Institutiones Administrationis

Journal of Administrative Sciences

Vol. 6 (2026)
No. 1

ISSN 2786–1929

SZÉCHENYI EGYETEM
UNIVERSITY OF GYŐR

Institutiones Administrationis

Journal of Administrative Sciences
Vol. 6 (2026) No. 1

Published by Széchenyi István University
UNIVERSITAS-Győr Nonprofit Ltd.
Hungary-9026 Győr, Egyetem tér 1.

Responsible for publishing: Eszter Tóth

Layout editor: Éva Harkai

ISSN (online) 2786–1929

Contents

Articles

Articles

Institutiones Administrationis – Journal of Administrative Sciences Vol. 6 (2026) No. 1, 6–23
DOI: 10.54201/iajas.207

Regulating the Regulators

Digital Platforms as Quasi-Administrative Authorities in China's Food Safety Governance

Yi Lu*

* Senior Lecturer of Law, School of Transnational Law, Peking University, Shenzhen, Guangdong Province, China. E-mail: yi.lu@pku.edu.cn

Abstract

China's 2026 Provisions on livestream e-commerce food safety vest digital platforms with core regulatory functions that are functionally equivalent to those of administrative agencies. These functions include licensing, inspection, standard-setting, and sanctioning. Platforms now screen vendor credentials, deploy algorithmic surveillance, and impose sanctions up to permanent blacklisting. These powers mirror core administrative functions, yet Chinese law classifies them as private "corporate responsibility." This classification displaces core administrative law constraints on platform power. Duties of reason-giving, procedural fairness, and the availability of judicial review are thereby excluded. This article exposes the accountability vacuum that results. Comparison with the EU Digital Services Act demonstrates that procedural safeguards need not compromise regulatory efficiency. This article proposes four interventions that establish minimum procedural safeguards for platform-mediated governance, including mandatory internal review, a duty to provide reasons, bridging mechanisms for dispute resolution, and technological due process embedded in platform architecture.

Keywords

platform regulation, food safety, privatisation of regulation, administrative law, technological due process, algorithmic accountability

1 Introduction

On 16 January 2026, China's State Administration for Market Regulation (SAMR) unveiled the Provisions on Supervision and Administration of Livestream E-Commerce Operators' Implementation of Food Safety Responsibilities (hereinafter: "2026 Provisions"), set to take effect on 20 March 2026 (SAMR, 2025). The regulation arrives as a direct response to the National People's Congress Standing Committee's October 2025 enforcement inspection, which concluded that "new-format food business risks are emerging continuously" and recommended sector-specific rules for livestream commerce (Standing Committee of the National People's Congress, 2025). SAMR's Food Safety Director framed the instrument's ambition in direct terms, stating that the platform must manage both "the goods in the livestream room" and

 Published online: 20 April 2026

"the persons in it" (Zuo, 2026). What distinguishes the 2026 Provisions is that they go beyond requiring platforms to refrain from facilitating food safety violations, conscripting platforms into the active business of regulation itself. Platforms must now screen vendor credentials, appoint dedicated food safety personnel, deploy algorithmic surveillance systems, and impose sanctions ranging from traffic restrictions to permanent blacklisting. Platforms are thereby called upon to perform functions historically regarded as the exclusive province of the administrative state, including licensing, inspection, standard-setting, and sanctioning.

This regulatory architecture raises a question that transcends Chinese food law: when a digital platform exercises such functions, is it discharging a contractual compliance obligation, or wielding delegated public power? If the latter, the procedural apparatus of administrative law, including hearing rights, reason-giving duties, and judicial review, ought to apply. If the former, affected vendors are left with only the thin protections of contract law. The question of how to maintain procedural discipline when regulatory authority migrates to private intermediaries is one of the enduring puzzles of modern administrative science, and China's 2026 Provisions offer an unusually advanced case study through which to examine it.

The article proceeds in six sections. Section 2 traces the institutional evolution from state monopoly to platform-mediated food safety governance. Section 3 argues that the 2026 Provisions create administrative power operating outside administrative law. The analysis examines the functional taxonomy of platform obligations, the classification problem under Chinese administrative law doctrine, and the resulting accountability vacuum. Section 4 introduces the EU Digital Services Act as a comparative reference, and it demonstrates that procedural safeguards are institutionally feasible and explores the factors that account for the divergence between Chinese and European approaches. Section 5 proposes minimum procedural safeguards compatible with digital-age regulatory efficiency, built on the distinction between algorithmic detection and administrative disposition. Finally, section 6 concludes by situating the Chinese experience within the broader global trajectory of platform-mediated governance.

2 Institutional Evolution: From State Monopoly to Platform-Mediated Governance

The doctrinal questions examined here arise from a specific regulatory trajectory. This section traces the institutional evolution of China's food safety governance from state-centered command-and-control to platform-mediated enforcement.

2.1 The Command-and-Control Baseline

For the first three decades of China's post-reform food safety regime, governance rested on a model of direct state intervention. Government agencies, principally the former State Food and Drug Administration and its local counterparts, held a near-monopoly on the core regulatory functions of licensing, inspection, standard-setting, and enforcement. The original Food Safety Law of 2009 (Standing Committee of the National People's Congress, 2009) codified this model, concentrating authority in a tripartite structure: the health administration department assumed overall coordination responsibility, while quality supervision, industry and commerce, and food and drug administration departments exercised segmented jurisdiction over production, distribution, and catering services respectively (Food Safety Law, arts. 4–6). The approach was classically command-and-control, with the state setting the rules, conducting the inspections, and imposing the penalties.

This model, however, suffered from structural deficiencies well documented in the Chinese regulatory literature (Li et al., 2014; Li, 2015; Lian & Cheng, 2014). Jurisdictional fragmentation among multiple agencies created gaps in coverage and inconsistencies in enforcement. Inspection capacity was chronically insufficient relative to the scale and geographical dispersion of China's food production system, which at its peak encompassed millions of food production, processing, and distribution enterprises. The regime was overwhelmingly reactive: inspections were typically triggered by consumer complaints or periodic campaigns rather than by continuous monitoring. Moreover, the penalties, while nominally severe, were inconsistently applied, generating what Chinese scholars have described as a pattern of "heavy legislation, light enforcement" (see, e.g., Wu et al., 2016).

2.2 The 2015 Revision and the Birth of Platform Liability

The comprehensive revision of the Food Safety Law in 2015 (Standing Committee of the National People's Congress, 2015) marked a turning point. Article 62 fundamentally shifted the regime by introducing, for the first time, a statutory basis for imposing food safety obligations on third-party online platforms (Food Safety Law, art. 62). The provision required platforms operating online food transaction services to verify the real name and food business licenses of vendors using their services, to establish dedicated channels for reporting food safety violations, and to cease providing services to vendors discovered to be operating without valid licenses. Where a platform failed to discharge these obligations, and the failure resulted in harm to consumers, the platform bore joint and several liability with the offending vendor (Food Safety Law, art. 131).

Article 62 was a significant conceptual departure. It acknowledged that the state alone could not regulate a food market that was rapidly migrating online (Xie et al., 2018). Yet the provision was also cautiously drafted. It imposed on platforms a duty to verify and a duty to report. Still, it stopped short of requiring platforms to conduct the kind of proactive, continuous surveillance that would characterize the later regulatory instruments. The platform, under this framework, was cast as an auxiliary to the state, a gatekeeper whose role was to channel information to the regulatory authorities, rather than as a regulatory agent in its own right (Hu, 2018). This auxiliary role, however, would prove to be a transitional arrangement. Subsequent regulatory instruments would progressively transform platforms from passive gatekeepers into active regulators.

2.3 Regulatory Escalation (2019–2025)

The period from 2019 to 2025 witnessed a rapid and sustained expansion of platform food safety obligations, driven by a combination of legislative reform, administrative rulemaking, and political pressure. The E-Commerce Law (Standing Committee of the National People's Congress, 2018), which came into force on 1 January 2019, reinforced the platform's gatekeeper role and established a general framework of platform liability for defective products and services sold through its infrastructure (E-Commerce Law, arts. 38, 83). In 2022, SAMR promulgated the Provisions on the Implementation of the Principal Responsibility System for Food Safety in Food Production and Business Enterprises, which introduced two mandatory positions, the Food Safety Director and the Food Safety Officer, and required enterprises to conduct daily risk assessments and maintain compliance logs through institutionalized "daily control, weekly investigation, and monthly coordination" mechanisms (State Administration for Market Regulation, 2022). Although these provisions applied to food businesses generally rather than to

platforms specifically, their institutional logic, the internalization of regulatory functions within the regulated entity, would prove foundational for what was to follow.

These regulatory developments, however, were largely designed for conventional e-commerce. In parallel, the growth of livestream commerce as a food distribution channel introduced regulatory challenges that the existing framework was not designed to address. Livestream food sales differ from conventional e-commerce in several respects that complicate regulatory oversight. The transaction is mediated by a charismatic host whose persuasive authority may exceed the consumer's capacity for independent evaluation. Product claims are made orally and ephemerally, rendering *ex post* verification difficult. And the supply chain may involve actors, MCN (Multi-Channel Network) agencies, freelance hosts, cross-platform aggregators, whose legal status and regulatory responsibilities are poorly defined. These structural vulnerabilities were not merely theoretical. High-profile enforcement actions, including cases involving celebrity endorsements and livestream hosts with tens of millions of followers, exposed significant food safety violations that existing regulatory tools had failed to prevent. Against this backdrop, the regulatory response accelerated. By October 2025, the National People's Congress Standing Committee's enforcement inspection of the Food Safety Law had concluded that the regulatory tools available for online food commerce were insufficient (Standing Committee of the National People's Congress, 2025). The inspection report recommended that sector-specific regulations be developed for online food sales, meal delivery services, livestream commerce, and chain restaurant operations, and that regulatory methods be upgraded from traditional to "smart" supervision capable of achieving "online-offline integrated governance."

2.4 The 2026 Provisions: a Regulatory Inflection Point

The Provisions on the Supervision and Administration of the Implementation of Principal Responsibility for Food Safety by Livestream E-commerce Operators, promulgated on 28 December 2025 and effective from 20 March 2026, represent the culmination of this evolutionary trajectory (State Administration for Market Regulation, 2025). They are, in both scope and specificity, qualitatively different from the earlier instruments. The 2026 Provisions impose upon platforms a suite of obligations that can be grouped into six categories, supplemented by an evidentiary rule. Each merits a brief examination.

First, platforms must appoint food safety management personnel, including a Food Safety Director and Food Safety Officers commensurate with the platform's transaction scale and risk profile, and conduct food safety training for all hosts engaged in food-related livestreaming for the first time, with a requirement to verify host identity and maintain training records (art. 7–8). Second, platforms must verify the business qualifications, food business licenses, and registration information of all livestream operators and establish compliance files updated on a biannual cycle (art. 6). Third, platforms must prepare and implement food safety risk control checklists that address vendor qualifications, prohibited food screening, and advertising compliance (art. 9). Fourth, platforms must establish an "intelligent monitoring – investigation dispatch – rapid response" mechanism utilizing technological monitoring and real-time surveillance (art. 9). Fifth, upon detecting a food safety violation, platforms must immediately halt the activity and report it to the regulatory authorities, and must impose graduated sanctions according to the severity of the breach (art. 12). These sanctions include warnings, traffic restrictions, livestream suspension, account closure, or blacklisting. Sixth, thirteen categories of food products are subject to an absolute prohibition on livestream sale (art. 18). These include food produced from non-food raw materials, food with added toxic or harmful substances, expired food, and food

from quarantine-failed livestock. Finally, as a complementary evidentiary rule, the Provisions expressly provide that technical monitoring records generated by the platform's surveillance systems are admissible as electronic evidence in administrative penalty proceedings (art. 26).

Taken individually, each of these obligations might be characterized as a discrete compliance requirement. Taken together, they constitute something more. They create a regulatory architecture in which the platform assumes the role of frontline inspector, standard-setter, evidence-gatherer, and enforcer for an entire segment of the food market. It is this aggregate transformation, rather than any single provision, that gives rise to the doctrinal questions examined in section 3.

3 Administrative Power in Private-Law Clothing

The previous section traced the institutional evolution that transformed platforms from passive gatekeepers into active regulatory agents. Section 3 subjects that transformation to doctrinal scrutiny. The central argument is that the 2026 Provisions vest platforms with powers functionally equivalent to those of administrative agencies, yet classify those powers as private contractual obligations, thereby exempting them from the procedural constraints that ordinarily attach to public authority.

3.1 A Functional Taxonomy of Platform Obligations

Administrative law scholarship has long recognized that the characterization of a function as "public" or "private" cannot rest on formal labels alone; it must be informed by an assessment of what the actor in question actually does. This functional approach, associated in the common law tradition with the work of scholars (e.g., Mashaw, 1983; Freeman, 2000; Metzger, 2003) and in the continental tradition with the doctrine of "functional equivalence" (Zweigert & Kötz, 1998), provides the analytical starting point for evaluating the 2026 Provisions. Applying this functional lens to the 2026 Provisions reveals the extent of the regulatory transformation. Recent scholarship has extended this functional analysis to digital platforms, characterizing them as "new governors" exercising quasi-public authority over online spaces (Klonick, 2018) and proposing "platform administrative law" as a framework for analyzing how administrative law principles might apply to platform governance (Schramm, 2024).

Consider the four classical functions of an administrative regulatory agency that have long defined the administrative state. These are controlling access to a regulated market, monitoring compliance with regulatory standards, setting or specifying the standards against which compliance is measured, and imposing sanctions for non-compliance. Under the 2026 Provisions, platforms perform each of these functions. The credential verification and biannual file-updating requirements are functionally equivalent to the licensing and renewal procedures administered by local market supervision bureaus. The "intelligent monitoring" obligation goes further than traditional government inspection. Whereas SAMR relies primarily on random inspections conducted under the "double-random, one-disclosure" method, a sampling approach that, by design, reaches only a fraction of regulated entities (State Council of the People's Republic of China, 2019), the platform's algorithmic surveillance operates continuously across the entirety of its vendor base. The risk control checklists that platforms are required to draft and implement serve a standard-specifying function, translating the general prohibitions of the Food Safety Law into operational criteria tailored to the livestream environment. And the

graduated sanction regime, ranging from traffic restriction to permanent blacklisting, mirrors the penalty ladder set out in the Administrative Penalty Law (Standing Committee of the National People's Congress, 2021) (art. 9), with blacklisting in particular potentially carrying economic consequences comparable to, or exceeding, the revocation of a business license.

The functional equivalence documented above, however, understates the transformation. In at least one critical respect, the platform's regulatory capacity surpasses that of the state. Government food safety inspectors are constrained by personnel shortages, jurisdictional boundaries, and the physical logistics of on-site inspection. Platform algorithms, by contrast, can monitor millions of transactions in real time, flagging anomalies in product descriptions, pricing patterns, and vendor behavior with a speed and comprehensiveness that no government agency can match. The 2025 NPC enforcement inspection report acknowledged this asymmetry implicitly when it recommended that regulatory authorities leverage "big data and artificial intelligence technologies" and transition from "traditional to smart supervision," in effect urging the state to borrow the platform's own regulatory infrastructure (Standing Committee of the NPC, 2025). Table 1 below synthesizes how the 2026 Provisions vest platforms with functions that are functionally equivalent to core regulatory powers.

Table 1.
Functional Taxonomy of Platform Obligations Under China's 2026 Provisions

Administrative Function	State Agency Mechanism	Platform Equivalent (2026 Provisions)	Functional Assessment
Market Access Control	Business license issuance and renewal by local Market Supervision Bureaus (MSBs)	• Credential verification, • Biannual file updating, and • Denial of platform access to unlicensed vendors (Art. 6)	Functionally equivalent to licensing
Compliance Monitoring	"Double-random, one-disclosure" spot inspections; periodic sampling campaigns	"Intelligent monitoring – investigation dispatch – rapid response" algorithmic surveillance across entire vendor base (Arts. 9–10)	Exceeds state capacity: continuous, comprehensive coverage
Standard Specification	National food safety standards (GB standards) promulgated by the National Health Commission	Platform-drafted risk control checklists translating statutory prohibitions into operational criteria for livestream environment (Art. 9)	Quasi-legislative: operationalizes abstract norms
Sanctioning	Administrative penalties under the Administrative Penalty Law: fines, license revocation, business closure orders	Graduated sanctions: warnings → traffic restriction → livestream suspension → account closure → permanent blacklisting (Art. 12)	Economically equivalent to license revocation

Source: Author's analysis based on the 2026 Provisions.

This article does not claim that platforms have acquired formal administrative-subject status under Chinese law, which is the question to be addressed in the next subsection. Rather, the functional analysis demonstrates that platforms exercise powers functionally equivalent to those of administrative agencies, and argues that this functional equivalence should trigger corresponding procedural constraints.

3.2 The Classification Problem in Chinese Administrative Law

The doctrinal question that follows is how Chinese administrative law classifies and constrains such exercises of public power by private actors. Chinese administrative law provides three doctrinal categories for classifying the exercise of public-interest functions by non-state actors.

The first is statutory authorization. Under Articles 2 and 26 of the Administrative Litigation Law (Standing Committee of the National People's Congress, 2017), an organization that has been authorized by a law, regulation, or rule to exercise administrative functions acquires the status of an administrative subject. Its acts are subject to administrative litigation, and the procedural requirements of the Administrative Penalty Law and the Administrative Compulsion Law apply in full. The second is administrative entrustment. Unlike statutory authorization, the delegate does not acquire independent administrative subject status. Under this arrangement, an administrative agency delegates the exercise of a specific power to another entity, which acts on the agency's behalf and under its supervision. The delegating agency retains legal responsibility for the entrusted acts, and affected parties may seek judicial review of the delegating agency's conduct (Administrative Litigation Law, art. 26). The third is a private obligation. Here, the analytical framework shifts entirely. The actor performs a function required by law, but the relationship between the actor and the persons affected by its conduct is governed by civil or commercial law. Disputes are resolved through civil litigation or contractual arbitration, and the procedural protections of administrative law do not apply.

Against this doctrinal backdrop, the classification chosen by the 2026 Provisions becomes significant. The regulatory framework places platform food safety enforcement squarely in the third category, characterizing the platform's obligations as the discharge of "principal responsibility". This is a term of art in Chinese regulatory discourse that connotes the regulated entity's own duty to ensure compliance, as distinct from the state's supervisory authority (SAMR, 2022). This classification is not an inadvertent lacuna but a deliberate institutional choice. By framing platform enforcement as a private obligation, the regulatory framework avoids the legal and procedural consequences that would attach if the same functions were formally characterized as delegated administrative power. The platform is empowered to screen, surveil, and sanction, but it is not burdened with the administrative law duties that ordinarily accompany such powers.

3.3 The Accountability Vacuum

Having established that Chinese administrative law doctrine classifies platform enforcement as a private contractual exercise rather than administrative action, this section examines the practical consequences of that classification. The practical consequences become most visible when the system produces errors. In a system that processes millions of transactions through automated surveillance, errors are inevitable. Available evidence, though not comprehensive, suggests a recurring pattern: automated enforcement systems impose significant commercial consequences on vendors based on algorithmic determinations that are neither explained nor effectively contestable.

Three cases reported to the E-Litigation Platform (*Diansubao*) illustrate the problem. In the first, a food vendor on *Douyin* (the Chinese version of TikTok) received a "false efficacy claim" determination from the platform's automated monitoring system, resulting in the restriction of his product-sharing functionality. The vendor contested the determination as unreasonable, but the platform's internal appeals process offered no meaningful avenue for review (China

E-commerce Research Centre, 2025). In the second, a vendor on *Douyin* selling decorative stickers was incorrectly flagged for selling "counterfeit lighters" based solely on a customer-uploaded image. The vendor's shop was closed, funds frozen, and repeated appeals rejected. The vendor characterized this conduct as "entrapment enforcement" (China E-commerce Research Centre, 2025). In the third, a food vendor on *Kuaishou* (a short-video platform similar to TikTok) had store credit points deducted following a buyer's malicious review, yet could locate no channel through which to contest the determination (China E-commerce Research Centre, 2024).

These are not isolated incidents. According to data compiled by the China E-commerce Research Centre, during the 2025 "618" Shopping Festival promotional period alone, *Douyin*'s e-commerce platform received 191 merchant complaints, with food products accounting for 17.5% of the affected categories. A 2024 report by the Peking University E-Commerce Law Research Centre observed that platforms now exercise "quasi-legislative, quasi-executive, and quasi-judicial" powers over merchants, yet without corresponding procedural accountability (Peking University E-Commerce Law Research Centre, 2024). These cases illuminate three specific procedural deficits that are structural rather than incidental to the regulatory design.

The first is the absence of hearing rights. Under the Administrative Penalty Law, as revised in 2021, administrative agencies are required to notify the affected party of the right to request a hearing before imposing penalties that involve significant consequences. These include relatively large fines, confiscation of relatively large amounts of illegal gains or property, lowering of qualification levels or revocation of licenses, ordering the suspension of production or business, ordering closure, restricting engagement in certain business operations, or other relatively heavy administrative penalties (art. 63). Blacklisting by a dominant platform can be economically equivalent to license revocation: for a vendor whose entire distribution channel runs through a single platform, permanent exclusion from that platform is, for all practical purposes, a death sentence for the business. Yet because the platform's action is classified as the exercise of a contractual right rather than an administrative penalty, no hearing obligation arises.

The second is the absence of a duty to state reasons. Administrative agencies, when taking decisions that adversely affect private parties, are required to set forth the factual basis and legal grounds for their action. This duty serves both an informational function, enabling the affected party to understand and respond to the case against it, and a disciplining function, forcing the decision-maker to articulate a rational basis for its action, thereby constraining arbitrariness. The automated notifications generated by platform monitoring systems satisfy neither function. They communicate the conclusion (restriction or suspension) without disclosing the reasoning, and they provide no basis on which the affected vendor can formulate a meaningful response.

The third, and most consequential, is the remedial vacuum. A vendor aggrieved by a platform enforcement decision faces a choice between two unsatisfactory options. It may file a civil lawsuit, in which case the court will treat the dispute as a contractual matter, evaluating the platform's conduct against the terms of the platform service agreement, terms that the platform itself has drafted and that typically confer broad discretion to take enforcement action. Alternatively, the vendor may attempt to file an administrative lawsuit, arguing that the platform's action constitutes the exercise of *de facto* administrative power. This path faces a threshold jurisdictional objection: under prevailing doctrine, the platform is not an "administrative subject", and its actions are therefore not amenable to administrative judicial review.

The result is a legal no-man's land. Vendors subject to platform enforcement are trapped in a jurisdictional gap where administrative law does not apply due to the platform's private status, and civil law provides no relief due to judicial deference to contractual autonomy.

Substantial economic harm inflicted through the exercise of quasi-public authority thus falls between the two pillars of the judicial system, leaving algorithmic power effectively unfettered and unaccountable.

3.4 The Algorithmic Dimension

The procedural deficits identified above are structural, but they are not static. The accountability problem is compounded by the role of algorithmic decision-making in the enforcement process. The use of algorithmic decision-making by platforms raises questions familiar to administrative law scholarship on automated governance (Coglianese & Lehr, 2017; Engstrom & Ho, 2020). The 2026 Provisions require platforms to deploy "intelligent monitoring" systems (art. 9), a term that, in practice, refers to machine-learning algorithms that scan product listings, livestream content, and transaction patterns for indicators of non-compliance. When such a system identifies a potential violation, it may trigger enforcement action, including traffic restriction, product delisting, or account suspension, without human intervention.

Algorithmic enforcement offers undeniable advantages in terms of speed and coverage. But it also introduces risks that the existing legal framework does not adequately address. Algorithms are trained on datasets that may contain biases or gaps, they apply probabilistic rather than deterministic logic, and their decision processes are, in most cases, opaque to the affected parties and, frequently, to the platform operators themselves (Pasquale, 2015). The 2026 Provisions' stipulation that "technical monitoring records may serve as evidence in administrative penalty proceedings" (art. 26) further elevates the stakes. This provision effectively transforms the platform's algorithm into a quasi-official evidence-gathering instrument whose outputs feed directly into government enforcement action, yet the algorithm itself is proprietary, its methodology is not subject to independent audit, and its error rate is not publicly disclosed. The evidentiary implication is significant. A vendor subjected to government penalties on the basis of platform-generated evidence has no procedural means of challenging the reliability of the algorithm that produced that evidence. In traditional administrative enforcement, the accuracy of the state's evidence-gathering methods is subject to scrutiny through the administrative litigation process. When the evidence-gathering function is outsourced to a private algorithm, this layer of accountability is lost.

Viewed through the lens of administrative law theory, what has emerged is a regulatory architecture in which public power is exercised through private instrumentalities, disciplinary consequences are imposed through automated systems, and the procedural safeguards that have traditionally mediated the relationship between regulatory authority and regulated subjects are absent. Rose-Ackerman's account of the regulatory state emphasizes that the legitimacy of delegated authority depends not only on the substantive rationality of regulatory outcomes but also on the procedural fairness of the processes by which those outcomes are reached (Rose-Ackerman, 2021). Approaching the question from a different but complementary angle, Mashaw's influential framework for evaluating administrative due process asks whether the decision-making process provides adequate accuracy, transparency, and opportunities for participation relative to the stakes involved (Mashaw, 1976; 1985). Measured against either standard, the platform enforcement regime established by the 2026 Provisions falls short, not because it is ineffective, but because its effectiveness is achieved by circumventing the procedural discipline that administrative law theory regards as constitutive of legitimate public authority.

4. Comparative Perspectives

The accountability deficits identified in the preceding analysis are not unique to the Chinese regulatory context. Other jurisdictions have confronted similar challenges in governing platform-mediated enforcement, and their responses offer instructive comparative reference points.

4.1 Functional Parallels

The European Union's Digital Services Act, Regulation (EU) 2022/2065 (hereinafter: "DSA"), which became fully applicable on 17 February 2024, provides a useful comparative reference point, not as a normative ideal but as an illustration of how a different legal tradition has addressed functionally similar problems. Like the Chinese 2026 Provisions, the DSA imposes on online platforms affirmative obligations to establish mechanisms for receiving notices of, and acting against, illegal content and products distributed through their services. The DSA's "notice-and-action" mechanism for illegal content (art. 16) is functionally analogous to the Chinese "intelligent monitoring – investigation dispatch – rapid response" framework: in both cases, the platform is required to receive or detect reports of illegality, evaluate them, and take enforcement action where appropriate. In both cases, the platform's enforcement decisions, including content removal, account restriction, or service termination, carry significant economic consequences for the affected parties.

A significant distinction, however, warrants acknowledgment. The DSA expressly disclaims any general obligation to monitor the information transmitted or stored, or to actively seek facts indicating illegal activity (art. 8), a principle rooted in fundamental rights protections (Frosio & Geiger, 2023). Platforms act upon receiving notice of potential illegality, but are not required to engage in proactive surveillance. The Chinese 2026 Provisions, by contrast, mandate platforms to establish "intelligent monitoring" systems capable of real-time, automated detection of non-compliance (arts. 9–10). This divergence in the triggering mechanism, reactive notice-and-action versus proactive algorithmic surveillance, carries significant implications for the procedural architecture that accompanies platform enforcement, as the following section demonstrates.

4.2 The Procedural Safeguards Divergence

The functional similarities documented above, however, mask a fundamental divergence in regulatory design. The critical institutional difference lies not in the substantive obligations imposed on platforms but in the procedural architecture that accompanies those obligations. The DSA constructs a three-layer system of procedural safeguards for parties affected by platform enforcement decisions (Bayer, 2022; Frosio & Geiger, 2023; Husovec, 2024).

At the first layer, Article 20 of the DSA requires platforms to provide an internal complaint-handling mechanism that is easily accessible, user-friendly, and capable of processing complaints in a timely, non-discriminatory, diligent, and non-arbitrary manner (DSA, art. 20 par. 3–4). Critically, the provision specifies that decisions on complaints must not be taken solely based on automated means: platforms must ensure that appropriately qualified staff are involved in the review process (art. 20 par. 6). This requirement directly addresses the opacity problem inherent in algorithmic enforcement by ensuring that automated decisions are subject to human oversight at the contestation stage (Bayer, 2022).

At the second layer, Article 21 establishes a framework for out-of-court dispute settlement through certified independent bodies. These bodies must possess the requisite expertise and

independence to resolve disputes impartially (par. 3 a–b), and their decisions are non-binding on both parties. Platforms are required to engage with the process in good faith, while affected parties retain the right to seek judicial review (par. 1–2). This mechanism provides an accessible, low-cost alternative to litigation, available to users free of charge or at a nominal fee (par. 5), that is specifically designed for the high-volume, relatively low-value disputes that characterize platform enforcement.

At the third layer, Article 21 itself expressly preserves the right of affected parties to seek judicial remedy against platform decisions in the courts in accordance with applicable law (par. 1(3)). This provision ensures that the internal and out-of-court mechanisms do not function as substitutes for judicial review but as supplements to it (Van Hoboken et al., 2023).

Against this procedural backdrop, the Chinese regulatory approach presents a striking contrast. The Chinese 2026 Provisions contain none of these three safeguards. The regulation addresses in meticulous detail the obligations that platforms owe to the state, including what they must monitor, how they must respond, and what records they must retain, but is essentially silent on the obligations that platforms owe to the vendors they discipline. There is no requirement for an internal complaint-handling mechanism, no provision for independent dispute resolution, and no express preservation of judicial remedy. The vendor's procedural rights, in short, are not merely underspecified. They are absent from the regulatory design. Table 2 provides a systematic comparison of the procedural safeguards available under each regime, revealing a stark accountability asymmetry.

Table 2.
Procedural Safeguards Comparison: EU DSA vs. China 2026 Provisions

Procedural Safeguard	EU Digital Services Act (Reg. 2022/2065)	China 2026 Provisions (SAMR)	Gap Assessment
Internal Complaint Mechanism	Art. 20: Mandatory internal complaint-handling; human review required; timely, non-discriminatory processing	No provision. Vendors lack any guaranteed channel for contesting platform enforcement decisions.	Critical gap
Independent Dispute Resolution	Art. 21: Out-of-court settlement through certified independent bodies; binding on platform, not on affected party	No provision. No independent body exists to review platform food safety enforcement actions.	Critical gap
Judicial Remedy	Art. 21: Express preservation of right to judicial review in courts of establishment or residence	De facto unavailable. Civil courts apply deferential contractual standard; administrative courts lack jurisdiction (platform is not an administrative subject).	Structural gap
Duty to State Reasons	Art. 17: Statement of reasons required for content moderation decisions, including specific rule and explanation of automated means	No provision. Automated notifications communicate conclusions without disclosing factual basis, standard applied, or reasoning.	Critical gap
Human Oversight of Algorithms	Art. 20(6): Complaint decisions may not be taken solely by automated means	No provision. "Intelligent monitoring" systems operate without mandated human-in-the-loop at any stage.	Critical gap

Note: ***Critical gap*** *= safeguard entirely absent;* ***Structural gap*** *= formally available but functionally inaccessible.*

Source: Author's compilation.

4.3 Understanding the Divergence

What accounts for this striking procedural asymmetry? The divergence reflects different institutional logics rooted in different administrative law traditions and political economies of platform governance. Three factors merit consideration.

The first factor is an efficiency-first regulatory philosophy combined with an outcome-oriented administrative culture. China's approach to platform food safety regulation responds to an urgent practical problem: the state's inspection capacity is structurally inadequate to supervise a food market that processes hundreds of millions of online transactions annually. Platform co-optation offers a mechanism for scaling regulatory capacity rapidly and at low marginal cost. From this perspective, procedural constraints introduce friction into the very mechanism the regulation is designed to create. Moreover, Chinese regulatory performance metrics are overwhelmingly output-oriented, focusing on the number of violations detected and the volume of non-compliant products removed, rather than process-oriented. Procedural quality does not feature as an independent criterion of regulatory success (Gao, 2015).

The second factor is the political economy of platform governance in post-2020 China. The comprehensive regulatory tightening campaign directed at major technology platforms, including antitrust enforcement, data governance regulation, and labor protection mandates, fundamentally altered the bargaining dynamic between regulators and platforms (Marco Colino, 2022). Platforms accepted expanded compliance obligations as part of maintaining their market position. In this environment, neither regulators nor platforms have strong incentives to advocate for vendor procedural rights.

The third factor lies in the distinct regulatory objects that the two regimes address. The DSA's procedural architecture was designed primarily for content moderation, a domain of inherent interpretive ambiguity that structurally justifies robust procedural safeguards. The Chinese 2026 Provisions, by contrast, address product safety involving material risks to human health, where regulatory standards are more determinate. In this domain, enforcement errors that fail to catch violations can cause physical harm, which may justify prioritizing speed over procedural formality. This distinction, combined with a deeply embedded institutional conviction that public health protection justifies subordinating individual procedural rights to collective safety outcomes, helps explain the divergent regulatory postures. The succession of food safety scandals that preceded the 2015 Food Safety Law revision, from melamine-tainted infant formula to recycled "gutter oil" (Pei et al., 2011), created a political environment in which any apparent softening of enforcement rigor carries significant legitimacy costs. The 2015 revision has been characterized as "the strictest food safety law ever," embodying the "four strictest" principles in standards, oversight, penalties, and accountability (Roberts & Lin, 2016).

Taken together, these factors constitute a coherent explanatory framework for the procedural divergence. The comparative analysis reveals that the delegation of regulatory functions to platforms is a global trend (Gorwa, 2019; Gorwa, 2024). The critical variable is not whether such delegation occurs. It does, and will intensify. The critical variable is whether it is accompanied by institutional mechanisms preserving minimum procedural fairness. The Chinese model's distinctive contribution lies precisely in its demonstration that efficiency gains from platform-mediated governance come at a measurable cost to procedural justice, a cost that, as the following section argues, need not be accepted as inevitable.

5. Procedural Proposals for Regulatory Reform

The doctrinal analysis in section 3 and the comparative study in section 4 have established both the necessity and feasibility of procedural safeguards for platform-mediated food safety governance. This section translates those insights into concrete institutional design.

5.1 The Conceptual Foundation: Detection versus Disposition

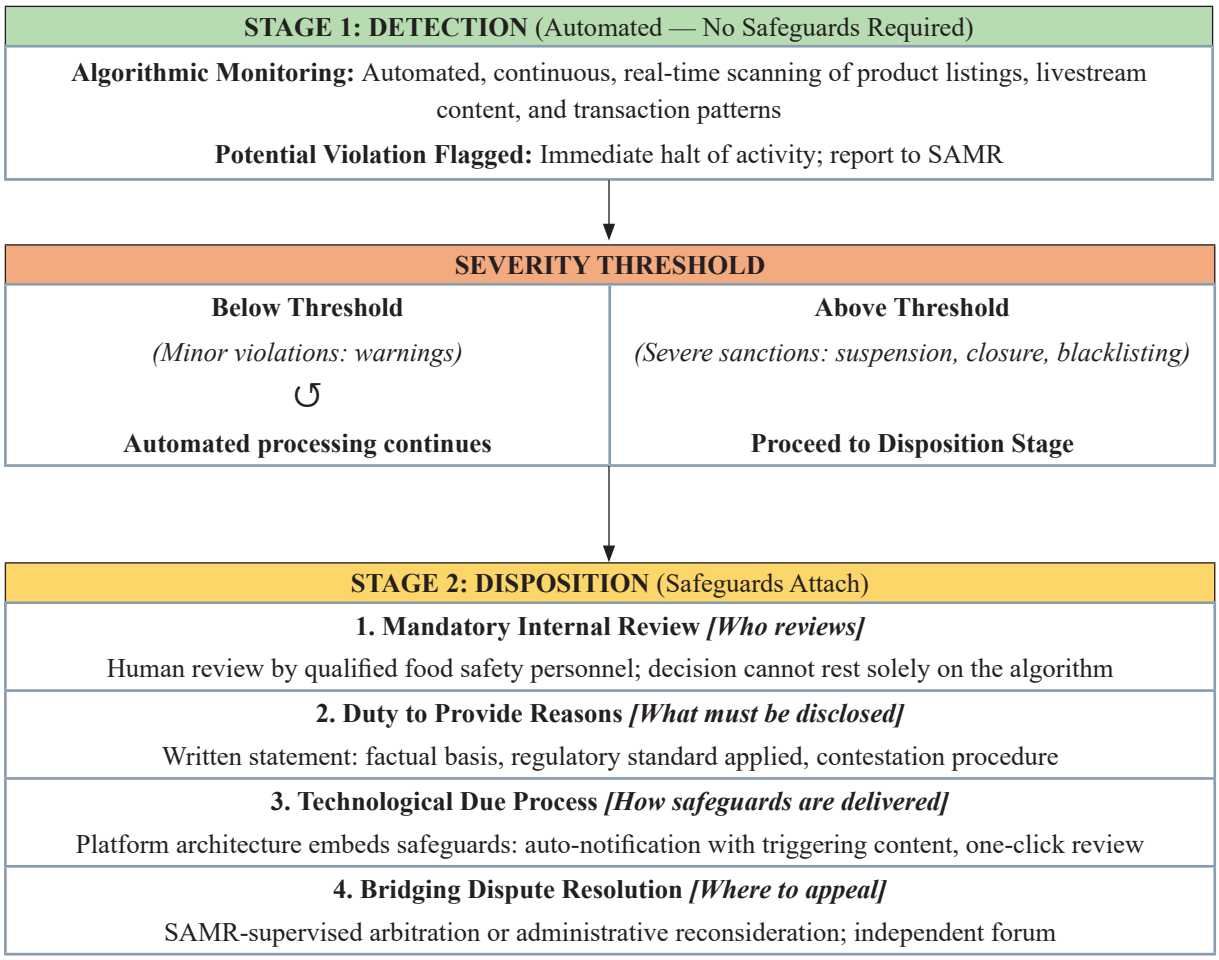
Procedural safeguards for platform-mediated governance are institutionally feasible, as the DSA demonstrates. However, the challenge lies in designing them for the Chinese context. The proposals in this section are guided by a single organizing principle. Minimum procedural safeguards should be calibrated to the operational realities of digital-age regulation. The objective is not to transpose the European model onto the Chinese regulatory landscape, but to identify the smallest set of procedural interventions that can meaningfully address the accountability vacuum identified in section 3 without materially impairing regulatory efficiency. Crucially, procedural safeguards need not operate at the detection stage. Requiring human review before every algorithmic flag would cripple the very efficiency that makes platform-mediated governance attractive. Rather, safeguards should attach at the disposition stage, where a detected anomaly translates into enforcement action imposing material economic harm. This distinction between detection (automated and instantaneous) and disposition (incorporating minimum procedural protections) is the key to reconciling efficiency with fairness. As scholars examining automated scoring systems have argued, due process protections must attach at the point where algorithmic outputs translate into consequential determinations affecting individual rights (Citron & Pasquale, 2014).

Although the 2026 Provisions outline a range of enforcement measures (art. 12), from warnings to permanent blacklisting, they do not distinguish between minor and severe sanctions with corresponding procedural consequences. This article proposes such a distinction. Minor sanctions should include warnings, function restrictions, and traffic restrictions. Severe sanctions should include livestream suspension, time-limited suspension, account closure, prohibition on re-registration, and blacklisting. In accordance with the proposed framework, procedural safeguards apply exclusively to the latter category, the economic consequences of which are commensurate with those of administrative license revocation.

5.2 The Regulatory Proposals: Four Procedural Interventions

Building on this detection-disposition distinction, four specific proposals follow. Each is designed to inject minimum procedural safeguards into platform enforcement without compromising the speed and scale that make algorithmic governance effective. Together, they address the three accountability deficits identified in section 3, including the absence of reason-giving, the unavailability of meaningful review, and the opacity of algorithmic determination. These are deficits that scholars have identified as endemic to algorithmic governance in the administrative state (Engstrom & Ho, 2020). Figure 1 below maps the four proposals onto the enforcement process. The framework preserves algorithmic efficiency at the detection stage while introducing procedural safeguards only at the disposition stage, when severe sanctions materialize.

Figure 1. Proposed "Detection–Disposition" Framework with Procedural Safeguards



Source: Author's illustration.

First, and most fundamental, is a mandatory internal review mechanism. Platforms should establish a dedicated channel through which vendors subject to enforcement actions exceeding a defined severity threshold, such as account suspension, blacklisting, or prolonged traffic restriction, may request human review by qualified personnel with food safety expertise. The mere existence of such a mechanism produces a disciplining effect on initial decision-making quality and a legitimizing effect on the system as a whole.

Second, complementing the review mechanism is a duty to provide reasons. When the threshold of severity is reached, platforms are required to provide affected vendors with a written statement. This statement should include the specific factual basis for the determination, namely the listing or transaction that triggered the decision. In addition, the statement must outline the regulatory standard that was applied and the procedure for contesting the decision. This requirement draws on the rationale underlying the Administrative Penalty Law (arts. 44, 45, 62). The cost is low, because algorithmic systems already record the triggering data points, while the benefit is substantial: reason-giving transforms opaque algorithmic outputs into contestable determinations.

Third, providing the delivery mechanism for the preceding safeguards is technological due process, which involves embedding procedural safeguards directly within platform algorithmic architecture (Citron, 2008). When a monitoring system triggers enforcement action meeting the severity threshold, it should automatically generate a structured notification identifying: (a) the

triggering content or transaction; (b) the regulatory standard applied; (c) the violation category; and (d) a one-click mechanism for initiating internal review. The algorithmic systems already record these data points as technical monitoring records admissible as electronic evidence (art. 26). The marginal cost of formatting them into intelligible notifications is negligible.

Fourth, addressing the jurisdictional gap that renders both civil and administrative review ineffective is a bridging mechanism for dispute resolution. The jurisdictional gap requires an institutional bridge. Vendors can neither obtain meaningful civil review under a deferential contractual standard nor access administrative review, which is jurisdictionally unavailable. Options include bringing disputes within existing administrative reconsideration mechanisms by designating SAMR as the reviewing authority, or establishing a SAMR-supervised arbitration panel modelled on China's online consumer dispute resolution mechanisms (Zheng, 2016).

