An Economic Interpretation of Construction Procurement Behaviour for the Commercial and Industrial Buildings

Chen-Yu Chang

Bartlett School of Graduate Studies
University College London

Submitted in fulfillment of the requirements for the degree of Doctor of Philosophy of University of London

December 2001
Abstract

This study applies the rationale of transaction cost economics to develop a positive theory for interpreting and predicting construction procurement behaviour. The client is assumed to be a decision-maker who chooses the optimal procurement method so as to maximise his/her profit function, i.e. (discounted expected) revenue minus the sum of transaction costs and production costs.

Three stereotypical and most popularly used forms of procurement methods are reviewed: traditional method (TM), design and build (DB) and management system (MS). By way of transaction cost reasoning, this study discovers that, with the assumption of fixed production costs across routes, revenue from the project, transaction costs from process specificity and from measurement difficulties constitute three pillars for comparing procurement methods. Moreover it is further claimed that it is to the presence of quasi rent, quality rent and design rent that we can ascribe the reasons why there are differentials in transaction costs between procurement routes.

From the theoretical view, this study puts forth the principle of an inconsistent trinity, according to which, to determine the most efficient procurement method, the client will inevitably face the trade-off between (1) fast delivery of the project, (2) single point of responsibility for design and construction (lower measurement costs) as well as (3) high flexibility in accommodating variations (lower transaction costs arising from process specificity).

Empirically, six variables are pointed out as relevant in affecting the client’s decision: (1) the client’s opportunity cost of time; (2) uncertainty of requirements; (3) costs of switching supplier; (4) the use of partnering or not; (5) the scale of the client; (6) degree of specialisation. With the help of a Logit model, six hypotheses relating to these variables are tested, as predictions regarding relative frequency of selection of procurement routes for projects of given sets of attributes.

The analysis of this study aims to provide a new way of thinking about the nature of the project coalition, risk management and the development of tools for aiding the client in selecting the right procurement route.
Definition Of Key Construction Terminologies In This Study

(1) **The client**: refers to the investor who has the right to determine the use of the land for serving a perceived market niche and to initiate a construction project to exploit this opportunity. Clients are divided into industrial and commercial.

(2) **Construction procurement**: refers to the construction client’s purchase of a capital project (specifically, commercial and industrial buildings).

(3) **Procurement route, procurement system**: used interchangeably in this study, stands for a mode of coordination, or a contractual arrangement by which the liabilities and rewards of all the client’s agents are defined. Three stereotypical types of procurement systems to be discussed are traditional method, design and build and management system.

(4) **Traditional Method (TM)**: by which the client has to retain an architect to complete the drawings and specifications of the project and then devolve the task of construction to the main contractor.

(5) **Design and Build (DB)**: by which the client elicits construction proposals (including detailed design and construction plan) drawn up to conform to scheme design by pre-qualified DB contractors and then devolve the task of detailed design and construction to a single most desirable D&B contractor.

(6) **Management System (MS)**: by which the client has to retain a construction management firm (management contractor), who will be in charge of coordinating the remaining participating parties of the project, including designer and trade contractors, and retain an independent designer, then divide the project into work packages and tender these sequentially so that design and construction can proceed in parallel.

(7) **Trade contractors (or work package contractors)**: the undertakers of construction work undertakers in the management system.

(8) **Subcontractors**: the undertakers of construction work undertakers in the traditional method as well as design and build; this name shows the relative relation between main contractor and sub-undertakers, i.e., the former sub-letting the construction work to the latter.
List of Acronyms and Abbreviation

CM: construction management

COMSP: variable standing for “COMplexity of SPspecialisation”

CV: consequence variable

DB: design and build

DEGPS: variable standing for “DEGree of Process Specificity”

DMA: direct measurement approach

GC: items of production costs in the grey area (cost items that are hard to be classified into production cost or transaction cost)

GS: governance structure

IMA: indirect measurement approach

MAUA: multi-attribute utility approach

MC: management contracting

MS: management system

OPPCT: variable standing for “OPPortunity Cost of Time”

PARTR: dummy variable standing for whether PARTneRing agreement is used or not

SCALC: variable standing for “SCALe of the Client”

TC: transaction costs

TCA: transaction cost approach

TCE: transaction cost economics

TM: traditional method

UNCER: variable standing for “UNCERTainty”
Table of Content

CHAPTER 1  INTRODUCTION

1.1 Motivation of study ................................................................. 1
1.2 Objective of study ........................................................................ 2
1.3 The problem, its significance and its scope ................................. 3
  1.3.1 The problem to be sorted out .............................................. 3
  1.3.2 The significance of the problem ........................................... 5
  1.3.3 The scope of the problem .................................................. 7
1.4 Orientation of this study ........................................................... 8
1.5 Plan of this study ....................................................................... 9

