Author/s:
Geng, L;Herath, N;Hui, FKP;Liu, X;Duffield, C;Zhang, L

Title:
Evaluating uncertainties to deliver enhanced service performance in education PPPs: a hierarchical reliability framework

Date:
2023-11-27

Citation:
Geng, L., Herath, N., Hui, F. K. P., Liu, X., Duffield, C. & Zhang, L. (2023). Evaluating uncertainties to deliver enhanced service performance in education PPPs: a hierarchical reliability framework. Engineering Construction and Architectural Management, 30 (9), pp.4464-4485. https://doi.org/10.1108/ECAM-10-2021-0896.

Evaluating Uncertainties to Deliver Enhanced Service Performance in PPP Schools: A Hierarchical Reliability Framework

Abstract

Purpose – This study aims to develop a hierarchical reliability framework for evaluating the service delivery performance of PPP schools during the long-term operations in an effective and efficient manner. The proposed framework is intended to be used by the government or education evaluators to measure and predict service performance when enough historical data is available.

Design/methodology/approach –The research design includes development and test phases. In development phase, firstly, three layers*, i.e.,* indicator, component, and system, in school service delivery are identified; then component reliability is computed through Advanced First Order Reliability Method (AFORM); Finally, the reliability of system is obtained by consideration of component weightings. In test phase, a PPP school example in Australia was set up, where the indicators were collected from contract documents and performance data were simulated under three assumptive scenarios. The scenario analysis used here is to reflect real-world situation that service performance could be affected by uncertainties and thus present different situations.

Findings–the example in the test phase shows good feasibility of the developed framework in evaluating uncertainties of service delivery performance for PPP schools. Also, it could identify the potentially underperforming services from the component level to system level at different timepoints, thus providing effective maintenance strategies.

Originality/value –The evaluation of education PPPs is challenging as it involves long-term measurement of various service components. This becomes further complicated due to multiple uncertainties. This study, thus, is valuable of the developed framework that could be a tool to ensure that reliability is maintained throughout the service life of education PPPs in the presence of uncertainty. This is the first attempt to quantitatively capture performance uncertainty in PPP school which assists decision-makers in developing efficient preventative maintenance strategies during long-term operations. Besides, dynamic weighting approach is preferred in this evaluation when considering time impact on the importance of different components.

Keywords Service delivery evaluation; Reliability; School PPPs; Dynamic weighting.
Paper type Research paper

1. Introduction

Public-private-partnerships (PPPs) have been broadly considered as both an innovative and effective way to achieve better value for money in infrastructure projects (Atmo et al. 2017; Duffield et al. 2020). Particularly, it lays more emphasize on the education infrastructure and services in Australia, with all focus on providing the framing structure through which to bring the public and private sectors together to complement each other's strengths in financing and provision of education services (Hueskes et al. 2017). Examples of these include but not limited to the New Schools PPP Project (2017), WA Schools PPP (2017), Partnerships Victoria in Schools Project (2008), New Schools Project I and II (2003, 2007). It is acknowledged as a preferred mechanism for school developments due to various benefits, such as better risk allocation (Chen et al. 2017; Wang et al. 2019), operation innovations (Carbonara and Pellegrino 2020; Lember et al. 2019; Saeed et al. 2018), value creation (Aarseth et al. 2016; Villani et al. 2017), and the whole-of-life considerations (Mohd-Rahim et al. 2018). The education PPPs involve lots of elementary and secondary school assets and service delivery, such as teaching facility, classroom, playground, and cleaning (Atmo et al. 2017; Liu and Wilkinson 2014). It plays a siginificant role for governments to improve society's welfare and equity. Consequently, PPPs are expected to be popular education service delivery vehicles in the foreseeable future. For instance, the database of infrastructure projects shows that an amount of $141.5 billion has been invested in 120 PPP projects until 2019 in Australia, and the education sector involved 46 of them (iMC 2020).

To achieve better infrastructure and service outcomes, PPPs have been extended beyond the construction into operation, otherwise known as "operator-led PPPs" (Karasavidis 2018). Education PPPs involve different stakeholders with a range of roles at key decision-making points throughout the long-operation project life. **Figure 1** describes the service delivery process of educational PPP's and its relationship to this research, *i.e.,* Research Focus. The research analysis began when the PPP contract is awarded, and delivery commences. The private sector partner begins to deliver on the contract. Implicit in this is their interpretation

of the contract which will determine what and how the service will be delivered. The government education department, upon awarding the contract will begin to gather key performance indicators which are connected to the service delivery payments (Oyedele 2013; Robinson and Scott 2009). This feeds into education departments evaluation of the service. There can be a gap between service delivered and the required service. If the gap is significant, the government is legal to defer or reduce service payments. The service evaluation process undertaken by the education department is the focus of this research paper. Thereby, both the government and the private sector expect all service can be well delivered (Fulbright 2021). However, the service delivery performance is subject to a mixture of internal and the external uncertainties (Wang et al. 2020; Zhang et al. 2019). The internal uncertainties include but not limited to long-term relationship, multi-stakeholders' participation, and external uncertainty include unstable resource investment, effort (Zhang et al. 2020) and changing environment. So, lots of research on PPPs focused on its performance measurement or evaluation. These have been used to improve and evaluate sets of delivered services combined to PPP contracts. For instance, Saeed et al. (2018) proposed an advanced framework by revising existing KPIs model. The European PPP Expertise Centre examine bunches of ex-ante evaluation practices in operations for PPP projects and how to better demonstrate service outcomes (EPEC 2020). Nevertheless, most measurements are qualitatively rather than quantitatively applied in PPPs or economic PPPs. Few studies capture the uncertainties in the service delivery process in the specific education PPPs.

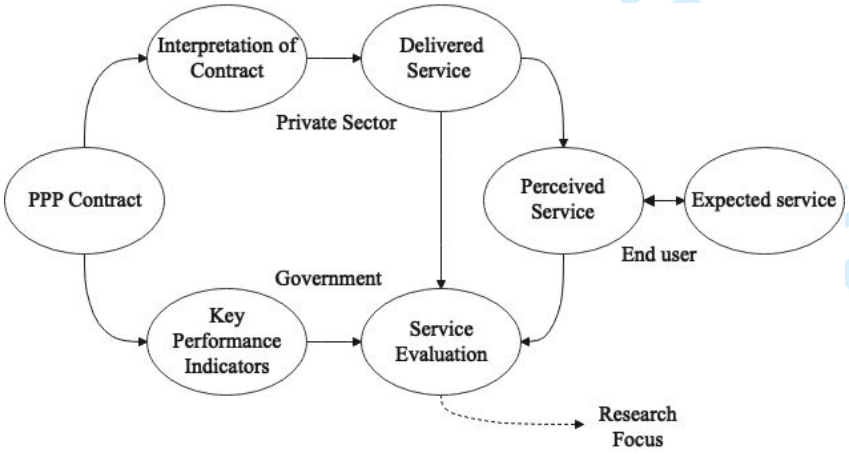

Figure 1 Service Delivery Process in Social PPPs (Geng et al. 2020)

To deal with this problem, a hierarchical evaluation framework combined with reliability theory is proposed. Reliability plays an important role in ensuring the high quality of products, especially in manufacturing industry (Mubarok et al. 2018). Reliability refers to probability of failure of a set of processes or system in producing a desired level of product quality. It helps to deal with uncertainty and make informed decisions (Ugurlu and Kumral 2020). In Maloney (2002) view, public service in construction can typically be considered as a product. The product being delivered to the education department is school buildings and there can also be service components, such as maintenance, cleaning. Thus, service evaluation in education PPPs can be focused on reducing the probability of failure in delivering quality services by ensuring a high reliability of service delivery. Particularly, the act of reliability evaluation for education PPPs service delivery can mitigate unforeseen events in uncertain environment.

As above, most previous research relevant to service performance evaluation in PPPs were qualitative or conceptual. There is a lack of studies using quantitative approaches on such a topic, especially in education PPPs. Therefore, this study is to develop a hierarchical reliability evaluation framework for service performance in education PPPs. And there are two innovative contributions. First, an innovative evaluation framework is developed using the current body of knowledge to evaluate service delivery performance reliability in education PPPs. Under this methodology, the service delivered by the private sector during the long-term operation in education PPPs could be quantitatively evaluated and predicted. Second, the developed framework is demonstrated through a typical PPP school project in Australia under different scenarios, which can be conveniently modelled. This can be used by the government evaluators as an effective and efficient tool to ensure that reliability is maintained throughout the service life of education PPPs in the presence of uncertainty. This will help optimize preventative strategies and performance by the education department.