Taken together, these four proposals constitute minimum procedural safeguards for platform-mediated enforcement. Crucially, none of them requires platforms to slow down algorithmic monitoring. Detection remains automated; enforcement remains swift. The proposals intervene only where swift enforcement produces contestable outcomes. In a system processing millions of enforcement decisions, even a modest rate of erroneous outcomes translates into substantial individual injustices. A procedural floor catching even a fraction of those errors produces a net gain in both legitimacy and accuracy.

6 Conclusion

China's 2026 Provisions on livestream e-commerce food safety mark a decisive shift in the global evolution of platform-mediated governance. By systematically vesting digital platforms with core regulatory functions, including market access verification, compliance monitoring, algorithmic surveillance, and graduated sanctioning, Chinese regulators have created a regulatory architecture that significantly expands the state's reach through private intermediaries. This article has argued that the resulting framework transforms platforms into quasi-administrative authorities exercising powers functionally equivalent to those of state agencies, yet operating entirely outside the procedural discipline of administrative law.

The accountability vacuum generated by this doctrinal misclassification is not merely a theoretical concern. The absence of hearing rights, the failure to provide reasons for enforcement decisions, and the remedial gap between civil and administrative adjudication produce concrete injustices for vendors subjected to algorithmic enforcement. As the comparative analysis has shown, the DSA demonstrates that procedural safeguards and regulatory efficiency are not, in this domain, a zero-sum trade-off. This article has proposed minimum procedural safeguards consisting of four interventions. These are mandatory internal review, a duty to provide reasons, bridging mechanisms for dispute resolution, and the embedding of procedural safeguards within platform architecture. The proposals are designed to restore procedural legitimacy without impairing the speed and scalability that make platform-mediated governance attractive to regulators.

The implications of this analysis extend beyond Chinese food law. The delegation of regulatory functions to digital platforms is a global phenomenon, observable across content moderation, financial services, healthcare, and an expanding range of other domains. For scholars and practitioners of administrative law, particularly those examining the privatization of public services in Europe, the Chinese experience offers both a *warning* and a *template*. The *warning* is that unconstrained platform-mediated enforcement erodes the procedural safeguards that have

historically legitimized the exercise of public authority. The *template* is that these safeguards can be preserved, and indeed must be preserved, if platform governance is to command the legitimacy that effective regulation requires. Whether China's regulatory authorities will move in this direction remains an open question. What is clear is that the institutional design choices made in this domain will shape not only the governance of food safety but the broader trajectory of platform regulation, in China and, increasingly, around the world.

References

Bayer, J. (2022). Procedural rights as safeguard for human rights in platform regulation. *Policy & Internet*, 14(4), 755–771. https://doi.org/10.1002/poi3.298

China E-commerce Research Centre [网经社电子商务研究中心]. (2024, November 5). *October 2024 China e-commerce platform merchant complaint data report* [*2024年10月中国电商平台商家投诉数据报告*]. Online: https://www.100ec.cn/zt/dsptts/

China E-commerce Research Centre [网经社电子商务研究中心]. (2025, April 7). *2025 Q1 China e-commerce platform merchant complaint data report* [2025年Q1中国电商平台商家投诉数据报告]. 163.com [网易]. Online: https://www.163.com/dy/article/JSIBCHD60514BOS2.html

Citron, D. K. (2008). Technological Due Process. *Washington University Law Review*, 85(6), 1249–1314. Online: https://tinyurl.com/4xnjnydr

Citron, D. K., & Pasquale, F. (2014). The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, 89(1), 1–33. Online: https://digitalcommons.law.uw.edu/wlr/vol89/iss1/2

Coglianese, C., & Lehr, D. (2017). Regulating by Robot: Administrative Decision-Making in the Machine-Learning Era. *Georgetown Law Journal*, 105(5), 1147–1223.

Engstrom, D. F., & Ho, D. E. (2020). Algorithmic Accountability in the Administrative State. *Yale Journal on Regulation*, 37(3), 800–854. Online: https://tinyurl.com/w3pdunnr

Freeman, J. (2000). The Private Role in Public Governance. *New York University Law Review*, 75(3), 543–675. Online: https://tinyurl.com/bd3k8wcr

Frosio, G., Geiger, C. (2023). Taking fundamental rights seriously in the Digital Services Act's platform liability regime. *European Law Journal*, 29(1–2), 31–77. https://doi.org/10.1111/eulj.12475

Gao, J. (2015). Pernicious Manipulation of Performance Measures in China's Cadre Evaluation System. *The China Quarterly*, 223, 618–637. https://doi.org/10.1017/S0305741015000806

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6), 854–871. https://doi.org/10.1080/1369118X.2019.1573914

Gorwa, R. (2024). *The Politics of Platform Regulation. How Governments Shape Online Content Moderation*. Oxford University Press. https://doi.org/10.1093/oso/9780197692851.001.0001

Van Hoboken, J., Quintais, J. P., Appelman, N., Fahy, R., Buri, I., & Straub, M. (Eds.). (2023). *Putting the DSA into Practice. Enforcement, Access to Justice and Global Implications.* Verfassungsblog. https://doi.org/10.17176/20230208-093135-0

Hu, J. (2018, December 19). *Analysis of innovation in internet food safety governance systems* [网络食品安全治理制度创新探析]. *China Pharmaceutical News*. Online: https://fzyjs.chinalaw.org.cn/portal/article/index/id/519.html

Husovec, M. (2024). *Principles of the Digital Services Act*. Oxford University Press. https://doi.org/10.1093/law-ocl/9780192882455.001.0001

Klonick, K. (2018). The New Governors. The People, Rules, and Processes Governing Online Speech. *Harvard Law Review*, 131(6), 1598–1670. Online: https://tinyurl.com/yx9kwejc

Li, J. (2015). From "unitary, unidirectional, and segmented" to "pluralistic, networked, and collaborative": The path to improving China's food safety regulatory mechanism [从"一元单向分段"到"多元网络协同"——中国食品安全监管机制的完善路径]. *Journal of Beijing Institute of Technology (Social Sciences Edition)* [北京理工大学学报（社会科学版）], 17(4), 93–97. Online: https://dx.doi.org/10.15918/j.jbitss1009-3370.2015.0413

Li, J., Yao, D., Li, S., & Wang, Q. (2014). Separate regulation or unified regulation: Organisational structure in food safety regulation [分头监管还是合并监管——食品安全监管中的组织结构]. *World Economy* [世界经济], 37(10), 165–192. Online: https://tinyurl.com/2uuvtfea

Lian, L., & Cheng, X. (2014). Explore the mode of regulation of food safety management in China [中国食品安全管理中监管模式探究]. *Journal of Agriculture* [农学学报], 4(5), 82–85. Online: https://doi.org/10.3969/j.issn.1007-7774.2014.05.020

Marco Colino, S. (2022). The Incursion of Antitrust into China's Platform Economy. *The Antitrust Bulletin*, 67(2), 237–258. https://doi.org/10.1177/0003603X221084152

Mashaw, J. L. (1976). The Supreme Court's Due Process Calculus for Administrative Adjudication in Mathews v. Eldridge: Three Factors in Search of a Theory of Value. *University of Chicago Law Review*, 44(1), 28–59. Online: https://chicagounbound.uchicago.edu/uclrev/vol44/iss1/5

Mashaw, J. L. (1983). *Bureaucratic Justice. Managing Social Security Disability Claims*. Yale University Press.

Mashaw, J. L. (1985). *Due Process in the Administrative State*. Yale University Press.

Metzger, G. E. (2003). Privatization as Delegation. *Columbia Law Review*, 103(6), 1367–1502. https://doi.org/10.2307/3593390

Pasquale, F. (2015). *The Black Box Society. The Secret Algorithms that Control Money and Information*. Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Pei, X., Tandon, A., Alldrick, A., Giorgi, L., Huang, W., & Yang, R. (2011). The China melamine milk scandal and its implications for food safety regulation. *Food Policy*, 36(3), 412–420. https://doi.org/10.1016/j.foodpol.2011.03.008

Peking University E-Commerce Law Research Centre [北京大学电子商务法研究中心]. (2024, December 25). *The governance rights of e-commerce platform operators and their exercise boundaries* [电商平台经营者的治理权利及其行使界限]. Online: https://www.100ec.cn/detail--6645539.html

Roberts, M. T., & Lin, C.-F. (2016). 2016 China Food Law Update. *Journal of Food Law & Policy*, 12(2), 238–272. https://doi.org/10.54119/jflp.fpcg7359

Rose-Ackerman, S. (2021). *Democracy and Executive Power. Policymaking Accountability in the US, the UK, Germany, and France*. Yale University Press. https://doi.org/10.2307/j.ctv1z9n1kq

Schramm, M. (2024). Platform Administrative Law. A Research Agenda. *SSRN*. https://doi.org/10.2139/ssrn.4898542

Standing Committee of the National People's Congress. (2009). *Food Safety Law of the People's Republic of China* [中华人民共和国食品安全法] (Order of the President of the People's Republic of China No. 9). Online: https://tinyurl.com/eenhr7vh

Standing Committee of the National People's Congress. (2015). *Food Safety Law of the People's Republic of China* [中华人民共和国食品安全法] (Order of the President of the People's Republic of China No. 2, as amended 2018, 2021, 2025). Online: https://tinyurl.com/bddbm5rn

Standing Committee of the National People's Congress. (2017). *Administrative Litigation Law of the People's Republic of China* [中华人民共和国行政诉讼法] (as amended). Online: https://www.spp.gov.cn/spp/fl/201802/t20180207_365256.shtml

Standing Committee of the National People's Congress. (2021). *Administrative Penalty Law of the People's Republic of China* [中华人民共和国行政处罚法] (as amended). Online: https://www.nmpa.gov.cn/xxgk/fgwj/flxzhfg/20250418144032179.html

Standing Committee of the National People's Congress. (2018). *E-Commerce Law of the People's Republic of China* [中华人民共和国电子商务法] (Order of the President No. 7). Online: https://tinyurl.com/5e9pr3wp

Standing Committee of the National People's Congress. (2025, October 27). *Report on the enforcement inspection of the Food Safety Law of the People's Republic of China* [全国人民代表大会常务委员会执法检查组关于检查《中华人民共和国食品安全法》实施情况的报告]. Online: http://www.npc.gov.cn/npc/c2/c30834/202510/t20251027_448926.html

State Administration for Market Regulation. (SAMR) (2022). *Provisions on the supervision and administration of the implementation of principal responsibility for food safety by enterprises* [企业落实食品安全主体责任监督管理规定] (SAMR Order No. 60). Online: https://www.gov.cn/gongbao/content/2022/content_5725279.htm

State Administration for Market Regulation. (SAMR) (2025, December 28). *Provisions on the supervision and administration of the implementation of principal responsibility for food safety by livestream e-commerce operators* (SAMR Order No. 120) [直播电商经营者落实食品安全主体责任监督管理规定]. Effective March 20, 2026. Online: https://tinyurl.com/mum5v8un

State Council of the People's Republic of China. (2019). Opinions on comprehensively implementing inter-departmental "double-random, one-disclosure" supervision in the field of market regulation [关于在市场监管领域全面推行部门联合"双随机、一公开"监管的意见] (Guo Fa [2019] No. 5). Online: https://www.gov.cn/gongbao/content/2019/content_5368520.htm

Wu, L., Wang, X., Yin, S., & Zhang, X. (2016). *Investigation report on China's food safety risk governance system and governance capacity* [中国食品安全风险治理体系与治理能力考察报告]. China Social Sciences Press.

Xie, J., Xu, L., & Tao, X. (2018). Analysis of food safety obligations of third-party network platforms [网络第三方平台的食品安全义务分析]. *China Food Safety Magazine* [食品安全导刊], (16), 32–34. Online: http://dx.chinadoi.cn/10.3969/j.issn.1674-0270.2018.16.016

Zheng, J. (2016). The Role of ODR in Resolving Electronic Commerce Disputes in China. *International Journal of Online Dispute Resolution*, 3(1), 41–68. https://doi.org/10.5553/IJODR/235250022016003001006

Zuo, Y. (2026, January 16). *New rules for livestream e-commerce food safety: Drawing red lines and establishing standards to target platform principal responsibility* [直指平台主体责任 新规为直播电商食品"划红线　立规矩"]. China News. Online: https://www.chinanews.com.cn/cj/2026/01-16/10553513.shtml

Zweigert, K., & Kötz, H. (1998). *Introduction to Comparative Law* (3rd ed., translated by T. Weir). Oxford University Press.

Institutiones Administrationis – Journal of Administrative Sciences Vol. 6 (2026) No. 1, 24–40
DOI: 10.54201/iajas.213

How Fundamental is the Right to Free Movement of Persons within the European Union?

Jeremy Julian Sarkin*, **Elena-Delia Bancu****

* Senior Research Fellow, School of Advanced Study, University of London, London, United Kingdom; Extraordinary Professor of Law, Department of Jurisprudence, University of Pretoria, Pretoria, South Africa. E-mail: JSarkin@post.harvard.edu

** Researcher, NOVA School of Law, Lisbon, Portugal. E-mail: edbancu@gmail.com

Abstract

The free movement of persons is a key pillar of European integration. However, whether it constitutes a fundamental right remains contentious. This article argues that the fundamental status of free movement has been interpreted inconsistently, and highlights the need to bridge the gap between rhetoric and reality by clearly affirming it as a fundamental right. The claim is that a legal redesign of this right is required. The article uses Rawls' theory of justice to formulate a clearer normative standard for evaluating and redesigning European law on free movement of persons. The proposed shift towards a human-rights-based, residence-anchored conception of free movement can thus be framed as an effort to align the European Union's basic structure more closely with Rawlsian principles: securing free movement as an equal basic liberty for all legally resident persons and restructuring mobility-related rules to promote, rather than undermine, fair equality of opportunity across the Union.

Keywords

free movement of persons, European residence, principle of equality of opportunity, fundamental rights, European Court of Justice, Rawls' theory of justice

1 Introduction

The free movement of persons is portrayed as one of the critical pillars and major achievements of European integration, largely due to its role in establishing the internal market (Barnard, 2016). However, its importance extends beyond this, significantly impacting the lives of citizens within the European Union (EU). The right, therefore, has an internal market and citizen-oriented scope, and may be viewed through both economic and human rights lenses. However, whether it constitutes a fundamental right remains unclear, though it is often assumed – especially since the introduction of the citizenship of the Union (the current Article 20, para. 1 of the Treaty on the Functioning of the European Union, henceforth: "TFEU") (Hyltén-Cavallius, 2020, 26; de Cecco, 2014, 388; Jacobs, 2007, 595; White, 2005, 886). In the case of the free movement of workers, it has largely been linked to its instrumental, though not exclusive, role in supporting the internal market (de Cecco, 2014; Tryfonidou, 2016). In addition, despite the great significance

 Published online: 03 June 2026

of free movement, the legal rules governing it (for the present purposes: Articles 20, 21 and 45 of the TFEU and Article 45 of the Charter of Fundamental Rights of the European Union, henceforth: "Charter") do not explicitly reference any principle to connect it to the core values of citizens, within its framework.

The 2020 pandemic created a context in which the balance between the functioning of the internal market and the protection of citizens' wider interests – particularly regarding the restoration of free movement – came to the forefront. This context raised questions about which fundamental values are better protected and the EU's capacity to safeguard citizens' broader interests on an equal footing with internal market protection. This directly impacts the characterization of free movement as a fundamental right and raises questions about how fundamental this right truly is within the EU.

This article provides a fresh analytical lens for assessing the free movement of persons as a fundamental right, starting from the discontinuity of its underlying values, as evidenced by current interpretations. On the one hand, there are the interpretations of European legal provisions and rulings by the European Court of Justice. On the other hand, there is the reality reflected in the legislative measures agreed at the Union level to facilitate the exercise of this right in practice. A key factor contributing to this discontinuity is that there are two different objectives – internal market and citizens' rights – associated with the free movement of persons. The problem is that these objectives often stand in opposition, leading to tensions that undermine the free movement as a fundamental right serving the interests of citizens. This was the case, for example, in the context of the pandemic. The reintroduction of border controls and the initial exceptions granted only to cross-border workers revealed a predominant focus on sustaining the internal market, while the broader interests of citizens were largely overlooked.

The article argues that a solution to the problem posed by the overlapping – and at times conflicting – internal market objectives and citizen-centered protection would be to redesign the legal framework of this right on an exclusive human-rights basis. This proposal introduces an innovative shift by framing free movement as a fundamental right, not contingent on economic and nationality-based conditions, rooted in human rights considerations, rather than internal market logic. It suggests using legally acquired residence – not Union citizenship or economic status – as the basis for accessing this right. This would ensure an equal opportunity for all citizens within the EU. It would also ensure an equitable balance between citizens' rights and Member States' welfare interests. Uniquely, it calls for legislative reform, rather than reliance on European Court of Justice jurisprudence, to embed this rights-based approach and provide a clearer, more inclusive legal framework. The envisaged solution guarantees the *continuum* of the movement and migration of people within the EU, exercised with the intention of enhancing their lives, and not of gaining access to a better social welfare system. The latter has emerged as a core basis of the opposition to viewing free movement as a right available to all.

The article is organized into six sections. The first section details the methodology employed, while the second introduces Rawls' theory of justice. Specifically, the article adopts a normative, Rawls-inspired framework that proposes to integrate the principle of equality of opportunity into a new approach to this right, using it as an analytical filter to identify the core values underlying citizens' movement. The third section examines key aspects surrounding the contextualization of the development of free movement as a fundamental right, closely linked to the European project's evolution. It is followed by a fourth section that turns to the actions of European institutions in relation to ensuring the exercise of this right. Together, these discussions help to assess the extent to which the fundamental rights dimension of free movement can transition

from rhetoric to reality. The fifth section considers a way forward for addressing this right – as the proposed solution – by redesigning its legal framework through a human rights perspective.

The conclusion brings together the findings of the article to further the discussion on how fundamental free movement is or might be thought to be within the EU. At a pragmatic level, strengthening this right as fundamental is linked to the need to consolidate the protection afforded to those who wish to exercise it. This means that any attempt to restrict this right must be viewed in the context of fundamental rights and the impact such measures would have on individuals (de Cecco, 2014, 398), rather than on the internal market. This article argues that efforts to bring the European project closer to its citizens, along with the socio-political context, require better addressing the fundamental rights dimension of free movement, and reconsidering the very basis of its legal framework.

2 Methodology

The article's methodology combines doctrinal legal analysis, historical interpretation, literature review, and normative theorizing to examine whether the free movement of persons in European law is or might be thought to be a fundamental right and to outline a rights-based redesign of its legal framework.

Regarding research design, the article employs a qualitative, doctrinal approach grounded in the interpretation and systematic analysis of EU primary and secondary law, the case law of the European Court of Justice, and relevant legal and political theory. The study is both descriptive and normative: it reconstructs the historical interpretation of free movement and puts forward a proposal to reconceptualize it as a human-rights-based fundamental right grounded in equality of opportunity.

The analysis draws on four main bodies of material: EU primary law (particularly Articles 20, 21, and 45 TFEU, as well as Article 45 of the Charter); EU secondary legislation (notably Regulation (EEC) No 1612/68 (1968) on freedom of movement for workers within the Community, henceforth: "Regulation 1612/68", Regulation (EU) No 492/2011 (2011) on freedom of movement for workers within the Union, henceforth: "Regulation 492/2011", and Directive 2004/38/EC (2004) on the right of citizens of the Union and their family members to move and reside freely within the territory of the Member States, henceforth: "Directive 2004/38"); European Court of Justice case law on free movement and Union citizenship; and academic literature on Union citizenship, internal market freedoms, and fundamental rights. These materials are supplemented by examinations of institutional practice and discourse from the European Commission, the European Parliament, and Member States, especially in response to the 2008 financial crisis and the 2020 pandemic.

The article reconstructs the historical development of the legal understanding of the free movement of persons, distinguishing pre- and post-Maastricht phases and the shift from an economic to a more rights-oriented framing, while noting the continued influence of internal-market logic. It then pursues a thematic analysis of academic debates, European Court of Justice jurisprudence, and European institutions' actions and discourse on free movement, especially during a crisis. The emphasis is placed on cases balancing human-rights arguments against market freedoms, and on whether practice reflects a fundamental-rights-based or a market-oriented conception of free movement. The article adopts a Rawlsian normative framework to guide future legislative changes related to free movement of persons, ensuring

equal access to fundamental rights and life opportunities for all individuals legally residing in the EU, regardless of their economic activity or Union citizenship status.

In developing its proposal for reforming free movement, the article integrates the doctrinal, historical, and normative analyses to outline a legislative blueprint for redesigning the EU's legal framework. This includes specifying how relevant Treaty provisions (Articles 20, 21, para.1, and 45 TFEU, and Article 45 of the Charter) and key secondary instruments could be reinterpreted or revised to reflect a rights-based understanding of mobility.

3 Rawls' Theory of Justice

Rawls' theory of justice offers a useful foundation for reconceptualizing the free movement of persons in the EU as a fundamental right rather than a primarily market-oriented freedom. Rawls articulates a model of a just basic structure grounded in two principles of justice chosen under fair conditions that remove knowledge of individuals' social positions, talents, or preferences (Rawls, 2016). This perspective illuminates the distributive and status-related consequences of how access to movement and residence is organized in the EU and allows the article to treat free movement as a matter of equal access to core opportunities rather than a by-product of economic integration. Scholars working at the intersection of the EU and fundamental rights frequently rely on Rawls (Klaser, 2021, 1–30), and his theory is widely regarded as "the most influential contemporary theory about justice" (de Boer, 2013, 149).

Rawls's framework is valuable because it structures a just society around two lexically ordered principles (Rawls, 1971, 54–56). The first guarantees a fully adequate scheme of basic liberties for all, including freedoms of movement and association – liberties that must be equal and cannot be sacrificed for aggregate economic benefit. The second regulates social and economic inequalities: they must be attached to positions open to all under conditions of fair equality of opportunity, and they must benefit the least advantaged. Applied to the EU, these principles call into question legal arrangements that frame cross-border mobility primarily as an economic freedom or a privilege tied to specific labor-market roles.

When applied to the EU's free movement regime, Rawls's first principle suggests that movement and residence across Member States should be understood as part of a broader set of basic liberties available to all individuals within the Union's jurisdiction. If free movement is to qualify as a fundamental right in Rawlsian terms, it must protect interests central to individuals' life plans – such as access to work, education, family life, and personal development in another Member State – regardless of their economic status. This supports an interpretation of free movement rooted in human-rights values and oriented towards the protection of persons, rather than the facilitation of cross-border economic activity.

Rawls' second principle, particularly the requirement of fair equality of opportunity (Rawls, 1991), directly engages with the dual structure of the current EU regime, which distinguishes sharply between economically active and inactive movers, and between Union citizens and other legally resident individuals. Fair equality of opportunity demands more than formal non-discrimination: it requires that individuals with similar talents and ambitions have comparable real prospects of accessing valuable life opportunities, irrespective of their social background. Transposed to free movement, this implies that the legal framework should avoid entrenching structural advantages for certain groups – such as workers with stable employment – while imposing restrictive conditions on others, such as students, economically inactive citizens, or family members, that limit their ability to relocate and integrate.

On this reading, a system that ties full access to movement-related rights to economic contribution or to a specific formal status (Union citizenship) risks violating fair equality of opportunity. It produces a stratified regime in which some individuals enjoy broad mobility and associated rights, while others face legal and practical barriers that constrain their life plans. A Rawlsian approach, therefore, strengthens the article's proposal to shift the basis of mobility-related rights from economic activity and formal citizenship to legally acquired residence. Treating all legally resident persons as potential holders of a fundamental right to free movement better aligns with the requirement that access to core opportunities should not depend on arbitrary features such as birth nationality or immediate employment status.

Rawls emphasizes that his principles apply to the "basic structure" of society – the central political, legal, and economic institutions that shape life prospects (Rawls, 1978). Although the EU is not a state, its legal order and internal market nonetheless function as a basic structure for those living under its jurisdiction: they allocate opportunities, distribute rights, and define membership boundaries in ways that deeply affect individuals' life chances. From a Rawlsian viewpoint, this basic structure should be evaluated based on whether its core freedoms, including the free movement of persons, are arranged to secure equal basic liberties and fair opportunities for all who are subject to it.

This perspective clarifies the normative tension at the heart of the current EU free movement regime. When free movement is framed primarily as a market freedom, restrictions are frequently justified by appeals to economic efficiency, fiscal protection, or labor-market management. A Rawlsian approach, by contrast, requires that any limitations on a fundamental right to move and reside be shown to respect basic liberties and fair equality of opportunity, giving priority to the protection of individuals' core interests over aggregate economic considerations (Rawls, 1975). This does not rule out all restrictions, but it does require that the starting point of legal design be the status of free movement as a basic liberty – not its utility for market integration.

By integrating Rawls' theory of justice into the analysis, the article articulates a clearer normative standard for evaluating and redesigning European law on the free movement of persons. The proposed shift towards a human-rights-based, residence-anchored conception of free movement can thus be framed as an effort to align the EU's basic structure more closely with Rawlsian principles: securing free movement as an equal basic liberty for all legally resident persons and restructuring mobility-related rules to promote, rather than undermine, fair equality of opportunity across the Union.

4 The Contextualization of Free Movement of Persons as a Fundamental Right

4.1 Free Movement of Persons – *Pre-Maastricht Treaty*

The free movement of persons was introduced in the original Community Treaty not from a human rights perspective, but as an economic instrument to facilitate the creation of the internal market (Alston & Weiler, 1998, 665; Barnard, 2016, 204). Nonetheless, within the complex context of the integration process, free movement gradually detached from the purely economic origin (Wollenschläger, 2011, 4–10).

What can be observed is that the free movement of workers was being reshaped as an individual and fundamental right at this early stage (Jacobs, 2007, 595–596; de Cecco, 2014, 386–387; Tryfonidou, 2017a, 76), primarily through secondary legislation and European Court of Justice rulings. In Regulation 1612/68, the preamble states that "freedom of movement

constitutes a fundamental right of workers and their families". This statement has been interpreted as setting the 'wider aims' of the free movement of workers (AG Jacobs Opinion, Case C- 344/87 *Bettray* 1989, para. 28),[1] as it clarifies "that labour is not, in Community law, to be regarded as a commodity and notably gives precedence to the fundamental rights of workers over satisfying the requirements of the economies of the Member States." (AG Jacobs Opinion, Case C- 344/87 *Bettray* 1989, para. 29) It follows from these legal interpretations that the early establishment of free movement of persons as a fundamental right was rooted in a broad understanding of the notion of "work" as being more than just an instrument for the internal market.

These secondary legal provisions laid the groundwork for several significant European Court of Justice rulings regarding the free movement of workers (de Cecco, 2014, 386–387; O'Leary, 2011, 505–506). Even though the rulings do not explicitly label this right as fundamental, they refer to it as "[…] one of the fundamental objectives of the Treaty" and recognize that "free access to employment is a fundamental right which the Treaty confers individually on each worker in the Community".[2]

Scholarly arguments for assessing the free movement of workers as an individual and fundamental right are rooted either in the idea of an emerging form of citizenship (Kochenov & Plender, 2012, 373) or in the existence of values beyond market-oriented ones associated with the movement of people (de Cecco, 2014, 387). While acknowledging the early expansion of free movement of persons interpretation, some scholars argue differently regarding the reasons for this, linking it more to the need to ensure the proper functioning of the internal market (Barnard, 2016, 204; Wollenschläger, 2011, 6–7). This article contends that what is often viewed and debated as early features of a fundamental rights dimension instead reflects the adaptation of an economic instrument – the free movement of persons – to a broader European context in which the internal market remained the predominant objective in relation to the movement of people. Therefore, neither of the arguments is persuasive enough to label free movement as a fundamental right, in the absence of a clear legal basis that directly connects this right to human rights interests inherent for the protection of citizens' core values.

4.2 Free Movement of Persons – *Post-Maastricht Treaty*

A formal development of the fundamental rights dimension of free movement began with the legal basis established by the Maastricht Treaty. It introduced a provision on Union citizenship and set the framework to include not only the economically active, but also the economically inactive Union citizens as beneficiaries of this right. However, this development is argued to convey the broader transformation of the European project and of fundamental rights discourse, which are not fully aligned with human rights considerations *per se*.

Thus, it is argued that the rhetoric of the free movement of persons as a fundamental right was part of the effort to transform the European project into a political community. This is because of the potential that the concept of fundamental rights has to provide legitimacy and

[1] Advocate General Jacobs Opinion, *Bettray* (1989) Case C- 344/87, ECLI:EU:C:1989:226, European Court of Justice.

[2] *Heylens* (1987) Case C-222/86, paras. 12, 14, ECLI:EU:C:1987:442, European Court of Justice. *Lawrie-Blum* (1986) Case C-66/85, para. 16, ECLI:EU:C:1986:284, European Court of Justice. *Bosman* (1995) Case C-415/93, paras. 93, 129, ECLI:EU:C:1995:463, European Court of Justice.

"[…] a moral grounding to a legal order which on its face was established principally to support the pursuit of economic goals […]" (de Búrca, 1995, 43). Framing the free movement of persons as a fundamental right could thus be seen as serving a purpose beyond the internal market's economic interests, contributing to the legitimation of the new European political community. However, such a perspective on free movement overlooks the essence of a fundamental right for citizens – specifically, its underlining core values and the opportunity of equal access to them. At the same time, a perspective of this kind reflects that the initial regulation of fundamental rights within the EU was primarily focused on ensuring the autonomy and supremacy of European law within Member States, rather than on the protection of human rights *per se* (de Búrca, 2011, 668; Williams, 2009, 563–564; Varju, 2014, 5, 7).

The Union citizenship concept had a significant contribution to how the free movement of persons is perceived as a fundamental right. Nonetheless, the extent of this impact depends on judicial interpretation, as the wording of the Treaty brings minimal details.

The interpretation evolved significantly beginning in 2001, with the European Court of Justice frequently affirming that "Union citizenship is destined to be the fundamental status of nationals of the Member States […]".[3] For some, this reading of Union citizenship laid the groundwork for the European Court of Justice's recognition of the right to free movement and residence as both an independent (Jacobs, 2007, 595) and a fundamental right (de Cecco, 2014, 388; White, 2005, 886) granted to all Union citizens. This recognition marked the shift in focus from economic considerations to citizens' rights (Spaventa, 2004; Kochenov & Plender, 2012; Nic Shuibhne, 2002). This article argues that this shift has supported the development of fundamental rights aspects within the free movement of persons, but it primarily occurred in the context of the European Court of Justice's interpretation of Union citizenship as the fundamental status of Member States' nationals, not as a direct emphasis on the protection of human rights *per se*.

It has been argued that the aim of Union citizenship is to establish "a general freedom of movement, independent from economic freedoms" (Kadelbach, 2010, 454). Where this article departs from other scholars is in its argument that, to achieve this goal, the design of a general free movement of persons should be grounded in human rights considerations, as this would reinforce its fundamental status. The goal is to prevent overlap with internal market objectives, while avoiding reliance solely on Union citizenship. This is because it is argued that, despite its powerful rhetoric, the Union citizenship as a basis for free movement limits, rather than ensures, its realization as a fundamental right. The *Rottmann* (Case C-135/08 2010),[4] *Zambrano* (Case C-34/09 2011) line of case law strengthens Union citizenship as a tool to protect citizens' rights associated with it, however in an attempt to depart from the requirement to have exercised the right to free movement. This undermines what free movement in itself could represent as a fundamental right for citizens and their equal access to core values in another Member State. In

[3] *Grzelczyk* (2001) Case C-184/99, para. 31, ECLI:EU:C:2001:458, European Court of Justice. *Baumbast, R* (2002) Case C- 413/99, para. 82, ECLI:EU:C:2002:493, European Court of Justice. *Garcia Avello* (2003) Case C-148/02, para. 22, ECLI:EU:C:2003:539, European Court of Justice. *Zhu and Chen* (2004) Case C-200/02, para. 25, ECLI:EU:C:2004:639, European Court of Justice. *Bidar* (2005) Case C-209/03, para. 31, ECLI:EU:C:2005:169, European Court of Justice. *Zambrano* (2011) Case C-34/09, para. 41, ECLI:EU:C:2011:124, European Court of Justice.

[4] *Rottmann* (2010) Case C-135/08, ECLI:EU:C:2010:104, European Court of Justice.

Mirin (Case C-4/23 2024)[5] the emphasis falls back on safeguarding the right to free movement as a core element of Union citizenship status. However, in this case, the free movement rationale is employed as a means to protect transgender rights.

There are, however, other shortcomings arising from the set-up of Union citizenship. These relate to issues concerning the personal scope of free movement of persons, which in practice goes beyond the limits of the legal status of Union citizenship. What is more is that despite the internal market and Union citizenship not being conflated, the legal reading of the latter concept has frequently resembled a return to the idea of market citizenship (Spaventa, 2017, 208, 221–222), where movement is conditioned on the exercise of an economic activity. This new approach treats Union citizenship "as a mere and minor *addition* [original emphasis] to national citizenship, rather than a true supranational status." (Spaventa, 2017, 224) In this narrower interpretation, "Union citizenship has lost much of its innovative potential" and adds little to the right to move and reside freely, as this right already existed for economically active individuals (Spaventa, 2017, 221). Alongside, there are situations in which the legal reading of Union citizenship has been confined to a minimalist interpretation according to which the free movement of persons exists only by virtue of the internal market and once a Member State withdrew from the European project, the rights of free movement were no longer available to the nationals of that Member States (Spaventa, 2023, 179). It is argued that these new readings may have a significant impact on the assessment of free movement of persons as a fundamental right, as they prioritize economic considerations and internal market objectives.

Legal and judicial developments related to the right to move and reside freely following the introduction of Union citizenship were consolidated into a single legal instrument (Guild et al., 2014, 2). However, Directive 2004/38 did not fully succeed in establishing a framework that ensures a consistent, human-rights-based approach to the right to move and reside freely – both in general and for economic purposes. This may be because the principles underlying Union citizenship – equal treatment and non-discrimination – and applicable to economically active citizens cannot be indefinitely extended to economically inactive ones, as they represent two different realities (Thym, 2019, 104). In addition, Directive 2004/38 did little to clarify the fundamental rights dimension of the free movement of persons, as many aspects related to free movement for economic reasons remain closely tied to the instrumental role of this right within the internal market.

The freedom of movement and residence is incorporated in Article 45, para.1 of the Charter as one of the rights of Union citizens. However, the Charter provisions outlining this right merely mirror the wording of Article 21, para. 1 of the TFEU – the right to move and reside freely – and provide no further details about the intrinsic values of citizens that they aim to protect. The Charter contains an enumeration of rights that arguably cements that there is no fundamental right of free movement beyond rhetoric, and does not extend the scope of rights to groups that were previously not included in the circle of beneficiaries of free movement rights. Moreover, the movement of workers under Article 45 of the TFEU continues to be framed primarily as a functional right serving internal market objectives. This is because Article 15, para.2 of the Charter recognizes only the right to work and seek employment as fundamental freedoms of Union citizens. By contrast, the freedom of movement for workers as a fundamental right is expressly articulated as such *merely* in the preamble of secondary legislation, notably Regulation 492/2011.

---

[5] *Mirin* (2024) Case C-4/23, ECLI:EU:C:2024:845, European Court of Justice.

To conclude, the fundamental rights dimension of free movement of persons remains a concept insufficiently defined by the legal framework. It emerges from the foregoing that the duality within the scope of this right – internal market objectives and citizens' interests – challenges its assessment as a fundamental right. This duality arises from the European project development and instruments such as Union citizenship, making it difficult to set a clear legal basis for defining the right as fundamental in human rights terms without overlapping with internal market considerations.

5 Exploring the Fundamental Rights Dimension of Free Movement in Action

This section argues that the actions undertaken by European institutions – the European Court of Justice, the European Parliament and the European Commission – to ensure the exercise of free movement of persons reflect the contextual particularities of the development of this right and the tendency to prioritize internal market objectives. An overview of these actions helps to assess whether the fundamental rights dimension of free movement, including that of workers, can transition from rhetoric to reality. This assumption rests on the argument that the values embedded in a fundamental right are reflected in European institutions' actions, and take precedence over other values, particularly economic ones (de Búrca, 1995, 44).

5.1 Free Movement in the European Court of Justice's Rulings

This article notes that the European Court of Justice's approach to the protection of free movement of persons as a fundamental right through a focus on human rights interests has been mixed because of the close connection between this right, citizens' interests, and internal market objectives, and the absence of a clearly defined EU concept of fundamental rights.

The European Court of Justice has tended to focus on internal market objectives by prioritizing the protection of fundamental freedoms (Douglas-Scott, 2011, 677; Kriki, 2020, 80–81) over fundamental rights. This occurred to the extent that some considered the protection of fundamental freedoms "[…] the leading rule to which other grounds may or may not qualify as exceptions." (Heliskoski, 2003, 441) This approach relates to the scholarly debates on the EU's lack of a "clearly developed, substantive sense of human rights" (Douglas-Scott, 2011, 678) and the fact that principles developed by the European Court of Justice jurisprudence, like respect for fundamental rights, are not grounded in coherent fundamental values (Williams, 2009, 560–561).

In contrast to the leading rule of fundamental freedoms protection, in cases such as *Gebhard* (Case C-55/94 1995)[6] and *Carpenter* (Case C- 60/00 2002)[7] – which touch on both citizens' rights and market considerations – the European Court of Justice has framed its reasoning to remove Member States' obstacles to freedom of movement from the perspective of citizens' interests, rather than that of the internal market. Likewise, in the *Nordic Info BV* ruling (Case C-128/22 2023),[8] the Court did not treat the travel restrictions as mere market obstacles. Instead, it interpreted them through the lens of fundamental rights – including the right to private/family life, freedom of movement, freedom to conduct a business (Case C-128/22 *Nordic Info BV* 2023,

[6] Gebhard (1995) Case C-55/94, ECLI:EU:C:1995:411, European Court of Justice.

[7] Carpenter (2002) Case C- 60/00, ECLI:EU:C:2002:434, European Court of Justice.

[8] Nordic Info BV (2023) Case C-128/22, ECLI:EU:C:2023:951, European Court of Justice.

paras. 92–95, 98). This approach aligns with a more explicit citizens-rights focus adopted by other European institutions later in the 2020 pandemic and moves beyond a purely internal-market proportionality analysis of travel restrictions.

Moreover, in rulings not involving conflicts between fundamental freedoms and rights, the Court approaches the free movement of persons from the perspective of the protection of citizens' rights. This was the case in *Sala*,[9] *Grzelczyk*, *Collins*,[10] *Garcia Avello*, *Bidar*, *Baumbast, R*, *Zhu and Chen*, *Zambrano*. However encouraging, in these cases, the European Court of Justice had gradually built on the reading of Union citizenship as the fundamental status of Member States' nationals, strengthening its role as a tool to protect citizens' rights at the European level and increasingly loosening its link to the actual exercise of the right to free movement.

An interesting aspect is noted in cases concerning the workers' movement. Notably, the European Court of Justice has appeared willing "to re-read the market freedoms in light of the status of Union citizenship", although this has been limited to economically active citizens (Tryfonidou, 2017b, 327). This was the case in the *Ritter-Coulais* (Case C-152/03 2006)[11] in which the Court interpreted the intent of the freedom of movement for workers to go beyond internal market interests towards facilitating the needs of Union citizens. This raises the question of whether the characterization of free movement as a fundamental right is conditional on the economic involvement of its beneficiaries.

In conclusion, the European Court of Justice largely confines to its initial approach to human rights discourse, whereby the scope of the application of the principle of respect for human rights "is to be determined by the preservation of the EU and in particular the construction of an internal market." (Williams, 2009, 566)

5.2 Free Movement in Other European Institutions' Actions

The actions taken at the European level by other European institutions to guarantee the exercise of the free movement of persons are mostly visible in specific circumstances, such as transnational challenges. Alongside, these circumstances question the extent to which the possible limitations on free movement align with the requirements as interpreted through the first paragraph of Article 52 of the Charter to respect the essence of the right to free movement and uphold the principle of proportionality. Relevant here are the effects of the 2008 financial crisis and the 2020 pandemic on the free movement of persons.

The 2008 financial crisis justified the temporary restrictions adopted by some Member States, such as Spain and the UK, later endorsed by European institutions, to limit access to their labor markets for Romanian and Bulgarian workers (Ludera-Ruszel, 2015, 166–167). These measures recalled, in effect, the restrictions introduced in 2007 when Romania and Bulgaria joined the EU and transitional arrangements temporarily limited their workers' free movement. A further study of changes in European institutions' discourse on the free movement of persons highlights the impact of the 2008 financial crisis (Roos & Westerveen, 2020). It notes that the European Commission's and the European Parliament's discourse shifted from treating the freedom of movement as an individual right to advocating for more restrictions on its exercise (Roos & Westerveen, 2020, 63, 75). Thus, it has been argued that "[t]he EU of equal rights and

[9] Sala (1998) Case C-85/96, ECLI:EU:C:1998:217, European Court of Justice.

[10] Collins (2004) Case C-138/02, ECLI:EU:C:2004:172, European Court of Justice.

[11] Ritter-Coulais (2006) Case C-152/03, ECLI:EU:C:2006:123, European Court of Justice.

opportunities for all European citizens is being challenged by the idea of the EU as a marketplace to which access is conditional." (Roos & Westerveen, 2020, 65) In other words, the essence of free movement of persons as a fundamental right has been challenged by its understanding as an economic instrument confined to workers' mobility.

The context of the 2020 pandemic highlighted the aspect of overlapping interests associated with the free movement of persons. At the European institutions' level, the initial response to this crisis was consistent with their obligation to ensure the internal market' functioning. They sought and implemented exceptions to travel restrictions with the primary aim to safeguard workers' mobility and access to the labor market (Guild, 2020, 3; Duić & Sudar, 2021, 36–37). This indicates that internal market interests are prioritized in addressing the exercise of free movement of persons.

Additionally, these initial actions raised questions about which values are better protected within the EU. In this sense, scholars questioned why, if exceptions to travel restrictions complying with public health standards were possible, similar measures were not taken on a broader basis (Ramji-Nogales & Goldner Lang, 2020, 597–598) to cover all aspects of the free movement of persons, beyond workers alone. Thus, an initial human-rights-based approach would have included the impact of travel restrictions on individuals, not just on the internal market. This is because the considerable measures taken to restrict the exercise of the free movement of persons effectively eroded the very essence of this right, were "[…] likely to undermine everyone's sense of belonging to the EU" (Pacces & Weimer, 2020, 288) and represented genuine challenges to the fundamental principles of the EU and human rights (Ramji-Nogales & Goldner Lang, 2020, 599).

