Bridging the Temporal Gap: A Comparative Analysis of Matrix Updating Techniques for the Portuguese Economy (2010-2023)

José Miguel Rebolho*
CeBER, Faculty of Economics, University of Coimbra
Luís Cruz
CeBER, Faculty of Economics, University of Coimbra
and
João Pedro Ferreira
Weldon Cooper Center for Public Service, University of Virginia

Abstract

Input-Output (I-O) matrices provide crucial insights into economic interdependencies but are frequently constrained by multi-year publication lags. This study assesses nine matrix updating procedures – including RAS, GRAS, and various optimization approaches (e.g., INSD, WSD) – to determine their efficacy in preserving intersectoral relationships within a small open economy characterized by high structural asymmetries. Utilizing Portuguese Supply and Use Tables (SUTs) from 2010 to 2023, we shift the focus from static reconciliation to dynamic forecasting across two resolution scales: macro-aggregated (88×80) and highly disaggregated, sparse matrices (431×125). Our empirical findings reveal that calculating multi-year averages minimizes discrepancies in aggregated matrices by leveraging structural memory, utilizing a single-year gap substantially improves accuracy in high-resolution models. In sparse environments, adapting swiftly to recent economic shocks proves superior to averaging out noise. Ultimately, we demonstrate that the Generalized RAS (GRAS) framework, anchored to the immediately preceding year, provides the most reliable methodology for projecting detailed regional models during periods of macroeconomic volatility.

Keywords: SUT Projection; GRAS; INSD; Matrix Updating; PReMMIA.

*Corresponding author. Email: jose.d.rebolho@fe.uc.pt. ORCID: 0000-0003-4800-699X.

1 Introduction

Originating from the foundational work of Leontief (1936), Input-Output (I-O) models remain indispensable for mapping multisectoral dependencies and quantifying economic shocks. By systematically linking industry supply chains, these frameworks allow policymakers to anticipate the economy-wide repercussions of localized disruptions. However, the operational effectiveness of I-O models is frequently undermined by temporal data constraints. National statistical agencies often operate with publication delays spanning several years. Such gaps become highly problematic during periods of severe global instability – such as the COVID-19 pandemic, geopolitical conflicts in Eastern Europe, or global supply chain realignments – where capturing real-time economic structure is vital.

To overcome missing data, matrix updating algorithms estimate current values using historical tables and known contemporaneous marginal totals. Historically, the literature categorizes these algorithms into biproportional techniques, most notably the classic RAS (Stone, 1961) and its generalized form GRAS (Junius and Oosterhaven, 2003), and mathematical optimization approaches that minimize target distances, such as SPAD or INSD (Jackson and Murray, 2004; Huang et al., 2008).

Despite decades of methodological evolution, there remains an empirical vacuum regarding which algorithm performs optimally when dealing with high-resolution sparse matrices, typical of regional economic models. Furthermore, traditional updating assessments generally rely on stable periods, leaving their robustness under modern macroeconomic volatility mostly untested.

This paper bridges this gap by systematically evaluating nine updating methods against the Portuguese Supply and Use Tables (**SUT**s) spanning 2010 to 2023. By comparing performance between highly aggregated macro-matrices and deeply granular micro-matrices, we explore whether the temporal baseline of an estimation – using an immediate past year versus a multi-year average – alters algorithmic efficiency.

2 Literature Review

The need to bridge the gap between historical I-O surveys and current economic reality has driven significant methodological advances in matrix updating. The biproportional RAS algorithm (Stone, 1961) established the foundation for this field due to its computational parsimony and ability to preserve macro-consistency. However, standard RAS is mathematically unequipped to handle negative elements natively or manage structurally sparse matrices without divergence (Möhr et al., 1987). To address this, the Generalized RAS (GRAS) method was developed (Junius and Oosterhaven, 2003; Lenzen et al., 2007), allowing the simultaneous adjustment of matrices containing both positive and negative values – a vital capability for handling inventory changes and net exports in standard national accounts.

