SOSYAL ARAŞTIRMALAR VE YÖNETİM DERGİSİ (SAYOD)
****
JOURNAL OF SOCIAL RESEARCH AND MANAGEMENT

E-ISSN: 2667-5897	https://dergipark.org.tr/tr/pub/sayod	**Paper Type:** Research Paper, **Makale Türü:** Araştırma Makalesi
Sayı:1, Eylül 2026	**Issue:**1, September 2026	**Received Date / Geliş Tarihi:** 18.08/2026 **Accepted Date / Kabul Tarihi:** 29.09.2026

CERTIFICATION INTENSITY AND NATIONAL LOGISTICS PERFORMANCE: A CROSS-COUNTRY ANALYSIS OF ISO MANAGEMENT SYSTEM STANDARDS AND THE WORLD BANK LOGISTICS PERFORMANCE INDEX

♦♦♦

SERTİFİKASYON YOĞUNLUĞU VE ULUSAL LOJİSTİK PERFORMANSI: ISO YÖNETİM SİSTEMİ STANDARTLARI VE DÜNYA BANKASI LOJİSTİK PERFORMANS ENDEKSİ ÜZERİNE ÜLKELER ARASI BİR ANALİZ

Atıf/ to Cite (APA): Öztürk Koçali, E. (2026). Certification intensity and national logistics performance: A cross-country analysis of ISO management system standards and the World Bank Logistics Performance Index. *Sosyal Araştırmalar ve Yönetim Dergisi*, (1), 89-110.

DOI: https://doi.org/10.35375/sayod.2021145

Emine ÖZTÜRK KOÇALİ *

ABSTRACT

Unlike what a bivariate correlation might suggest, ISO certification intensity is not a reliable independent predictor of national logistics performance when controlling for income and institutional capacity. To evaluate the claim, I regress the World Bank's 2023 Logistics Performance Index (LPI) on the 2022 ISO Survey of Management System Standard Certifications and World Bank development and governance data for 127 countries. There exists a strong relationship between per capita ISO 9001 certificates and LPI without controls ($b = 0.225$, $p < .001$, $R^2 = .613$); yet, after controlling for income, trade openness, government effectiveness, and population, the effect of certification intensity drops by 81% to $b = 0.042$ ($p = .0503$), and the residual becomes insignificant by any normal standards. Two more reasons warrant caution. First, because both government effectiveness and LPI are based on perceptions, the regression is specification-sensitive; when government effectiveness is removed from the list of controls, significance is restored ($b = 0.064$, $p = .021$). Second, specification testing shows no statistical difference between the current model and the same specification using the two-year-mismatched certification (interaction $p = .22$). After Holm correction for multiple testing across the six LPI dimensions, only international shipments remain significantly related to certification intensity (adjusted $p = .041$). According to the 2022 ISO Survey, the number of ISO 9001 certificates was 1,265,200 globally, compared with 521 ISO 28000 certificates, for a ratio of about 2,429 to 1; total numbers also differ substantially across survey editions because reports are voluntary and issued by national certification bodies.

Keywords: Logistics Performance Index, Certification intensity, Supply chain management standards.

ÖZ

İki değişkenli bir korelasyondan çıkarılabilecek sonucun aksine; gelir ve kurumsal kapasite değişkenleri kontrol edildiğinde, ISO sertifikasyon yoğunluğu, ulusal lojistik performansı açısından güvenilir ve bağımsız bir öngörücü değildir. Bu iddiayı değerlendirmek amacıyla, 127 ülkeye ait verileri kullanarak Dünya Bankası'nın 2023 Lojistik Performans Endeksi'ni (LPI), 2022 ISO Yönetim Sistemi Standardı Sertifikasyonları Anketi sonuçları ve Dünya

* Lecturer, Istanbul Gelisim University, Istanbul Gelisim Vocational School, Logistics Program, emozturk@gelisim.edu.tr, ORCID ID: https://orcid.org/0000-0001-8975-9459

Bankası'nın kalkınma ve yönetişim verileriyle regresyon analizine tabi tuttum. Kontrol değişkenleri olmaksızın, kişi başına düşen ISO 9001 sertifikası sayısı ile LPI arasında güçlü bir ilişki bulunmaktadır ($b = 0,225$, $p < 0,001$, $R^2 = 0,613$); ancak gelir, dışa açıklık, kamu yönetimi etkinliği ve nüfus değişkenleri kontrol edildiğinde, sertifikasyon yoğunluğunun etkisi %81 oranında azalarak $b = 0,042$ ($p = 0,0503$) seviyesine gerilemekte ve kalan etki, genel kabul görmüş standartlara göre istatistiksel olarak anlamsız hale gelmektedir. Bu sonucun temkinli yorumlanmasını gerektiren iki neden daha mevcuttur. İlk olarak, hem kamu yönetimi etkinliği hem de LPI algılara dayalı veriler olduğundan, regresyon modeli değişken seçimine (spesifikasyona) duyarlıdır; nitekim kamu yönetimi etkinliği kontrol listesinden çıkarıldığında istatistiksel anlamlılık yeniden sağlanmaktadır ($b = 0,064$, $p = 0,021$). İkinci olarak, model spesifikasyon testleri, mevcut model ile aynı değişken yapısına sahip ancak iki yıllık zaman uyumsuzluğuna (gecikmeli veriye) sahip sertifikasyon verilerinin kullanıldığı model arasında istatistiksel bir fark olmadığını göstermektedir (etkileşim $p = 0,22$). Altı LPI boyutuna ilişkin çoklu testler için Holm düzeltmesi uygulandıktan sonra, yalnızca uluslararası sevkiyatların sertifikasyon yoğunluğu ile anlamlı bir ilişki içinde kaldığı görülmektedir (düzeltilmiş $p = 0,041$). 2022 ISO Anketi verilerine göre, dünya genelindeki ISO 9001 sertifikası sayısı 1.265.200 iken ISO 28000 sertifikası sayısı 521'dir; bu da yaklaşık 2.429'a 1'lik bir orana karşılık gelmektedir. Ayrıca, ulusal sertifikasyon kuruluşlarının raporlama süreçlerinin gönüllülük esasına dayanması nedeniyle, toplam sertifika sayıları anket dönemleri arasında önemli farklılıklar göstermektedir.

Anahtar Kelimeler: Lojistik Performans Endeksi, Sertifikasyon yoğunluğu, Tedarik zinciri yönetimi standartları.

1. INTRODUCTION

Quality management standards and logistics performance are usually studied as separate literatures. One stream examines why organizations and countries adopt ISO management system standards, treating certification counts as an outcome to be explained by economic development, trade exposure, and institutional pressure (Guler et al., 2002; Corbett, 2006; Rodriguez-Arnaldo & Martínez-Lorente, 2020). A second stream treats the World Bank's Logistics Performance Index (LPI) as an outcome in its own right, explaining cross-country variation in trade facilitation through infrastructure investment, governance quality, and macroeconomic fundamentals (Jayathilaka et al., 2022). The two streams rarely meet. Certification research treats logistics, at most, as one adopting sector among many. In contrast, LPI research treats quality management systems, at most, as a background institutional condition rather than a measured variable.

Logistics performance is at the core of trade policy, not its periphery. Even a one-day delay before the shipment leads to a decrease in the state's trade volume (Djankov et al., 2010), and the development of the Logistics Performance Index (LPI) has been partly intended to make it possible to compare the resulting trade cost between countries (Hausman et al., 2013). However, while this issue may seem minor, it becomes hard to ignore when comparing certification statistics. According to the 2022 ISO Survey, there are 1,265,200 valid ISO 9001:2015 certificates worldwide; it is a generic quality management system for organizations in any industry. In contrast, ISO 28000:2022, the unique document for supply chain security management in the ISO portfolio, has only 521 valid certificates, creating a 2,429-to-1 ratio. This same difference appears in every ISO Survey with country-level statistics (Section 4.1, Figure 1). It cannot be explained if quality governance in logistics relies on sector-specific instruments.

This study will investigate these two issues. The first question is: Is there a systematic relationship between certification intensity, defined as the density of ISO management system certificates per population in a country, and the country's logistics performance as measured by the Logistics Performance Index (LPI)? If yes, which dimensions of the LPI mediate the

relationship between these two variables? The second question is: Is the ubiquitous presence of the generic standard (ISO 9001), coupled with the absence of the sector-specific standard (ISO 28000), evidence of a decoupling of institutionalized quality signaling and sector-specific governance? This is a phenomenon that has been anticipated in neo-institutional theories of standards adoption (DiMaggio & Powell, 1983). Still, it has never been empirically compared with a generic and a sector-specific standard in the same industry. Both questions share a caveat: the number of ISO 9001 and ISO 28000 certifications considered below is the economy-wide total for all organizations in a country, not the number of these organizations in the logistics sector, since no such information exists at the country level. Therefore, the investigation highlights the national aggregate picture of the distribution between generic and sector-specific certification. We propose sectoral decoupling as a possible explanation for this pattern.