CHAPTER 2  STATE OF ART IN THE STUDY OF CONSTRUCTION

PROCUREMENT BEHAVIOUR .......................................................... 12

2.1 The guideline approach ............................................................. 13
  2.1.1 Introduction ........................................................................ 13
  2.1.2 A critical assessment ........................................................... 15
2.2 The statistical approach ............................................................. 16
  2.2.1 Introduction ........................................................................ 16
  2.2.2 A critical review ................................................................. 19
2.3 Multi-attribute utility approach .................................................. 20
  2.3.1 Fundamentals of multi-attribute utility approach ................. 20
  2.3.2 Multiple-attribute utility approach for procurement route selection .... 22
  2.3.3 Where’s the problem? ......................................................... 25
  2.3.4 Conclusions of Sec.2.3 ......................................................... 36
2.4 Multidisciplinary perspectives .................................................... 37
2.5 Ending remarks of this chapter: where is the niche of this study? ........ 38

CHAPTER 3  AN INTRODUCTION TO TRANSACTION COST

REASONING .................................................................................... 41

3.1 The genesis of transaction cost analysis ...................................... 41
  3.1.1 The rationale of Coase(1937) .............................................. 41
  3.1.2 The methodology in Coase(1937) ......................................... 42
  3.1.3 Some reflections on Coase(1937) ......................................... 45
3.2 Discrete structural analysis ......................................................... 46
  3.2.1 A general illumination ....................................................... 46
  3.2.2 A Methodological Comparison of the DMA and IMA ............ 48
3.3 What Are Transaction Costs? .................................................... 53
  3.3.1 Why transaction costs exist at all? ...................................... 53
3.3.2 The concept of transaction costs .................................................................55
3.3.3 A Systematic Presentation Of Second-Round Information Problems ..........57
3.3.4 The nature of TC1 and TCII .................................................................58
3.3.5 The ways of modelling transaction costs ..................................................60
3.3.6 Which Way To Go: DMA V.S. IMA .......................................................62

3.4 Concluding remarks ..................................................................................64

4 THE DOCTRINE OF TRANSACTION COST ECONOMICS ......................65
4.1 Williamson’s three-stage comparative institutional analysis .......................65
4.2 The asset specificity approach .....................................................................67
   4.2.1 The unit of analysis: ‘transaction’ .........................................................67
   4.2.2 A four-step scheme for operationalising transaction costs ....................68
   4.2.3 Behavioural assumptions interaction structure in TCE ............................69
   4.2.4 Interaction structure in TCE .................................................................70
   4.2.5 Itemisation of transaction costs: a process perspective .........................71
   4.2.6 Operationalisation ...............................................................................72
4.3 The measurement approach .........................................................................74
   4.3.1 Alchian and Demsetz’s theory of team ..................................................74
   4.3.2 Measurement costs and organisations ....................................................75
4.4 Comparison of asset-specificity and measurement approach .....................76
   4.4.1 Ontological differences .........................................................................76
   4.4.2 Why measurement costs are downplayed in TCE .................................77
4.5 Dimensionalisation of governance structures ............................................81
4.6 Alignment of governance structures with transaction dimensions .............83

5 AN EXAMINATION OF THE BASIC PREMISES OF TCE .......................86
5.1 Assumption 1: Invariability of revenue across GSs .....................................86
5.2 Assumption 2: Sunk costs of the lump-sum investment ...............................87
5.3 Assumption 3: Exogenous design ...............................................................90
5.4 Assumption 4: Invariability of production costs .........................................91
   5.4.1 Three lines of critiques .........................................................................92
   5.5.2 The analysis .........................................................................................92
5.5 An examination of the basic premises of TCE in the context of construction
   procurement .....................................................................................................96
   5.5.1 Three stereotypical categories of procurement systems .......................96
   5.5.2 The assumption of revenue invariability ..............................................99
   5.5.3 The assumptions of irreversibility of lump-sum investments and continuous
       production .......................................................................................................99
5.5.4 The assumption of exogenous design ...................................................... 101
5.5.5 The assumption of invariability of production costs ............................. 103

5.6 A summary of chapter five ............................................................................. 106

CHAPER 6 A THEORY OF CONSTRUCTION PROCUREMENT BEHAVIOUR ................................................................. 109
6.1 An overview of the theory ........................................................................... 109
   6.1.1 The basic logic ......................................................................................... 109
   6.1.2 A brief overview of the model ................................................................. 110
6.2 The process perspective in the context of construction ......................... 111
   6.2.1 Context I: looking forward ................................................................. 112
   6.2.2 Context II: dynamic process of construction ...................................... 113
   6.2.3 Context III: equilibrium process ......................................................... 121
   6.2.4 Why process perspective? ................................................................. 122
6.3 Efficiency of procurement systems .............................................................. 124
   6.3.1 The nature of construction transaction .............................................. 124
   6.3.2 Efficiency of procurement systems .................................................... 125
6.4 Operationalisation of construction transaction(I): revenue .................... 127
6.5 Operationalisation of construction transaction(II): transaction cost ...... 128
   6.5.1 Identification of transaction costs in construction ............................ 129
   6.5.2 The basic reasoning of operationalising transaction ....................... 133
   6.5.3 A transaction cost theory of exchange ............................................. 134
   6.5.4 Process specificity in construction .................................................... 142
   6.5.5 Quality rent and design rent in construction .................................... 148
   6.5.6 The interaction effect of quasi rent, quality rent and design rent ..... 150
6.6 Dimensionalisation of procurement routes ............................................... 151
   6.6.1 Delivery time ....................................................................................... 153
   6.6.2 Adaptability to client-originated design changes .............................. 154
   6.6.3 Degree of efficacy in assuring construction quality ....................... 159
   6.6.4 Implications ......................................................................................... 160
6.7 Determination of the efficient procurement route .................................... 163
   6.7.1 Opportunity cost of time ................................................................. 163
   6.7.2 Degree of process specificity ......................................................... 166
   6.7.3 Measurement difficulties ............................................................... 167
   6.7.4 Concluding remarks of this section ................................................. 169
6.8 Conclusion .................................................................................................. 171
CHAPTER 7 EMPIRICAL INVESTIGATIONS OF CONSTRUCTION

