Built Environment Project and Asset Management

How performance-based are China's public-private partnership road projects? A KPI framework and scoring method

Xiaoli Shi
Chang'an University, Xi'an, China
Yongjian Ke
University of Technology Sydney, Sydney, Australia
Xuelian Wu
Chongqing Rail Transit Design and Research Institute Co., Ltd., Chongqing, China
Xiaotian Gong
PowerChina Roadbridge Group Co., Ltd., Beijing, China, and
Jingxiao Zhang
Chang'an University, Xi'an, China

Received 21 August 2025
Revised 10 December 2025
27 February 2026
Accepted 18 May 2026

Abstract

Purpose – This study aims to develop a practical method for evaluating how Performance-Based Contracting (PBC) features are embedded in the design of Public-Private Partnership (PPP) road contracts in China. It addresses a gap in understanding how performance expectations are structured, specified, and enforced during contract formulation.

Design/methodology/approach – A qualitative document analysis was conducted using all 26 road PPP projects recorded in China's National PPP Project Information System. A unified three-tier Key Performance Indicator (KPI) framework was developed through iterative coding and consolidation of contract content. A structured scoring method was then created to evaluate contracts across three dimensions: breadth (coverage of performance domains), depth (granularity of indicators), and clarity and enforceability (precision and legal operability). The approach was applied to a representative case to demonstrate its practical utility.

Findings – The study presents two core outputs: a benchmark KPI framework and a conceptual scoring method aligned with PBC features. These tools enable systematic and transparent evaluation of how well PPP contracts articulate and support performance expectations. The case application illustrates how contractual provisions can be assessed for structural coverage, operational detail, and enforceability.

Originality/value – This paper contributes a replicable framework for diagnosing performance system design in infrastructure contracts. It advances the literature by translating PBC features into measurable contract features and provides practical tools for cross-project comparison, contract template improvement, and policy alignment. The approach supports more consistent integration of performance logic into PPP governance.

Keywords Public-private partnerships, Performance-based contracting, Key performance indicators, Contract evaluation, Infrastructure governance

Paper type Research article

1. Introduction

Public-Private Partnerships (PPPs) in China refer to long-term contractual collaborations between the government and private capital for developing infrastructure and delivering

Funding: This research was supported by the China Scholarship Council (202306560107).

Built Environment Project and Asset Management
Emerald Publishing Limited
e-ISSN: 2044-1258
p-ISSN: 2044-124X
DOI 10.1108/BEPAM-08-2025-0235

public services (Ke *et al.*, 2024). The evolution of China's PPPs has spanned through four distinct stages (Cheng *et al.*, 2020, 2024a). From 1984 to 2002, the central government prioritised attracting foreign investment, primarily through Build-Operate-Transfer (BOT) model. Between 2003 and 2013, local governments and state-owned enterprises became the dominant players in implementation. A period of rapid expansion followed between 2014 and 2023, often described as the "PPP fever." However, the pace and scale of growth raised concerns over project quality and financial discipline, leading to a formal pause in new project approvals. In November 2023, a new PPP mechanism was launched, introducing stricter entry criteria and greater emphasis on performance-based delivery. This renewed policy emphasis on service outcomes and long-term value, which places stricter demands on PPP contract design and highlights the need for systematic, performance-oriented evaluation tools.

Although policy has clearly shifted toward performance-based approaches, significant challenges remain in translating this principle into executable contracts. Performance-Based Contracting (PBC) provides a key theoretical framework for this transformation. PBC emphasizes linking payments to measurable outcomes rather than prescribed inputs or processes, featuring core design elements such as clear outcome specifications, robust performance measurement, and performance-linked payments (Selviaridis and Wynstra, 2015). This approach aligns closely with the underlying logic of recent PPP policy in China.

While performance specification, monitoring, and the link between payments and service delivery are widely recognised as central to PPPs (Cao *et al.*, 2024; Ke *et al.*, 2024; Su *et al.*, 2023), research on performance management often remains disconnected from contract design practice. Although existing studies have examined output specifications (Lam and Javed, 2015), objectives and indicators (Yuan *et al.*, 2009), and measurement systems (Liu *et al.*, 2018; Wang *et al.*, 2023), there has been limited progress in understanding how performance systems are structured and embedded in PPP contracts (Ma *et al.*, 2023). As a result, it remains unclear whether these contracts reflect key features such as outcome orientation, measurability, enforceability, and incentive alignment. The absence of a structured assessment tool makes it difficult to verify whether contemporary contract practices have genuinely transitioned from "process compliance" to "output orientation," and limits evidence-based improvements to contract templates and policy evaluation. This exposes a key methodological gap between China's performance-oriented PPP policy and the actual design of PPP contracts: how can a PPP contract be objectively and systematically evaluated as a "well-designed" performance-based agreement? To address this gap, this study formulates the following core research question: How can a systematic diagnostic tool be developed and applied to assess the degree to which the textual design of road PPP contracts in China embodies the core features of PBC?"

To fill this gap, this study develops a reusable framework and scoring method to translate the abstract principles of PBC into quantifiable diagnostics for road PPP contracts. It addresses a critical gap by enabling policymakers, agencies, and investors to assess whether contract design effectively embeds performance expectations, thereby improving contract quality, value for money, and risk-informed decision-making. The study's innovation lies in creating the first systematic, document-based tool that operationalizes PBC features into measurable dimensions, constructs a full-sample benchmark KPI framework from 26 Chinese road PPP contracts, and links macro-level policy intent with micro-level contract evaluation, offering a transparent and scalable approach for research, practice, and policy assessment.

The paper is structured as follows. Section 2 reviews the conceptual foundations of PBC features, their application in infrastructure sectors, and how performance management has been addressed in PPP research. Section 3 outlines the research methodology. Section 4 presents the KPI framework and scoring method developed from the sample. Section 5 applies these tools to a representative case. Section 6 discusses the findings, draws implications, and concludes with suggestions for future research.

2. Literature review

A systematic search using terms such as "Performance-Based Contracting", "outcome-based contracts", "pay-for-performance", "infrastructure PPPs" and "performance measurement systems" was conducted. Included studies focused on the design of contract performance provisions, including outcomes, KPIs, service standards, and performance-linked incentives, within infrastructure projects, particularly road PPPs. Studies addressing only contract execution, behavioural dynamics, or post-contract performance were excluded. This strategy ensured a focused literature base and highlighted gaps in methods for systematically evaluating how PBC principles are embedded in contract design.

2.1 Performance-based contracting (PBC)

PBC is a procurement approach that links payment to the achievement of measurable outcomes rather than compliance with prescribed inputs or processes (Selviaridis and Wynstra, 2015). Although terminology varies across sectors, related concepts such as "outcome-based contracts" (Ng *et al.*, 2009), "pay for performance" (Tzani *et al.*, 2022), and "performance-based logistics" (Aytekin and Korucuk, 2024) share a common principle: supplier compensation is tied to the delivery of verified results. Despite contextual differences, these approaches reflect a shift from activity-based service delivery to performance-focused accountability. Rather than a fixed model, PBC is better understood as a design philosophy that emphasises aligning objectives, transferring performance risk, and incentivising innovation without prescribing how services must be delivered (Glas and Kleemann, 2017). As such, PBC provides a conceptual lens for examining whether contracts are designed in a way that enables accountability, encourages results, and balances control and autonomy. In this study, PBC is not treated as a predefined standard but as a set of core features that can be used to evaluate the structure and intent of performance provisions in contracts.

The literature identifies a broad range of features that collectively define the institutional logic of PBC, encompassing outcome specification, performance measurement, payment conditionality, risk allocation, supplier autonomy, transparency, and relational governance. Together, these features articulate a shift in contract philosophy from prescriptive input control to results-oriented accountability. Yet, from the perspective of contract theory, these features operate at different stages of the contractual relationship. Elements such as risk allocation, supplier discretion, and relational mechanisms primarily influence the behavioural and managerial dimensions of contract execution, reflecting the governance capacity of PBC in practice. By contrast, a subset of PBC features directly shapes how the performance logic is codified at the contract formation stage. Three design-oriented dimensions are particularly central. Outcome specification defines the intended service results rather than prescribed inputs, providing the normative anchor for accountability and value creation (Selviaridis and Wynstra, 2015; Hypko *et al.*, 2010). . Performance measurement systems operationalize these expectations through structured, verifiable indicators that ensure traceability between objectives and observed outcomes (Glas and Kleemann, 2017; Glas *et al.*, 2018). Payment conditionality links performance to financial consequences, creating incentive alignment and reinforcing compliance through measurable contingencies (Ng *et al.*, 2009).