Section 2 provides an extensive overview of the literature on diffusion and LPI determinants. Neither body of literature, in isolation, considers whether an increase in the number of certificates adds value to explaining a country's logistics performance beyond the contributions of income and trade openness. This paper makes contributions in the following four ways. First, it measures the density of certification in terms of the three ISO management systems that are used most broadly in the field of logistics management: ISO 9001, ISO 14001, and ISO 45001, plus ISO 28000, as opposed to the measure of just one standard. Second, it scales certificate numbers by both population and GDP, then examines their relationship with LPI using multivariate regression, controlling for income, trade openness, and government effectiveness, and uses the same reference year in which the certification statistics are compiled, whereas previous cross-country analyses used different years. Third, it evaluates if the association between standard and performance is uniform across income categories, thus extending the finding of Rodríguez-Arnaldo and Martínez-Lorente (2020) on the heterogeneity of the ISO 9001 diffusion effect. Fourth, it examines ISO 28000 to interpret the results: a small marginal effect compared to ISO 9001 is expected given the relatively low diffusion of this standard worldwide, and the paper treats it as meaningful information about logistics quality governance.

To answer the above-stated questions, we use a cross-section of data that combines LPI 2023 scores (obtained from an expert survey conducted in the last months of 2022) with ISO Survey 2022 certification numbers and development/governance indicators provided by the World Bank for the same reference year, covering 127 countries. First, the main finding contradicts what one might expect from a simple correlation. While ISO 9001 density explains more than half of cross-country variation in the LPI, the relationship decreases by about 81% to a non-significant level ($p = .0503$) when controlling for income, trade openness, population size, and government effectiveness. In addition, two more tests limit the remaining relationship rather than increase its significance. On one hand, because government effectiveness and LPI are both largely based on perception surveys of international experts, part of the attenuated effect may result from the measurement construction process itself rather than from the institutional channel; on the other hand, formal testing reveals that the same-year effect cannot be distinguished statistically from an equivalent specification with mismatched two-year certification measure. It should be noted that neither of these problems is crucial for making the key decoupling statement of this paper: national quality governance in logistics, to the extent that it is internationally certified at all, is implemented through a generic management standard and not through the special ISO standard for the sector, and the discrepancy between the two is quite substantial and consistent over survey editions.

2. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

2.1. Neo-Institutional Theory and the Diffusion of Management Standards

The neo-institutional approach explains why formal organizational routines become widespread among firms, and even across countries, without requiring efficiency. According to the classic approach, Meyer & Rowan (1977) suggest that organizations institutionalize "rationalized myths" within their structures to gain legitimacy, resources, and survival, and these structures are not necessarily connected to daily activities. The "logic of confidence" serves as an alternative to coordination and control, typical of an integrated structure. Expanding upon the foundations laid by previous research, DiMaggio & Powell (1983) identified three ways in which this convergence takes place: coercive isomorphism, when organizations are forced into adopting a practice by powerful external agents (government, parent companies, or major trading partners); mimetic isomorphism, when organizations mimic successful organizations; and normative isomorphism, when shared professional standards and training diffuse practices through occupational networks. None of these approaches requires that adopting a practice will improve the organization's technical performance. Legitimacy, not efficiency, becomes the theory's goal.

In their seminal study using empirical evidence to test the institutional theory of certification, Guler et al. (2002) were the first to use a large sample approach to test this theory with respect to ISO 9000. Using panel data on certification in 85 countries between 1993 and 1998, they showed that coercive isomorphism, transmitted through foreign multinational subsidiaries in a country as well as trade relationships with prior adopters, and mimetic isomorphism, transmitted through trade partners who are structurally equivalent to one another, predict a country's ISO 9000 adoption rates. The findings of the Guler et al. (2002) study, that is, the idea that certification diffusion is more dependent on international network position than on any direct measure of a country's quality management capability, form the empirical grounding behind the idea that the number of certificates is an institutional indicator, not a technical one. This hypothesis is tested directly in the current study through performance outcomes rather than additional diffusion measures. At the firm level, Boiral (2003), based on an analysis of ISO 9000-certified firms, complements Guler et al.'s country-level results using the "iron cage" metaphor of DiMaggio and Powell. Employees and managers in such firms repeatedly describe certification as a ritualistic process of audits overlaid on existing practice, rather than a transformation in how work is performed.

2.2 Signaling and Supply Chain Coercion

On the other hand, some literature takes the alternative position that adopting ISO 9001 does not stem from imitation at all, but rather from a rational decision to signal quality to overcome information asymmetry in the supply chain. Connelly et al. (2011) develop the necessary and sufficient conditions under which an expensive and observable action serves as a credible signal of the quality that can otherwise not be observed. Namely, the signal must be sufficiently costly to deter imitation by lower-quality senders, and receivers must be able to understand it without access to the private information. The third-party certification scheme, close to the one Spence (1973) originally formulated for the labor market, fits this framework well. Corbett (2006) tests the widely accepted assumption that buyers' need for certification drives ISO 9000 adoption further up the supply chain by breaking the assumption into four testable hypotheses and finding evidence for supply-chain diffusion based on firm adoption timing data. Terlaak and King (2006) develop the signaling approach directly. Using an eleven-year panel of United States manufacturing facilities, they show that certified plants grow faster after certification, that operational improvements in d cannot explain this growth, and that the effect is stronger when the rs have difficulty improving the supplier's quality.

These two mechanisms, institutional isomorphism and market signaling, are not mutually exclusive, and both predict the same observable pattern at the country level. Certification density should rise with a country's integration into international trade and with the presence of powerful trading partners who reward or require it, regardless of whether certification

changes the underlying quality of what is produced. Neither mechanism, on its own, predicts whether the resulting stock of certificates is associated with a country's realized logistics performance the way an outside observer, such as an LPI survey respondent, would assess it. That is the question this paper adds to the literature.

2.3. Determinants of Certification Diffusion and of Logistics Performance

Two bodies of country-level determinants research bear directly on the control variables used here. Rodriguez-Arnaldo and Martínez-Lorente (2020), analyzing ISO 9001 diffusion across countries from 2009 to 2018, find that economic development predicts certification density strongly among lower-income countries but that the relationship attenuates, and in some specifications reverses sign, among high-income countries where certification is already broadly saturated: evidence of a nonlinear, income-conditional diffusion process rather than a uniform one. Separately, the LPI-determinants literature has established GDP and trade openness as leading correlates of logistics performance and, through it, of trade volumes. Hausman et al. (2013), in one of the founding empirical treatments of the relationship, show that logistics performance has direct and indirect effects on trade volumes and export competitiveness; Martí et al. (2014) similarly find the LPI a significant positive determinant of bilateral trade flows using a gravity-model approach; and Jayathilaka et al. (2022), analyzing 142 countries nearly a decade later, find that GDP and the LPI jointly and positively predict trade volumes, with the effect strongest in Asia, Europe, and Oceania. Çelebi (2019) adds that this trade-promoting effect is itself income-conditional, with low-income economies capturing disproportionately larger gains from a comparable logistics improvement than high-income economies do, a heterogeneity pattern structurally similar to the one Rodriguez-Arnaldo and Martínez-Lorente document for certification diffusion. The mechanism is at least partly about time: Djankov et al. (2010) show that each additional day of delay before shipment reduces trade by roughly 1 percent, a direct estimate of the cost the LPI's timeliness and customs components are designed to proxy. Rezaei, van Roekel, and Tavasszy (2018), surveying 107 academics and logistics professionals with the Best Worst Method, further show that professionals do not weight the LPI's six components equally in practice even though the World Bank's published composite treats them as such, a methodological caution relevant to any study, including this one, that uses the overall LPI score as a single dependent variable. None of this literature treats certification as a variable in the model.

Read together, these two literatures imply that certification intensity and logistics performance share common upstream determinants: income, trade exposure, and the institutional capacity that both facilitates trade integration and is measured, imperfectly, by governance indicators such as the World Bank's government effectiveness index (Kaufmann, Kraay, & Mastruzzi, 2011). A bivariate association between certification intensity and the LPI is therefore uninformative on its own. It could reflect a genuine contribution of certified process quality to logistics outcomes, or it could simply reflect the fact that both variables are downstream of the same national conditions. Separating these possibilities requires the multivariate, controlled specification developed in Section 3.

2.4. The Sector-Specific Standard as a Decoupling Diagnostic

As discussed in Section 1 above, the certification densities mentioned in this section are economy-wide, not limited to the logistics industry; therefore, this discussion should be taken as an indication in support of sectoral institutional decoupling, not as an empirical example of it. With this qualification in mind, if certification density is indeed dictated by governance requirements in the respective sectors, one would expect the use of the logistics-specific tool, ISO 28000, among trading economies with logistics sectors. It does not: the 2022 ISO Survey records 521 valid ISO 28000 certificates worldwide, against 1,265,200 for the generic ISO

9001 standard (Section 4.1). Bromley and Powell (2012), revisiting the decoupling concept forty years after Meyer and Rowan's original formulation, argue that contemporary decoupling increasingly takes the form of a gap between means and ends rather than a simple gap between stated policy and day-to-day practice: organizations adopt elaborate procedural apparatus whose connection to the substantive outcome it is nominally meant to produce becomes progressively harder to trace. The ISO 9001–ISO 28000 asymmetry fits this "means-ends decoupling" pattern closely: an elaborate, near-universal certification apparatus coexists with a barely adopted, purpose-specific instrument, and the connection between the former and actual logistics-sector security outcomes is exactly the kind of traceability gap Bromley and Powell describe. Two additional examples from parallel sectors support this notion. According to Boiral (2007), the implementation of ISO 14001 has largely been characterized as "a rational myth," pursued to gain legitimacy and market access rather than to drive environmental change. Similarly, Mirtsch, Pohlisch, and Blind (2025), when analyzing ISO/IEC 27001, argue that certification is often used where regulation is less developed and is done in combination with ISO 9001.

