Chapter 1

Architecting Robust Anomaly Detection in Econometrics

Madhu Babu Amarappalli
https://orcid.org/0009-0009-2723-0256
Deloitte Consulting LLC, USA

ABSTRACT

Anomalous observations are no longer rare irregularities in econometric data, but recurring features of modern empirical environments shaped by measurement heterogeneity, nonstationary, and structural change. In such settings, conventional estimation and inference can become fragile, allowing a small fraction of atypical observations to exert disproportionate influence on estimates, model selection, and uncertainty quantification. This chapter develops an econometric framework for robust anomaly handling that foregrounds robust estimation and inference preservation under contamination, synthesizing core ideas from robust econometrics bounded influence, high-breakdown resistance, and contamination models with modern anomaly detection tools. Rather than treating anomaly detection as a preprocessing step, we frame it as an inference-aware component of the econometric workflow and highlight the risks introduced by data-dependent filtering, trimming, or down-weighting.

1. INTRODUCTION

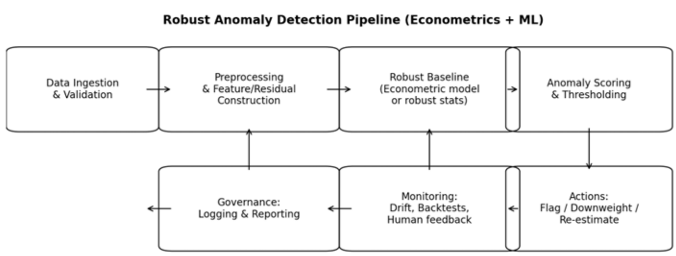
Econometric analysis is increasingly conducted in data environments characterized by scale, heterogeneity, and fragility. Administrative microdata, high-frequency financial records, platform logs, and sensor streams allow researchers to study

DOI: 10.4018/979-8-3373-8297-5.ch001

economic behavior at unprecedented granularity and often in near real time. These gains, however, come with a structural cost: modern datasets are routinely shaped by heterogeneous measurement systems, nonstationary behavior, and rare but economically consequential events. As a result, the assumptions that underlie many classical estimators thin-tailed errors, stable regimes, and largely uncontaminated measurement are frequently strained or violated. Observations that would appear as isolated "outliers" in textbook examples may instead reflect pipeline failures, reporting changes, policy shocks, strategic responses to incentives, or genuine structural disruptions. In contemporary empirical work, anomalous observations are therefore not peripheral irregularities; they are recurring features of the data-generating process with the potential to materially affect estimation, interpretation, and inference.

Robust econometrics provides a principled framework for addressing these challenges by explicitly controlling estimator sensitivity to contamination. Influence functions formalize local robustness by quantifying how small perturbations in the data affect estimates, directly linking robustness to bias control and estimator stability (Hampel et al., 1986). Breakdown points capture global robustness by measuring the largest fraction of contaminated observations an estimator can tolerate before yielding arbitrarily misleading results, thereby defining worst-case reliability (Hampel et al., 1986; Maronna et al., 2006). These concepts are not merely theoretical: they motivate operational choices such as bounded-influence loss functions, robust scale estimation, and high-breakdown regression methods. Huber's M-estimation framework illustrates the core bias–variance tradeoff underlying robustness: by tempering the influence of extreme residuals, one can substantially reduce sensitivity to contamination while retaining high efficiency under well-behaved data (Huber, 1964). Robustness, in this sense, is not about discarding anomalies, but about preventing a small subset of observations from silently dominating econometric conclusions.

Figure 1. Robust anomaly detection pipeline



In modern empirical pipelines, robustness cannot be treated as an after-the-fact diagnostic applied only at the final estimation stage. Many consequential decisions

Related Content

Use of Big Data Analytics by Tax Authorities
Brendan Walker-Munro (2022). *Research Anthology on Big Data Analytics, Architectures, and Applications (pp. 1388-1412).*
www.irma-international.org/chapter/use-of-big-data-analytics-by-tax-authorities/291043

COVID-19 Sentiments and Impact on Stock Market Prices
Chandra Prayaga, Krishna Devulapalli, Lakshmi Prayagaand Aaron Wade (2021). *International Journal of Data Analytics (pp. 40-58).*
www.irma-international.org/article/covid-19-sentiments-and-impact-on-stock-market-prices/285467

A Review of Data Governance Definitions and Emerging Perspectives
Uma G. Guptaand San Cannon (2020). *International Journal of Data Analytics (pp. 30-47).*
www.irma-international.org/article/a-review-of-data-governance-definitions-and-emerging-perspectives/258919

Comprehensive Analysis of State-of-the-Art CAD Tools and Techniques for Chronic Kidney Disease (CKD)
Mynapati Lakshmi Prasudha, Rakesh Kasumollaand Deepak Sukheja (2021). *International Journal of Big Data and Analytics in Healthcare (pp. 1-12).*
www.irma-international.org/article/comprehensive-analysis-of-state-of-the-art-cad-tools-and-techniques-for-chronic-kidney-disease-ckd/287605

A Survey on Comparison of Performance Analysis on a Cloud-Based Big Data Framework
Krishan Tuli, Amanpreet Kaurand Meenakshi Sharma (2022). *Research Anthology on Big Data Analytics, Architectures, and Applications (pp. 1865-1875).*
www.irma-international.org/chapter/a-survey-on-comparison-of-performance-analysis-on-a-cloud-based-big-data-framework/291070