PROCUREMENT BEHAVIOUR .................................................................................172

7.1 Preview of empirical investigations ..........................................................172

7.2 In search of observable variables ...............................................................173

7.2.1 The principle of search ...........................................................................173

7.2.2 An example in the literature ...................................................................174

7.2.3 Factors affecting revenue difference ......................................................174

7.2.4 Factors affecting transaction costs from process specificity (I) - DEGSP ..175

7.2.5 Factors affecting transaction costs from process specificity (II-1) - UNCER
................................................................................................................................176

7.2.6 A digression: the role of trust .................................................................176

7.2.7 Factors affecting transaction costs from process specificity (II-2): partnering
agreement ...............................................................................................................178

7.2.8 Factors affecting transaction costs from process specificity (II-3): prospect of
future works ...........................................................................................................180

7.2.9 Factors affecting transaction costs from measurement difficulties: COMSP
...............................................................................................................................181

7.3 The empirical methodology: reduced-from analysis ...............................181

7.4 Conditions that three procurement systems are best suited ....................182

7.4.1 The analysis ...............................................................................................182

7.4.2 Uncertainty ...............................................................................................182

7.4.3 Partnering agreement ...............................................................................184

7.4.4 The scale of the client ...............................................................................185

7.4.5 Summary ....................................................................................................186

7.5 Development of Hypothesis .....................................................................186

7.5.1 Average opportunity costs of time (OPPC) ............................................186

7.5.2 Uncertainty (UNCER) .............................................................................187

7.5.3 Cost of switching (DEGSP)......................................................................187

7.5.4 Partnering agreement (PARTR) ................................................................187

7.5.5 The scale of the client (SCALC) ...............................................................188

7.5.6 Complexity of specialisation (COMSP) ..................................................188

7.6 Testing of hypotheses ................................................................................188

7.6.1 Econometric model ..................................................................................188

7.6.2 A brief account of logit model .................................................................190

7.6.3 How to verify hypotheses by way of logistic model ...............................194

CHAPTER 8 DEVELOPMENT OF QUESTIONNAIRE AND SURVEY DESIGN
10.2 The contributions of this study for TCE theorists ............................................. 240
10.2.1 Process specificity ............................................................................................ 241
10.2.2 Design rent ....................................................................................................... 242
10.2.3 Quality rent ...................................................................................................... 245

10.3 The contributions of this study for construction management theorists.. 245
10.3.1 Set out a complete framework for the application of TCE to construction 245
10.3.2 Place the role of the client at the central stage of analysis ............................ 246
10.3.3 Provide a new paradigm for analyzing the construction process ............... 248

10.4 Implications of this study .................................................................................. 251
10.4.1 Procurement route selection as strategy of first-order economising .......... 251
10.4.2 Implications for construction risk management .......................................... 254
10.4.3 Implications for the procurement route selection techniques .................... 255
10.4.4 Implications for the study of construction disputes .................................. 255
10.4.5 Implications for the philosophy of construction management ................. 256

10.5 Remarks on the limitations of this study ....................................................... 257
10.5.1 The assumption of ex ante efficiency ............................................................. 257
10.5.2 Lack of better measures for variables ............................................................. 258
10.5.3 The effect of TC₁ on TC₁I ............................................................................... 259

10.6 Suggestions for policies .................................................................................... 260
10.7 Suggestions for further studies ........................................................................ 262
10.7.1 Organisational studies in construction ......................................................... 262
10.7.2 Improve procurement system selection technique ........................................ 263

Bibliography ............................................................................................................. 271

Appendix: Questionnaire ......................................................................................... 280
Catalogue Of Figures

