NOTE

A critical review of the application of TCE in the interpretation of risk allocation in PPP contracts

Chen-Yu Chang

Bartlett School of Construction and Project Management, University College London, 1–19 Torrington Place, London, WC1H 7HB, UK, chen-yu.chang@ucl.ac.uk

Abstract

Along with the rise of public–private partnerships (PPPs) as a mainstream procurement system, we have seen a growing interest in studies of risk allocation in these projects. One of the serious academic endeavours is to apply both transaction cost economics (TCE) and the resource-based view (RBV) to explain risk allocation patterns found in PPP projects. The existing literature along these lines is deficient in three aspects: inappropriate choice of unit of analysis; poor specification of governance structure; and misinterpretation of asset specificity. A way for improvement is to analyse risk allocation in the context of PPP procurement in its entirety.

Keywords: competence, procurement, public–private partnerships, risk analysis, transaction cost.

Introduction

In recent decades, we have seen an increasing number of governments around the world join the bandwagon of public–private partnerships (PPPs) in an effort to improve efficiency in the provision of public services. Risk allocation plays a pivotal role in the pursuit of value for money in PPP procurement (Her Majesty’s Treasury, 2006). That is why risk allocation in PPP projects has emerged as a vibrant research topic. In the existing literature on this topic, some researchers use transaction cost economics (TCE) as theoretical grounding for their explanation of risk allocation patterns found in PPP projects. As will be elaborated later, the methodologies employed in these papers are flawed. This line of inquiry has yielded a number of papers in leading construction management journals, such as Jin (2010) and Jin and Doloi (2008). As these papers share a common theoretical core, for ease of exposition, ‘JD Model’ is used in the following discussion to indicate the arguments of these papers. If these deficiencies of argument are left unattended, it could impede the emergence of a better analytical model for a proper understanding of risk allocation in PPP projects.

To set the stage for the JD Model in the third section of this paper, the rubric of transaction cost reasoning is elaborated in the second section. Those two sections are aimed at evaluating the JD Model against the TCE framework and discussing the implications for the development of a theoretical understanding of risk allocation in PPP projects in the future.

Nature of governance structures in TCE

According to Williamson’s own definition, governance structure is referred to as ‘the institutional matrix in which the integrity of a transaction is decided’ (Williamson, 1996, p. 378). In TCE empirical studies, the focus is normally placed on three common types of
governance structures: market, hierarchy, and hybrid. Why do governance structures matter in the comparison of efficacy of economic organizations? Unlike neoclassical economists who envisage trading can proceed in a frictionless world through complete contracting, transaction cost economists acknowledge the costs and differential competencies of governance structures in regulating transaction relationships. In principle, governance structures are meant to solve three fundamental problems: first, how to increase the total gain from trade by eliciting traders' best efforts; second, how to facilitate best response to economic changes during the contracting period; third, how to reduce the impact of disputes. Williamson (1991) posits that governance structures can solve these problems with varying efficacy (Williamson, 1991, p. 281):

1. Incentives and controls: The advantage of the market is that, under full exposure to market risks, traders have strong incentives to perform to their best. However, under hierarchy, suppliers become employees and the output of the workers is pooled, which more or less insulates them from bearing the full cost of shirking. To keep underperformance in check, the employer then needs to ramp up monitoring and motivate the workers through career progression. The choice between market and hierarchy is tantamount to a choice between high-powered incentive (market) and low-powered incentive (hierarchy).

2. Adaptability: In instantaneous market transactions, traders are under pressure to make timely and opportune responses to market changes in order to stay in the game. However, this so-called 'spontaneous adaptability' would be inhibited where transactions have duration and involve lump-sum investments in specific assets (Williamson, 1996, Chapters 4 and 6). Once specific investments are sunk, the investing party will forfeit some flexibility to switch the assets to an alternative use. The losses arising from redeployment of the assets would leave him/her in a position vulnerable to exploitation in renegotiations. When one party makes an opportunistic move, the other will not normally back down without resistance, making it inevitable for both of them to incur transaction costs in settling disputes. For transactions entailing specific investments, coordinated adaptability by way of authority offers a more efficient solution for holdup problems.

3. Contract law: TCE is built on the assumption of contract incompleteness, namely it considers the cases where unforeseen events might drift the contractual relationship out of the governable range and cause frictions between traders. The cost for resolving disputes varies with the regime of contract law to which disputes can be referred. In market transactions, disputes need to be resolved through a costly process (e.g. litigation, arbitration). By contrast, firms can settle internal disputes (say, disagreement between two departments of a company) in a more efficient way because of forbearance law that allows them to exercise fiat to reach final settlement without recourse to the courts as the final resort.

