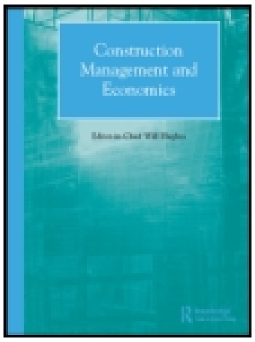

Construction Management and Economics

ISSN: 0144-6193 (Print) 1466-433X (Online) Journal homepage: https://www.tandfonline.com/loi/rcme20

Reversal of bargaining power in construction projects: meaning, existence and implications

Chen-Yu Chang & Graham Ive

Construction Management and Economics (August 2007) **25**, 845–855

Reversal of bargaining power in construction projects: meaning, existence and implications

CHEN-YU CHANG[1]* and GRAHAM IVE[2]

[1] *Bartlett School of Construction and Project Management, University College London, UK*
[2] *Bartlett School of Graduate Studies, University College London, UK*

Received 21 July 2006; accepted 9 December 2006

Reversal of bargaining power arising from asset specificity is important for the understanding of hold-up problems. Various types of asset specificity have been identified in different transaction contexts, but a previously unidentified or unnamed type is developed here: process specificity. Numerous widely used financial and contractual preventive measures in construction practice can be justified as responses to this problem of process specificity. These include bonds and retentions. However, these measures have limitations. Specifically, the client-led change orders cannot be completely averted and when they occur, the pricing of additional work largely relies on negotiation, implying that bargaining power determines the result. Consequently, the hold-up problem remains a managerial issue. To mitigate this problem, clients should choose a procurement system by aligning project attributes with the procurement system characteristics.

Keywords: Bargaining power, transaction cost, opportunism, asset specificity, procurement

Introduction

For many practitioners who have worked on the client side, it is not an unusual observation that change of mind after the signing of contracts may cost the client dearly (Chan and Yeong, 1995). The use of various different financial instruments, such as surety bonds, to protect client interests is also commonplace (Hughes *et al.*, 1998; Russell, 2000). For practitioners, these experiences might be taken for granted. However, for researchers these phenomena can be regarded as a stable equilibrium outcome of an evolutionary process within which participating parties in the construction project interact with each other. The exploration of these phenomena can reveal clues to understanding the peculiar economic nature of the construction process. Some prior attempts have been made to examine the contractual relationship of the construction process from the perspective of transaction cost economics (TCE) (Winch, 1989, 2001; Doree, 1997). This study aims to bring this line of inquiry forward by putting forth a model to explain the reversal of bargaining power between client and contractor between the pre-contract and post-contrast stages. Here bargaining power indicates the extent to which one party may yield to terms which benefit the other. Bargaining power reversal originates from the sunk cost of an investment specific to a transaction. The prevalent existence of this fact has been observed in a wide range of transactions, and has been well documented in the literature on transaction cost economics (Shelanski and Klein, 1995; Boerner and Macher, 2002). The purpose of this study is to demonstrate that this problem also exists in construction transactions in the form of process specificity and many widely used financial and contractual preventive measures can be justified as responses to this problem. However, client-led change orders cannot be entirely avoided. When they happen, the pricing of additional work may need to be determined through negotiation if it cannot be sorted in accordance with an agreed schedule of rates. How price negotiation will wind up largely depends on bargaining power. As a result, the hold-up problem is still a managerial concern. To mitigate this problem, the client should choose a procurement system by aligning project attributes with procurement system characteristics.

*Author for correspondence. E-mail: chen-yu.chang@ucl.ac.uk

ISSN 0144-6193 print/ISSN 1466-433X online
http://www.tandf.co.uk/journals
DOI: 10.1080/01446190601164113

Theoretical importance of bargaining power reversal

In the last decades, transaction cost economics has led to fascinating developments in the economic analysis of organizations (Williamson, 1985, 1996). The application of this theory has crossed the border of economics and permeated contiguous disciplines, including sociology, law and management, to name a few. There are two main reasons for the success of this theory:

(1) The use of the 'transaction' as the unit of analysis makes the theory widely applicable.
(2) This theory is able to identify the key determinants of transaction costs (in TCE terms, this is called operationalization), making the theory refutable.

Two key concepts in TCE are central to the argument of this paper about reversal of bargaining power: the idea that a 'fundamental transformation' occurs when a contract is signed, and the idea of asset specificity.

Prior to signature of a contract, most construction buyers face large numbers of potential suppliers, and most suppliers face large numbers of potential clients. Generally, some clients have opportunities to obtain resource-specific rents from their projects, while competition between contractors drives down contract prices to levels that contain no economic rent for the contractor. Thus the client appears to be able to appropriate the economic rent of the project (the excess of its value over its cost of resources used). The balance of bargaining power at this point can be said to lie with the client, therefore. Thereafter, post-contract, the situation for both actors changes to one of a degree of bilateral dependence. This fundamental transformation is not the same thing as (does not always give rise to) process specificity. It is however a necessary precondition for it. Process specificity derives from specific features of the construction process when performed under the contracting system (as opposed to the speculative development system, where the builder owns the site and the customer buys a completed building): differences between pre-contract and final design, deriving from the fact that design is endogenous to each construction project, and is developed iteratively through that process; the peculiar dependence of the value of a set of construction activities comprising a project on the completion of that set; the fact that the project comprises a process of building upon a fixed site, belonging to the client; and the fact that the client pays for work in process of construction, and does not only pay for a completed product. In combination, these features give the client a persistently weak bargaining position throughout post-contract stages.