In parallel, optimization-based techniques emerged as a distinct family of solutions. Instead of iterative scaling, these approaches frame updating as a constrained mathematical programming problem. Jackson and Murray (2004) formalized methodologies like Sign-Preserving Absolute Differences (SPAD) and Sign-Preserving Squared Differences (SPSD), demonstrating their utility when strict adherence to original flow directions is required. Subsequent innovations sought to eliminate optimization bias; Huang et al. (2008) introduced the Improved Normalized Squared Differences (INSD) algorithm, which frequently outperforms classical techniques in minimizing structural distortion. Furthermore, weighted formulations (WSD, IWSD) were proposed to prioritize the accuracy of critical, high-volume transactions (Lahr, 1998).

Empirically, comparative studies suggest that algorithmic performance is highly context-dependent. Studies by Temurshoev et al. (2011) and Révész (2023) indicate that while GRAS handles standard constraints exceptionally well, optimization approaches like INSD are highly competitive when structural shifts are pronounced. Nevertheless, the overwhelming majority of empirical validations have been conducted on small, dense, aggregated symmetric tables (e.g., 60×60). Sparse, rectangular **SUT**s at high resolutions remain under-researched, leaving open questions about the scalability and practical reliability of these methods for granular regional modeling contexts, such as those utilized in Portugal – e.g., the PReMMIA framework (Ferreira et al., 2025).

3 Methodology

To robustly assess updating techniques, we employ a consistent notation where $\mathbf{Z}$ represents the Transaction Matrix, $\mathbf{F}$ the Final Demand Matrix, and $\widetilde{\mathbf{Z}}$ the Expanded Transaction Matrix. Target margins are defined by vectors $\mathbf{u}$ (row sums) and $\mathbf{v}$ (column sums).

3.1 Algorithmic Frameworks

Biproportional Scaling: The traditional RAS algorithm iteratively adjusts rows and columns using multiplicative scaling factors ($\mathbf{r}$ and $\mathbf{s}$) such that $\overline{\widetilde{\mathbf{Z}}}^{*} = \mathbf{R} \cdot \widetilde{\mathbf{Z}} \cdot \mathbf{S}$. Because standard RAS fails with negative numbers, GRAS (Junius and Oosterhaven, 2003) circumvents this by mathematically splitting the baseline into distinct positive ($\mathbf{P}$) and negative ($\mathbf{N}$) matrices, applying biproportional adjustments independently, and recombining them.

Optimization Procedures: We tested several optimization models utilizing the `Ipopt` nonlinear solver (Wächter and Biegler, 2006). For instance, the SPAD and SPSD models strictly minimize absolute and squared deviations, respectively:

$$\min \sum_{i,j} \left| \widetilde{z}_{i;j} - \overline{\widetilde{z}}_{i;j} \right| \quad \text{or} \quad \min \sum_{i,j} \left(\widetilde{z}_{i;j} - \overline{\widetilde{z}}_{i;j} \right)^2 \tag{1}$$

subject to row/column margin constraints and strict sign preservation.

Normalized variants, notably INSD, minimize proportional deviations to prevent large sectors from exclusively dominating the objective function:

$$\min \sum_{i=1}^{m} \sum_{j=1}^{n} |\widetilde{z}_{i;j}| \cdot \left(\frac{\overline{\widetilde{z}}_{i;j}}{\widetilde{z}_{i;j}} - 1 \right)^2 \tag{2}$$

Weighted approaches (WSD, IWSD) incorporate base-year magnitudes as weights, explicitly forcing tighter fits on historically large flows.

Before applying the updating algorithms, the raw Portuguese SUTs required rigorous preprocessing to ensure mathematical stability and algorithmic convergence. This predominantly involved the systematic handling of structural zeros, which can trigger undefined operations in proportion-based metrics (such as MAPE) or normalized algorithms (like NSD). To mitigate this, algorithms were programmed to enforce strict zero-preservation constraints where applicable, ensuring that cells historically devoid of transactions were not artificially inflated by the updating process. Division-by-zero errors during evaluation were bypassed by neutralizing undefined percentage calculations.

The rationale for deploying a multi-metric validation framework stems from the inherent biases present in individual statistical measures. Scale-dependent metrics, such as MAD and RMSE, accurately capture absolute volumetric differences, disproportionately penalizing large nominal errors. However, they can obscure structural shifts in smaller sectors. Conversely, scale-independent metrics like MAPE, alongside the associative Pearson Correlation Coefficient (CORR), map the structural fidelity and linear alignment of the estimated matrices independent of gross transaction sizes. This triangulated approach prevents methods that simply fit well to aggregate GDP constraints from masking severe micro-level distortions.