Collectively, these three features, i.e. outcome specification, performance measurement, and payment conditionality, constitute the design-oriented core of PBC. They represent the contractual mechanisms through which abstract performance principles are institutionalized in enforceable provisions. Accordingly, this study focuses on these features to examine how performance expectations are encoded within the textual structure of PPP contracts.

PBC has been widely adopted across a range of public and private service sectors, including health and social services (Petersen *et al.*, 2006), defense and aerospace (Ng *et al.*, 2009), and education and workforce development (Lu, 2016). These diverse applications demonstrate the adaptability of the PBC approach, which allows contracts to be tailored around sector-specific goals and operational conditions. Although PBC is well established in these domains, its application in infrastructure delivery, particularly through PPPs, presents additional challenges and design considerations. The next section explores how PBC thinking has been interpreted and implemented in infrastructure projects, with a focus on the road sector.

2.2 PBC in infrastructure projects

The application of PBC in infrastructure has been discussed in a limited but growing body of literature, particularly in the context of road and bridge maintenance. These studies explicitly adopt terms such as PBC, performance-based maintenance contracting (PBMC), or output-based contracts, and focus primarily on asset preservation and service quality assurance in long-term maintenance arrangements (Sultana *et al.*, 2013; Wirahadikusumah *et al.*, 2015; Tseng and Yang, 2024). Much of this work centres on roads, where standardised service outcomes, such as pavement condition, ride quality, or lane availability, are relatively well suited to performance-based models. Studies from Indonesia, Australia, and China, e.g. Wirahadikusumah *et al.* (2015), Wu *et al.* (2025), highlight the value of aligning payment with verifiable service outcomes but caution that successful implementation depends on clear metrics, institutional capacity, and adaptive contract design.

Beyond these explicitly PBC-labelled studies, a much larger body of literature addresses performance requirements in infrastructure projects without using PBC terminology. For instance, Flyvbjerg *et al.* (2003) and Odeck (2004) document widespread cost overruns in road infrastructure and attribute them partly to weak performance forecasting and lack of accountability mechanisms. Chan *et al.* (2004) and Zhao *et al.* (2010) explore success factors in construction and energy projects, emphasising the role of well-defined quality standards and measurable outputs. Others propose tools such as fuzzy logic systems (Poveda and Fayek, 2009) or stakeholder-aligned metrics (Doloi, 2013) to support continuous performance monitoring and adaptive decision-making. Although these studies do not refer to PBC, their focus on performance outcomes and measurement systems is consistent with performance-based logic.

A comparison of the two strands of literature reveals important complementarities and gaps. The first group provides direct conceptual and implementation models of PBC in infrastructure, but is largely confined to maintenance-focused projects with limited empirical scope. The second group shows that performance management is a widespread concern in infrastructure delivery, yet it rarely addresses how performance expectations are embedded at the contract level. Together, the two bodies of work suggest that while performance-based thinking is relevant and increasingly applied, there is no established method for systematically evaluating how it is operationalised through infrastructure contract design. This gap reinforces the value of applying a structured PBC lens to the analysis of contractual performance systems.

The next section builds on this foundation by reviewing how performance management has been studied specifically in the context of PPPs, where contracts are central to long-term service delivery.

2.3 Structuring performance in PPP contracts

In PPPs, contract design plays a central role in defining, monitoring, and enforcing long-term performance obligations. Recent studies have converged on three interrelated instruments used to structure these obligations: output specifications, KPIs, and service standards (Javed *et al.*, 2013). Output specifications describe the desired outcomes of service delivery, such as availability, safety, or user satisfaction, rather than prescribing how services should be delivered (Lam and Javed, 2015). KPIs translate these outputs into quantifiable measures that can be tracked throughout the contract period (Yuan *et al.*, 2009). Service standards define the acceptable levels or thresholds for each KPI, distinguishing compliance from underperformance (Robinson and Scott, 2009). Together, these three instruments form a layered performance system: output specifications guide the strategic intent, KPIs provide measurable metrics, and service standards establish enforceable benchmarks. While this study focuses specifically on the design and benchmarking of KPI systems, these are understood as the operational core of a broader performance framework structured through all three instruments.

The enforceability of performance expectations depends on how well contractual provisions define consequences for underperformance. Several studies describe how contracts operationalise enforcement through financial and procedural mechanisms. These include penalty deductions, performance-linked payments (Robinson and Scott, 2009), step-in rights, and in some cases, termination clauses (Opawole, 2018; Song *et al.*, 2018). However, enforceability is not always straightforward. Germani and Zeller (2021) point out that ambiguous drafting and uneven monitoring capacity often limit the effectiveness of formal penalties. Furthermore, practical enforcement may be constrained by political considerations, especially when the private operator is delivering essential public services (Galilea and Medda, 2010). The literature suggests that enforceability is strongest when response mechanisms are clearly defined, proportionate, and supported by objective performance data (Liu *et al.*, 2018).

Designing performance expectations into contracts is only the first step. Effective implementation also depends on how performance is measured, reported, and managed throughout the project lifecycle (Liu *et al.*, 2018). Several studies emphasise that performance monitoring acts as the bridge between contractual expectations and enforcement (Wang *et al.*, 2023). Despite Robinson and Scott (2009) describe a reporting system and Liu *et al.* (2018) develop a process management life cycle performance measurement system, the literature reports persistent challenges such as disputes over data accuracy, limited oversight capacity, and vague enforcement triggers (Liu *et al.*, 2014, 2015). As noted by Koppenjan *et al.* (2022), the practical effectiveness of performance systems depends not only on their design but also on how clearly roles, responsibilities, and response procedures are defined in the contract.

2.4 Gap in knowledge

Despite increasing policy emphasis on performance-based delivery, a persistent gap remains in the literature: there is no established method to evaluate whether PPP contracts are designed to reflect the core features of PBC. Existing research provides limited insight into how outcome expectations, performance indicators, and enforcement mechanisms are jointly structured within contractual performance systems, leaving unclear how performance logic is embedded in long-term service delivery. This gap is especially salient in China, where recent reforms have clarified policy-level performance expectations, yet contractual provisions still vary widely in structure and clarity. Addressing this gap, this study distinguishes itself by examining the contract design phase of PBC rather than focusing solely on implementation or outcomes. It introduces a systematic, document-based framework that operationalizes core PBC principles—outcome orientation, measurable performance, and incentive alignment—allowing objective evaluation of how performance expectations are embedded in contracts. By bridging explicit PBC literature with broader performance management research, it extends the theoretical lens to a wider set of infrastructure contracts. This approach addresses a key methodological gap, challenges assumptions that contract effectiveness can be inferred without design analysis, and enables large-scale, comparable assessments.

3. Methodology

3.1 Research philosophy and design

This study adopts a pragmatist research philosophy, reflecting its aim to develop a practical method for evaluating how PBC features are embedded in PPP contract design. Pragmatism is suited to problem-driven research that priorities useful outcomes over strict adherence to a single epistemological tradition (Saunders *et al.*, 2019). Given that the data source of this study consists of complex, unstructured legal contract texts, and the objective is to uncover their underlying performance logic and structural patterns, qualitative content analysis was identified as the most appropriate research method.

A qualitative document analysis is employed to examine how performance expectations are structured within contractually defined indicator systems. PPP contracts are treated as formal artefacts that encode outcome objectives, measurement arrangements, and enforcement conditions. Document analysis is particularly appropriate for interpreting meaning, structure, and intent embedded in written materials (Bowen, 2009). The research is based on a complete sample of 26 PPP road projects currently recorded in the National PPP Project Information System, established in 2024. This ensures comprehensive coverage of current PPP practice in the road sector.