This paper is structured as follows. **Section 2** presents an overview of education PPPs service measurement research or practices, regarding reliability evaluation. **Section 3** explains the procedure of the service evaluation framework. Then a case example of a PPP school project in Australia is analysed and discussed in **Section 4** to demonstrate its practical use. Finally, this paper closes with a summary and suggestions for future research in **Section 5**.

2. Reliability and Performance Evaluation for Education PPPs: Literature review

In the context of evaluation, reliability is defined as the probability that a component or an entire system will perform its function for a specified period, when operating in its design environment (De Carlo 2013). Reliability has different meanings in different applications. It can exist in a product or service by way of the likelihood of mission success, availability to perform function, dependability, quality over time and the durability (Kececioglu 2002). American Society for Quality (ASQ) is an international professional organization dedicated to advancing research and training in field of service quality. It explains difference between quality and reliability. Quality shows how well an object performs its proper function, while reliability shows how well this object maintains its original level of quality over time through various and uncertain conditions (ASQ 2021).

Performance evaluation is considered important to both the public and private sector to implement successful PPPs (Liu et al. 2015), as it provides timely reports and feedback on where something needs to be modified or improved (Liu et al. 2019). Strategic objectives institute the baseline of performance evaluation, such as the achievement of best value, efficiency, cost, quality, time (Solomon and Young 2007). For example, many performance evaluations target the cost as the main indicator (Cappellaro and Ricci 2017; Liang and Wang 2019; Yuan et al. 2009; Zheng et al. 2018). However, as operator-led projects, the value of service delivered in the education PPPs is far more than the above one-shot evaluations or economic benefits but the long-term quality to provide social benefit, which would be essentially influenced by the uncertainties (Government 2010).

Thus, ensuring the reliability over long-time operation could be considered as one of main targets and distinct characteristic between the traditional construction projects and education PPP projects.

Furthermore, the question of delivering quality service and how to evaluate its performance has been at the heart of PPP projects in recent decades (Doloi 2012; Yuan et al. 2009). In industry, periodic performance audits by a third party, such as National Audit Office (NAO) in the UK, Auditor General in Australia, are always employed to measure if service is delivered at the required level based on the guidelines and best practices in International

Standards of Supreme Audit Institutions (ISSAI 2007). Key performance indicators (KPIs) are playing an important and popular role in such audits or performance evaluation (Haponava and Al-Jibouri 2012), which have been used as benchmarking tools for service delivery evaluation (NAO 2003). To better address service outcomes is essentially the nature of improving sets of comprehensive, efficient KPIs for evaluations. For example, within a lifecycle perspective, Liu and his colleagues in Western Australia have made many efforts on developing a process-based performance measurement framework, which identified performance indicators in the phases of planning and design, procurement and partnership (Liu et al. 2018, 2015, 2017; Love et al. 2017). For education PPPs, Saeed et al. (2018) developed an enhanced operation performance measurement framework for PPP school project by optimizing the performance indicators. Though there is an emphasis on KPIs, their use has been criticised for providing no insight to improve performance (Bassioni et al. 2004; Kagioglou et al. 2001).

Considering the special characteristics and changing environments, long-term operation of education PPPs suffer from pervasive uncertainties over future performance (Zhang et al. 2020). Uncertainty is one of the inherent project characteristics (Cleden 2012; Padalkar and Gopinath 2016; Perminova et al. 2008; Pich et al. 2002; Zarei et al. 2017), and being an unavoidable issue. Cruz and Marques (2013) emphasized PPP arrangements are particularly vulnerable to uncertainties due to complex and large-scale systems. For the education PPPs, service in operation comprises educational programmes, community activities and facility management, which embrace a range of individual service components, such as operational, security, cleaning, safety, utility, maintenance and repair services (Grimsey and Lewis 2004). In addition, school building developments involve diverse groups of stakeholders including central procuring agencies, principals, teachers, parents, students and wider community. The broad spread of stakeholders is likely to result in inconsistency towards provision of school services, complicating the decision-making process at all stages (Liu and Wilkinson 2014). It also aggravates the uncertainties in service delivery performance, when there is not a single factor but a rather complex set of factors between the public and private sector to take an effect.

On the other hand, some considerations are given to the qualitative analysis on efficiency and effectiveness of strategies to deal with uncertainties. House (2016) provides a dynamic vision on rules, relationships and regulations that make PPPs adapt to possible external shocks. Dewulf and Garvin (2019) conducted two investigations exploring the role of governance in addressing uncertainty in PPPs. One of which identified contractual mechanisms for risk sharing and the other proved the necessities of adjusting relational mechanisms to respond to unforeseen events. Flexible contracts have also been suggested to be an efficient way to cope with the uncertainty in PPPs (Cruz and Marques 2013; Zhang et al. 2020). However, few have paid attention to the quantitative approaches to analyse such uncertainties for the service performance over time, especially for education PPPs. Only a few studies have shed light on this field, such as Yuan et al. (2019) who measured operation performance of PPP rental housing by a comprehensive fuzzy-AHP evaluation.

It has become increasingly clear from the literature that, current performance evaluation practices usually provide a static qualitative profile, thus unable to fully grasp uncertainties or accurately predict trends. It has been observed that few studies track the dynamic service performance over long operations, and researchers generally have not examined service performance regarding uncertainties in education PPPs. Moreover, since the service delivery in education PPPs can be influenced by evolving uncertainties, a time dimension must be considered to understand the performance of the education public service. Thus, it is argued that the first key issue of performance evaluation in education PPPs is considering and quantifying the uncertainty and as well as improve the service delivery performance over long-term operations. To do this, an effective evaluation method is imperatively needed to better understand and measure the service performance of education PPPs.

Furthermore, reliability measurement has been an innovative way to evaluate performance in the field of service-oriented architecture and relevant systems, with regard to complex and dynamic uncertain environments. Singh et al. (2007) has studied reliability as the probability of no failure in decision-making processes. Using reliability is beneficial in reflecting state of knowledge for treatment of uncertainty in evaluations (Aven and Zio 2011), and effectively providing a picture of complex system likely behaviour, which is affected by

range of uncertainties (Johansson et al. 2013). In project management field, Wang et al. (2019) developed the reliability methodology to analyse the appropriateness of contracts guarantee or warranty provisions. Ghodoosi et al. (2021) estimated the probability of project failure by using reliability to quantify the uncertainty of schedule and cost overruns. A detailed review of reliability can be found in Chen et al. (2021); Cheng et al. (2018); Cheng et al. (2019); Cheng et al. (2021); Cheng et al. (2021); Ditlevsen (1979); Geng et al. (2020); Ghodoosi et al. (2021); Homaei and Najafzadeh (2020); Lai et al. (2013); Mahadevan and Haldar (2000); Mubarok et al. (2018); Sun et al. (2021). It should be noted that reliability has been playing an important role in assessing the performance uncertainty in fields such as manufacturing, structural engineering, supply chain management, and computing. And it also gets popular in project management, especially due to changing environment. Nevertheless, no substantial research has been done to address specifically this aspect into service performance in social PPPs.

Regarding the previously described background, the scope of this study is to address the reliability of service delivery in the long-operation education PPPs to quantify performance uncertainties, allowing for identifying potentially underperforming services and development of effective preventative maintenance strategies. A methodological hierarchical reliability evaluation framework is thus developed in next section.

3. Hierarchical Reliability Evaluation Framework

This study aims to evaluate the service delivery performance in terms of reliability over time and identify the key challenges for the PPP delivery in building, management and operation of school sectors. The main methodology involves two qualitative and quantitative parts. The qualitative part involves establishing a performance evaluation system via identifying performance indicators according to the output specifications of school PPP projects. A hierarchical structure is used here to comprehensively evaluate overall service performance as ranges of individual service components are included in the school PPP projects. The qualitative part involves the reliability analysis via the data collection of the above performance indicators. It could reflect the state of knowledge for treatment of uncertainty in evaluations (Aven and Zio 2011), and effectively provide a picture of complex system

behaviour, which is affected by uncertainties (Johansson et al. 2013). Similarly in quality-of-service requirements in manufacturing industry, reliability analysis could be employed here to ensure the service delivery process will perform as expected despite uncertainties. Four steps have been detailed illustrated as following for the service performance evaluation in school PPPs:

Step 1*: *Establish hierarchical evaluation system. The purpose of this step is to identify the service compositions in PPP school projects based on the output specifications. Three layers are used here, *i.e.,* indicator, component, and system (**Figure 2**). KPIs ($K_{i,j}$) derived from PPP school output specifications act as the indicator layer that are used for monitoring the service components(C_i), which could be further grouped into the upper-layer service system performance (P). Such a hierarchical system could better grasp overall service delivered in the context of school PPP projects, and could be easily adapted to any alternative project schemes (Wang et al. 2016).