Process specificity is a particular type of a more general phenomenon—asset specificity. Asset specificity is the most important of the three key determinants of transaction costs identified in the literature, both theoretically and empirically (Williamson, 1985). The explanatory power of asset specificity comes from a chain of reasoning:

(1) During the pre-contract stage, traders are free to choose where they want to allocate their resources, such as monetary capital, human resources and technology; meanwhile, the incentive attracting traders to go into a contractual relationship lies in the expectations of each party on the potential for non-negative profits from the transaction.
(2) If a specific investment is required at the outset of a transaction, the difference between the value of this investment and its value in alternative uses may be appropriated by the non-investing transactor to some extent during the post-contract stage.
(3) It is posited that the larger the value lost arising from the switching of the investment to the second best use, the more rent can be appropriated and the more likely that the vulnerability will be exploited.
(4) Opportunistic behaviour will intensify with the magnitude of increase in appropriable quasi rent, causing disputes and associated transaction costs (such as legal costs, schedule delays and additional preventive measures).
(5) Organizational forms (such as firm or market) play the role of scaling factors due to their differential competence in dealing with different types of transactions. That is, an appropriate governance structure can mitigate the severity of rent-seeking behaviour and thus reduce transaction costs; in construction the relevant forms of governance structure of a project are the types of procurement system. An incorrectly chosen procurement system may worsen the situation.
(6) The differential performance of an organizational form in governing different types of transaction leads to refutable hypotheses that can be used to test the robustness of transaction cost based reasoning in interpreting the selection of organizational forms.

TCE has been empirically demonstrated to provide a strong basis for explaining a multitude of issues regarding organizational choices such as vertical integration, franchise and long-term contract (Shelanski and Klein, 1995; Joskow, 1991). Asset specificity plays a key role in explaining the choice of organizational forms. However, construction contracting differs from general manufacturing production in four respects:

(1) *Discontinuous workload*: This takes two forms. First, within a single project, there is discontinuous demand for each trade. Second, while total output of a firm comprises only a small number of projects, the tendering system means a firm cannot predict how many projects it will win, or what the resource requirements of these projects will be. Over a set of projects total per-period demand for each resource is unstable and unpredictable. These are important factors explaining subcontracting, absence of vertical integration and low levels of physical and human asset specificity.[1]

(2) *Mobile production site*: In construction, there is no fixed production factory. Instead the production site changes with each project.

(3) *Less specific lump sum production investment at the outset of a project*: In current practice, most of the plant and machinery employed in construction projects are hired (or in any event are intended to outlast the project), and thus are generally designed for general purposes.

(4) *Endogenous design*: Construction is an order-to-build type production, meaning that each project has a unique design and that design itself is a part of the production process.

These differences make it legitimate to examine whether the key determinants of transaction costs identified in the TCE literature remain relevant to construction. Furthermore, if not all of these previously identified determinants are relevant, the question as to what determinants are important for transaction costs in construction should be explored.

Its existence

Basic set-up

The construction procurement route governs the process of organizing project participants. After the agreement is entered into, the estimated costs are gradually realized. A complete description of the dynamic process of a construction transaction comprises two parts: accounting based cost–benefit analysis, and opportunity cost based cost–benefit analysis. To be involved in a project coalition, all participants must spend ‘real’ resources in exchange for expected profits, while simultaneously losing the opportunity to switch already dedicated resources to alternative uses without incurring costs. These two systems are complementary, one showing how much money has been spent and earned, and the other indicating how wisely the economic value of the resources managed has been exploited. In principle, opportunity costs can help decision makers choose the correct means of achieving the goal of increasing accounting surplus. Moreover, the change of opportunity costs and accounting costs affects the interaction relation between trading parties in the course of the transaction.

Analytically, the condition under which the client is willing to initiate a project is the presence of non-negative economic rent at the outset, namely, the client rent (CR) is no less than zero.

$$\mathrm{CR} = \mathrm{V} - \mathrm{P} \geq 0 \tag{1}$$

where V denotes the value of the project to the client and P represents the total payment to the agents involved for covering their production costs. The delay problem faced by the client in the event of project disruption is a key concern (delay reduces realized V below expected V).

Client’s quasi rent in the construction process

The first scenario is that, were the project to be discontinued, then as a last resort the client alters the final use of the partly completed project. In this case, the quasi rent ($\mathrm{CQR_1}$) of the client is

$$\mathrm{CQR_1} = (\mathrm{V} - \mathrm{V_C}) - \mathrm{P_u} \tag{2}$$

where V_C denotes the return of the completed part of the project in its alternative best use (if left in that state), and P_u represents the payment for the uncompleted part of the construction work. Thus, equation (2) stands for the cost–benefit appraisal of the remaining part of the project. If CQR is positive, it is better for the client to complete the project (transaction) and obtain the intended project value, rather than leave the building part-finished, define a new project to use this asset, and thus change the building’s use (and value).