With a comparable logic of bundling applied to the adoption of ISO 28000, while the adoption, in absolute numbers, is low, it can be expected to mirror the adoption of ISO 9001 reasonably well, wherever it occurs, despite the overall rate being too low for it to expect a statistically significant independent relationship with the LPI. This makes it possible to treat ISO 28000 not as another control variable added automatically, but as an actual case study: a sector-specific standard developed specifically for the industry in question, which is adopted at three orders of magnitude lower frequency than the general ISO 9001 standard, which provides descriptive evidence in line with the predictions of decoupled neo-institutionalism (Meyer & Rowan, 1977; DiMaggio & Powell, 1983). However, it does not demonstrate the theory in the logistics industry, since the ISO 9001 certification numbers mentioned above are economy-wide. Testing the weak-regulation compensation logic – for instance, whether the adoption of ISO 28000 is higher in cases where supply chain security regulation is relatively weak would require a cross-country proxy for supply chain security regulation strength, which unfortunately is not available for the 127 countries covered in this study; it seems like an interesting area for further research.

2.5. Mixed Empirical Evidence on Management Standards' Substantive Effects

This paper's decoupling argument should not be read as claiming that ISO management system standards never have substantive effects. The empirical literature on that narrower question is genuinely divided, which helps inform how to interpret the present results. Sampaio et al. (2009), reviewing roughly one hundred ISO 9001 certification studies, and Heras-Saizarbitoria and Boiral (2013), synthesizing the broader meta-standards literature spanning ISO 9001 and ISO 14001, both note that most studies rely on survey methodology and self-reported perceptions and produce frequently contradictory conclusions, an unusually opinion-laden evidence base for a topic that ought to be empirically tractable; both call for research designs less dependent on respondents' own perceptions of whether certification "worked."

Levine and Toffel (2010) represent the strongest empirical case for a substantive effect. Using matched-sample data on nearly 1,000 California firms with objective outcomes, survival, sales growth, employment growth, and injury rates, rather than survey-based perceptions, they find ISO 9001 adopters had substantially lower organizational death rates and faster subsequent growth than matched non-adopters. This is a genuine substantive effect estimated from administrative rather than perceptual data, and it sits uncomfortably next to the decoupling account emphasized elsewhere in this literature review. The tension need not be resolved here: Levine and Toffel's result concerns firm survival and growth in a single U.S. state, while the present paper concerns country-level logistics-sector performance as assessed

by international professionals. Both can be true simultaneously: certification can carry real firm-level operational value in some contexts while, in the aggregate, functioning as a poor proxy for a country's logistics-sector performance once development and governance are accounted for (Section 4.2). This paper focuses on the latter claim, not a general verdict on whether ISO 9001 "works."

2.6. Hypothesis Development

Drawing on the above mechanisms, four hypotheses provide the framework for the empirical work:

H1: With controls on income, openness to trade, government effectiveness, and population, there is a significant positive relationship between the degree of ISO 9001 certification and a country's LPI rating, consistent with both the supply chain coercion theory (Corbett, 2006) and the cross-border signaling theory (Terlaak & King, 2006).

H2: The correlation between the two variables, identified in H1, will be most prominent in the LPI sub-component "logistics competence and quality" and not evenly distributed among the six LPI sub-components. This hypothesis is based on construct validity: ISO 9001 certification concerns process quality at the firm level and is therefore closest to perceptions of service competence. Infrastructure quality and customs efficiency, in contrast, are mostly determined by public spending and regulation.

H3: There will be income-specific differences in the correlation strength. Rodriguez-Arnaldo and Martínez-Lorente (2020) found that a country's level of economic development predicts ISO 9001 diffusion strongly only among lower-income countries, with the relationship attenuating among high-income countries where certification is already broadly saturated. A parallel pattern is expected here: the marginal logistics-performance return to additional certification intensity should be larger where certification density is currently low (lower- and middle-income countries) and smaller where it already approaches saturation (high-income countries).

H4: Certification Intensity of ISO 28000, estimated together with ISO 9001 certification intensity and the macroeconomic and governance controls in the underlying model, is not expected to have a statistically significant independent association with a country's LPI score under the intensity specification. This null finding is considered a priori expected for the same reason as the lack of a statistically significant independent effect of certification on the country-level indicator of institutional environment (Section 4.2.2): the low level of global adoption of this standard (Section 2.4). This study uses economy-wide ISO 9001 certification rather than sector-specific standards (Section 1), which also helps explain this descriptive asymmetry. Nonetheless, Section 4.5 will reiterate why this finding alone cannot prove the non-existence of a relationship between the sector-specific standard and logistics performance.

3. DATA AND METHOD

3.1. Data Sources

The dependent variable is drawn from the World Bank's 2023 Logistics Performance Index, published as Connecting to Compete 2023 (Arvis et al., 2023). The 2023 wave is based on an expert survey of international logistics professionals fielded in the final months of 2022, yielding assessments for 139 economies across six components (customs, infrastructure, ease of arranging international shipments, logistics competence and quality, tracking and tracing, and timeliness) together with an overall composite score; this study uses both the overall score and the six components, the latter needed to test H2. The LPI 2023 wave is also the final edition of the World Bank's original expert-survey-based methodology: from 2025 the index has been restructured around shipment-level tracking data under a new LPI 2.0 framework, so the present analysis documents a specific, now-closed measurement regime. The certification

variables are drawn from the ISO Survey of Management System Standard Certifications, 2022 edition, which reports the number of valid certificates for each country as
Three aspects of data quality in the ISO Survey should be considered when interpreting the results. First, certification body participation is optional for each edition, and the 2023 edition shows a sharp fall in reported certificates compared with 2022 and 2024 (see Figure 1). According to the ISO Survey documentation, this drop is due to low reporting, not decertification (ISO, 2024); Section 5 explores the implications of this phenomenon for certificate-count credibility. Second, ISO 28000 certificate count data were affected by a reporting error in 2023 and are therefore unavailable for 2023; data for 2022 and 2024 are unaffected. Third, the LPI is a perception-based survey, not a transaction-level measure, and is suitable for assessing a signaling mechanism.

3.2 Sample and Variable Construction

The 2023 LPI covers 139 economies. Out of these, 138 are included in the merged database used for this research. In contrast, the one that has been omitted at the merging point is Taiwan, China, since it is not included in the World Bank's development and governance datasets (World Development Indicators and Worldwide Governance Indicators), which are mandatory for every control variable in the present study, and no alternative source has been found. Besides, eleven more economies have been omitted for having missing World Bank control variable data in 2022: Democratic Republic of Congo (government effectiveness missing), Cuba (GDP per capita missing), Grenada, Guyana, Jamaica, Lao PDR, Liberia, Nigeria, Papua New Guinea, and Trinidad and Tobago (all – trade openness), and Yemen (both GDP per capita and trade openness missing). Thus, the present study includes 127 economies, spanning all four World Bank groups and seven regions. The omissions mainly involve smaller, poorer economies with incomplete World Development Indicators coverage, and the low-income group is the most affected (Table 5).
The number of certificates is normalized by population size (certificates per million people), rather than using raw numbers, to eliminate the effect of population size on certification efforts. As a result of the significant right-skewness of the certificate intensity, which is characterized by a few economies, mostly smaller, trade-oriented European countries having much higher certificate numbers per capita compared to the rest, all certification intensity variables are log-transformed as ln(1+certificates per million). Therefore, the main independent variable is the logarithm of one plus ISO 9001 certificates per million in 2022. We use logarithms of per capita GDP and population to address skewness and variance-stability issues.

3.3 Model Specification and Estimation

The core specification is an ordinary least squares regression:

$$LPIi = \beta0 + \beta1 \cdot \ln(1+ISO9001pm)i + \beta2 \cdot \ln(GDPpc)i + \beta3 \cdot TradeOpeni + \beta4 \cdot GovEffi + \beta5 \cdot \ln(Pop)i + \varepsilon i \quad (1)$$

The estimation methodology involves two stages: (i) estimation of the LPI index as the dependent variable to test H1 and (ii) estimation of the six LPI subindices as the dependent variables to test H2. We run three regressions consecutively: M1, a bivariate regression to estimate the unconditional association; M2, which includes log GDP per capita to capture the income effect; and M3, the full regression, which includes trade openness, government effectiveness, and log population in addition to M2. In all estimates, heteroskedasticity-robust (HC3) standard errors are used instead of classical OLS standard errors, since the assumption of constant error variance does not hold in these models; robust standard errors protect inference from heterogeneity in error variance, regardless of its source. HC3 standard errors are preferred over HC1/HC2 because they perform slightly better in small to moderate samples. A Breusch-Pagan test for the preferred specification fails to reject homoskedasticity

(LM = 9.14, p = .104, Section 4.6). Nonetheless, we do not assume heteroskedasticity is absent.