3.2 Experimental Design: Gap vs. Mean Analysis

To analyze temporal stability, the base matrix $\widetilde{\mathbf{Z}}_{[t]}$ was established using two distinct strategies:

1. **Gap-Based Estimation**: Projecting the target matrix using a single historical year. For brevity, we report *Gap 1* (immediate past year, prioritizing fast shock adaptation) and *Gap 5* (a five-year lag representing common statistical publication delays).

2. **Average-Based Estimation (Means)**: Constructing a synthetic baseline by averaging the **SUT**s of the preceding years. This tests whether smoothing historical volatility provides a superior structural anchor compared to single-year observations.

Methodological validation relied on five standard metrics: Mean Absolute Deviation (MAD), Weighted Absolute Differences (WAD), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and the Pearson Correlation Coefficient (CORR).

4 Results

4.1 Macro-Aggregated Matrices (88x80)

The performance analysis on standard aggregated matrices reveals clear stratification among the algorithms. GRAS and INSD consistently lead all accuracy metrics, managing to securely restrict major deviations to inherently volatile categories, such as Changes in Inventories. Biproportional RAS functions as a competent alternative but demonstrates marginally higher residual variation.

Figure 1: Heatmap of **SUT** Estimation Errors for 2022 with a 1-Year Reference Gap – INSD

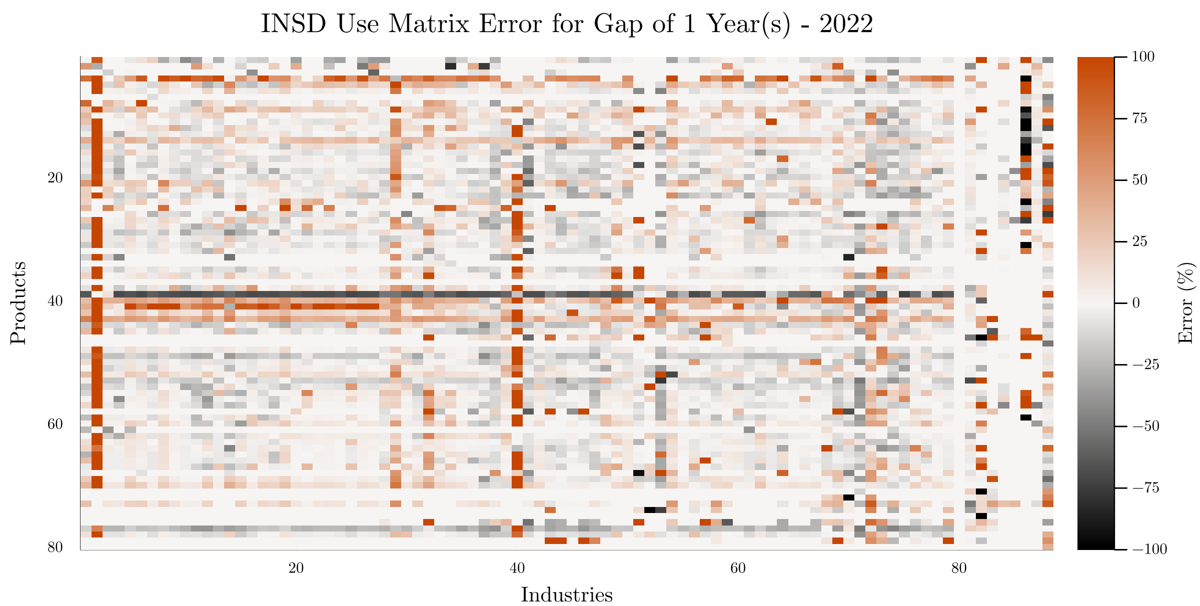
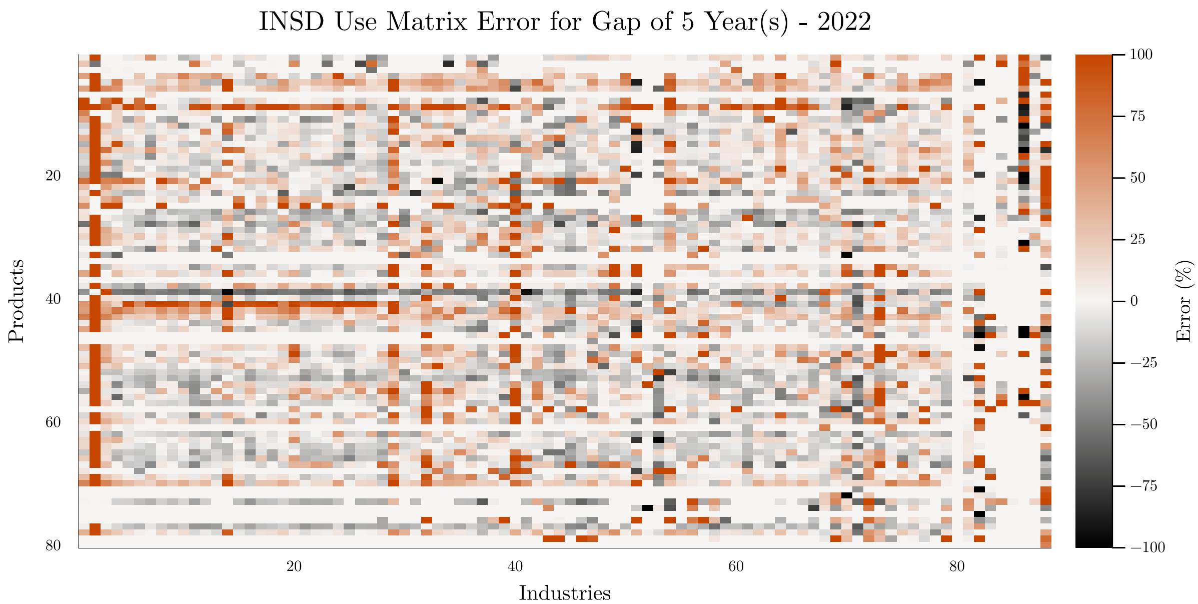