At this point, the original payment scheme can be divided into two parts: (1) payment made (P_m) and (2) payment still to be made (P_u). These two parts can be regarded as the client evaluation of the project in terms of costs. For ease of exposition, it is assumed that the ex ante economic rent of the client is zero namely,

$$\mathrm{V} = \mathrm{P_m} + \mathrm{P_u} \tag{3}$$

Substituting equation (3) into equation (2), we can get

$$\mathrm{CQR_1} = \mathrm{P_m} - \mathrm{V_C} \tag{4}$$

Normally, the value of a partly completed project is appreciably lower if the project is not completed to serve the planned function, i.e. $\mathrm{P_m} \gg \mathrm{V_C}$. Provided $\mathrm{CQR_1} > 0$, the client will stick to the original contract.

Thus, $P_m - V_C$ is a source of an appropriable rent. This means that the contractor can appropriate some of this rent, by renegotiating the price of P_u upwards, above C_u (the cost of completing the construction).

In the second scenario, the client may choose another option, namely finding a replacement contractor to take over from the original contractor. In this case, the quasi rent CQR_2 of the client equals the cost of switching, which can be defined as the additional cost for restoring the production flow. More exactly, the magnitude of process specificity can be expressed as:

$$CQR_2 = v \times (t_1 + t_2) + C_e + C_{rp} + C_d \quad (5)$$

where t_1 and t_2 stand for the time required to find a replacement agent and the additional time required for the replacement agent to complete the unfinished project; v is the average cost of time. Therefore, the first term indicates the potential loss associated with the opportunity costs of time in the event of the project being disrupted. Moreover, C_e denotes the extra costs due to repeated set-up on the construction site;[2] C_{rp} indicates the risk premium requested by the replacement contractor for the uncertain quality of work of the first contractor. Meanwhile, the final term, C_d, indicates the cost incurred due to difficulties in identifying liabilities between the original and replacement contractors, given the discovery of defects in the interface between them.

Thus, a measure of the lock-in effect on the client side can be defined as $CQR = \text{Min}\ \{CQR_1, CQR_2\}$.[3] Figure 1 graphically represents this function. CQR_1 is indicated by a vertical line, i.e. a constant because $P_m \gg V_C$ and, *at any point of time in the course of the project*, the corresponding P_m is exogenously given, only to do with the agreed payment schedule. In contrast, CQR_2 will change as design/technology characteristics of the project are endogenously determined, and thus

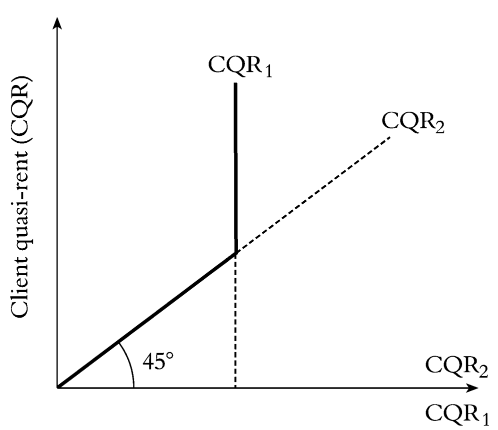

Figure 1 Client's quasi rent at one point of time during construction process

would range widely in value. When $CQR_1 \leqslant CQR_2$, $CQR = CQR_1$, i.e. it is a straight line going through the point of origin with a unity gradient. Moreover, when $CQR_1 \geqslant CQR_2$, $CQR = CQR_2$. For a project chosen based on sound feasibility study, scrapping a project halfway will be considerably more costly than replacing the contractor. Consequently, CQR_2 denotes the binding condition.

Contractor's quasi rent in the construction process

The contractor's hold-up problem involves fixed investment, which complicates the analysis. To facilitate the following analysis, a simplified form[4] of the (expected) contractor rent (BR) at the outset is expressed as

$$BR = P_C - \sum_{t=t_1}^{t_w} (VC_t + I_t) \quad (6)$$

where P_C denotes the total payment for undertaking the construction work; VC_t represents the variable cost, such as project overheads, wages paid to project personnel and payments to subcontractors; I_t is the lump sum investment for machinery and equipment or other durable resources. Finally, the subscript t indicates the cost incurred at the period of t, and t_w is the project duration.

At any point of time in the post-contract stage, such as t_1, the quasi rent of the contractor becomes

$$BQR = P_u - \sum_{t=t_1}^{t_w} (VC_t + S_t) \quad (7)$$

where P_u denotes the payment to be received (or unpaid by the client) and S_t represents the ex post opportunity cost of the lump sum investment. Assume the ex ante profit is zero and the payment is made on schedule, then the ex ante expectation of the payment to be received from t_1 to t_w becomes

$$P_u = \sum_{t=t_1}^{t_w} (VC_t + I_t) \quad (8)$$

Substituting equation (7) into equation (6) gives

$$BQR = \sum_{t=t_1}^{t_w} (I_t - S_t) \quad (9)$$

Equation (8) clearly shows that the contractor quasi rent comes from the difference between the ex ante and ex post opportunity cost of lump sum investment. Nonetheless, generally most high-valued equipment and machinery in construction are removable and not specific to particular projects, so BQR is not expected to be significant relative to the whole project. As a consequence, compared with the appropriable quasi rent on the client side (CQR_2), the contractor faces a considerably less severe lock-in problem.