Fig. 1-1 Coevolutionary Process of the Construction Contracting Markets .......................5
Fig. 1-2 Research Agenda for the analysis of construction procurement behaviour ..........11
Fig. 2-1 The reasoning of multi-attribute utility approach ...............................................22
Fig. 2-2 The reasoning structure of multi-attribute approach for procurement route selection ..........................................................26
Fig. 2-3 An account of the elements in MAUA in the context of procurement route selection ..........................................................28
Fig. 2-4 The relation between consequence variables in MAUA literature and the goals of construction management .........................................................30
Fig. 2-5 The roles of three consequence variables .............................................................31
Fig. 2-6 Procurement route research in three contexts .....................................................37
Fig. 2-7 The evolving path of the methods used in the analysis of construction procurement routes ................................................................................39
Fig. 3-1 The value chain of a production process ..............................................................42
Fig. 3-2 Transaction costs savings due to the emergence of the firm ................................44
Fig. 3-3 The determination of the boundary of a firm .........................................................44
Fig. 3-4 The flow chart of implementing DMA and IMA ..................................................51
Fig. 3-5 A structural presentation of information problems in a transaction: .................58
Fig. 4-1 Three-stage Scheme for Comparative Organisational Analysis .......................66
Fig. 4-2 Logic of asset-specificity branch of transaction cost theory ...............................76
Fig. 4-3 Logic of measurement branch of transaction cost theory ....................................77
Fig. 4-4 Buy or make decision ..........................................................................................79
Fig. 4-5 Comparative Production and Governance Costs .................................................85
Fig. 4-6 The Production Chain of the final product ..........................................................87
Fig. 5-2 The benefit and cost flow of a capital goods investment ....................................89
Fig. 5-3 Three types of production costs ..........................................................................93
Fig. 5-4 Comparative Production and Governance Costs .................................................96
Fig. 5-5 (a) Traditional system .........................................................................................97
Fig. 5-5 (b) Design and build ............................................................................................97
Fig. 5-5 (c) Management system .......................................................................................97
Fig. 5-6 Contractual structures within three procurement routes .................................98
Fig. 5-7 Two types of decision sequences of design and construction ............................102
Fig. 5-8 The effect of economies of scale .......................................................................105
Fig. 6-1 The trilateral relation between procurement routes, decision function and observable factors ...............................................................109
Fig. 6-2 A modified trilogy for the analysis of construction procurement routes ..........110
Fig. 10-4 Two modes of specialisation between designer and contractor ...............243
Fig. 10-5 The client’s learning mechanism for procurement system selection ........247
Fig. 10-6 The contractor’s learning mechanism for organisation selection ..........248
Fig. 10-7 The logic of claiming procurement system selection as strategy ...........253
Fig. 10-8 A Measure of project risk ........................................................................259
Fig. 10-9 The effects of trust on the selection of governance structures ...............261
Fig. 10-10 The further development of the theory espoused in this study ..........263
Catalogue Of Tables

Table 2-1 Conditions to which procurement routes are most suited .................. 14
Table 2-2 The measure of variables in Molenaar and Songer (1998) .................. 17
Table 2-3 Documented priority variables affecting the client’s decision on 
procurement systems ....................................................................................... 23
Table 2-4 A simplified version of procurement route evaluation table in Love et. al. 
(1998) .............................................................................................................. 24
Table 2-5 Comparison of delivery time through three procurement routes .......... 32
Table 2-6 The client’s desired consequences and their proxy variables ............ 33
Table 3-1 Comparison of two classical cases of transaction cost analysis .......... 45
Table 3-2 Items of transaction costs in the different stages of transacting ......... 55
Table 3-3 A classification system of transaction costs ....................................... 59
Table 3-4 A Summary table for the comparison of $TC_I$ and $TC_{II}$ .............. 61
Table 3-5 Comparison of advantage and disadvantage of the DMA and IMA .... 63
Table 4-1 A comparison of search costs and measurement cost influenced by 
integration .......................................................................................................... 80
Table 5-1 Relevance of three lines of attack to the three sets of alternative governance 
structures ........................................................................................................... 106
Table 5-2 Comparison of five assumptions in TCE as applied to construction 
procurement system selection .......................................................................... 107
Table 5-3 The modifications of TCE for the analysis of construction procurement 
systems .............................................................................................................. 108
Table 6-1 A comparison of process perspective and contract perspective in the 
analysis of construction procurement systems .............................................. 123
Table 6-2 A comparison of measurement cost in three procurement systems .... 130
Table 6-3 Payoff matrix of buyer and producer’s strategy selection ................. 141
Table 6-4 Sources of quasi rent arising from seven types of asset specificity ...... 146
Table 6-5 Comparison of procurement systems in terms of contractual characteristics 
.................................................................................................................................. 151
Table 6-6 Comparison of delivery time through three procurement routes ....... 154
Table 6-7 The effect of contractual arrangement on process specificity ............ 156
Table 6-8 A comparison of adaptability to design changes of procurement systems 158
Table 6-9 Micro-economic characteristics of three main procurement routes .... 160
Table 6-10 Micro-economic features of main procurement routes .................. 160
Table 7-1 The frequency distribution of a hypothetical example ....................... 192
Table 7-2 The actual observations and predicted results of a hypothetical example 193
Table 7-3 Expected results of logit estimation ................................................. 195
Table 7-4 Expected results of logit estimation ................................................................. 195
Table 7-5 Expected results of logit estimation ................................................................. 195
Table 8-1 The measurement of variables in TCE .......................................................... 200
Table 9-1 Proportion of procurement routes used in the sample projects ................. 217
Table 9-2 A basic framework for construction organisation strategy ....................... 218
Table 9-3 Descriptive Statistics of the years of experience of the respondents ........... 219
Table 9-4 An example of a normalised variable .............................................................. 221
Table 9-5 An illustration of the calculation of Cronbach’s alpha ................................. 222
Table 9-6 Basic statistics and distribution of dependent variables ............................... 223
Table 9-7 The estimation of coefficients in Eq.9-3~Eq.9-6 .......................................... 224
Table 9-8 A summary of empirical results ................................................................. 228
Table 9-9 The test of fitness of the model specification ................................................ 225
Table 9-10 The test of fitness of the model specification ........................................... 230
Table 10-1 An assessment of three stereotypical procurement systems ..................... 233
Table 10-2 A comparison of dimensions of three procurement systems .................... 235
Table 10-3 Comparison of three types of asset specificity ......................................... 242
Table 10-4 Comparison of two modes of specialisation between design and construction ........................................................................................................ 244
Table 10-5 A Comparison of Process Perspective and Contract Perspective .......... 250
Table 10-6 Comparison of strategy of governance and strategy of contract .............. 252
Table 10-7 A summary table for the comparison of three approaches ....................... 270
Chapter 1 Introduction