Figure 2: Heatmap of **SUT** Estimation Errors for 2022 with a 5-Year Reference Gap – INSD



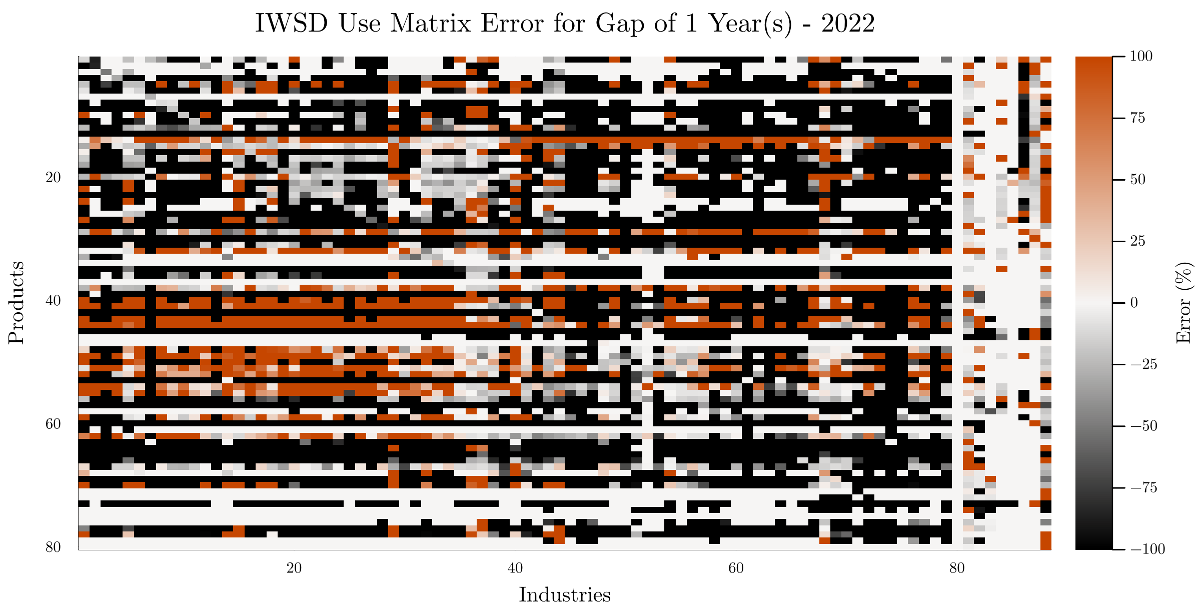
Conversely, distance-based and weighted optimizations struggle noticeably. Algorithms like SPAD, SPSD, WSD, and IWSD generate significantly higher MAPE and weighted discrepancies. The absolute enforcement of zero-deviations or purely squared penalties in these methods appears overly rigid during periods of structural evolution, frequently redistributing major imbalances into smaller sectors and distorting the underlying technical coefficients.

