

Closing the Street-Level Discretion Gap: A Bureaucratic Calibration Framework for Algorithmic Decision Support in Means-Tested Benefit Administration

Nico Thompson

Professor

Erasmus University Rotterdam

Burgemeester Oudlaan 50, 3062 PA Rotterdam, Netherlands

Drew Evans

Associate Professor

University of Toronto

78 Queen's Park, Toronto, Ontario M5S 2C5, Canada

Riley Smith

PhD

Australian National University

Acton Campus, Chifley Business School, Kingsley Street, Acton ACT 2601, Australia

Abstract. Algorithmic decision-support systems (ADSS) increasingly mediate frontline caseworker determinations in means-tested public benefit programs, yet institutional frameworks governing street-level discretion calibration remain underdeveloped. This study examines the structural misalignment between automated eligibility scoring and human administrative judgment in welfare administration across three OECD jurisdictions. Drawing on a mixed-methods design combining administrative microdata analysis ($n = 84,630$ casefiles) and semi-structured interviews with 62 frontline administrators, we identify systematic discretion compression—a condition in which ADSS output constrains caseworker agency below legislatively mandated thresholds. Our findings demonstrate that procedural legitimacy deficits arise when algorithmic outputs are treated as binding rather than advisory. We propose a Bureaucratic Calibration Framework (BCF) operationalizing tiered override protocols, auditability benchmarks, and discretion-preservation mandates, offering actionable governance pathways for public administrators navigating automation-accountability tensions.

Keywords: algorithmic decision-support systems, street-level bureaucracy, administrative discretion, means-tested benefit administration, procedural legitimacy, automation-accountability gap, frontline public administration, welfare state governance, bureaucratic calibration

Introduction

The integration of algorithmic decision-support systems into public benefit administration represents one of the most consequential—and least institutionally resolved—governance transformations of the contemporary welfare state. As of 2025, more than forty national governments across the OECD have deployed some form of automated eligibility

determination or risk-scoring infrastructure within their social protection architectures, ranging from predictive analytics in unemployment insurance to machine-learning-assisted fraud detection in housing benefit programs. While proponents emphasize gains in processing efficiency, cost containment, and inter-rater consistency, a growing body of evidence from administrative law, public management, and sociotechnical systems research indicates that the unreflective embedding of algorithmic outputs into casework workflows produces measurable distortions in the exercise of street-level discretion—a foundational principle of Lipskian bureaucratic theory that remains constitutionally and legislatively protected in most democratic jurisdictions. The concept of street-level discretion, first systematized by Lipsky (1980) and subsequently elaborated through decades of comparative public administration scholarship, holds that frontline administrators must retain meaningful decisional latitude to contextualize rule application to individual circumstance. This latitude is not merely a practical accommodation to case complexity; it constitutes the primary mechanism through which administrative systems achieve substantive, as opposed to merely formal, procedural justice. When algorithmic decision-support systems are operationalized in ways that effectively supplant rather than inform this discretion—whether through interface design, organizational pressure, or inadequate override protocols—a condition we term discretion compression emerges. Discretion compression carries profound implications for beneficiary outcomes, institutional legitimacy, and the rule-of-law obligations of public agencies. Despite extensive normative commentary on the ethics of algorithmic governance, the public administration literature lacks an operationalizable, empirically grounded framework for diagnosing and correcting discretion compression at the institutional level. Existing contributions tend to address either the macro-level policy design of algorithmic accountability regimes or the micro-level behavioral responses of individual caseworkers, leaving a critical meso-level institutional gap—specifically, the organizational protocols, supervisory architectures, and auditability standards that determine whether algorithmic tools function as decision aids or de facto decision-makers. This paper addresses that gap directly. We proceed as follows: Section 2 reviews the theoretical literature on street-level bureaucracy, algorithmic governance, and procedural legitimacy. Section 3 describes our mixed-methods research design and the three comparative case jurisdictions. Section 4 presents quantitative findings on discretion compression incidence and its correlation with caseworker override rates. Section 5 reports qualitative findings from administrator interviews concerning perceived autonomy and institutional accountability norms. Section 6 introduces the Bureaucratic Calibration Framework and details its four operational components. Section 7 discusses policy implications and governance recommendations for public administrators and legislative oversight bodies. Section 8 concludes with directions for future comparative research.

This is a preliminary version. To read the full version of the article, please purchase a subscription.

References

1. Toni, D., Vangel, D., & Ljuben, K. (2021). (De) Localization of Arbitration-Online v. Offline Arbitration-E-Arbitration in Time of Covid 19. *Iustinianus Primus L. Rev.*, 12, 1.
2. Toshevska-Trpchevska, K., Lei, W. C., Kocev, L., & Tonovska, J. (2024). REGIONAL TRADE AGREEMENTS: TOOL TO PROMOTE CLIMATE-SMART AND SUSTAINABLE TRADE. *InterEULawEast: Journal for the international and european law, economics and market integrations*, 11(2), 217-244.