The subsequent measures and actions by European institutions reflected an emphasis on human rights considerations and the broader impact of travel restrictions on citizens' lives – a change also evident in the Court's approach, as highlighted in the case previously discussed, the *Nordic Info BV*. Although not the norm, these are important steps towards a human-rights-based approach to the free movement of persons. Making the shift effective requires revising the legal framework for this right.

To sum up, the various actions and measures taken by European institutions – whether on a case-by-case basis or in response to transnational challenges – largely prioritize internal market interests in addressing the free movement of persons. In light of the argument raised at the beginning of section five, European institutions' actions reflect little about the free movement of persons as a fundamental right of citizens through a focus on the wider aspects of human rights interests.

6 A Way Forward to Deal with the Free Movement of Persons in the European Union

This article helps determine that, in relation to the free movement of persons, there are aspects that sometimes overlap. These aspects relate to the various interests – citizens and internal market – associated with this right. Consequently, due to the different components of free movement, there remains a discontinuity in terms of its underlying principles. There is also a difference in how free movement is understood and protected. On the one hand, there are the interpretations of European law and European Court of Justice rulings; on the other hand, there is the practice of the EU, as indicated by the measures it has taken to protect the exercise of this right. This situation raises the question of the extent to which the free movement of persons, including workers, as a fundamental right can transition from rhetoric to reality and be protected in practice from the perspective of citizens rather than the internal market interests.

A solution and way forward to deal with this would be to redesign the European legal framework of free movement of persons, primarily through the lens of human rights considerations. The argument holds that while free movement as a fundamental right originates in Union citizenship, the aim is to extend beyond it. In this light, a better-regulated legal framework should not focus on further expanding the principles of equal treatment and non-discrimination based on Union citizenship, but rather on establishing an equitable approach between the citizens' rights and Member States' welfare interests. The human-rights-based rationale views free movement as a fundamental entitlement enjoyed by all individuals legally residing within a given space, grounded in human rights considerations and operating independently of internal market objectives. The citizenship rationale, by contrast, is inherently selective and more readily aligned with internal market logic: it grants free movement only to those falling within the category of Union citizens, whose mobility is often seen as contributing to the functioning of the internal market. The human-rights-based rationale aims to structure the free movement of persons as a right that is not divided into categories of beneficiaries or defined through the lens of the rights to which it may or may not give access.

This could be achieved through new legislation – rather than European Court of Justice rulings – that explicitly establishes free movement of persons as a fundamental right that is unconditionally available to citizens. This right would be grounded in the principle of equality of opportunity as a guiding filter, ensuring that citizens can access fundamental rights in a host Member State, prioritizing their individual interests over those of the internal market. The access would, however, be limited by the requirement that citizens are legally residing in the EU. Legally acquired residence would therefore become a key criterion for grounding this right and the principles underlying it. This approach takes into account the hierarchy of norms within the EU at present, which determines, according to Article 21, para. 1 of the TFEU, that the right to free movement is "[…] subject to the limitations and conditions laid down in the Treaties and by the measures adopted to give them effect." Thus, these existing limitations and conditions would need to be reconsidered by primary and secondary legislation. The aspects of the solution previously outlined are detailed below.

This article departs from other scholars' approaches by aiming to redesign the free movement of persons through a new lens – one grounded in human rights considerations that are intrinsic to protecting citizens' core values. These core values can be broadly defined as "[…] those ends deemed worth pursuing" (Williams, 2009, 559) and "[…] as intersubjectively shared preferences." (Habermas, 1996, 255) They correspond not only to the core values embedded in the right to free movement of persons, but also to the fundamental rights that citizens are entitled to, with equal access, once they have exercised their right to free movement within the EU.

The method for concretely defining citizens' core values draws on other scholars' approaches (de Boer, 2013, 159) and is guided by a normative framework rooted in Rawls' theory of justice (Rawls, 1971, 54–56, 60). Accordingly, the principle of equality of opportunity is proposed as an effective tool for understanding citizens' core values. This principle is to be interpreted broadly – not confined to economic opportunities or a person's economic standing in society (de Boer, 2013, 166–167). This principle helps outline a free movement genuinely grounded in citizens' interests rather than internal market objectives. Hence, the article argues that the redesigned free movement of persons framework must incorporate equality of opportunity – as a filter to ensure emphasis on intrinsic human rights considerations tied to citizens' core values.

The proposed solution for redesigning the free movement of persons legal framework aims to broaden the personal scope of this right beyond the current limitations set by the Treaty and Union citizenship, to include citizens legally residing in the EU as beneficiaries. This approach

rests on the principle that a right's fundamental nature should not depend on the nationality of its beneficiaries. Moreover, the approach reflects the current reality and resonates with the ongoing scholarly debate about Union citizens' rights already extended to third-country nationals provided they reside legally in a Member State (Kochenov, 2013, 106, 132; de Cecco, 2014, 397; Raucea, 2016; van den Brink, 2019, 28–29). The proposal is innovative in that it involves amending the Treaty provisions and seeks to firmly establish the residence requirement within primary law.

Legally acquired residence is argued to be a better criterion for determining free movement of persons beneficiaries because it also clears the traditional distinction between economically active and inactive Union citizens and reflects the European Court of Justice's tendency to progressively dilute the requirement to actively participate in an economic activity (Tryfonidou, 2016, 83). However, the concept of legally acquired residence must be defined more broadly than the framework set by the European Court of Justice in some rulings,[12] which cautiously limited it largely to the situations listed in Directive 2004/38 (Carter & Jesse, 2018, 1181, 1193–1202). Achieving a broader definition implies considering Member States' conditions for access to nationally based rights, while respecting their perspectives in setting legal residence criteria without imposing economic requirements, such as participation in the host state's labor market. Their involvement in adopting the new legislation helps address concerns about making the free movement of persons a fundamental right. This is because the legal residence requirement establishes a criterion for equal treatment other than Union citizenship, particularly relevant in relation to citizens' access to social benefits.

The aim is to integrate the aspects of the solution outlined in this paper by amending primary legal provisions – Articles 20, 21, para.1, and 45 of the TFEU, and Article 45 of the Charter. By amending Article 20 of the TFEU, the primary legislation would define Union citizenship as a legal status, but distinct from the specific rights granted to citizens in the EU. Consistent with this, the legal framework governing the right to move and reside freely – under Article 21 para.1 of the TFEU – should apply to all individuals legally residing in the EU. It should also incorporate the principle of equality of opportunity – as a guiding framework for clearly identifying the core values informing citizens' free movement. In addition, Article 45 of the Charter must reinforce the underlying rationale of the free movement of persons – as set out in the amended version of Article 21 para.1 of the TFEU. The idea of disconnecting Union citizenship from its associated rights, and redefining eligibility based on legally acquired residence, is currently under debate among European scholars (van den Brink, 2019, 28).

Moreover, this paper's arguments for reforming the free movement legal framework prioritize citizens' rights protection over internal market objectives and equal access to fundamental rights across all Member States. It is maintained that the right to free movement should be treated separately from the specific rights it grants access to. Consequently, it is contended that Article 45 of the TFEU should be amended to explicitly address only the right to work, as a right that reflects one of the citizens' core values, equally accessible to all citizens legally residing in the EU.

Changes to primary legislation would require corresponding amendments to EU secondary legislation to align with and effectively implement the new legal framework for free movement

[12] Dano (2014) Case C-333/13, ECLI:EU:C:2014:2358, European Court of Justice. García-Nieto (2016) Case C-299/14, ECLI:EU:C:2016:114, European Court of Justice. Alimanovic (2015) Case C-67/14, ECLI:EU:C:2015:597, European Court of Justice.

of persons. Moreover, existing requirements related to economic activity in EU secondary legislation would need to be reconsidered in relation to the exercise of free movement, as they are incompatible with the very notion of a fundamental right.

In view of the foregoing, the protection of citizens' core values must be at the heart of European legislation regulating the free movement of persons, encompassing all citizens legally residing in the EU. To achieve this, the European Commission must offer a legislative proposal to initiate changes to the legal framework governing this right. Such reform would establish a solid legal basis for addressing the question of how fundamental the right to free movement of persons within the EU is, or might be perceived to be.

7 Conclusion

To address the central question posed in this article, it has been necessary to situate the free movement of persons within the context of the European project's consolidation. As discussed in the third section, interpretations based on European legislation and European Court of Justice case law have favored the understanding of the free movement of persons as a fundamental right at an early stage. However, this development counts largely as the adaptation of the economic instrument – the free movement of persons – to the need of ensuring the functioning of the internal market. In addition, the close link between the fundamental rights dimension of free movement and Union citizenship – despite its powerful rhetoric – may create limitations in the assessment of this right as fundamental. This is the case due to the often market-oriented and minimalist interpretations of Union citizenship, the limitations its structure imposes on the personal scope of free movement, and the efforts to strengthen this concept as a source of rights no longer contingent on the exercise of free movement. On the other hand, as shown in the fourth part, European institutions' actions in relation to ensuring the exercise of the free movement of persons highlight the tendency to prioritize internal market objectives in association with this right.

The solution proposed by this article envisages the redesign of the free movement of persons legal framework through the legislative process, thereby establishing it as a fundamental right unconditionally available to citizens and ensuring them an equal opportunity to access fundamental rights in a host Member State. However, this access would be limited by the requirement that citizens are legally residing in the EU. The solution approaches the free movement of persons from a perspective grounded in human rights considerations, intrinsic to protecting citizens' core values, that would reinforce its status as a fundamental right.

It is argued that a legislative proposal, integrating aspects of the solution discussed in this article, should be put forward by the European Commission to initiate amendments, either to European secondary legislation or the Treaty provisions. The aim is to provide a solid legal basis for addressing how the issue of free movement of persons as a fundamental right, within the EU, should be handled. This proposal ought to employ legally acquired residence as the basis for determining who benefits from free movement, while integrating the principle of equality of opportunity as a safeguard to ensure that the right is grounded in human rights rather than driven by internal market objectives. This legislative proposal would enable the opportunity to address existing limitations and conditions on the right to free movement through legislation. It would also enable reaching an equitable balance between citizens' rights and Member States' welfare interests.

Acknowledgement

The Portuguese Fundação para a Ciência e a Tecnologia (FCT) provided funding, under award 2020.09127.BD.

References

Alston, P., & Weiler, J. H. H. (1998). An 'Ever Closer Union' in Need of a Human Rights Policy. *European Journal of International Law*, 9(4), 658–723. https://doi.org/10.1093/ejil/9.4.658

Barnard, C. (2016). *The Substantive Law of the EU. The Four Freedoms* (5th ed.). Oxford University Press. https://doi.org/10.1093/he/9780198830894.001.0001

Carter, D., & Jesse, M. (2018). The "Dano Evolution": Assessing Legal Integration and Access to Social Benefits for EU Citizens. *European Papers – A Journal on Law and Integration*, 3(3), 1179–1208. https://doi.org/10.15166/2499-8249/266

De Boer, N. J. (2013). Fundamental Rights and the EU Internal Market: Just how Fundamental are the EU Treaty Freedoms? A Normative Enquiry Based on John Rawls' Political Philosophy. *Utrecht Law Review*, 9(1), 148–168. https://doi.org/10.18352/ulr.219

De Búrca, G. (1995). The Language of Rights and European Integration. In J. Shaw, & G. More (Eds.), *New Legal Dynamics of European Union* (pp. 29–54). Oxford University Press.

De Búrca, G. (2011). The Road not Taken: The European Union as a Global Human Rights Actor. *The American Journal of International Law*, 105(4), 649–693. https://doi.org/10.5305/amerjintelaw.105.4.0649

De Cecco, F. (2014). Fundamental Freedoms, Fundamental Rights, and the Scope of Free Movement Law. *German Law Journal*, 15(3), 383–406. https://doi.org/10.1017/S2071832200018964

Douglas-Scott, S. (2011). The European Union and Human Rights after the Treaty of Lisbon. *Human Rights Law Review*, 11(4), 645–682. https://doi.org/10.1093/hrlr/ngr038

Duić, D., & Sudar, V. (2021). The Impact of Covid-19 on the Free Movement of Persons in the EU. *EU and Comparative Law Issues and Challenges series*, 5, 30–56. https://doi.org/10.25234/eclic/18298

Guild, E. (2020). Covid-19 Using Border Controls to Fight a Pandemic? Reflections from the European Union. *Frontiers in Human Dynamics*, 2, 1–5. https://doi.org/10.3389/fhumd.2020.606299

Guild, E., Peers, S., & Tomkin, J. (2014). *The EU Citizenship Directive. A Commentary*. Oxford University Press. https://doi.org/10.1093/oso/9780198849384.001.0001

Habermas, J. (1996). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (trans. W. Rehg). The MIT Press.

Heliskoski, J. (2003). Fundamental Rights versus Economic Freedoms in the European Union: Which paradigm?. In J. Petman, & J. Klabbers (Eds.), *Nordic Cosmopolitanism Essays in International Law for Martti Koskenniemi* (pp. 417–443). Brill/Nijhoff. https://doi.org/10.1163/9789004482043_023

Hyltén-Cavallius, K. (2020). *EU Citizenship at the Edges of Freedom of Movement.* Hart Publishing. https://doi.org/10.5040/9781509937288

Jacobs, F. G. (2007). Citizenship of the European Union–A Legal Analysis. *European Law Journal*, 13(5), 591–610. https://doi.org/10.1111/j.1468-0386.2007.00385.x

Kadelbach, S. (2010). Union Citizenship. In A. von Bogdandy, & J. Bast (Eds.), *Principles of*

European Constitutional Law (revised 2nd ed.) (pp. 443–478). Hart Publishing. https://doi.org/10.5040/9781474202169

Kochenov, D. (2013). The Essence of EU Citizenship Emerging from the Last Ten Years of Academic Debate: Beyond the Cherry Blossoms and the Moon?. *International and Comparative Law Quarterly*, 62(1), 97–136. https://doi.org/10.1017/S0020589312000589

Kochenov, D., & Plender, R. (2012). EU Citizenship: From an Incipient form to an Incipient Substance? The Discovery of the Treaty Text. *European Law Review*, 37(4), 369–396. Online: https://ssrn.com/abstract=2137680

Klaser, K. (2021). A Theory of Justice of John Rawls as Basis for European Fiscal Inion. *Review of Economics and Institutions*, 11(1–2), Article 4. http://doi.org/10.5202/rei.v11i1-2.328

Kriki, L. (2020). Freedoms or Rights? A Court Deciding While Comfortably Numb. *Nordic Journal of European Law*, 3(2), 58–81. https://doi.org/10.36969/njel.v3i2.22390

Ludera-Ruszel, A. (2015). Free Movement of Workers as an Instrument of Creation of the European Common Market. *European Integration Studies*, 9, 161–169. https://doi.org/10.5755/j01.eis.0.9.12804

Nic Shuibhne, N. (2002). Free Movement of Persons and the Wholly Internal Rule: Time to Move On?. *Common Market Law Review*, 39(4), 731–771. https://doi.org/10.54648/5095075

O'Leary, S. (2011). Free Movement of Persons and Services. In P. Craig, & G. de Búrca (Eds.), *The Evolution of EU Law* (2nd ed.) (499–545). Oxford University Press. Online: https://hdl.handle.net/1814/16314

Pacces, A. M., & Weimer, M. (2020). From Diversity to Coordination: A European Approach to Covid-19. *European Journal of Risk Regulation*, 11(2), 283–296. https://doi.org/10.1017/err.2020.36

Raucea, C. (2016). European Citizenship and the Right to Reside: 'No One on the Outside has a Right to be Inside?'. *European Law Journal*, 22(4), 470–491. https://doi.org/10.1111/eulj.12196

Ramji-Nogales, J. & Goldner Lang, I. (2020). Freedom of movement, migration, and borders. *Journal of Human Rights*, 19(5), 593–602. https://doi.org/10.1080/14754835.2020.1830045

Rawls, J. (1971). *A Theory of Justice*. The Belknap Press of Harvard University Press. https://doi.org/10.2307/j.ctvjf9z6v

Rawls, J. (1975). Fairness to Goodness. *The Philosophical Review*, 84(4), 536–554. https://doi.org/10.2307/2183853

Rawls, J. (1978). The Basic Structure as Subject. In A. I. Goldman, & J. Kim (Eds.), *Values and Morals. Essays in Honor of William Frankena, Charles Stevenson, and Richard Brandt* (pp. 47–71). Springer Netherlands. https://doi.org/10.1007/978-94-015-7634-5_4

Rawls, J. (1991). Justice as Fairness: Political Not Metaphysical. In J. A. Corlett (Ed.), *Equality and Liberty. Analyzing Rawls and Nozick* (pp. 145–173). Palgrave Macmillan. https://doi.org/10.1007/978-1-349-21763-2_10

Rawls, J. (2016). A Theory of Justice. In L. May (Ed.), *Applied Ethics* (pp. 21–29). Routledge. https://doi.org/10.4324/9781315097176

Roos, C., & Westerveen, L. (2020). The conditionality of EU freedom of movement: Normative change in the discourse of EU institutions. *Journal of European Social Policy*, 30(1), 63–78. https://doi.org/10.1177/0958928719855299

Spaventa, E. (2004). From Gebhard to Carpenter: Towards a (non-)economic European Constitution. *Common Market Law Review*, 41(3), 743–773. https://doi.org/10.54648/cola2004023

Spaventa, E. (2017). Earned Citizenship – Understanding Union Citizenship through Its Scope. In D. Kochenov (Ed.), *EU Citizenship and Federalism. The Role of Rights* (pp. 204–225). Cambridge University Press. https://doi.org/10.1017/9781139680714.009

Spaventa, E. (2023). Brexit and the Free Movement of Persons: What is EU Citizenship Really About?. In N. Nic Shuibhne (Ed.), *Revisiting the Fundamentals of the Free Movement of Persons in EU Law* (pp. 158–185). Oxford University Press. https://doi.org/10.1093/oso/9780198886273.003.0006

Thym, D. (2019). The Failure of Union Citizenship Beyond the Single Market. In R. Bauböck (Ed.), *Debating European Citizenship* (pp. 101–106). Springer. https://doi.org/10.1007/978-3-319-89905-3_20

Tryfonidou, A. (2016). *The Impact of Union Citizenship on the EU's Market Freedoms*. Hart Publishing. https://doi.org/10.5040/9781474202282

Tryfonidou, A. (2017a). Free Movement of Persons Through the Lenses of 'Discrimination' and 'Restriction'. In M. Andenas, T. Bekkedal, & L. Pantaleo (Eds.), *The Reach of Free Movement* (pp. 57–83). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-195-1_3

Tryfonidou, A. (2017b). The Federal Implications of the Transformation of the Market Freedoms into Sources of Fundamental Rights for the Union Citizen. In D. Kochenov (Ed.), *EU Citizenship and Federalism. The Role of Rights* (pp. 316–340). Cambridge University Press. https://doi.org/10.1017/9781139680714.014

Van den Brink, M. (2019). EU citizenship and (fundamental) rights: Empirical, normative, and conceptual problems. *European Law Journal*, 25(1), 21–36. https://doi.org/10.1111/eulj.12300

Varju, M. (2014). *European union human rights law: The dynamics of interpretation and context*. Edward Elgar Publishing. https://doi.org/10.4337/9781781951743

White, R. C. A. (2005). Free Movement, Equal Treatment, and Citizenship of the Union. *International and Comparative Law Quarterly*, 54(4), 885–905. https://doi.org/10.1093/iclq/lei041

Williams, A. T. (2009). Taking Values Seriously: Towards a Philosophy of EU Law. *Oxford Journal of Legal Studies*, 29(3), 549–577. https://doi.org/10.1093/ojls/gqp017

Wollenschläger, F. (2011). A New Fundamental Freedom beyond Market Integration: Union Citizenship and its Dynamics for Shifting the Economic Paradigm of European Integration. *European Law Journal*, 17(1), 1–34. https://doi.org/10.1111/j.1468-0386.2010.00536.x

Institutiones Administrationis – Journal of Administrative Sciences Vol. 6 (2026) No. 1, 41–57
DOI: 10.54201/iajas.210

Democratic Re-legitimation and its Limits in European Social Dialogue

Functional Representation, Institutional Balance and Regulatory Outsourcing in EU Governance

Nikolaos Gaitenidis*

* Postdoctoral Researcher, School of Political Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece.
e-mail: nikgait@yahoo.com

Abstract

This article examines the European Social Dialogue as a distinctive mechanism in EU governance, institutionalized under Articles 154 and 155 TFEU, which allows social partners to negotiate agreements that may become binding EU law. Challenging traditional views of it as neo-corporatism or a deliberative complement to democracy, the study reconceptualizes it as "managed regulatory outsourcing", where norm-drafting is delegated to organized interests but remains subject to Commission discretion, Council adoption, and limited parliamentary oversight. Drawing on theoretical frameworks of democratic legitimacy (input, output, throughput), institutional balance, and functional representation, the analysis highlights how this process redistributes authority without resolving the EU's democratic deficit. Key case law, such as *UEAPME* and *EPSU*, underscores executive gatekeeping and constitutional constraints. Post-*EPSU* developments, including the 2023 Council Recommendation, reflect procedural expansion amid a shift to soft-law instruments, reinforcing hybrid legitimacy but diffusing accountability. The article concludes that while social dialogue innovates participatory governance, it reconfigures rather than eliminates legitimacy tensions, emphasizing the need for stronger accountability in EU multilevel structures.

Keywords

European Social Dialogue, democratic legitimacy, institutional balance, functional representation, legitimacy, social partners

1 Introduction

The European Social Dialogue occupies a distinctive place in EU governance. Institutionalized under Articles 154 and 155 TFEU, it enables social partners, at cross-industry and sectoral levels, to conclude agreements that may become binding EU law through a Council decision. It is thus not merely a policy instrument but a formal law-making channel embedded in the Union's legislative framework and subject to institutional balance (Dorssemont & Van Malleghem, 2021).

 Published online: 06 May 2026

Its emergence marked a shift toward more flexible and non-hierarchical governance, reflecting the limits of traditional decision-making in addressing complex social demands and national diversity (Stephanou, 2006). As part of broader "new modes of governance", it combines soft law, private actor participation, and multi-level coordination (Gatto, 2006). Since its formalization in 1985 and consolidation through the Treaty of Amsterdam, it has been viewed as an alternative regulatory pathway (Trubek & Trubek, 2007). Linked to Article 28 of the Charter of Fundamental Rights, it has been interpreted either as supranational neo-corporatism (Bisson, 2022) or as contributing to EU democratic legitimacy and deliberative governance (Smismans, 2003; 2007a). However, the post-Lisbon constitutional order, the eurozone crisis, the European Pillar of Social Rights, and increasing politicization have reshaped EU governance and legitimacy debates (Schmidt, 2019; Kelemen, 2020). Contemporary scholarship views EU legitimacy as dynamic and structurally evolving (Bellamy, 2019; Schmidt, 2019), yet less attention has been paid to how Commission discretion affects the legitimacy implications of social partner involvement.

Based on the above, this article inquires how European Social Dialogue reshapes democratic legitimacy within the EU's constitutional framework, and whether it constitutes a form of democratic re-legitimation or rather a reconfiguration of the Union's democratic deficit. It argues that European Social Dialogue is best understood as a constitutionally managed form of regulatory outsourcing. While norm-drafting is delegated to organized interests, legislative authority remains contingent on Commission assessment and Council adoption, with limited parliamentary involvement. The system extends regulatory initiative beyond the ordinary legislative procedure without altering the institutional hierarchy, thereby reconfiguring rather than eliminating the EU's democratic deficit. As such, European Social Dialogue constitutes a mode of democratic re-legitimation with ambivalent effects: it broadens participation by incorporating organized socio-economic actors, yet Commission discretion and weak parliamentary oversight fragment accountability. Legitimacy is thus redistributed across electoral, functional, and executive channels, without necessarily being strengthened.

In terms of its limitations, I must note that this article does not engage in depth with broader debates on constitutional pluralism. While such approaches are relevant for analyzing tensions between EU law and national systems, the present study focuses on the internal institutional structure of European Social Dialogue and its implications for legitimacy within the EU framework. Adopting a doctrinal and conceptual approach, the paper analyses the legal framework governing European Social Dialogue, focusing on Articles 154–155 TFEU and relevant case law. The analysis is complemented by a conceptual examination of legitimacy and representation in EU governance. By combining doctrinal interpretation with theoretical analysis, the article evaluates how social dialogue operates within the EU's institutional structure and its implications for democratic legitimacy.

2 Theoretical Foundations: Representation, Legitimacy, and Institutional Balance

Understanding European Social Dialogue requires situating it within the EU's broader framework of democratic representation and legitimacy. Rather than relying on a single model, the EU combines multiple modes of representation that interact within a multilevel system (Benz, 2015). This section does not offer an exhaustive account of legitimacy theories, but focuses on those most relevant to the constitutional and governance dynamics of European Social Dialogue.

2.1 The "Legitimacy" Debate

At the European supranational level, Svetlozar (2004) distinguishes four sources of democratic legitimacy: (1) Input Legitimacy: The most direct form, derived from the European Parliament's elected representatives, citizen participation, and transparent decision-making processes. It forms the basis of the EU's representative democracy model. (2) Output Legitimacy: Legitimacy based on results, stemming from citizens' support due to the effectiveness of decision-making and implementation. (3) "Borrowed" Legitimacy: Indirect legitimacy transferred from the democratic institutions of member states to the EU level. The reference to national constitutional traditions is its most fundamental expression (Skandamis, 2006). (4) Constitutional Legitimacy: Legitimacy derived from European and international law and the rule of law principle (Grimm, 2017).

Throughput legitimacy (Schmidt, 2013; 2020) is especially useful for analyzing complex multi-level systems like the EU. It focuses on the quality of governance processes and includes normative criteria such as efficacy, accountability, transparency, inclusiveness, and openness. These ensure fair, responsive, and procedurally sound governance. Unlike input or output legitimacy, throughput legitimacy cannot easily compensate for weaknesses in the other two: strong procedures cannot fully offset low citizen engagement or poor results, and flawed processes can undermine trust in both participation and performance (Stephenson, 2023).

Since the post-Lisbon period, particularly after the eurozone crisis, EU governance has increasingly relied on executive power, policy coordination, and regulatory networks (Schmidt, 2019; Kelemen, 2020). The European Commission plays a central role in these processes by designing and managing participation and consultation, filtering and translating inputs into legislative proposals for the European Parliament and the Council. While this enhances policy coordination, it also concentrates discretion in unelected actors, reinforcing concerns about the EU's democratic deficit (Stephenson, 2023; Rangone, 2022).

Against this background, the key issue is not simply whether European Social Dialogue complements the ordinary legislative procedure, but how it reshapes authority within the EU's institutional framework. This can be assessed through its effects on input (participation), output (effectiveness), and throughput (procedural quality and accountability), which together form the analytical framework of this article.

2.2 The Institutional Balance Principle

While institutional balance is traditionally understood as ensuring compliance with the Treaty-based allocation of powers, it is analytically relevant for European Social Dialogue because it structures and limits the transformation of social partner agreements into binding EU law, particularly through Commission discretion. The principle requires each EU institution to act within its conferred competences and prohibits encroachment on others' prerogatives, as established in *Meroni v High Authority*[1] and reaffirmed in *Parliament v Council*[2] and Article 13(2) TEU. It operates both as a safeguard of institutional equilibrium and as a mechanism of adaptive evolution within the EU framework. Koutrakos (2019) further highlights its interaction with the

[1] Meroni & Co., Industrie Metallurgiche SpA v High Authority of the European Coal and Steel Community, Case 9/56, ECLI:EU:C:1958:7 (Court of Justice of the European Communities, 1958).

[2] European Parliament v Council of the European Communities, Case C-70/88, ECLI:EU:C:1990:217 (Court of Justice of the European Communities, 1990).

duty of sincere cooperation, whereby the Court enforces procedural fidelity while maintaining balanced inter-institutional dynamics, including in cases such as *Commission v Council*.[3]

In the context of European Social Dialogue, institutional balance preserves social partner autonomy under Article 152 TFEU while maintaining the Commission's monopoly of legislative initiative. Agreements concluded under Article 155 TFEU may be implemented autonomously or through a Council Directive, but legislative transformation remains subject to Commission discretion. This limits the capacity of social dialogue to generate binding EU law and aligns it more closely with structured participation than co-legislation (Carré & Steiert, 2022). More broadly, institutional balance enables executive filtering of participatory inputs, making Commission discretion decisive for whether negotiated outcomes acquire legal force. From a legitimacy perspective, this conditions the impact of functional participation on EU law-making and shapes the relationship between input and throughput legitimacy.

2.3 Electoral and Functional Representation

The democratic legitimacy of the EU rests on a dual foundation that combines electoral and functional representation, reflecting its complex multilevel character.

Electoral representation forms the traditional core of democratic legitimacy in the EU. Based on a territorial logic, authority derives from citizens via the European Parliament and from Member States in the Council (Cotter, 2025). This embodies representative democracy at both national and supranational levels. However, the EU's supranational structure has created persistent concerns about a "democratic deficit", including limited accountability, political distance from citizens, insufficient parliamentary empowerment, and weak voter-decision-making links (Svetlozar, 2004; Persson & Lindgren, 2008; Cotter, 2025). In a system of dispersed authority and multilevel governance, electoral representation remains essential but structurally incomplete.

Functional representation complements electoral representation by incorporating organized socio-economic actors into EU decision-making based on the interests and sectors they represent, rather than territorial constituencies (Knodt et al., 2011). The European Social Dialogue is a prime example: its legitimacy derives from organized membership, sectoral expertise, and bargaining capacity. Such mechanisms are often viewed as strengthening input legitimacy by directly involving affected stakeholders (Smismans, 2007a; Obradovic, 2006). However, functional representation faces challenges. The strength and inclusiveness of organized interests vary considerably across Member States due to differences in membership density, internal democracy, and bargaining power (Curtin & Meijer, 2006; Cárdenas Domínguez et al., 2025). It is thus mediated, sectoral, and uneven, lacking the universal accountability of electoral channels. As Piattoni (2011) notes, the EU features a hybrid system in which territorial (electoral) and functional logics coexist and intersect without clear hierarchical dominance. Functional inclusion, therefore, supplements electoral representation without replicating its broad-based accountability, resulting in a distinctive, and sometimes tension-filled, blend of representational forms in EU governance.

European Social Dialogue institutionalizes functional representation, but without replicating the accountability mechanisms of electoral representation, making it a central site for examining legitimacy trade-offs.

---

[3] Council of the European Union v. European Commission, Case C-660/13, ECLI:EU:C:2016:616 (Court of Justice of the European Union July 28, 2016).

2.4 Hybridization and Accountability Tension

The EU's multi-level governance model represents a hybrid that combines different types of legitimacy and both territorial and functional forms of representation. The European Social Dialogue exemplifies this intersection: it institutionalizes functional representation and appears to strengthen input legitimacy by embedding organized socio-economic actors in norm production. At the same time, it operates under executive coordination by the European Commission and the principle of institutional balance. The Commission retains significant evaluative discretion, especially in transforming negotiated agreements into binding EU law, as confirmed in *European Federation of Public Service Unions (EPSU) v Commission*.[4]

This hybridization creates structural tension. While functional participation expands access to decision-making, the Commission's executive power acts as a filter for turning agreements into legislation. Electoral representation remains limited, as the European Parliament does not exercise full co-legislative authority in this domain (Smismans, 2007a; Obradovic, 2006).

Kröger and Friedrich (2013) note that the EU lacks a clear hierarchy between electoral/territorial and functional channels of representation. When multiple modes coexist, democratic accountability and responsibility become diffused. The core tension is therefore not merely between participation and exclusion, but between participation and accountability. Inclusion of organized interests can enhance deliberative capacity, yet it also raises concerns about transparency and representativeness (Curtin & Meijer, 2006; Bellamy, 2019).

2.5 Social Dialogue as Managed Regulatory Outsourcing

Within this hybrid mode of legitimacy and representation, European Social Dialogue can be re-conceptualized as a form of managed regulatory outsourcing under institutional balance constraints.

It is "outsourced" because law-drafting in specific policy areas is delegated to organized interests outside the ordinary legislative procedure. It is "managed" because executive institutions retain decisive authority over legislative transformation. Under Article 155(2) TFEU, the Commission decides whether a negotiated agreement will be submitted for Council implementation. The Court has clarified that this discretion is substantive, not merely procedural (*EPSU v Commission*). The principle of institutional balance anchors institutional prerogatives in Treaty-based allocations (Koutrakos, 2019).

This model goes beyond traditional corporatism, which implies a redistribution of regulatory authority between the state and organized interests (Sargent, 1985). Instead, the Social Dialogue is embedded in the EU framework, where the European Commission holds ultimate executive power. It operates as a hierarchical yet multi-level system with institutionally centralized authority.

The result is a legitimacy model distributed across electoral, functional, and procedural channels. Social dialogue enhances input participation and may improve output effectiveness (Cattaneo, 2018), but it functions within executive-coordinated structures that limit its regulatory capacity. Democratic re-legitimation thus unfolds under constitutional constraints defined by institutional balance.

---

[4] European Federation of Public Service Unions (EPSU) v European Commission, Case C-928/19 P, ECLI:EU:C:2021:656 (Court of Justice of the European Union, 2021).

Taken together, these elements frame the European Social Dialogue not as a simple extension of democratic participation, but as a structured interaction between functional representation and executive control, where legitimacy gains come with accountability trade-offs. The central claim is that the Social Dialogue neither replaces electoral representation nor creates corporatist co-decision. It institutionalizes a hybrid form of regulatory production in which participatory inclusion and executive control coexist. Consequently, the EU's democratic deficit is not eliminated but re-configured through the redistribution of authority across overlapping governance modes.

3. The "Constitutional" Dynamics of the European Social Dialogue

European Social Dialogue is a Treaty-based procedure under Articles 154 and 155 TFEU. These provisions define the scope and limits of social partner involvement in EU law- and policy-making, reflecting a balanced allocation of roles among the Commission, the Council, and indirectly the European Parliament. This framework highlights the tension between participatory democracy and institutional hierarchy, with the Dialogue operating within the limits of EU social policy competence. Article 153 TFEU allows the Union to support and complement Member States' actions in specific areas, while excluding key domains such as pay and the right to strike. These constraints explain the reliance on social partners to negotiate norms where direct EU legislation is restricted, while reinforcing the Commission's and Council's role in ensuring agreements remain within Union competences (Obradovic, 2006).

At the European level, cross-sectoral social dialogue is conducted by organizations representing broad interests across the economy, labor market, and social policy. Owing to their general scope, these actors occupy a central position in EU governance and attract sustained attention from the EU institutions and academic scholarship. Representation is structured through cross-sectoral "umbrella" organizations, which aggregate numerous national and sectoral associations. The principal actors include the European Trade Union Confederation (ETUC), BusinessEurope (formerly UNICE), SMEunited (formerly UEAPME), and SGI Europe (formerly CEEP), alongside sectoral social partner organizations operating in specific industries (Gaitenidis, 2025).

3.1 Articles 154–155 TFEU : Structured Participation Under Limits

Article 154 TFEU imposes on the Commission the duty to promote consultation of social partners at the EU level and to facilitate dialogue between them. Before submitting proposals in the field of social policy, the Commission must consult social partners regarding both the possible direction and the content of contemplated action. Social partners may then express opinions or recommendations, and they may also inform the Commission of their wish to initiate negotiations under Article 155 TFEU (Bisson, 2022).

Article 155 TFEU provides that agreements concluded at the Union level may be implemented either: (1) In accordance with the procedures and practices of the social partners and Member States (autonomous implementation), or (2) In matters covered by Article 153 TFEU, at the joint request of the signatory parties, by a Council decision on a proposal from the Commission. This dual pathway grants social partners a distinctive role in norm production. They may negotiate contractual relations and agreements, and these may acquire binding EU law status. Yet this "legislative" potential is limited to the Commission's right of initiative and the Council's decision-making authority (Bisson, 2022).

Therefore, according to the Treaty, social partners may initiate negotiations, but institutional power ultimately returns to the "formal" EU institutions if no agreement is reached or if implementation through Council Decision is pursued. The autonomy of social partners is thus constitutionally limited.

The representativeness of social partners at the EU level cannot be taken for granted. Although cross-sectoral organizations aggregate national and sectoral actors, their membership density, internal governance, and bargaining capacity vary significantly across Member States, leading to uneven participation and influence. Sectoral imbalances further reinforce these disparities, as some industries are well organized while others remain fragmented or weakly represented (Gaitenidis, 2025; Kerckhofs et al., 2021; Larsson et al., 2025). These asymmetries are reflected in broader tensions within EU social policy, illustrated by *Viking*[5] and *Laval*,[6] as well as debates on the Minimum Wage Directive,[7] where national industrial relations traditions produced divergent responses to EU intervention (Ozols et al., 2025). Social dialogue, therefore, does not function as a uniformly representative channel, but reflects and reproduces structural differences across Member States and sectors.

3.2 Institutional Balance and Executive Gatekeeping

In the context of social dialogue, the principle of institutional balance operates as a constraint for the social partners' authority. Although the Commission is obliged to consult social partners, its right of legislative initiative is not displaced. In addition, Article 155 TFEU does not eliminate the Commission's authority in social policy. The power to initiate is shared, in procedural terms, but not transferred. Thus, the European Commission remains the primary "gatekeeper" of legislative transformation.

The executive character of this arrangement is reinforced by supervisory functions. Before proposing implementation by Council decision, the Commission and the Council may verify the representativeness of the signatory parties, the compatibility of the agreement with Union law, and its potential consequences. If collective representativeness is insufficient, implementation may be refused. This review capacity confirms that social partner autonomy operates under institutional oversight (Velluti, 2022).

The European Parliament's position is weaker in this process as, according to Article 155(2) TFEU, the European Parliament is informed, but does not exercise co-legislative authority. It is this procedural differentiation that distinguishes the European Social Dialogue from the ordinary legislative procedure and raises constitutional questions regarding democratic legitimacy. While the European Social Dialogue is often justified as a response to the democratic deficit, the minimal Parliamentary involvement complicates that claim (Bisson, 2022).

---

[5] International Transport Workers' Federation and Finnish Seamen's Union v Viking Line ABP and OÜ Viking Line Eesti, Case C-438/05, ECLI:EU:C:2007:772 (Court of Justice of the European Union, 2007).

[6] Laval un Partneri Ltd v Svenska Byggnadsarbetareförbundet, Svenska Byggnadsarbetareförbundets avdelning 1, Byggettan and Svenska Elektrikerförbundet, Case C-341/05, ECLI:EU:C:2007:809 (Court of Justice of the European Union, 2007).

[7] Directive (EU) 2022/2041 of the European Parliament and of the Council of 19 October 2022 on adequate minimum wages in the European Union, Pub. L. No. L 275/33 (2022). Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022L2041

3.3 The UEAPME Case: Procedural Recognition Without Autonomy

The Court's case-law clarifies this structure in *UEAPME v Council*.[8] The General Court acknowledged that European Social Partners occupy a specific procedural position within the Union's decision-making framework. According to the Court, consultation under Articles 154–155 TFEU is not discretionary but mandatory.

At the same time, the Court did not elevate social partners to autonomous constitutional actors. Representativeness must be demonstrated in relation to the subject matter, and "sufficient collective representativeness" may suffice even where other organizations could claim participation (Smismans, 2007a; Franssen & Jacobs, 1998). The Court thus suggested that participation by adequately representative organizations could satisfy the principle of participation alongside parliamentary representation (Franssen & Jacobs, 1998). However, this recognition remained procedural rather than constitutive: social partners gained constitutional visibility, but not independent law-making authority, as their role derives from and remains bounded by the Treaties (Franssen & Jacobs, 1998).

UEAPME therefore consolidated procedural status without granting autonomy. Social dialogue became constitutionally recognized, but not constitutionally self-standing. In legitimacy terms, UEAPME confirms that participation in European Social Dialogue is structured through functional representation rather than electoral accountability. While the Court recognizes the procedural role of sufficiently representative organizations, it does not establish a direct link between participation and democratic representation in the electoral sense. This reinforces the mediated and indirect nature of input legitimacy within the social dialogue framework.

3.4 The *EPSU* Case: Executive "Gatekeeping" Constitutionalized

The structural limits of this arrangement became explicit in *EPSU v Commission*. The Court held that the Commission is not obliged to submit a proposal to the Council under Article 155(2) TFEU, even where social partners jointly request it. The Commission retains its exclusive right of initiative under Article 17 TEU and enjoys broad discretion in deciding whether to proceed, including assessing political, economic, and social considerations. This discretion extends beyond verifying representativeness and legality to evaluating the appropriateness of EU action, while judicial review remains limited (Article 296 TFEU).

The judgment confirms that European Social Dialogue does not entail a shared legislative initiative, but operates as a request mechanism, with the transformation of negotiated outcomes dependent on executive approval (Velluti, 2022). Carré and Steiert (2022) interpret *EPSU* as a constitutional redefinition of social dialogue, subsuming Article 155(2) TFEU under the general logic of Commission initiative and rejecting its characterization as a form of quasi-co-decision. The ruling thus reaffirms institutional balance by anchoring social dialogue within an executive hierarchy: participatory inclusion does not displace Commission authority.

The constitutional structure of European Social Dialogue is therefore characterized by duality. While Articles 154 and 155 TFEU institutionalize functional participation and grant social partners a structured role in norm production, institutional balance preserves executive initiative and Council authority while limiting parliamentary involvement. Social dialogue is

[8] Union Européenne de l'artisanat et des petites et moyennes entreprises (UEAPME) v Council of the European Union, Case T-135/96, ECLI:EU:T:1998:128 (Court of First Instance, 1998).

thus constitutionally recognized but hierarchically constrained. From a legitimacy perspective, the *EPSU* case highlights the limits of input-based participation. Although social partners contribute to norm-drafting, Commission discretion constrains their influence on legislative outcomes, making throughput legitimacy decisive as executive control determines whether participatory inputs translate into binding law.