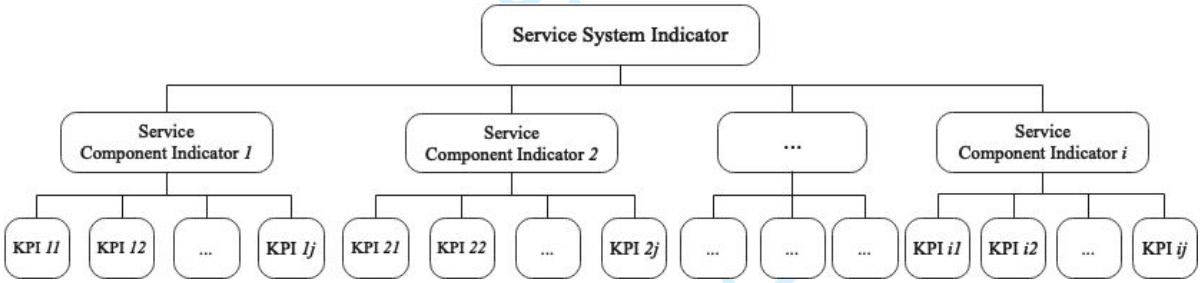

Figure 2 Hierarchical Structure of Service Evaluation System

Step 2: *Calculate service components reliability.* In the evaluation system, the performance of service components will be presented as the terms of reliability that depends on the failure limit state, where the delivered service does not reach performance standards. The higher the reliability, the lower probability of failure. The target reliability index for service components could be 4.0, which corresponds to the zero probability of failure. Two performance variables, *i.e.,* actual performance $P_{i,t}^{a}$, and performance standard $P_{i,t}^{*}$ are selected as input variables because of the service characteristic in PPP school projects. In contrast, for economic PPPs, the focus is often on revenue (Cabero Colín et al. 2017; Pantelias and Zhang 2010) or cash

flow (Lai et al. 2013) as key variables. The performance function of the failure limit state of the service component (C_i) can be described by the equation (1):

$$C_{i,t} = f(P_i) = P_{i,t}^{a} - P_{i,t}^{*}, \forall i = 1,2,...m \#(1)$$

Where $i \epsilon [1,2,\cdots,m]$, m represents the total number of service components in social PPPs service delivery system; and t refers to the year when the service evaluation is conducted.

Advanced First Order Reliability Methods (AFORM) are applied to calculate the components' reliability. The overall principle of AFORM is based on the theory that reliability can be estimated by the approximation of a performance function following the conversion of the random variables in terms of their moments (Ivanhoe et al. 2020). It is a semi-probabilistic method devised to evaluate the reliability of a system and has been widely used in reliability assessment because of computational efficiency and ease of implementation. The AFORM process in this study is illustrated in **Figure 3**.

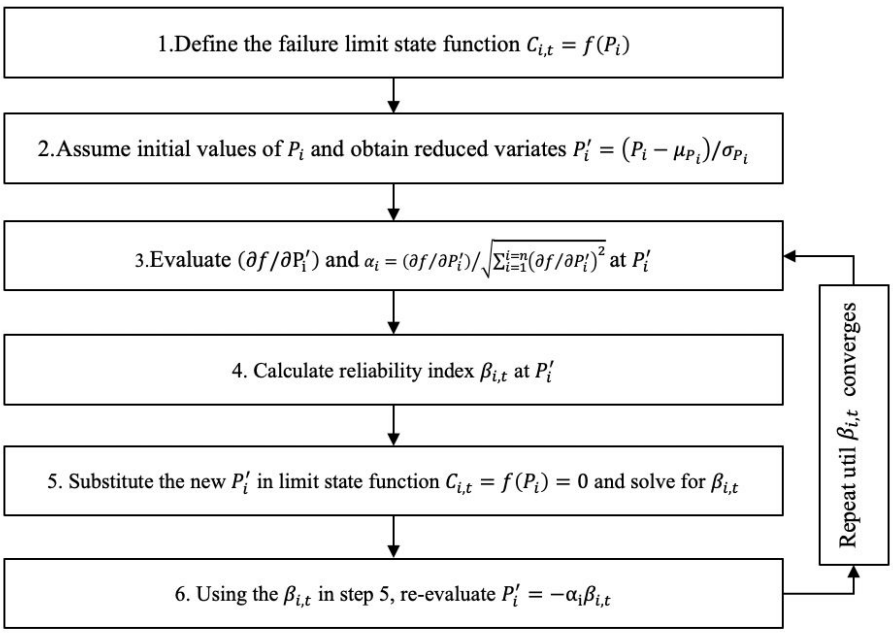

Figure 3 AFORM Process for Service Component Reliability

Thus, the reliability of components can be calculated by equation (2):

$$\beta_{i,t} = \frac{\mu_{i,t}^{a} - \mu_{i}^{*}}{\sqrt{\sigma_{i,t}^{a\,2} + \sigma_{i}^{*\,2}}}, \forall i = 1,2,\ldots m \#(2)$$

Where $\mu_{i,t}^{a}$,$\sigma_{i,t}^{a}$ and $\mu_{i,t}^{*}$,$\sigma_{i,t}^{*}$ are the mean and standard deviation of performance variables in equation (1); $\beta_{i,t}$ is reliability of i_{th}component at the timepoint t. The performance variables are assumed to follow a normal distribution, *i.e.*, $P_{i,t}^{a} \sim N(\mu_{i,t}^{a}, \sigma_{i,t}^{a})$, and $P_{i,t}^{*} \sim N(\mu_{i,t}^{*}, \sigma_{i,t}^{*})$. In the case where variables are not normally distributed, the reliability can be estimated by replacing equivalent normal mean $(\mu_{i,t}^{a})^{e}$and standard deviation $(\sigma_{i,t}^{a})^{e}$ iteratively, until the new variable becomes the normally distributed (Ghodoosi et al. 2021).

Step 3: *Evaluate the reliability of overall service performance*. The reliability of overall service delivery depends on weighting the reliability of each service component. Weighting (w_i) of service components accounts for their contributions in overall performance reliability, Thus, the overall reliability, *i.e.,* β_t, could be obtained by the equation (3).

$$\beta_t = \sum_{i=1}^{i=m} w_i * \beta_{i,t} \#(3)$$

In this step, the dynamic weighting approach determined through entropy method is preferred to indicate the time factor on the importance of different service components.

Step 4: *Provide strategic suggestions for PPP school service improvements.* The reliability of overall service delivery at *t* can be analysed and predicted using historical performance data via repeating ***step 2*** and ***step 3***. If reliability β_t is over 4, the government could consider future similar co-operative projects with this same service provider as satisfactory service performance. Otherwise, potentially underperforming service component should be analysed, thus providing the development of preventive maintenance strategies for improving overall service delivery.

The developed framework is illustrated in **Figure 4**.