The strengths and weaknesses of governance structures in these three dimensions mean none of them commands an absolute advantage over the other two in efficiency terms, rendering it economical to align governance structures with transaction attributes (asset specificity, frequency and uncertainty). This is the fundamental reason why TCE has strong predictive power for real-life choices of governance structures. As a result, the extent to which governance structures under study could impinge upon the efficiency of transactions is a critical factor for the success of any attempt to apply TCE to a new context.
A synopsis of the JD Model
In Jin and Doloi (2008), risk allocation in PPP projects is conceptualized as an exchange of risk management responsibilities between the government and the project company. They argue that how much risk one party should take depends on the attributes of risk management responsibilities under negotiations at tender. These attributes are of four types: the trading counterparty’s capability in managing risks, prior trading history with the party (proxied as frequency), environment uncertainty and the party’s risk management commitment (tantamount to behavioural uncertainty of TCE). Different proportions of risk transfer from the government to the private sector are deemed as governance structures. Different degrees of risk transfer are treated as equivalent to ‘make’ (0%), ‘make and buy’ (between 0% and 100%) and ‘buy’ (100%) decisions, respectively. It is argued that the four types of attributes jointly determine the proportion of risks desirable for the government to transfer (which is a step corresponding to the alignment of governance structures with transaction attributes in TCE).

Three critiques of the JD Model
For TCE to work in the explanation of risk allocation decisions, the JD Model conceives of risk allocation as ‘a transaction of risk management responsibility between potential risk bearers’ (Jin and Doloi, 2008, p. 711) and posits that ‘different allocation strategies are in fact different governance structures of risk management’ (Jin and Doloi, 2008, p. 711). What this construct actually captures is ‘proportion of risk management task transferred from public to private partner for risk X’ (Jin and Doloi, 2008, p. 712). As this is the cornerstone of the model, a central question is whether risk allocation can qualify as a governance structure. With reference to the definition of governance structure in TCE and its role in transaction cost reasoning, the answer is in the negative.

The first issue is concerned with the choice of unit of analysis. In the JD Model, exchange of ‘risk management responsibility’ is treated as a transaction. It is true that PPP procurement involves an exchange of risk management responsibility and that there is consensus in government guidance (Office of Government Commerce, 2002, 2009; Her Majesty’s Treasury, 2006), practical guides (Simon et al., 1997; Institute of Civil Engineers, 1998; Project Management Institute, 2000; Walewski and Gibson, 2003) and textbooks (Yescombe, 2002; Smith et al., 2007; Winch 2010) that risks should be assigned to the parties who can best manage/control them. Risk-managing capability can affect risk-bearing cost and thus should be a key determinant of risk allocation. Misallocation of project risks therefore would damage value for money for the whole procurement. But does it mean ‘risk management responsibility’ is an appropriate unit of analysis? In TCE, ‘[a] transaction occurs when a good or service is transferred across a technologically separable interface’ (Williamson, 1996, p. 379). By this definition, it is not wrong to think of trading of ‘risk management responsibility’ as a transaction. However, an essential qualifier for a transaction being a unit of analysis is that it can be mediated by governance structures (markets, hybrids, hierarchies). As will be discussed in the second point, the so-called governance structures in the JD Model are actually the outcomes of negotiations for ‘risk management responsibility’ at tender, serving none of the theoretical functions of governance structures. Without genuine governance structures, this calls into question whether the unit of analysis has been well chosen.

Second, the focus of the JD Model is concerned with the agreed-upon assignment of
responsibility between traders for handling risk events. In the world of incomplete contracting, ex ante agreement may not be maintained ex post. The issue of concern should be the mechanism through which the loss/gain will be allocated between parties after risks eventuate. An essential condition for ‘risk management responsibility’ to be tradable is that one party’s promise of retaining responsibility can be enforced by a third party. Enforceability depends considerably on how well the risk management responsibility can be drafted into contractual clauses and linked to payment. Without a contract as the enforcement mechanism, evading the consequence of a risk event that one party agrees to take on ex ante will go unpenalized, making the ex ante promise no more than a goodwill gesture. Who eventually bears the impact of the risk depends not only on the pre-agreed allocation, but also on contractual arrangement, payment mechanism and financial risk management instruments used. The efficacy of a combination of these factors in securing the attainment of a pre-agreed risk allocation should be endogenized in the decision of risk allocation.