4 Towards a Managed Regulatory Outsourcing Mode of Governance

The constitutional architecture analyzed above makes clear that European Social Dialogue cannot be accurately described either as a form of co-decision or just as a form of deliberative complement to representative democracy, giving rise to several competing potential interpretations.

4.1 Beyond Corporatism

The European Social Dialogue has often been interpreted through a corporatist lens. The institutionalized negotiation between employers and trade unions, the possibility of transforming agreements into binding law, and the relative marginalization of parliamentary mediation have encouraged characterizations of the system as "neo-corporatist" or as a form of "governance of private interests" (Schmitter, 1994; Smismans, 2007a; Bisson, 2022). However, classical corporatism presupposes a horizontal redistribution of regulatory authority between the state and organized interests. In such models, social partners exercise autonomous norm-producing power alongside or in partnership with public institutions (Molina & Rhodes, 2002; Dobbins & Labanino, 2023). The EU framework does not correspond to this structure.

As demonstrated by the *EPSU* case, social partner agreements do not bind the European institutions automatically. Legislative transformation remains contingent upon the European Commission's exclusive power of initiative and the Council's adoption. This hierarchical oversight is structurally inconsistent with corporatist autonomy, as the European Social Dialogue therefore does not displace authority but rather operates within Treaty-based limitations (Delfino, 2020).

4.2 Beyond Deliberative Democracy

A second interpretation frames the European Social Dialogue as a deliberative complement to representative democracy. In this respect, it enhances input legitimacy by incorporating organized socio-economic actors into law and policy-making (Smismans, 2007a; Heidbreder, 2012). The Court has suggested that participation through sufficiently representative social partners may satisfy the principle of participation alongside parliamentary representation (Franssen & Jacobs, 1998).

Yet this equivalence is normatively and institutionally unclear. Unlike the European Parliament, social partners are not elected bodies. Their internal democratic structures vary, their membership density differs across Member States, and sectoral asymmetries persist (Smismans, 2007b; Cárdenas Domínguez et al., 2025). As a result, functional representation is indirect.

Furthermore, the transformation phase excludes Parliament from full co-legislative participation. When agreements are implemented by Council Decision under Article 155(2) TFEU, Parliament is informed but does not exercise equal legislative authority like during

the ordinary legislative procedure. This procedural asymmetry complicates the claim that the European Social Dialogue simply supplements electoral representation. Instead, it partially displaces parliamentary mediation, and participation may be expanded, but under hierarchical conditions.

4.3 The Limits of the Co-Legislative Narrative

The idea that the European Social Dialogue represented an "alternative legislative procedure" has been decisively weakened by the *EPSU* case. As Carré and Steiert (2022) argue, the Court effectively redefined Article 155(2) TFEU as a request mechanism comparable to Articles 225 and 241 TFEU, reaffirming the Commission's initiative monopoly. Social partners may trigger a procedural opportunity, but they cannot compel legislative action.

In this respect, the European Social Dialogue differs fundamentally from the ordinary legislative procedure. In co-decision, the European Parliament and the Council jointly exercise legislative authority. In the European Social Dialogue, substantive drafting may originate outside the institutional triangle, but legal force rests on the European Commission and the Council. The view of the European Social Dialogue as "co-legislator" is therefore obscured.

4.4 Defining Managed Regulatory Outsourcing

The European Social Dialogue can be better conceptualized as managed regulatory outsourcing. While this framing draws on longstanding analyses in EU governance literature, such as the hybrid hard/soft law mechanisms in new modes of governance (Trubek & Trubek, 2007), the "shadow of hierarchy" that conditions social partners' autonomy through the threat of unilateral Commission action (Smismans, 2008), and post-*EPSU* critiques of "corporatism without autonomy" that highlight the subordination of negotiated outcomes to executive discretion (Carré & Steiert, 2022; Dorssemont & Van Malleghem, 2021), it offers a distinct lens.

More specifically, managed regulatory outsourcing foregrounds the post-Lisbon and post-*EPSU* constitutional reality by emphasizing the deliberate, hierarchical relocation of substantive norm-drafting to functional actors under irreversible executive supervision, the structural sidelining of parliamentary co-legislation in the binding pathway, and the resulting redistribution of legitimacy across fragmented channels rather than its resolution.

This model consists of five interrelated features. First, substantive norm-drafting is delegated to organized interests. Social partners negotiate framework agreements, draft regulatory standards, and shape policy content outside the ordinary legislative procedure (Smismans, 2007b; Trubek & Trubek, 2007). The production of Council Directive 96/34/EC of 3 June 1996 on the framework agreement on parental leave concluded by UNICE, CEEP and the ETUC, Council Directive 97/81/EC of 15 December 1997 concerning the Framework Agreement on part-time work concluded by UNICE, CEEP and the ETUC, and Council Directive 1999/70/EC of 28 June 1999 concerning the framework agreement on fixed-term work concluded by ETUC, UNICE, and CEEP illustrates this delegation of drafting capacity. Second, the Commission retains the monopoly of legislative initiative. Even where social partners conclude an agreement and request implementation, the Commission decides whether to submit a proposal. The *EPSU* judgment confirms that this discretion is substantive and constitutionally protected, extending beyond procedural checks to assessments of political opportunity, subsidiarity, proportionality, and the Union's general interest. Third, the Council exercises formal adoption authority. The binding force derives not from social partner negotiation alone, but from institutional

endorsement through Council Decision. Fourth, the European Parliament is structurally limited in the transformation phase. Unlike in the ordinary legislative procedure, Parliament does not act as co-legislator when agreements are implemented under Article 155(2) TFEU. And fifth, accountability becomes diffused. Norms are drafted by organized interests, filtered by the executive, adopted by the Council, and implemented through multi-level structures often dependent on national actors. Responsibility is distributed across functional, executive, and territorial channels without a single, clearly identifiable center of democratic control.

Outsourcing, in this sense, does not imply abdication of public authority. It describes the relocation of substantive drafting capacity to actors external to the institutional legislature, under conditions defined and supervised by executive institutions. It is "managed" because the delegation is conditional, reversible, and embedded within hierarchical oversight, ensuring that participatory expansion occurs without displacing the EU's institutional balance or fully alleviating the democratic deficit.

4.5 Post-Crisis Governance and the Soft Law Shift

The post-2008 transformation of EU governance reinforces this interpretation. The production of binding legislative agreements has declined, while autonomous and non-binding instruments have become more prominent (Bisson, 2022; Cárdenas Domínguez et al., 2025). The rise of coordination mechanisms, including the European Semester, has embedded social dialogue within recurring executive-led policy cycles (Pecinovsky, 2018; Schmidt, 2019).

Soft-law instruments, process-oriented texts, joint opinions, and monitoring frameworks increasingly define the practical output of social dialogue. From a functional perspective, such instruments may foster learning, trust-building, and horizontal coordination (Larsson et al., 2020). Yet they further distance regulatory production from formal parliamentary oversight.

The shift from binding directives to autonomous implementation and coordination-based governance illustrates a broader move from shared legislative production toward regulated participation within executive steering frameworks (Bogg & Dukes, 2013; Cárdenas Domínguez et al., 2025). As such, social dialogue is being integrated more deeply into managed governance structures. And this trajectory confirms that the central dynamic is not the expansion of co-decision but the consolidation of conditional delegation.

4.6 Legitimacy Redistribution Through Outsourcing

Conceptualizing social dialogue as managed regulatory outsourcing clarifies its democratic ambivalence. Input legitimacy may be enhanced through functional inclusion. Output legitimacy may benefit from negotiated compliance and stakeholder ownership (Cattaneo, 2018). However, throughput legitimacy becomes dependent on executive filtering and multi-level coordination. Authority is redistributed: (1) From Parliament to organized interests in drafting, (2) From public consultation to negotiated compromise, (3) From visible legislative arenas to procedural frameworks shaped by the Commission.

The democratic deficit is therefore not resolved but re-allocated. The European Social Dialogue institutionalizes a hybrid governance structure in which substantive regulatory work is outsourced to social partners under the limits of executive control and institutional hierarchy. In this configuration, participation expands, but autonomy remains conditional on institutional balance.

5 Post-*EPSU* Governance Reinforcement: Procedural Expansion Without Power Rebalancing

The period following the *EPSU* judgement indicates an effort to consolidate social dialogue within the EU's broader policy architecture. It is noteworthy, however, that post-*EPSU* reforms increase procedural participation or political visibility, but they do not recalibrate constitutional authority.

A central milestone is the Council Recommendation of 12 June 2023 on strengthening social dialogue in the EU.[9] The Recommendation calls for improved conditions for collective bargaining, systematic and timely involvement of social partners in employment and social policy, and enhanced access to information and institutional support. As Hromadska (2025) notes, it frames social dialogue broadly across bipartite and tripartite formats and emphasizes its relevance at multiple governance levels.

The same pattern is visible in the Val Duchesse Social Partner Summit of January 2024 and the subsequent Tripartite Declaration for a Thriving European Social Dialogue. The Declaration situates social partners at the center of responses to labor shortages, skills gaps, and the green and digital transitions. It envisages strengthened monitoring practices, enhanced coordination, and the development of a Pact for European Social Dialogue, formally launched in 2025. These initiatives elevate social dialogue politically and symbolically. They integrate social partners more explicitly into strategic Union priorities (Cárdenas Domínguez et al., 2025; Galetto & Neri, 2026).

The growing integration of social dialogue into the European Semester further illustrates this trajectory. Since the post-crisis consolidation of economic governance, social partners have become increasingly involved in consultation, monitoring, and reporting linked to Country-Specific Recommendations and employment coordination (Cárdenas Domínguez et al., 2025; Guisset et al., 2025). The 2024 Action Plan on labor and skills shortages confirms this embedding within recurring Commission-led governance cycles (European Commission, 2024).

A similar pattern emerges in the European Pillar of Social Rights. Principle 8 emphasizes consultation, collective bargaining, and worker participation, while the accompanying Action Plan highlights the role of social dialogue in managing restructuring, economic transitions, and changing work patterns within the EU's broader regulatory framework (European Commission, 2021).

However, these developments remain non-binding and do not alter Articles 154–155 TFEU. Political elevation thus does not entail constitutional repositioning. At the same time, empirical research points to a shift toward non-binding and autonomous instruments, with fewer legislative agreements under Article 155(2) and greater reliance on soft-law coordination (Bisson, 2022; Cárdenas Domínguez et al., 2025). While such forms may enhance learning and coordination (Larsson et al., 2020), the decline in binding outcomes remains constitutionally significant.

European Social Dialogue has therefore entered a phase of procedural expansion without power rebalancing: its institutional visibility has increased, but its constitutional position has not. This development confirms the article's core claim that social dialogue operates as a form of managed regulatory outsourcing, in which participatory inclusion expands while executive authority remains predominant.

---

[9] Council Recommendation (EU) 2023/C 185/01 of 12 June 2023 on strengthening social dialogue in the European Union, OJ 2023 C 185/1.

6 Democratic Legitimacy Revisited: Re-legitimation or Accountability Displacement?

The analysis has demonstrated that European Social Dialogue is constitutionally embedded, procedurally reinforced, and structurally managed within an executive-centered governance framework. The remaining question is normative: does this configuration strengthen democratic legitimacy, or does it merely relocate its tensions? To answer this, it is useful to reconsider the legitimacy effects of social dialogue across three dimensions: input, output, and throughput.

6.1 Input Gains

The European Social Dialogue clearly enhances input legitimacy in a functional sense. In policy areas directly affecting socio-economic actors, such inclusion can increase the responsiveness of regulation and ensure that expertise and sectoral knowledge inform legislative design. Particularly in complex labor market contexts, negotiated drafting may produce more context-sensitive norms than purely parliamentary deliberation. Yet input gains remain uneven. Social partners derive their authority from organizational membership structures that vary across Member States and sectors (Curtin & Meijer, 2006; Cárdenas Domínguez et al., 2025). Functional representation is therefore indirect. Moreover, parliamentary participation in the Article 155(2) process is limited, meaning that electoral channels are not reinforced but partially displaced. Input legitimacy is therefore broadened, but not democratically equalized.

6.2 Output Gains

Social dialogue may also contribute to output legitimacy. Negotiated agreements can foster compliance, reduce implementation conflicts, and enhance ownership among stakeholders. Where actors become co-authors of regulatory standards, regulatory acceptance and effectiveness may increase (Cattaneo, 2018). Empirical studies further suggest that effectiveness is often measured by participants in terms of trust-building, coordination, and learning rather than exclusively through binding legislative outcomes (Larsson et al., 2020).

However, the post-crisis decline in legislative transformation and the increasing reliance on autonomous or soft-law instruments (Bisson, 2022; Cárdenas Domínguez et al., 2025) demonstrate that output effectiveness becomes more diffuse. Output gains are therefore plausible, but they are structurally contingent and unevenly distributed.

6.3 Throughput Gains and Accountability Risks: Re-legitimation or Re-distribution?

The most significant legitimacy tension arises at the level of throughput legitimacy. As demonstrated above, the European Social Dialogue operates within a governance structure characterized by institutional hierarchy. This layered configuration diffuses accountability. Regulatory content originates in private negotiations, is filtered by executive assessment, formally endorsed by the Council, and implemented through multi-level arrangements.

The European Social Dialogue represents a project of democratic re-legitimation. It expands participatory inclusion, institutionalizes functional representation, and may enhance regulatory effectiveness. Its procedural reinforcement in the post-*EPSU* period confirms its political centrality. In this respect, accountability becomes dispersed across functional, executive, and territorial channels.

The resulting chain of responsibility is neither purely electoral nor purely functional but fragmented, resulting in the democratic deficit not disappearing or being reduced. In effect, the European Social Dialogue re-distributes the democratic deficit from "participation deficit" to "accountability slippage" and diffused responsibility.

7 Conclusion: Innovation Under Constraint

The European Social Dialogue is often portrayed as a democratic corrective within the EU's constitutional order, and I have argued that it rather constitutes a distinct governance mode instead, something which I have described as managed regulatory outsourcing. Under Articles 154 and 155 TFEU, norm-drafting is delegated to organized socio-economic actors, yet legislative transformation remains dependent on executive initiative and institutional balance. In this framework, the Commission retains discretion, the Council formalizes adoption, and parliamentary participation is structurally limited. The result is neither privatized regulation nor shared legislative authority, but a hybrid architecture in which participation expands within constitutional constraints. And this arrangement reflects a deeper constitutional ambivalence: while the Social Dialogue strengthens functional representation and may enhance regulatory responsiveness, it preserves hierarchical control and diffuses accountability. The democratic deficit is therefore not eliminated but reconfigured, i.e., participation widens, yet responsibility fragments.

My historical analysis showed that the Social Dialogue has evolved cyclically, alternating between phases of autonomy and executive steering. The post-*EPSU* trajectory confirms a shift toward the latter. Although institutional visibility and procedural integration have increased, constitutional authority has not been reallocated. As such, reinforcement has occurred within existing limits rather than through structural rebalancing.

More broadly speaking, the Social Dialogue illustrates a transformation in EU governance. As executive coordination and multilevel steering expand, participatory mechanisms are increasingly embedded in managed frameworks rather than autonomous democratic processes. The future of EU legitimacy may thus depend less on expanding participation than on strengthening accompanying accountability structures. Innovation, in this context, operates under constraint. The European Social Dialogue exemplifies the Union's experimentation with new forms of representation while remaining anchored in a constitutional order that prioritizes institutional balance and executive initiative. Its central challenge is not merely to deepen participation, but to align it with coherent and transparent democratic responsibility.

Accordingly, the European Social Dialogue embodies, I believe, both democratic innovation and structural ambivalence, revealing the promise and tensions of hybrid legitimacy within the EU constitutional framework.

References

Bellamy, R. (2019). *A Republican Europe of States. Cosmopolitanism, Intergovernmentalism and Democracy in the EU*. Cambridge University Press. https://doi.org/10.1017/9781139136303

Benz, A. (2015). Multilevel Governance in the European Union: Loosely Coupled Arenas of Representation, Participation, and Accountability. In S. Piattoni (Ed.), *The European Union: Democratic Principles and Institutional Architectures in Times of Crisis* (pp. 201–218). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198716273.003.0010

Bisson, L. S. (2022). European social dialogue: History, characteristics, and perspectives. *Herald of the Russian Academy of Sciences*, 92(Suppl. 7), S660–S666. https://doi.org/10.1134/S1019331622130147

Bogg, A., & Dukes, R. (2013). The European Social Dialogue: From Autonomy to Here. In N. Countouris, & M. Freedland (Eds.), *Resocialising Europe in a Time of Crisis* (pp. 466–492). Cambridge University Press. https://doi.org/10.1017/cbo9781107300736.028

Cárdenas Domínguez, F., Fernández García, M., & Molinero Gerbeau, Y. (2025). Revisiting European social dialogue: A systematic literature review. *Open Research Europe*, 5, 309. https://doi.org/10.12688/openreseurope.21020.1

Carré, P., & Steiert, M. (2022). Social Europe without social dialogue: Decision of the Court of Justice of the European Union in C-928/19 P European Federation of Public Service Unions. *European Constitutional Law Review*, 18(2), 315–333. https://doi.org/10.1017/S1574019622000177

Cattaneo, A. (2018). Social dialogue and the regulatory power of governing bodies. *The International Sports Law Journal*, 17(3–4), 119–127. https://doi.org/10.1007/s40318-018-0122-2

Cotter, J. (2025). Democracy Manifest? Ensuring the EU Legislature's Democratic Legitimacy in the Face of National-Level Autocratisation. *European Papers*, 10(2), 489–515. https://doi.org/10.15166/2499-8249/842

Curtin, D., & Meijer, A. J. (2006). Does transparency strengthen legitimacy? A critical analysis of European Union policy documents. *Information Polity*, 11(2), 109–122. https://doi.org/10.3233/ip-2006-0091

Delfino, M. (2020). The reinterpretation of the principle of horizontal subsidiarity in European social law. *Biblioteca''20 Maggio'* (pp. 111–119). Vol. 1. University of Catania. Online: https://tinyurl.com/4pbffpyn

Dobbins, M., & Labanino, R. P. (2023). Corporatism and neo-corporatism. In M. Grasso, & M. Giugni (Eds.), *Elgar Encyclopedia of Political Sociology* (pp. 111–114). Edward Elgar. https://doi.org/10.4337/9781803921235.00037

Dorssemont, F., & Van Malleghem, P.-A. (2021, September 21). *Sounding the death knell for EU social dialogue?: The EPSU judgment as a turning point for EU labour law*. Verfassungsblog. https://doi.org/10.17176/20210921-205411-0

European Commission. (2021). *The European Pillar of Social Rights Action Plan*. Publications Office of the European Union. https://doi.org/10.2767/45

European Commission. (2024, March 20). *Labour and skills shortages in the EU: An action plan (COM(2024) 1507).* European Commission. Online: https://ec.europa.eu/commission/presscorner/detail/en/ip_24_1507

Franssen, E., & Jacobs, A. (1998). The Question of Representativity in the European Social Dialogue. *Common Market Law Review*, 35(6), 1295–1312. https://doi.org/10.54648/199393

Gaitenidis, N. (2025). *Ευρωπαϊκή διακυβέρνηση και κοινωνία των πολιτών*. [European Governance and Civil Society] Papazisis Publishers.

Galetto, M., & Neri, S. (2026). The new European sectoral social dialogue for social services: The "bite" challenge, the coordination potential and the future of care. *European Journal of Industrial Relations.* Advance online publication. https://doi.org/10.1177/09596801261423744

Gatto, A. (2006). Governance in the European Union: A legal perspective. *Columbia Journal of European Law*, 12(2), 487–516. Online: https://tinyurl.com/y4e5w6vr

Grimm, D. (2017). *The Constitution of European Democracy*. Oxford University Press. https://doi.org/10.1093/oso/9780198805120.001.0001

Guisset, A., Lenaerts, K., & Vangeel, N. (2025). *How the EU shapes national social dialogue: A qualitative analysis of social partner" involvement in the Recovery and Resilience Plans and Territorial Just Transition Plans*. HIVA– KU Leuven, Research Institute for Work and Society. Online: https://tinyurl.com/yj273p2c

Heidbreder, E. G. (2012). Civil society participation in EU governance. *Living Reviews in European Governance*, 7(2), 5–42. https://doi.org/10.12942/lreg-2012-2

Hromadska, N. (2025). European standards of social dialogue: Challenges and prospects of implementation in Ukraine. *Public Administration and Regional Development*, 28, 449–470. https://doi.org/10.34132/pard2025.28.07

Kelemen, R. D. (2020). The European Union's authoritarian equilibrium. *Journal of European Public Policy*, 27(3), 481–499. https://doi.org/10.1080/13501763.2020.1712455

Kerckhofs, P., Krokovay N., & Gustafsson, A.-K. (2021, October 27), *Representativeness of the European social partner organisations: Live performance sector*. Eurofound. Online: https://tinyurl.com/yvh44ynk

Knodt, M., Greenwood, J., & Quittkat, C. (2011). Territorial and Functional Interest Representation in EU Governance. *Journal of European Integration*, 33(4), 349–367. https://doi.org/10.1080/07036337.2011.582279

Koutrakos, P. (2019). Institutional balance and sincere cooperation in treaty-making under EU law. *International and Comparative Law Quarterly*, 68(1), 1–33. https://doi.org/10.1017/S0020589318000350

Kröger, S., & Friedrich, D. (2013). Democratic representation in the EU: Two kinds of subjectivity. *Journal of European Public Policy*, 20(2), 171–189. https://doi.org/10.1080/13501763.2013.746112

Larsson, B., Galetto, M., Weber, S., Bechter, B., & Prosser, T. (2020). What's the point of European sectoral social dialogue? Effectiveness and polycontexturality in the hospital and metal sectors. *Industrial Relations Journal*, 51(5), 410–426. https://doi.org/10.1111/irj.12305

Larsson, B., Heffernan, M., & Trif, A. (2025). *Trust in relations between unions and employers in Europe: Final report*. Online: https://tinyurl.com/4v6zje3j

Molina, O., & Rhodes, M. (2002). Corporatism: The past, present, and future of a concept. *Annual Review of Political Science*, 5, 305–331. https://doi.org/10.1146/annurev.polisci.5.112701.184858

Obradovic, D. (2006). Civil Society and the Social Dialogue in European Governance. *Yearbook of European Law*, 24(1), 261–329. https://doi.org/10.1093/yel/24.1.261

Ozols, E., Hristov, S. I., & Paster, T. (2025). Unions divided? Trade union attitudes towards the European Union's Directive on Adequate Minimum Wages. *Economic and Industrial Democracy*, 46(2), 372–396. https://doi.org/10.1177/0143831X241245007

Pecinovsky, P. (2018). EU economic governance and the right to collective bargaining: Part I. Standard and extreme governance and the indicators and limits of the right to collective bargaining. *European Labour Law Journal*, 9(4), 374–388. https://doi.org/10.1177/2031952518817567

Persson, T., & Lindgren, K. O. (2008). *Input and output legitimacy: Synergy or trade-off?* Paper presented at the Swedish Political Science Association Meeting, Uppsala University.

Piattoni, S. (2011). The Problematic Coexistence of Functional and Territorial Representation in the EU. *Journal of European Integration*, 33(4), 369–384. https://doi.org/10.1080/07036337.2011.579747

Rangone, N. (2022). Improving consultation to ensure the European Union's democratic legitimacy: From traditional procedural requirements to behavioural insights. *European Law Journal*, 28(4–6), 154–171. https://doi.org/10.1111/eulj.12439

Sargent, J. A. (1985). Corporatism and the European Community. In W. Grant (Ed.), *The political economy of corporatism* (pp. 229–253). Macmillan. https://doi.org/10.1007/978-1-349-18041-7_9

Schmidt, V. A. (2013). Democracy and Legitimacy in the European Union Revisited: Input, Output and 'Throughput'. *Political Studies*, 61(1), 2–22. https://doi.org/10.1111/j.1467-9248.2012.00962.x

Schmidt, V. A. (2019). Politicization in the EU: Between national politics and EU political dynamics. *Journal of European Public Policy*, 26(7), 1018–1036. https://doi.org/10.1080/13501763.2019.1619189

Schmidt, V. A. (2020). Conceptualizing legitimacy: Input, output, and throughput. In *Europe's crisis of legitimacy: Governing by rules and ruling by numbers in the Eurozone* (pp. 25–55). Oxford University Press. https://doi.org/10.1093/oso/9780198797050.003.0002

Schmitter, P. C. (1994). The Future Euro-Polity and its Impact upon Private Interest Governance within Member-States. *Droit et Société*, 28, 659–675. Online: https://tinyurl.com/4ff6dttk

Skandamis, N. (2006). *Το Παραδειγμα της Ευρωπαϊκής Διακυβερνησης: Μεταξυ Κυριαρχιας και Αγορας. [The example of European governance: Between sovereignty and market]* Ant. N. Sakkoulas.

Smismans, S. (2003). European Civil Society: Shaped by Discourses and Institutional Interests. *European Law Journal*, 9(4), 473–495. https://doi.org/10.1111/1468-0386.00187

Smismans, S. (2007a). The European social dialogue between constitutional and labour law. *European Law Review*, 32(3), 341–364.Online: https://orca.cardiff.ac.uk/id/eprint/3000/

Smismans, S. (2007b). The Shadow of Hierarchy over Self-regulation in the European Union: The Case of the European Social Dialogue. In J.-C. Graz, & A. Nolke (Eds.), *Transnational Private Governance and its Limits*. Routledge. https://doi.org/10.4324/9780203939338

Smismans, S. (2008). The European social dialogue in the shadow of hierarchy. *Journal of Public Policy*, 28(1), 161–180. https://doi.org/10.1017/s0143814x08000822

Stephanou, C. A. (2006). La nouvelle gouvernance européenne: bilan et perspectives. *Cahiers de droit européen*, 42(5–6), 581–627.

Stephenson, P. (2023). Exploring the Throughput Legitimacy of European Union Policy Evaluation: Challenges to Transparency and Inclusiveness in the European Commission's Consultation Procedures and the Implications for Risk Regulation. *European Journal of Risk Regulation*, 14(2), 351–370. https://doi.org/10.1017/err.2023.33

Svetlozar, A. (2004). *The EU "crisis of legitimacy" revisited: Concepts, causes, and possible consequences for European politics and citizenry*. [Unpublished manuscript].

Trubek, M., & Trubek, L. G. (2007). New Governance and Legal Regulation: Complementarity, Rivalry or Transformation. *Columbia Journal of European Law*, 13, 539–564. Online: https://tinyurl.com/5ha9ce43

Velluti, S. (2022). The European Social Dialogue as a source of EU legal acts following EPSU: Collective bargaining and industrial relations get lost in translation. *Common Market Law Review*, 59(3), 871–888. https://doi.org/10.54648/cola2022055

Institutiones Administrationis – Journal of Administrative Sciences Vol. 6 (2026) No. 1, 58–70
DOI: 10.54201/iajas.183

Improving Social Welfare through Tax Incentives for Philanthropy

Sangkyun Park*

* Retired economist. The author has served as an economist at the U.S. Office of Management and Budget (Washington DC, USA) and Federal Reserve Banks (New York, New York and Saint Louis, Missouri, USA). E-mail: sangkyun99@gmail.com

Abstract

To be justified, charity tax incentives should improve social welfare. This paper illustrates the importance of allocating public funds to worthier causes in evaluating the effectiveness of charity tax incentives and proposes a tax-incentive scheme (new scheme) that can substantially improve the allocation of funds. The new scheme consists of a refundable equal-percentage tax credit for all donors, a tax on the receipts of nonprofit organizations (NPOs), and a tax exemption for NPOs based on the number of donors. Tying the tax exemption to the number of donors should improve the allocation of public funds, provided that worthier causes attract more donors. In addition, the new scheme can increase philanthropic contributions, improve the operational efficiency of NPOs, and produce more intangible benefits, such as social cohesion and giving culture. The new scheme is also fairer in that it treats low-income and high-income donors equally.

Keywords

philanthropy, charitable giving, tax expenditures, tax incentives

1 Introduction

Many countries provide tax incentives for philanthropic activities (charity tax incentives). Since philanthropy is socially desirable, it is hard to argue against promoting it. That does not automatically justify charity tax incentives, however. To be justified, charity tax incentives should improve social welfare by promoting worthy causes (public goods) effectively. The literature on charity tax incentives loosely refers to activities producing social benefits as public goods.

To evaluate the social benefits of charity tax incentives, one needs to consider several factors: the total amount of spending on public goods (social spending), the composition of social spending, the operational efficiency of nonprofit organizations (NPOs) receiving charitable contributions relative to that of governments, and intangible benefits (e.g., strengthening social cohesion and promoting sharing culture).

The amount of social spending depends on the responsiveness of private contributions to charity tax incentives. Charity tax incentives increase private contributions but reduce tax receipts (incur tax expenditures). To increase the amount of social spending, the increase in

 Published online: 09 April 2026

private contributions must be larger than the decrease in tax receipts (treasury efficient). The public goods chosen by NPOs are not the same as those chosen by governments. Holding the amount of spending and other factors constant, NPOs must choose worthier public goods to improve social welfare. The operational efficiency of NPOs affects the net amount of social spending. Widespread philanthropic activities might produce intangible benefits.

It is hard to determine either conceptually or empirically whether charity tax incentives improve social welfare. The responsiveness of private contributions to the tax incentives is an empirical issue, but the empirical estimates vary widely across studies (Han et al., 2024). The worthiness of each public good is subjective and can hardly be quantified. The complexity of government operations and a large number of NPOs make it difficult to compare the operational efficiencies of governments and NPOs. By definition, intangibles are not measurable.

The contributions of this paper are highlighting the importance of allocating social spending to worthier public goods (allocational efficiency) and proposing ways to improve the effectiveness of charity tax incentives (new scheme). Despite its importance, previous studies have not paid much attention to the allocational efficiency of charity tax incentives. This paper illustrates that the effect of charity tax incentives on allocational efficiency can be much more important than treasury efficiency when the social benefit differs across public goods and the marginal benefit from each public good decreases with the amount of spending on it.

Although it is hard to tell whether or not charity tax incentives improve social welfare, it is more manageable and practical to explore how to improve the effectiveness of charity tax incentives. With a careful design, the government can substantially increase the effect of charity tax incentives on social welfare and make the tax breaks fair. The key design principle adopted in this paper is to utilize the private sector's strength in allocating resources, that is, to rely on market mechanisms. The new scheme has the following features: a refundable equal-percentage tax credit for all donors (uniform tax credit), a tax on the receipts of NPOs, and a tax exemption for NPOs based on the number of donors. The uniform tax credit would be fair in that it treats low-income and high-income donors equally. The tax on NPOs would enable the government to control the tax expenditure. Tying the tax exemption to the number of donors should improve the allocation of social spending, provided that worthier public goods attract more donors, and more strongly urge NPOs to improve operational efficiency. The new scheme might also produce more intangible benefits by increasing the number of donors.

The rest of this paper is organized as follows. The next section reviews the arguments for and against charity tax incentives. Section 3 shows that allocating funds to worthier public goods can be more important than raising a larger amount of funds. Section 4 presents a new tax-incentive scheme that can more effectively improve social welfare, and Section 5 concludes the inquiry.

2 Debates on Charity Tax Incentives

Many studies (e.g., OECD, 2020; Colinvaux, 2013; Colinvaux et al., 2012) evaluate the rationales for charity tax incentives. The rationales include increasing the provision of public goods, producing intangible benefits, decentralizing decision-making, and adjusting income bases. The theoretical justifications for these rationales are controversial, and the empirical results are mixed at best. As such, OECD (2020, 8) asserts, "[t]here is no single generally accepted rationale for the preferential tax treatment of philanthropy."

The private sector underproduces public goods because the producer of a public good captures just a fraction of its social value. To increase the production of public goods, the government may provide public goods directly or subsidize the private producers of public goods. Provided that all public goods are equally worthy, the subsidy for a public good is justified if the government is unable to provide it directly or the subsidy is more cost-effective than the direct provision.

Some legal restrictions prevent governments from being involved in certain activities. The U.S. government, for example, cannot directly aid religious institutions. One can argue that religion is a socially desirable activity promoting good citizenship and generosity. Such social benefits could justify indirect subsidies through charity tax incentives. Legal restrictions exist for a reason, however. The separation of church and state may help maintain political stability and religious freedom. The government may also be restricted from allocating funds to help illegal aliens. While it is humane to help needy people regardless of their citizenship status, aiding illegal aliens can encourage illegal immigration. Thus, it is debatable whether bypassing legal restrictions produces a positive net benefit.

Charity tax incentives increase social spending if they are treasury efficient. The treasury efficiency of charity tax incentives is largely an empirical question, and the empirical results are mixed at best. According to OECD (2020), taxpayers in many cases are relatively unresponsive to charity tax incentives, suggesting charity tax incentives may not be treasury efficient. The literature routinely produces highly divergent estimates of the responsiveness (Han et al., 2024). Researchers face many difficulties. To obtain a convincing estimate, they need a large number of observations, but governments do not change charity tax incentives often. And responses to charity tax incentives can depend greatly on economic and social circumstances (Karlan & List, 2007). Marginal tax rates vary across donors, and the complexity of tax codes makes the calculation of marginal tax rates difficult. Marginal tax rates can even be endogenous because households can increase contributions enough to move into a lower tax bracket (Meer & Priday, 2020).

The importance of treasury efficiency is debatable also at the conceptual level. According to Kaplow (2024), treasury efficiency can even be irrelevant in a general optimal income taxation framework that incorporates the adjustment of the tax schedule. For welfare analyses, a tax-and-transfer regime offering a special provision for philanthropic giving can be equivalent to a regime offering no special provision when the tax schedule is adjusted to produce the same revenue in the two regimes.

Charity tax incentives can produce intangible benefits by making philanthropic contributions more common. People may feel and behave better when they see more philanthropic activities. The literature describes this intangible benefit using various terms: promoting a giving culture, strengthening civil society, promoting a more altruistic and cooperative society, and developing better citizenship. These are just possibilities which cannot be verified. It would be imprudent to justify charity tax incentives largely based on these abstract notions. It is not even clear whether people view larger amounts/numbers of contributions driven by tax incentives more favorably than smaller amounts/numbers of contributions purely driven by benevolence. Some people might even resent big contributions that wealthy individuals make for the purpose of tax planning.

It is practically impossible for governments to evaluate the worthiness of each public good accurately and allocate funds based on the evaluation. Governments cannot observe the preferences of all individuals. Even if they could, political outcomes would hardly be the same as analytical solutions. Given the limitations of governments, decentralized decisions can improve the outcome. The hurdle is high, however. Worthier public goods would probably receive more

funds if donors care only about social benefits, as opposed to personal benefits. In this case, giving decisions could function as voting for the worthiness of each public good. The outcome of fair voting is likely to be better than that of a bureaucratic or a political process. Giving decisions are not fair voting, however. Many types of contributions, such as giving to alma mater or ideology-driven organizations, reflect personal interests. More importantly, larger donors exercise larger numbers of votes (more influence on the allocation of tax expenditures). In effect, charity tax incentives allow wealthy individuals to purchase votes. Unless large donors care more about social benefits than small donors do, disproportionate influence of large donors might produce a socially undesirable outcome.

To truly decentralize allocation decisions, the government should restructure the tax incentives. Reich (2018) argues that philanthropy can be a threat to democracy because big philanthropy is often an exercise of power. In the U.S., philanthropy policies make the threat more serious by facilitating the conversion of private assets into public influence for wealthy individuals. In restructuring philanthropy policies, he suggests, the government should aim to decentralize power in the production of public goods.

Another rationale for charity tax incentives is an adjustment to the income tax base; the income tax base should be the amount available for taxpayers' consumption, but the amount contributed is unavailable for donors' own consumption. This argument has a historical background. According to Colinvaux et al. (2012), when the U.S. Congress enacted the tax deduction for philanthropic contributions in 1917, only four years after the income tax, its concern was that taxpayers giving away their income might have less money to pay tax. Many economists may dispute this rationale.

When individuals spend on something voluntarily, they must derive utility from the spending. For example, Saez (2004) treats the philanthropic contribution as a variable generating utility in his model. The main criterion for the tax deductibility should be the essentiality of a spending category. Based on this criterion, the child tax credit is reasonable, but not the deduction of philanthropic contributions. Parents must raise their children. Spending on childcare is unavoidable ethically and legally. Perhaps, one can refine the rationale somewhat. When the government provides few essential goods (e.g., social safety nets and education), philanthropic contributions are essential from the society's perspective, although not from the donors' perspective. This may have been the case in the early 1900's when governments provided few social safety nets, but not anymore.

Charity tax incentives, especially those of the U.S., are subject to several criticisms (see OECD, 2020 for the review of charity tax incentives around the world). The tax breaks are regressive and undemocratic (OECD, 2020). Most low-income people do not itemize deductions, and those people do not get a tax break for philanthropic contributions. Even among those who itemize deductions, the tax benefit disproportionately increases with the marginal tax rate and hence with income. As a result, wealthy people have a disproportionate influence over public funds. But the priorities of big donors may not be consistent with those of society in general.

Typically, high-income donors favor higher education, arts, and culture, while low-income donors favor religion and welfare (Atkinson, 1997; Rabin, 1966). Husock (2024) reports that large foundations have their own agendas which are ideology-driven in many cases. In addition, the U.S. tax code gives donors ample room for tax planning and abuse (Colinvaux, 2013).

In sum, it is unclear whether charity tax incentives produce a net social benefit. The tax incentives have several rationales, but all of them are disputable. Moreover, charity tax incentives have side effects.

3 Treasury Efficiency versus Allocational Efficiency

In evaluating the effect of charity tax incentives on social welfare, previous studies focus largely on treasury efficiency. The total amount of social spending determines social welfare if all public goods have the same and constant effect on social welfare. It is more reasonable, however, to assume that public goods are not the same and that the marginal benefit of each public good decreases with the amount of spending on it. Under these assumptions, the allocation of funds across public goods can be much more important than the amount of funds. According to Saez (2004), government and private contributions are rarely perfect substitutes; for example, private and public schools do not provide exactly the same services and are not attended by the same public. Allocational efficiency is especially important when the NPOs' selections of public goods are dramatically different from the governments' selections. This section presents simple examples to illustrate these points.

Let us assume that there are 3 public goods: *PG1*, *PG2*, and *PG3*. *PG1* and *PG2* are equally good, and *PG3* is inferior. For every public good, the marginal social benefit decreases with the amount of spending on it.

Let x be the spending on the social good. The marginal benefit from *PG1*,

$$\begin{aligned} f(x) &= 10 - 0.01x && \text{for } 0 < x \leq 1{,}000 \\ &= 0 && \text{otherwise} \end{aligned} \tag{1}$$

Integrating $f(x)$, the total benefit from *PG1*,

$$\begin{aligned} F(x) &= \int f(x)dx = 10x - 0.005x^2 && \text{for } 0 < x \leq 1{,}000 \\ &= 0 && \text{for } x \leq 0 \\ &= F(1000) && \text{for } x > 1{,}000 \end{aligned} \tag{2}$$

The marginal benefit from *PG2*,

$$\begin{aligned} g(x) &= 10 - 0.01x && \text{for } 0 < x \leq 1{,}000 \\ &= 0 && \text{otherwise} \end{aligned} \tag{3}$$

The total benefit from *PG2*,

$$\begin{aligned} G(x) &= \int g(x)dx = 10x - 0.005x^2 && \text{for } 0 < x \leq 1{,}000 \\ &= 0 && \text{for } x \leq 0 \\ &= G(1000) && \text{for } x > 1{,}000 \end{aligned} \tag{4}$$

The marginal benefit from *PG3*,

$$\begin{aligned} h(x) &= 3 - 0.01x && \text{for } 0 < x \leq 300 \\ &= 0 && \text{otherwise} \end{aligned} \tag{5}$$

The total benefit from *PG3*,

$$\begin{aligned} H(x) &= \int h(x)dx = 3x - 0.005x^2 && \text{for } 0 < x \leq 300 \\ &= 0 && \text{for } x \leq 0 \\ &= G(300) && \text{for } x > 300 \end{aligned} \tag{6}$$

Without charity tax incentives, the government spends $1,000 on public goods, and NPOs spend $200 on public goods. With charity tax incentives, the tax revenue decreases by $100, so the government spends $900 on public goods. The government makes the allocation first, and NPOs follow.

3.1 Case One

In *Case 1*, both the government's allocation of funds and the NPOs' allocation of funds are socially optimal.

For the allocation to be optimal, the benefit from the last dollar must be the same across the public goods that deserve to be funded (high-benefit public goods). Since $f(x)$ and $g(x)$ are identical, *PG1* and *PG2* must receive the same amount of funds. *PG3* becomes a high-benefit public good only after large allocations to *PG1* and *PG2* make their marginal benefits equal to the benefit from the first dollar spent on *PG3*.

Without charity tax incentives, the government allocates $500 to *PG1*, $500 to *PG2*, and $0 to *PG3* because *PG3* is a low-benefit public good ($f(500) = g(500) = 5 > h(1) = 2.99$). NPOs allocate $100 to *PG1*, $100 to *PG2*, and $0 to *PG3* because *PG3* remains to be a low-benefit public good ($f(600) = g(600) = 4 > h(1) = 2.99$). The social benefit without charity tax incentives (baseline) totals $8,400 ($F(600) + G(600) = \$4,200 + \$4,200$).

With charity tax incentives, the government allocates $450 to *PG1*, $450 to *PG2*. The social benefit from the government's spending is $6,975 ($F(450) + G(450) = \$3487.5 + \$3487.5$). The $8,400 baseline is matched if NPOs add $150 to *PG1* and $150 to *PG2*, that is, if the $100 tax expenditure is matched by a $100 increase in private contributions (treasury neutral).

In this case, treasury efficiency solely determines the effectiveness of charity tax incentives. Social welfare stays the same if charity tax incentives are treasury neutral, improves if they are treasury efficient, and worsen it if they are treasury inefficient.

3.2 Case Two

In *Case 2*, the government's allocation is socially optimal, and the NPOs' allocation is socially suboptimal.