Practice responding to hold-up problems in construction

The previous section demonstrates that in real world situations a post-contract bargaining power reversal will take place in the construction process, placing the client at a disadvantage. Following reasoning similar to that used in Klein *et al.* (1978), many measures taken by the client can be understood as responses to prevent this disadvantage being exploited.

Suppose that the client and contractor initially enter into a legal agreement without any financial protection. Since a contract cannot cover all possible contingencies, some clients will find that their vulnerability may be exploitable under certain circumstances, such as:

- the contractor refuses to rectify defects found in the course of the project;
- the contractor asks for additional payment due to inflation or other cost overruns and threatens to deliberately go bust (i.e. threatens ex post to shift inflation and cost risk on to the client);
- the contractor falls behind schedule and either refuses to speed up or demands extra payment due to the higher costs of overtime labour (i.e. threatens ex post to shift schedule risks on to the client);
- the contractor may pick out design errors and charge unreasonable rates for correcting them.

Once these problems become recurrent difficulties for the client, some protective actions will be taken. Practical solutions to these problems include:

(1) *Financial protection*: The client may demand that the contractor provide various bonding instruments, such as performance bonds, parent company guarantees or retention bonds;
(2) *Contractual protection* (Mitropoulos and Howell, 2003):

- legally 'force account' clauses provide the client with powers to direct the contractor to proceed and to pay only when the work is completed satisfactorily;
- the client may devise clauses preferable to their interests, such as the right to interpret design drawings;
- the client may include liquidated damages or early completion reward clauses to incentivize the contractor to deliver prompt or early completion.

Even when these protective measures are in place, the problems arising from the vulnerability of the client are still outstanding owing to three reasons:

(1) *Financial protection is not free*: If the construction project is treated as a system, the bonding cost can be considered rent dissipation (value added leaked to the bank), thus reducing productivity. Apparently, this cost falls upon the contractor, but these extra expenses will drive up the bidding price and be borne by the client, so its (insurance) cover must be limited.
(2) *Financial protection has limited efficacy*: Disciplining the contractor using financial protection may be less effective for three reasons:

- the contractor can minimize the negative effects of payment retention on cash flow by taking the strategy of paying subcontractors only after receiving payment. An analogy involving a reservoir can be developed. This situation resembles controlling outflow in response to inflow to keep a stable reservoir reserve. If the contractor can easily transfer the pressure of maintaining positive cash flow to subcontractors, the effectiveness of retention in disciplining the contractor will be weakened;
- to have positive cash flow as early as possible in the course of the project, the contractor is likely to 'frontload' the payment by setting the unit prices of cost items incurred during the early stage higher and evening out the unbalanced payment by reducing the unit prices of items incurred during the later stage. This strategy makes the contractor's received payments grow ahead of expended costs, implying that retention is not as effective as it appears under this condition;
- the strategy of payment retention is at best a pre-emptive instrument, and cannot totally prevent potential contractor opportunism. The reasoning is as follows. Retained payment is not money that the client is entitled to unilaterally decide whether to pay. In the event of disputes, the final resolution has to go through a pre-agreed mechanism, such as third party arbitration or the court. The key point is still whether the client can afford the losses arising from this process. If not, the bargaining power of the client deriving from holding a part of the payment due will lose a lot of its pre-emptive function.

(3) *Contractual protection is incomplete*: The incompleteness of contract comes from three reasons:

- costs of drafting clauses: only those contingencies that can be foreseen and described in a low-cost way can be stipulated in the contract, otherwise, the costs of contract drafting will be unacceptably high;
- enforcing clauses is costly: only the contingencies in which the third party has access to sufficient information can be enforced, narrowing the scope of the enforceable condition;
- a trade-off must be made between completeness and rigidity: the greater the detail in which the contract clauses are formulated, the less flexibility exists for the trading parties to accommodate subsequent changes. The contract is sometimes deliberately left incomplete to maintain flexibility for dealing with unanticipated events.

Contractual clauses thus may be able to seal some loopholes in the contract, but an alarming number of loopholes remain. Situations arising from client ignorance or omissions are particularly problematic. Classifying contingencies according to two dimensions, as follows, provides a means of understanding the nature of construction contracts:

(1) *Third party enforceability*: The effectiveness of third party dispute resolution depends largely on the existence of objective yardsticks. Disagreements on materials quality or escalating material costs are relatively easy to resolve since material quality standards and the extent to which material costs have risen are public information. In contrast, it is difficult to determine who should be responsible for the defects of completed work. The worse cases are those above the dotted line in Figure 2. For example, the third party arbitrator may have difficulty in reaching a verdict on the pricing of change orders arising from client-induced design errors as the situation often varies substantially. Similarly, if the client changes their mind during construction, it becomes difficult for the third party to determine compensation of the contractor for the additional work.

(2) *Information asymmetry between traders*: The more information that one party gets hold of relative to their counterpart over disputes, the more likely that the vulnerable party will be exploited. Information asymmetry in the area of material quality is relatively slight as the client can obtain information by random inspection. However, information asymmetry relating to escalated costs is more difficult to overcome since the terms and conditions of the procurement agreement between the contractor and the material supplier are normally a commercial secret. Moreover, a similar level of information asymmetry exists for completed quality, since the task undertaker invariably has more information about the quality of the work.[5] The worse cases are still associated with determining the fair price of additional work. The actual costs are always withheld by the contractor. The client may try to work out a baseline via comparison with previously completed similar work, but more often than not the benchmark is unreliable because factors such as design, site condition and business cycle can markedly change its reliability.