The unit of analysis is the performance indicator system within each contract. The study follows an inductive–deductive process: a three-level KPI framework is developed inductively from the observed contract content, followed by the conceptual development of a scoring method to assess the presence of key PBC features.

To establish a robust foundation, a systematic literature review was conducted before analyzing PPP contracts, focusing on scholarship at the intersection of PBC theory, infrastructure project performance, and PPP contract design. The search was performed in Web of Science and Scopus, supplemented by the China National Knowledge Infrastructure (CNKI) database to include relevant Chinese-language studies. The following search string was used for English databases (2000–early 2025):

(“performance-based contracting” OR “outcome-based contract” OR “pay for performance” OR “performance-based logistics”) AND (“public-private partnership” OR PPP OR “infrastructure project” OR “road project”) AND (“performance indicator” OR KPI OR “key performance indicator” OR “contract design” OR “performance measurement” OR “service standard”)

A corresponding Chinese search string was applied in CNKI. Titles and abstracts were screened to include studies on the design of performance provisions in infrastructure or PPP contracts, while excluding those focused solely on post-contract execution, behavioural dynamics, or pure project finance. Snowballing from reference lists further identified relevant literature. This systematic process ensured a comprehensive conceptual basis for developing the study’s analytical framework.

3.2 Data collection

Data for this study were obtained from the National PPP Project Information System, an official database established by the National Development and Reform Commission in 2024. The dataset includes all 26 PPP road projects recorded as of March 2025 and verified in June 2025 (see Appendix Table A1). All projects are toll roads under the BOT model, including two Rehabilitate-Operate-Transfer (ROT) projects consistent with the BOT framework. The contracts and performance annexes are publicly available, ensuring comprehensive and unbiased coverage of current PPP practices in China’s road transport sector..

For each project, the contract document and annexes were reviewed to identify performance-related provisions. Contract content was extracted, cleaned, and segmented into structured records for analysis. Common formatting issues and variations in terminology were resolved through manual review to ensure consistency. This prepared dateset formed the basis for the development of the KPI framework and the subsequent scoring method, as detailed in the next two sections.

3.3 Development process of the KPI framework

The KPI framework was developed through a two-step process that combined inductive content analysis with deductive validation. Drawing on all 26 national PPP road contracts, original performance indicators were systematically coded to reflect real-world industry practice. Structural validity was ensured through iterative coding following the mutually exclusive and collectively exhaustive (MECE) principle, producing a logically distinct and comprehensive framework. The empirically derived structure was then refined using policy guidelines (Caijin [2020] No. 13) and PBC literature, ensuring alignment with policy intent and theoretical foundations.

Step 1: Extraction and Normalisation.

All available KPIs were extracted from each contract document, including annexes or attachments where indicator tables were presented. Some projects employed a multi-level indicator structure, typically including primary categories, subcategories, and specific indicators. These three levels were disaggregated and standardised to enable cross-project comparison. Where necessary, variations in terminology were harmonised to reflect consistent meaning. This step also involved identifying equivalent indicators presented under different names and normalising their structure and function.

Step 2: Categorisation and Framework Construction.

Following extraction, indicators were grouped into a hierarchical framework based on shared functional meaning. The top-level categories were aligned with China's official PPP performance guidance (Caijin [2020] No. 13), which requires evaluation across three dimensions: output, outcome, and management. A fourth dimension, i.e. incentive-based indicators, was added to reflect features specific to PBC, such as granted autonomy and outcome-driven innovation. Within each top-level category, subcategories and third-level indicators were developed to capture specific operational, technical, and public value domains observed in the contracts.

The outcome of this process is a benchmark KPI framework that consolidates indicator systems from all reviewed contracts into a unified three-level architecture. It does not assign weights or interpret performance quality but provides a neutral basis for evaluating how contractually defined indicators are organised and represented. Importantly, the framework is intended to remain adaptive and will be refined as more PPP road projects are added to the national database. This framework forms the foundation for the scoring method developed in the next section.

3.4 Development process of the scoring method

To evaluate the extent to which contractually defined performance systems embody key features of PBC, a scoring method was developed based on the benchmark KPI framework. The method enables structured and comparative assessment across projects, focusing on how performance expectations are defined, organised, and operationalised within each contract. Its development combined theoretical deduction with empirical calibration using real contract data. The core framework, comprising breadth, depth, clarity, and enforceability, was derived from the fundamental principles of PBC theory on sound contract design. Building on this conceptual foundation, specific scoring rules were iteratively tested and inductively refined through analysis of actual contract texts, ensuring that the theoretical dimensions were translated into objective, transparent, and consistently applicable measurement criteria.

Step 1: Definition of Scoring Dimensions

The first step involved identifying relevant aspects of contract design that reflect PBC features. Drawing on the architecture of the KPI framework and literature on PBC, three scoring dimensions were defined: breadth of KPI coverage, depth of indicator detail, and clarity and enforceability of performance definitions. Each dimension was selected to capture a specific attribute of how performance is structured within the contract. Breadth reflects the scope of performance areas addressed. Depth captures the granularity and detail of performance expectations. Clarity and enforceability relate to the definitional strength and operational utility of individual indicators.

Step 2: Development of Measurement Rules

Each dimension was then translated into a measurable rule to ensure consistency across projects. Breadth was measured as the proportion of sub-categories (Level 2 indicators) from the benchmark framework that appear in each contract. Depth was assessed as the proportion of third-level indicators (Level 3) used under each included category. Clarity and enforceability were operationalised through a four-level classification of each indicator: vague or incomplete, defined, measurable, and enforceable. These categories were designed to reflect increasing degrees of contractual specificity and alignment with PBC features.

Step 3: Design of the Scoring Format

Finally, a percentage-based scoring format was developed to enable comparability across projects of different size and complexity. Each contract receives an independent score between 0 and 100% for each of the three dimensions. In addition to these individual scores, an overall PBC alignment score is calculated as the average of the three dimensions. The scoring method is diagnostic in nature and intended to support consistent assessment across a diverse set of PPP projects.

The scoring method is conceptual in nature but designed to be transparent, repeatable, and adaptable. It serves as a tool for assessing contract design quality from a performance-based perspective and is applied in Section 5 to a representative case.

4. Results: KPI framework and scoring method

4.1 Benchmark KPI framework

Based on a complete sample of 26 Chinese road PPP contracts, this study develops a three-level KPI framework to represent current practice in structuring contractual performance obligations. The framework adopts a hierarchical architecture comprising four top-level categories (Level 1), 13 subcategories (Level 2), and 42 specific indicators (Level 3), as shown in Table 1. The framework focuses on the operational phase of PPP delivery, where lifecycle costs are most concentrated and performance expectations are most fully articulated in contract documents. In road PPPs, this phase is critical because it governs long-term service delivery and forms the basis for linking payments to outcomes (Cherkos and Jha, 2021).

The four Level-1 categories reflect both national policy guidance and the conceptual foundations of PBC. The first three categories, i.e. output, outcome, and management, are directly drawn from the official PPP performance evaluation guidelines (Caijin [2020] No. 13). The output category includes indicators that capture operational deliverables directly under the private partner’s control, such as service quality, maintenance execution, and safety management. These indicators reflect routine service delivery responsibilities and are essential for ensuring baseline performance. The outcome category encompasses broader public value dimensions, including social and environmental impacts, financial resilience, and satisfaction of government and end users. These indicators move beyond basic service provision to assess the long-term value and externalities generated by the project. The management category covers institutional and procedural aspects such as organizational structure, financial discipline, information disclosure, and record-keeping. It reflects the importance of transparency and internal control in maintaining contract compliance and enabling effective oversight.

The fourth category, incentive, captures contract features that extend beyond government-mandated requirements, reflecting the advanced logic of PBC (Abu Samra *et al.*, 2017). Unlike the first three dimensions, output, outcome, and management that set baseline compliance thresholds, the incentive dimension highlights mechanisms that actively encourage private partners to exceed expectations, fostering flexibility, innovation, and value creation. By rewarding enhanced performance and adaptive service delivery, it aligns with the PBC principle of shifting from input control to outcome-based incentives (Sumo *et al.*, 2016), embodying a forward-looking, value-generating approach to PPP contract design.