As in *Case 1*, the government spends $500 each on *PG1* and *PG2*. Suppose that NPOs spend only on *PG1*. In this case, the social benefit without charity tax incentives is $8,300 ($F(700) + G(500) = \$4,550 + \$3,750$). With charity tax incentives, the social benefit from the government's spending is $6,975 ($F(450) + G(450) = \$3487.5 + \$3487.5$). To match the $8,300 baseline, the benefit from NPOs' spending must be $1,325 ($8,300 - $6.975), and hence the total benefit from *PG1* must be $4,812.5. The amount of spending needed to generate $4812.5 from *PG1* rounds up to $807 ($F(806) = \$4811.82$ and $F(807) = \$4813.755$). Thus, to breakeven, the tax subsidy of $100 must increase private contributions by $357. To improve social welfare, charity tax incentives must be extremely treasury efficient.

Now, suppose that NPOs spend only on *PG3*. Then the baseline is $7,900 ($F(500) + G(500) + H(200) = \$3,750 + \$3,750 + \$400$). To match the baseline, the benefit from the NPOs' spending on *PG3* must increase to $925 ($F(500) + G(500) + H(200) - F(450) - G(450) = \$7,300 – \$6,975$). The maximum benefit from *PG3* ($H(300)$), however, is only $450. Thus, in this case, charity tax incentives decrease social welfare, regardless of treasury efficiency.

3.3 Case Three

In *Case 3*, the government's allocation is socially suboptimal, and the NPOs' allocation is socially optimal.

The government spends only on *PG1*, and NPOs spend on *PG2* to mitigate the suboptimal allocation by the government. In this case, the baseline is \$6,800 ($F(1000) + G(200) = \$5{,}000 + \$1{,}800$). With charity tax incentives, the social benefit from the government spending is \$4,950 ($F(900)$). To match the baseline, the social benefit from the NPOs' spending on *PG2* must increase to \$1,850 (6800 - 4950). The amount of spending needed to generate \$1,850 from *PG2* rounds up to \$207 ($G(296) = \$1{,}847.82$ and $G(207) = \$1{,}855.76$). In this case, social welfare improves if private contributions increase by only \$7 in response to \$100 in tax breaks. Charity tax incentives can be extremely treasury inefficient and still improve social welfare.

These examples highlight the importance of allocational efficiency. The effect of charity tax incentives on social welfare depends both on treasury efficiency and allocational efficiency. In the circumstances where the marginal benefit varies widely across public goods, allocational efficiency can easily dominate treasury efficiency. Charity tax incentives that are highly treasury efficient can fail to improve or even worsen social welfare if the NPOs' selections of public goods are inferior to the government's selections. The converse is also true.

4 Improving the Effectiveness of Charity Tax Incentives

The effectiveness of charity tax incentives depends on allocational efficiency, treasury efficiency, and the size of intangible benefits. Allocational efficiency, despite its importance, has been understudied. This section discusses ways to improve the effectiveness of charity tax incentives, with an emphasis on allocational efficiency.

Governments around the world regulate NPOs to ensure that they duly serve social purposes. An obvious way to improve the effectiveness of charity tax incentives, therefore, is to improve rules and regulations. The difficulty of quantifying the outcomes of philanthropic activities, however, limits governments' ability to improve the effectiveness of charity tax incentives through rules and regulations. In the U.S., few rules are particularly designed to improve allocational efficiency, which is very difficult to measure. At the federal level, the Internal Revenue Service grants tax-exempt status to NPOs and regulates them based on broad rules: NPOs must operate for exempt purposes, including charity, religion, education, and science; the rules prohibit NPOs from distributing their earnings to their shareholders or individuals; and the rules restrict NPOs from many political or lobbying activities. States oversee NPOs for their operational integrity. Neither the federal government nor the states scrutinize the relative worthiness of NPOs' activities.

There are some rules that can indirectly contribute to improving allocational efficiency. The U.S. government, for example, gives a more generous tax break to public charities than that to private foundations. The tax deduction limit is higher for donors giving to public charities than for those giving to private foundations, and public charities avoid the investment income tax imposed on private foundations. Provided that activities of private foundations are more likely to be steered by personal interests of a few large donors, the tax rules favoring public charities may improve allocational efficiency.

The tax rules also treat cash contributions and in-kind contributions differently. The tax deductibility limit is higher for cash contributions than for in-kind contributions. Possibly,

donors contributing cash are more serious about social benefits than those who contribute an unwanted vehicle. Those who are more serious about social benefits may choose a better NPO that spends money more wisely. These tax rules might help improve allocational efficiency somewhat, but probably not significantly. Furthermore, most rules have loopholes. For example, donating appreciated assets, such as stocks and real estate, is often more tax-efficient than selling and donating cash, as it enables donors to avoid the capital gains tax.

To significantly improve the effectiveness of philanthropic tax incentives, the government would have to link the tax incentives to social benefits more explicitly and strongly. Considering the difficulty of quantifying social benefits, the task is impractical. To quantify the social benefits of philanthropic activities, the government would need to learn the socio-economic conditions of beneficiaries and externalities from NPOs' activities, and develop innovative analytical tools.

A natural alternative to analytical approaches is to employ market mechanisms. In essence, charity tax incentives are a public-private partnership. The government provides some funds to NPOs in the form of tax expenditures, and NPOs provide public goods using the funds. As I have argued earlier (Park, 2019), the key to a successful public-private partnership is to identify the tasks that the private sector can do better and delegate those tasks to the private sector with proper incentives.

Currently, the U.S. government delegates the task of allocating tax expenditures to the private sector in an incorrect way that fails to utilize the strength of the private sector sufficiently. The private sector efficiently allocates resources through market mechanisms, which aggregate the decisions of a large number of participants. The current arrangement, however, lets a small number of large donors manage a huge amount of tax expenditures. The allocation decisions of those large donors may not be better than those of the government. If they base their decisions on analyses, their decisions may be comparable to those of a well-functioning government. If they base their decisions on ideology and personal interests, their decisions may be comparable to those of a dysfunctional government.

To take advantage of market mechanisms, the government should delegate the task of allocating tax expenditures to the general public. An ordinary person may choose an NPO based on a guess about the NPO's worthiness. In most cases, the size of the error may be larger for a guess than for a sophisticated analysis. Some people may greatly overestimate the worthiness, and some others may greatly underestimate it. When a large number of those estimates are aggregated, however, overestimates should largely offset underestimates. Thus, the average of a large number of guesses may be much more accurate than an analytical estimate. Furthermore, personal interests of donors will have an insignificant influence on the allocation of tax expenditures when allocation decisions are decentralized among a large number of people. Since personal interests vary widely across individuals, any particular NPO is unlikely to benefit significantly from personal interests.

This section presents a charity tax incentive scheme (new scheme) that delegates the task of allocating tax expenditures mostly to the general public and hence takes advantage of market mechanisms more fully. Although the main goal of the new scheme is to improve allocational efficiency through market mechanisms, it could also improve treasury efficiency, incentivize NPOs to improve their operational efficiency, and increase intangible benefits. In addition, the tax break would become fairer.

The new scheme features a uniform tax credit for all donors, a tax on NPOs' receipts, and a tax exemption tied to the number of donors. Obviously, the uniform tax credit would make the tax break fairer. In percentage terms, low-income people and high-income people would get the same tax break. More importantly, the uniform tax credit plays a critical role in improving

allocational efficiency. In a sense, the new scheme is a voting system; NPOs getting more votes get a larger tax break. To obtain a fair voting outcome, the voting system should give all eligible voters the same incentive to vote. The uniform tax credit would substantially equalize the voting incentive.

Tying the tax exemption for NPOs to the number of donors is a way to delegate the task of allocating tax expenditures to the general public, as opposed to a few large donors. It would be a transition from the vote-per-dollar system to the vote-per-person system. A simple way to tie the tax exemption to the number of donors is to grant a fixed amount of exemption per donor up to the total amount of receipts. Suppose that a NPO has 1,000 donors and that the per-donor exemption amount is $100. Then the allowable exemption amount is $100,000 (100 x 1,000). For the calculation of the exemption amount, it does not matter whether a donor contributed $10 or $10,000. Thus, if a NPO receives $10,000 from only one donor, the exemption is only one percent of the receipt. If a NPO receives $10,000 from 200 donors, on the other hand, the allowable exemption (100 X 200 = $20,000) exceeds the total receipt, and the exemption is 100 percent of the receipts.

Tying the tax break to the number of donors should improve allocational efficiency if the market mechanisms described above operate properly. Measuring the worthiness of a public good is subject to large errors because many relevant variables, such as the utility of beneficiaries and the value of externalities, are unobservable. Thus, analytical estimates may not be much better than simple guesses. Market mechanisms can be much more reliable in this circumstance. Donors may rank public goods based on their guesses about the worthiness of each public good and choose highly ranked public goods. A single guess is likely to be less accurate than an analytical estimate, but the average of guesses is likely to be more accurate than an analytical estimate when the number of observations is large. This market mechanism makes the number of donors a fairly reliable measure of the worthiness of a public good. The new scheme employing the market mechanism should steer tax expenditures towards worthier causes.

As usual, there would be some implementation issues. Counting a very small contribution like $1 (micro-donation) as one vote could be problematic. Many micro-donors might not be serious about the worthiness of public goods. In addition, micro-donations might make it easy for NPOs to manipulate the donor count. To assure the reliability of the donor count as a measure of worthiness, the government might need to set a floor on the contribution amount as a seriousness check. As a compromise, the government might aggregate micro-donations. If the floor would be $50, for example, the government might count 50 one-dollar contributions as one donor.

Another issue is how to link the donors' tax break to the tax exemption for NPOs. Foreigners would not be eligible for the tax credit. People who had dropped a few dollars into the Salvation Army's Red Kettle might not claim the tax credit. To be fair, the government should neither tax nor count as votes those contributions that did not incur tax expenditures. Those contributions would be revenue-neutral transactions between private parties. Implementation issues could be challenging, but in this era of big data, they should be resolvable.

The new scheme would prevent large donors from commanding a large amount of tax expenditures. This restraint is important if large donors are more likely to have a strong personal interest. Many alumni of Ivy League schools contribute a large amount to their alma mater, probably to enhance the prestige of their alma mater and secure the legacy preference for their children. Some rich people generously support a specific genre of art or music. Under the new scheme, the tax exemption for those contributions would be small, unless the general public joined the rich donors. Overall, the influence of personal interests would be very limited

because personal interests vary widely across individuals. The new scheme is a desirable form of democracy that rules out the tyranny of the majority and the tyranny of the wealthy. Each group would take its fair share; NPOs receiving 60 percent of votes would manage 60 percent of tax expenditures, and NPOs receiving 40 percent of votes would manage 40 percent of tax expenditures. Neither the majority nor the wealthy would be allowed to control a disproportionate share of tax expenditures.

The new scheme can also improve the treasury efficiency. Suppose that the government structures the new scheme such that the tax expenditure is the same as that of the current tax structure of the U.S. The government can recoup the cost of the tax credit through the tax on NPOs to any desired extent. The government can even predetermine the tax expenditure in the budget and allocate the tax exemption to each NPO based on its share of donors (the number of its donors / total number of donors). Thus, it is not a problem to set the tax credit at a generous level. If the government sets the tax credit at the highest marginal income tax rate, the tax break will be the same as before for the donors in the highest tax bracket and larger for all other donors.

Imagine that a donor is contemplating about "purchasing" a public good. There are two relevant prices for the donor: the after-tax price for the donor disregarding the tax on NPOs (list price) and the price that fully reflects both the tax credit for the donor and the tax on NPOs (impact price). Assuming that the fair market price of the public good is 1, the list price is 1-t_1, where t_1 is the tax credit as a share of the contribution. The impact price is $\frac{1-t_1}{1-t_2}$, where t_2 is the effective tax rate on the NPO's receipts; the donor pays 1-t_1 for 1-t_2 unit of the public good.

Unambiguously, donors making contribution decisions based on the list price would contribute more. For donors making contribution decisions based on the impact price, the new scheme would have an ambiguous effect on the treasury efficiency; low-income donors would contribute more, and high-income donors would contribute less. The relative importance of the list price and the impact price is not really testable. However, there are several reasons why the new scheme might not reduce high-income donors' contributions significantly.

Many high-income donors are motivated by tax planning and publicity. For example, corporations make philanthropic contributions to increase their profit (Navarro, 1988). What matters for those donors is the list price. Some large donors may be motivated by personal interests, such as gaining influence over certain matters. Provided that their interests are strong, those donors may not reduce their contributions significantly when the impact price rises. It is even possible that they increase their contributions. A charitable contribution motivated by personal interests is like the purchase of a consumption good. The contribution will increase if the demand for influence is inelastic. If the price of beef rises, consumers may buy less beef, but they may spend more money on beef. Among those donors who are purely altruistic, the effect of uneven changes in the impact price across tax brackets should be neutral. In addition, high-income donors could lower the impact price by switching to more popular public goods. Such switching would mitigate the effect of a high impact price on the treasury efficiency, while further improving the allocational efficiency. All in all, the new scheme is likely to improve the treasury efficiency.

The new scheme might strengthen the incentives of NPOs to improve their operational efficiency. The government might provide standardized performance measures, such as operating expenses and the composition of beneficiaries, so that donors could evaluate NPOs more easily and choose more efficient ones. Under the new scheme, efficient NPOs could keep a larger portion of receipts by attracting a larger number of donors, in addition to raising more funds.

The new scheme might produce larger intangible benefits. The tax credit should increase the number of donors by incentivizing small donors more strongly. Widespread giving should strengthen social cohesion and promote a giving culture more effectively.

One may argue that prominent contributions by superrich people have stronger effects on social cohesion and giving culture. Contributions by superrich people are more conspicuous and may positively influence the public perception about social justice. The new scheme is unlikely to make superrich people stop contributing. Several tycoons, such as George Peabody (1795–1869) and Andrew Carnegie (1835–1919), established a philanthropic foundation in the era of no income tax. Perhaps, superrich people might contribute a smaller amount. A reduced amount might not necessarily reduce the positive influence on the public perception. A smaller amount of donation perceived to be motivated by pure goodwill could be more inspiring to the general public than a larger amount suspected of tax management.

In sum, by tying the tax exemption for NPOs to the number of donors, the new scheme would effectively delegate the task of selecting worthier public goods to the general public. Given that it is very difficult to estimate the worthiness of public goods analytically, market mechanisms averaging the estimates of a large number of participants may be more reliable than analyses by the government or a few large donors. The new scheme employing market mechanisms should improve the allocational efficiency of charity tax incentives and might also improve their treasury efficiency. In addition, the new scheme would more strongly urge NPOs to improve their operational efficiency, it would be fairer, and it might produce larger intangible benefits.

5 Conclusion

To be justified, charity tax incentives should improve social welfare. It is hard to tell whether charity tax incentives improve social welfare, however. Charity tax incentives can improve social welfare in several ways: increasing the amount of social spending, allocating public funds to worthier public goods, reducing the cost of delivering public goods, and generating more intangible benefits. None of these possibilities is convincingly supported. Empirical findings about the treasury efficiency are mixed at best. Estimating the allocational efficiency is nearly impossible because the worthiness of each public good is subjective and can hardly be quantified. The complexity of government operations and a large number of NPOs make it difficult to compare the operational efficiencies of governments and NPOs. By definition, intangibles are not measurable.

This paper has looked at a more manageable and practical issue, namely, how to improve the effectiveness of charity tax incentives, rather than whether or not charity tax incentives improve social welfare. The main contributions are illustrating the importance of allocational efficiency and proposing a new scheme that might significantly improve the allocational efficiency of charity tax incentives. In evaluating the effect of charity tax incentives on social welfare, the allocational efficiency can matter much more than the treasury efficiency under the following conditions: the government and NPOs select different public goods; the social benefit differs across public goods; and the marginal benefit from each public good decreases with the amount of spending on it. If the NPOs' selection of public goods is more socially optimal than the government's selection, charity tax incentives that are treasury inefficient can still improve social welfare. The converse is also true.

The new scheme has the following features: a uniform tax credit for all donors, a tax on the receipts of NPOs, and a tax exemption for NPOs based on the number of donors. In effect, the

government delegates the task of choosing worthy public goods to the general public by tying the tax exemption for NPOs to the number of donors. Provided that worthier public goods attract more donors, the new scheme would substantially improve the allocation of tax expenditures. The new scheme could also positively affect the treasury efficiency of charity tax incentives, the operational efficiency of NPOs, and the generation of intangible benefits. In addition, the new scheme would be fairer in that it offered the same tax break to low-income and high-income donors.

Since philanthropic activities involve highly complex elements, such as externalities and intangible benefits, it is exceptionally difficult to estimate their effects analytically. The new scheme relies on market mechanisms coordinating a large number of participants, as an alternative to analytical approaches. Given the complexity of philanthropic activities, any reform proposal would be debatable. At the minimum, however, the new scheme offers a new perspective for policy debates and future research.

Acknowledgement

The author thanks three anonymous reviewers for their helpful comments and suggestions.

References

Atkinson, R. (1997). Theories of the Federal Income Tax Exemption for Charities: Thesis, Antithesis and Synthesis. *Stetson Law Review*, 27(2), 395–432. Online: https://ir.law.fsu.edu/articles/376/

Colinvaux, R. (2013). Rationale and Changing the Charitable Deduction. *Tax Notes*, 138, 1453–1459. Online: https://scholarship.law.edu/scholar/108/

Colinvaux, R., Galle, B., & Steuerle, E. (2012). *Evaluating the Charitable Deduction and Proposed Reforms.* The Urban Institute. Online: https://tinyurl.com/2u7yud6t

Han, X., Hungerman, D. M., & Ottoni-Wilhelm, M. (2024). Tax Incentives for Charitable Giving: New Findings from the TCJA. *National Bureau of Economic Research Working Paper*, No. 32737. https://doi.org/10.3386/w32737

Husock, H. (2024, November 15). *JD Vance is right: Reduce the power of big foundations to help charity*. The Hill. Online: https://tinyurl.com/d99vjmuy

Kaplow, L. (2024). Optimal Income Taxation and Charitable Giving. *Tax Policy and the Economy*, 38, 123–162. https://doi.org/10.1086/730054

Karlan, D., & List, J. A. (2007). Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment. *American Economic Review*, 97(5), 1774–1793. https://doi.org/10.1257/aer.97.5.1774

Meer, J., & Priday, B. A. (2020). Tax Prices and Charitable Giving: Projected Changes in Contributions under the 2017 Tax Cuts and Jobs Act. *Tax Policy and the Economy*, 34, 113–138. https://doi.org/10.1086/708172

Navarro, P. (1988). Why Do Corporations Give to Charity?. *The Journal of Business*, 61(1), 65–93. https://doi.org/10.1086/296420

OECD. (2020, November 26). Taxation and Philanthropy. *OECD Tax Policy Studies*, No. 27. https://doi.org/10.1787/df434a77-en

Park, S. (2019). A Critical Look at Public-Private Partnerships. In *Public Policy: The Second Best, Political Compromise, and Social Welfare* (pp. 9–30). Amazon Kindle.

Rabin, E. H. (1966). Charitable Trusts and Charitable Deductions. *New York University Law Review*, 41(5), 912–925. Online: https://tinyurl.com/3fx8ec8y

Reich, R. (2018). *Just Giving: Why Philanthropy Is Failing Democracy and How It Can Do Better*. Princeton University Press. https://doi.org/10.2307/j.ctvc77jz8

Saez, E. (2004). The Optimal Treatment of Tax Expenditures. *Journal of Public Economics*, 88(12), 2657–2684. https://doi.org/10.1016/j.jpubeco.2003.09.004

Institutiones Administrationis – Journal of Administrative Sciences Vol. 6 (2026) No. 1, 71–86
DOI: 10.54201/iajas.172

Tort Law as Risk Allocation Contract

Some Critical Remarks on the Regulatory Deterrence Model

Tze-Shiou Chien*

* Adjunct Research Professor, Institutum Iurisprudentae, Academia Sinica, Taipei, Taiwan. Adjunct Professor, College of Law, National Chengchi University, Taipei, Taiwan. E-mail: tschien@alum.sinica.edu.tw

Abstract

The standard economic analysis of law has modelled tort liabilities as the state's instruments to reduce accident costs, comprising prevention costs and damages. This regulatory deterrence model (RD model), however, does not fit with the private autonomous nature of tort law (PA model). This is why the RD model has to resort to numerous *ad hoc* elements to explain it. Nevertheless, given that private law is not some policy-instrument; rather, it is a complex system of norms with courts being independent adjudicators of legal disputes arising, amongst others, from contractual arrangements, legal dogmatics, rather than standard economics, seems to be a more appropriate approach to the economic analysis of private law. As such, in this paper, I am presenting an economic analysis of tort law that is closely associated with the PA model.

Keywords

tort liability, accident costs, liability rules, private autonomy, complexity system

1 Introduction

Accidents cannot be fully anticipated either because the prevention costs are too high or, more importantly, because the damages are less worthy than those acts causing the accidents (Coase 1960). Tort law, therefore, should assign the responsibility for the damage to the party whose benefits are less because this formula is *ex ante* beneficial for the injurer and the injured and then could be seen as their accident-risk-contract (the private autonomous model, henceforth: PA model) (Cheung 1992; 1998). The standard economic analysis of tort law, however, models tort law as a regulatory instrument to deter behaviors that do not align with some preexisting safety standards (the regulatory deterrence, henceforth: RD model). But there are several deficiencies in this model.

First, the remedy of tort law is compensatory, which is inconsistent with the function of deterrence. Second, there is a built-in mechanism in the risk-allocation contract model that induces the parties to take optimal safety measures because the less safe measures a party takes, the more likely it is that the party is liable for the damages, as the benefits of its act (B) exceed the expected damages (PL). Therefore, a regulatory safety prescription is not necessary. Third, the preexisting safety standards are external to the injurer when it acts, causing damage. And fourth, the court as an independent adjudicator is not institutionally fit for prescribing safety measures.

 Published online: 24 April 2026

The misuse of the RD model in analyzing tort law has resulted from the very foundation of economic analysis of law – externality theory. In this theory, coercive governmental measures have been required to correct market failures (externalities). As Coase (1960) has shown that, on the one hand, the externalities usually are not failures; on the other hand, coercive governmental measures have their own problems, the costs of which might be higher than the benefits gained. As the tort liability is imposed only after accidents occur and a more or less fault responsibility is the principle for paying compensatory damages, the RD model does not fit for analyzing the tort liability. In practice, tort law is a part of private law of which private autonomy is the principle and default rules are their contents. Therefore, tort law is the implicit contractual arrangement by the injure and the victim of allocating accident risks.

In the following section, I expose the major problems the RD model encounters in explaining tort law, such as the reasonable person standard problem, the harmful precautions problem, the strict liability problem, the causation problem, the good Samaritan problem, the unknown risks problem, and the empirical inadequacy problem. Afterwards, I am putting forward the model of tort liability as a contractual arrangement of risk allocation (the PA model), and I will demonstrate its explanatory power for resolving those problems encountered by the RD model. In section four, I am examining the claim that economic analysis of law is a scientific study of law which underlines the misuse of the RD model and the missing of the PA model.

2 The Regulatory Deterrence Model

The standard tort liability model has been constructed on the assumption that the injurer and the injured party can *ex ante* "jointly"[1] take precautions to prevent the injury and the liability imposed will induce the injurer to take optimal precaution by which the sum of precaution costs of the injurer and the injured and the expected damages will be minimized (Brown, 1973; Shavell, 1987; Landes & Posner, 1987; Cooter & Ulen, 2016; Cooter & Porat, 2014; Schäfer & Ott, 2022). There are, however, some fundamental problems with the premises of this understanding of tort liability, which ought to be laid open and deconstructed before any constructive argument on my part could be introduced. As such, this section is dedicated to roadmap the most widespread faulty assumptions of the RD model and to highlight their essential inconsistencies.

2.1 The Reasonable Person Standard Problem

The most obvious problem of the RD model is its inconsistency with "the reasonable person standard". According to this model, the more the injurer's precaution costs are, the lower its liability standards.[2] In other words, the less capable of preventing the injury the injurer is, the less likely the injurer is liable for the injury, whereas "the reasonable person standard" doctrine states that the injurer's lesser capacity of taking precautions is not a reason for immunity from

[1] The injurer and the injured act independently, but whether they are negligent is assessed on the standards calculated interactively. $L(x,y) = p(x,y)D + A(x) + B(y)$ is the social costs function of the parties' precautions (Landes & Posner 1987, 59).

[2] In the equation $L(x,y) = p(x,y)D + A(x) + B(y)$, the optimal precautions" x*and y* can be found by taking the first partial derivatives of L with respect to x and y and setting the resulting expressions equal to zero. This requires that x* and y* satisfy the following conditions: $Ax = -pxD$ and $By = -pyD$". (Landes & Posner, 1987, 59–60).

its otherwise negligence liability. Landes and Posner (1987, 123–131) argue that the courts' costs of investigating into the less capable injurer's capability would justify the courts seeing the less capable injurer as a standard one. But this argument misses the point. The courts cannot accept such reasoning, it can only accept the *fact* of the injurer's lesser capability, but cannot find any merit in it. Hence, the courts' information about the injurer's capability is not an issue here.

Shavell (1987, 75) would object, however, that, in some cases, a "moderate" level of care that the less capable injurer cannot meet would discourage them from engaging in the socially harmful activity at the beginning. But this is not a good argument, either. Firstly, the courts have never produced, according to the best of my knowledge, such a reasoning. Secondly, the fact that the injury had been inflicted would inevitably mean that the deterrence had failed.[3] Thirdly, resorting to this outside model factor of benefits of activity would discredit the model *qua* model. Fourthly, the factor of benefits of activity would also apply to reasonable persons, and such cannot be allowed to engage in those activities whose benefits are less than the damages done, even if they take optimal care.

Furthermore, there is an information problem for determining the injurer's and the victim's reasonable (negligence) standard from an *ex post* point of view. The determination of the injurer's reasonable standard depends on the victim's reasonable standard, while the determination of the victim's reasonable standard depends on the injurer's reasonable standard. This would put too much information burden on the courts to determine the reasonable (negligent) standards of the injurer and the injured, respectively. Meanwhile, the injurer and the injured would deviate from the standard even if the opposing party makes a small error (Bayern 2023, 47–57).

Cooter and Porat (2000), on the other hand, argued that the courts should take into account the risk to the injurer in determining the negligence standard; otherwise, the standard would be lower than the social optimal. While this is correct under the *ex ante* regulatory model, determining tort liability is an *ex post* investigation of the cause of a specific realized event, not an unrealized abstract standard. And the negligence rule is derived from the private autonomy principle, therefore, in determining whether the injurer should be liable or not, the courts should assume that the injurer has already taken the risk to itself into account and would only consider the injurer's activity's external impacts on others.

2.2 The Harmful Precautions Problem

According to some RD model understandings, those benefits that do not flow to the injurer would be excluded, and thus sometimes would make the injurer take self-optimal but socially inefficient care. Perry (2023) called this negative externality "harmful precautions". Consider the example of driving. To drive fast for taking a patient to the hospital for emergency treatment should allow the driver to drive faster than normally, but not according to this RD model.

Intentional torts present the opposite problem. In most intentional torts, the precaution costs are negative, i.e., resources spent on intentionally causing damages, and thus the injurer should not be liable according to this model. To tackle the problem, Landes and Posner (1987, 149–189) introduced an additional factor, namely that the damage to the injured party should be larger than the gain of the injurer. But this has a "utility monster" problem. Ultimately, they resort to the reason that the injurer can but does not transact under low transaction costs, and therefore, the injurer is liable; some rather awkward explanation.

---

[3] Note this doctrine has only been invoked in actual cases in which accidents did occur and therefore it means that the deterrence has failed in these specific cases.

Furthermore, on the injured party's side, under this *ex ante* RD model, the potential damage which might have occurred due to the injurer's dangerous activities has not been included in the calculation of the liability standard, and thus the standard would be too low from the perspective of social efficiency (Bayern 2023, 30–31).

2.3 The Strict Liability Problem

Under the RD model, a tort regime of negligence rule or strict liability would make no difference in inducing the injurer to take optimal care. But if it is so, the distinction between the negligence rule and strict liability would be misleading. This problem is exposed through Cooter (1984), who puts forward a prices/sanctions model in which negligence rule is the sanction and strict liability is the price. He argues that when imperfection in setting the liability standard is more significant, strict liability should be adopted, while the imperfection in assessing damages is more significant, negligence rule should be adopted. But as Grady (2009) pointed out, the discontinuity of liability facing the injurer in the negligence rule cannot exist due to the rule of factual causation, and therefore this prices/sanctions model must be regarded as unsound. Furthermore, firstly, the term "prices" is misleading. "Taxes" would be more accurate because the "prices" herein are set by the courts rather than the spontaneous order of market transactions. Secondly, a strict liability as *ex post* liability is not an appropriate instrument for enforcing these "prices/taxes" regimes. Take car driving as an example. Speeding tickets are "prices/taxes", a strict liability which is initiated only after damages have been done,and is still a sanction *tout court*. Thirdly, as the above subsection 2.2 shows, the negligence rule would take the benefits inflowed to third party by the concerning injurer's activity into account in determining the negligence standard, while strict liability would induce the injurer to disengage in this kind of socially beneficial activities.

Shavell (1980) argues that strict liability is a better tort regime for those cases in which the "activity level effect" exists. But as Garoupa and Ulen (2013) point out, the effect is not necessarily a monotonically increasing relationship between the probability of an accident's occurring and the quantity of the underlying risky activity. Even in the cases where this activity level effect exists, they correctly argued that *ex ante* regulations are much more effective than an *ex post* liability regime to address this problem.

The fundamental problem of this comparison of negligence rule and strict liability, however, lies in its methodological flaw: in this comparison, the negligence rule is, indeed, a legal rule (whatever its contents) while "strict liability" is a legal result, similar to the predicate "not liable". And if the courts assessed tort liability totally based on effectiveness, legal rules such as the negligence rule would be superficial because the courts would just rule the injurer not liable upon not having induced it to take optimal precautions or the injurer liable upon having induced it to take optimal precautions. The meaningful comparison between negligence rule and strict liability, therefore, has to be based on strict liability as a legal rule which has a defined scope of application. This is why strict liability as a legal doctrine has been confined to specific torts such as products liability and abnormally dangerous activities.[4]

[4] In civil law system, fault is the rule while strict liability is exception.

2.4 The Causation Problem

Next, there is Cooter's (1987) thesis, built on the ideas of Bertrand Russell, according to which cause is disappearing in economic models of tort law. Instead, the mathematical relationships among such variables as the probability of accidents, the harm inflicted, and optimal precautions tend to dominate and supplement these kinds of discussions. However, factual causation and proximate cause are important doctrines in tort law. What is more, from a practical point of view, as Grady (1989) pointed out, the injured party usually would not allege that the injurer does not meet the due care standard and therefore is negligent; rather, the victim would allege those precautions the injurer does not take which would have prevented the damage from occurring. Being positive theorists of law, legal economic analysts have to explain this phenomenon.

Landes and Posner (1987, 228–255), as usual, introduced an extra factor of "care of complying the safety standard" to save this "positive" model of tort law. In their model, the usual safety standard, below which the injurer is judged to be liable for damages, is closely associated with the "care of complying the safety standard", and, as such, causation is largely left out, being kept for those accidents which would have happened regardless of the "care of complying the safety standard".

Shavell (1987, 105–110) has demonstrated that those accidents not caused by the injurer would have no effect on determining the optimal care. He also pointed out that, similar to the incentive induced by strict liability, the liability regime without this causation requirement would have no effect on the injurer's taking optimal care. However, he put forward some affirmative reasons for the causation doctrine. First, he claims that the injurer would not engage in the activities, regardless of taking due care, because it would bear "crushing" liability. The problem with this argument is that it is an *ad hoc* argument, again, similar to his explanation of the reasonable person standard above, in subsection 2.1. Second, the causation requirement would reduce the number of cases brought into courts, and thus, administrative costs would be reduced.

The real problem, however, is that this *ex ante* regulatory model does not fit with *ex post* compensatory tort law. The causation requirement only makes sense retrospectively for finding those parties having engaged in an unreasonable risky activity causing some damages in concrete cases. In short, causation in tort law is what Hayek (1945) said "the knowledge of the particular circumstances of time and place" contrasting with "scientific knowledge". That is, the *ex post* nature of tort liability allows the injurer to dynamically adjust his act to changing risks.

2.5 The Good Samaritan Problem

Others argue that the less costly the precaution is, the higher the liability standard, leading to the conclusion that everyone should be a good Samaritan. This is, however, inconsistent with the legal requirement of duty of care in tort liability. Landes and Posner (1987, 142–148) have argued that the duty of a good Samaritan would induce "the injurer" to take substitutive activities which are absent of "coerced rescue risks" and incidentally would reduce the socially not optimal activities. Again, an extra factor, substitutive activities, has sneaked in to justify the model. In my opinion, they should have incorporated these substitute activities into the model from the very beginning. The omission of this factor makes the investigation of strict liability partial and misleading: strict liability compared with negligence would not induce the injurer to take too much precautions. The more serious analytical problem, however, is that it is impossible for the potential injurers *ex ante* to know which activities are prone to encountering coerced rescue risk, and therefore, when taking substitutive activities to avoid those risks is a

viable option. Hence, in this situation, liability rules do not induce people's conduct and thus are meaningless, even harmful, to transfer risks with no benefits.

2.6 The Insurance Problem

From a deterrence standpoint, the insurance the injurer procures would dilute its incentive to take optimal care, and therefore, the injurer should not be allowed to procure insurance for its liability. In fact, third-party insurances are prevalent and sometimes compulsory as well. Some countries, such as New Zealand, even take certain accidents out of the tort law regime and set up a comprehensive fund for compensation without considering whether the injurer is liable or not.

Whether the insurer will, indeed, insure the liability depends on the insurer's ability to prevent the insured's (i.e., the injurer) from committing moral hazards. This has become an issue; however, only insurance for liability had been allowed by the (efficient) law based on the assumption that the insurance would not dilute the injurer's incentive to take optimal care. Otherwise, a calculation of trade-off between the degree of dilution and the benefits of insurance has to be done at least from the viewpoint of social optimization. The injurer's judgment-proof argument has the same problem. If so, this would put too much burden on the courts as adjudicators of individual cases and thus would not be a part of private law.

2.7 The Unknown Risks Phenomenon

Nevertheless, some damages, in terms of type or amount, are too uncertain to be *ex ante* calculated for optimal precautions, such as environmental pollutions, nuclear disasters, or other mass torts. Therefore, the *ex ante* regulatory model cannot apply to these damages. A strict liability based on deterrence, giving incentives to take optimal precautions, is superficial. In cases of environmental pollution, many causes of pollutions are unknown until recently. Therefore, there are no "optimal" precautions as such. Similarly, in cases of nuclear disasters, either the precautions are optimal until the disasters occur, or maximum damages have been granted by special laws. Hence, these phenomena also lay some fundamental inconsistencies open with the RD model understandings.

2.8 The Empirical Inadequacy Problem

Finally, there is the so-called empirical inadequacy problem, which consists in tort liability being a remedy, not a sanction. Therefore, the empirical studies of the effectiveness of tort liability are meaningless in the application of law (cf. Arlen, 2021). Empirically speaking, however, in terms of verifying this deterrence model, no case of negligence is apparent, and injurers have taken optimal care in all strict liability cases. These have never been empirically determined, though.

But this is vehemently opposed by some, such as Eric Posner (2021), who put forth an opposite "futile argument", saying that empirical studies would prove that the imposition of tort liability would have no effect on reducing the number of accidents or would induce otherwise detrimental defensive acts due to the contractual nature of tort liability. Shao and Weng (2023) also experimentally demonstrated that while there is some deterrence effect of punitive damages for intentional torts, there is no such effect of compensatory damages for the negligence rule. This should not be surprising because the negligence rule, in contrast to intentional torts, applies to those scenarios wherein the injurer should have recognized the unsocial risk but failed to do so when engaging in the activity, hence causing accidents. While the risk never came to the

mind of the injurer, the liability derived therefrom would have no effect regardless of whether the injurer engages in the activity.

Gary Schwartz (1994), in a comprehensive survey of economic analysis of tort law literature, found that those fine-tuned economic theories predict very poorly the consequences of various tort rules in practice, while there did exist a weak deterrence effect of the general tort liability. But even this latter finding rather "perform[s] a function similar to that of advertising and other promotional activities in the normal product market" (Coase, 1994, 25–28) because the so-called deterrence effects of the general tort liability have been measured without a causal theory.

3 The Private Autonomous Model

The fundamental problem with the RD model, besides its inconsistencies laid open above, is its intrinsic association with the so-called "externality" theory, i.e., the law should *prima facie* intervene to eliminate those "external effects" which someone's action has on others. But as Coase (1987, 20–30) has pointed out, the value of damage might be less than the benefits of the action. Moreover, the administrative costs of law's intervention might be larger than the gap between the greater value of damage and the benefits of the action. In both cases, the law's intervention would make things worse than before. In short, the externality theory is *prima facie* unsound even in terms of "efficiency" (Chien, 2013; 2016). Yet, even more important it is that this legal effectiveness approach distracts legal economists from inquiring into more relevant scientific issue, such as the fundamental question of why the law is needed? (Pearl & Mackenzie, 2018) The legal effectiveness is only an incidental consequence of the law. The law's effectiveness has rarely been determined by the black letter law alone (Basu, 2018).

Tort law, indeed, is concerned with those damages occurring among people who do not contract for allocating the risks. This "market failure", however, should not immediately lead to the conclusion that a coercive law is needed. There are two reasons for this. First, the *ex post* attribute of tort liability indicates that the deterrence has failed. Second, its attribute of compensating damages rather than penalty (non-monetary or criminal fines) indicates it is not a sanction which is the necessary instrument backing up the law's deterrence effect. That is why Coase (1987, 148) could say that "the lack of a contract" and "the gap in a contract" are the same problem because of high transaction costs.

Private law is autonomy-based and thus its function is to facilitate transactions rather than deter behaviors (Chien, 2025; Hayek, 1945; Zywicki & Sanders, 2008). Tort law, as part of private law is, therefore, simulating a risk-allocating contractual provision which the injurer and the injured would have agreed. This mutually beneficial risk-allocating contractual provision is as follows: the benefits of engaging in the activity (B) should not be less than the expected loss (PL) caused by the activity. In legal terms, the injurer will be liable for compensating the loss if the unrealized benefits of not engaging in the activity are less than the resultant reduced expected loss, i.e., B<PL. This formula implies that the less capable the injurer is, or the more dangerous the activity is, or the less beneficial the activity is, the higher the liability standards are (Chien, 2016). Take driving as example. The less capable driver, such as short sighted, would reduce more expected damage than normal driver would once they reduce their driving speed.

As such, the Hand formula B<PL has been misunderstood by most legal economists. Judge Hand, in the decision *U.S. v. Caroll Towing* (159 F.2d 169 (2d Cir 1947)), did not rule for the injurer based on the victim's negligence of leaving no body on board *per se* which would otherwise have prevented the sinking of the barge. Rather, Judge Hand ruled that the victim

cannot justify his leaving no body on board and thus B is zero meanwhile this happened at war time and in winter which increases the accident-occurring rate P of boat collision and therefore B<PL. It follows that the victim is contributorily negligent.

In contrast to the RD model, this formula would coherently and concisely explain tort law. The reasonable person standard would pose no problem because the injurer with less capability under the formula should bear a higher liability standard, and thus the defense claim of less capability cannot be sustained. The benefits of activity are the factors determining the liability standard; therefore, the harmful precautions problem would not arise. When the activity is abnormally dangerous, i.e., P close to 1, the injurer is almost certainly liable, and thus a strict liability regime emerges. When P is close to 0, in which the harm has not been caused by the activity, the formula would dictate that the injurer is not liable. The contributory negligence which the victim or others contribute would reduce the effectiveness of the injurer's precautions, i.e., P, and thus the injurer might no longer be liable. Furthermore, the damages could be adjusted to P, and therefore comparative negligence regime arises due to the *ex post* compensation as a risk allocation contractual arrangement. The good Samaritan problem is also dismissed, as the harm would still occur even if the injurer is not engaged in the (rescue) activity.

There is another efficiency advantage in *ex post* tort liability compared with *ex ante* regulatory measures. The *ex post* nature of tort liability means the injurer can use what Hayek termed "the knowledge of particular circumstances of time and place" such as speeding up on a clear road without increasing accident risks.

It is puzzling, then, that there would be no tort liability cases if the liability standard had been correctly determined, even on the comparison of activity benefits rather than on precaution costs. There are two explanations for this. According to the first viable option, most humane decisions are not deliberate and thus not always precise (Kahneman, 2011). The other solution would be that people make errors in the process of living (as learning) (Rizzo & Whitman, 2018). In both cases, a tort liability of compensating loss would not over-deter people from engaging in socially beneficial risky activities.

4 Economic Analysis of Law as Legal Science

As Zerbe (2014) pointed out, legal economists should abandon the unreal Walrasian world and the quasi-real world of market failures and externalities, namely, the world of the RD model, and instead focus on institutions reducing transaction costs. In other words, institutions are contractual arrangements reached by people to reduce dissipation of rent (Cheung, 1998). In the world of law, this means that people would trust the courts to delimit rights to resolving disputes, which makes mutual benefits unrealized. Thus, firstly, the courts have to impartial and. secondly, they have to efficiently apply law, resulting in like cases decided alike, different cases decided differently. Therefore, legal dogmatics is required for the courts' credibility and thus the law's effectiveness (Summers, 2005; Basu, 2018). Furthermore, against the criteria of consilience, simplicity, and analogy for scientific theory choice (Thagard, 1978), the PA model, rather than the RD model, is the better explanatory theory of tort law.

4.1 Tort Liability as Prices?

Robert Cooter and Thomas Ulen (2016, 3), in their widely used textbook, *Law and Economics*, claimed that "[e]conomics provided a scientific theory to predict the effects of legal sanctions

on behavior. To economists, sanctions look like prices, and presumably, people respond to these sanctions much as they respond to prices […] Economics has mathematically precise theories (price theory and game theory) and empirically sound methods (statistics and econometrics) for analyzing the effects of the implicit prices that laws attach to behavior." On the theoretical side, this puts the cart before the horse. In terms of science, legal theory should be concerned with "why the law" not "how effective the law". On the empirical side, correlation is not causation (Pearl & Mackenzie, 2018; Shang & Williamson, 2023), and the quantification cannot be value-neutral.