Testing H3 requires re-estimating the full model for each World Bank income group, in addition to the pooled model that multiplies the certification variable by an income high/low dummy. Because the low-income country group has a small sample size (N = 11), which prevents estimation of a reliable model with five predictors, the results for this group are presented descriptively only. To test the validity of H4, we run the model with ISO 28000 intensity as a single predictor; then we add the controls, followed by ISO 9001 intensity and the controls. To check the robustness of the ISO 28000 coefficient estimate to two possible factors, two additional tests are run: (i) exclusion of the single largest user of ISO 28000 (Colombia); (ii) transformation of the ISO 28000 continuous intensity variable into a dichotomous variable indicating ISO 28000 usage. Since all six regressions testing the validity of H2 use the same predictor, the p-values are supplemented by Holm-adjusted p-values. Six robustness tests are performed: (1) use of a composite variable of the three generic standards (ISO 9001, 14001, and 45001) instead of the ISO 9001 standard; (2) using GDP instead of the population as the denominator for measuring certificate intensity because ISO certificates are issued to organizations not to individuals, and thus population is a poor measure of exposure to the risk (as discussed more in Limitations of Section 5.7); (3) removing Germany, which is the largest certificate issuing country in the sample, as a check of leverage; (4) including World Bank region fixed effect; (5) excluding government effectiveness from the list of control variables in order to test the contribution of government effectiveness in generating the attenuation effect reported in Section 4.2; and (6) dropping all observations where the value of Cook's distance is larger than the usual 4/n cut-off point, which may be considered as a potentially influential observation rather than outlier because the exclusion is based on regression leverage not the incorrectness of the original data. In addition, as a sensitivity test to the temporal difference, we stack the same-year (2022) and mismatched-year (2024) certificate intensities in a pooled regression with a year indicator and an interaction between year and certification. Standard errors are clustered by country, as is common practice when comparing coefficients across years, but this comparison should be treated as a sensitivity to the measurement year rather than to causality.

4. RESULTS

4.1. Sample, Certification Landscape, and the Reliability of Certificates' Counts

The analyzed sample consists of 127 economies, broken down as follows: 53 high-income, 38 upper-middle-income, 25 lower-middle-income, and 11 low-income economies. Descriptive statistics are shown in Table 1. The average score on the LPI scale is 3.041 with a standard deviation of 0.594; the minimum value is 1.9, and the maximum is 4.3. ISO 9001 certificates' intensity has the following parameters: mean = 234.6 per million inhabitants, pronounced right skew, standard deviation = 302.0, median = 79.3, and maximum = 1,596.5 (for Italy). ISO 28000 intensity is quite low, at 0.257 certificates per million inhabitants. Of the 127 economies, 73 (57.5%) have no ISO 28000 certificate.

Table 1. Descriptive statistics (N = 127, reference year 2022)

Variable	Mean	SD	Min	Median	Max
LPI overall	3.041	0.594	1.900	3.000	4.300
LPI customs	2.844	0.624	1.500	2.700	4.200
LPI infrastructure	2.966	0.727	1.700	2.800	4.600
LPI international shipments	2.967	0.509	1.800	3.000	4.100
LPI logistics competence	3.074	0.644	1.800	2.900	4.400
LPI tracking and tracing	3.090	0.676	1.600	3.000	4.400
LPI timeliness	3.270	0.567	2.100	3.200	4.300
ISO 9001 certificates per million	234.624	302.033	0.196	79.282	1,596.512
ISO 9001+14001+45001 per million	384.247	483.365	0.475	111.865	2,198.897
ISO 28000 certificates per million	0.257	0.621	0.000	0.000	3.498
ln(1 + ISO 9001 per million)	4.101	2.065	0.179	4.386	7.376
ln(1 + ISO Main-3 per million)	4.539	2.144	0.389	4.726	7.696
ln(1 + ISO 28000 per million)	0.160	0.325	0.000	0.000	1.504
GDP per capita (current US$)	21,165.5	26,249.6	357.3	9,281.3	123,719.7
ln GDP per capita	9.138	1.413	5.878	9.136	11.726
Trade openness (% of GDP)	102.546	66.740	2.699	88.701	412.177
Government effectiveness	0.134	1.416	−3.119	0.074	3.002
Population	51,885,242	180,812,066	92,840	10,437,805	1,425,423,212
ln population	16.255	1.656	11.439	16.161	21.078

Note: *Certification variables refer to 2022, consistent with the reference period of the LPI 2023 (conducted September – November 2022). Monetary values are stated in terms of current US Dollars.*

Figure 1 shows the 2022 data used in the main analysis, along with the subsequent 2023 and 2024 editions of the ISO Survey. The global number of ISO 9001 certifications decreased from 1,265,200 in 2022 to 836,946 in 2023 (about 34%) and then increased to 1,474,118 in 2024 (about 76%). Because it is a stock variable, such changes usually proceed relatively slowly without decertification. The same holds for ISO 14001 and ISO 45001 – the former decreases by about 43% from 2022 to 2023 and then grows by about 125% from 2023 to 2024; the latter decreases by about 53% and then increases by about 193% during the same periods of time. The ISO Survey itself notes that the decrease in 2023 is due to a drop in the number of certification bodies included in the survey, not decertification (ISO, 2024). This issue matters in Section 5, where national certificate numbers reported in any given edition of the survey are partly determined by accreditation-system behavior.

The asymmetry between the generic standard (ISO 9001) and the sector-specific standard (ISO 28000) from Section 1 is supported by the 2022 data: 1,265,200 ISO 9001 certificates globally versus 521 ISO 28000 certificates, for a ratio of approximately 2,429 to 1. Almost half of the global total of ISO 28000 certificates (Colombia – 181, Italy – 42, India – 26, China – 20) is accounted for by four countries.

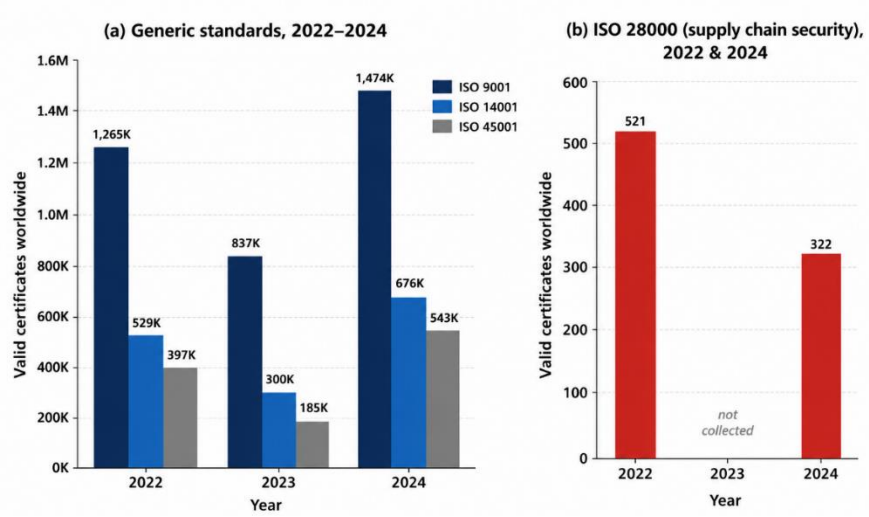


Figure 1. Global ISO management-system certificates by standard and year, 2022–2024.

Table 2 reports bivariate correlations. Overall LPI correlates strongly with government effectiveness (r = .859), log GDP per capita (r = .816), and log ISO 9001 intensity (r = .783). Certification intensity is itself strongly correlated with development (r = .853) and institutional quality (r = .774), which foreshadows the attenuation reported next.

Table 2. Pearson correlation matrix (N = 127)

	LPI	ln9001	ln28000	lnGDPpc	Trade%	GovEff	lnPop
LPI overall	1.000	.783	.332	.816	.347	.859	.100
ln(1+ISO9001 pm)	.783	1.000	.400	.853	.397	.774	−.044
ln(1+ISO28000 pm)	.332	.400	1.000	.327	.370	.324	−.192
ln GDP per capita	.816	.853	.327	1.000	.403	.854	−.142
Trade openness %GDP	.347	.397	.370	.403	1.000	.410	−.448
Government effectiveness	.859	.774	.324	.854	.410	1.000	−.111
ln population	.100	−.044	−.192	−.142	−.448	−.111	1.000

Note: *"ln9001" and "ln28000" are ln(1 + certificates per million population). |r| ≥ .17 significant at p < .05.*

4.2. Certification Intensity and Overall Logistics Performance (H1)

Table 3 presents the sequential specifications with HC3-robust standard errors, and Figure 2 plots the underlying bivariate relationship. The bivariate model (M1) returns a coefficient of 0.225 on ln(1 + ISO 9001 per million), p < .001, explaining 61.3 percent of the variance in overall LPI. A one-unit increase in the logged measure, roughly a 2.7-fold increase in certificates per capita, corresponds to a 0.225-point gain on the five-point LPI scale. Adding log GDP per capita (M2) reduces the coefficient to 0.092, a 59.2 percent attenuation, though it remains significant (p = .002). Introducing the full control set (M3) reduces it further to 0.042, a further 54.2 percent fall, for cumulative attenuation of 81.3 percent from M1 to M3. The M3 coefficient is not significant at conventional levels (SE = 0.022, t = 1.96, p = .0503, 95% CI [−0.0001, 0.084]). Sections 4.6 and 5.2 report two further checks (a formal same-year-versus-mismatched-year test and a measurement-validity concern about the government-effectiveness control) that temper this residual further rather than strengthen it.