Table 1. Three-level KPI framework for road PPP projects

Categories	Sub-categories	Indicators
1 Output	1.1 Project Operation	1.1.1 Highway service quality 1.1.2 Service area service level 1.1.3 Toll service level 1.1.4 Road network information service 1.1.5 Road section operation monitoring and dispatching 1.1.6 Breakdown and rescue service management 1.1.7 Overload management 1.1.8 Asset disposal 1.1.9 Complaint channels 1.1.10 Performance bond 1.1.11 Insurance coverage
	1.2 Project Maintenance	1.2.1 Maintenance plan submission 1.2.2 Maintenance management system 1.2.3 Road condition inspection completion status 1.2.4 Road appearance and cleanliness 1.2.5 Time-limited maintenance measures implementation for traffic safety 1.2.6 Subgrade Condition Index (SCI) 1.2.7 Pavement Quality Index (PQI) 1.2.8 Bridge and Tunnel Structure Condition Index (BCI) 1.2.9 Traffic Facilities Condition Index (TCI) 1.2.10 Highway Maintenance Quality Index (MQI) 1.2.11 Maintenance project completion status 1.2.12 Facility equipment integrity rate
	1.3 Safety Assurance	1.3.1 Safety production management system and responsibility implementation 1.3.2 Emergency management 1.3.3 Number of safety production liability accidents 1.3.4 Safety management of maintenance operations
2 Outcome	2.1 Social Impact	2.1.1 Major litigation, public opinion events, and mass incidents 2.1.2 New employment opportunities created
	2.2 Ecological Impact and Environmental Protection	2.2.1 Energy conservation and emission reduction performance 2.2.2 Environmental protection
	2.3 Sustainability	2.3.1 Daily operation risk resilience 2.3.2 Financial sustainability
	2.4 Satisfaction	2.4.1 Government department satisfaction 2.4.2 Public satisfaction
3 Management	3.1 Organizational Management	3.1.1 Organizational structure and personnel management
	3.2 Financial Management	3.2.1 Financial management system 3.2.2 Fund usage compliance
	3.3 Archives Management	3.3.1 Archives management status
	3.4 Information Disclosure	3.4.1 Information disclosure status
	3.5 Institutional Management	3.5.1 Maintenance management, contract management systemetc
4 Incentive	4.1 Social Honors	4.1.1 The project was commended and awarded honorary titles for its outstanding performance in adopting “Four New” technologies, innovative operation and management models, and safety management, which led to enhanced efficiency and reduced operational costs

Source(s): Authors’ own work

This benchmark framework is not intended as a static model but as a structured representation of current practice. As the National PPP Project Information System continues to expand and more contracts are made available, the framework can evolve to incorporate emerging indicator types, refined definitions, and additional performance dimensions. Its hierarchical framework allows for modular updates without compromising consistency, supporting both longitudinal analysis and adaptation to shifting policy or sectoral priorities. In this way, the framework offers a foundation for ongoing refinement and broader application in future contract evaluations.

4.2 Scoring method

The unified KPI framework introduced in Section 4.1 provides the analytical basis for evaluating the extent to which individual PPP contracts reflect core features of PBC. The scoring method focuses not on actual project outcomes, but on how performance expectations are defined and operationalized within each contract. Each contract is assessed across three dimensions: breadth, depth, and clarity and enforceability.

Breadth refers to the extent to which a contract addresses the range of performance domains outlined in the benchmark framework. It is measured by calculating the proportion of Level-2 subcategories covered in each contract. A higher score indicates broader coverage of key functional areas. For this dimension, a higher score is clearly preferable, as it reflects a more comprehensive performance system that attends to a wider set of operational, outcome-based, and management-related responsibilities. This breadth ensures that performance monitoring extends across all critical aspects of service delivery, rather than concentrating on a narrow subset of priorities, thereby supporting balanced contract execution and more holistic oversight.

Depth captures the granularity of indicator systems, as measured by the proportion of Level-3 indicators present under the performance domains that a contract includes. This dimension reflects how fully each addressed performance area is articulated through specific, actionable indicators. A higher score often signals greater operational clarity, but it does not necessarily imply better contract design. Excessive detail concentrated in a few subcategories may suggest over-engineering, while moderate but evenly distributed depth, particularly when combined with broad coverage, may indicate a more coherent and balanced structure. To avoid penalising contracts for domains they do not cover, the depth score is calculated only across Level-2 subcategories actually present in the contract. This approach ensures that the depth dimension remains focused on indicator refinement, independent of breadth. Moreover, as the benchmark framework evolves with additional project inputs, the reference framework itself may shift. This implies that there may be an optimal range for depth scores, and values significantly above this range could signal unnecessary complexity rather than improved performance definition. Therefore, depth should be interpreted in relation to both coverage balance and benchmark evolution.

Clarity and enforceability refer to how precisely each performance indicator is defined and the extent to which it is supported by measurable targets and enforceable mechanisms. This dimension captures the operational utility of Level-3 indicators, i.e. whether they are vague or actionable, and whether they provide a credible basis for monitoring and enforcement. Each present indicator is assessed using a four-tier classification (see Table 2): (1) vague or incomplete (0.0), (2) defined only (0.5), (3) defined and measurable (0.75), and (4) defined, measurable, and enforceable (1.0). For each contract, the clarity score is calculated as the total points assigned to all present Level-3 indicators divided by the maximum possible points (i.e. the number of such indicators multiplied by 1.0), then converted to a percentage. This approach ensures that the score reflects the quality of definition among indicators that are actually used in the contract, rather than penalising for absence.

Together, the three scores provide a multidimensional perspective on how performance-based China's PPP road projects are. Although the three scores serve primarily diagnostic purposes, an overall PBC alignment score calculated as the average can support simplified benchmarking across projects. This aggregate score may be helpful for policy analysis, cross-project comparison, and identifying contracts where alignment with PBC features is limited or

Table 2. Four-tier scale for clarity and enforceability

Score	Criteria	Description
1.0	Defined + Measurable + Enforceable	The indicator is clearly worded, includes specific targets or thresholds, and is explicitly linked to penalties, bonuses, or performance reviews
0.75	Defined + Measurable	The indicator is well-defined and quantifiable, but no enforcement mechanism is specified
0.5	Defined only	The indicator is specific but lacks a measurable target or enforcement mechanism
0.0	Vague or Incomplete	The indicator is too general, lacks measurable targets, or is poorly formulated

Source(s): Authors' own work

uneven. Practitioners can refer to the algorithm flowchart in Figure 1 to conduct the aforementioned analysis.

5. Case study: application of the scoring method

To demonstrate the practical application of the KPI framework and scoring methodology developed in Section 4, the G208 project is presented as an illustrative case following the process outlined in Figure 1. The example illustrates how a real PPP contract can be systematically transformed into structured, quantifiable evaluation results, validating the method's feasibility and providing a clear practical reference for readers.

5.1 Case overview

The G208 Reconstruction Project in Mengjin District, Luoyang City, is a toll road PPP implemented under a BOT model. The concession period is 34 years, comprising a 4-year preparatory and construction phase followed by a 30-year operational phase.

The concession agreement, included in the project's bidding documents, consists of 15 chapters and 90 thematic clauses. Chapter 7, titled "Operations and Services," sets out 12 clauses addressing general requirements for operations, maintenance, and service provision. Chapter 9, "Specific Regulatory Matters and Measures," specifies that an operational performance evaluation must be conducted annually on a rolling 12-month basis.

The operational performance indicators that underpin this evaluation are detailed in Appendix Table A1 of the concession agreement. For the purposes of this study, these indicators have been extracted and reorganised as Table 3, which serves as the basis for applying the scoring method presented in Section 4.2.

5.2 Contract assessment results

Table 4 presents the mapping between the performance indicators specified in the G208 project and the corresponding benchmark indicators in Table 1. Each matched benchmark indicator is identified using its hierarchical code. While the wording of indicators in the case differs from the terminology used in the benchmark, interpretive alignment was achieved using the "Evaluation Criteria" and "Scoring Methodology" outlined in Table 3. These two columns guided the mapping of case indicators by assessing their intent, structure, and level of specificity, ensuring consistent correspondence with the benchmark.