To justify using economics for legal studies, Cooter and Ulen have doubly misused the term "price", the *raison d'état* of economists. In ordinary language, sanctions could be said as prices that the rule violators have to pay. But in economics, price has a specific meaning of consequence or mechanism of voluntary transactions. This is exactly the opposite of sanctions. As Hayek (1945) pointed out, the (rule of) law has emerged in human actions, which is similar to the price system (Zywicki & Sanders, 2008). The law, as a price guiding people's behavior, is derived from the law's stabilizing of people's expectations. To make sure the law is being obeyed, sanctions on lawbreakers, indeed, are needed. But they are derived from the primary law and therefore cannot be the core of legal studies. In economic analysis of tort liability, this sanctions-centered approach has one further defect, namely, liabilities of compensation cannot be sanctions.

Viewing laws as sanctions also makes many empirical legal studies irrelevant for legal studies because the interdependence exists between theory and empirical analysis. Although legal empirical analysts have expanded their testing scope to evidence about the informational environment and evidence prompting expansion of the potential set of options, and thus are closer to real decision-making (Arlen, 2021), they still miss the target of legal studies, that is, law as a price (i.e., norm or rule).

Cost-benefit analysis is an economist's tool to assess "legal policies". Economists claim that this analysis is objective, scientific, and thus value-neutral. However, as Bernard Harcourt (2018) pointed out, this is impossible. One reason why the precautions/deterrence model has failed is its avoidance of comparing the benefits of conflicting activities, which are mostly determined by social norms. For example, Judge Learned Hand said no benefits lost (B=0) because there is no good reason for leaving nobody onboard, which would have reduced the damage and therefore B<PL.

Coase (1994), in his essay, *How Should Economist Choose?*, pointed out the common problems of economic theorizing. On the theoretical side, he said: "[b]ut a theory is not like an airline or bus timetable. We are not interested simply in the accuracy of its predictions. A theory also serves as a base for thinking. It helps us to understand what is going on by enabling us to organize our thoughts." (Coase, 1994, 16–17) On the empirical side, he said most empirical studies do not test the competing theories while only measure the effects of the selected theory, which can be done with the advancement of computational power and statistics. What is more, "if you torture the data enough, nature will always confess." (Coase, 1994, 37)

4.2 Tort Law as Liability Rules?

Calabresi and Melamed (1972, 1128), in their now classic article, concluded that "[l]egal scholars, precisely because they have tended to eschew model building, have often proceeded in an ad hoc way, looking at cases and seeing what categories emerged. But this approach also affords only one view of the Cathedral. It may neglect some relationships among the problems involved

in the cases which model building can perceive, precisely because it does generate boxes, or categories".

From the perspective of whether transactions have been based on voluntariness, Calabresi and Melamed (1972) created three categories of legal rules, i.e., property rules, liability rules, and rules of inalienability. This is a great insight, particularly for legal economists who usually only pay attention to the laws governing market transactions. But as Epstein (1997) pointed out, Calabresi and Melamed have overemphasized the role of liability rules, considering them either in contrast to property rules or devoid of institutional contexts.

Even so, it is incorrect to assume that Calabresi and Melamed (1972) regarded tort liability as belonging to the category of liability rules. In their opinion, the liability rules referred to those rules allowing entitlements to be taken by paying a price determined not by their holders. Tort liability, on the other hand, is usually not an absolute one. Rather, it is a standard of liability for damage that has to be found, such as intentional, negligent, or engaging in dangerous activities. Putting tort liability into the category of liability rules would distract the investigation of the most important issue, namely, what is the liability standard and why is it so? Consequently, the market failure paradigm or externality theory, whose deficiencies have been exposed by Coase (1960), has been sneaked in for the reason of tort liability.

As such, Calabresi and Melamed (1972) have attributed the cause of liability rules to high transaction costs, which block market transactions based on property rules. Therefore, they have eradicated the public law/private law divide and have put tort liability and eminent domain into the same category of liability rules. But, firstly, as Krier and Schwab (1995) pointed out, the concept of high transaction costs is too vague to be operationalized. Secondly, eminent domain is part of public law, which imposes constraints on the government's exercise of the taking power (Epstein 1997, 2111–2120). This dimension would be lost in putting eminent domain in the category of liability rules.

Elaborating on their critics, Calabresi (2016, 117–130) characterizes in his more recent work the liability "rule" as a hybrid of market (the property rule) and command rule (inalienability) to implement what he called "social democratic values". But this argument should be rejected because of its tautology. As such, this argument can justify any amount of damages: when the payments are equal to market price, it is a substitute for market; when the payments are above market price, it is a penalty for inalienability; when the payments are lower than market price, it is an assessment for collective values. More importantly, this argument is institution-free and thus legal classification, such as contract/tort, civil/criminal, or public law/private law, would become meaningless (Rose, 1997); that is, no legal constraints being imposed on rule-choosers. Although Calabresi (1991) invoked the cost-effectiveness used by Coase (1960) to further elaborate on this research question, a proper legal classification is still a prerequisite (Kaitan & Steel, 2023).

4.3 The Institution of Courts

The standard law and economics approach regards law as policy and thus implicitly assumes that the courts are institutionally fit for making a correct cost-benefit analysis. But this cannot be further from the truth (Komesar, 1994; 2001). Not only does the "efficiency" of common law not derive from judges consciously making society-wide cost-benefit analysis (P. Rubin, 2005), but the society-wide cost-benefit analysis would also undermine that very law and thus its "efficiency" (Buchanan, 1974). The economic reasoning by Judge Richard Posner in the *Indiana Harbor* (916 F. 2d 1174 (7th Cir. 1990)) is a good example to illustrate this problem.

In *Indiana Harbor*, the RD model, which Professor Posner has pioneered has led him as a judge to mischaracterize relevant legal issues (Rosenberg, 2007; Sykes, 2007) and misinterpret the reasons for the abnormally dangerous activity doctrine (Bogus, 2023). In that model, strict liability is needed when the activity the injurer engaged in is socially harmful, but it would not be prevented by the negligence rule, which requires the injurer to take optimal care only. In that decision, Judge Posner had done his own investigation and found that the transportation of the dangerous chemical by rail, even through populous areas, is the inevitable choice for the defendant under the then technological conditions and thus the defendant would not have been deterred from engaging in its concerned activity even under strict liability.

In individual cases, judges' investigation into whether the concerned activity is the inevitable choice under constraints, however, would make strict liability meaningless, overstretch the courts' capacity, and complicate the adjudication procedure. Brilliant Judge Posner's investigation might be, but it still could be wrong. For ordinary judges, it would be even harder to get it right. More seriously, the result of this kind of economy-wide investigation, even under professional economists, would still be uncertain and thus would *de facto* deprive the application of strict liability due to the victim bearing the burden of proof. As Rosenberg (2007) pointed out, strict liability is a price that would automatically induce the injurer to take the optimal level of activity, and thus, the courts' inquiry into the issue is superfluous.

One might add that Posner (2003) once criticized Hayek's law and economics precisely for its formalism. As Zywicki and Sanders (2008) pointed out, Posner holds an unrealistic assumption of judicial knowledge and ignores the "efficiency" of rules emerging spontaneously, which facilitates the use of knowledge of society. In the like manner, Bogus (2023) argued, by the vivid example of Jurassic Park, that the reason for strict liability for abnormally dangerous activity is the society acceptance of the abnormally dangerous activities but with the condition that the damage should be compensated once the damage occurred without inquiring into whether negligent or not because the deal between the injurer and the victim has been made under the condition the injurer takes optimal care.

4.4 Legal Dogmatics as the Scientific Foundation of Law

The economic analysis of law has arisen to counter the formalism of doctrinal legal studies (Rubin, 2017). Focusing on investigating the effectiveness of law and thus using quantitative analysis, however, only distracts the scientific studies of legal doctrines: the systematization of legal rules. As Holmes (1897, 62), the pioneer of the law and economics, claimed: "The number of our predictions when generalized and reduced to a system is not unmanageably large. They present themselves as a finite body of dogma which may be mastered within a reasonable time. […] We could reconstruct the corpus from them if all that went before were burned. The use of the earlier reports is mainly historical […]".

This systematization of law would reduce the legal costs of legislation, adjudication, administration, and learning to improve the survival rate of the rule of law, competing with other dispute-resolving measures such as violence, money, or power. And as a scientific study, it has the logical form of the inference to the best explanation (Amaya, 2018; Michelon, 2019; Chao, 2019).

From this perspective, particularly concerning the consilience of the competing principles, whether the RD model or the PA model is a better one for explaining tort liability is an easy case. As I have demonstrated above, unlike the PA model, the RD model cannot, in any meaningful way, explain any tort law doctrine without introducing external elements.

Tort law, as liability rules discussed above in subsection 4.2, poses a different problem. Although the tripartite framework covers a whole legal system and thus superficially has much consilience, the rules that are the *explanandum* do not "belong together", and thus the framework cannot explain these rules (Michelon, 2019, 889–894). For example, tort liability and eminent domain have different normative meanings: the former being autonomy-based private law and the latter being power-restricted public law. Therefore, liability rules cannot be the principle which can explain both tort liability and eminent domain.

The RD model was built to compare the effectiveness of the no liability, negligence rule, and strict liability, although they do not belong together. No liability is not a rule, and strict liability also is not a rule, but an opposite side of no liability, while the negligence rule is a rule properly so-called. The result of applying the negligence rule is either the predicate liable (strict liability) or not (no liability). Therefore, the RD model would be more appropriate to be viewed as seeing the effects of errors in implementing the negligence rule. No liability designates those cases wherein the liability should be imposed, but actually it is not, while strict liability designates those cases wherein the liability should not be imposed, but actually it is.

But how does this relate to legal economics being a science? Friedman (2009) argued that the predictive power rather than the reality of the assumption is the criterion by which a model should be assessed as scientific or not. The assumption of *ex ante* contracting of the private autonomous model is not real, but it has much predictive power. He further clarified that "[t]o avoid confusion, it should perhaps be noted explicitly that the 'predictions' by which the validity of a hypothesis is tested need not be about phenomena that have not yet occurred, that is, need not be forecasts of future events; they may be about phenomena that have occurred but observations on which have not yet been made or are not known to the person making the prediction." (Friedman, 2009, 9)

In the PA model, the fault principle makes not only the separate independent torts of intentional harms, negligence, and strict liability but also the various doctrines of compensation, causation, contributory negligence, duty of care, and others a unified liability rule, becoming known to the legal complexity, i.e., legal practitioners, educators and scholars (Khaitan & Steel, 2023). This is truly scientific.

As Amaya (2018) pointed out, "[i]n the philosophy of science, explanatory coherentism has been advocated as a main alternative to the dominant Bayesian approach to theory choice." For an economic analysis of law, this failure of quantitative analysis, however, has less to do with technical problems concerning constraints on data, computing capacity, or estimation techniques (Levmore, 2021). Rather, it has to do with the normative nature of law, which defies quantitative analysis. Furthermore, the standard economic models of law usually lack mechanisms linking economic efficiency to law and thus further reduce their scientific credibility (Shang & Williamson, 2023). Although the evolutionary models based on either the demand side or the supply side of litigations might mitigate this problem (P. Rubin, 2005), they could only explain common law as a process, and thus they could not shed light on individual legal doctrines, which are the core of legal studies, i.e., legal dogmatics (Summers, 2005).

Aarnio (1994) argued that legal dogmatics is a social science in the sense of studying society as existing in phenomena. He said that "legal dogmatics expressly interprets linkages (norms) through which people's personal relationships become legalized and in this way sociated […] and thus legal dogmatics is a study of these normative 'reciprocating mechanism'." (Aarnio, 1994, 16) Legal dogmatics has two tasks, namely interpreting legal texts and systematizing legal norms which cannot be empirically confirmed or falsified, but can only be judged by their coherence. It is, thus, the above-mentioned missing mechanism of scientific explanation.

The function of law is to let people accept the courts' judgments of right-delineation and thus avoid resorting to non-legal dispute resolutions. A systematization of law and a consistent and coherent interpretation of law would minimize the costs of legislating, adjudicating, and learning law and thus increase the law's chances of winning over violence, money, and power in the arena of dispute resolution.

4.5 The Criteria for Theory Choice

Finally, against the criteria for the theory choice (Thagard, 1978), the PA model is better than the RD model for explaining tort law. Regarding the criterion of consilience, the PA model is not inconsistent with the liability insurance, unknown risk, and legal dogmatics, which the RD model does not cover. In the criterion of simplicity, the PA model does not add *ad hoc* hypotheses as the RD model does for explaining concepts such as the reasonable person standard or intentional torts. Last and not least, the criterion of analogy, the PA model is better than the RD model because tort law is a part of private law, of which private autonomy is a core principle, whereas regulatory deterrence is the function of public law.

5 Conclusion

The RD model is an *ex ante* regulatory model, while the tort law is an *ex post* legal mechanism to allocate the "inevitable" risks. The RD model sees legal liability as an incentive to induce the injurer and the victim to take optimal precautions to minimize the total costs of precautions and risks. If this is the function of tort liability, there would be no accidents that occur for which anyone is responsible because the tort law has already given perfect incentives to people; therefore, no person, being rational, would take under-optimal precautions and then bear tort liability for damages. Empirically speaking, the volumes of tort cases are significant, if not dominant.

By contrast, the PA model of tort liability has been constructed on an implicit contract and ordinary human nature. Accidents, by definition, imply the transaction costs being too high to contract an allocation of risk. The law, therefore, simulates the implicit contract that the injurer and the victim would have agreed on. This provision of allocating risk would be that the injurer's interests unrealized equal the reduced expected damage due to the injurer's disengaging the conflicting activity. This is the equation of $B=PL$. When $B<PL$, the injurer will be liable for the damage. Judging from the human nature of fast thinking, $B<PL$ sometimes would happen; therefore, tort cases arise.

In this study, I intended to emphasize that the subject matter of economic analysis of law studies should be law, not economics. Otherwise, Buchanan's (1959; 1974) comment of "good economics but bad law" or Mestmäcker's (2007) "a legal theory without law" would happen. Law is a complexity system in which private law and public law have different kinds of rent-dissipation to be reduced. Therefore, the RD model of public law cannot be used to analyze tort liability of private law, something befitting the PA model instead.

References

Aarnio, A. (1994). Is Legal Science a Social Science? In W. Krawictz, N. MacCormick, & G. H. von Wright (Eds.), *Prescriptive Formality and Normative Rationality in Modern Legal Systems: Festschrift for Robert S. Summers* (pp. 3–17). Duncker & Humbolt. https://doi.org/10.3790/978-3-428-07895-0

Amaya, A. (2018). Coherence and Systematization in Law. In G. Bongiovanni, G. Postema, A. Rotolo, G. Sartor, C. Valentini, & D. Walton (Eds.), *Handbook of Legal Reasoning and Argumentation* (pp. 637–672). Springer. https://doi.org/10.1007/978-90-481-9452-0_22

Arlen, J. (2021). The Essential Role of Empirical Analysis in Developing Law and Economics Theory. *Yale Journal on Regulation*, 38(2), 480–502. Online: https://tinyurl.com/4uwew2ca

Basu, K. (2018). *The Republic of Beliefs: A New Approach to Law and Economics*. Princeton University Press. https://doi.org/10.2307/j.ctv10qqzs3

Bayern, S. (2023). *The Analytical Failures of Law and Economics.* The Cambridge University Press. https://doi.org/10.1017/9781009159203

Bogus, C. T. (2023). Why *Indiana Harbor* Is the Worst Torts Decision in American History. *Connecticut Law Review*, 55(3), 649–680. Online: https://tinyurl.com/2hwv97za

Buchanan, J. M. (1959). Positive Economics, Welfare Economics, and Political Economy. *Journal of Law and Economics*, 2, 124–138. https://doi.org/10.1086/466556

Buchanan, J. M. (1974). Good Economics. Bad Law. *Virginia Law Review*, 60(3), 483–492. https://doi.org/10.2307/1072438

Brown, J. P. (1973). Toward an Economic Theory of Liability. *Journal of Legal Studies*, 2(2), 323–349. https://doi.org/10.1086/467501

Calabresi, G. (1991). The Pointlessness of Pareto: Carrying Coase Further. *The Yale Law Journal*, 100(5), 1211–1237. https://doi.org/10.2307/796691

Calabresi, G. (2016). *The Future of Law and Economics: Essays in Reform and Recollection.* Yale University Press. https://doi.org/10.12987/yale/9780300195897.001.0001

Calabresi, G., & Melamed, A. D. (1972). Property Rules, Liability Rules, and Inalienability: One View of the Cathedral. *Harvard Law Review*, 85(6), 1089–1128. https://doi.org/10.2307/1340059

Chao, H.-K. (2019). Inference to the Best Model of the Consumption Function. *History of Political Economy*, 51(3), 493–513. https://doi.org/10.1215/00182702-7551888

Cheung, S. N. S. (1992). On The New Institutional Economics. In L. Werin, & H. Wijkander (Eds.), *Contract Economics* (pp. 48–65). Blackwell.

Cheung, S. N. S. (1998). The Transaction Costs Paradigm. 1998 Presidential Address Western Economic Association. *Economic Inquiry*, 36(4), 514–521. https://doi.org/10.1111/j.1465-7295.1998.tb01733.x

Chien, T.-S. (2013). The Legal Meaning of Coasean Economics. *Peking University Law Journal*, 1(1), 197–214. https://doi.org/10.5235/205174813807351636

Chien, T.-S. (2016). The Evolution of Tort Law in Taiwan. A Doctrinal-Economic Interpretation. In Y. Chang, W. Shen, & W. Wang (Eds.), *Private Law in China and Taiwan: Legal and Economic Analyses* (pp. 156–181). Cambridge University Press. https://doi.org/10.1017/9781316658925.007

Chien, T.-S. (2025). Economic Analysis of Private Law in Taiwan: A Return to the Private Autonomy Principle. *Asian Journal of Law and Economics*, 16(1), 87–101. https://doi.org/10.1515/ajle-2024-2004

Coase, R. H. (1960). The Problem of Social Cost. *Journal of Law and Economics*, 3, 1–44. https://doi.org/10.1086/466560

Coase, R. H. (1987). *The Firm, the Market, and the Law*. University of Chicago Press.
Coase, R. H. (1994). *Essays on Economics and Economists*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226051345.001.0001
Cooter, R. (1984). Prices and Sanctions. *Columbia Law Review*, 84(6), 1523–1560. https://doi.org/10.2307/1122472
Cooter, R. (1987). Torts as the Union of Liberty and Efficiency: An Essay on Causation. *Chicago-Kent Law Review*, 63(3), 523–551. Online: https://tinyurl.com/55uvdbze
Cooter, R. D., & Porat, A. (2000). Does Risk to Oneself Increase the Care Owed to Others? Law and Economics in Conflict. *The Journal of Legal Studies*, 29(1), 19–34. https://doi.org/10.1086/468062
Cooter, R. D., & Porat, A. (2014). *Getting Incentives Right: Improving Torts, Contracts, and Restitution*. Princeton University Press. https://doi.org/10.23943/princeton/9780691151595.001.0001
Cooter, R. D., & Ulen, T. (2016). *Law and Economics*. (6th ed.). Addison-Wesley. Online: https://tinyurl.com/mvwf2f4c
Epstein, R. A. (1997). The Clear View of the Cathedral: The Dominance of Property Rules. *Yale Law Journal*, 106(7), 2091–2120. https://doi.org/10.2307/797162
Friedman, M. (2009). The methodology of positive economics (1953). In U. Mäki (Ed.), *The Methodology of Positive Economics: Reflections on the Milton Friedman Legacy* (pp. 3–44). Cambridge University Press. https://doi.org/10.1017/CBO9780511581427.002
Garoupa, N., & Ulen, T. S. (2013). The Economics of Activity Levels in Tort Liability and Regulation. In T. J. Miceli, & M. J. Baker (Eds.), *Research Handbook on Economic Models of Law* (pp. 33–53). Edward Elgar. https://doi.org/10.4337/9781781000151.00007
Grady, M. F. (1989). Untaken Precautions. *Journal of Legal Studies*, 18(1), 139–156. https://doi.org/10.1086/468143
Grady, M. F. (2009). Unavoidable Accident. *Review of Law and Economics*, 5(1), 177–231. https://doi.org/10.2202/1555-5879.1302
Hayek, F. A. (1945). The Use of Knowledge in Society. *American Economic Review*, 35(4), 519–530. Online: https://www.jstor.org/stable/1809376
Holmes, O. W., Jr. (1897). The Path of the Law. *Harvard Law Review*, 10(8), 457–478.
Kahneman, D. (2011). *Thinking, Fast and Slow*. Allen Lane.
Khaitan, T., & Steel, S. (2023). Areas of Law: Three Questions in Special Jurisprudence. *Oxford Journal of Legal Studies*, 43(1), 76–96. https://doi.org/10.1093/ojls/gqac025
Komesar, N. K. (1994). *Imperfect Alternatives. Choosing Institutions in Law, Economics, and Public Policy*. University of Chicago Press.
Komesar, N. K. (2001). *Law's Limits. The Rule of Law and the Supply and Demand of Rights*. Cambridge University Press.
Krier, J. E., & Schwab, S. J. (1995). Property Rules and Liability Rules: The Cathedral in Another Light. *New York University Law Review*, 70, 440–483. Online: https://tinyurl.com/3w9pknu5
Landes, W. M., & Posner, R. A. (1987). *The Economic Structure of Tort Law*. Harvard University Press. https://doi.org/10.4159/harvard.9780674864030
Levmore, S. (2021). The Eventual Decline of Empirical Law and Economics. *Yale Journal on Regulation*, 38(2), 612–632. Online: https://tinyurl.com/mwbpb3zn
Mestmäcker, E.-J. (2007). *A Legal Theory Without Law. Posner v. Hayek on Economic Analysis of Law*. Mohr Siebeck. https://doi.org/10.1628/978-3-16-151072-4
Michelon, C. (2019). The Inference to the Best Legal Explanation. *Oxford Journal of Legal Studies*, 39(4), 878–900. https://doi.org/10.1093/ojls/gqz021

Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books.

Perry, R. (2023). Harmful Precautions. *Notre Dame Law Review*, 99(1), 153–202. Online: https://scholarship.law.nd.edu/ndlr/vol99/iss1/4/

Posner, E. A. (2021). The Boundaries of Normative Law and Economics. *Yale Journal on Regulation*, 38(2), 657–677. Online: https://tinyurl.com/49zf657u

Posner, R. A. (2003). *Law, Pragmatism, and Democracy*. Harvard University Press. https://doi.org/10.4159/9780674042292

Rizzo, M. J., & Whitman, G. (2018). Rationality as a Process. *Review of Behavior Economics*, 5(3–4), 201–219. https://doi.org/10.1561/105.00000098

Rose, C. M. (1997). The Shadow of The Cathedral. *The Yale Law Journal*, 106(7), 2175–2200. https://doi.org/10.2307/797165

Rosenberg, D. (2007). The Judicial Posner on Negligence versus Strict Liability: *Indiana Harbor Belt Railroad Co. American Cyanamid Co*. *Harvard Law Review*, 120(5), 1210–1222. Online: https://tinyurl.com/3zazax9u

Rubin, P. H. (2005). Micro and Macro Legal Efficiency: Supply and Demand. *Supreme Court Economic Review*, 13, 19–34. https://doi.org/10.1086/scer.13.3655299

Rubin, E. L. (2017). From Coherence to Effectiveness: A Legal Methodology for the Modern World. In R. van Gestel, H.-W. Micklitz, & E. L. Rubin (Eds.), *Rethinking Legal Scholarship: A Transatlantic Dialogue* (pp. 310–350). Cambridge University Press. https://doi.org/10.1017/9781316442906.009

Schäfer, H.-B., & Ott, C. (2022). *Economic Analysis of Civil Law*. (2nd ed., translated by M. Braham). Edward Elgar. https://doi.org/10.4337/9780857935076

Schwartz, G. T. (1994). Reality in the Economic Analysis of Tort Law: Does Tort Law Really Deter. *UCLA Law Review*, 42(2), 377–444.

Shang, Y., & Williamson, J. (2023). *Evidential Pluralism in Social Sciences*. Routledge. https://doi.org/10.4324/9781003143000

Shao, J.-H., & Weng, M.-H. (2023). The Deterrence Effect of Tort Damages: An Empirical Study via Public Bad Experiment.［邵靖惠，翁明宏，侵權行為損害賠償之預防目的—以法律實證分析為中心] *National Chengchi Law Review* [政大法學評論], 173, 157–226. https://doi.org/10.53106/102398202023060173003

Shavell, S. (1980). Strict Liability versus Negligence. *The Journal of Legal Studies*, 9(1), 1–25. https://doi.org/10.1086/467626

Shavell, S. (1987). *Economic Analysis of Accident Law*. Harvard University Press. https://doi.org/10.4159/9780674043510

Summers, R. S. (2005). *Form and Function in a Legal System: A General Study.* Cambridge University Press. https://doi.org/10.1017/cbo9780511511066

Sykes, A. O. (2007). Strict Liability versus Negligence in *Indiana Harbor. The University of Chicago Law Review*, 74, 1911–1931. Online: https://tinyurl.com/664w68t9

Thagard, P. R. (1978). The Best Explanation: Criteria for Theory Choice. *The Journal of Philosophy*, 75(2), 76–92. https://doi.org/10.2307/2025686

Zerbe, R. O. (2014). *Economic Efficiency in Law and Economics*. Edward Elgar. https://doi.org/10.4337/9781843761488

Zywicki, T. J., & Sanders, A. B. (2008). Poser, Hayek, and the Economic Analysis of Law. *Iowa Law Review*, 93(2), 559–603. Online: https://tinyurl.com/52yuya4h

Institutiones Administrationis – Journal of Administrative Sciences Vol. 6 (2026) No. 1, 87–104
DOI: 10.54201/iajas.216

Identity, Integration, and Loyalty in Estonia: A Relational Comparative Analysis of Immigrant Integration

Ivan Polynin*

* Junior Research Fellow, School of Governance, Law and Society, Tallinn University, Tallinn, Estonia. E-mail: ivan.polynin@tlu.ee

Abstract

This article presents a relational comparative analysis of Estonian integration, building directly on Marju Lauristin's cluster analysis and national defense index in the Estonian Integration Monitoring 2023. Drawing on relational integration theory and John W. Berry's acculturation framework, it maps Lauristin's clusters onto Berry's four strategies and traces ethnic differences in trust, belonging, support for NATO, and military aid to Ukraine. The analysis identifies important limitations in Lauristin's symmetric design: even when integration is framed as a two-sided process, asymmetrical power relations persist, and acculturation strategies continue to collapse into assimilation, separation, or marginalization. The post-2022 security context amplified these pre-existing cleavages. Vetik's national unity-versus-equal-rights imperatives remain central, yet the majority has claimed the moral high ground by framing strong support for Ukraine as a loyalty test. And the relational lens shows that this securitization of integration has proved counterproductive, widening trust gaps across the whole society.

Keywords

relational integration, immigrant integration, loyalty, Estonia, Russian-speaking minority, securitization

1 Introduction

This study is a practical attempt to resolve a challenge faced by every researcher at a certain stage of their career: the transition from theory to empirical analysis. For some areas of the social sciences, this transition is more difficult than for others; one such area is research on immigrant integration, which is also sometimes called migrant integration or ethnic minority integration. The language behind the phrase "immigrant integration" already implies a single direction to the process, or some invisible force that takes migrants and inserts them into a larger society (Favell, 2019). For that reason, some scholars, including myself, have reservations about using the term but prefer to adhere to it due to the lack of a better alternative. For example, the notion of "social integration" excessively stretches the concept, making it too broad for a meaningful comparison (Polynin, 2024).

 Published online: 25 May 2026

In contrast, postcolonial scholars tend to resolve the problem of the term "immigrant integration" in a more drastic fashion, namely by calling for its complete repeal (Schinkel, 2018; Favell, 2019). They argue that research on integration should not be reduced to finding better methods for turning migrants into non-migrants, as this reinforces racism and prejudice and places migrants at the bottom of an imaginary social hierarchy (Schinkel, 2018; Favell, 2019). Despite the radical nature of the postcolonial critique, it raises highly important questions, such as, how can integration be studied in a non-reductionist way? How can it be reflected as a two-way process? Or what is there to integration besides a pathway for ethnic minorities losing themselves in the majority?

I address these questions in a relational way by analyzing the case of immigrant integration in Estonia. This article proceeds in three main parts. First, it outlines the theoretical departure from Alba's mainstream approach toward Klarenbeek's relational integration framework and introduces the key analytical concepts employed in the study: the distinction between assimilation and integration, Berry's acculturation strategies, and Vetik's loyalty-versus-equal-rights imperatives. Second, the empirical analysis is divided into two steps. It first examines Marju Lauristin's integration clusters in the Estonian Integration Monitoring 2023 and maps them onto Berry's acculturation strategies. It then conducts a relational comparative analysis of the 2023 Integration Monitoring and the 2025 Public Opinion Monitoring Survey, evaluating the components of Lauristin's national defense index, such as trust, belonging, support for NATO, and military aid to Ukraine against the loyalty-versus-equal-rights imperatives. The article concludes with a discussion of the unreflected shift toward a loyalty paradigm in Estonian integration research and policy and its broader implications for Estonian society.

2 Theoretical Framework and Methodology of Relational Integration

2.1 Critique of Mainstream Approaches: Assimilation versus Integration

Notwithstanding my previous critique of the mainstream approach to the theory of immigrant integration (Polynin, 2024), particularly represented by Alba (2024) and Statham and Foner (2024), I emphasize that it also makes an important theoretical statement by picturing integration as a two-way process (Polynin, 2024, 9). However, the key problem of Alba's stance reveals itself when moving from the theoretical to the empirical level and requires deeper consideration. Alba and others conflate assimilation and integration by using them interchangeably. Thereby, assimilation is defined as "the decline of an ethnic distinction and its corollary cultural and social differences" (Alba & Nee, 2003), while integration is "approximately equivalent" (Alba, 2024, 28) to it. Moreover, Alba emphasizes that assimilation does not require the total loss of ethnic identity. Cultural practices and ethnic markers can remain intact, even as their significance in shaping social interactions diminishes over time (Alba, 2020, 146). In other words, under the lens of this approach, the ethnic differences do not disappear; they simply stop mattering.

As I have already argued, using integration and assimilation interchangeably completely misrepresents what these processes mean (Polynin, 2024, 2). But why is this problematic both empirically and normatively? The answer is quite straightforward: when one appeals to dictionaries[1] or to alternative influential sources, assimilation means forcing a migrant to give up their heritage,

[1] https://dictionary.cambridge.org/dictionary/english/integration

culture, and identity, while integration is defined by the opposite: an attempt to retain them and to use them as the basis for participation in the larger society (Berry, 1970). These are not the same processes, and without drawing a clear line between them, researchers lose any effective way to explain the real-life problems that migrants and the larger society are going through.

To illustrate this point, let us assume that we take Alba's conceptualization of assimilation for granted. If assimilation means "entry into the societal mainstream" (Alba, 2024, 28) where the latter is defined as "a part of the society where the influence of ethnoracial origins on social interaction and position is greatly reduced" (Alba, 2024, 28; Alba & Nee, 2003, 12), then the entire logic of integration as a two-way process is completely lost. In this scenario, researchers focus exclusively on how migrants enter the societal mainstream, bearing the full burden of change. Moreover, the analytical need to distinguish assimilation from integration arises as soon as we introduce discrimination and prejudice into the equation. For example, Alba's (2024) focus on individual outcomes (e.g., intermarriage, job access) might overlook cases where migrants adopt mainstream behaviors but still face discrimination due to unchangeable traits such as an accent or race, which suggests incomplete assimilation rather than failed integration.

From a normative point of view, the equalization of integration and assimilation on the premise that they lead to a similar result is inherently cynical. If assimilation is reduced to the mere erasure of ethnic differences, as Alba's framework suggests, it dismisses the questions of why the change occurs and at what cost – whether it stems from coercion, cultural suppression, or voluntary adaptation. As a result, even when tackling the decline of ethnic distinctions, Alba's framework overlooks essential questions, such as: how do these differences vanish? What role do government policies, e.g., immigration laws or citizenship requirements, play in driving this process? Or to what degree is change imposed rather than chosen?

By sidestepping these issues, Alba treats assimilation as an inevitable, almost natural phenomenon, stripped of its political context, which is often manifested in negative forms such as discriminatory rhetoric, legalized segregation, or structural inequalities that shape migrant experiences. I must emphasize that while Alba's contribution is clearly not cynical in intention, his focus on outcomes glosses over the real struggles and trade-offs involved in the relationship between migrants and the majority. So, the question evidently arises: what can be done about it?

2.2 Relational Methodology: From Theory to Data

This article is much more than an argument in favor of the difference between integration and assimilation, as this particular goal may be achieved while pursuing a solution to a larger problem. As I have demonstrated (Polynin, 2024), the postcolonial and relational critiques of the mainstream approach to immigrant integration agree that the mainstream reduces integration to the study of migrants becoming non-migrants, thereby placing them by default in a subordinate position. Similarly, both postcolonial and relational approaches highlight the nation-state-centered logic of the mainstream, calling it methodological nationalism or civic integrationism, respectively (Klarenbeek, 2024; Favell, 2019).

The term "civic integrationism" describes one of the two ways in which this logic is expressed in mainstream scholarship. On the one hand, it means producing policy-oriented research that shapes official government integration discourse. Postcolonial researchers such as Favell call this particular form of knowledge production "thinking-for-the-state", which is separated from an autonomous sociology that remains unaffected by the government's influence (Favell, 2019, 4). On the other hand, the research questions posed within the civic-integrationist discourse traditionally revolve around protecting the values of the majority, national security, and the

survival of the nation (Klarenbeek, 2024, 235), which makes the entire approach politically loaded and provokes objections from opposing scholars.

I have already presented a detailed review of the postcolonial and relational critiques of the mainstream approach (Polynin, 2024), but in both respects, my analysis was mostly theoretical. In this article, I would like to draw on real empirical material and studies from the case of Estonia, one of the most economically successful post-Soviet states, having an uneven integration history with its large Russian-speaking minority that accounts for roughly one-quarter of its 1.3 million population.[2] Nonetheless, I aim to go beyond presenting a single case study of a unique regional experience and instead offer a deeper and more broadly comparable insight into how theoretical concepts affect the way we study integration. In other words, this article aims to present a transparent, traceable, and replicable way of translating theory into practice, something with which even established scholars often struggle.

I would like to call the methodological approach adopted in this study a relational comparative analysis, grounded in Klarenbeek's relational integration theory (Klarenbeek, 2019; 2024). This framework shifts the focus from traditional one-way integration models, studying how migrants lose the salience of their identity, to a new approach focusing instead on how the majority and minorities negotiate the social boundaries affected by migration.

Relational scholars in general conceptualize societal relations as hierarchical and contested, emphasizing asymmetries in power, recognition, and social standing (Emirbayer, 1997; Tilly, 2005; Vetik, 2019; Klarenbeek, 2024) rather than individual migrant outcomes alone (e.g., language proficiency, economic welfare, etc.). Klarenbeek's approach builds on this by advancing a strong normative dimension: it acknowledges that migrants might end up in a subordinate position in society, but clearly states that imposing such inferiority is morally wrong (Klarenbeek, 2024, 234). She urges that integration problems be treated as part of a larger picture of relational inequality expressed in the relations through which society distributes material goods, freedoms, and opportunities (Klarenbeek, 2024, 238).

Klarenbeek (2024, 238) claims that, in its current form, relational integration theory lacks a sociological component and therefore offers no empirical explanation of how integration really happens. I connect the methodological struggles of her approach to the evasive notion of relations. It is not impossible, but it is highly difficult, to rely on relations as a unit of analysis, because they are always dynamic and therefore difficult to capture, measure, or express in any empirical form, leaving us with no on-the-ground understanding of the integration process. The mainstream approach offers an easy solution for measuring benchmarks, the most popular being language skills, family ties, or socio-economic positions. Even though these indicators are important, they do not constitute the notion of integration in its entirety.

As I have argued before (Polynin, 2024, 4), achieving adaptation or the skills needed to survive and function in a certain society does not necessarily mean being integrated. Migrants may suffer exclusion from the host society due to racial, ethnic, religious, or any other form of prejudice, preventing them from developing a sense of belonging and a common identity with the majority. That is why I argue that, in order to capture relations, we need to look at what is happening with collective identities and the integration strategies associated with them.

That being said, I must emphasize that relational integration should not repeat the postcolonial dismissal of the mainstream approach. While Favell's (2019; 2022) idea of "planetary integration", understood as total equality and diversity, is an interesting theoretical exercise, nation-states

[2] https://www.stat.ee

remain the primary entities where integration takes place. Migration and integration policies form and are being formed by an ever-moving discourse that includes public polls, election results, media narratives, and, among other things, sociological research: research that in some cases is produced independently and in others is commissioned by the government itself.

To dismiss any research on the grounds of its association with the government as completely non-autonomous is as much of a mistake as to entirely ignore educational and economic achievements as a component of integration. Therefore, in my interpretation, the task of the relational approach is to acknowledge the potential influences and power relations that shape both the integration process and our perception of it.

In order to bridge the gap between relational theory and its practical application, I apply comparative analysis, but not in the traditional form of Mill's methods of agreement or difference. Instead of variables, I look for relations, juxtaposing indicators and conceptualizations of integration across ethnic groups, primarily Estonians and Estonian Russophones (the Russian-speaking population of Estonia). The analysis also incorporates a reflective dimension, as I argue that examining how integration is studied can yield as much insight as the data from quantitative surveys themselves.

Through this approach, I seek to uncover unreflected shifts in how Estonian authorities and the associated research conceptualize integration, particularly the increasing emphasis on loyalty and the use of attitudes toward the war in Ukraine as a proxy. An important part of this analysis is to check the results of the comparison against Berry's acculturation strategies – assimilation (adopting the host culture while abandoning one's own), integration (maintaining heritage while engaging with the host culture), separation (preserving heritage while avoiding the host culture), and marginalization (detaching from both) – within the relational lens. This enables an examination of how these individual and group-level strategies intersect with societal expectations of loyalty, security, and boundary negotiation, particularly in Estonia's unique post-Soviet context (Berry, 1997; 2019; Grigoryev et al., 2023; Vetik, 2019).

3 Integration Clusters and Acculturation Strategies

3.1 Data Sources and Lauristin's Symmetric Cluster Design

For the purpose of this article, I draw on two sources of survey data. First, the Estonian Integration Monitoring 2023 (hereinafter: EIM 2023), a comprehensive, theoretically grounded study conducted by Kantar Emor for the Ministry of Culture of Estonia. With a representative sample of just over 1,500 respondents (Kultuuriministeerium, 2023, 6), EIM covers a variety of topics ranging from integration typologies, language proficiency, citizenship shifts, trust in institutions, economic well-being, security, and loyalty. Second, the Public Opinion Monitoring Survey's 25th wave, commissioned by the State Chancellery, is a representative quarterly poll of 1,418 respondents (Riigikantselei, 2025a, 4) that reveals broad societal attitudes on economic coping, feelings of security, support for Ukraine, institutional trust, and media habits.

Where relevant, the article also draws on the results of the September wave of the same survey for supplementary longitudinal comparisons (Riigikantselei, 2025b). Both surveys are longitudinal and have relatively comparable sample demographics (by gender, age, income, regional variation, etc.) and include representatives of all primary ethnic groups living in Estonia, including Estonians, Russian speakers, and other nationalities. The primary difference, however, lies in their scope and focus. EIM 2023 belongs to a series of analyses of ethnic integration processes and features normative and empirical innovations such as comparing integration

clusters across groups and addressing post-Soviet challenges like segregation and polarization. In contrast, the Public Opinion Monitoring Survey (henceforth: Monitoring Survey) has little to no theoretical foundation and offers cross-sectional snapshots of general public sentiments across diverse topics, which enables useful comparisons but lacks EIM's specialized emphasis on integration morphologies.

My reasoning for this choice is multi-layered and is largely dictated by the scope of this article. Both EIM and the Monitoring Survey contain overlapping sections that address trust in government institutions, feelings of security in Estonia, and attitudes toward supporting Ukraine. By comparing the responses and the reasoning behind the questions, it becomes easier to understand not only the dynamics between integration and security but also the relationship between more fundamental theoretical research, such as EIM, and more narrowly focused empirical surveys such as the Monitoring Survey.

It must also be noted that EIM is a massive multi-section study conducted by multiple researchers, so the discussion of each segment may warrant a separate article. That is why, guided by the aforementioned theoretical and methodological choices, I pay closer attention to the relevant sections of chapters 2 (Lõimumisprotsessi edukus ja Eesti identiteet) and 3 (Lõimumise klasteranalüüs) of EIM, written by Marju Lauristin (Kultuuriministeerium, 2023, 16–47). Her theoretical approach, as formulated in these sections of EIM, and specifically the ties between the integration clusters and the national defense index, is of particular interest to this study.

The cluster analysis in EIM 2023 represented an evolved methodological tool first introduced in 2011 by Lauristin and Vihalemm (Kultuuriministeerium, 2011) to characterize integration processes through typologies. These typologies combined key metrics such as language proficiency, citizenship, and societal participation into five distinct clusters that captured diverse integration patterns beyond isolated indicators. Although the 2023 version was built on refinements introduced in 2015 (Kultuuriministeerium, 2015), it abandoned the previously introduced rigid criteria in order to implement a new logic for forming the clusters. Lauristin emphasized that before EIM 2023, Estonians and Russian-speaking respondents had been treated asymmetrically. In her words, Estonians used to be framed as the dominant party initiating integration policy, while the Russian-speaking population of Estonia was left with the role of the target of this policy rather than one of the actors (Kultuuriministeerium, 2023, 29). Through EIM 2023, Lauristin pursued the goal of correcting these former asymmetries by treating Estonians and non-Estonians as equal partners rather than automatically assigning the former a position of power.

She also justified that choice by noting that, over the course of the past thirty years, one of the consequences of the integration of Russophones has been the growth in the number of Estonian citizens who speak Russian as their mother tongue. This development has led to the internal differentiation of the Russian-speaking minority through economic stratification and in terms of professional and political activity. Therefore, the new symmetrical approach would better reflect the new equilibrium (Kultuuriministeerium, 2023, 29).