Table 3. ISO 9001 intensity and overall LPI: sequential OLS models (HC3-robust SE)

Predictor	M1	M2	M3
ln(1+ISO9001 per million)	0.225*** (0.015)	0.092** (0.029)	0.042† (0.022)
ln GDP per capita	—	0.229*** (0.042)	0.096* (0.042)
Trade openness (% GDP)	—	—	0.001 (0.0004)
Government effectiveness	—	—	0.228*** (0.035)
ln population	—	—	0.084*** (0.018)
N	127	127	127
R^2	.613	.694	.819
Adjusted R^2	.610	.689	.811
F	219.50***	157.83***	101.46***

Note: *Unstandardized coefficients; HC3 robust standard errors in parentheses. * p < .05, ** p < .01, *** p < .001. For the coefficient of M3 for the certification variable (p = .0503), an exact figure is provided since it falls very close to the significance level of 5%.*

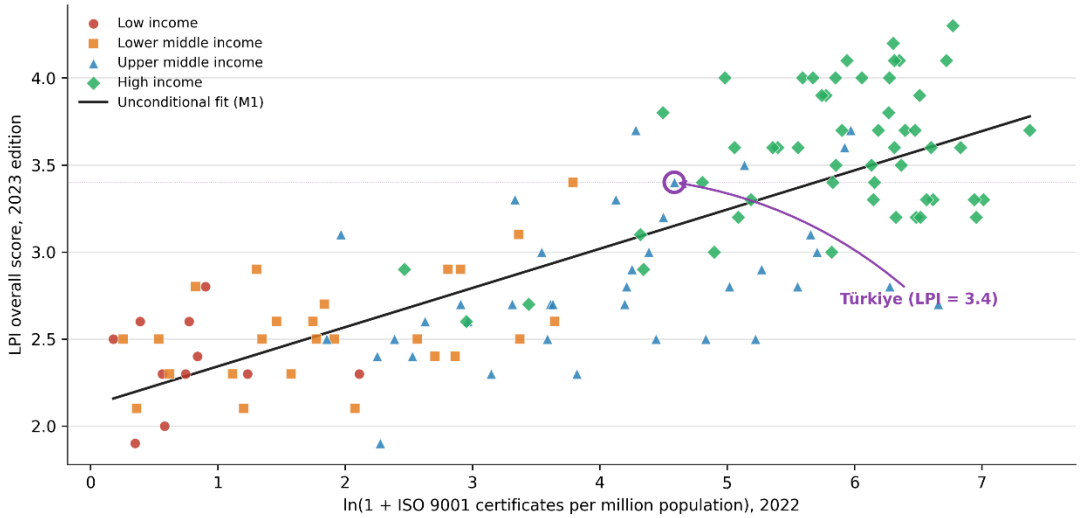


Figure 2. Intensity of ISO 9001 certification and performance of the logistics sector (N = 127, same year).

In the full model, government effectiveness has the highest standardized coefficient (β = 0.545, p < .001), followed by log population (β = 0.234, p < .001) and log GDP per capita (β = 0.228, p = .024). The standardized effect size of the certification variable (β = 0.146) is lower than that of the previous three variables and statistically insignificant. VIFs are low (certification intensity = 3.98; log GDP per capita = 5.74; government effectiveness = 3.87; trade openness = 1.56; log population = 1.32), so attenuation cannot be attributed to severe collinearity among the controls in the classical sense. Still, the VIF for GDP per capita is substantial (5.74), and the coefficients and their standard errors from the correlated development and governance controls must be interpreted cautiously, not as independent estimates. Section 5.2 provides an additional methodological explanation for why the control for government effectiveness might absorb more than just an institutional dimension. Running the regression of government effectiveness on certification intensity, controlling for log GDP per capita, we obtain the positive but marginally significant coefficient (b = 0.116, p = .095).

4.3. Component-Level Decomposition (H2)

H2 claims that any significant relationship will be limited to the logistics competence and quality subcomponent. Table 4 below provides estimates of the full regression for each of the six LPI components separately, with Holm-adjusted p-values reported alongside the

uncorrected results to address multiple testing across six correlated hypotheses on the same regressor. The empirical results falsify the hypothesis on both counts. The logistics competence variable has the lowest coefficient among all six LPI components (b = 0.024, β = 0.075, unadjusted p = .368), and it remains non-significant even after adjusting for multiplicity. International shipments show the strongest relationship (b = 0.066, β = 0.267, unadjusted p = .007) and are the only LPI component that passes Holm's adjustment (p = .041). The timeliness (b = 0.062, p = .029) and tracking and tracing (b = 0.059, p = .057) LPI components lose significance after adjustment (Holm-adjusted p = .147 and .228). At the same time, customs and infrastructure are non-significant irrespective of adjustment. H2 is falsified. After adjustment for multiple testing, only the international shipments component shows a significant relationship, which we discuss further in Section 5.3.

Table 4. Component-level estimation results for the entire (coefficient on ln(1+ISO9001 per million))

LPI component	b	SE	t	raw p	Holm p	β (std.)	R^2
International shipments	0.066	0.024	2.70	.007**	.041*	.267	.656
Timeliness	0.062	0.028	2.18	.029*	.147	.225	.713
Tracking and tracing	0.059	0.031	1.90	.057	.228	.179	.778
Overall LPI	0.042	0.022	1.96	.050	—	.146	.819
Logistics competence	0.024	0.026	0.90	.368	1.000	.075	.761
Infrastructure	0.015	0.026	0.56	.578	1.000	.042	.799
Customs	0.013	0.025	0.53	.595	.595	.044	.800

Note: *Holm corrected p-values are provided for the six LPI component regressions (the Overall LPI line is omitted since it represents the H1 regression from Table 3 and is not an orthogonal LPI component).* p < .05, ** p < .01. The only remaining significance is for international shipping.*

4.4. Income-Group Heterogeneity (H3)

Table 5 shows the results from the subsamples. The point estimates from the upper-middle-income group (0.094) and the lower-middle-income group (0.074) are larger than those from the high-income group (0.056), but not necessarily in a monotonic way, since the lower-middle-income group estimate is smaller than that from the upper-middle-income group. None of the three point estimates is statistically significant. The combined interaction term ISO9001 × high-income, where high-income countries are coded as 1 and all other countries as 0, has a small effect and is not statistically significant (b = 0.041, p = .468). If interpreted literally, the interaction effect should be positive, indicating a higher slope for high-income economies. This would contradict the results from the separate regressions above. The subgroup regression models allow all coefficients to vary freely across income groups, not just the ISO9001 slope. However, in the pooled interaction model, we impose equality of the coefficients on GDP per capita, trade openness, government effectiveness, and population across income groups, allowing only the ISO9001 slope and the intercept to vary. The coefficient on log GDP per capita in the pooled model (0.047, nonsignificant) differs substantially from estimates without subgroup restrictions; thus, the constraint affects the results. This means that the sign of the interaction term cannot be taken as evidence independent of the above constraint. Both strategies share a key feature: the interaction term is not statistically significant, regardless of sign. The only model with a statistically significant outcome is the pooled non-high-income subsample (N=74, b=0.070, p=.043), which combines the middle-income groups ad hoc to increase estimation power above that of the individual groups. Because this combination is not specified formally, the interaction test does not support H3. At best, the larger estimates of the non-high-income economies may suggest an indicative (not confirmed) pattern.

Table 5. Heterogeneity by income group in the relationship between ISO 9001 and LPI.

Specification	N	b	SE	p	R^2
High income	53	0.056	0.049	.260	.754
Upper-middle income	38	0.094	0.053	.078	.621
Lower-middle income	25	0.074	0.123	.546	.405
Low income	11	—	—	descriptive only	—
Pooled: ISO9001 × high-income (interaction term)	127	0.041	0.057	.468	.834
Pooled, non-high-income only	74	0.070	0.035	.043	.572

4.5. ISO 28000 as a Decoupling Diagnostic (H4)

Table 6A presents three models exploring ISO 28000 adoption. While the bivariate coefficient is positive and significant (b = 0.606, p = .001), the model explains merely 11.0 percent of the variance. With the inclusion of controls, the coefficient declines to 0.120 and becomes insignificant (p = .136). Including both standards together (M9) shows that ISO 9001 has a coefficient of 0.036 (p = .096), while ISO 28000 has 0.090 (p = .264); hence, in this continuous model, neither standard shows a significant relation with the dependent variable once development and governance are controlled. A p-value above .05 means the null hypothesis cannot be rejected in this sample; it does not necessarily mean that there is no relationship between the variables. The following two additional analyses explain why this distinction matters here. Removing Colombia where ISO 28000 certifications account for over one third of global ISO 28000 certifications shows that the coefficient for ISO 9001 changes little (b = 0.037, p = .083) while that for ISO 28000 becomes closer to significance (b = 0.146, p = .054, N = 126); thus, the null finding is not due to Colombia's oversized effect that would make it more significant, and actually Colombia inclusion slightly dampens the effect. Using a dichotomous ISO 28000 indicator instead of the intensity score changes the picture because the indicator shows a significant relation with the LPI (b = 0.133, p = .021). The complete results for both of these checks are presented in Table 6B. In 2022, out of the 127 surveyed economies, fifty-four (42.5 percent) had at least one ISO 28000 certificate, which represents a higher share compared to other editions of the survey but still leaves us with an extreme concentration in the possession of such certificates: Columbia alone owns one third of all ISO 28000 certificates, followed by Italy, India and China—three large and diverse economies that have a very small number of ISO 28000 certificates along with thousands of ISO 9001 certificates. These findings support only the continued specification of H4: no statistically significant independent effect emerges even after including control variables, and the raw adoption rate of the sector-specific ISO standard remains three orders of magnitude lower than that of the generic standard, regardless of the survey edition. The significant binary indicator suggests that the lack of a statistically significant effect does not mean ISO 28000 certification is irrelevant to logistics performance; it means that any additional intensity beyond initial adoption does not appear to affect logistics performance in this sample. No equivalence test was performed here; therefore, the lack of a continuous effect is stated as an inability to reject the null hypothesis.