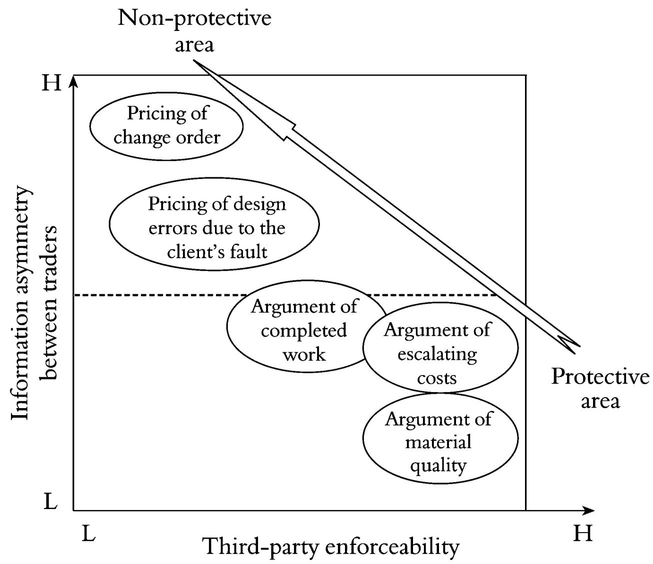

Figure 2 Indicative cases in construction where contract can protect and cannot (completely) protect

The right-lower corner of Figure 2 indicates the area in which contingencies can be protected by financial or contractual measures, while the contingencies located in the left-upper corner are outside the protective coverage of these measures. This area is also the main reason why hold-up problems continue to be managerial concerns in construction.

Some cases in practice

The existence of client post-contract vulnerability can also be examined in relation to cases involving renegotiation. The private cases are hard to obtain because the concessions made due to hold-up are likely to be hidden rather than revealed to the public. Fortunately, some of the disputes happening in high-profile infrastructure projects have either received extensive coverage in the press or have been reviewed by the UK government auditor, the National Audit Office, both of which provide ample facts for us to look at hold-up problems in construction. For the first category, the Channel Tunnel project is a case at issue. Chang and Ive (forthcoming) conducted a detailed case study to examine a major dispute between Eurotunnel, the concessionaire, and TML, a conglomerate of 10 contractors undertaking that project, claiming that disputes arising out of cognitive dissonance over the works unspecified in the contract can trigger opportunistic intent originating from process specificity and make the situation even worse. For the second category, a number of NAO reports that have investigated problematic government capital projects show that the UK government has encountered considerably more difficulties in securing good terms of trade than can be achieved during the pre-contract stage. In specific projects, such as the nuclear submarine facility construction project at Devonport (NAO, 2002), change orders have put the UK Ministry of Defence in a weak position to bargain down the huge extra payment claimed by the contractor. In certain cases, for example the renegotiation of the Royal Armouries Museum in Leeds (NAO, 2001a), the reversal of bargaining power extends to the operating stage. The government was forced to take back the responsibility of operating the loss-making parts of the project to avoid the default of the PFI company due to significantly lower than expected ticket revenue, even though the government was not exposed to the downside risk of revenue shortfall in the original contract. The worst case in terms of client post-contract vulnerability is evidenced in the Channel Tunnel Rail Link project (NAO, 2001b), where the government salvaged the crippled concessionaire, Continental Railways Limited (CRL), by giving this company further assistance in terms of loan facilities and government guaranteed bonds. While the government technically avoids increasing the present value of direct grants, the potential call on the loan support would be up to £400 million and there would be an enormous obligation of government backed bonds (up to £3.7 billion in total) in the event of CRL defaulting (NAO, 2005).

Its uniqueness

Types of asset specificity in the literature

In TCE, the lock-in effect may be caused by different types of asset specificity (Williamson, 1983):

(1) *Site specificity*: A 'cheek-by-jowl' relationship between the buyer and the seller is often based on a desire to save transportation and inventory costs. An example of asset specificity is the construction of a steel plant near an ore mine. Resorting to other mine supplies becomes much more costly after the factory begins construction, which may subject the plant to exceptional price increases from the mine owner.

(2) *Physical asset specificity*: For transactions involving a substantial lump sum investment in transaction-specific machinery or equipment, the producer can recover only a small part of that investment (i.e. salvage value) if the buyer terminates the transaction unilaterally.

(3) *Human asset specificity*: In transactions involving an investment in relationship-specific human capital through the learning-by-doing process, this investment is of little value when personnel with special expertise are transferred to other transactions.

(4) *Dedicated asset specificity*: Besanko *et al.* (1996) provide a useful definition for this term. They refer to 'an investment in plant and equipment made on behalf of a particular buyer' and therefore '[w]ithout the promise of that particular buyer's business, the investment would not be profitable'. The benefit of doing so is to tap economies of scale. That is, the assets themselves have no idiosyncratic qualities (unlike the cases of site and physical specificity), but the quantity of assets has been aligned for a specific transaction. The threat of orders being discontinued can leave the producer with excess capacity and thus in a weak position to bargain with the buyer.