The breadth analysis identifies 9 out of 13 Level-2 subcategories from the benchmark framework as being represented in the G208 project, resulting in a breadth coverage score of 69%. While some correspondences, such as those related to financial management, do not fully align with the benchmark's depth or specificity, they reflect the presence of relevant

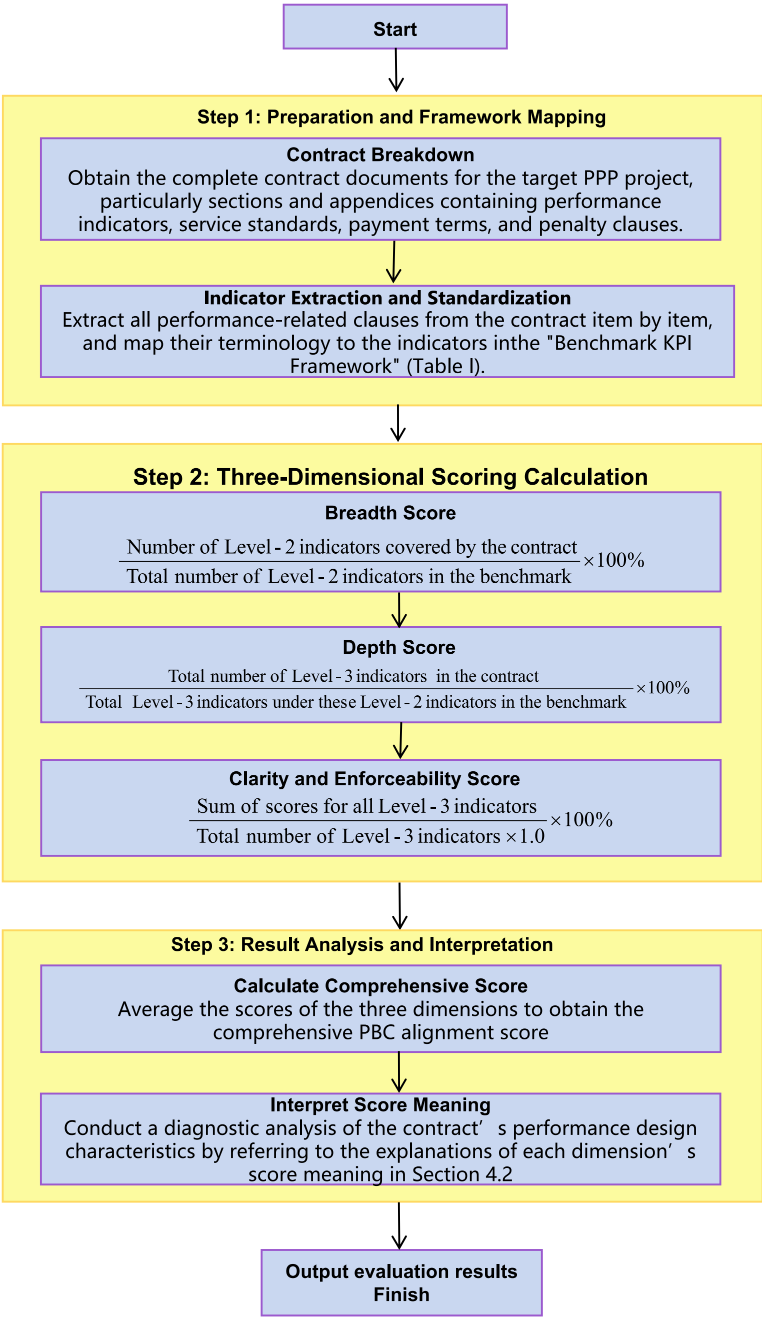

Figure 1. Implementation flowchart of the scoring process. Source: Authors' own work

domains and are therefore counted in the breadth dimension. According to the benchmark in Table 1, the G208 project lacks indicators in four subcategories: sustainability, organisational management, institutional management, and incentive mechanisms. However, relevant

Table 3. Operational performance indicators for the G208 reconstruction project

Categories	Sub-categories	Indicators	Score	Evaluation criteria	Scoring methodology
Output	Project Operation	Road Network Information Services	5	Timely and accurate release of travel and road condition information	3 points awarded for information releases that are timely, accurate, and compliant with regulatory requirements. Each instance of untimely or inaccurate release results in a deduction of 1 point, up to a maximum deduction of 5 points
		Station Area Management	5	1. Clean and tidy toll plaza environment without debris or garbage; facilities intact, clean, and undamaged. 2. Normal operation of toll lane equipment and facilities	1. 0.5 points deducted per instance of unclean plaza, debris, damaged safety islands, zebra crossing paint peeling, canopy damage, or improper barrier placement (max 2 points) 2. 1 point deducted per unresolved equipment malfunction or network outage (max 3 points)
		Toll Management	10	1. Implementation of "Five Disclosures" (toll standards, approving authority, supervision hotline, social commitments, toll operator) 2. Standardized toll records management and strict adherence to approved charging standards	1. 1 point deducted per missing disclosure item (max 2 points) 2. 3 points deducted for irregular record-keeping 3. 5 points deducted for non-compliance with approved standards
	Project Maintenance	Maintenance System	10	1. Annual/monthly maintenance planning and implementation 2. Emergency response plans for disasters/special weather 3. Complete inspection records with signatures 4. Timely post-rain inspections and repairs 5. 24-h issuance of repair orders for identified defects	1. 1 point deducted per non-compliance item (max 2 points) 2. 1 point deducted per non-compliance (max 2 points) 3. 1 point deducted per incomplete/unverified record (max 2 points) 4. 1 point deducted per non-compliance (max 2 points) 5. 2 points deducted per failure (max 2 points)
		Maintenance Quality	25	Comprehensive evaluation of road maintenance quality meeting contractual requirements, with rectification plans for deficiencies	Full score awarded when Pavement Quality Index (PQI), Subgrade Condition Index (SCI), Bridge/Tunnel Condition Index (BCI), and Traffic Facilities Index (TCI) all ≥80. Otherwise, no points awarded
		Maintenance Inspections	14	1. Smooth pavement without defects 2. Intact bridge facilities 3. Well-maintained roadside vegetation 4. Complete ancillary facilities	1. 0.5 points deducted per pavement defect (max 5) 2. 0.5 points deducted per bridge defect (max 5) 3. 0.5 points deducted per vegetation issue (max 1) 4. 0.5 points deducted per facility defect (max 3)
	Safety Assurance	Safety Management System	4	1. Established safety organization with dedicated personnel 2. Comprehensive safety regulations	1. 1 point deducted per missing organizational element (max 2) 2. 1 point deducted per regulatory deficiency (max 2)
		Accident Handling	8	1. Compliance with accident reporting requirements 2. Proper investigation and reporting of major incidents	1. 2–4 points deducted for reporting violations 2. 1–4 points deducted for investigation/reporting failures
		Emergency Plans	4	1. Comprehensive emergency plans with resource inventories 2. Regular emergency drills	1. 1 point deducted per planning deficiency (max 2) 2. 2 points deducted for drill non-compliance

(*continued*)

Table 3. Continued

Categories	Sub-categories	Indicators	Score	Evaluation criteria	Scoring methodology
Outcome	Ecological Impact	Energy Conservation	1	1. Adoption of energy-efficient lighting 2. ≥80% asphalt recycling rate	0.5 points deducted per unmet criterion
		Environmental Penalties	2	Compliance with environmental regulations	1 point deducted per penalty received (max 2)
	Social Impact	Public Sentiment	2	1. Effective complaint resolution 2. Avoidance of media exposure 3. Prevention of major incidents	Full deduction for major exposures; partial deductions for minor issues
		Travel Assurance	2	Implementation of traffic maintenance plans	1 point deducted per significant traffic disruption
	Satisfaction	Public Satisfaction	3	Public satisfaction survey results	1 point base score; 0.1 points deducted per 1% below 80% satisfaction
Management	Regulatory Compliance	Tax Compliance	2	Timely tax filing	Full deduction for non-compliance
	Records Management	Archiving System	1	1. Complete archival system 2. Dedicated personnel 3. Proper document control	0.5 points deducted per deficiency (max 1)
	Information Disclosure	Timely Disclosure	2	Policy-compliant information sharing	Full deduction for non-compliance