1.1 Motivation of Study

This study arises out of my long-term curiosity about what are the fundamental principles governing construction behaviour. Despite construction's long-standing development in the human history and its indispensable role in any economy, we still lack a core theory that can be used as a foundation for spawning an all-encompassing theory suitable for the (microeconomic) characteristics of construction [Ofori, 1994]. Notwithstanding a large number of researches under the heading of economics of building, economics of construction and the like, their goals, in the main, focus on some aspects of engineering economic problems, the relationship between construction activities and the whole economy, input-output analysis of construction production, or how to apply neoclassical economics to deal with economic behaviour in construction, e.g., Hillebrandt, 2000. Just as Williamson's breakthrough contributions originate in his dissatisfaction with the state of the art of organisation theory, this study claims that construction economics should be restructured in a more systematic way that reflects the very reason why economics is enlarging its territory [Coase, 1978]. To wit, this study highlights the necessity of generating a holistic view on the functioning of construction industry. However, how to select a proper instrument from the toolbox of economics? Presumably the best policy for head-on confrontation with this thorny issue is to re-examine the nature of the problem, and then decide the point of departure.

Simply speaking, construction economic activities include the set of human economic behaviour engaging in construction. Two key terms need to be decomposed for further analysis: human economic behaviour and construction. First of all, in term of methodological individualism [Hodgson, 1986], every economic agent is an entity with autonomy of decision. All of the (associated) intentional decisions will interact under an institution, e.g. market, and result in unintentional consequences. The exploration of this phenomenon is considered as the overriding task of a social scientist [Popper, 1965]. Following this line of exploration, two issues appear to be of foremost relevance: behavioural assumptions; and the relation between individual decisions and an institution. The signifi-
cance of the former lies not only in its role in model building, but also in changing the ontological nature of subject matter. The latter is concerned with the interaction between system and its subsystems. Both of them are cornerstones for generating refutable economic theory.

Another key term is construction. Construction activities are unique in the human economies. A word on this uniqueness is often thought of as a necessary step to help readers follow the analysis of a study of construction behaviour. As early as the 1960s, Turin has embarked on related research. Until the present, except for few studies, our grasp of the economic nature of construction is still limited to orthodox perspectives and cannot keep abreast of the progress of economic theory. As a result, if we desire to pick up this topic as the theme of a study, on the one hand, re-examination of some characterisation of construction is necessary, and on another, tailoring the established theories to fit in these features seems unavoidable.

1.2 Objective of Study

This study identifies itself as an attempt to put organisation theory in construction under review for the future development of a more ambitious and comprehensive enterprise, reorientation of the microanalysis of economic behaviour in construction. Evidently, this broader aim is far beyond the scope that a Ph.D. dissertation can handle. Then, where should we depart? The analysis of the client’s procurement behaviour is a sensible starting point because the construction product specification is customer-made, so the client’s preference on material, design style, and the like, decisively impinges on the direction of development of technology and industry structure. This observation implies that the decisions of the clients will in aggregate shape what the construction industry will be like. As a consequence, we claim that, regarding the study of construction behaviour, the first priority we should place is on the analysis of the client’s procurement behaviour. As will be shown in Chapter 2, the state of affairs on this topic still lack a systematic theory by

---

1 Ivč (1995) should be, as far as I know, the first essay to approach construction problems by the lens of different approach, evolutionary economics.

2 In no doubt, the reality is not so simple and we will go back to this issue later.
which a coherent logic can be found. As a consequence, an attempt is made in this study to modify Williamson's transaction cost theory so as to accommodate some unique features of construction. This study hopes to obtain the goal of providing a novel way of thinking the nature of construction behaviour, ferreting out a suitable methodology, and framing an explanatory model for the analysis of the client's procurement behaviour.