These types of asset specificity are not directly relevant to transactions in the construction process. An item identified in Masten *et al.* (1991), temporal specificity, infuses new blood to this concept. Examining the distinctive features of construction operations revealed that 'delays become a potentially effective strategy for exacting price concessions' when 'the timely performance [of material suppliers] is critical' (Masten *et al.*, 1991). Put another way, the presence of hold-up problems does not necessarily involve the specificity of durable investments. As long as there are huge losses due to discontinuity of production flow, the party whose tasks are on the critical path in the production process will be in an advantageous position to increase the price to seize the rents generated by temporal specificity. A similar concept has been applied to analysis of the bulk shipping industry (Pirrong, 1993).

It is claimed that 'time and space factors in shipping markets may create "temporal specificity" that encourages costly haggling between shippers and carriers over quasi rents if they rely on spot markets'. In this case, only a vessel with special apparatus can transport certain commodities. If there is only one suitable ship in the nearby harbour, the sender may be subject to temporal hold-up problems, the severity of which depends on the transportation costs and time delay associated with finding another available vessel in the same harbour. In Hubbard (1999) an attempt is made to conduct a quantitative test of the relation between temporal specificity and contractual arrangements in the trucking industry, and concludes that 'the contractual implications of hold-up-based theories extend beyond circumstances where specific investments are large and sunk'. This assertion can bolster our confidence in identifying other possible sources of asset specificity.

Comparison

This section compare process specificity with other types of asset specificity identified in the TCE literature, and shows a case for its uniqueness. Seven types of asset specificity are scrutinized in Table 1, and provide a basis for understanding the sources of quasi rent in terms of options. Put explicitly, the vulnerable party has two options in responding to the hold-up demand: first, giving up the original transaction and switching the invested resources to the best alternative use with the loss L_1; second, finding a replacement trading party to complete the original transaction with loss L_2. A rational decision maker should select the option with the lower loss. The magnitude of this loss can be defined as a measure of degree of asset specificity (DAS). That is,

$$\mathrm{DAS} = \min\{L_1, L_2\} \tag{10}$$

Classifying seven types of asset specificity enables the identification of two types of asset specificity:

Type I: $L_1 > L_2$ (so that L_2 is the relevant measure of DAS)

This type of asset specificity originates from the increased costs associated with switching from the original trading partner to an alternative one. Site specificity, dedicated asset specificity, temporal specificity and process specificity are examples of this type of asset specificity.

Type II: $L_1 < L_2$ (so that L_1 is the relevant measure of DAS)

For this type of asset specificity, including physical asset specificity and human asset specificity, the losses result from switching dedicated resources to alternative uses. L_1 is much less than the loss resulting from finding a replacement trading partner L_2 because no other parties want to employ the resources in the same way, and thus L_2 would be relatively large. For example, in the case of General Motors and Fisher Body, Fisher Body needs to make an investment on dies and stamping machines specific to the requirements of GM (i.e. physical asset specificity). No other

Table 1 Sources of quasi rent arising from seven types of asset specificity

Types of asset specificity	Example	Vulnerable party	Sources of quasi rent
Physical asset specificity	Car assembler and car body manufacturer	Car body manufacturer	Value depreciation caused by switching the specific investment, including stamping machines and dies, to other uses
Site asset specificity	Ore mining company and steel plant	Steel plant	Increased costs in transportation and inventory when transportation and inventory costs associated with switching to alternative suppliers
Dedicated asset specificity	International Systems and IBM[a]	International Systems	Costs of maintaining excess capacity when the original transaction is cancelled
Human asset specificity	Salesperson and marketing company	Marketing company	Costs of training salespersons with company-specific knowledge in the event of experienced salespersons leaving the company
Temporal specificity—timing	Contractor and material suppliers	Contractor	Opportunity costs of delays caused by postponed materials delivery
Temporal specificity—space	Cargo sender and cargo carrier	Cargo sender	Opportunity costs of delay caused by finding a replacement cargo carrier in the neighbouring harbour
Process specificity	Construction client and contractor	Client	Increased costs of switching to a replacement contractor for resuming the disrupted project

Notes: [a] International Systems has invested in a production line for making integrated circuits for IBM. The value of this investment will be greatly reduced if IBM stops to make orders.

car manufacturers will buy this type of car body ($L_2 \approx \infty$), but the machinery still has a salvage value, thus reducing L_1 (for example as scrapped steel).

Another example is human asset specificity. For a marketing company where the on-job training of its salesmen is designed only for the products that the company is selling, the value of the investment in a salesman's training will be gone along with his leaving of the company (thus L_1 equals his/her training cost). Similarly, the knowledge of new salesman hired from outside will also be of little value to the company and thus hiring new salesmen will lead to additional costs of training (L_2). L_2 is slightly higher than L_1 as training normally becomes more expensive over time. For these two types of asset specificity, transaction termination means that the salvage value of the sunk investment will be recovered.

