Robust GREG and Calibration Estimators in Official Statistics

Addressing Systematic Errors in Auxiliary Variables

1. Executive Summary

In official statistics, the **Generalized Regression Estimator (GREG)** and the **Calibration Estimator** are standard tools for improving estimation precision and ensuring consistency between sample surveys and population registers. Seminal work by **Särndal, Swensson, and Wretman (1992)** and **Deville and Särndal (1992)** established the framework where sample weights are adjusted so that weighted auxiliary totals exactly match known population totals.

However, a critical vulnerability exists: standard calibration assumes the auxiliary population totals ($\mathbf{t}_x$) are error-free. When these "known" totals contain **systematic (non-random) errors**—due to register outdatedness, definitional mismatches, or coverage errors—forcing exact calibration introduces bias and instability. This summary reviews robust methods designed to compensate for such errors, specifically **Penalized (Ridge) Calibration** and **Relaxed (Soft) Calibration**.

2. The Standard Framework and its Vulnerability

The Deville and Särndal Paradigm (1992)

The standard calibration estimator seeks a new set of weights w_k that are close to the design weights d_k (usually $1/\pi_k$) while satisfying the constraint:

$$\sum_{k \in s} w_k \mathbf{x}_k = \mathbf{t}_x$$

where $\mathbf{t}_x$ is the vector of *assumed* perfect population totals derived from a census or administrative register.

The Problem of Systematic Auxiliary Error

In modern official statistics, $\mathbf{t}_x$ is rarely perfect. Systematic errors arise from:

- **Time Lags:** The register refers to the population on January 1st, but the survey is conducted in June.
- **Definitional Differences:** The register defines "unemployed" differently than the ILO definition used in the survey.
- **Register Coverage:** The administrative source may suffer from under-coverage (e.g., undocumented residents) or over-coverage (e.g., emigrants not removed from the list).

If $\mathbf{t}_x$ contains a systematic bias Δ, forcing $\sum w_k \mathbf{x}_k = \mathbf{t}_x$ essentially "calibrates the survey to the wrong target," shifting the bias from the register into the survey estimates.

3. Robust Methods for Systematic Auxiliary Errors

Research has moved beyond simple outlier treatment (robustness to y-outliers) to address **robustness to auxiliary error** (robustness to x-errors).

A. Penalized "Ridge" Calibration

Originating from **Chambers (1996)** and further developed by **Rao and Singh (1997, 2009)** and **Beaumont and Bocci (2008)**, this approach relaxes the requirement for exact consistency.

Instead of a constrained minimization, the problem is formulated as a penalized minimization:

$$Q = \sum_{k \in s} \frac{(w_k - d_k)^2}{d_k} + \lambda \left(\sum_{k \in s} w_k \mathbf{x}_k - \mathbf{t}_x \right)^T C^{-1} \left(\sum_{k \in s} w_k \mathbf{x}_k - \mathbf{t}_x \right)$$

- **Mechanism:** The estimator balances the distance between old and new weights against the "calibration error" (the distance between the weighted sample total and the register total).
- **Result:** The weights do *not* exactly reproduce $\mathbf{t}_x$. If $\mathbf{t}_x$ is suspected of having systematic error, the parameter λ allows the estimator to drift away from the suspect control total, letting the survey data "speak for itself" rather than being forced into a biased mold.

- **Benefit:** This provides a trade-off between consistency (matching the register) and bias reduction (ignoring a flawed register).

B. Relaxed (Soft) Calibration with Tolerances

Proposed in recent years (e.g., **Guggemos and Tillé 2010**, **Rupp 2018**), this method replaces the equality constraint with an inequality constraint based on a tolerance δ:

$$\mathbf{t}_x - \boldsymbol{\delta} \leq \sum_{k \in s} w_k \mathbf{x}_k \leq \mathbf{t}_x + \boldsymbol{\delta}$$

- **Application:** This is particularly useful when $\mathbf{t}_x$ is derived from a *model* or a *previous survey* rather than a census.
- **Systematic Error Handling:** If a statistical agency knows that a register has a systematic lag error of roughly 1%, they can set $\boldsymbol{\delta}$ accordingly. The solver finds weights that fall within this "truth interval" without over-adjusting to a precise, erroneous point.

C. Bias-Corrected GREG (Measurement Error Models)

Work by **Statistics Netherlands (CBS)** and **Statistics Canada** has explored scenarios where $\mathbf{t}_x$ is estimated via a time-series model (e.g., for monthly unemployment figures).

- **The Method:** The variance of the GREG estimator is adjusted to account for the variance in $\mathbf{t}_x$.
- **Bias Correction:** If the systematic error (bias) in the auxiliary variable is estimable (e.g., via a coverage survey), a bias correction term $-\sum \hat{B}_{register}$ is added to the standard GREG estimator.

4. Summary of Key Differences

Feature	Standard Calibration (Deville & Särndal)	Robust/Ridge Calibration (Chambers/Rao)
Goal	Exact consistency with register	Balance consistency vs. Mean Squared Error
Constraint	$\sum w_k x_k = t_x$ (Hard)	$\sum w_k x_k \approx t_x$ (Soft/Penalized)

Assumed t_x	Perfect (Truth)	Imperfect (Contains error/bias)
Risk	Imports register bias into survey	Potential cosmetic inconsistency
Best for	High-quality, up-to-date Censuses	Outdated registers or definitional mismatches

5. Selected References

- **Deville, J.-C., & Särndal, C.-E. (1992).** Calibration estimators in survey sampling. *Journal of the American Statistical Association*. (The foundational text).
- **Chambers, R. L. (1996).** Robust case-weighting for multipurpose establishment surveys. *Journal of Official Statistics*. (Introduces the ridge/penalized concept).
- **Rao, J. N. K., & Singh, A. C. (1997).** A ridge-shrinkage method for range-restricted weight calibration in survey sampling. *Proceedings of the Section on Survey Research Methods*.
- **Beaumont, J.-F., & Bocci, C. (2008).** Another look at ridge calibration. *Metron*.
- **Särndal, C.-E. (2007).** The calibration approach in survey theory and practice. *Survey Methodology*. (Reflects on the limitations of exact calibration).