGES OF MIXED-MODE

- **Worrying signals around the world.**
  - Social composition of samples becoming more problematic than ever, especially online (we often struggle to cover the most precarious segments of our societies).
  - The lack of representation is not necessarily visible through the usual stratification and control criteria (region, rural/urban, socio-demos, etc.).
  - Especially hard for sub-national representation: online panels not designed to provide (enough) localized sample. RDD does not always enable proper targeting of localized respondents.

- **Mixed mode seems to increasingly emerge as a good option** to address representation and measurement tradeoffs. But **several specific challenges** make it difficult to implement:
  - How to calibrate the share of each data collection mode? What kind of sample procedure can help to decide what the share of each data collection mode should be?
  - Beyond the usual quota control, **what kind of benchmarks should be used to evaluate the representation of the mixed-mode solution?**
There is limited knowledge on how to calibrate the share of each data collection mode, and decisions are frequently taken following some kind of rule of thumb: experience over theory.

Our Serbian team has explored using the Generalized Reduced Gradient (GRG) method to optimize the sample blend. This method minimizes the distance between two distributions (e.g., predicted vs. actual results of an election), and can be used to optimize the proportion of each mode.

The Serbian experience looks like a solid base for a more formalized approach of sample blend calibration, and we are evaluating the use of the GRG method in more countries, and in different research contexts, beyond election research.

We are working with Ipsos Global Science Org. to evaluate – and possibly improve - the existing procedure, investigate the interaction with weighting schemes, and finally produce a robust, scientifically sound methodology to piece together the multiplicity of elements that the sample design must assemble to reconstitute the reality of target population.

The idea is to define a two-step procedure: weighting first within each mode, then compute an optimal weight between modes, as this approach appears to be more robust when samples from different data sources are paired or at least weighted to make them comparable on relevant criteria (e.g., region * rural/urban, gender * age, employment status, etc.).
## OPTIMIZING SAMPLE BLEND. A CASE FROM CANADA [1/2] (ex post facto)

### CAN FED 2015

<table>
<thead>
<tr>
<th>Party</th>
<th>ONLINE col %</th>
<th>CATI col %</th>
<th>ACTUAL (TARGET)</th>
<th>SOLVER</th>
<th>sq error</th>
<th>chi distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Conservative Party</td>
<td>28</td>
<td>35</td>
<td>31,9</td>
<td>33,1</td>
<td>1,3</td>
<td>0,0</td>
</tr>
<tr>
<td>The Liberal Party</td>
<td>39</td>
<td>38</td>
<td>39,5</td>
<td>38,3</td>
<td>1,5</td>
<td>0,0</td>
</tr>
<tr>
<td>The New Democratic Party</td>
<td>25</td>
<td>20</td>
<td>19,7</td>
<td>21,4</td>
<td>2,8</td>
<td>0,1</td>
</tr>
<tr>
<td>The Bloc Québécois</td>
<td>4</td>
<td>4</td>
<td>4,7</td>
<td>4,0</td>
<td>0,5</td>
<td>0,1</td>
</tr>
<tr>
<td>The Green Party</td>
<td>5</td>
<td>4</td>
<td>3,4</td>
<td>4,3</td>
<td>0,8</td>
<td>0,2</td>
</tr>
</tbody>
</table>

| Unweighted                  | 0,60         | 0,40       |                 |        |          |              |
| Weight from Solver          | **0,28**     | **0,72**   |                 |        |          |              |

### CAN FED 2019

<table>
<thead>
<tr>
<th>Party</th>
<th>ONLINE col %</th>
<th>CATI col %</th>
<th>ACTUAL (weights 2015)</th>
<th>SOLVER</th>
<th>sq error</th>
<th>chi distance</th>
<th>Call Decision</th>
<th>sq error</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Conservative Party</td>
<td>30</td>
<td>36</td>
<td>34,4</td>
<td>34,3</td>
<td>0,0</td>
<td>0,0</td>
<td>33</td>
<td>2,0</td>
</tr>
<tr>
<td>The Liberal Party</td>
<td>32</td>
<td>31</td>
<td>33,1</td>
<td>31,3</td>
<td>3,3</td>
<td>0,1</td>
<td>31</td>
<td>4,4</td>
</tr>
<tr>
<td>The New Democratic Party</td>
<td>20</td>
<td>16</td>
<td>15,9</td>
<td>17,1</td>
<td>1,5</td>
<td>0,1</td>
<td>18</td>
<td>4,4</td>
</tr>
<tr>
<td>The Bloc Québécois</td>
<td>7</td>
<td>6</td>
<td>7,7</td>
<td>6,3</td>
<td>2,0</td>
<td>0,3</td>
<td>7</td>
<td>0,5</td>
</tr>
<tr>
<td>People’s Party</td>
<td>2</td>
<td>4</td>
<td>6,5</td>
<td>6,6</td>
<td>0,0</td>
<td>0,0</td>
<td>6</td>
<td>0,3</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>4</td>
<td>1,6</td>
<td>3,4</td>
<td>3,4</td>
<td>2,1</td>
<td>3</td>
<td>2,0</td>
</tr>
</tbody>
</table>

| Unweighted                  | 0,65         | 0,35       |                     |        |          |              |
| Solver 2015                 | **0,28**     | **0,72**   | **1,00**            | **2,6**| **10,3** | **13,5**    |