At this aggregated scale, a clear pattern regarding the historical baseline emerges: utilizing a multi-year average (*Means*) generally outperforms relying on a single temporal gap. Averaging effectively smooths temporary stochastic noise within the broad sectors, providing a stable foundation that algorithms like GRAS and INSD leverage to minimize global errors.

A closer examination of the sector-specific error distributions reveals that the superiority of GRAS and INSD largely derives from their proportional adjustment logic. Unlike SPAD, which forces absolute adjustments and often severely distorts smaller intermediate flows to satisfy marginal totals, GRAS gracefully scales coefficients. As a result, the inevitable errors stemming from major shocks (e.g., sudden shifts in energy inputs or specific service sectors) remain localized rather than bleeding indiscriminately across the transaction matrix.

Furthermore, the temporal degradation of accuracy is distinctly visible when comparing the Gap 1 and Gap 5 scenarios. Even within the more stable macro-aggregated matrices, estimating an I-O table over a five-year horizon produces substantially wider error tails across all algorithms. While the Means strategy remains optimal overall at this resolution, when practitioners are forced to use a single baseline year, Gap 1 dramatically curtails systemic divergence. This suggests that while minor noise can be averaged out, the accumulation of permanent structural drift over half a decade renders Gap 5 estimations highly vulnerable to cumulative distortion.

Figure 3: Heatmap of **SUT** Estimation Errors for 2022 with a 1-Year Reference Gap – IWSD



4.2 High-Resolution Sparse Matrices (431x125)

When shifting the evaluation to the deeply disaggregated Portuguese matrices, algorithmic behavior transforms. While GRAS and INSD maintain their dominance (sustaining CORR metrics above 0.998), standard RAS exhibited structural collapse during the highly volatile 2021 period, showing an inability to resolve conflicting constraints driven by pandemic recovery anomalies.

The temporal baseline dynamic inverted completely. In the 431×125 environment, the *Gap*

1 specification drastically outperformed the *Means* strategy across the leading algorithms. In highly sparse data sets, multi-year averaging mathematically fabricates artificial linkages – if an intermittent transaction occurred two years ago but ceased entirely last year, averaging forcefully drags a phantom transaction into the baseline. Thus, in high-resolution, the immediate historical structure (*Gap 1*) provides the truest proxy for the current technological state, maintaining matrix sparsity and preventing mathematical contamination.

The catastrophic failure of the standard RAS algorithm during the 2021 estimation merits specific attention. The post-COVID economic recovery induced extreme exogenous shocks to specific final demand constraints, most notably violent fluctuations in gross capital formation, trade disruptions, and inventory swings. Standard RAS, which lacks native mechanisms to reconcile conflicting mixed signs without severe ad-hoc modifications, entered oscillatory divergence. This illustrates that as matrices grow more disaggregated and real-world shocks intensify, older biproportional tools simply lack the mathematical elasticity required to force convergence without shattering the underlying I-O structure.

Beyond pure accuracy, the computational dynamics of estimating massive sparse matrices heavily favor biproportional methods over nonlinear optimizations. While INSD matched GRAS in producing near-perfect correlation coefficients, its reliance on heavy interior-point numerical solvers (`Ipopt`) results in substantially larger computational overhead. Given that GRAS relies on computationally parsimonious iterative multiplicative scaling, it effectively matches the structural fidelity of top-tier optimization routines while circumventing their steep processing costs. This makes GRAS an inherently more practical engine for executing high-frequency updates on massive regional datasets.

5 Discussion and Conclusion

This comprehensive evaluation of matrix updating procedures across Portuguese SUTs spanning 2010 to 2023 highlights severe operational limits and establishes optimal protocols for regional economic modeling. Our findings reveal that calculating multi-year averages minimizes discrepancies in aggregated matrices by leveraging structural memory, utilizing a single-year gap substantially improves accuracy in high-resolution models.