1.3 The problem, its significance and its scope

1.3.1 The problem to be sorted out

The main theme in the study is *why the client chooses the procurement route that he does in practice*. If the problem is addressed this way, what is relevant here is not the process of decision-making, but the outcome of selection. Getting closer to this issue, we can realise that the problem of the client procurement behaviour, in fact, involves more than statically choosing a suitable procurement route. Many dynamic or processual factors shouldn’t be omitted, which will made be clear in Chapter 4 and 6. This study proposes that the original statement of the problem can be reformulated: *how does the client organises a project coalition in an economising way*. Two related issues should be brought to the fore for further clarification: *what are the key determinants that make a procurement route preferable to a client*, and *what is a project coalition*.

First, now that this study intends to make sense of the existing procurement practice, an attempt inevitably should be made to identify the factors with greatest explanatory power. Inspired by Williamson(1987,1988), this study develops a transaction cost economising approach to the analysis of construction process and singles out the most essential variables.

As far as the reformulated problem is concerned, the nature of the procurement decision might be reduced to *the selection of mode of coordination*, though which all participants, such as designer, contractors and so on, can be organised in a least costly way. The nexus of contractual relationships between them is quite similar to a coalition, which indicates formal, long-term alliances between firms that link aspects of their business but fall short of merger [Porter & Fuller, 1986]. However, there are at least three points of
difference to distinguish project coalition from corporate coalition: (1) mode of coordination, (2) horizon of planning and (3) feedback mechanism. The implications resulting from each of them merit attention in the economic analysis.

Roughly speaking, three stereotypical modes of coordination working in the contemporary society include: price, authority and trust [Bradach and Eccles, 1989]. Ordinary corporate coalition is by and large classified under the heading of trust, whereas the classification of the project coalition ranges from pure price mechanism to different degree of mixture of three modes.

As regards the time horizon over which decision is made, it is relatively determinate for the case of project coalition, approximately equivalent to the sum of planning and construction time, whereas the maintenance of corporate coalition is mainly determined by performance, indicating its duration of collaboration being indeterminate within the range from a couple of months to several decades. The most direct effect of this difference is on the calculation of costs and benefits from the coalition.

The last difference, feedback mechanism, a mechanism for transmitting information on historical performance to the next decision, has bearing on the creation, maintenance and breakdown of a coalition. It seems safe to say that the mechanism functions better in a corporate coalition due to its continuity of transmission. In other words, it implies that the past record is more genuinely reflected in the corporate coalition than in project coalition. In fact, the root of the problem identified as “the rigidities of flexibility” in Winch’s term (Winch, 1996) principally stems from the malfunction of this mechanism in construction.

By way of this simple analysis, we can find that these differences render construction behaviour hard to square with received approaches in economics. It seems indispensable to go back one step to contemplate philosophical foundations, such as ontological and epistemological analysis, before methodological construction.
1.3.2 The significance of the problem

1.3.2.1 In theory

While there is no doubt that in the real world, the factors that the client needs to take into account in determining a procurement route are tremendously large and the means that can be employed are diverse [Masterman, 1992], there should exist regularities lying behind the logic of the procurement route selection. This study will make a concerted effort to explore the mechanism leading to these regularities. The value of this undertaking not only lies in helping make crucial practical problems that practitioners "could feel but could not tell" made explicit in a systematic way [Williamson, 1990], but also provides the possibility of developing an integrated theory for the microeconomic analysis of contracting markets in construction.

Fig. 1-1 Coevolutionary Process of the Construction Contracting Markets

To develop a systematic approach for the analysis of construction development on the basis of methodological individualism, we need to forge the link between individual pro-
jects and the whole industry. A tentative conceptual assertion is made here that the long-term equilibrium distribution of organisation forms in construction is determined by a coevolutionary process of (1) selection of procurement routes, and (2) selection of contractor's organisation form, shown as Fig. 1-1. According to the illuminating claim by Groak and Ive(1986) that “economic, technical and organisational factors should be considered in a unified way for the study of building process”, a more complete framework should take the evolution of technology into consideration. However, that would make analysis much more complicated. Here we take an expedient, which means technology is ruled out from the framework but the implications of coevolution for innovation of technology are emphasised.

The logic of this coevolutionary process is this: the clients and contractors are located in a process of mutual adjustment, which moves toward to an equilibrium state. The client wants to take a least costly procurement route to complete the project; the contractor prefers utilising an advantageous form of organisation to cope with unpredictable workload. These two individual evolutions are directed by the performance of the project. For example, a client who originally favours design-and-build on the basis of his own cost-benefit analysis will revert to another route if experiences of implementation are frequently unhappy. If a majority of the clients feels the same way, the fragmented structure, the separation of designers from construction firms, is a sensible response to this tendency. In the evolutionary trail of this sort, resources, in fact, are allocated by way of two interacting processes: selection of procurement route and construction project process. The former one guides production resources, including physical capital and intangible human capital, to flow to producers with viable organisations. This determines what form of organisation will expand and dominate. The latter one is responsible for channelling project resources to divisions of a main contractor or subcontractors, the mechanism singled out in Winch(1987). Obviously, construction process plays a subsidiary role of redistribution of production resources. It is argued that the major allocation mechanism, selection of procurement route, will decisively influence the configuration of the construction industry in terms of resource allocation. What is more important is that the outcome of resource allocation in turn affects the nature of dynamic development and long-
term performance of production. This is the very issue that this study is concerned with.