Applying the optimal combination of sample sources’ weights computed on Federal election 2015 to the blend of Federal election 2019 ... brings a sum of square errors lower than the one computed on the released “call poll”. Interesting enough, initial weights (before calibration) and optimal weights are very different. The present case clearly shows how strong the mode effect can be.
## OPTIMIZING SAMPLE BLEND. A CASE FROM CANADA [2/2] (ex post facto)

<table>
<thead>
<tr>
<th>CAN FED 2019</th>
<th>ONLINE col %</th>
<th>CATI col %</th>
<th>ACTUAL TARGET</th>
<th>SOLVER sq error</th>
<th>chi distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Conservative Party</td>
<td>30</td>
<td>36</td>
<td>34,4</td>
<td>34,2</td>
<td>0,1</td>
</tr>
<tr>
<td>The Liberal Party</td>
<td>32</td>
<td>31</td>
<td>33,1</td>
<td>31,3</td>
<td>3,2</td>
</tr>
<tr>
<td>The New Democratic Party</td>
<td>20</td>
<td>16</td>
<td>15,9</td>
<td>17,2</td>
<td>1,8</td>
</tr>
<tr>
<td>The Bloc Québécois</td>
<td>7</td>
<td>6</td>
<td>7,7</td>
<td>6,3</td>
<td>1,9</td>
</tr>
<tr>
<td>The Green Party</td>
<td>7</td>
<td>6</td>
<td>6,5</td>
<td>6,6</td>
<td>0,0</td>
</tr>
<tr>
<td>People’s Party</td>
<td>2</td>
<td>4</td>
<td>1,6</td>
<td>3,4</td>
<td>3,2</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>1</td>
<td>0,8</td>
<td>1,0</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Unweighted | 0,65 | 0,35 | COEF sum | 1,00 |
Weight from Solver | 0,31 | 0,69 |

<table>
<thead>
<tr>
<th>CAN FED 2021</th>
<th>ONLINE col %</th>
<th>CATI col %</th>
<th>ACTUAL TARGET</th>
<th>SOLVER sq error</th>
<th>chi distance</th>
<th>Call Decision</th>
<th>sq error</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Conservative Party</td>
<td>30</td>
<td>35</td>
<td>34,0</td>
<td>33,5</td>
<td>0,3</td>
<td>0,0</td>
<td>32</td>
</tr>
<tr>
<td>The Liberal Party</td>
<td>35</td>
<td>28</td>
<td>33,0</td>
<td>30,1</td>
<td>8,2</td>
<td>0,2</td>
<td>31</td>
</tr>
<tr>
<td>The New Democratic Party</td>
<td>21</td>
<td>20</td>
<td>18,0</td>
<td>20,3</td>
<td>5,3</td>
<td>0,3</td>
<td>21</td>
</tr>
<tr>
<td>The Bloc Québécois</td>
<td>7</td>
<td>8</td>
<td>8,0</td>
<td>7,7</td>
<td>0,1</td>
<td>0,0</td>
<td>7</td>
</tr>
<tr>
<td>The Green Party</td>
<td>3</td>
<td>3</td>
<td>2,0</td>
<td>3,0</td>
<td>1,0</td>
<td>0,5</td>
<td>3</td>
</tr>
<tr>
<td>People’s Party</td>
<td>3</td>
<td>6</td>
<td>5,0</td>
<td>5,1</td>
<td>0,0</td>
<td>0,0</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>1</td>
<td>1,0</td>
<td>1,3</td>
<td>0,1</td>
<td>0,1</td>
<td>1</td>
</tr>
</tbody>
</table>

Unweighted | 0,60 | 0,40 | COEF sum | 1,00 |
Solver 2019 | 0,31 | 0,69 |

- **Applying** the optimal combination of sample sources’ weights computed on Federal election 2019 to the blend of Federal election 2021 ...

- brings a sum of square errors lower than the one computed on the released “call poll”.

- Interesting enough, initial weights (before calibration) and optimal weights are very different.

- The present case clearly shows how strong the mode effect can be.
IDENTIFYING RELEVANT BENCHMARKS, TO EVALUATE SAMPLE REPRESENTATION BEYOND USUAL SOCIO-DEMOS

- As already mentioned, the lack of representation of our sample is not necessarily detectable through the usual stratification and control criteria (region, rural/urban, socio-demos, etc.).