For macro-aggregated matrices (88×80), sectors are inherently dense, and multi-year averaging (*Means*) serves as an optimal mechanism to filter out transient economic noise. However, applied to high-resolution, sparse matrices (431×125), this exact same averaging technique drags outdated structural zeroes into the current projection. In these granular environments, leveraging exclusively the immediate past (*Gap 1*) ensures rapid *Shock Adaptation*, allowing the model to quickly internalize supply-chain shifts and localized volatility without polluting the coefficient structure.

Algorithmically, this study confirms a harsh stratification. Distance-minimizing tools (SPAD, SPSD) and weighted variants (WSD) lack the flexibility required to navigate modern macroeconomic shocks, yielding disproportionate errors. Consequently, for practitioners needing to maintain high-fidelity, high-resolution regional I-O frameworks (such as PReMMIA), the optimal strategy relies on the Generalized RAS (GRAS) method paired with a *Gap 1* baseline. This combination offers unparalleled computational stability, natively handles mixed-sign constraints without subjective interventions, and ensures the highest degree of structural fidelity even when standard accounts lag reality.

Acknowledgments

This work has been funded by national funds through FCT – *Fundação para a Ciência e a Tecnologia, I.P.* –, through project *UID/05037/2025* with *DOI* *10.54499/UID/05037/2025*, *2023.02227.BDANA* with *DOI* *10.54499/2023.02227.BDANA*, and and Integrated Project RN21 – Innovation in the Natural Resin Sector to Strengthen the National Bioeconomy, co-financed by the Environmental Fund through Component 12 –Promotion of the Sustainable Bioeconomy – Investment TC-C12-i01 – Sustainable Bioeconomy – Notice No. *01/C12-i01/2021* and No. *02/C12-i01/2021*, of the European funds allocated to Portugal by the Recovery and Resilience Mechanism (RRP), under the European Union's (EU) Recovery and Resilience Facility (RRF), as part of Next Generation EU, for the period 2021–2026.

References

Ferreira, J.-P., Cruz, L., Raposo, V., Rebolho, J. M., Lopes, L., Teotónio, C., Barata, E., and Sargento, A. (2025). Role of distinct export strategies in regional growth. A multi-regional input-output approach (PReMMIA). *European Planning Studies*, 33(4):601–621.

Huang, W., Kobayashi, S., and Tanji, H. (2008). Updating an Input-Output Matrix With Sign-Preservation: Some Improved Objective Functions and Their Solutions. *Economic Systems Research*, 20(1):111–123.

Jackson, R. W. and Murray, A. T. (2004). Alternative input-output matrix updating formulations. *Economic Systems Research*, 16(2):135–148.

Junius, T. and Oosterhaven, J. (2003). The Solution of Updating or Regionalizing a Matrix with both Positive and Negative Entries. *Economic Systems Research*, 15(1):87–96.

Lahr, M. L. (1998). A Strategy for Producing Hybrid Regional Input-Output Tables. In *12th International Conference on Input-Output Techniques*, New York. International Input-Output Association.

Lenzen, M., Wood, R., and Gallego, B. (2007). Some Comments on the GRAS Method. *Economic Systems Research*, 19(4):461–465.

Leontief, W. W. (1936). Quantitative Input and Output Relations in the Economic Systems of the United States. *The Review of Economics and Statistics*, 18(3):105.

Möhr, M., Crown, W. H., and Polenske, K. R. (1987). A Linear Programming Approach to Solving Infeasible RAS Problems. *Journal of Regional Science*, 27(4):587–603.

Révész, T. (2023). A Not Sign-Preserving Iteration Algorithm for the 'Improved Normalized Squared Differences' Matrix Adjustment Model. *Central European Journal of Operations Research*, 31:49–71.

Stone, R. (1961). *Input-Output and National Accounts*. Organisation for Economic Co-operation and Development.

Temurshoev, U., Webb, C., and Yamano, N. (2011). Projection of Supply and Use Tables: Methods and Their Empirical Assessment. *Economic Systems Research*, 23(1):91–123.

Wächter, A. and Biegler, L. T. (2006). On the Implementation of an Interior-Point Filter Line-Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming*, 106(1):25–57.