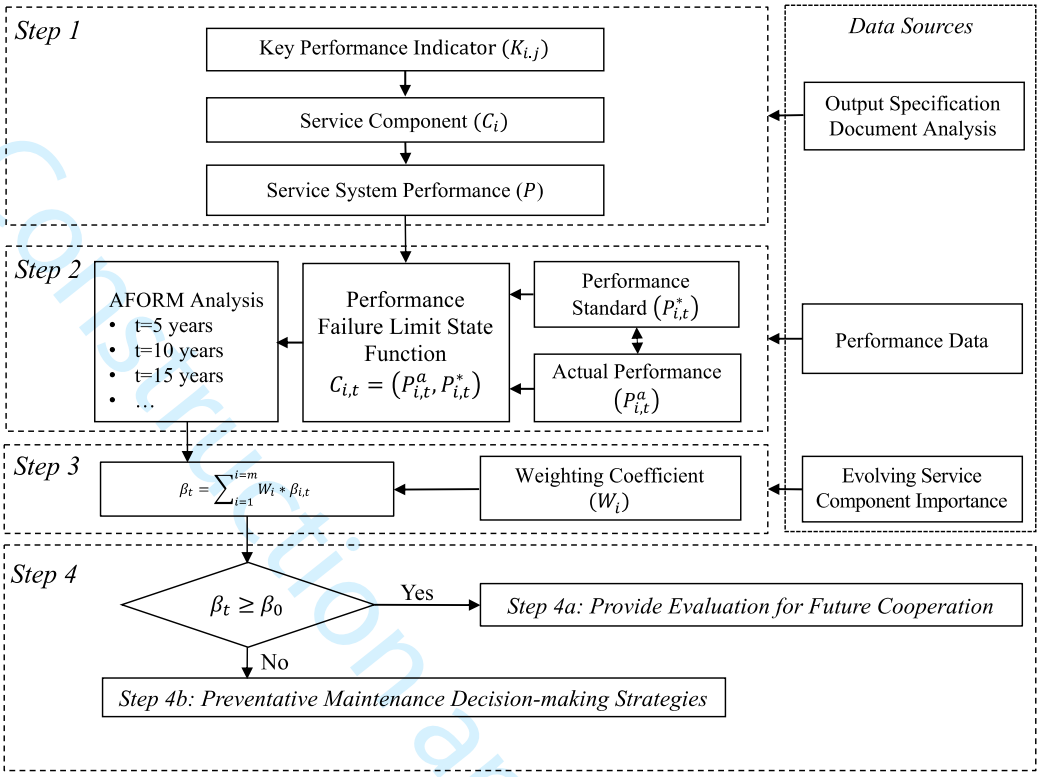

Figure 4 Hierarchical Reliability Evaluation Framework

4. PPP School Case Example

4.1 Background Information

The Victorian government in Australia recognizes the importance of education as it is both an essential criterion for social justice and economic growth. The high population growth forecast calls for many new PPP school facilities that aim to provide community services, primary, secondary, and special educations. As an example, the Victorian government firstly entered a PPP arrangement in the school sector in 2008, named *Partnerships Victoria in Schools Project* (*PViS*), with Project Co, *i.e.,* the PPP partner, to design, build, operate, and maintain 12 government schools for a 25-year period. The PPP schools were undertaken in compliance with National PPP policies and supplementary Victorian PPP guidance (DEECD 2009), and focused on exemplary educational design, innovative community hubs, and the support of inclusive education for all learners. The whole of life service delivery approach of the PPPs will free up the time of principals and managers, enabling educators to focus on

learners rather than maintenance and repair of their schools. The *PViS* is selected here as the main basis for the theoretical case example in this study to demonstrate the feasibility and effectiveness of the developed framework. *PViS* represents ecologically sustainable features in core PPP school services in the community in key growth area of Melbourne.

4.2 Scenario Development

In this case example, the penalty points approach is applied to track the service performance. That is, associated number of penalty points of KPIs for service quality failure accounting for individual service components are identified and agreed upon during contract negotiation. Then the accrued number of penalty points is compared with a contractually agreed baseline, which is usually based on built-in tolerance. If the accrued points are greater than baseline total, corresponding abatement percentages will be imposed on the 'Project Co'. Thus, the performance penalty points of KPIs accrued at the evaluation time t, and their performance baseline act as the two main performance variables, i.e., P^a and P^*, for the service evaluation. However, due to confidentiality reasons, the historical data for penalty points (P^a) was not available. To tackle this practical problem, scenario analysis is used to predict the possible occurrences of the KPI penalty points and their consequences.

Because of various uncertainties influencing the service performance, i.e., whether the KPI reaches the requirements or not, it is necessary to develop various scenarios to evaluate the different situations. In this scenario development, four endogenous variables, *i.e.,* a) resource investment level; b) effort level (Zhang et al. 2020); c) the number of enrolment numbers; and d) the length of operated year, are assumed to have obvious impacts on service delivery system over time in PPP schools. In other words, the more resource investment level and the more effort 'Project Co' put into this case; the better service performance will be delivered. The more students are enrolled and the longer the operations, the worse service delivery performance. Three scenarios derived from the assumptions based on endogenous variables are therefore set to simulate the performance data and validate the feasibility of the proposed framework. The main goal of the scenario analysis is to reflect the real-world situation that service delivery performance will be affected by many uncertainties and thus can be present different situations.

Scenario 1: the resource investment level and effort level that Project Co puts into this school case are so high that the negative impact of increasing enrolment number and operated time could be mitigated. The risk remains unchanged during 25-year operation.

Scenario 2: the resource investment level and effort level that Project Co puts into this school case are not enough to mitigate negative impact of enrolment numbers and the operated time. The risk is assumed to increase on a basis of the increasing enrolment numbers during 25-year operation.

Scenario 3: the resource investment level and effort level that Project Co puts into this school case are not enough to mitigate the negative impacts of enrolment numbers and operated time. The risk is assumed to increase on a basis of the increasing enrolment numbers during 25-year operation, also weighting of service components evolves over time.

Table. 1 shows the assumptions in the three scenarios.

Table.1. Three Scenarios Assumptions

Scenario No.	Assumptions	
	Risk	**Weighting**
1	Ignore the risk change due to the increase in enrolment numbers and time operating in the high resource investment level and effort level, *i.e.,* $r_t = r_0$	Fixed
2	Consider the risk change due to the increase in enrolment numbers in the low resource investment level and effort level, *i.e.,* $r_t = r_{t-1} \cdot k \cdot {}^{n_t}/_{n_{t-1}}$	Fixed
3	Consider the risk change due to the increase in enrolment numbers in the low resource investment level and effort level, i.e., $r_t = r_{t-1} \cdot k \cdot {}^{n_t}/_{n_{t-1}}$	Evolves over time

r_0 represents the initial service performance risk at $t = 0$, assumed 4% in this example; n represents the enrolment numbers of PPP school. A P^a dataset is then simulated via binomial distribution in MATLAB based on the endogenous variables under the scenario assumptions. Thus, the value of $K_{i,j}^a$ is generated as 1 or 0 according to risk profiles in the three scenarios,

where 1 refers to this KPI does not reach requirement in output specifications and thereby corresponding penalty point will be assigned to this KPI, while 0 means no penalty points.

4.3 Data Acquisition

Qualitative data in this case example is obtained from *PViS* contract documents, which covers output specifications and project agreements for 12 PPP schools (DEECD 2009). It supports the establishment of a hierarchical service evaluation system in *Step 1*, as KPIs and their components are identified. Quantitative data used for service delivery performance reliability evaluation comes from a binomial simulation of penalty points of KPIs under the different scenario assumptions. The data of endogenous variable, like enrolment numbers, is predicted through Autoregressive Integrated Moving Average (ARIMA) by capturing historical data from official websites (e.g., *https://www.myschool.edu.au/*). The quantitative data supports reliability analysis in *step 2* and *step 3*. The data processing flow chart is illustrated in **Figure 5.**

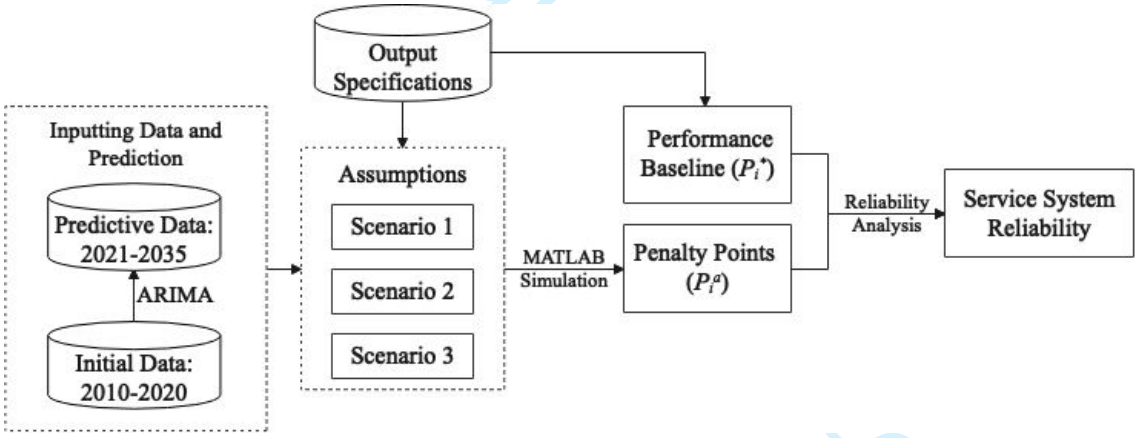

Figure 5 Data Processing Flow Chart

4.4 Results and Discussion

4.4.1 *Step 1: Establish hierarchical evaluation system*

The service delivered in *PViS* is composed of 15 service components, *i.e.,* C_1: Performance monitoring and reporting; C_2: Management; C_3: Staff development; C_4: Policy and Strategy; C_5: Partnerships and resources; C_6: Failure event, end-user Satisfaction and material; C_7:

Table 2 Service Components and Indicators

System	Components	Indicator Descriptions
Service delivery system (P)	C_1: Performance Monitoring and Reporting	Including response to service failure, performance reports (4 indicators)
	C_2: Management	Including reasonable systems, controls to safeguard property, clearly defined proposed changes (3 indicators)
	C_3: Staff and Development	Including no staff shortage and continuous professional development (7 indicators)
	C_4: Policy and Strategy	Including the Project Co complies with all health and safety requirements (19 indicators)
	C_5: Partnerships and Resources	Including Project Co provides incidental equipment and consumables for provision of services and must liaise with all relevant stakeholders (7 indicators)
	C_6: Failure Event, End-user Satisfaction and Material	Including avoiding more than 6 same service failure in any rolling 3 months, undertake user satisfaction survey and students cannot have access to the stored materials (3 indicators)
	C_7: Help Desk Service	Including service facilitating the carrying out of school service on 24/7 (21 indicators)
	C_8: Building Management Service	Including the integrity and functionality of the building fabric, building services, utility services systems security, public safety, equipment and plant which comprise the facility (32 indicators)
	C_9: Utilities Management Service	Including the procurement, management and maintenance of electricity, communications system, renewable energy sources, gas water (16 indicators)
	C_{10}: Waste Management Service	Including collection, management and disposal of the full range of waste that minimizes the impact on the environment (5 indicators)
	C_{11}: Ground and Gardens Maintenance Service	Including maintenance for roads, pathways, signage, lighting, carpark (17 indicators)
	C_{12}: Pet Control Service	Including emergency pest control advice, pest deterrents, capture, disposal, the use of pest control chemicals (6 indicators)
	C_{13}: Cleaning Service	Including scheduled cleaning, Emergency cleaning and deep cleaning to maintain a clean, healthy and safe environment for all users (11 indicators)
	C_{14}: Security Service	Including monitoring and alarms, secure perimeter to the site and facility and access to facilities (20 indicators)
	C_{15}: Churn Management Service	Including equipment management for specific requests, mail and portable appliance testing (7 indicators)

Help desk service; C_8: Building management service; C_9: Utilities management service; C_{10}: Waste management service; C_{11}: Ground and gardens maintenance service; C_{12}: Pest control service; C_{13}: Cleaning service; C_{14}: Security service; C_{15}: Churn management service. The details of their 178 indicators could be found in DEECD (2009). The evaluation system is shown in **Table 2.**

4.4.2 *Step 2: Evaluating service components reliability*

The penalty points of 15 service components could be calculated by the summation of their KPIs penalty points depending on weather they reach performance standard or not. In this example, the performance data, *i.e.*, penalty points, is generated by binomial distribution via MATLAB under three scenario assumptions: 1) the penalty points of KPIs are generated under unchanging service risk, *i.e.*, $r_t = r_0 = 4\%$; 2) the penalty points of KPIs are generated under changing service risk due to the increase of enrolment numbers in schools, i.e., $r_t = r_{t-1} \cdot {}^{n_t}/_{n_{t-1}} (k = 1)$; 3) the penalty points of KPIs are generated under changing service risk due to enrolment numbers increase, *i.e.*, $r_t = r_{t-1} \cdot {}^{n_t}/_{n_{t-1}} (k = 1)$ and dynamic weightings is considered. Besides, the abatement cost will occur when the performance is under the 85% satisfaction. Hence, it is assumed that $P_{i,t}^{*} = 0.15 * total\ penalty\ points$ is the baseline when comes to the service component reliability. For example, a total of 27 penalty points in C_1 means that when the penalty points are over 4, the abatement cost will be accrued on this component. The reference values, such as $r_0 = 4\%$, $k = 1$, and 85% are just to reflect the prerequisites for the validation of the proposed framework, which could be modified when put into the real case. Based on the equation (2), the component reliabilities are illustrated in **Figure 6.**

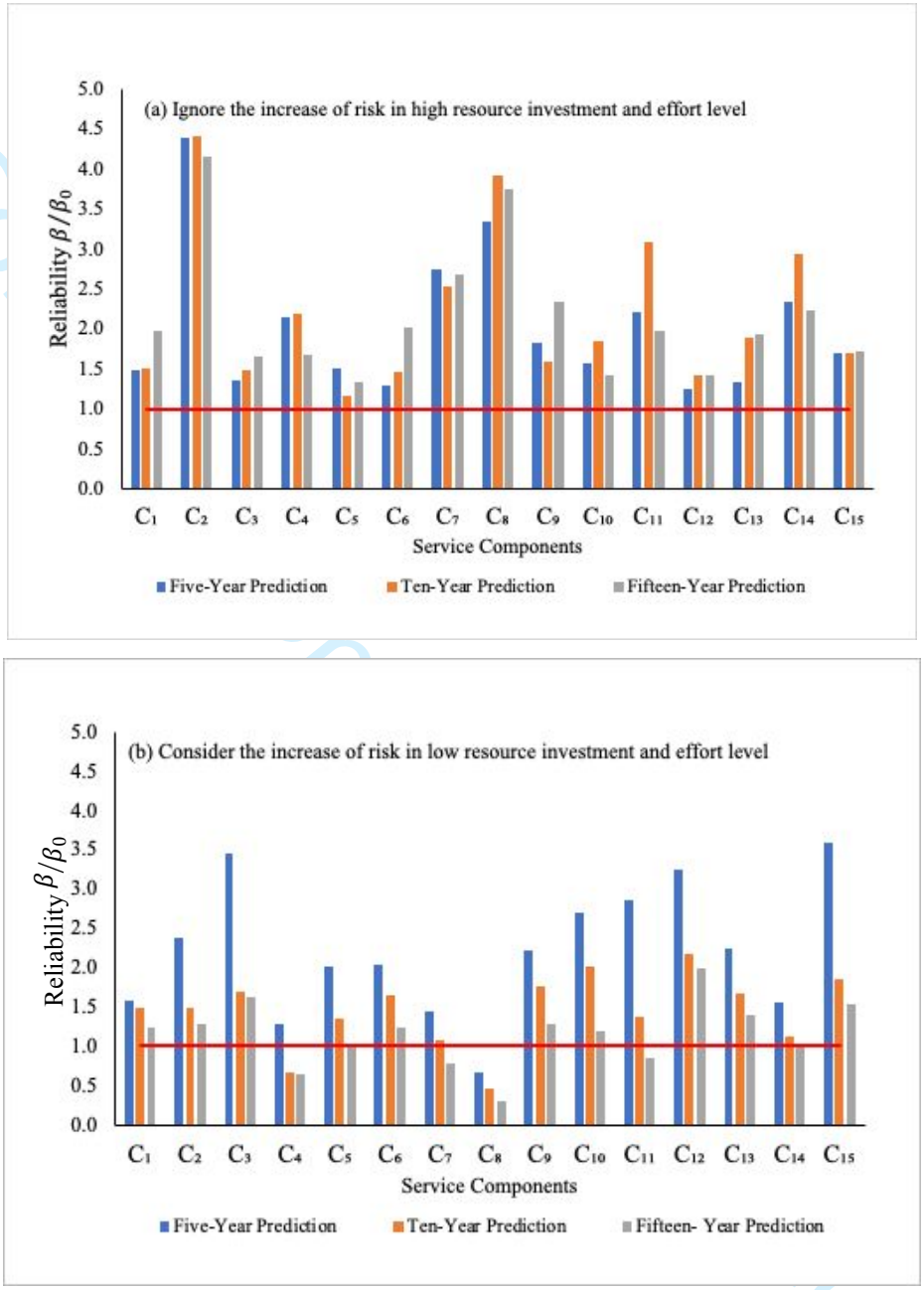

Figure 6 Effect of Enrolment Numbers on Long-term Performance of each Service Component a) Ignore the increase of risk in high resource investment level and effort level; b) Consider risk increase in low resource investment level and effort level

Figure.6 shows the reliability of 15 service components under the high and low resource investment level and effort level scenarios. Since the target reliability index $\beta_0 = 4$, then $\beta_i/\beta_0 < 1$ could be a criterion that means the i_{th} service component has the probability to be

abated. Some key information could be derived. Firstly, the reliability of service components is generally higher than 1 in scenario of high resource and effort level, meaning the service would be delivered well, if Project Co invests enough resources and effort. Otherwise, the reliability will decrease due to the increasing enrolment numbers. In reliability theory, the lower reliability, the higher probability of failure, which is also consistent with the increasing risk scenario under the increasing enrolment numbers in scenario low effort and resources level. While unlike this case, the service risk level is normally unknown in real cases, then a reliability index can be derived from actual performance data. This provides an intuitive and comprehensive indicator of service performance taking into consideration the uncertainties coming from complex internal and external factors. Secondly, for this theoretical example (**Figure 6**), the lowest reliability points to problems in components C_4, and C_8, which are policy & strategy, and building management service. Similarly, Qin et al. (2021) has said that policy uncertainty is the main cause of PPP failure, especially for the firm-level investment. Political interference is also seen as one of the top-ranked risk factors in PPP infrastructure projects (Ameyaw and Chan 2015). Building management service is highly related to end-user's satisfaction, and makes a contribution to the educational performance (Wijayasundara et al. 2020). Thus, the result provides a reminder where the focus should be given in current project and also for the future similar cases to keep a high reliable service delivery, when the resources and effort is limited.