- Beyond the usual quota control, **additional points of reference must be found to evaluate the quality of representation our sample provides.**

- **The adoption of a set of benchmarks** covering structures of attitudes, readership and more generally media behaviors, social positions related to the structuration of the employment can open new perspectives of analysis and provide decisive points of reference to evaluate the quality of representation of our samples.
AN EXAMPLE OF POSSIBLE BENCHMARK FROM FRANCE

For a rapid adoption of the « class of employment », to better apprehend casualization of labor and to identify precarious segments of society

- In 2020 the French National Institute of Statistics (INSEE) significantly redefined its categorization of the occupational status, promoting an innovative solution.

  - On one side, the detailed list of occupations was updated to better reflect the employment environment and trends, still keeping a good level of comparability over time with the previous nomenclature (at aggregated level, 13 professional categories).

  - One the other side, a social categorization was built to complement the historical one - the “class of employment” – which brings a frame of analysis for the social positions, taking in consideration the emerging divides in the structure of occupations.

Webinaire instituts.pdf
AN EXAMPLE OF POSSIBLE BENCHMARK FROM FRANCE

For a rapid adoption of the « class of employment », to better apprehend casualization of labor and to identify precarious segments of society

Addresses the need to better represent the divides which have a growing impact in the structuration of the employment and the determination of the social structure:

- salaried employees vs. independent workers, highly vs. less qualified, regular vs. casual workers, working in the public vs. private sector, in production vs. service orientation.
- It aims to support the analysis of the transformation our society has been going through the last decades.

The adoption of the “class of employment” frame, combined with a better categorization of our territories can open a new perspective of analysis and a decisive benchmark for the quality of representation of our samples.
OPERATIONAL CHALLENGES OF MIXED PHONE AND WEB

- Need two operations’ teams: phone and online
  - Two separate internal launch processes: kick-off meetings, planning, work-schedules, etc.
  - Two programming teams

- Different questionnaire formats.
  - Preestablished formats for online surveys to deal with allocation and reallocation needs.
  - Reading versus listening to question wording
  - Read all choices, but show them all as well?
  - Different programming approach and software.
Fieldwork management differs by mode

- Stratification of email invites: timing and subsample management.
- Random daily phone calls vs daily quota management online
- Yet: balanced final sample across modes

Necessitates daily planning and management of phone calls and emails.
IN PRACTICE: LONGER FIELD TIMES AND MORE EXPENSIVE

- Online survey generally cheaper
  - More expensive means less frequent.
  - Fewer questions by phone than online
  - Equal samples, or uneven?

- Two teams means more quality control time and complex data tabulation.
  - Two data files to merge post facto
  - Phone fieldwork generally longer
  - Complex weighting challenges (as discussed)

- Not possible to do for smaller firms without combined fieldwork capabilities.
A DIFFERENT KIND OF MIXED-MODE: THE KNOWLEDGE PANEL®
MIXED MODE APPLIED TO RECRUITMENT AND FIELDWORK?

- KnowledgePanel® is an online panel recruited through probabilistic means.
  - In the U.S. and U.K: recruited via scientifically developed addressed-based sampling.
  - Panel is more representative of the overall population.
  - Offline respondents provided with tablet and internet access.
  - Used for political surveys in the U.S. for past two years.

- Brings significant advantages.
  - No mode effect for fieldwork
  - Offline representation
  - Faster fieldwork compared to mixed-mode or phone

- More R&D needed to better understand impact on voting intentions.
ART AND
SCIENCE OF
POLLING

5
Mixed-mode surveys require a greater level of in-house expertise.

Researchers need to fully understand the strengths and limitations of the modes they combine.

Statistical theory helps us to measure and reduce sampling error. Real life is more complex.

The "art" of the pollster, survey practitioner, is to appreciate and understand:
- the causes generating representation and measurement errors
- their relative importance
- their effects on survey results
- the relative costs of the measures aimed to mitigate the errors

Practicing such an "art" implies more error modeling, combined with a deep understanding of the sources of error.

Mixed-mode can be part of the framework, as it can help improving sample coverage and mitigating measurement error, as long as:
- We can adopt a robust procedure to calibrate the share of each data collection mode
- We identify country specific benchmarks to evaluate the representation of the mixed-mode sample
CONTRIBUTORS

- Oliviero Marchese, Global Director, Election Research
- Gilbert Saporta, Emeritus Professor, CNAM, Paris
- Rich Timpone, Head of Global Science Org., Ipsos
- Marko Uljarevic, Public Affairs Director, Ipsos Serbia
- Aleksandar Zoric, Operations Director, Ipsos Serbia
THANK YOU