While the framework set out here is not the core theory of the study, it is hoped that it will effectively assist the reader to assess the value of this study in terms of its role in further theoretical development. We want to enhance the understanding of the dynamic evolution of construction industry and answer Ive’s inspiring question [Ive, 1995]: why do not the ordinary economic forces of self-interest and competition work sufficiently in construction to generate spontaneously a process of competitive innovation and hence continuous improvement in client value-for-money? A positive theory of the client’s procurement behaviour is an indispensable part as far as a complete explanatory framework for construction behaviour is concerned.

1.3.2.2 In practice
Successful strategies turn on the application of consistent principles to constantly changing business conditions [Besanko et. al.(1996)]. Similarly, the proper selection of the procurement route results from the recognition that the performance of a procurement route is simultaneously determined by different types of variables, rather than simply by the client’s preference. The discovery of these principles will greatly enhance the quality of decision. However, in essence, this study is not intended as a managerial study, so the intent is not to generate an apparatus for sorting out some specific management problems, but to reveal implicit mechanisms governing the decisions. The development of transaction cost economics (TCE) is good auxiliary evidence to elucidate the transformation from a positive study to its normative application. The emergence of TCE aims to provide a theory for the analysis of governance. After sufficient empirical data verify the robustness of this theory, original refutable hypotheses become guidelines for managing transaction relations [Rubin, 1990]. Whether this transformation will take place in the case of this study, of course, depends on to what extent the derived hypotheses can be borne out by real data. Undoubtedly, this expectation provides partial impetus to set this study in motion.
1.3.3 The scope of the problem

The variety of the client ranges from public authorities, professional client firms (like property developers), non-profit-seeking organisations, to once-and-for-all clients. There is no denying that the client’s motives for initiating a project are variegated. For a modeller, the decision variables that can be accommodated in the model must be restricted to a tractable level, as must the objective that the client is assumed to pursue. These considerations seem to make it sensible that the focus of this study is only on the client for commercial and industrial buildings in that these two categories of clients reveal strong proclivity to enlarge investment profit as much as possible. That is to say, a standard economic behavioural assumption, optimisation of an objective profit function, is most likely to hold under this circumstance.

Besides, another dimension, planning horizon, is worthy of notice. Kay (1998) rightly points out that it is necessary to examine hybrid governance, such as joint venture, alliances and networks, at different levels. Take McDonald as an example. The merits of its procurement strategy that stresses standardisation and modularisation to reduce planning and construction time cannot be rationalised unless you take the operation of the whole company into account [Cox & Townsend, 1997, 1998a, 1998b]. In view of the dangers that an unclear level of analysis will bring about, this study will concentrate on procurement decisions made on the basis of project by project. That is, the range of factors considered in this study is confined to those directly relating to the individual project. The issues associated with strategic procurement path at the company level or at the programme level (involving several sub-projects) is not the concern of this study.

1.4 Orientation Of This Study

In the face of such an unmanageable reality, in terms of still dominant Popperian perspectives in economic methodologies [Maki, 1993], what is required maybe is to look for the determinants of a phenomenon and take advantage of logical inference to derive refutable hypotheses for testing them with empirical data and verifying the plausibility of the theory. Arriving at the goal of a scientific theory is possible by way of different approaches. Pursuant to Williamson (1993), the progression of a scientific theory of organisation
might need to go through four stages: informal analysis, preformal analysis, semiformal analysis and fully formal analysis. Coase (1937) and Williamson (1975) belong to the first two stages, respectively. Williamson's classic *The Economic Institutions of Capitalism* is classified as a work of the third stage. Finally, Williamson hopes that the formal model of "the theory of incomplete contracting in its entirety" can emerge. This present study, principally, belongs to semiformal analysis. For ease of exposition, mixed use of formal language (mainly, mathematics) and informal language is made, dependent upon which one is suitable. It is worthwhile to note that the foremost task of this study is to devise a refutable theory instead of building up a fully formal model.

1.5 Plan of this study

After justifying the significance of the core issue of this study and delineating its scope, the next task is to show how the ultimate goal can be reached.

To help make my stance explicit, I would like to draw a metaphor. Imagine the generation and application of economic knowledge as some form of specialisation of a manufacturing industry, e.g. chemical industry. Theoretical economists play the roles like those of intermediate goods suppliers, who aspire to promote their product for enhancing academic status and influence on society. Applied economists, who merely are interested in a relatively restricted scope of subject matter with more or less different analytic properties, are like buyers for intermediate goods in the sense that they have got discreetly to find a suited method and assess its appropriateness and usefulness out of a bundle of competing paradigms so as to sort out their own problems. Their output can be consumer goods for government or practitioners or another sort of intermediate goods for other potential followers. To avoid being weeded out, applied economists on the one hand need to justify the tool they employ and on the other, ought to show potential users the validity and fruitfulness of their theories. This is the principle in formulating the research agenda of this study.