Source(s): Authors' own work

Table 4. Mapping of G208 project indicators to benchmark KPI framework

Categories	Sub-categories G208	Benchmark	Indicators G208	Benchmark
Output	Project Operation	1.1	Road Network Information Services	1.1.4
			Station Area Management	1.1.2
			Toll Management	1.1.3
	Project Maintenance	1.2	Maintenance System	1.2.1;1.2.3;1.3.2
			Maintenance Quality	1.2.6;1.2.7;1.2.8;1.2.9
			Maintenance Inspections	1.2.4
	Safety Assurance	1.3	Safety Management System	1.3.1
			Accident Handling	1.3.3
			Emergency Plans	1.3.2
Outcome	Ecological Impact	2.2	Energy Conservation	2.2.1
			Environmental Penalties	2.2.2
	Social Impact	2.1	Public Sentiment	2.1.1
			Travel Assurance	2.1.1
	Satisfaction	2.4	Public Satisfaction	2.4.2
Management	Regulatory Compliance	3.2	Tax Compliance	3.2.2
	Records Management	3.3	Archiving System	3.3.1
	Information Disclosure	3.4	Timely Disclosure	3.4.1

Source(s): Authors' own work

clauses within the concession agreement address these domains implicitly. For instance, Article 42 ("Quality and Service Standards") and Article 50 ("Party B's Innovation and Proactive Efficiency Enhancement") stipulate requirements related to institutional systems and innovation practices. These provisions underscore the importance of broader performance domains, even when they are not explicitly itemised within the formal evaluation criteria.

The G208 project covers 20 Level-3 indicators across the 37 indicators located within the nine Level-2 subcategories it addresses, resulting in a depth score of 54%. This score reflects the extent to which performance expectations are specified through detailed and actionable indicators in the domains actually covered by the contract. Several indicators are directly defined, including those related to toll services (1.1.3), maintenance quality (1.2.6–1.2.9), and emergency management (1.3.2). Notably, in sub-category 1.1 (Project Operation), only 3 out of 11 benchmark indicators are present, indicating a shortcoming in how operational performance is detailed. While the formal indicator system appears limited, additional contract clauses implicitly align with benchmark indicators. For instance, Articles 19, 42, 44, 45, and 47 correspond to indicators such as insurance coverage (1.1.11), asset disposal (1.1.8), and maintenance quality standards (1.2.10). These provisions suggest a deeper performance logic embedded within the contract text, even when not presented as standalone indicators.

The G208 project demonstrates a high level of clarity and enforceability across its performance indicators. All 20 Level-3 indicators identified are clearly defined, with explicit assessment objectives, measurable thresholds, and associated penalty rules, as outlined in Table 3. For instance, full marks are granted for PQI, SCI, BCI, and TCI scores of 80 or above, while scores below this threshold receive no credit. In addition, Article 56 specifies evaluation methods such as document review, on-site inspection, and stakeholder feedback, and links overall assessment results to financial penalties ranging from 500,000 to 2 million yuan per point below defined thresholds. Contracts falling below a score of 60 for two consecutive years may be terminated. Based on these provisions, the clarity and enforceability score is calculated as: $(20 \times 1)/20 \times 100 = 100\%$. This result indicates that all performance indicators are unambiguous, measurable, and backed by enforceable mechanisms, reflecting strong alignment with PBC features and a proactive approach to performance-based management.

5.3 Interpretation and reflections

Taken together, the G208 project's breadth score (69%), depth score (54%), and clarity and enforceability score (100%) yield an overall PBC Alignment Score of 74%. This reflects a reasonably strong alignment with PBC features. The project demonstrates wide coverage of performance domains and a high degree of indicator clarity and enforceability. However, the moderate depth score points to opportunities for enhancing the detail and granularity of performance expectations across all subcategories. Notably, many performance indicators are not explicitly listed within the contract's formal KPI framework but are embedded across various clauses. While these dispersed references support a more comprehensive performance logic, integrating such indicators directly into the KPI system would improve transparency, facilitate systematic monitoring, and strengthen contractual accountability.

The assessment of the G208 project highlights several important opportunities to strengthen the design of its performance evaluation system. While the contract demonstrates clear alignment with PBC features in terms of indicator clarity and enforceability, other areas merit targeted improvement. For instance, the operational subcategory (1.1) includes only 3 of 11 benchmark indicators, suggesting an under-specified performance framework in one of the most critical domains. More generally, the moderate depth score (54%) reflects limited granularity across several subcategories, which may constrain the precision of performance monitoring and weaken the guidance provided for service delivery.

A further limitation lies in the disconnection between formal KPI listings and performance-related clauses embedded elsewhere in the contract. The analysis revealed that many requirements aligned with benchmark indicators were present in the text but not explicitly included in the KPI framework. Had these qualitative clauses been quantified and incorporated into the formal indicator system, the breadth score would have risen from 69% to 85%, and the depth score to 78%. This finding reinforces a practical recommendation: integrating such clauses directly into the KPI framework would improve transparency, enhance systematic monitoring, and strengthen contractual accountability.

6. Conclusion

This study developed two key tools to evaluate how PBC features are embedded within road PPP contracts: a benchmark KPI framework and a structured scoring method. The framework offers a standardised, three-tiered architecture for organising performance expectations across four functional categories, while the scoring method provides a replicable approach to assessing the breadth, depth, and enforceability of performance provisions. Together, these outputs enable systematic analysis of contract design alignment with PBC features. The application of the method to the G208 project illustrated its practical utility, demonstrating how the framework and scoring criteria can be used to diagnose strengths and identify design gaps in real contracts.

This study contributes to the body of knowledge by creating the first systematic, document-based, and quantitative tool to diagnose how PBC features are embedded in Chinese PPP road contracts, bridging a gap between abstract policy principles and contract-level evaluation. The resulting KPI framework provides a practical diagnostic tool for key stakeholders involved in PPP road projects. For practitioners and contract drafters, it serves as a benchmark checklist for evaluating and comparing contracts, identifying gaps (e.g. missing sustainability metrics) and enabling structured self-assessment through scoring. For instance, a low breadth score indicates incomplete coverage, while high clarity with low depth suggests well-defined but operationally simplistic indicators. Benchmarking across contracts helps identify industry-wide weaknesses and supports establishing minimum design standards. For researchers, the framework provides a consistent coding scheme for large-scale comparative studies of PPP contracts across sectors or time, with the scoring method supplying replicable variables for empirical analysis.

This study also recognises its limitations. The benchmark KPI framework is based on a specific dataset of 26 PPP road projects and reflects the current stage of contract evolution in China. As additional projects enter the national database, the reference framework will need to be revised to capture emerging practices and new performance areas. Similarly, the scoring method focuses on design provisions rather than actual implementation or outcomes. Moreover, future research should implement the proposed evaluation method across all the projects in the project database to examine the status of PBC in PPP roads and inform the refinement of PPP contract performance design in China.

Appendix

Table A1. Basic information of 26 projects

No	Project name	Mileage length (km)	Location (province)	Operation mode	Concession period (year)
1	G4 Beijing-Hong Kong-Macao Expressway Yanglousi (Xiang-E boundary) to Yueyang Longwan Section Expansion Project Bundle G4 Beijing-Hong Kong-Macao Expressway Yueyang Longwan-Changsha Guangfu Section Expansion Project	150.099	Hunan	ROT	33
2	G4 Beijing-Hong Kong-Macao Expressway Zhuzhou Wangshiwan (Zhuting)-Leiyang Dashi Section Expansion Project Bundles Guidong to Xintian (Ningyuan) Expressway Chenzhou-Chenzhou section	189.236	Hunan	ROT/ BOT	35
3	Yan ’an East Ring Expressway	57.16	Shaanxi	BOT	33
4	Changsha-Anhua Section of Changji Adjustment Highway	123.208	Hunan	BOT	32
5	Wushan to Guandu Section of Xuanhan-Kaizhou-Yunyang-Wuxi-Wushan Expressway	18.125	Chongqing	BOT	34
6	S35 Jingtai-Jingyuan Section of Li County Expressway	144.29	Gansu	BOT	35
7	Beiliu-Huazhou Baowei Expressway (Guangxi Section)	75.641	Guangxi	BOT	33
8	Mengjin Boundary Reconstruction Project of Luoyang City of G208 Second Xi Line (including newly built Yellow River Highway and Railway Bridge Lead)	20.979	Henan	BOT	34
9	G341 Huangdao-Haiyan Highway Zhongchuan-Heqiao Section and S01 Lanzhou Economic Circle Loop Heqiao-Haishiwan Section	167.074	Gansu	BOT	35
10	G571 Subei-Ruoqiang Highway Aksai to Lapei Spring (Ganxinjie) Section	230.629	Gansu	BOT	34
11	S02 Xuzhou to Shangqiu Expressway Suzhou Section	74.168	Jiangsu	BOT	33