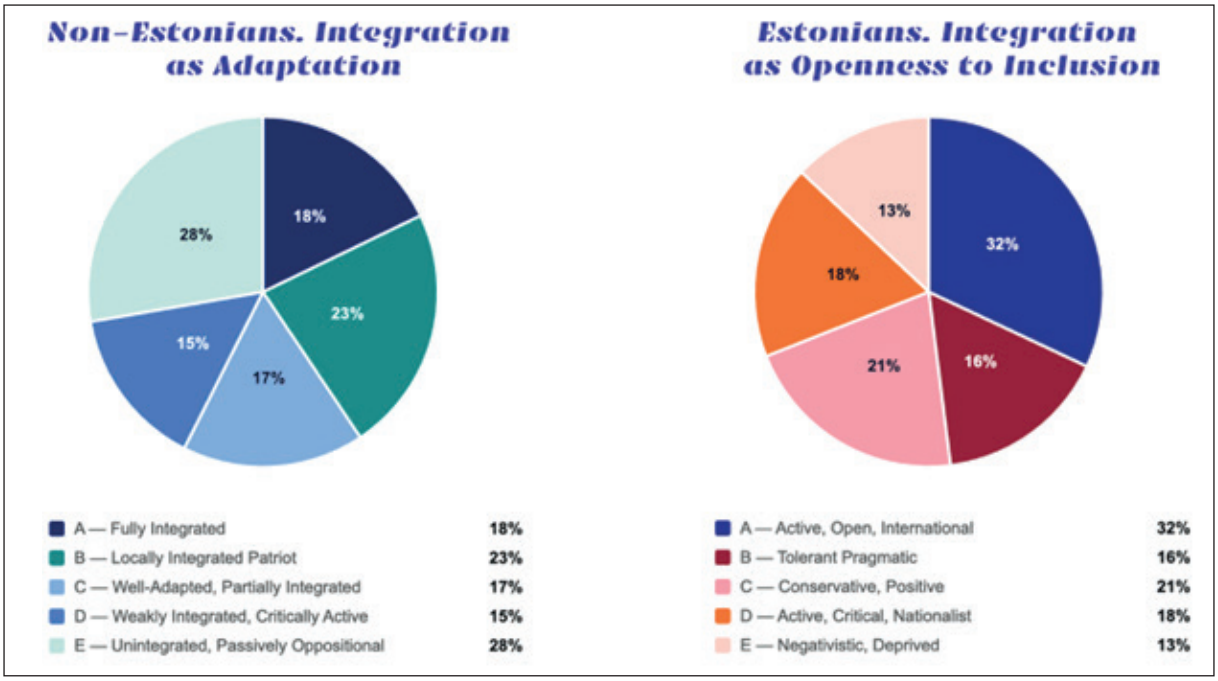
Even though I deeply share the sentiment and direction of striving for relational equality in integration research, I reckon that Lauristin's methodology achieved slightly different goals than initially intended. First and foremost, treating Estonians and non-Estonians symmetrically blurs the goal of relational equality. I definitely agree with framing everyone engaged in integration as equal partners, whether in research or policy, but bypassing the dominant position of the majority and the structural barriers faced by the minority hinders the analysis rather than helps it. By accounting for the asymmetrical power relations between the majority and the minority in any country, we give meaning to existing problems and provide crucial context for why respondents choose particular answers.

For the Estonian case, in his previous works, Vetik (2019) presented a valuable example of Russian-speaking and Estonian respondents giving an affirmative answer to the question of whether everyone should have equal opportunities in the nationwide labor market. He noted that for an Estonian, answering "yes" means giving something up from their group's dominant position, while for a Russophone, it means laying claim to something not yet achieved, striving for truly equal rights. Consequently, the same answers led to quite distinct integration strategies (Vetik, 2019, 411–412). For Lauristin's analysis, this meant that the chosen symmetrical approach would still reveal the same asymmetries but without reflecting on why they came about.

3.2 Cluster Profiles for Non-Estonians and Estonians

Upon comparing the logic behind the cluster formation and the results it yielded, I concluded that the clusters remained rather asymmetrical despite efforts to the contrary. Lauristin grouped respondents into five distinct typologies using a mix of shared and group-specific measurements. The common indicators applied to both Estonians and non-Estonians (primarily Russian speakers) included economic and social capital (e.g., life satisfaction, income adequacy, societal influence), Estonian identity and state relations (e.g., institutional trust, voting activity), integration-related changes and attitudes (e.g., discrimination perception, cultural participation, tolerance toward diversity), and stances on reforms (e.g., openness to change, protest involvement, defense attitudes) (Kultuuriministeerium, 2023, 30).

Figure 1. Integration clusters for Non-Estonians and Estonians[3]



Source: Kultuuriministeerium (2023, 30–39)

[3] The percentages for non-Estonians round to 101% as in the original report.
The description of the clusters "Integration as adaptation" and "Integration as Openness to Inclusion" are the interpretations introduced in this study, not the original analysis.

For non-Estonians, the additional unique factor measuring integration was Estonian language proficiency, while Estonian citizenship was treated more as a demographic criterion that added no integration scores by itself – a rather thoughtful choice that goes against the mainstream paradigm automatically assuming Estonian citizens are more integrated. For ethnic Estonians, the "inclusion clusters" (kaasamisklastrid) focused even more on readiness to facilitate non-Estonians' inclusion and recognition as co-participants in societal decision-making (Kultuuriministeerium, 2023, 38–39). Thus, as promised, Lauristin's approach sought to highlight relational dynamics by treating both groups as active agents rather than positioning Estonians solely as initiators and non-Estonians as passive targets.

While, in contrast to mainstream top-down models, Lauristin's cluster analysis genuinely represents integration as a two-sided process (Kultuuriministeerium, 2023, 29), I argue that it could not have achieved full symmetry in its approach to both groups, and, moreover, that it did not need to. The key asymmetry in the cluster division is that, for non-Estonians, integration was framed as adaptation, relying on indicators such as skills and achievements, while for Estonians, integration was formulated as openness to inclusion (Kultuuriministeerium, 2023, 30). The non-Estonian clusters, though not openly stated by Lauristin in such terms, may be separated into three "well-adapted" and two "disengaged" ones (Kultuuriministeerium, 2023, 31).

The three well-adapted clusters account for 58% overall: the fully integrated cluster A (18%) was marked by high language use, strong Estonian identity, institutional trust, political/cultural activity, and a young, highly educated demographic that predominantly consumes Estonian media (Kultuuriministeerium, 2023, 31). The second cluster of well-adapted individuals, locally integrated patriots (cluster B, 23%), featured strong state loyalty and local ties but lower political participation and consisted of older respondents with mixed citizenship and a larger number from the Ida-Virumaa region (Kultuuriministeerium, 2023, 31). Finally, the partially integrated cluster C (17%) was characterized by pragmatic multilingualism, positive life satisfaction, and moderate political caution (Kultuuriministeerium, 2023, 31).

The remaining 42% of the survey respondents fell into the "disengaged" groups: the weakly integrated, critically active cluster D (15%) primarily consisted of individuals with Estonian citizenship (60%) who nevertheless displayed low trust in state institutions and higher-than-average discrimination sensitivity (Kultuuriministeerium, 2023, 32). Lauristin also emphasized that, despite being small in number, this group had an elevated presence on social media (Kultuuriministeerium, 2023, 32). The largest unintegrated, passively oppositional cluster E (28%) was defined by Soviet-era mindsets, minimal participation, low identity, and heavy reliance on Russian media among older, lower-educated pensioners (Kultuuriministeerium, 2023, 32).

Estonians received different treatment, as their clusters were grouped based on openness to inclusion, that is, attitudes such as the readiness to facilitate minority participation in the life of society and to guarantee equal access to goods and services (Kultuuriministeerium, 2023, 38). This approach yielded a more favorable division, with 69% of respondents comprising three groups that expressed a certain degree of support for integration (Kultuuriministeerium, 2023, 38). The dominant active, open, international cluster A (32%) was formed by younger, educated city dwellers who actively engage with diversity, multilingual contacts, and cultural richness while maintaining a strong Estonian identity (Kultuuriministeerium, 2023, 39). This was followed by the tolerant pragmatic cluster B (16%), which showed moderate indicators on most issues; however, because this cluster included multiple Russian speakers, it demonstrated a notable uncertainty regarding Estonian identity (Kultuuriministeerium, 2023, 39). Overall, this cluster expressed a positive pragmatic attitude toward the inclusion of other nationalities, hence its name. The last pro-integration cluster was the conservative positive cluster C (21%), with the

majority being older rural respondents who trust the Estonian state, value language and traditions, and support inclusion primarily in non-political spheres (Kultuuriministeerium, 2023, 39).

The remaining 31% of Estonian respondents, or two groups, demonstrated that they were somewhat or decisively against integration (Kultuuriministeerium, 2023, 38). The active, critical nationalist cluster D (18%) consisted mostly of urbanites who, in contrast to their counterparts in cluster A, demonstrated nationalist beliefs, primarily emphasizing language skills and loyalty (Kultuuriministeerium, 2023, 39). This group, which had a predominantly middle-aged public-sector employee demographic, was critical of state national policies, including education and integration reforms, with the sole exception of defense policy and military support for Ukraine (Kultuuriministeerium, 2023, 39). Interestingly, the critical nationalists indicated the same low level of trust in Estonian state institutions as cluster D in the non-Estonian section (Kultuuriministeerium, 2023, 41). The smallest negativistic deprived cluster E (13%), composed of residents of small towns and rural areas, reflected strong economic pessimism, as 20% of them reported difficulties coping. They also emerged as the most dissatisfied group of respondents, expressing the opinion that Estonians, in comparison with other nationalities, are treated unfairly when it comes to the distribution of public goods and opportunities (Kultuuriministeerium, 2023, 39). This group showed extremely high levels of distrust toward all public institutions, demonstrating that a negative attitude toward integration is part of a larger negative predisposition.

3.3 Discussion and Mapping onto Berry's Acculturation Strategies

The integration cluster approach helps to process a large volume of accumulated data, and it is important to stress that in this particular article, I analyzed only a small portion of the available empirical material from EIM 2023. Nevertheless, this specific portion is of utmost importance. From the normative point of view, although Lauristin's cluster analysis offers a two-way perspective on integration and thus avoids the one-sided burden on minorities, the unreflected asymmetrical power relations present several problems. Non-Estonians are judged on personal adaptation metrics, while Estonians are evaluated on attitudinal openness, which highlights relational hierarchies because intention unsupported by action will always yield more favorable results.

Another important conclusion, which reflects problems with both Estonian integration policy and the way we analyze it, is that the current framework leaves little room for new migrants. The integration clusters reflect a *status quo* of more than thirty years, in which Russophones are still perceived as a migrant population even after multiple generations since Estonia regained its independence. In their current form, the clusters render long-term residents as migrants and render new migrants almost invisible, since the profile descriptions make it difficult to place recent arrivals into the same categories as long-term residents.

Concerning other metrics, despite the nuanced insights into intra-group variations, such as critical citizens among non-Estonians or skeptical nationalists among Estonians, the method risks bias and stereotyping. When reading descriptions such as "unintegrated," "negativistic," or "critical nationalist," one may easily divide the clusters into "good" and "bad" ones. For minorities, this occurs by conflating financial struggles with integration failure and ignoring structural barriers and regional disparities. For the majority, it potentially stigmatizes the "deprived" clusters of Estonians by associating them exclusively with being poor and conservative. The cluster analysis translates their struggles into an assumed barrier to integration, while in reality, this is not always necessarily the case.

For example, arguably the biggest integration-related scandal of recent years ended with the President blocking an unlawful bill from the Riigikogu aimed at closing the Orthodox

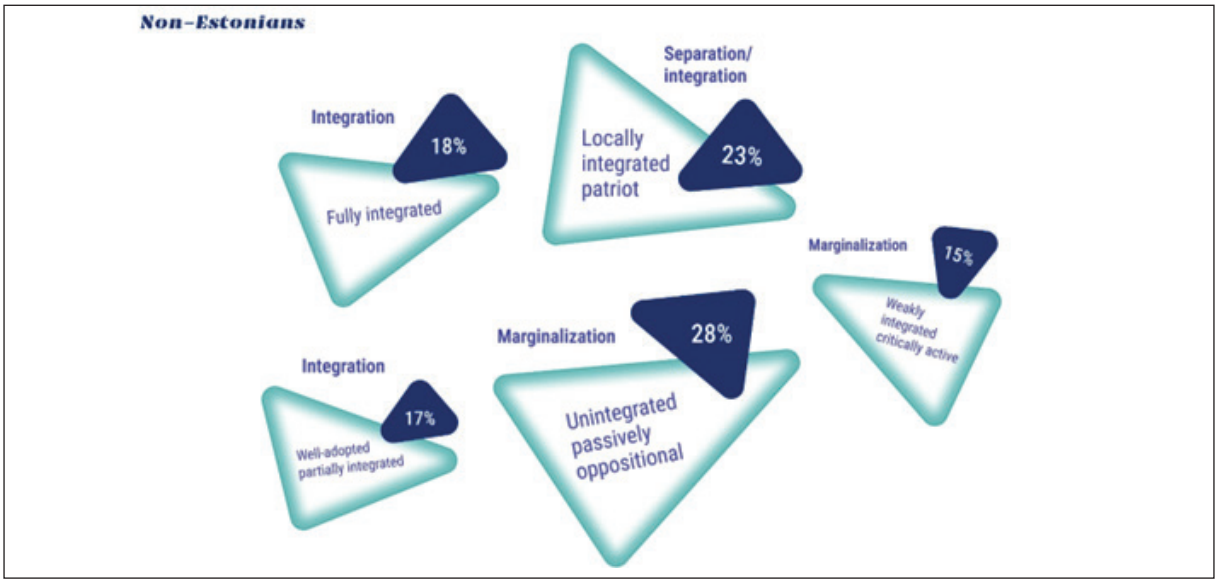
Church, a move connected to the liberal political parties that formed the government coalition (Wright & Kook, 2025). This particular example also illustrates a further implication: the use of distrust toward the government as a measure of failed integration should be treated with certain reservations, because government decisions may not necessarily affect all demographics equally, and a policy that causes resentment in one group may not provoke it in the other.

The final aspect that could be improved by adding the relational component to the cluster analysis is dynamics. Lauristin described the variability in identities in much finer detail than the scope and scale of this article would have allowed. She also framed the discussion of relational hierarchy in terms that allowed us to go further and easily build on these ideas. However, the integration clusters by themselves are static and offer only a snapshot of the integration process in 2023. Moreover, they present a somewhat more idealistic picture than the situation really is.

To illustrate this, let us map the EIM 2023 clusters onto Berry's acculturation strategies (Berry, 1997; 2017). Berry's integration hypothesis remains one of the most widely validated frameworks for distinguishing how individuals and groups balance heritage-culture maintenance with host-society engagement. Let us therefore try to map out what kind of acculturation strategies each cluster is undertaking, based on its description. Also, as we have data on the attitudes of the majority as well, it is possible to infer which type of acculturation strategy they support.

For non-Estonians, Cluster A (Fully Integrated, 18%) aligns with integration: high Estonian language use, identity, trust, and participation coexist with multilingualism and retained features of minority identity. At the same time, Cluster B (Locally Integrated Patriots, 23%), even in its name, reflects separation much more than integration: strong local loyalty and identity are maintained within geographically and linguistically segregated spaces (especially Ida-Virumaa), with minimal broader societal engagement. Cluster C (Well-Adapted Partially Integrated, 17%) also approximates integration through pragmatic multilingualism and life satisfaction, but falls short in trust and political participation. Clusters D and E fall into marginalization: D (Weakly Integrated Critically Active, 15%) via attempted self-driven assimilation that has not brought the desired reception from the wider society (high citizenship and language proficiency yet low trust and elevated discrimination sensitivity), and E (Unintegrated Passively Oppositional, 28%) via complete detachment from both host institutions and active heritage networks.

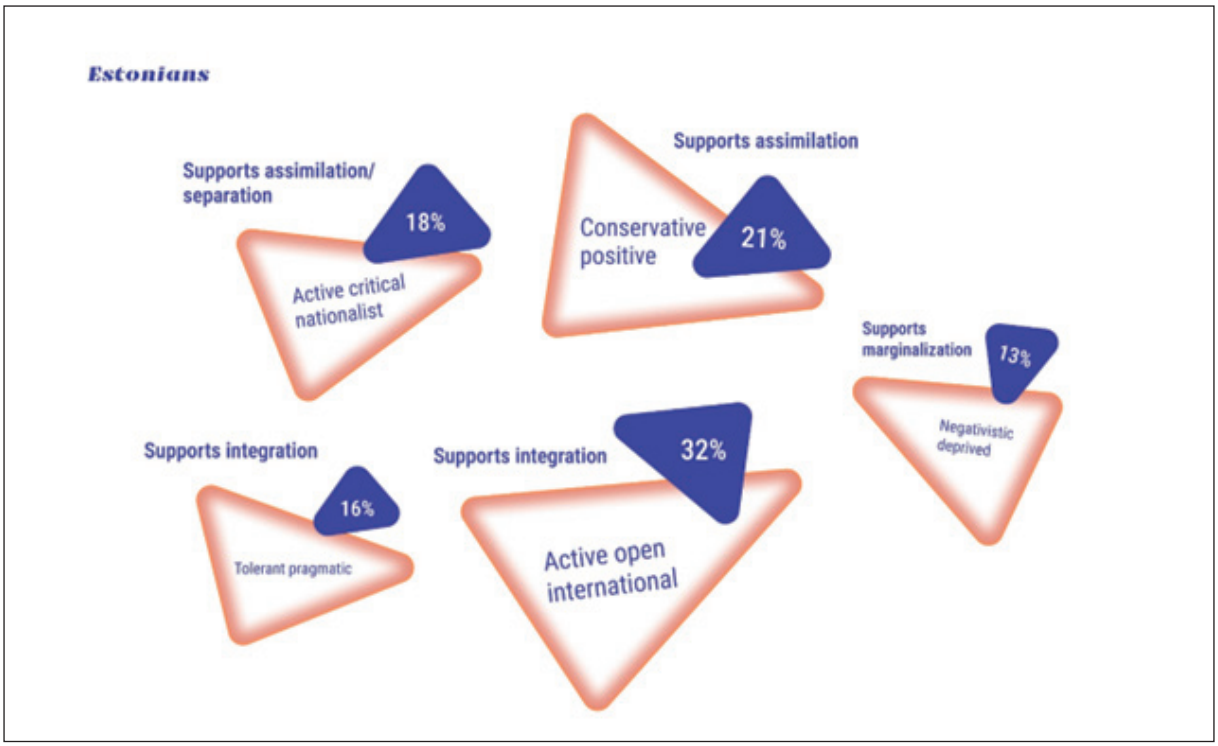
Figure 2. Lauristin's Integration clusters alongside Berry's acculturation strategies for Non-Estonians



Source: adapted from Kultuuriministeerium (2023); Berry (1997; 2019)

For Estonians, the mapping shifts to their openness to inclusion: Clusters A (Active Open International, 32%) and B (Tolerant Pragmatic, 16%) support minority integration. Cluster C (Conservative Positive, 21%), however, rather tilts toward assimilation rather than integration by prioritizing language and traditions over other parameters. Cluster D (Active Critical Nationalist, 18%) combines assimilationist demands with separationism, expressed in the distrust toward government policies and its declared course on bringing the majority and minorities together. The final cluster E (Negativistic Deprived, 13%), itself marginalized, declares support for the marginalization of other groups. On the whole, this mapping of acculturation strategies for both groups demonstrates that what is presented as integration often collapses into assimilation (for the ethnic nationalism-inspired Estonians) or separation (among non-Estonians who remain locally anchored and among Estonians who resist cultural mixing) when rendered from static into dynamic. In other words, integration clusters in their current form need considerable revision should they move further in the relational direction.

Figure 3. Lauristin's Integration clusters alongside Berry's acculturation strategies for Estonians



Source: adapted from Kultuuriministeerium (2023); Berry (1997; 2019)

4. Loyalty, Trust, and Security Attitudes

4.1 Loyalty vs Demand for Equal Rights

Loyalty is a difficult concept to capture, but it boils down to supporting something regardless of whether the terms for that support are favorable or not. For a country, the question of loyalty becomes especially pressing when it feels threatened and concerned about its security, which is definitely true for Estonia. Since the day Russia attacked Ukraine, Estonian authorities, driven

by security concerns, have taken multiple actions to increase unity and security. Within the field of integration, Vetik (2019) described that most relations and subject positions are formed between the two imperatives.

From the majority's side, this takes the form of a demand for loyalty expressed as national unity based on ethnic roots and the need to protect the Estonian language and culture. The minority, however, has contraposed to this the demand for equal rights and integration on equal footing, instead of forced assimilation. That demand, as Vetik (2019, 410–413) noted, used to give the minority the "moral high ground". Therefore, I argue that there are two major tendencies in the field of integration. First, the war in Ukraine did not create a new, unique situation within the Estonian integration field, but rather amplified all the cleavages and added a security context. Second, even though the loyalty-versus-equal-rights imperatives have remained, support for military aid to Ukraine has become a way for the majority to change the balance and secure the new moral high ground.

Lauristin, in EIM 2023, introduced a new measure, the national defense index, which she created to reflect the increased security threats and societal divisions caused in Estonian society by the war in Ukraine. The index included the following indicators: defense readiness, trust in the Estonian state and law enforcement structures, support for military aid to Ukraine, and support for NATO. The resulting combined index divided the respondents into five almost equal categories (18–22%), ranging from 1 to 5 points on readiness to defend Estonia (Kultuuriministeerium, 2023, 26). However, for the purposes of this article and for the sake of comparison between EIM 2023 and the Monitoring Survey (Riigikantselei, 2025a), I will use the actual indicators rather than the aggregated index number. Let us compare these indicators and draw insights into how the national defense index is formed and what its components allow us to see.

4.2 Trust in State Institutions

Table 1. Side by side comparison of the trust towards Estonian institutions in the Estonian Integration Monitoring (2023) and the Monitoring Survey (AAS, Dec. 2025)

Institution	Group	EIM 2023 (%)	AAS, Dec. 2025 (%)	Change (%)
Riigikogu (Legislature)	Estonians	42	41	-1
	Non-Estonians	36	25	-11
Government	Estonians	52	43	-9
	Non-Estonians	37	24	-13
President	Estonians	80	76	-4
	Non-Estonians	50	58	8
Police	Estonians	88	86	-2
	Non-Estonians	84	76	-8
Defense Forces	Estonians	89	87	-2
	Non-Estonians	64	54	-10

Source: Kultuuriministeerium (2023); Riigikantselei (2025a)

Trust in Estonian state institutions is sharply divided along ethnic lines, and these divisions have either considerably deteriorated since EIM 2023 or reflect differences in sampling methods. Trust in both the Riigikogu and the Government stands at low levels, with only 33% of Estonians and 22% of non-Estonians expressing confidence in 2025. This marks a steep decline for the Government compared with EIM 2023 (from 52% to 33% among Estonians and from 37% to 22% among non-Estonians), while the drop is more moderate for the Riigikogu (from 42% to 33% and from 36% to 22%, respectively). Estonians continue to show relatively high confidence in the President (76%, down slightly from 80% in EIM 2023) and especially in the Police (84%, down from 88% in EIM 2023). Non-Estonians trust these institutions more than the Government and the Riigikogu, with the figures standing at 44% for the President (down from 50%) and 76% for the Police (down from 84%). The Estonian Defense Forces remain highly trusted by Estonians, showing a confident 85% result in the latest September survey (a modest decline from 89% in EIM 2023). Non-Estonians (primarily Russophones) show markedly lower confidence, with only 49% expressing trust in the army (down substantially from 64% in EIM 2023), revealing a persistent and widening 36-percentage-point ethnic gap (Riigikantselei, 2025a, 7–8; Lauristin, 2023, raw crosstables).

When it comes to drawing conclusions, trust is highly important in evaluating integration. However, a researcher must be careful not to overemphasize its components. As integration ultimately happens between people, interpersonal or inter-community trust – which forms the foundation of social capital – should take priority over trust in the state. Another issue is that trust in politicians, particularly in the government and legislature (currently low across all ethnic groups), is often temporary and heavily dependent on the prevailing political situation. In this particular case, the decline likely signals dissatisfaction with recent reforms as much as deepening polarization between the coalition and the opposition. Nonetheless, the ethnic gap remains pronounced, showing that the pre-existing loyalty versus equal-rights imperatives are still at work. The high trust in the police shows that non-Estonians are confident that law enforcement is ready to defend their rights. In contrast, the more than 30-percentage-point difference in trust in the Defense Forces – with Estonians rallying around the army while Russophones show markedly lower confidence – may signify not a lack of loyalty (which may be implicitly inferred from the combined national defense index in EIM 2023), but rather a lack of belief that the army would be motivated or capable of defending the border regions should a war with Russia ever take place. Therefore, I argue that integration research and policy could treat non-Estonians' lower trust in the Defense Forces not as a loyalty deficit but as a signal for targeted confidence-building measures.

4.3 Sense of Belonging and Societal Inclusion

Table 2. Percentage of positive reactions to the "I feel I belong to Estonian society" in the Estonian Integration Monitoring (2023) and the Monitoring Survey (AAS, Dec. 2025)

Metric	Group	EIM 2023 (%)	AAS, Dec. 2025 (%)	Change (%)
Belonging	Estonians	89	85	-4
	Non-Estonians	76	59	-17

Source: Kultuuriministeerium (2023); Riigikantselei (2025a)

Even relying solely on the results of the Monitoring Survey collected between 2022 and 2025, the sense of belonging to Estonian society has declined noticeably for everyone over the three-year period. Among all respondents, the proportion who fully or rather agree with the statement "I feel I belong to Estonian society" dropped from 86% in March 2022 to 76% in December 2025. The three-year (2022–2025) decline is especially pronounced among non-Estonians, falling from 73% to 59%, while Estonians experienced a milder decrease from around 93% to 85% (Riigikantselei, 2025b, 48; Riigikantselei, 2025a, 30).

The comparison of the Monitoring Survey of December 2025 with the Integration Monitoring 2023 only reinforced these results, showing a 4% drop in the feeling of belonging for Estonians (from 89% in EIM 2023 to 85% in AAS, Dec. 2025) and a considerable 17% fall for non-Estonians (from 76% in EIM 2023 to 59% in AAS, Dec. 2025). This follows an even deeper drop observed in September 2025 (to 81% for Estonians and 45% for non-Estonians), showing that the sense of belonging may occasionally fall even lower than the latest Monitoring Survey reports (Riigikantselei, 2025b, 48; Riigikantselei, 2025a, 29; Kultuuriministeerium, 2023, raw crosstables). That allows us to conclude that the erosion of belonging is not merely a short-term fluctuation but a definite trend captured by the Monitoring Survey series and EIM alike, which has been accelerating since 2022 and affecting the minority the most.

Such problems with belonging are usually the result of exclusionary policies, securitization, structural barriers, or polarized rhetoric dominating public discourse. In turn, it is safe to infer that belonging affects all other relevant results, from public trust to narrower integration parameters. This decline in belonging among both Estonians and non-Estonians demonstrates that securitizing integration has been counterproductive. Rather than building greater unity under the post-2022 security pressures, pressing harder for loyalty in policy and research has actively eroded the sense of inclusion across ethnic lines, hitting the minority hardest but also pulling the majority down.

4.4 Support for NATO and military aid to Ukraine

Table 3. Trust in NATO and Support for NATO membership in the Estonian Integration Monitoring (2023) and the Monitoring Survey (AAS, Dec. 2025)

Metric	Group	EIM 2023 (%) Trust	AAS, Dec. 2025 (%) Trust	Change (%)	AAS, Dec. 2025 (%) Support for membership in NATO
Support for NATO	Estonians	81	74	-13	91
	Non-Estonians	41	30	-11	48

Source: Kultuuriministeerium (2023); Riigikantselei (2025a)

In contrast to the Monitoring Survey, measuring both trust in NATO and support for membership in the organization, EIM measures only trust, making it the sole basis for comparison. For both groups, Estonians and non-Estonians alike, trust in NATO has declined: it dropped 13 percentage points among Estonians (from 81% in EIM 2023 to 74% in AAS, Dec. 2025) and 11

percentage points among non-Estonians (falling to a record-low 30% in AAS, Dec. 2025 from 41% in EIM 2023). At the same time, longitudinal analysis of the Monitoring Survey alone shows that support for NATO membership has remained consistently high among Estonians throughout 2022–2025, fluctuating between 86% and 94 % (with the December 2025 figure at 91 %). In stark contrast, support among non-Estonians started at only 26 % in January 2022, rose to a peak of 53 %% in September 2023, and has since stabilized around 48–50% in 2025 (Riigikantselei, 2025a, 35).

Such a difference between trust in NATO and support for its membership may be situational and influenced by external factors, specifically tensions between the US and its allies over Greenland. On the one hand, the profound 43-percentage-point gap in support for NATO membership between Estonians and non-Estonians is an alarming sign of fundamental perception differences between the communities. On the other hand, earlier iterations of the Monitoring Survey reveal that since 2022, support for NATO membership among non-Estonians has almost doubled, rising from 26% to 48–50% in 2025 (Riigikantselei, 2025b, 35; Riigikantselei, 2025a, 13; Lauristin, 2023, raw crosstables). It also shows that security can serve as a point of convergence of interests, but under the conditions of an intensive loyalty-demanding model, its potential remains limited.

Table 4. Support for continued military aid to Ukraine in the Estonian Integration Monitoring (2023) and the Monitoring Survey (AAS, Sep. 2025)

Metric	Group	EIM 2023 (%)	AAS Sept, 2025 (%)	Change (%)
Support for continued military aid to Ukraine	Estonians	83	78	-5
	Non-Estonians	45	24	-21

Source: Kultuuriministeerium (2023); Riigikantselei (2025b)

Support for continued military aid to Ukraine reveals one of the deepest and most persistent divides in Estonian society. It is also important to note that, because the December version of the Monitoring Survey does not present an ethnic breakdown for these figures, the table above relies on the September 2025 survey.

Among Estonians, backing has remained consistently high and stable, never dropping below 68% across all survey waves since 2022 and reaching 78% in September 2025. Still, this represents a 5-percentage-point drop compared with 83% in EIM 2023. In sharp contrast, non-Estonians show markedly lower support that has hovered around 17–24% from mid-2023 through September 2025. Accordingly, opposition to military aid is high among non-Estonians (often 40–50%), while for Estonians it rarely rises above 15–20%. It is important to note that this enduring gap reported by the Monitoring Survey demonstrates a serious cleavage in security attitudes between the two communities.

Nonetheless, the EIM 2023 data painted a more optimistic picture, with 83% of Estonians and 45% of non-Estonians supporting continued military aid to Ukraine; figures for non-Estonians that were twice as high as those recorded in the Monitoring Survey of September 2025 (Riigikantselei, 2025b, 34–35; Lauristin, 2023, raw crosstables). This sharp ethnic divide on military aid to Ukraine has allowed the majority to establish a new test for loyalty in the current security environment. By framing strong support for Ukraine as a moral and security

imperative, Estonians have effectively claimed the moral high ground that the minority once held through demands for equal rights. As a result, the loyalty imperative has gained significant strength, shifting the balance in the integration field in favor of the majority's position, altering the previously achieved balance in interethnic relations, and sidelining the minority's quest for equal rights.

5 Conclusion: Unreflected Shift Toward the Loyalty Paradigm and Its Limits

Due to the horrors of the war in Ukraine, Russia's proximity, and the existential threat it poses, the demand for loyalty, already embedded in Estonia's integration field, has gained fresh moral justification and political urgency. The current mainstream framing in both research and public policy rests on the necessity of supporting Ukraine in its just struggle against subjugation and tyranny. While this imperative itself is not in question, closer scrutiny of the loyalty discourse reveals important analytical shortcomings.

First, the loyalty paradigm places the primary burden of proof on the Russophone minority, creating a vicious circle in which members of that community must constantly signal their allegiance in the public sphere. Those who are already the most marginalized are thereby further ostracized rather than persuaded or assisted to integrate. Integration researchers often treat this paradigm as self-evident and fail to reflect on its implications. Someone unfamiliar with the Estonian context might therefore be surprised to read Lauristin's conclusion on the national defense index in EIM 2023 that the responses of non-Estonians show no clear link between citizenship and loyalty, and that readiness to defend Estonia does not depend on whether a person holds Estonian, Russian, or undetermined citizenship (Kultuuriministeerium, 2023, 26). The unspoken question is: before whom, exactly, is Professor Lauristin defending the loyalty of Russian passport holders?

Second, this unreflected shift toward the loyalty paradigm has encouraged researchers and policymakers to define integration, categories, and even people in largely negative terms. It is easier to identify what loyalty is not (lack of support for military aid to Ukraine, lower trust in the Defense Forces, or a declining sense of belonging) than to articulate what positive loyalty to the Estonian state actually entails. The same tendency appears in the integration cluster analysis, where clusters are implicitly divided into "good" and "bad" on the basis of political attitudes. Conservative nationalists facing economic hardship, for example, are framed as obstacles to integration, even though the empirical reality is more complex. As the mapping onto Berry's acculturation strategies demonstrated, many of the behaviors observed in both majority and minority clusters collapse into assimilation (on the Estonian side) or separation and marginalization (on the non-Estonian side) rather than genuine integration.

Third, the loyalty approach is accompanied by a striking absence of any practical account of how loyalty is supposed to be cultivated. Integration research simultaneously renders Russophones both hyper-visible (as the primary object of loyalty tests) and invisible (by lumping long-term, multi-generational residents together with recent migrants under the single label "non-Estonians"). The empirical material examined here – the sharp drop in the sense of belonging among non-Estonians, the persistent ethnic gaps in trust in the Defense Forces, and the deep divide in support for military aid to Ukraine – shows that loyalty expectations are being directed precisely at those Russian-speaking residents who have lived in Estonia for generations.

Yet these residents occupy exactly the same position in the imaginary relational hierarchy as recent migrants. This raises several questions: should all newcomers be subjected to identical loyalty demands? At what point does loyalty become a realistic expectation, for instance, for a Brazilian arriving to Estonia? Are EU migrants granted long-term residence also expected to bear arms and defend Estonia? And should integration programs designed to facilitate the return of Estonian emigrants from abroad be re-evaluated through the lens of loyalty?

The relational promise made at the beginning of this article, i.e., that integration needs to be studied as a contested negotiation of social boundaries rather than as a one-way adaptation of minorities, finds clear empirical grounding here. Equalizing assimilation and integration blinds us to how these processes actually unfold in practice. The same goes for ignoring the asymmetrical power relations between majority and minority.

Moreover, the mixing of integration, assimilation, marginalization, and separation can occur even when a researcher conducts normatively proper work that genuinely treats integration as a two-sided process, as Lauristin did in EIM 2023. A scholar need not be thinking for the state when producing research commissioned by the government, yet the aims and logic of the nation-state still shape, and are shaped by, its key assumptions. In Estonia, and I suspect in many other contexts, this has produced a clear shift toward the loyalty paradigm. The relational comparative analysis developed in this article, however, allows us to recognize that the shift has taken place instead of following it blindly. One does not necessarily have to support the resulting change in the moral high ground of the Estonian integration field, but examining it through a relational lens certainly helps us see it.

References

Alba, R. (2020). *The Great Demographic Illusion. Majority, Minority, and the Expanding American Mainstream*. Princeton University Press. https://doi.org/10.1515/9780691202112

Alba, R. (2024). Culture's role in assimilation and integration: the expansion and growing diversity of U.S. popular culture. *Journal of Ethnic and Migration Studies,* 50(1), 27–46. https://doi.org/10.1080/1369183X.2023.2213046

Alba, R., & Nee, V. (2003). *Remaking the American Mainstream. Assimilation and Contemporary Immigration*. Harvard University Press.

Berry, J. W. (1970). Marginality, Stress and Ethnic Identification in an Acculturated Aboriginal Community. *Journal of Cross-Cultural Psychology,* 1(3), 239–252. https://doi.org/10.1177/135910457000100303

Berry, J. W. (1997). Immigration, Acculturation, and Adaptation. *Applied Psychology,* 46(1), 5–34. https://doi.org/10.1111/j.1464-0597.1997.tb01087.x

Berry, J. W. (2019). *Acculturation*. Cambridge University Press. https://doi.org/10.1017/9781108589666

Emirbayer, M. (1997). Manifesto for a Relational Sociology. *American Journal of Sociology,* 103(2), 281–317. https://doi.org/10.1086/231209

Favell, A. (2019). Integration: Twelve propositions after Schinkel. *Comparative Migration Studies,* 7, Article 21. https://doi.org/10.1186/s40878-019-0125-7

Favell, A. (2022). *The integration nation: Immigration and Colonial Power in Liberal Democracies*. Polity.

Grigoryev, D., Berry, J. W., Stogianni, M., Nguyen, A. M. D., Bender, M., & Benet-Martínez, V. (2023). The integration hypothesis: A critical evaluation informed by multilevel meta-analyses of three multinational datasets. *International Journal of Intercultural Relations,* 97, 101897. https://doi.org/10.1016/j.ijintrel.2023.101897

Klarenbeek, L. M. (2019). Reconceptualising 'integration as a two-way process'. *Migration Studies,* 9(3), 902–921. https://doi.org/10.1093/migration/mnz033

Klarenbeek, L. M. (2024). Relational integration: from integrating migrants to integrating social relations. *Journal of Ethnic and Migration Studies,* 50(1), 233–256. https://doi.org/10.1080/1369183X.2023.2259038

Kultuuriministeerium [Ministry of Culture]. (2011). *Eesti ühiskonna lõimumismonitooring 2011 [Estonian society integration monitoring 2011].* Online: https://www.kul.ee/sites/default/files/documents/2020-10/integratsiooni_monitooring_2011.pdf

Kultuuriministeerium [Ministry of Culture]. (2015). *Eesti ühiskonna lõimumismonitooring 2015 [Estonian society integration monitoring 2015].* Online: https://www.ibs.ee/wp-content/uploads/Integration-Monitoring-Estonia-2015.pdf

Kultuuriministeerium [Ministry of Culture]. (2023). *Eesti ühiskonna lõimumismonitooring 2023 [Estonian society integration monitoring 2023].* Online: https://www.kul.ee/sites/default/files/documents/2024-04/EIM%202023%20aruanne.pdf

Lauristin, M. (2023). Raw crosstables from the Estonian Integration Monitoring 2023 [unpublished raw data]. Personal communication.

Polynin, I. (2024). Relational integration/assimilation? A critical dialog with postcolonial and mainstream perspectives. *Frontiers in Political Science,* 6, Article 1493637. https://doi.org/10.3389/fpos.2024.1493637

Riigikantselei [Republic of Estonia Government Office]. (2025a). *Avaliku arvamuse seireuuringu raport. 25. uuringulaine. 15.-22. detsember 2025 [Public Opinion Monitoring Survey. 25th wave. 15-22. December 2025]*. Online: https://riigikantselei.ee/sites/default/files/documents/2026-01/2025%2012%20AA%2025%20seire%20raport%20-%20avaldamiseks.pdf

Riigikantselei [Republic of Estonia Government Office]. (2025b). *Avaliku arvamuse seireuuring. 24. serie. 17.-21. sptember 2025 [Public Opinion Monitoring Survey. 24th wave. 17-21. September 2025].* Online: https://riigikantselei.ee/sites/default/files/documents/2026-03/2025%2009%20AA%2024%20seire%20raport%20-%20avaldamiseks.pdf

Schinkel, W. (2018). Against 'immigrant integration': for an end to neocolonial knowledge production. *Comparative Migration Studies,* 6, Article 31. https://doi.org/10.1186/s40878-018-0095-1

Statham, P., & Foner, N. (2024). Assimilation and integration in the twenty-first century: where have we been and where are we going?. *Journal of Ethnic and Migration Studies,* 50(1), 4–26. https://doi.org/10.1080/1369183X.2023.2293537

Tilly, C. (2005). *Identities, Boundaries and Social Ties*. Paradigm Publishers. https://doi.org/10.4324/9781315634050

Vetik, R. (2019). National Identity as Interethnic (De)mobilization: A Relational Approach. *Ethnopolitics, 18*(4), 406–422. https://doi.org/10.1080/17449057.2019.1613065

Wright, H., & Kook, U. (Eds.) (2025, July 3). *President sends church influence law back to Riigikogu for a second time.* ERR News. Online: https://news.err.ee/1609736706/president-sends-church-influence-law-back-to-riigikogu-for-a-second-time

Institutiones Administrationis – Journal of Administrative Sciences Vol. 6 (2026) No. 1, 105–119
DOI: 10.54201/iajas.204

Power and Narratives in the Era of Behavioral Government

Manuel C. Ortiz de Landázuri*, **Alejandro Cid Moreno****

* Associate Professor, Department of Philosophy, University of Navarra, Pamplona, Spain. E-mail: mcruz@unav.es
** Research Assistant, Marketing and Business Ethics, IESE University of Navarra, Barcelona, Spain. E-mail: icidmoreno@alumni.unav.es

Abstract

Behavioral government employs tools like nudges to steer behavior toward certain practices, raising ethical concerns if the underlying narratives are not transparent or inclusive. In this context, it has been debated whether these types of policies are a source of manipulation or rather are practices inherent to a libertarian paternalism. This article attempts to answer this question from an analysis of power. In contrast to theorizations of power in terms of force or pure domination, it is possible to establish a theoretical framework that understands power in an inclusive manner, as a capacity for order, and from there establish a rational framework for evaluating the narratives implicit in this type of policies. We propose that, within a framework of dialogue and shared values, power need not be coercive to be effective. The discussion concludes that narratives, when aligned with the common good and grounded in clear reasoning, legitimize behavioral government as an ethical and collaborative approach to addressing contemporary social challenges.

Keywords

choice architecture, libertarian paternalism, manipulation, nudges

1 Introduction

In recent years, there has been a growing interest in behavioral aspects that affect not only economics but also politics and government. There are already many institutions within governments that deal with behavioral policies, such as the Behavioural Insights Team (BIT) in the UK, the Behavioural Economics Team of the Australian Government (BETA) in Australia, the Office of Behavior and Public Policy (OCPP) in Spain, or the behavioral policy unit of the World Bank. And, in line with these initiatives, by "behavioral government" we are referring to decision-making based on unconscious and not entirely rational tendencies in terms of efficiency. It is for this reason that some authors have emphasized in recent decades the role played by unconscious aspects in decision-making (Ariely, 2008) and public opinion (Sunstein, 2019).

