Conjoint Experiments*

Public Opinion Analytics Lab (POAL) Methods Brief Series

Thomas S. Robinson

10/11/2025

Executive Summary

- Conjoint experiments present multi-attribute *profiles*—e.g., policy packages, products—whose values (or *levels*) are randomised
- You can estimate the causal effect of levels by randomising them across profiles (e.g. the effect on policy support of raising taxes relative to holding them constant)
- These effect estimates are averages over the design's attribute distribution and are therefore *design-dependent*—effects are meaningful only relative to the set of attributes and their assignment in your design
- Conjoint profiles should have 5–8 realistic attributes, each with 2–5 levels, avoid implausible/impossible combinations, and give respondents multiple choice tasks.

1 Introduction

Would you rather vaccinate the old or the young first? What about a bus driver versus a solicitor? Would you be more likely to vaccinate the younger person first if they were also a bus driver? Political choices involve many dimensions, and we often want to separate their individual effects. While testing one dimension at a time can yield simple results, those effects may not generalise to richer, real-world contexts. For example, only asking if one wants to vaccinate the old or young first would not factor in how the recipient's occupation, likelihood to transmit the disease, and general health vulnerability factor into that consideration.

Conjoint experiments tackle this estimation issue by randomising *multiple* aspects of profiles simultaneously. They let analysts estimate the marginal impact of a feature—like age—averaged over the distribution of other facets of the profile (e.g., occupation). They are especially useful for choice-based contexts like policymaking and elections, though care is needed in how to interpret their results. Moreover, and unlike focus groups, they can be performed at large scale and quickly.

*

2 The method

In conjoint experiments, we construct *profiles* described by *attributes*, each taking one of several *levels*. In a policy setting, attributes might be the various dimensions one wants to pledge over. For example, a government funding attribute might have the levels 'increase income tax', 'increase borrowing', and 'maintain current funding balance'. In candidate settings, attributes may be features of the individual (both demographic and/or their political stances on issues).

In the survey, respondents (typically) see two profiles side-by-side with attribute levels randomised according to the design (ideally independently and uniformly). Figure 1 shows an example of a simple conjoint vignette, where respondents are asked to *choose* which policy package they prefer (a binary measure). Since assignment is random, differences in choices across profiles can be **causally attributed** to specific attribute levels.

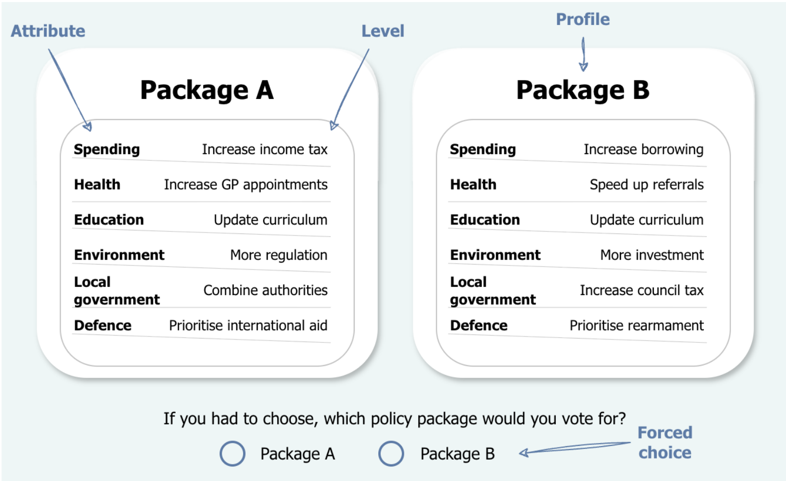

Figure 1: An example of a two-profile conjoint experiment vignette

A simple way to analyse conjoint data is to calculate the **marginal mean (MM)** of a level—the probability of choosing a profile when a specific attribute-level is shown:

$$\text{MM(Level)} = \mathbf{E}[\,Y \mid \text{Attribute} = \text{Level}\,]$$

with all other attributes averaged over the design.

Marginal means are *not* treatment effects, but they give a good indication of whether, on average, people are more or less likely to choose profiles with a certain attribute-level. MMs are also more straightforward to compare when looking at results across subset groups.

To understand the *causal effect* of an attribute level, we typically estimate the **average marginal component effect (AMCE)**: the average *change* in the probability of choosing a profile when an attribute is set to a specific level, relative to a reference value, and averaged over the design's randomisation of all other attributes. Averaging over the distribution of other attributes is important because an individual may react differently to the same level depending on what else is being shown.

In practice, assuming all attribute-levels are independently and uniformly randomised, AMCEs can be estimated with a linear probability model:

$$\Pr(\text{Choose profile}) = \alpha + \sum_{\text{Att.,Level}} \beta_{\text{Att.,Level}}\, D_{\text{Att.,Level}} + \varepsilon,$$

where $\beta_{\text{Att., Level}}$ is the AMCE for a non-reference level, $D_{\text{Att., Level}}$ is a binary indicator of whether the attribute-level was shown in the profile. You should cluster standard errors by respondent, and can include task fixed effects.

It is essential to be precise about the generalisability of estimated AMCEs. These effects are conditional on the specific attributes shown; adding or dropping other attributes may change the estimates. Moreover, since the AMCE is estimated as an average across other attributes, the distribution of attributes matters greatly. For example, under independent and uniform assignment, the probability of being young and single equals that of being old and single. In the real world, these two conditional probabilities are unlikely to be the same. This affects how we interpret the AMCEs because it may be that some levels are much less or more likely than others, yet are given equal weight under the design when estimating the treatment effect.

3 Implementation guidelines

Recommended Practice

1. **Number of attributes/levels** Aim to include enough attributes that the choice is meaningful and realistic, but not so many that the subject will find it hard to parse the information. 5–8 attributes, each with 2–5 levels, is a good starting point.
2. **Pose multiple choices per subject** Detecting AMCEs with magnitudes of 2–4 percentage points often needs $>$10,000 profile evaluations (respondents $\times$ profiles per task $\times$ tasks). Evidence suggests posing 10–15 tasks is acceptable with good design.
3. **Avoid impossible combinations** It is possible to restrict the design to prevent impossible combinations—like a 20 year old with 40 years prior experience—but you will need to adapt the estimation (e.g., using the known assignment probabilities). It is often simpler to design attributes where all possible combinations are plausible
4. **Mind generalisation.** Interpret effects as averaging over your design's attribute distribution; avoid policy counterfactuals you did not design for.

4 Further reading

Hainmueller, J., D. Hopkins, and T. Yamamoto. 2014. "Causal Inference in Conjoint Analysis." *Political Analysis.*

Leeper, T., S. Hobolt, and J. Tilley. 2020. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis.*

Duch, R. *et al.* 2021. "Citizens from 13 countries share similar preferences for COVID-19 vaccine allocation priorities." *Proceedings of the National Academy of Sciences.*

de la Cuesta, B., N. Egami, and K. Imai. 2022. "Improving the External Validity of Conjoint Analysis." *American Journal of Political Science.*