4.4.3 Step 3: Evaluate the reliability of overall service performance

The reliability of overall service performance can be calculated by aggregating component's reliability and their weightings through equation (3). In the scenario 1 and 2, the components are assumed to be equally important, *i.e.,* $w_i = 1/15$; In scenario 3, the dynamic weighting of components over time is considered, which is determined through the entropy method. In this method, the greater difference in indicator values of samples, the smaller the information entropy, and the greater weighting of the indicator. The entropy method is advantageous in eliminating the influence of subjective factors on the weight of each influencing factor, which makes the evaluation of the system fairer and objective (Ivanhoe et al. 2020). The overall reliability is presented in **Figure 7.**

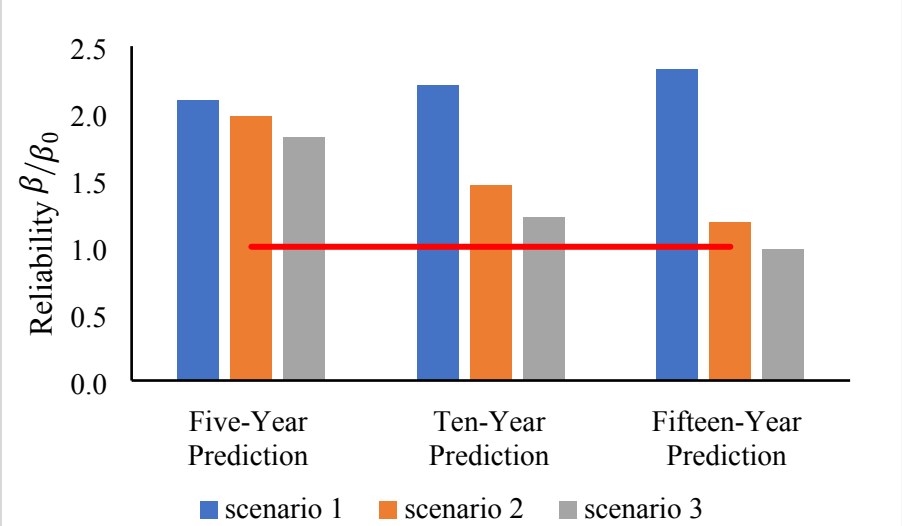

Figure 7 Overall Service Reliability

Figure 7 shows the overall reliability in scenario 1, *i.e.,* ignore the increase of risk due to enrolment numbers in high resources investment level and effort level, is always over 1, while the overall reliability is in a decreasing trend in scenario 2 and 3, when considering the increase of risk due to enrolment numbers in low resources investment level and effort level. It is noted that the reliability in the scenario 3 is lower than that in scenario 2, this is because in this service delivery system, the more unstable penalty points the components have, the more critical it should be, thus the weighting will be dynamic over time as the penalty points change. In scenario 3, dynamic weighting is exactly considered by tracking the penalty points of each component at each year. In this case, the critical service components are C_4: policy and strategy, and C_8: building management service. They accrue the most unstable on-going penalty points in 12 school samples. Therefore, their weightings should be correspondingly increased, which leads to a lower reliability in overall service performance compared to the fixed weighting. This provides us an objective dynamic evaluation rather than overestimated result. In addition, to improve service delivery performance, some efficient measures, like more inspections, stronger punishment, and professional support on the critical components, should be taken into account to improve their reliability, thus improving overall service delivery performance.

Though the assumptions in scenarios are used for validating the proposed service reliability framework only, the reference values, like k, r_t, w_i have an important relation to the evaluation result. Therefore, a simple parametric study is conducted to explore the influence of different parameter values on results. In **Figure 8,** When $k = 2$, the reliability has already been below the target value (red line), which represents the high uncertainty in such a high-risk assumption. **Figure 9** shows service risk has a more obvious impact than the weighting on the result, taking the weighting of C_8 as an example.

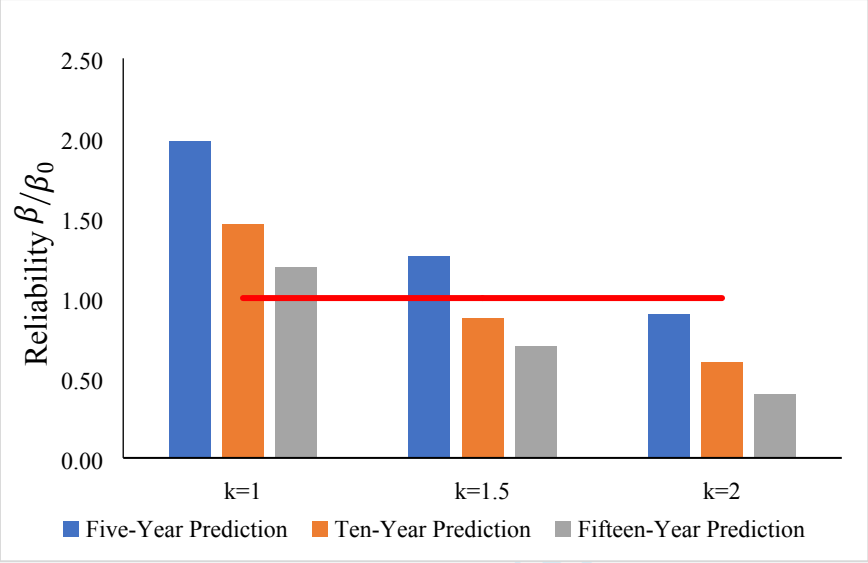

Figure.8 Overall Reliability under Assumptions k=1,1.5 and 2 in Scenario 2

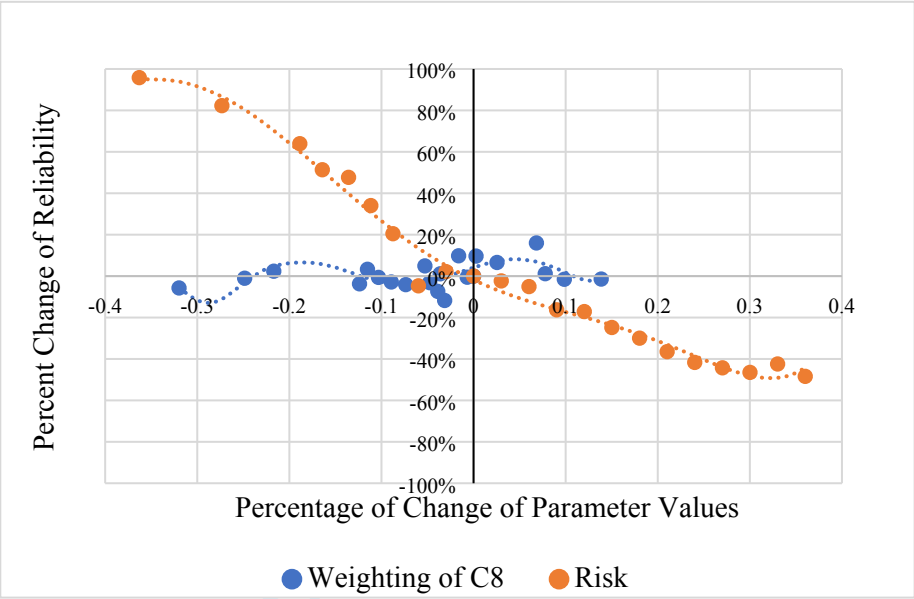

Figure.9 Sensitivity Analysis of r_t, w_8

4.4 Discussion

Effective service performance evaluation has been the critical part for both the government and private sector to successfully implement a PPP project. For education PPPs, the service performance over a long period of 25 or 30 years has been particularly important to achieve better value for money. However, complex, uncertain internal and external factors introduce numerous negative impacts on service delivery, result in unqualified service, unpredictable renegotiation costs or society complaints. Regarding the uncertainties in service delivery in education PPPs, a hierarchical framework with reliability analysis is developed in this study. In this evaluation, reliability of critical service components and overall service delivery can all be considered. This developed framework provides an innovative quantitative perspective for managing uncertainties based on prediction of historical KPIs data. The long timespan of asset service operations is one of the most special features of education PPPs when compared with traditional projects and thus deserves to be considered in detail. Thus, this framework also enables tracking dynamic uncertainties in future time predictions through providing the performance reliability index, when enough historical data can be provided. The theoretical example set up in *PViS* background proved the feasibility of the developed framework in assessing dynamic performance trend over the long operations period. In other words, when