(*continued*)

Table A1. Continued

No	Project name	Mileage length (km)	Location (province)	Operation mode	Concession period (year)
12	S20 Changfeng-Gushi Expressway Huoqiu-Anhui-Henan Boundary Section Project	60.832	Anhui	BOT	33
13	S24 Hangzhou-Hefei Expressway Xuancheng Section Phase I Project	84	Anhui	BOT	34.5
14	S40 Tongling Section of Ningguo-Zongyang Expressway	19.725	Anhui	BOT	34
15	Highway Project from Jinan Interchange to Aiweiergou to Alehui Town of Line S103	111.8449	Xinjiang	BOT	33
16	Chaideng-Tuk Class I Highway	100.1	Inner Mongolia	BOT	33
17	Highway Engineering of Xiaozaikou to Damiao Section of National Highway 210	20.647	Inner Mongolia	BOT	33
18	Tiantangzhai Branch of He County to Xiangyang Expressway	45.7	Anhui	BOT	33
19	Wuhan-Huangmei Expressway (Wuhan-Jinchun Section)	109.061	Hubei	BOT	28
20	Song County Section of Yongcheng-Lingbao Expressway (Luoyang Territory)	73.672	Henan	BOT	34
21	S20 Jiayuguan-Ruoqiang Expressway Jiayuguan-Aksai Section	344.377	Gansu	BOT	35
22	S10 Zhangye to Mazong Mountain Expressway Zhangjiu Junction to Mazong Mountain Section, S44 Jinta to Yumen Highway and S63 Mazong Mountain to Liuyuan Highway	578.526	Gansu	BOT	34
23	Beidaihe New District to Beijing-Qinhuangdao Expressway Section of Qinhuangdao-Shenyang Expressway (Beidaihe New District Branch of Beijing-Qinhuangdao Expressway)	60.517	Hebei	BOT	28
24	Qi County to Lishi Expressway Project	95.995	Shanxi	BOT	30.25
25	Tai Lau Pagoda to Kam Kip Expressway and Baomui Link	87.4	Shaanxi	BOT	32.917
26	Qianjiang Channel and Connection Project North Connection Project	11.415	Zhejiang	ROT	28

Source(s): Authors' own work

References

Abu Samra, S., Osman, H. and Hosny, O. (2017), "Optimal maintenance and rehabilitation policies for performance-based road maintenance contracts", *Journal of Performance of Constructed Facilities*, Vol. 31 No. 1, 04016075, doi: 10.1061/(asce)cf.1943-5509.0000928.

Aytekin, A. and Korucuk, S. (2024), "Evaluating performance-based logistics in manufacturing through polytopic fuzzy SWARA: a criterion assessment approach", *Journal of Engineering Management and Systems Engineering*, Vol. 3 No. 2, pp. 65-71, doi: 10.56578/jemse030201.

Bowen, G.A. (2009), "Document analysis as a qualitative research method", *Qualitative Research Journal*, Vol. 9 No. 2, pp. 27-40, doi: 10.3316/qrj0902027.

Cao, Y., Li, H. and Su, L. (2024), "A dynamic performance-based payment mechanism for public-private partnership projects: an integrated model for principal-agent and multi-objective optimization models", *International Journal of Strategic Property Management*, Vol. 28 No. 2, pp. 116-129, doi: 10.3846/ijspm.2024.21183.

Chan, A.P., Scott, D. and Chan, A.P. (2004), "Factors affecting the success of a construction project", *Journal of Construction Engineering and Management*, Vol. 130 No. 1, pp. 153-155, doi: 10.1061/(asce)0733-9364(2004)130:1(153).

Cheng, Z., Ke, Y., Yang, Z., Cai, J. and Wang, H. (2020), "Diversification or convergence: an international comparison of PPP policy and management between the UK, India, and China", *Engineering, Construction and Architectural Management*, Vol. 27 No. 6, pp. 1315-1335, doi: 10.1108/ecam-06-2019-0290.

Cheng, Z., Li, Y., Liu, T. and Wang, H. (2024a), "Public private partnership with Chinese characteristics: a critical review and theoretical framework", *Public Works Management and Policy*, Vol. 29 No. 4, pp. 518-538, doi: 10.1177/1087724x231203373.

Cherkos, F.D. and Jha, K.N. (2021), "Drivers of road sector public-private partnership adoption in new and inexperienced markets", *Journal of Construction Engineering and Management*, Vol. 147 No. 3, 04020186, doi: 10.1061/(asce)co.1943-7862.0002004.

Doloi, H. (2013), "Cost overruns and failure in project management: understanding the roles of key stakeholders in construction projects", *Journal of Construction Engineering and Management*, Vol. 139 No. 3, pp. 267-279, doi: 10.1061/(asce)co.1943-7862.0000621.

Flyvbjerg, B., Skamris Holm, M.K. and Buhl, S.L. (2003), "How common and how large are cost overruns in transport infrastructure projects?", *Transport Reviews*, Vol. 23 No. 1, pp. 71-88, doi: 10.1080/01441640309904.

Galilea, P. and Medda, F. (2010), "Does the political and economic context influence the success of a transport project? An analysis of transport public-private partnerships", *Research in Transportation Economics*, Vol. 30 No. 1, pp. 102-109, doi: 10.1016/j.retrec.2010.10.011.

Germani, A. and Zeller, K. (2021), "Concessionaire's performance assessment in public-private partnerships", *European Procurement and Public Private Partnership Law Review*, Vol. 16 No. 2, pp. 125-130, doi: 10.21552/epppl/2021/2/6.

Glas, A.H. and Kleemann, F.C. (2017), "Performance-based contracting: contextual factors and the degree of buyer supplier integration", *Journal of Business and Industrial Marketing*, Vol. 32 No. 5, pp. 677-692, doi: 10.1108/jbim-04-2016-0065.

Glas, A.H., Henne, F.U. and Essig, M. (2018), "Missing performance management and measurement aspects in performance-based contracting: a systematic process-based literature analysis of an astonishing research gap", *International Journal of Operations and Production Management*, Vol. 38 No. 11, pp. 2062-2095, doi: 10.1108/ijopm-09-2016-0571.

Hypko, P., Tilebein, M. and Gleich, R. (2010), "Benefits and uncertainties of performance-based contracting in manufacturing industries: an agency theory perspective", *Journal of Service Management*, Vol. 21 No. 4, pp. 460-489, doi: 10.1108/09564231011066114.

Javed, A.A., Lam, P.T. and Zou, P.X. (2013), "Output-based specifications for PPP projects: lessons for facilities management from Australia", *Journal of Facilities Management*, Vol. 11 No. 1, pp. 5-30, doi: 10.1108/14725961311301448.

Ke, Y., Cheng, Z., Zhang, J. and Liu, Y. (2024), "Making sense of the definition of public-private partnerships", *Built Environment Project and Asset Management*, Vol. 14 No. 1, pp. 4-21, doi: 10.1108/bepam-01-2023-0009.

Koppenjan, J., Klijn, E.-H., Verweij, S., Duijn, M., van Meerkerk, I., Metselaar, S. and Warsen, R. (2022), "The performance of public–private partnerships: an evaluation of 15 years DBFM in Dutch infrastructure governance", *Public Performance and Management Review*, Vol. 45 No. 5, pp. 998-1028, doi: 10.1080/15309576.2022.2062399.