Table 6A. ISO 28000 specifications

Model	Predictor	b	SE	p	R^2
M7 (bivariate)	ln(1+ISO28000 pm)	0.606	0.185	.001**	.110
M8 (+ controls)	ln(1+ISO28000 pm)	0.120	0.080	.136	.817
M9 (both standards + controls)	ln(1+ISO9001 pm)	0.036	0.022	.096†	.821
M9 (both standards + controls)	ln(1+ISO28000 pm)	0.090	0.081	.264	.821

Note: *† p < .10, ** p < .01. All models include N = 127 observations; models M8 and M9 account for the entire set of controls, including log GDP per capita, trade openness, government effectiveness, and log population. The signification marks conform to the standard convention of the paper: † p < .10, * p < .05, ** p < .01, *** p < .001.*

Table 6B. ISO 28000 sensitivity checks (full specifications)

Model	Predictor	b	SE	95% CI	p	N	R^2
M9 (baseline)	ln(1+ISO9001 pm)	0.036	0.022	[−0.006, 0.079]	.096†	127	.821
M9 (baseline)	ln(1+ISO28000 pm)	0.090	0.081	[−0.068, 0.248]	.264	127	.821
M9 (excl. Colombia)	ln(1+ISO9001 pm)	0.037	0.022	[−0.005, 0.079]	.084†	126	.824
M9 (excl. Colombia)	ln(1+ISO28000 pm)	0.146	0.076	[−0.002, 0.294]	.054†	126	.824
M9 (binary ISO28000)	ln(1+ISO9001 pm)	0.031	0.022	[−0.012, 0.073]	.157	127	.827
M9 (binary ISO28000)	ISO28000 ≥1 (binary)	0.133	0.058	[0.020, 0.247]	.021*	127	.827

Note: *Full-specification (M9) sensitivity checks for H4, reported in coefficients, standard errors, 95% confidence intervals, and model fit; all models include the full control set (ln GDP per capita, trade openness, government effectiveness, ln population). † p < .10, * p < .05. The 'excl. Colombia' and 'binary ISO28000' rows correspond to the two sensitivity analyses discussed in Section 4.5.*

4.6. Robustness and Temporal Sensitivity

Table 7 gives six robustness tests for the M3 specification. Substituting the ISO 9001 certificate with the three-standard composite (ISO 9001 + 14001 + 45001) yields a coefficient similar in magnitude but marginally insignificant (b = 0.037, p = .084). When the intensity of certificates is recalibrated using GDP instead of population (that is, the number of the composite standards divided by the current-price GDP measured in billions of US$ [certificates per billion USD]), the raw coefficient becomes even smaller (b = 0.034, p = .145). Since the two denominators calibrate the variable on different scales, it is not possible to compare the raw coefficients; hence standardized coefficients are presented: β = .134 for the population-calibrated specification and β = .063 for the GDP-calibrated specification, which means that the GDP-calibrated variable has an objectively weaker relationship than the smaller raw one. If we exclude Germany, the largest certification economy in the sample, the coefficient remains virtually unchanged and just below the .05 level of significance (b = 0.042, p = .049, N = 126). Adding World Bank region fixed effects slightly increases the coefficient but does not change its significance (b = 0.051, p = .065). Once we remove government effectiveness from the control set, which reveals the covariate's role in explaining the bias presented in Section 4.2, the coefficient becomes statistically significant (b = 0.064, p = .021). This means that one covariate alone explains the unconditional correlation, and Section 5.2 discusses the rationale. The exclusion of the ten observations having the Cook's distance value above the threshold of 4/n = .0315 (Luxembourg, Fiji, Libya, Guinea-Bissau, Central African Republic, Solomon Islands, Antigua and Barbuda, Mauritius, Bosnia and Herzegovina, and Russian Federation) results in a coefficient value that does not differ greatly from the previous one yet does not become statistically significant (b = 0.035, p = .075, N = 117). The observations are classified as potentially influential cases rather than outliers because the removal is driven by regression leverage, not by errors in the data. Diagnostic tests confirm the validity of the base model specification – the Breusch-Pagan test does not reject homoskedasticity (LM = 9.14, p = .104), and the Shapiro-Wilk test does not reject

normality (W = .990, p = .472). Neither test result can be used as proof of the absence of heteroskedasticity or non-normality.

Table 7. Robustness checks (coefficient on the certification measure)

Specification	N	b	SE	p
R1. ISO Main-3 composite (9001+14001+45001)	127	0.037	0.022	.084*
R2. GDP-scaled denominator	127	0.034	0.023	.145
R3. Excluding Germany	126	0.042	0.021	.049*
R4. Region fixed effects	127	0.051	0.028	.065†
R5. Without government effectiveness	127	0.064	0.028	.021*
R6. Excluding influential observations (Cook's D > 4/n)	117	0.035	0.020	.075†

Note: *† p < .10, * p < .05. Baseline for comparison is Table 3, Model 3 (b = 0.042, p = .0503, N = 127). Significance symbols follow the same convention throughout the paper: † p < .10, * p < .05, ** p < .01, *** p < .001.*

As a formal check on temporal sensitivity, Table 8 pools the same-year (2022) and mismatched-year (2024) certification measures into a single model with a year indicator and a certification-by-year interaction term, using standard errors clustered by country and holding 2022 controls fixed for both years. The 2022 coefficient in this pooled specification (b = 0.036, p = .083) is close to the 2024 coefficient it is compared against, and the interaction term (the formal test of whether the two years' slopes differ) is small and not significant (b = −0.004, p = .220, 95% CI [−0.009, 0.002]). Using temporally mismatched certification data remains the wrong design choice in principle, and the same-year point estimate is directionally larger than the mismatched-year estimate from the original 2024-based specification (b = 0.027, p = .293). Still, the difference between the two is not itself statistically distinguishable from sampling noise. Same-year alignment remains the correct default for this analysis. It should not be read as having driven a large, confirmed empirical change relative to the mismatched-year alternative.

Table 8. Pooled same-year vs. mismatched-year test (country-clustered SE, N = 254 country-year observations)

Term	b	Cluster-robust SE	p
ln(1+ISO9001 pm), 2022 (base)	0.036	0.021	.083†
Year = 2024 (indicator)	0.010	0.014	.494
ln(1+ISO9001 pm) × Year 2024 (interaction)	−0.004	0.003	.220

Note. Control variables (log GDP per capita, trade openness, government effectiveness, and log population; all for 2022) have been included in the estimation but are not shown in the table; clustered standard errors are reported. N = 254 observations of country-years refer to 127 countries observed over two years (2022 and 2024) rather than 254 different countries. † p < .10; the interaction term serves as a test of whether the certification-LPI slope varies across the same-year and mismatched-year estimations, and fails to reach statistical significance.

4.7. Türkiye in Comparative Context

Türkiye's 2022 case highlights the disparity between ISO 9001 intensity and the LPI, which inspired this research. An LPI of 3.4 (rank 38 of 139 economies, the official World Bank ranking) for Türkiye goes hand in hand with the ISO 9001 intensity of 97.0 certificates per million (lower than the upper-middle-income group average of 125.7) and 6 ISO 28000 certificates- a low number indeed, but not zero, contrary to most of its income-group peers. Türkiye's government effectiveness (−0.279) is close to the average for its income group.

5. DISCUSSION

5.1. Certification Counts Track Institutional Development, Not Logistics Capability

However, H1 fails to reach any statistically significant level of association, as the full-model coefficient does not rise above b = 0.042 on a five-point scale, despite an 81 percent reduction in the uncontrolled coefficient after controlling for income, trade openness, government effectiveness, and population. Two other robustness checks (Sections 4.6, 5.2) further weaken the statistical significance. In fact, this appears to be a specification-sensitive effect rather than a robust null result, since, as shown in Section 5.2, dropping government effectiveness from the controls restores the statistical significance of ISO 9001 certification stock (b = 0.064, p = .021). At the same time, the latter is highly correlated with the LPI (r = .859). Taking this into consideration, the best conclusion is that ISO 9001 certificate stock, as measured here, is sufficiently closely related to institutional development that no independent relationship with logistics performance is detectable in this sample. This conclusion is consistent with the observation signature predicted by neo-institutionalist theory, but it does not constitute definitive proof that the measure is a proxy for the latter. Indeed, as DiMaggio & Powell noted, organizations adopt formal organizational structures to gain legitimacy rather than efficiency. Furthermore, Guler, Guillén & Macpherson showed that the cross-national diffusion of ISO 9000 certification followed trade and network ties rather than efficiency. Thus, the logic follows: diffusion is institutionally based, and the resulting certificate stock shows little independence from sectoral performance, which this standard supposedly promotes.