actual historical performance data becomes available, it will be able to provide more clues in predicting service performance and developing preventative strategies. Also, as an evaluation tool, the developed framework could help government in the evaluations concerning decisions around future cooperation with the same service provider.

5. Conclusions

Most previous research relevant to service performance evaluation in PPPs have mainly been qualitative or conceptual. The impact of uncertainties on the service delivery over the long operations has largely been ignored. There is a lack of studies using quantitative approaches on such a topic, especially in education PPPs. An accurate way of knowing such impacts on the service delivery could provide decision-makers with hints on how to improve service performance and determining the private sector's ability in delivering such PPP services.

In this study, the research gap is filled by developing a hierarchical reliability evaluation framework for service performance in education PPPs. The reliability index can be used to measure if the delivered public-school services reach the performance standard are required. To illustrate the developed framework, a theoretical example is used where the qualitative data derives from *PViS* contract documents, while quantitative performance data is generated in MATLAB by assuming different scenarios. It shows that the service components, policy and strategy, building management service, have the lowest performance reliability of the 15 service components. In addition, the reliability of overall performance has a decreasing trend when considering the increase of risk due to enrolment numbers. Furthermore, dynamic weighting is preferred as providing a more objective evaluation result. The example shows the proposed framework could be used by practitioners in industry as a decision tool to identify, analyse and assess the reliability in social PPPs service delivery system over long period operations. This will increase the likelihood of success in a complex environment.

For the contributions, in theory, a performance evaluation index, reliability, is constructed in this study to process the dynamics of service delivery uncertainty over the PPP operations. Project performance functions for service components and system are also formulated. These enhance the integration of domain knowledge on performance evaluation and improvements. The developed framework is tested with real-world data to simulate practical conditions. In

addition, the proposed model can measure the impact of uncertainties on service performance and predict in future similar projects. It also provides useful guidance for government to understand the ability of the private sector in delivering PPP school service.

Admittedly, in demonstrating the proposed framework, the unavailability of actual KPIs data remains a main limitation. Current KPI models should provide fundamental performance data that could be efficiently used by government or education departments through the proposed framework to further evaluate the impact of uncertainty in service delivery performance. Also, this framework theoretically takes all relevant factors as an uncertainty system and only considers the linear impacts of specific variables, like resource investment level and effort level. Future studies could further explore quantitative consideration of these two variables and incorporate them into the current framework to obtain an instructive relationship between such uncertainty and performance reliability. Finally, the proposed framework examined PPP school projects in an Australian context, but it could be applied generally to any social PPP context by adjusting the indicators in the service delivery evaluation system.

Reference

Aarseth, O. A., Urdal, V. M., Bjørberg, S., Støre-Valen, M., and Lohne, J. (2016). "PPP in Public Schools as Means for Value Creation for User and Owner." *Procedia - Social and Behavioral Sciences*, 226, 398-406.

Ameyaw, E. E., and Chan, A. P. (2015). "Risk ranking and analysis in PPP water supply infrastructure projects: an international survey of industry experts." *Facilities*.

ASQ (2021). "what is reliability." <https://asq.org/quality-resources/reliability>.

Atmo, G. U., Duffield, C., Zhang, L., and Wilson, D. I. (2017). "Comparative performance of PPPs and traditional procurement projects in Indonesia." *International Journal of Public Sector Management*.

Aven, T., and Zio, E. (2011). "Some considerations on the treatment of uncertainties in risk assessment for practical decision making." *Reliability Engineering & System Safety*, 96(1), 64-74.

Bassioni, H. A., Price, A. D., and Hassan, T. M. (2004). "Performance measurement in construction." *Journal of management in engineering*, 20(2), 42-50.

Cabero Colín, F., Sánchez Soliño, A., and Lara Galera, A. L. (2017). "Default and abandonment option in motorway concessions." *Journal of Infrastructure Systems*, 23(1), 04016022.

Cappellaro, G., and Ricci, A. (2017). "PPPs in health and social services: a performance measurement perspective." *Public Money & Management*, 37(6), 417-424.

Carbonara, N., and Pellegrino, R. (2020). "The role of public private partnerships in fostering innovation." *Construction Management and Economics*, 38(2), 140-156.

Chen, C., Yan, T., and Chen, P. "Research on Risk Allocation Model in PPP Projects: Perspectives from the SOEs." *Proc., International Conference on Management Science and Engineering Management*, Springer, 1559-1572.

Chen, S., Duffield, C., Miramini, S., Raja, B. N. K., and Zhang, L. (2021). "Life-cycle modelling of concrete cracking and reinforcement corrosion in concrete bridges: A case study." *Engineering Structures*, 237, 112143.
Cheng, C., Zhang, L., and Thompson, R. G. (2018). "Reliability analysis for disaster waste management systems." *Waste Management*, 78, 31-42.
Cheng, C., Zhang, L., and Thompson, R. G. (2019). "Reliability analysis of road networks in disaster waste management." *Waste Management*, 84, 383-393.
Cheng, C., Zhu, R., Thompson, R. G., and Zhang, L. (2021). "Reliability analysis for multiple-stage solid waste management systems." *Waste Manag*, 120, 650-658.
Cheng, C., Zhu, R., Thompson, R. G., and Zhang, L. (2021). "Reliability analysis for multiple-stage solid waste management systems." *Waste Management*, 120, 650-658.
Cleden, M. D. (2012). *Managing project uncertainty*, Gower Publishing, Ltd.
Cruz, C. O., and Marques, R. C. (2013). "Flexible contracts to cope with uncertainty in public–private partnerships." *International journal of project management*, 31(3), 473-483.
De Carlo, F. (2013). "Reliability and maintainability in operations management." *Operations Management*, 1, 32.
DEECD (2009). "Partnerships Victoria in Schools Project Summary." DEECD, ed.Melbourne.
Dewulf, G., and Garvin, M. J. (2019). "Responsive governance in PPP projects to manage uncertainty." *Construction Management and Economics*, 38(4), 383-397.
Ditlevsen, O. (1979). "Narrow reliability bounds for structural systems." *Journal of structural mechanics*, 7(4), 453-472.
Duffield, C., Saeed, A., and Tamburro, N. (2020). *Measuring the value and service outcomes of Social Infrastructure PPPs in Australia and New Zealand*.
EPEC (2020). "Review of the European PPP Market in 2019."
Fulbright, N. R. (2021). "Infrastructure/PPP Course Handbook - Level 2." ALSF Academy, https://alsf.academy/resources/infrastructureppp-course-handbook-level-2, 168.
Geng, L., Herath, N., Zhang, L., Kin Peng Hui, F., and Duffield, C. (2020). "Reliability-Based Decision Support Framework for Major Changes to Social Infrastructure PPP Contracts." *Applied Sciences*, 10(21), 7659.
Ghodoosi, F., Bagchi, A., Hosseini, M. R., Vilutienė, T., and Zeynalian, M. (2021). "Enhancement of Bid Decision-Making in Construction Projects: A Reliability Analysis Approach." *Journal of Civil Engineering and Management*, 27(3), 149-161.
Government, N. (2010). "National Infrastructure Plan." Infrastructure, ed.https://www.treasury.govt.nz/sites/default/files/2017-12/nip-mar10.pdf, 143.
Grimsey, D., and Lewis, M. K. (2004). "The governance of contractual relationships in public–private partnerships." *Journal of corporate citizenship*(15), 91-109.
Haponava, T., and Al-Jibouri, S. (2012). "Proposed system for measuring project performance using process-based key performance indicators." *Journal of management in engineering*, 28(2), 140-149.
Homaei, F., and Najafzadeh, M. (2020). "A reliability-based probabilistic evaluation of the wave-induced scour depth around marine structure piles." *Ocean Engineering*, 196, 106818.
House, S. (2016). "Responsive regulation for water PPP: Balancing commitment and adaptability in the face of uncertainty." *Policy and society*, 35(2), 179-191.
Hueskes, M., Verhoest, K., and Block, T. (2017). "Governing public-private partnerships for sustainability An analysis of procurement and governance practices of PPP infrastructure projects." *International Journal of Project Management*, 35(6), 1184-1195.
iMC, W. (2020). "Australia PPP Market 2020." InfraPPP reports.
ISSAI, I. S. o. S. A. I. (2007). "Guidelines on best practice for the audit of public/private finance and concessions."