The following considers whether process specificity is intrinsically different from those types in the literature. The concept of process specificity is put forth to describe the interdependence in the construction process between client and contractor. The losses resulting from resuming a disrupted project are the ultimate source of process specificity. Switching costs are a general term for such losses. In fact, all types of asset specificity listed in Table 1 occur for this reason. The difference between them lies in the reasons for the losses. For type-II asset specificity, the value of the sunk investment may fall sharply if its original use changes, so continuing the transaction is critical to full recovery of the investment. For type-I asset specificity, the relevant option is to find a replacement trading partner to complete the transaction because it is more costly to give up the original use of the sunk investment (i.e., $L_1 > L_2$). However, various reasons can be identified by going further to examine why L_1 is so high for the type I asset specificity:

(1) *Sunk physical investment*: For site specificity and dedicated asset specificity, the desire to recover the sunk lump sum physical investment makes it worthwhile to bear additional costs (for example transportation and inventory costs in the former case and excess capacity in the latter case).

(2) *Embedded in a large transaction*: In a temporal-specific transaction, it is more desirable for the vulnerable party to find an alternative trading partner if the original partner demands excessive extra charges, because the ongoing transaction in which the held-up transaction is embedded is considerably more valuable than the held-up transaction itself. For example, for a cargo sender, the cost of transportation may be relatively trivial compared with the cargo value. Moreover, for a contractor, the value of the whole construction project normally significantly outweighs that of an individual building material or subcontract. Consequently, being embedded in a large transaction is the ultimate reason for making $L_1 > L_2$ in the case of temporal specificity.[6]

(3) *Built on a fixed site*: In the case of process specificity, the client will resort to an alternative contractor, in the event of the project being disrupted, because scrapping the project involving the unfinished building that has been fixed on site is much more costly. That is, if the original project is abandoned, not only does the building lose value, but so does the land (which has been diverted from other uses with higher value). The fact the building cannot be moved explains much of the size of $V - V_C$. The building as it is might correspond well to the requirements of some other client—but that client needs the building to be elsewhere.

The comparison conducted above demonstrates that the origin of process specificity is intrinsically different from that of other types of specificity because the product is fixed on land provided by the client. If the builder owned the land, it would develop the project at their own expense; the customer would then buy a completed building, and the problem of process specificity hold-up would not arise. The high costs of giving up on the project halfway are the ultimate reason why the client is in a vulnerable position.

Managerial implications

The presence of post-contract power reversal would trigger a series of rent-seeking behaviours when unforeseen events occur, in turn leading to the waste of (otherwise unnecessary) resources in dealing with unproductive activities. Improving efficiency is a crucial managerial task. Management needs to pay attention to this area, but the best means of doing this remains uncertain. The hold-up problem is based on a fundamental assumption that some unforeseen contingencies must exist. However, this assumption does not imply that nothing can be done. Though management suffers from limited ability to foresee the occurrence of all possible contingencies, it still appears reasonable to assume that an experienced decision maker should be able to predict the likelihood of occurrence of the most important contingencies. If this is the case, preventive actions should be taken. In nature, the possibility of post-contract renegotiation cannot be entirely avoided, so the best that can be done is to govern the project by an appropriate procurement system. Chang (2002) showed that three main categories of construction procurement systems—the traditional system (design-then-tender), design–build and management system, display differential competencies in dealing with the projects with different attributes. Alignment of project attributes with procurement systems is a means of mitigating the negative impacts of the hold-up problem. Another way is to embed one-off transactions into part of a series of transactions by partnering to build trust and enable the increased losses associated with behaving opportunistically to serve as a pre-emptive measure (Chang, 2006). Similarly, other trust-related factors, such as reputation, can be fitted into the model presented in this study. Incorporating these factors can enrich the transaction cost-based model and broaden the scope over which it can be applied.

Conclusion

Post-contract bargaining power reversal frequently occurs in the business world. It provides the advantaged party with a strong proclivity to appropriate quasi rent and provoke a bout of rent-seeking behaviour that increases transaction costs. The main purpose of this study is to present an account of why this problem also exists between clients and builders in construction projects developed under a contracting system, and explains why the origin of this problem differs from those identified in the general literature. In managing construction projects, this issue has to be taken seriously. Prudent selection of procurement system and the use of partnering can help mitigate the potential hazards arising from this problem. The rationale presented in this study gives procurement system selection a central role to play in practice and in the future research agenda.