Among the sample countries, Italy has the highest number of ISO 9001 certifications (1,596.5 per million) and the 18th position on the LPI list. Singapore has almost half of Italy's density (871.9 per million) but ranks first on the list. Similarly, Bulgaria and Slovenia both have more than 1,000 certifications per million but rank far away from the top 30.

5.2. Confounder or Channel? A Caution About the Preferred Specification

Government effectiveness correlates with certification intensity (0.774) and the Logistics Performance Index (LPI) (0.859). Excluding government effectiveness from the control variables (Table 7, R5) makes the certification variable statistically significant (b = 0.064, p = .021). The finding leads to an unresolved issue due to the cross-sectional nature of the study – the quality of governance may be a confounding factor in the association between certification and LPI since the capacity of the state affects both the certificates and logistics performance, or because of the co-evolution of the certification intensity and accreditation institutions for some reasons. A secondary regression of government effectiveness on certification intensity, controlling for income level, yields a positive but insignificant coefficient (b = 0.116, p = .095, Section 4.2).

A third option complicates both interpretations. The World Bank's Government Effectiveness indicator (Kaufmann et al., 2011) consists of roughly thirty underlying components, most of which are surveys of international business people and experts—a basis which partly overlaps with the LPI sample of logistics professionals. Thus, the two indicators may partially measure the same construct: "how international professionals perceive this country." In that case, including government effectiveness as a control would remove some of the LPI's measurement error and, hence, the 81 percent attenuation described in Section 4.2. This dataset does not have another governance indicator that would be less dependent on business perception surveys; hence, this limitation is mentioned. The natural next step would be to re-estimate Table 3 using such a governance indicator.

The final problem concerns directionality. The signaling and coercion mechanisms introduced above in Section 2.2 rest on the assumption that certification comes first and helps produce performance. Yet, neither statistical evidence nor temporal considerations support such ordering in the data. Statistically, a country with a well-functioning logistics system may

attract more certified multinationals in the first place, reproducing the unconditional M1 correlation through a performance-to-certification causal pathway. Moreover, the measure used to proxy for the extent of certification does not come before the outcome variable it is meant to predict: the survey for LPI 2023 was administered between September and November 2022, while the ISO Survey reports certificate stocks as of 31 December 2022, two to three months after the end of the LPI fieldwork. Thus, the measure of certification intensity is captured not even contemporaneously, but after the fact, with respect to the outcome variable – further making it impossible to conclude the former's causal priority. The study design provides no means to distinguish between the certification-to-performance and performance-to-certification causal pathways, or to rule out simultaneous operation of both. Signaling and coercion mechanisms should therefore be understood as a theoretically possible story that fits the data, rather than as a mechanism inferred from the data.

The relationship examined in this paper between certification intensity and logistics performance is only correlational. There are three possible interpretations of the data presented here: (a) government effectiveness confounds the relationship, (b) there is an unidentified relationship between certification intensity and governance capacity that moves both variables simultaneously, or (c) performance drives some part of the relationship with certification intensity. The current study does not use mediation analysis, instrumental variable techniques, or longitudinal data to arbitrate between these alternative explanations. The language above does not support any conclusions about such methodological approaches. The one bound we can draw from the evidence presented is that the full model coefficient ($b = 0.042$; p-value > 0.1) is the most conservative estimate of the relationship after controlling for the confounds examined in this study; even the second most conservative estimate in Model 2 ($b = 0.092$) is still considerably lower than the unconditional relationship (Model 1; $b = 0.225$). None of these three interpretations provides a reason to consider the number of national certifications a lever for boosting logistics performance.

5.3. Where the Residual Variation Sits

Once again, the component analysis yields results that seem to contradict the paper's main hypothesis, and the general pattern narrows further after adjustment for multiple testing. The sub-component labeled logistics competence and quality, which is supposed to represent the signal for firm-level process quality, demonstrates one of the weakest correlations among all components examined, both before and after correction. Of the three sub-components having the smallest p-values, only one – international shipments (defined as competitiveness of the shipments) – remains significant after Holm adjustment (Table 4; adjusted $p = .041$), while timeliness and tracking and tracing sub-components are not. A possible explanation of such a correlation could be the fact that ISO 9001 certification works in the international logistics industry not as a signal about better service quality, but rather as a transactional credential, used as a signal by shippers and freight forwarders when assessing unknown counterparts in their international business transactions (Terlaak & King, 2006), and as a sign of being a part of an integrated and formalized supply chain where buyer pressure serves as a tool of upstream diffusion (Corbett, 2006). This interpretation is offered as consistent with, not established by, a single corrected coefficient. As the H1 result it accompanies, it is subject to the same certificate-volatility caveat developed in Section 5.5: if a single ISO Survey edition's counts are a noisy composite of certified practice and reporting behavior, that noise affects the international-shipments estimate exactly as much as it affects every other coefficient in this paper, including the ones the paper treats more skeptically. No component-level result in Table 4 should be read as more measurement-reliable than another merely because it happens to clear a significance threshold.

5.4. A Fragile but Directionally Consistent Income Gradient

The upper- and lower-middle-class income group point estimates exceed the high-income group estimate, but not necessarily in a strictly monotonic fashion. In contrast, a pooled estimate of the non-high-income coefficients is significant under conventional criteria (b = 0.070, p = .043). This pattern may, at first glance, align with findings from Rodriguez-Arnaldo and Martínez-Lorente (2020), who found certification diffusion most related to development in lower-income countries; here, we extend the interpretation from the diffusion process itself to the diffusion-performance relationship. However, one must exercise great caution since the interaction effect coefficient itself is still small and statistically insignificant (b = 0.041, p = .468), while the three-way split of income groups (the most direct test of H3 as formulated) is statistically insignificant in all three subsamples.

5.5. The Sector-Specific Standard, and What Certificate Volatility Reveals About the Measure Itself

The evidence on ISO 28000 adds value to the paper's descriptive contribution; the longitudinal angle adds another dimension that a single-edition analysis would not capture. As a supply chain security management standard last reviewed in 2022, ISO 28000 provided 521 valid certificates globally in 2022, compared with 1,265,200 for the general standard ISO 9001; the same disparity remains in the 2024 edition (322 certificates). In 2022, fifty-seven percent of the sampled economies lacked an ISO 28000 certificate, a figure that varies moderately in subsequent Survey editions.

Another, secondary and unexpected consequence reiterates the article’s general warning about the role of certificates as a measure of performance. First of all, the global numbers of all three generic standards display significant variation in their values between the editions of the ISO Survey of 2022, 2023, and 2024: the number of ISO 9001 decreases by approximately 34 percent from 2022 to 2023 and then rises by roughly 76 percent from 2023 to 2024; the variation for ISO 14001 and ISO 45001 is even greater, namely −43/+125 percent and −53/+193 percent, respectively. According to the ISO Survey’s own documentation, these variations are caused by different participation by reporting certification bodies rather than by any rapid decertification process (ISO, 2024). It would be unreasonable to believe that the stock of certificates decreased by one third and then increased by three quarters in two years; therefore, such fluctuations show that the total numbers for a particular edition reflect only a rather inexact mixture of certified practice and reporting behavior. The second finding supports the article's broader warning from another angle: the number of certificates for each nation is influenced by institutional and administrative factors, such as the completeness of the reporting process in a given year, which have no clear connection to logistics performance. Nevertheless, this variation is not merely corrected. Consistent reporting by a country's certification bodies likely signals its institutional capacity, which supports the paper's H1 finding.

5.6. Implications

In studies of management standards, these results run counter to using country-level certificate counts from any particular survey edition as proxies for the diffusion of management practice. Volatility of reporting of year-to-year counts, discussed in Section 5.5, presents a second reason to be skeptical about using cross-country counts from any particular edition in addition to the marker-of-institutional-capacity argument in Section 5.1. Research estimating the performance consequences of certification should use within-country variation, multiple-edition averaging, firm-level performance metrics, or methods that allow direct use of the certification adoption date.

This represents an association pattern, not a comparison of the effectiveness of policy implementation in both programs: the analysis does not compare them, and it does not claim

that shifting the program budget from the certification subsidy to institutional capacity development would yield better results. Rather, the pattern shows that when the budget is limited, one should not immediately regard the subsidy on certification as more relevant to logistics performance because government effectiveness is more strongly associated with LPI in all estimations carried out in the present work. Certification can still provide the market-access advantages described in Section 5.3, regardless of its overall performance.

The persistent difference between ISO 9001 and ISO 28000 across three survey iterations, despite the 2022 revision intended to update this industry standard, raises questions about the ability of an industry-focused tool within a sphere controlled by a general standard. The results of the binary-indicator measure in Section 4.5 show that this standard still has some effectiveness, though not in adoption; rather, adoption seems linked to logistics performance.