The third problem of the JD Model is associated with the application of the concept of asset specificity. It claims ‘the most important specified assets in any RM [risk management] service transaction would be the organizational RM capability’ (Jin and Doloi, 2008, p. 710). Jin and Doloi attribute asset specificity to two sources: risk management routine and risk management mechanism. In what sense can we regard risk management routine as a specific asset? The authors submitted ‘the longer the history of a RM routine (a transaction-specific asset), the more productive activities have been carried out, which means more alternative use has been made of the asset, and accordingly, the less specific the RM routines become’ (Jin and Doloi, 2008, p. 710). The risk management mechanism in the JD Model encompasses the four standard components of a risk management framework: risk identification, risk analysis, risk response and risk monitoring/control. The completeness and experience in the implementation of this system is labelled as ‘specific assets maturity’ (Jin and Doloi, 2008, p. 710). Odd terminology aside, a fundamental problem rests with failure to acknowledge the distinction between investment specific to an individual transaction and specific assets in general terms. The reason why assets specificity offers the greatest explanatory power for TCE reasoning is that its presence is correlated with the magnitude of transaction costs. The reasoning runs as follows. Some transactions require lumpy investments, which, once sunk, would lose most of their value, were they redeployed to an alternative use. The investing party has much more to lose than the other, were the transaction to stall, and thus would be subject to holdup demands when renegotiations arise. One party’s intention to take advantage of another’s vulnerability would prompt fierce disputes and, as a consequence, cause the ex post costs of maladaptation. In the presence of asset specificity, accommodating changes through the market becomes costly. In this regard, hierarchy has the advantage in that adaptation can be implemented relatively easily by way of administrative orders. The increase in specificity of investments in a transaction would tip the scale of comparative desirability from the market to the hierarchy at certain point. The causal link between asset specificity and choice of governance structures is a key to making TCE arguments refutable. When applying the same logic to examine two types of specific assets in the JD Model, we can easily identify the reasons why they do not qualify as cases of transaction-specific investment:

1. Generally speaking, these intangible assets are not invested specifically for a
particular PPP project. For example, risk management routine is measured by the public/private partner’s experience in managing risk X on a scale of 1 to 5, and risk management mechanism by the maturity of the public/private partner’s identification, analysis, response planning, and monitoring and control mechanisms for risk X. Without the ‘loss-of-value outside this transaction’ element of the investment, the link between asset specificity and holdup problems (and thus transaction costs) is feeble.

2. Even if, in rare cases, parties need to make investments in building risk management capability for a project, how significant would that investment be relative to the capital value of the project? Since in practice transferring risk to a third party (e.g. insurance company or financial risk management instrument vendors) is common, the investment really required seems rather trivial.

Discussion
As has been argued, the weaknesses of the JD Model indicate its application of TCE is flawed. Drawing on the RBV alone would make the model look theoretically sounder, but the downside is that it could then only manage to scratch the surface of the issue.

It seems useful to take a broader view on risk allocation in PPP procurement. Governments put value for money as top priority in negotiations with PPP investors, with the aim of striking an optimal balance between risk transfer and its cost through a proper design of payment mechanism. Incentives in PPP contracts are mainly created through the link between performance indicators set against output specification and variable payments tied to them.

While risk allocation is an important input in the value for money assessment, optimal risk allocation alone cannot guarantee value for money. Risk allocation patterns should be understood in the context of PPP procurement in its entirety. In a world of incomplete contracting, governance structures regulating the assignment of the impacts of risk events should warrant higher weighting than agreed-upon risk allocation in the explanation of risk allocation arrangement.

Conclusions and recommendations
The JD Model reviewed in this note is a credible attempt to enhance our understanding of risk allocation in PPP contracts. However, the model suffers from three deficiencies, namely: inappropriate choice of unit of analysis, poor specification of governance structure, and misinterpretation of asset specificity. Addressing these critiques is an essential step towards a better application of TCE to the study of PPP contracts in general, and risk allocation in PPP contracts in particular.

This research suggests improvements can be made in the future, in three ways:

1. Unit of analysis: Risk allocation is an integral part of PPP transactions. The optimality of risk allocation cannot be judged in its own right, but needs to be viewed in the context of value for money of the whole PPP project.
2. Operationalization of transaction attributes: PPP projects may entail specific investments, but, to correctly identify the source of asset specificity, researchers should consider taking a more rigorous approach. An example is demonstrated in Chang and Ive (2007) where a special form of asset specificity, process specificity, was discovered through examination of the evolution of the client’s and the contractor’s quasi rent during the construction period.
3. Nature of governance structures: A PPP transaction consists of a web of contractual relationships. A fundamental problem of the JD Model is that it treats PPP transactions as if they were insurance transactions. This methodological position is ill-conceived because it ignores the fact that PPP projects involve huge efforts and resources from the contractor to deliver output. In future studies, researchers should re-examine the nature of PPP governance. Transaction cost, payoff rights and property rights are all expected to offer important theoretical angles for understanding PPP governance.

Acknowledgement
I am grateful for the insightful comments from Graham Ive on the early drafts of this paper.

References


Institute of Civil Engineers (1998) RAMP: Risk Analysis and Management for Projects, Thomas Telford, Trowbridge.


---

1 The word ‘specified’ in this context seems to carry the same meaning as ‘specific’ in the TCE literature, but this usage should be avoided.

ii See Gibbons (2005) for an excellent comparison of these three perspectives.