Notes

1. 'The construction industry has some particular characteristics, which make it substantially different from other industries, especially manufacturing. The main distinctive feature of construction is the nature of the final product, characterized by its uniqueness, immobility and variety (Eccles, 1981; Masten *et al.*, 1991). First, construction output is elaborated as a project production system, so it can be adapted to the buyer, location and use. The uniqueness of the output does not recommend the use of specialized assets for each project … Second, … the productive assets have to move to the location of the product and not the other way around … The third particular feature is the wide variety of final products … A final remarkable feature of the construction industry is the economic importance of each unit of product. Each project or contract usually represents an important percentage of the transactors' operations … demand is of a discrete nature, the sum of only a few projects' (Gonzalex *et al.*, 2000).
2. This cost includes the costs of installing construction machinery, administrative facilities (for example a temporary office), management hardware (for example computers), and so on.
3. A formal model developed by MacLeod and Malcomson (1993) is based on a similar idea. 'The contract specifies a price and, possibly, breach penalties. At stage 1 each agent chooses a level of investment … at stage 2 realization of a random variable … determines both the value of trade between the two parties and their alternative trading opportunities … During stage 3, trade may occur. Agents may trade at the contract price (or) refuse to trade at all or contract to trade with a third party instead, either of which may involve breach penalties … They may also renegotiate … but this requires both parties to agree … trade occurs over time rather than at a single fixed date. This … is important analytically. It allows delay to be a credible threat, which alters the bargaining equilibrium and has interesting implications for the form contracts take.'
4. As discounting does not affect the analysis result, it is not considered for the purpose of simplification.
5. Unlike in the case of material quality, random inspection cannot effectively eliminate information asymmetry regarding construction quality due to two reasons: (1) some inspections must be performed in time, i.e. before the next stage starts; otherwise the quality will be very difficult to check afterwards; (2) the grey area is greater for construction work quality than for construction material quality.
6. Due to the apparently close similarity between process specificity and temporal specificity, further accounts are given to clarify their differences:
   (1) In Masten *et al.* (1991), temporal specificity takes place in the transaction between contractor and building material supplier, while process specificity is for the transaction between contractor and client.
   (2) Temporal specificity tries to capture the threats caused by delaying critical path activities. Timing is critical, so Masten *et al.* use temporal to describe this concept. However, process specificity tries to capture the persistent weak bargaining position of the client at post-contract stage.
   (3) The exploitation of temporal specificity is mainly about timing, while the vulnerability of process specificity won't be a real threat unless the events in the non-protective area of Figure 1 occur.

References

Besanko, D., Dranove, D. and Shanley, M. (1996) *The Economics of Strategy*, John Wiley & Sons, London.

Boerner, C.S. and Macher, J.T. (2002) Transaction cost economics: an assessment of empirical research in the social science. Mimeo, Hass School of Business, University of California, Berkeley, CA.

Chan, A.P.C. and Yeong, C.M. (1995) A comparison of strategies for reducing variations. *Construction Management and Economics*, **13**, 467–73.

Chang, C.Y. (2002) An economic interpretation of construction procurement behaviour for commercial and industrial buildings. PhD dissertation, Bartlett School of Graduate Studies, University College London.

Chang, C.Y. (2006) An empirical investigation of construction procurement system selection: a transaction cost explanation. Working Paper, The Bartlett School of Construction and Project Management, University College London.

Chang, C.Y. and Ive, G.J. (forthcoming) Process specificity: the case of the Channel Tunnel. *International Journal of Project Management*.

Doree, A.G. (1997) Construction procurement by Dutch municipalities. *Journal of Construction Procurement*, **3**(3), 78–88.

Gonzalex, M., Arrunada, B. and Fernandez, A. (2000) Causes of subcontracting. *Journal of Economic Behaviour and Organization*, **42**(2), 167–87.

Hubbard, T.N. (1999) How wide is the scope of hold-up-based theories? Contractual form and market thickness in trucking. NBER Working Paper 7347.

Hughes, W., Hillebrandt, P. and Murdoch, J. (1998) *Financial Protection in the UK Building Industry: Bond, Retention and Guarantees*, E & FN Spon, London.

Joskow, P. (1991) Asset specificity and the structure of vertical relationships: empirical evidence. *Journal of Law, Economics, and Organization*, **4**(1), 95–117.

Klein, B., Crawford, R.A. and Alchian, A.A. (1978) Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics*, **21**, 297–326.

MacLeod, W.B. and Malcomson, J.M. (1993) Investments, holdup, and the form of market contracts. *American Economic Review*, **83**(4), 811–37.

Masten, S.E., Meehan, J.W. and Snyder, E.A. (1991) The costs of organization. *Journal of Law, Economics and Organization*, **7**, 1–25.

Mitropoulos, P. and Howell, G. (2003) Closure to 'model for understanding preventing and resolving project disputes'. *ASCE Journal of Construction Engineering and Management*, **129**(2), 232–3.

NAO (2001a) *The Re-negotiation of the PFI-type Deal for the Royal Armouries Museum in Leeds*, HC103, Session 2000–01, The Stationery Office, London.

NAO (2001b) *The Channel Tunnel Rail Link*, HC302, Session 2000–01, The Stationery Office, London.

NAO (2002) *The Construction of Nuclear Submarine Facilities at Devenport*, HC638, Session 2002–3, The Stationery Office, London.

NAO (2005) *Progress on the Channel Tunnel Rail Link*, HC77, Session 2005–06, The Stationery Office, London.

Pirrong, S.C. (1993) Contracting practices in bulk shipping markets: a transaction cost explanation. *Journal of Law and Economics*, **36**(2), 937–76.

Russell, J.S. (2000) *Surety Bonds for Construction Contracts*, ASCE Press, Reston, VA.

Shelanski, H.A. and Klein, P.G. (1995) Empirical research in transaction cost economics: a review and assessment. *Journal of Law, Economics, and Organization*, **11**(2), 335–61.

Williamson, O.E. (1983) Credible commitments: using hostages to support exchange. *American Economic Review*, **73**, 519–40.

Williamson, O.E. (1985) *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, The Free Press, New York.

Williamson, O.E. (1996) *The Mechanisms of Governance*, Oxford University Press, New York.

Winch, G. (1989) The construction firm and the construction project: a transaction cost approach. *Construction Management and Economics*, **7**, 331–45.

Winch, G. (2001) Governing the project process: a conceptual framework. *Construction Management and Economics*, **19**(8), 799–808.