Ivanhoe, R., Wang, L., and Kolios, A. (2020). "Generic framework for reliability assessment of offshore wind turbine jacket support structures under stochastic and time dependent variables." *Ocean Engineering*, 216, 107691.
Johansson, J., Hassel, H., and Zio, E. (2013). "Reliability and vulnerability analyses of critical infrastructures: comparing two approaches in the context of power systems." *Reliability Engineering & System Safety*, 120, 27-38.
Kagioglou, M., Cooper, R., and Aouad, G. (2001). "Performance management in construction: a conceptual framework." *Construction management and economics*, 19(1), 85-95.
Karasavidis, S. (2018). "'Build and operate' increasingly common in social infrastructure."
Kececioglu, D. (2002). *Reliability engineering handbook*, DEStech Publications, Inc.
Lai, J., Zhang, L., Duffield, C., and Aye, L. (2013). "Economic risk analysis for sustainable urban development: validation of framework and decision support technique." *Desalination and Water Treatment*, 52(4-6), 1109-1121.
Lember, V., Petersen, O. H., Scherrer, W., and Ågren, R. (2019). "Understanding the relationship between infrastructure public–private partnerships and innovation." *Annals of Public and Cooperative Economics*, 90(2), 371-391.
Liang, Y., and Wang, H. (2019). "Sustainable Performance Measurements for Public–Private Partnership Projects: Empirical Evidence from China." *Sustainability*, 11(13).
Liu, H. J., Love, P. E., Sing, M. C., Niu, B., and Zhao, J. (2019). "Conceptual framework of life-cycle performance measurement: Ensuring the resilience of transport infrastructure assets." *Transportation Research Part D: Transport and Environment*, 77, 615-626.
Liu, J., Love, P. E., Davis, P. R., Smith, J., and Regan, M. (2015). "Conceptual framework for the performance measurement of public-private partnerships." *Journal of Infrastructure systems*, 21(1), 04014023.
Liu, T., and Wilkinson, S. (2014). "Using public-private partnerships for the building and management of school assets and services." *Engineering, construction and architectural management*.
Mahadevan, S., and Haldar, A. (2000). "Probability, reliability and statistical method in engineering design." *John Wiley & Sons*.
Maloney, W. F. (2002). "Construction product/service and customer satisfaction." *Journal of construction engineering and management*, 128(6), 522-529.
Mohd-Rahim, F., Abd-Rahim, M., Zainon, N., Chuing, L., and Abd-Samad, Z. (2018). "Project life cycle risk of public-private partnership (PPP) projects for construction sustainability." *Journal of Design and Built Environment*, 39-53.
Mubarok, K., Xu, X., Ye, X., Zhong, R. Y., and Lu, Y. (2018). "Manufacturing service reliability assessment in cloud manufacturing." *Procedia CIRP*, 72, 940-946.
NAO (2003). "The operational performance of PFI prisons." The stationery office London.
Oyedele, L. O. (2013). "Avoiding performance failure payment deductions in PFI/PPP projects: model of critical success factors." *Journal of Performance of Constructed Facilities*, 27(3), 283-294.
Padalkar, M., and Gopinath, S. (2016). "Are complexity and uncertainty distinct concepts in project management? A taxonomical examination from literature." *International Journal of Project Management*, 34(4), 688-700.
Pantelias, A., and Zhang, Z. (2010). "Methodological framework for evaluation of financial viability of public-private partnerships: Investment risk approach." *Journal of infrastructure systems*, 16(4), 241-250.
Perminova, O., Gustafsson, M., and Wikström, K. (2008). "Defining uncertainty in projects–a new perspective." *International journal of project management*, 26(1), 73-79.

Pich, M. T., Loch, C. H., and Meyer, A. D. (2002). "On uncertainty, ambiguity, and complexity in project management." *Management science*, 48(8), 1008-1023.
Qin, S., Luo, W., and Wang, Y. (2021). "Policy uncertainty and firm-level investment: Evidence from public-private partnership market in China." *Applied Economics Letters*, 1-7.
Robinson, H. S., and Scott, J. (2009). "Service delivery and performance monitoring in PFI/PPP projects." *Construction Management and Economics*, 27(2), 181-197.
Saeed, A. M., Duffield, C., and Hui, F. K. P. (2018). "An enhanced framework for assessing the operational performance of public-private partnership school projects." *Built Environment Project and Asset Management*, 8(2), 194-214.
Singh, V. P., Jain, S. K., and Tyagi, A. "Risk and reliability analysis: a handbook for civil and environmental engineers." American Society of Civil Engineers.
Solomon, P. J., and Young, R. R. (2007). *Performance-based earned value*, J. Wiley & Sons.
Sun, D., Huang, C., Cao, Z., Wu, K., and Zhang, L. (2021). "Reliability assessment of concrete under external sulfate attack." *Case Studies in Construction Materials*, 15, e00690.
Ugurlu, O. F., and Kumral, M. (2020). "Reliability-based performance analysis of mining drilling operations through Markov chain Monte Carlo and mean reverting process simulations." *SIMULATION*, 96(7), 593-604.
Villani, E., Greco, L., and Phillips, N. (2017). "Understanding value creation in public-private partnerships: A comparative case study." *Journal of Management Studies*, 54(6), 876-905.
Wang, B., Cui, Q., and Zhang, S. (2019). "Evaluation of the contract reliability for alternative infrastructure project delivery: a contract engineering method." *Frontiers of Engineering Management*, 6(2), 239-248.
Wang, H., Liu, Y., Xiong, W., and Song, J. (2019). "The moderating role of governance environment on the relationship between risk allocation and private investment in PPP markets: Evidence from developing countries." *International Journal of Project Management*, 37(1), 117-130.
Wang, T., Wang, S., Zhang, L., Huang, Z., and Li, Y. (2016). "A major infrastructure risk-assessment framework: Application to a cross-sea route project in China." *International Journal of Project Management*, 34(7), 1403-1415.
Wang, Y., Wang, Y., Wu, X., and Li, J. (2020). "Exploring the Risk Factors of Infrastructure PPP Projects for Sustainable Delivery: A Social Network Perspective." *Sustainability*, 12(10).
Wijayasundara, M., Zhang, L., and Duffield, C. (2020). "Relating building space to performance outcomes–a methodology to explore the relationship." *Journal of building engineering*, 32, 1-17.
Yuan, J., Li, W., Xia, B., Chen, Y., and Skibniewski, M. J. (2019). "Operation performance measurement of public rental housing delivery by ppps with fuzzy-ahp comprehensive evaluation." *International Journal of Strategic Property Management*, 23(5), 328-353.
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*, 27(3), 253-270.
Zarei, H., Hui, K., Duffield, C., and Wang, G. (2017). "The Risk of Power Imbalance in Project Delivery: A Study of Large Victorian Public Infrastructure Projects."
Zhang, H., Yu, L., and Zhang, W. (2020). "Dynamic performance incentive model with supervision mechanism for PPP projects." *Engineering, Construction and Architectural Management*, 27(9), 2643-2659.
Zhang, L., Sun, X., and Xue, H. (2019). "Identifying critical risks in Sponge City PPP projects using DEMATEL method: A case study of China." *Journal of cleaner production*, 226, 949-958.

Zheng, S., Xu, K., He, Q., Fang, S., and Zhang, L. (2018). "Investigating the sustainability performance of ppp-type infrastructure projects: a case of China." *Sustainability*, 10(11), 4162.