5.7. Limitations

Five restrictions apply to these conclusions. First, despite the alignment of the ISO Survey with the LPI by year, certificate stock statistics from the ISO Survey pertain to 31 December 2022, two to three months after the September to November 2022 survey window of the LPI; thus, the certification variable does not have a clear precedence over the outcome in question, and the analysis of the difference between 2022 and 2024 in Section 4.6 can be regarded only as a measurement-year sensitivity check, not as a test of the causal time ordering. Second, ISO Survey participation is voluntary, and as shown in Sections 4.1 and 5.5, it varies enough across editions to significantly affect the country totals; this carries into single-edition estimates, like those provided in this study. Third, the perception-based nature of the government-effectiveness control means that the model cannot separate the effects of income, state capacity, outcome measurement overlap, and certification (Section 5.2). Fourth, the theory rests on certification → performance processes such as signaling and coercion, which cannot be separated from a performance → certification logic in which more effectively governed logistics sectors have more certifications. Fifth, certification affects organizations rather than people, meaning that country population is not the ideal basis for measuring the intensity of certification; GDP-weighting in Table 7 (R2) tries to mitigate the issue, but the economically or internationally active firm population would be preferred; unfortunately, it is not available for this analysis at the country level.

6. CONCLUSION

This study asked whether national intensity of ISO management system certification is associated with national logistics performance, using the LPI 2023 edition, the ISO Survey 2022 edition (chosen to match the LPI's reference year), and World Bank development and governance indicators for 127 economies.

The unadjusted association is quite strong, but the adjusted association is not significant, and two other tests tend to disconfirm rather than confirm it. First, the density of ISO 9001 on its own explains 61 percent of cross-national variance in the LPI; yet, the coefficient falls by 81 percent once income, openness, population, and government effectiveness are controlled for, leaving the residual association ($b = 0.042$, $p = .0503$), which is formally indistinguishable from the otherwise equal specification with mismatched years. The residual may partly reflect overlap in measurement construction between the government-effectiveness control and the LPI outcome (Section 5.2). In addition, the result is sensitive to specification: exclusion of government effectiveness from the list of controls recovers significance ($b = 0.064$, $p = .021$, Table 7, R5). Hence, the result should be regarded as specification-sensitive rather than robustly zero.

The test of no effect (H1), cluster effect (H2), and income group interaction effect (H3)—the three hypotheses that can be tested in a single statistical test—are not supported by the corresponding test of the full model: the coefficient of the full model is not statistically

significant (H1); the predicted clustering effect is absent, as the effect holds only for international shipments and the Holm-corrected coefficients for domestic shipments and for the pooled shipment type (all shipments) are insignificant; and the income interaction is insignificant, with the effect in non-high-income subsample (H3) being a post-hoc result. Only the continuous ISO 28000 specification supports H4, but the binary ISO 28000 indicator remains significant (Section 4.5).

The descriptive results emerge as one of the most consistent patterns across the estimations and survey-iteration analyses. Activity related to certification in quality management is more likely to follow a generic standard than a sector-specific one: 1,265,200 ISO 9001 certifications versus 521 ISO 28000 certifications in 2022. This gap does not narrow during the observation period (it persists in 2024 and changes little after the single revision of the standard). Because the number of ISO 9001 certificates is economy-wide and not specific to the logistics sector (Section 1), this result is regarded as evidence of the described decoupling in quality management practices; however, it is not considered direct evidence of decoupling in the logistics sector. Moreover, a second, unexpected finding further highlights the problems with certificate numbers: world totals for the three generic standards under consideration demonstrate significant changes from one survey iteration to another: they decrease by about 34 to 53 percent between 2022 and 2023, and increase by about 76 to 193 percent between 2023 and 2024.

National certificate numbers, available in the year closest to the outcome in question, should be treated as a noisy, participation-sensitive index of institutional development rather than a performance indicator. They correlate highly with the same national characteristics, which indicate the level of institutional development, do not contribute once institutional development is measured directly, and vary enough between survey iterations to make cross-country comparisons for any single year impossible. Interpreting institutional prevalence at the national level as evidence of managerial practice or a reason to adopt the policy could substitute an unstable institutional symptom for the determinant of performance.

DATA AND CODE AVAILABILITY

We compiled and estimated the merged country-level dataset (LPI 2023, ISO Survey 2022, and 2022 World Development Indicators/Worldwide Governance Indicators) in Python 3.12, along with the scripts used to generate every table and figure in this work. Data merging and transformations were performed using pandas (2.x), and all OLS, HC3-robust, cluster-corrected standard errors, and Holm-corrected regressions were estimated with statsmodels (0.14). Cook's distance and variance inflation factors are based on statsmodels outlier and collinearity tests. Figures were made using matplotlib. The merged dataset and scripts are available upon request from the corresponding author.

REFERENCES

Arvis, J.-F., Ojala, L., Shepherd, B., Ulybina, D., & Wiederer, C. (2023). *Connecting to compete 2023: Trade logistics in an uncertain global economy — The Logistics Performance Index and its indicators.* World Bank.

Boiral, O. (2003). ISO 9000: Outside the iron cage. Organization Science, 14(6), 720–737. https://doi.org/10.1287/orsc.14.6.720.24873

Boiral, O. (2007). Corporate greening through ISO 14001: A rational myth? Organization Science, 18(1), 127–146. https://doi.org/10.1287/orsc.1060.0224

Bromley, P., & Powell, W. W. (2012). From smoke and mirrors to walking the talk: Decoupling in the contemporary world. Academy of Management Annals, 6(1), 483–530. https://doi.org/10.1080/19416520.2012.684462

Çelebi, D. (2019). The role of logistics performance in promoting trade. Maritime Economics & Logistics, 21(3), 307–323. https://doi.org/10.1057/s41278-017-0094-4

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. Journal of Management, 37(1), 39–67. https://doi.org/10.1177/0149206310388419

Corbett, C. J. (2006). Global diffusion of ISO 9000 certification through supply chains. Manufacturing & Service Operations Management, 8(4), 330–350. https://doi.org/10.1287/msom.1060.0120

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147–160. https://doi.org/10.2307/2095101

Djankov, S., Freund, C., & Pham, C. S. (2010). Trading on time. The Review of Economics and Statistics, 92(1), 166–173. https://doi.org/10.1162/rest.2009.11498

Guler, I., Guillén, M. F., & Macpherson, J. M. (2002). Global competition, institutions, and the diffusion of organizational practices: The international spread of ISO 9000 quality certificates. Administrative Science Quarterly, 47(2), 207–232. https://doi.org/10.2307/3094804

Hausman, W. H., Lee, H. L., & Subramanian, U. (2013). The impact of logistics performance on trade. Production and Operations Management, 22(2), 236–252. https://doi.org/10.1111/j.1937-5956.2011.01312.x

Heras-Saizarbitoria, I., & Boiral, O. (2013). ISO 9001 and ISO 14001: Towards a research agenda on management system standards. International Journal of Management Reviews, 15(1), 47–65. https://doi.org/10.1111/j.1468-2370.2012.00334.x

International Organization for Standardization. (2023). The ISO Survey of management system standard certifications 2022. https://www.iso.org/the-iso-survey.html

International Organization for Standardization. (2024). The ISO Survey of management system standard certifications 2024. https://www.iso.org/the-iso-survey.html

Jayathilaka, R., Jayawardhana, C., Embogama, N., Jayasooriya, S., Karunarathna, N., Gamage, T., & Kuruppu, N. (2022). Gross domestic product and logistics performance index drive the world trade: A study based on all continents. PLOS ONE, 17(3), Article e0264474. https://doi.org/10.1371/journal.pone.0264474

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The worldwide governance indicators: Methodology and analytical issues. Hague Journal on the Rule of Law, 3(2), 220–246. https://doi.org/10.1017/S1876404511200046

Levine, D. I., & Toffel, M. W. (2010). Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers. Management Science, 56(6), 978–996. https://doi.org/10.1287/mnsc.1100.1159

Martí, L., Puertas, R., & García, L. (2014). The importance of the Logistics Performance Index in international trade. Applied Economics, 46(24), 2982–2992. https://doi.org/10.1080/00036846.2014.916394

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340–363. https://doi.org/10.1086/226550

Mirtsch, M., Pohlisch, J., & Blind, K. (2025). Certification as a compensation mechanism for weak regulation? Exploring the diffusion of the international standard ISO/IEC 27001 for information security management. Computers & Security, 162, Article 104774. https://doi.org/10.1016/j.cose.2025.104774

Rezaei, J., van Roekel, W. S., & Tavasszy, L. (2018). Measuring the relative importance of the logistics performance index indicators using Best Worst Method. Transport Policy, 68, 158–169. https://doi.org/10.1016/j.tranpol.2018.05.007

Rodriguez-Arnaldo, O., & Martínez-Lorente, A. R. (2020). What determinants influence the diffusion of ISO 9001 by countries? The TQM Journal, 33(1), 223–246. https://doi.org/10.1108/TQM-03-2020-0055

Sampaio, P., Saraiva, P., & Rodrigues, A. G. (2009). ISO 9001 certification research: Questions, answers and approaches. International Journal of Quality & Reliability Management, 26(1), 38–58. https://doi.org/10.1108/02656710910924161

Spence, M. (1973). Job market signaling. The Quarterly Journal of Economics, 87(3), 355–374. https://doi.org/10.2307/1882010

Terlaak, A., & King, A. A. (2006). The effect of certification with the ISO 9000 Quality Management Standard: A signaling approach. Journal of Economic Behavior & Organization, 60(4), 579–602. https://doi.org/10.1016/j.jebo.2004.09.012

World Bank. (2024). World Development Indicators and Worldwide Governance Indicators [Data sets]. https://data.worldbank.org/