

Quality of BCS data Task Force 1 - Sample Frames – Summary & Key Results

The terms of reference for taskforce 1 were:

Analysis of sampling frames across institutes: appropriateness and comprehensiveness of sampling frames, theoretical considerations, empirical evidence on links with data volatility and bias

We collected data on four key metrics concerned the frame properties of 104 business surveys across 26 countries and 4 sectors (industry, services, retail and construction):

- **Institute type** (academic, business association, private body, statistical institute, other public bodies)
- **Frame type** (bought list, internally compiled list, national register, private register, combination of frame types)
- **Effective frame size**, both in absolute terms and as a proportion of the business population being covered by the survey
- **Update frequency** – the typical number of years before the frame is updated

The impact of these properties of the sampling frames were then analysed for their impact upon survey volatility and tracking performance, as measured by months for cyclical dominance (MCD) and correlation against a reference series respectively).

We took four main approaches to the analysis:

- Cross-checks between frame characteristics
- Headline descriptive statistics on average MCD and correlation across surveys with particular frame characteristics
- Scatter-plots of frame size/coverage and update frequency
- Regression analysis of volatility and tracking performance measures on frame characteristics.

Our key findings are as follows:

- **There are significant links between institute and frame type and frame coverage/updating frequency**, making it difficult to dis-entangle their independent effects on volatility and tracking performance
- **Frame size and updating frequency explain only a small part of the differences in volatility and tracking performance between surveys**, which means that omitted variables rendered a detailed regression analysis impracticable
- Regarding the determinants of volatility:
 - **Updating frequency is a key determinant of survey volatility.**
 - However, it explains only part of the variation in survey MCDs – **updating frequency can be thought of as ‘necessary, but not sufficient’ for low volatility**
 - Frame size appears to have a lesser influence – though a small sub-set of surveys with very large absolute frame sizes do have low volatility
- Regarding the determinants of tracking performance:
 - **Frame coverage over about 35% is associated with a stronger tracking performance**
 - Below 35%, the relationship is less clear. **Frame coverage can be thought of as ‘sufficient, but not necessary’ for strong correlation**
 - The relationship with updating frequency looks relatively weak