Lam, P.T. and Javed, A.A. (2015), "Comparative study on the use of output specifications for Australian and UK PPP/PFI projects", *Journal of Performance of Constructed Facilities*, Vol. 29 No. 2, 04014061, doi: 10.1061/(asce)cf.1943-5509.0000554.

Liu, J., Love, P.E., Smith, J., Regan, M. and Sutrisna, M. (2014), "Public-private partnerships: a review of theory and practice of performance measurement", *International Journal of Productivity and Performance Management*, Vol. 63 No. 4, pp. 499-512, doi: 10.1108/ijppm-09-2013-0154.

Liu, J., Love, P.E., Smith, J., Regan, M. and Palaneeswaran, E. (2015), “Review of performance measurement: implications for public–private partnerships”, *Built Environment Project and Asset Management*, Vol. 5 No. 1, pp. 35-51, doi: 10.1108/bepam-12-2013-0070.

Liu, J., Love, P.E.D., Smith, J., Irani, Z., Hajli, N. and Sing, M.C.P. (2018), “From design to operations: a process management life-cycle performance measurement system for Public-Private Partnerships”, *Production Planning and Control*, Vol. 29 No. 1, pp. 68-83, doi: 10.1080/09537287.2017.1382740.

Lu, J. (2016), “The performance of performance-based contracting in human services: a quasi-experiment”, *Journal of Public Administration Research and Theory*, Vol. 26 No. 2, pp. 277-293, doi: 10.1093/jopart/muv002.

Ma, L., Hu, Y., Zhu, L. and Ke, Y. (2023), “Are public-private partnerships still an answer for social infrastructure? A systematic literature review”, *Frontiers of Engineering Management*, Vol. 10 No. 3, pp. 467-482, doi: 10.1007/s42524-023-0249-1.

Ng, I.C., Maull, R. and Yip, N. (2009), “Outcome-based contracts as a driver for systems thinking and service-dominant logic in service science: evidence from the defence industry”, *European Management Journal*, Vol. 27 No. 6, pp. 377-387, doi: 10.1016/j.emj.2009.05.002.

Odeck, J. (2004), “Cost overruns in road construction–what are their sizes and determinants?”, *Transport Policy*, Vol. 11 No. 1, pp. 43-53, doi: 10.1016/s0967-070x(03)00017-9.

Opawole, A. (2018), “Penalty mechanisms for enforcing concessionaire performance in public–private partnership contracts in Nigeria”, *Journal of Engineering, Design and Technology*, Vol. 16 No. 1, pp. 161-191, doi: 10.1108/jedt-04-2017-0030.

Petersen, L.A., Woodard, L.D., Urech, T., Daw, C. and Sookanan, S. (2006), “Does pay-for-performance improve the quality of health care?”, *Annals of Internal Medicine*, Vol. 145 No. 4, pp. 265-272, doi: 10.7326/0003-4819-145-4-200608150-00006.

Poveda, C.A. and Fayek, A.R. (2009), “Predicting and evaluating construction trades foremen performance: fuzzy logic approach”, *Journal of Construction Engineering and Management*, Vol. 135 No. 9, pp. 920-929, doi: 10.1061/(asce)co.1943-7862.0000061.

Robinson, H.S. and Scott, J. (2009), “Service delivery and performance monitoring in PFI/PPP projects”, *Construction Management and Economics*, Vol. 27 No. 2, pp. 181-197, doi: 10.1080/01446190802614163.

Saunders, M.N.K., Lewis, P. and Thornhill, A. (2019), *Research Methods for Business Students*, 8th ed., Pearson, Harlow.

Selviaridis, K. and Wynstra, F. (2015), “Performance-based contracting: a literature review and future research directions”, *International Journal of Production Research*, Vol. 53 No. 12, pp. 3505-3540, doi: 10.1080/00207543.2014.978031.

Song, J., Hu, Y. and Feng, Z. (2018), “Factors influencing early termination of PPP projects in China”, *Journal of Management in Engineering*, Vol. 34 No. 1, 05017008, doi: 10.1061/(asce)me.1943-5479.0000572.

Su, L., Cao, Y. and Li, H. (2023), “Performance-based payment structural design for infrastructure PPP projects”, *International Journal of Strategic Property Management*, Vol. 27 No. 2, pp. 133-145, doi: 10.3846/ijspm.2023.19180.

Sultana, M., Rahman, A. and Chowdhury, S. (2013), “A review of performance based maintenance of road infrastructure by contracting”, *International Journal of Productivity and Performance Management*, Vol. 62 No. 3, pp. 276-292, doi: 10.1108/17410401311309186.

Sumo, R., van der Valk, W., Duysters, G. and van Weele, A. (2016), “Using performance-based contracts to foster innovation in outsourced service delivery”, *Industrial Marketing Management*, Vol. 59, pp. 12-24, doi: 10.1016/j.indmarman.2016.05.029.

Tseng, C.-C. and Yang, J.-B. (2024), “The benefit evaluation framework of using performance-based contracting in urban road maintenance”, *Journal of Civil Engineering and Management*, Vol. 30 No. 2, pp. 133-148, doi: 10.3846/jcem.2024.20491.

Tzani, D., Stavrakas, V., Santini, M., Thomas, S., Rosenow, J. and Flamos, A. (2022), “Pioneering a performance-based future for energy efficiency: lessons learnt from a comparative review analysis of pay-for-performance programmes", *Renewable and Sustainable Energy Reviews*, Vol. 158, 112162, doi: 10.1016/j.rser.2022.112162.

Wang, D., Wang, X., Wang, L., Liu, H.J. and Jia, X. (2023), "A performance measurement system for public–private partnerships: integrating stakeholder influence and process trans-period effect", *International Journal of Productivity and Performance Management*, Vol. 72 No. 1, pp. 137-155, doi: 10.1108/ijppm-08-2020-0408.

Wirahadikusumah, R., Susanti, B., Coffey, V. and Adighibe, C. (2015), "Performance-based contracting for roads–experiences of Australia and Indonesia", *Procedia Engineering*, Vol. 125, pp. 5-11, doi: 10.1016/j.proeng.2015.11.002.

Wu, X., Shi, X., Gong, X., Wang, X. and Chen, W. (2025), "Methodologies for highway maintenance cost-effectiveness: a review", *Journal of Traffic and Transportation Engineering*, Vol. 12 No. 5, pp. 1508-1531, doi: 10.1016/j.jtte.2024.04.011.

Yuan, J., Zeng, A.Y., Skibniewski, M.J. and Li, Q. (2009), "Selection of performance objectives and key performance indicators in public–private partnership projects to achieve value for money", *Construction Management and Economics*, Vol. 27 No. 3, pp. 253-270, doi: 10.1080/01446190902748705.

Zhao, Z.-Y., Zuo, J., Zillante, G. and Wang, X.-W. (2010), "Critical success factors for BOT electric power projects in China: thermal power versus wind power", *Renewable Energy*, Vol. 35 No. 6, pp. 1283-1291, doi: 10.1016/j.renene.2009.09.016.

Further reading

Alsharqawi, M., Abu Dabous, S. and Zayed, T. (2024), "Designing a fuzzy-based framework for implementing performance-based contracts in bridge asset management", *Innovative Infrastructure Solutions*, Vol. 9 No. 4, p. 126, doi: 10.1007/s41062-024-01437-1.

Cai, J., Lin, J., Yang, Z., Zhou, X. and Cheng, Z. (2021), "Retro or Renewal: an assessment of PPP management and policy in China since 2014", *Public Works Management and Policy*, Vol. 26 No. 4, pp. 359-380, doi: 10.1177/1087724x20970955.

Cheng, Z., Zhu, Y., Wang, H. and Ke, Y. (2024b), "Developing a suitability assessment model for public private partnerships: a case in urban China", *Humanities and Social Sciences Communications*, Vol. 11 No. 1, p. 1215, doi: 10.1057/s41599-024-03733-8.

Selviaridis, K. and Norrman, A. (2014), "Performance-based contracting in service supply chains: a service provider risk perspective", *Supply Chain Management: International Journal*, Vol. 19 No. 2, pp. 153-172, doi: 10.1108/scm-06-2013-0216.

Corresponding author
Yongjian Ke can be contacted at: yongjian.ke@uts.edu.au