Narratives can be defined as structured stories and representations that give shape to a series of events, ideas, or experiences in a meaningful way. They serve as a framework to organize information and communicate meaning. In this article, we want to show the importance of narratives in the exercise of power in the light of what, in recent years, has been called "behavioral

 Published online: 12 June 2026

government". To do so, we first give some examples of narratives present in policies based on behavioral government. In this way, we show that behind these measures always underlie a narrative that serves for the constitution of power in a double dynamic: on the one hand, behavioral government is based on a previously admitted narrative, from which information is processed. On the other hand, behavioral governance generates a new narrative through new policies (reframing). The question arises, then, whether behavioral governance decisions hide a manipulative intention or only top-down guidance.

In this article, we address this question through the analysis of power and try to delimit what we mean by manipulation. Power has often been equated with the use of violence or domination, but from other perspectives, it is possible to understand power as a source of order, and thus establish parameters that ensure shared communicative processes and narratives based on good reasons. The risk of behavioral governance lies in the elaboration of narratives that appeal only to unconscious emotions and ignore other important aspects, such as coherence and shared values.

A particularly controversial aspect in this respect is whether behind the practices of behavioral governance lies a form of unconscious manipulation and domination. The central tension in behavioral governance lies in its dual nature: it can simultaneously serve as a tool for the collective good and as a mechanism for manipulation. Moving beyond the binary debate of whether 'nudging' is inherently manipulative or paternalistic, the ethical distinction can be found in the role of narratives as a cognitive and political bridge. Employing a Dual Process Theory framework (Kahneman, 2011), it becomes evident that while choice architecture primarily targets System 1 (intuitive processing), the presence of a transparent and shared narrative allows for System 2 (deliberative processing) engagement. This prevents the exercise of power from becoming a 'clandestine' influence and transforms it into a legitimate 'power-to' – a collective capacity where citizens can consciously endorse policy goals.

Thus, narratives are not mere rhetorical additions but the very condition that ensures the agency of the subject within the era of behavioral government. In this article, we explore the role of the narrative in establishing a behavioral framework that affects the decision-making of both rulers and ruled in order to discuss the basis of a soft "libertarian paternalism". To this end, we first of all show, through some examples, in what behavioral governance consists, and reflect on the extent to which it is supported by certain narratives. Secondly, we analyze what constitutes power and what role narratives play in its exercise. Finally, we establish a theoretical proposal of the exercise of power through narratives in relation to the mechanisms implicit in behavioral governance. In the light of the narrative paradigm, it can be observed that power is developed through narrative mediations in which both the rulers and the ruled are involved.

And we are considering power in behavioral governance, because it is the most subtle and efficient form of soft power. This type of power is exercised through implicit narratives and could be seen as a form of manipulation. Literature on power has often emphasized coercive power ("power over" others), overlooking power as the capacity to achieve certain ends, or "power to" (Battegazzorre, 2017). Power, thus understood, is an intrinsically negative concept, signifying that "A affects B in a manner contrary to B's interests" (Lukes, 1974, 27). Structural power, from this perspective, manifests as a set of predominant values, beliefs, rituals, and institutional procedures that systematically benefit certain individuals and groups at the expense of others (Bachrach & Baratz, 1970, 43).

We find this equation of power and domination over others in many authors, sometimes understanding that domination is an act of getting someone to do something he or she would not want to do (Dahl, 1957), at other times understanding that, in the act of domination, there

is a force that leads to acting against one's own will (Wartenberg, 1990, 85). For example, according to Pettit (1996, 578), "one agent dominates another if and only if he or she has a certain power over that other: in particular, the power to interfere in the affairs of the other and to inflict a certain damage". As we shall see at the end of the article, this view of domination as the imposition of harm on another is the one that has permeated much of the discussion on the manipulation that behavioral governance can entail.

However, these analyses of power sometimes overlook that power is often exercised as follows: "A gets B to do something that is good for him and for the community." Power as dominance has usually been understood as A being able to impose on B something that he or she would not really want to do, but this presupposes that we are individuals who have interests prior to any form of political community. Many times, both A and B participate in certain shared narratives, and A's power lies in B's participation in a good, common to both (Parsons, 1967, 308). In this sense, power-to (power as capacity) is more important than power over others, because power over others makes sense only if it helps to achieve some ends. From Arendt's (1998, 200) perspective of power, it can be understood that not only individuals but also communities are subjects of power to the extent that there are shared narratives. Thus, behavioral governance need not be seen as manipulation, as long as clear intercommunicative processes are established. As such, Fisher's (1989) narrative paradigm may thus prove useful in establishing criteria for a good story: background values, internal coherence, and being based on good reasons.

In this article, we will first present a vision of power that goes beyond the framework of domination in order to understand how narratives allow for the construction of non-repressive and non-manipulative forms of power. Next, we will analyze specific characteristics of behavioral governance and nudges. Finally, we will examine how these types of practices can be understood through certain underlying narratives, and by what criteria they can be said to constitute manipulation or not.

2 Power and Intermediations

First, to comprehend behavioral governance, one must understand that it is a form of power and, as such, is constituted through certain narratives. However, power remains a diffuse and difficult concept to define. Here, we argue that power should be understood in general terms as *dynamis* or the capacity to accomplish things, and within human communities, as the capacity to order and shape actions toward an objective. In this sense, although power is linked to domination, the latter need not be repressive or violent; rather, it can be based on intermediations and shared narratives that integrate different agents around a common project, rooted in certain accepted values.

Literature on power has typically focused on its coercive aspects or understood domination in terms of violence and imposition (Lukes, 1974, 27; Bachrach & Baratz, 1970, 43; Wartenberg, 1990, 85; Pettit, 1996, 578). In this respect, even a significant figure like Foucault (1980), who clearly established the relationship between power and discourse, often frames power in terms of domination.

Due to Foucault's considerable influence on the understanding of power, we will briefly analyze the link he establishes between power and discourse before highlighting some inadequacies of his coercive model. Following this, and drawing on Han's (2005) critique of the purely coercive model of power, we advocate for a vision aligned with the perspectives

of Parsons and Arendt: power as the capacity to integrate forces around a shared objective or narrative. Power is the capacity to achieve some ends, and it grows in as much there is a shared narrative. Thus, it will be shown that behavioral government can be understood as a practice that is not necessarily manipulative.

As for Foucault, he focused a good part of his studies on the relationship between truth and power, and affirmed that power should be analyzed in its expressions, and one should not try to define it, since one would then be trapped in the nets of the modern science that he is criticizing (Foucault, 1997, 13–14). Foucault holds that power and truth are two linked concepts: power builds truth; truth legitimizes power. In this binary structure, power-knowledge truth is nothing more than a construct. Through his analysis of the constitution of sciences, he tries to show that truth is the result of a persuasive discourse that generates control. In the background, there is only power in terms of control and domination (Foucault, 1997, 15–30). And the role of reason in the constitution of power is to generate truth in terms of persuasion.

According to Foucault (1980, 109–116; 2014, 113), the modern state operates as a means of controlling the social body through norms and legal structures. The ruler wields power through violence and coercion to impose their will, with sovereignty shifting from control over territory to control over a population, akin to managing a natural body (Foucault, 2009, 23). This view of power has been very influential in the way we approach power relations, not only within society but also in international relations. However, although Foucault understands power as creative energy, the coercive model weighs heavily in his model.

In contrast to Foucault, we have in recent years the position of Byung-Chul Han (2005), who argues that real power has more to do with persuasion than with disciplinary methods of control and must be understood as the capacity to expand one's will to another's will. Han (2005) states that, although power has traditionally been interpreted through a coercive-negative model, greater power lies in influencing another's future rather than hindering it. Power does not consist solely in the capacity to use violence or impose punishments. Although that capacity may seem important to ensure order, it is equally true that respect for the law arises from the recognition of a juridical and social order that deserves to be upheld. Therefore, "the one who is only able to set his decision with the use of force or a negative sanction has little power" (Han, 2005, 25). Power configures networks and opens fields between people, in which different individual powers are organized toward a determined direction (Han, 2005). The essence of power, as Han observes, is not violence or domination, but the energy that helps to develop the political community. In this sense, there is no real power without communication and intermediations (Arendt, 1998).

On the other hand, as some scholars have argued (Porter, 2009; Sassen, 2006), it might be more accurate to say that power is not exercised by a single actor but is better understood through "assemblages" when analyzing governance. There are no independent and autonomous actors; rather, communities are governed through a network of institutions and practices in which actors move and play their roles. "Originally used in art, an assemblage is created out of disparate elements, each of which has meanings or purposes that might be quite unrelated to the other elements, but which together are brought into a new relationship with one another to create an arrangement with its own distinctive meaning or purpose" (Porter, 2009, p. 89). In this framework, there is no single sovereign power but a social epistemic network that generates forms of power through norms, practices, and institutions (Adler & Bernstein, 2005).

Governance takes place through a variety of interconnected networks. Rather than conceptualizing different institutions as sovereign powers, one should think in terms of networks of power involving states, non-state actors, and various intermediaries. These networks operate through systems of inclusion and exclusion procedures (Castells, 2011) and entrench their power

through communicative processes. For example, a standard defines the specific way in which a group of people is interconnected within a network. It acts as the shared norm or practice that enables network members to access one another (Grewal, 2008).

Power can be described as creative energy, or a "capacity to do", which is actualized through networks and intermediations, with narratives playing a critical role in this process. From the perspective of Fisher's (1989) narrative paradigm, humans primarily comprehend through stories rather than strictly logical discourses. These narratives incorporate imaginative elements and, more importantly, values and "good reasons" that resonate with individuals. Fisher identifies several key characteristics of the narrative paradigm, emphasizing that human communication is inherently narrative. He argues that, throughout history, rationality has often been reduced to logic, neglecting other forms of understanding, such as storytelling. Within stories, values are communicated through the narration of facts, and narrative rationality can be assessed based on coherence, fidelity, implicit values, and compelling reasons.

From this perspective, power is most effective when its narrative is well-constructed: when it reflects shared values and provides compelling, fact-based reasons, as narratives act as tools of unity and persuasion, enabling collective action. As Arendt (1998, 244) points out, power arises from shared narratives expressed through actions and words. Therefore, control or coercion without dialogue lacks the legitimacy of real power.

These narratives are shaped by various institutions and processes, reflecting broader social, cultural, and political contexts. In recent years, global narratives have gained prominence, influencing individual decision-making and policy development. Examples include ecological sustainability, gender equality, and bio-sanitary protection, which form part of global epistemic communities (Haas, 1992). These narratives establish codes and norms that structure power relations on a global scale, operating through networks rather than traditional hierarchies (Castells, 2011).

The thesis we defend here is that the narrative paradigm allows us to understand the exercise of power in the behavioral era not necessarily as a form of manipulation and despotic domination, but as the ordered orientation through shared narratives.

3 What is Behavioral Governance?

Traditional conceptions of governance often prioritize the explicit legislative and fiscal deliberations of political actors. However, contemporary regulatory practice suggests that individuals are increasingly governed not through formal prohibitions, but through subtle interventions in their decision-making environments. By configuring the "choice architecture", ranging from the strategic placement of urban infrastructure to the framing of information in public spaces, governments exert a form of influence that bypasses conscious deliberation, systematically steering behavioral outcomes without resorting to coercive legal frameworks.

Behavioral governance concerns the processes of information assimilation and decision making so that various actors and population segments act in a certain way (Gofen et al., 2021). These are public policies aimed at generating an architecture of choice (Thaler & Sunstein, 2008). This type of governance assumes rationality from above, insofar as it is assumed that agents do not possess complete information when making decisions and need a little nudge or framework for action to make correct choices.

As some scholars have shown (Battaglio et al., 2019; Moseley & Thomann, 2021; Simon, 1978), individuals often tend to make decisions that are not the best precisely because of deficiencies in

the process of assimilating and interpreting information. We often act not on the basis of reality itself, but because of an inadequate perception of reality (Rittel & Webber, 1973; Thomann et al., 2018). Behavioral governance, then, can be understood as the set of psychological processes aimed at affecting the behavior of individuals (see Galizzi, 2017; Grimmelikhuijsen et al., 2017). In this framework, nudges, understood as small nudges or *stimuli* that unconsciously guide the activity of individuals, have gained special interest (Thaler & Sunstein, 2008; Sunstein, 2014).

The policy narrative, as a discursive tool, not only shapes public perception of policies but also profoundly influences the behavior of citizens and policymakers. In this framework, the concepts of "nudging" and "choice architecture", analyzed by Lehner et al. (2015), and the conception of the state as a "market maker", proposed by De la Cruz Prego (2021), highlight the intersection between policy design, human behavior, and sustainability. Lehner et al. (2015) present nudging as a behavioral approach to overcome cognitive limitations and human biases, promoting socially desirable decisions without coercion. These strategies have been successfully applied in domains such as energy efficiency and food consumption, demonstrating that tools such as social norms or changes in default choice are highly effective in changing specific behaviors. However, their effectiveness varies by context, underscoring the importance of culturally and socially tailored design.

In this line of thought, De la Cruz Prego (2021) argues that the state must assume an active role as an "architect of markets" to achieve the Sustainable Development Goals (SDGs). This approach implies not only regulating existing markets but also redesigning them to promote environmental sustainability and social equity objectives. The integration of behavioral tools, such as those proposed by Lehner et al. (2015), complements traditional interventions by incorporating knowledge of human biases and behaviors in the design of public policies. Both approaches converge on the need for policy narratives that legitimize these interventions, since a policy narrative not only explains and justifies policies to the citizens but also influences how policymakers perceive their own rationale and decision-making. This emphasizes the importance of integrating behavioral tools with robust institutional design and narratives that reinforce public trust in the policy process.

In the next section, we are going to show some examples of policies based on behavioral governance to demonstrate two things: on the one hand, there is always a narrative underlying these policies to shape the architecture of choice; on the other hand, it is not only the governed who are affected by these narratives, but also those who design public policies.

4 Behavioral Governance and Public Policy

Political narratives impact the formulation and reception of public policies, not only through large-scale decision-making, but sometimes through mechanisms that generate different scenarios and guide the behavior of individuals in certain directions. Let us look at some examples, then.

The automatic enrollment in pension schemes in the UK serves as a quintessential case study. This resulted in a significant increase in participation from 61% to 88% in four years, adding more than 5 million new savers (Halpern & Sanders, 2016). This shift was based on the design of default policies that use inertia to encourage financially responsible decisions. The narrative behind this policy framed the lack of savings as a "failure" of individual behavior, justifying state intervention as a tool to correct this bias. When analyzed through the narrative-power framework, this policy does not merely function as a mechanical default setting; rather,

it relies on a latent narrative of "long-term security" and "intergenerational responsibility". While critics might view this as a bypass of individual will, this critique must account for the fact that non-enrolment equally constitutes a choice architecture, one that carries its own implicit narrative, namely, that saving for retirement is an individual responsibility requiring active initiative. The relevant distinction, therefore, is not between a "neutral" default and an "interventionist" one, but between narratives that operate transparently and those that remain unexamined.

Both enrolment and non-enrolment frames shape behavior; what differs is whether the underlying values are made visible and open to public scrutiny. The success of the nudge is contingent upon its alignment with a shared social narrative that legitimizes the state's role in securing future welfare. From a behavioral symmetry perspective, this narrative omits that the experts designing these policies may also be influenced by optimism or overconfidence biases. While the policy is effective in terms of outcomes, it perpetuates an institutional reliance on individual decisions, limiting the space for citizen autonomy.

Madrian and Shea (2001) have analyzed how automatic enrollment in 401(k) deferred compensation plans affects savings behavior in a large U.S. corporation. Prior to implementing this policy, employees had to actively enroll to participate, resulting in a participation rate of 37%. After the introduction of automatic enrollment, this figure increased to 86%, even with no change in benefits or plan terms. In addition, 61% of automatically enrolled employees maintained the default contribution (3%) and funding allocation (100% in the money market fund), evidencing strong behavioral inertia. These results highlight how defaults not only simplify decisions but also act as implicit suggestions perceived as recommendations. It is a clear example of the power of biases, such as change aversion and procrastination, in policy design (for an overview of such behavioral tools in public policy, see Chetty, 2015). However, it raises ethical questions about the use of these mechanisms, as they shape decisions by exploiting cognitive limitations rather than encouraging more informed participation. As we will see below, the question is whether or not the unconscious orientation of individuals toward certain decisions is manipulation.

Similarly, interventions designed to increase fiscal compliance through social proof are often interpreted solely as psychological triggers. In Hallsworth et al.'s (2017) field experiments with the UK tax authority, messages telling recipients that "nine out of ten people pay their tax on time", and, more strongly, that they were currently in the small minority who had not yet paid, raised payment rates by around five percentage points and accelerated several million pounds in tax revenue (Hallsworth et al., 2017). However, within the conceptual framework of this article, these interventions are understood as attempts to reinforce a "narrative of fairness". By highlighting peer behavior, the state is not merely manipulating a bias, but is instead invoking a collective narrative where tax compliance is a shared contribution to the common good. Under this interpretation, the efficacy of the nudge depends on whether the citizen perceives this narrative as a legitimate expression of "power-to" (collaborative capacity) rather than an arbitrary imposition.

Another clear example of behavioral policy can be found in the nutritional traffic light, which has been promoted by many governments (although not in a prescriptive manner) and is having a major impact on food consumption (Osman & Thornton, 2019).

However, the narrative underlying behavioral government policies is not always the right one. Sometimes policies have been implemented based on erroneous assumptions, as for example in the case of predicting problematic alcohol consumption (Miller et al., 2024). Contrary to what was thought, delay discounting (the decrease in the value of a reward as its receipt is delayed) is not as significant a factor as the expectation of receiving rewards under temporal uncertainty,

or, in other words, human beings usually act more out of fear of losing than out of expectation of gaining.

One controversial example can be found in the letters that governments send to women over 50 to prevent breast cancer through mammography (with a specific appointment already in the letter). Although the letter states that mammography reduces mortality by 20%, this is actually an exaggerated figure (Gigerenzer, 2014). What this measure aims to do is to increase the rate of mammography, but not to teach people. The narrative underpinning policies based on these findings emphasizes the need to "correct" impulsive behaviors through interventions designed to improve decision-making. However, from a behavioral symmetry perspective, these policies could amplify expert design biases, particularly if paternalistic approaches are prioritized over participatory approaches.

As discussed above, political narratives not only shape citizen perception but also the rationality of policymakers. These narratives, combined with cognitive biases such as confirmation or authority bias, generate a cycle in which narratives and policy decisions reinforce each other. This is seen in the examples mentioned above: the use of social norms in tax charters has increased tax compliance by 5% (Hallsworth et al., 2017), or policies such as automatic enrollment in 401(k) plans have leveraged inertia and change aversion to improve citizens' savings (Madrian & Shea, 2001). Behind these behavioral policies, one can read a narrative in which it is understood that the state must ensure that citizens contribute to the common good through taxes, energy savings, food health, and concern for the ecosystem. These narratives, where good and bad appear, are not generated directly by those who take measures, but are developed through communication networks in which many agents and processes intervene (Castells, 2011).

5 Behavioral Governance: Manipulation or the Need for Order?

Thaler and Sunstein (2003) have been the main proponents of nudging and behavioral government, coining the concept of "libertarian paternalism", a view that grants freedom to individuals while assigning the state a role in channeling individual wills toward the common good. This perspective arises from the recognition that individuals often lack sufficient information to make optimal decisions. Since all policies inevitably influence individual choices, the state must at least nudge individuals through small motivations to achieve outcomes beneficial to both themselves and society, such as contributing to the common good, improving health, or achieving energy savings. This approach seeks to guide decisions toward positive outcomes while respecting freedom of choice, acknowledging the biases inherent in both individuals and policy design.

In contrast, the Behavioral Symmetry Theory developed by Thomas (2019) critiques the assumption that policymakers are inherently rational and benevolent agents, unlike citizens, who are often portrayed as subject to cognitive biases. This theory challenges the presumption that experts are free from systematic errors, arguing that behavioral interventions amplify the risks of manipulation if policymakers' own biases are not acknowledged. This critique aligns with the concept of the policy narrative, which structures the perception of problems and solutions for both citizens and policymakers. As the Behavioural Insights Team report emphasizes, these narratives not only justify specific interventions but also shape how governments prioritize problems and evaluate their effectiveness (Hallsworth et al., 2018). In sum, Thomas's (2019) Behavioral Symmetry Theory posits that both policymakers and citizens are influenced by cognitive biases, thereby questioning the superior rationality often attributed to experts.

It is important to note, however, that Thomas's critique does not require policymakers to be perfectly rational in order for nudge policies to serve citizens' interests. The relevant question is not whether experts are free of cognitive bias, but whether their informed judgements improve outcomes relative to unassisted individual decision-making. Consulting domain experts, whether physicians, economists, or public health specialists, is a rational strategy precisely because expertise reduces, even if it cannot eliminate, systematic error. Moreover, democratic accountability mechanisms – parliamentary oversight, freedom of information, and independent policy evaluation – serve as institutional correctives when expert judgement fails. Acknowledging the cognitive limitations of policymakers strengthens, therefore, rather than undermines, the case for transparent and externally scrutinized behavioral governance.

Thus, we have come to a key question, namely, whether there is an objective behind these practices to guide while preserving individual freedom, or do they represent ideological manipulation? Libertarian paternalism, as proposed by Thaler and Sunstein (2003), acknowledges that both citizens and policymakers are subject to cognitive biases and conditioned by the architecture of decisions. It suggests designing policies that respect individual freedom while nudging people toward choices that benefit their personal and collective well-being. This framework accepts the inevitability of decision conditioning, advocating for public policies that guide decisions toward social welfare without unduly restricting autonomy. However, as White (2013) has argued, libertarian paternalism presupposes that individuals are irrational and must be guided, potentially overlooking alternative motivations that might drive their actions. This raises an important question: is behavioral governance an exercise in manipulation based on arbitrary narratives, or is it better understood as an exercise of structured power? To address this, we will examine the foundations of power and its relationship to narrative.

The thesis we will maintain here is that these practices are always the product of a narrative or help to construct a narrative, so they can never be neutral. However, the fact that there is a narrative does not mean that there is manipulation. A correct approach to power shows that having power is not the same as exercising violence. Social organization inherently requires mechanisms of authority and coordination. The challenge lies in ensuring that such authority does not devolve into manipulation or coercive dominance. This transition is prevented when power is embedded within a shared and transparent narrative, established through robust dialogic processes. While critics such as White argue that individual decision-making remains independent of unconsciously assimilated frameworks, it is argued here that agency always operates within a narrative perspective – one often internalized without explicit awareness. This cognitive reality applies equally to policy designers, who operate from within settled institutional narratives that precede their own agency.

To understand how narratives transform behavioral governance from mere manipulation into a legitimate exercise of power, it is essential to ground the analysis in Dual Process Theory (Kahneman, 2011). Standard behavioral interventions primarily target System 1 – the fast, instinctive, and emotional brain – by exploiting cognitive biases to bypass deliberation. Critics argue that this "by-passing" constitutes manipulation because it excludes the individual's agency. However, narratives function as a bridge to System 2 – the slower, more deliberative and logical mind. While a nudge may provide the initial impulse (System 1), a transparent and shared narrative provides the reasons and values that allow System 2 to endorse the action. By providing a narrative framework, the government moves beyond "clandestine" influence and offers a justification that the citizen can rationally evaluate and consciously adopt as part of a collective "power-to" (capacity), rather than a coercive "power-over".

At this point, we can see how behavioral governance does not necessarily have to be a form of manipulation. Although manipulation remains a somewhat diffuse concept (Wilkinson, 2013) due to the multiple methods through which it is carried out (Noogle, 2026), we could define manipulation as a behavior that seeks to direct someone's conduct toward the interest of the manipulator, concealing true intentions and nullifying the other person's capacity for free and informed choice. Whereas in persuasion both parties usually know the objective and can benefit, in manipulation, there is an instrumentalization of the person; they are turned into a means to an external end.

Contrary to those who believe that behavioral governance and, specifically, nudges are a form of manipulation (Grüne-Yanoff, 2012), it can be argued that they are instead a type of situational persuasion when they meet certain requirements. Specifically, behavioral governance practices can be subjected to a disclosure test (Dowding & Oprea, 2024). According to these authors, a communicative practice (or a nudge) is not manipulation if it meets two rules: (1) it offers a Suitable Reason: the influencer must believe that the reason offered is a good reason for the other person to act; and (2) there is Testimonial Honesty: the influencer must truly believe in the truth of what they are saying or presenting.

The hypothetical disclosure test poses a counterfactual question, namely: would the subject continue to act in the same way if the influencer openly revealed their intentions and the methods being used? If the subject proceeds, the influence is minimal or non-existent, because the subject accepts the reasons provided as valid for their own decision-making. If the subject changes their mind or feels indignant, it is because a high degree of manipulation exists, as the behavior depended entirely on concealment. For example, if a school cafeteria places fruit at eye level so that children eat better and, upon being told, "we have placed the fruit here so you see it first and eat healthily", the child says "oh, okay" and continues to choose the fruit, there is low manipulation. In contrast, if the child feels deceived or coerced upon discovering the trick, the degree of manipulation is greater.

Dowding and Oprea's (2024) proposal is especially relevant when analyzing how shared narratives legitimize behavioral governance, precisely because a narrative is "shared" and legitimate not because the people repeat it, but because it passes the Hypothetical Disclosure Test. If, upon revealing the narrative behind the nudge, the citizen recognizes it as their own and, based on acceptable reasons, then there is no manipulation, but rather a coordinated exercise of power.

From Kahneman's (2011) perspective, this represents an exploitation of System 1: the manipulator leverages our automatic responses and intuitive biases so that System 2 – our rational watchdog – does not even wake up.. Dowding and Oprea's (2024) test acts here as the ultimate intruder detector: if, upon turning on the lights of transparency and revealing the trick, the subject would abort the mission, we are facing pure manipulation.

Now, if this is the case, it is crucial for agents who design behavioral governance policies and practices to first consider which shared narratives their own practices are based on, as it is only from these shared narratives that different behavioral strategies can be viewed as persuasion rather than manipulation.

6 Narrative Paradigm and the Generation of an Unconscious Narrative

If the thesis of the narrative paradigm is correct, human beings do not primarily understand or make decisions based on logical arguments or rational discourses. Instead, we rely on narratives that we unconsciously assimilate, carrying implicit values with which we identify and which

shape our actions. The strength of behavioral governance lies in its ability to subtly connect specific individual actions to collective narratives, such as "together we all contribute to the common good", "we must protect the planet", or "taking care of your health benefits general welfare". These narratives act as anchors that orient behavior without explicit coercion, fostering a sense of shared purpose and responsibility.

Consider issues like sustainability and ecology. Here, narratives construct a moral framework where citizens are cast as "heroes" taking active steps to save the planet, while those who disregard environmental concerns are subtly positioned as "villains" threatening collective survival. This dynamic drives behaviors such as recycling, reducing energy consumption, and supporting green policies. Similarly, in the domain of nutrition, narratives oppose industrial practices that prioritize profit over health and environmental sustainability. They encourage individuals to resist "easy consumption" of highly processed foods, advocating instead for balanced diets that align with ethical values and long-term well-being. These examples highlight how narratives create a framework for behavior, tying individual actions to larger societal goals through shared values and implicit moral codes.

Power, as exercised through behavioral governance, cannot exist without narrative. In this context, narratives are often more unconsciously assimilated than in other forms of governance. Importantly, the relationship between power and narrative in behavioral governance affects not only individuals but also policymakers, who are themselves influenced by the narratives they absorb. For instance, global narratives around climate change or public health often shape policy priorities, creating frameworks that guide decision-making at institutional levels. These frameworks are not neutral but embed specific values and visions of the common good, reinforcing the power of shared narratives in structuring both governance and public behavior.

For these reasons, behavioral governance should not necessarily be viewed as manipulative. Manipulation implies the use of coercion or dominance to achieve a benefit at someone else's expense, often without transparency or consent. In contrast, the exercise of authority through behavioral governance can introduce an order based on good reasons, fostering behaviors that align with the common good. Human beings are interdependent and frequently lack the information or resources to guide behavior autonomously. In such cases, behavioral governance provides a framework for orientation, offering subtle guidance that respects individual agency while promoting collective well-being. However, this approach relies on ethical considerations, ensuring that narratives are transparent, inclusive, and directed toward the shared benefit of society as a whole.

Behavioral governance, when grounded in shared narratives and directed toward the common good, represents a tool that can align individual behavior with collective goals. It leverages the human tendency to find meaning through stories, creating a powerful mechanism for shaping societal dynamics while maintaining the delicate balance between guidance and autonomy.

As such, we claim that behavioral government is not necessarily a form of manipulation that removes freedom from citizens, because life itself demands that those who govern will have to make decisions for us and try to organize society in the best possible way. Power is violent when it goes against our own good, but there is a common good that demands that someone try to establish it through narratives. The risk of behavioral government is that it becomes a manipulative form of politics. In order to avoid excesses it is necessary to have certain mechanisms of transparency and control, such as external audits, but above all, it is necessary that the narratives underlying these practices are based on shared values and are open to dialogue.

7 Conclusions

Throughout this article, we have explored the theoretical foundations of the exercise of power in behavioral governance to highlight several key issues. First, power is always exercised through implicit narratives, and in the case of behavioral governance, it involves orienting behavior through unconscious mechanisms. These mechanisms work by connecting individuals to shared values and collective goals, embedding decisions within a broader narrative framework that subtly guides action.

Given that power operates through shared communicative networks and processes, the logic of behavioral governance need not be seen as manipulation, provided that the underlying narratives are oriented toward the common good and not merely toward achieving individual benefits. Power is not an inherent property of structures, as Dowding (2019, 9) wants us to believe, but is realized through the communities generated within these structures. These communities become subjects of power when shared values and narratives guide governance and decision-making processes. Narratives, therefore, serve as the foundation for collective action, fostering cohesion and legitimacy in governance practices.

The narrative paradigm emphasizes that there is logic and rationality embedded in discourses that do not conform to formal logical structures. Fisher (1989) argues that human understanding is inherently narrative, shaped by stories that carry implicit values and good reasons. These narratives are assimilated unconsciously and influence our decisions in profound ways. For instance, narratives about sustainability, health, or social responsibility shape consumer behavior, encouraging choices that align with broader societal goals. Far from being irrational, these decisions reveal a hidden logic rooted in shared narratives and collective values. Ultimately, the unconscious elements that drive behavior should not be dismissed as irrational but understood as part of a broader rationality that incorporates narrative-based reasoning. Behavioral governance, when grounded in ethical narratives oriented toward the common good, can serve as a powerful tool for fostering collaboration and social cohesion. By recognizing the role of narratives in shaping decisions, we can better understand how power operates within behavioral governance frameworks and ensure that it supports equitable and sustainable outcomes.

In conclusion, the ethical viability of behavioral government depends on its ability to engage both cognitive systems of the citizenry. While choice architecture inevitably influences System 1 through intuitive and automatic heuristics, it is the shared narrative that provides the necessary bridge to System 2 deliberation. By embedding nudges within transparent, inclusive narratives, the exercise of power shifts from a "power-over" (manipulation of the unconscious) to a "power-to" (collective capacity). A narrative that is publicly accessible and rationally defensible ensures that behavioral interventions are not mere techniques of subversion, but tools for agency that respect the individual's capacity to consciously endorse the common good.

To move from theory to practice, behavioral governance must adopt institutional mechanisms that guarantee narrative openness. This includes (1) narrative audits, where independent bodies evaluate the ethical alignment of policy frameworks: a concept rooted in the tradition of frame-critical analysis in public policy (Schön & Rein, 1994). And (2) it must endorse a publicity principle, which mandates the disclosure of the underlying values of every nudge to invite System 2 scrutiny, grounded in the Kantian principle of publicity and developed within the Rawlsian concept of public reason (Rawls, 1993). Institutionalizing narrative audits ensures, on the one hand, that the justificatory framework of a policy is subject to external scrutiny, preventing the slide from persuasion into manipulative propaganda. On the other hand, a "publicity principle" mechanism requires that every behavioral intervention be accompanied by a clear statement

of its narrative goals, allowing System 2 engagement through public contestation (here the hypothetical disclosure test could serve as an effective tool for discerning when a practice aligns with a shared narrative**,** Dowding & Oprea, 2024). These mechanisms provide the "watertight" link between cognitive science and democratic legitimacy, ensuring that power remains a collaborative capacity.

Finding an appropriate form of behavioral governance represents one of the critical challenges of the twenty-first century. In an era where digital media and algorithmic systems play an increasingly influential role, governments must not only regulate these technologies but also provide guidance that aligns individual behavior with societal goals. This guidance, however, must avoid the pitfalls of manipulation. The key lies in ensuring that behavioral governance is exercised transparently, through participatory communication processes that foster shared narratives oriented toward the common good. By grounding governance in ethical principles and inclusivity, it is possible to harness the potential of behavioral governance without compromising individual autonomy or trust.

As societies become more interconnected, the role of narratives in shaping collective action becomes ever more significant. Behavioral governance, when rooted in transparent practices and shared values, has the potential to guide communities toward sustainable and equitable futures. The challenge, therefore, is not to abandon behavioral governance but to refine its methods to ensure that it supports collaboration, equity, and the pursuit of the common good.

References

Adler, E., & Bernstein, S. (2005). Knowledge in power: the epistemic construction of global governance. In M. Barnett, & R. Duvall (Eds.), *Power in Global Governance* (pp. 294–318). Cambridge University Press. https://doi.org/10.1017/CBO9780511491207.013

Arendt, H. (1998). *The Human Condition*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226924571.001.0001

Ariely, D. (2008). *Predictably Irrational*. Harper Collins.

Bachrach, P., & Baratz, M. S. (1970). *Power and Poverty: Theory and Practice*. Oxford University Press.

Battaglio, P. R., Belardinelli, P., Bellé, N., & Cantarelli, P. (2019). Behavioral Public Administration ad fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations. *Public Administration Review,* 79(3), 304–320. https://doi.org/10.1111/puar.12994

Battegazzorre, F. (2017). Some reflections on the power to/power over debate. *Journal of Political Power*, 10(3), 274–285. https://doi.org/10.1080/2158379X.2017.1382169

Castells, M. (2011). A Network Theory of Power. *International Journal of Communication,* 5, 773–787. Online: https://ijoc.org/index.php/ijoc/article/view/1136

Chetty, R. (2015). Behavioral Economics and Public Policy: A Pragmatic Perspective. *American Economic Review,* 105(5), 1–33. https://doi.org/10.1257/aer.p20151108

Dahl, R. A. (1957). The concept of power. *Behavioral Science,* 2(3), 201–215. https://doi.org/10.1002/bs.3830020303

De la Cruz Prego, F. (2021). The Role of the State in the 2030 Agenda: An Economic Policy Approach. *Revista de Economía Mundial,* 63, 97–116. https://doi.org/10.33776/rem.vi63.7149

Dowding, K. (2019). *Rational Choice and Political Power*. Bristol University Press. https://doi.org/10.51952/9781529206340

Dowding, K., & Oprea, A. (2024). Manipulation in politics and public policy. *Economics and Philosophy*, 40(3), 685–710. https://doi.org/10.1017/S0266267124000063

Fisher, W. (1989). *Human Communication as Narration: Toward a Philosophy of Reason, Value, and Action*. University of South Carolina Press. https://doi.org/10.2307/j.ctv1nwbqtk

Foucault, M. (1980). Truth and power. In C. Gordon (Ed.), *Power/Knowledge: Selected Interviews and Other Writings, 1972–1977* (pp. 109–133). Pantheon Books.

Foucault, M. (1997). *Il faut défendre la société: Cours au Collège de France, 1976*. Gallimard/Seuil.

Foucault, M. (2009). *Security, Territory, Population. Lectures at the Collège de France, 1977–1978*. Palgrave Macmillan. https://doi.org/10.1057/9780230245075

Foucault, M. (2014). The politics of health in the eighteenth century. *Foucault Studies,* 18, 113–127. https://doi.org/10.22439/fs.v0i18.4654

Galizzi, M. M. (2017). Behavioral aspects of policy formulation: experiments, behavioral insights, nudges. In M. Howlett, I. Mukherjee, & S. Fraser (Eds.), *Handbook of Policy Formulation. Handbooks of Research on Public Policy* (pp. 410–429). Edward Elgar Publishing. https://doi.org/10.4337/9781784719326.00034

Gigerenzer, G. (2014). Breast cancer screening pamphlets mislead women. All women and women's organizations should tear up the pink ribbons and campaign for honest information. *British Medical Journal,* 348, g2636. https://doi.org/10.1136/bmj.g2636

Gofen, A., Moseley, A., Thomann, E., & Weaver, R. K. (2021). Behavioural governance in the policy process: introduction to the special issue. *Journal of European Public Policy,* 28(5), 633–657. https://doi.org/10.1080/13501763.2021.1912153

Grewal, D. S. (2008). *Network Power. The Social Dynamics of Globalization*. Yale University Press. https://doi.org/10.12987/9780300145120

Grimmelikhuijsen, S., Jilke, S., Olsen, A. L., & Tummers, L. G. (2017). Behavioral Public Administration: Combining Insights from Public Administration and Psychology. *Public Administration Review,* 77(1), 45–56. https://doi.org/10.1111/puar.12609

Grüne-Yanoff, T. (2012). Old wine in new casks: libertarian paternalism still violates liberal principles. *Social Choice and Welfare,* 38, 635–645. https://doi.org/10.1007/s00355-011-0636-0

Haas, P. M. (1992). Epistemic communities and international policy coordination. *International Organization,* 46(1), 1–35. https://doi.org/10.1017/S0020818300001442

Halpern, D., & Sanders, M. (2016). Nudging by Government: Progress, Impact, & Lessons Learned. *Behavioral Science & Policy,* 2(2), 53–65. https://doi.org/10.1177/237946151600200206

Hallsworth, M., Egan, M., Rutter, J., & McCrae, J. (2018). Gobierno conductual: Aplicación de las ciencias del comportamiento para mejorar la toma de decisiones de los gobiernos. Behavioural Insights Team. Online: https://www.bi.team/publications/gobierno-conductual/

Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148, 14–31. https://doi.org/10.1016/j.jpubeco.2017.02.003

Han, B.-C. (2005). *Was ist Macht?*. Reclam.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Lehner, M., Mont, O., & Heiskanen, E. (2015). Nudging – A promising tool for sustainable consumption behaviour?. *Journal of Cleaner Production,* 134, 166–177. https://doi.org/10.1016/j.jclepro.2015.11.086

Lukes, S. (1974). *Power: A Radical View*. Macmillan.

Madrian, B. C., & Shea, D. F. (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *The Quarterly Journal of Economics,* 116(4), 1149–1187. https://doi.org/10.1162/003355301753265543

Miller, P., Kable, J. W., & Lempert, K. M. (2024). Willingness to wait outperforms delay discounting in predicting drinking severity. *Journal of Experimental Analysis of Behavior,* 122(3), 247–258. https://doi.org/10.1002/jeab.4210

Moseley, A., & Thomann, E. (2021). A behavioural model of heuristics and biases in frontline policy implementation. *Policy and Politics,* 49(1), pp. 49–67. https://doi.org/10.1332/030557320X15967973532891

Osman, M., & Thornton, K. (2019). Traffic light labelling of meals to promote sustainable consumption and healthy eating. *Appetite,* 138, 60–71. https://doi.org/10.1016/j.appet.2019.03.015

Parsons, T. (1967). *Sociological Theory and Modern Society*. Free Press.

Pettit, P. (1996). Freedom as Antipower. *Ethics,* 106(3), 576–604. https://doi.org/10.4135/9781446215272.n18

Porter, T. (2009). Global Governance as Configurations of State/Non-State Activity. In J. Whitman (Ed.), *Palgrave Advances in Global governance* (pp. 87–104). Palgrave Macmillan. https://doi.org/10.1057/9780230245310_5

Rawls, J. (1993). *Political Liberalism*. Columbia University Press.

Schön, D., & Rein, M. (1994). *Frame Reflection: Toward the Resolution of Intractable Policy Controversies*. Basic Books.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences,* 4(2), 155–169. https://doi.org/10.1007/BF01405730

Sassen, S. (2006). *Territory, Authority, Rights: From Medieval to Global Assemblages*. Princeton University Press. Online: http://www.jstor.org/stable/j.ctt7sx98

Simon, H. (1978). Rationality as Process and as Product of Thought. *The American Economic Review,* 68(2), 1–16. Online: https://www.jstor.org/stable/1816653

Sunstein, C. R. (2014). Nudging: A Very Short Guide. *Consumer Policy,* 37, 583–588. https://doi.org/10.1007/s10603-014-9273-1

Sunstein, C. R. (2019). *Conformity. The Power of Social Influences*. New York University Press. https://doi.org/10.18574/nyu/9781479896585.001.0001

Thaler, R. H., & Sunstein, C. R. (2003). Libertarian Paternalism. *American Economic Review,* 93(2), 175–179. https://doi.org/10.1257/000282803321947001

Thaler, R. H., & Sunstein, C. R. (2008). *Improving Decisions About Health, Wealth, and Happiness*. Yale University Press.

Thomann, E., van Engen, N. A. M., & Tummers, L. (2018). The necessity of discretion: A behavioral evaluation of bottom-up implementation theory. *Journal of Public Administration Research and Theory, 28*(4), 583–601. https://doi.org/10.1093/jopart/muy024

Thomas, M. D. (2019). Reapplying behavioral symmetry: public choice and choice architecture. *Public Choice,* 180, 11–25. https://doi.org/10.1007/s11127-018-0537-1

Wartenberg, T. E. (1990). *The Forms of Power. From Domination to Transformation*. Temple University Press.

White, M. D. (2013). *The Manipulation of Choice. Ethics and Libertarian Paternalism*. Palgrave Macmillan.

Wilkinson, T. M. (2013). Nudging and Manipulation. *Political Studies*, 61, 341–55. https://doi.org/10.1111/j.1467-9248.2012.00974.x

CONTENTS

ARTICLES