As shown in Fig. 1-2, this study constitutes ten chapters. After the introductory remarks of chapter 1, chapter 2 is a review of existing studies associated with construction procure-
ment behaviour, from which we want to point out what different insights this study can generate. Chapter 3 purports to shed light on the nature of transaction cost, its origin, its categories and its implications for economic analysis. In chapter 4, the mainstream theory of transaction cost economics (TCE) upheld by Williamson is systematically introduced for purposes of outlining the logic underlying the following chapters. Though it has been widely accepted in the economic theory of organisation, we should not stretch its analytic power too much without caution about its limitations. In Chapter 5, an effort is made to examine four basic premises hidden behind the TCE’s reasoning and analyse whether they are applicable to construction and, if not, how to advance along the basic reasoning path of TCE by incorporating the characteristics of construction. After finding a possible solution to make TCE framework square with construction, we set out an analytical model in chapter 6 for interpreting construction procurement behaviour. Chapter 7 elaborates upon an empirical model where all relevant variables are identified with reference to the theoretical analysis in chapter 6 and completes the process of deriving the testable hypotheses. In chapter 8, the focus is diverted to the problem of variable measurement. Chapter 9 presents the results of empirical investigation. The implications of the present study and conclusions of our findings are in chapter 10.
Chapter 2
Review the existing studies on construction procurement behaviour

Chapter 3
An illumination of the concept of transaction costs

Chapter 4
An analytical framework for transaction cost reasoning

Chapter 5
Examine the basic premises for applying transaction cost theory to construction

Chapter 6
Building up a new paradigm for the study of procurement behaviour

Chapter 7
An empirical model for exploring construction procurement behaviour

Chapter 8
Development of questionnaire and survey design

Chapter 9
The analysis of empirical results

Chapter 10
Conclusions and Suggestions
The value and necessity of a Ph.D. research, in large part, depends on its novelty in providing new insights into reality. Before jumping to judgement, we have to realise what kind of knowledge has been generated, in what sense it has value and to what extent it is still insufficient to our satisfaction.

As far as the number of research papers is concerned, the study of procurement routes is an active academic research topic, evidenced by the fact that from 1970 onwards, there are nearly 900 pieces of work aiming their focus on this issue [Franks, 1998]. Notwithstanding that their purposes are somewhat diverse, the theme relating to the exploration of best procurement practice seems to have attracted particular attention. Against this backdrop, it is quite clear that the significance of procurement route selection in affecting the outturn of the project implementation is widely recognised. However, the existing literature doesn’t provide a satisfactory theory to solve the core research problem of this study - “the rationalisation of construction client’s procurement route selection behaviour”. Before canvassing the reason, it is essential for us to get a whole picture of how this problem is currently handled. Four lines of inquiry will be critically assessed in turn in this chapter. The first approach uses the most preliminary induction method to discover the conditions suitable for each procurement route through the observations of project implementation. These conditions can be used as guidelines for assistance in procurement route selection, so this method is labeled the guideline approach. The second one, the statistical approach, is an attempt to utilize statistical methods to improve the crude inference quality of the guideline approach. The third one pursues the aim of providing normative advice for the client’s reference by applying the multi-attribute utility approach to develop a decision method, so it is named after its methodology, the multi-attribute utility approach (MAUA). The last approach doesn’t exhibit a coherent line of exploration, but instead tries to expose the problem through the lens of different perspectives. Thus they are categorized into the group of the multidisciplinary approach. We will take a look at these approaches in turn.
2.1 The guideline approach

2.1.1 Introduction

The increasing number of construction procurement guides being published in recent decades signifies a prevailing awareness of the importance in finding a suitable procurement route in practice [e.g. Franks, 1993; Naoum, 1991; RICS, 1996; Knochar and Sanvido, 1998]. It implies that a well-founded theory of procurement route selection will be beneficial to the improvement of the client's decision quality. However, what choice can be said to be right? A direct idea springing up to mind may be to consult with a panel of experts who have experience on the issue in question, and pool their opinions together to form a reference base. Relying upon the mechanism of learning by doing is one of the most primitive means of knowledge accumulation. The repeated occurrence of effect and end of a phenomenon can give observers a certain level of confidence to infer regularities that would have predictive power and that the robustness of this inference will be reinforced as the number of samples under review are expanding. Through past experience in adopting a specific route, some conditions suitable for the use of a route can be obtained. For example, Masterman (1992) offers guidelines for selecting procurement systems, summarized in Table 2-1. These rule-of-thumb criteria can be a good starting point for the exploration of the fundamental mechanism behind the client's decision. Let's condense Table 2-1 into a number of principles:

1. Traditional procurement system

(i) Due to the availability of complete drawings and specifications, potential contractors are allowed to compete on bidding prices, which is likely to drive the winning price down. Moreover, quality is easier to be ensured on the basis of the same reason.

(ii) Separation of design and construction makes fast tracking infeasible; that is, the client has to bear the opportunity costs of longer total time of design and construction.
<table>
<thead>
<tr>
<th><strong>Traditional System</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) projects where competitive bids are required in order to ensure financial accountability and minimise tender price; time is not of the essence;</td>
</tr>
<tr>
<td>(b) where there is a need to ensure that high standards of quality and functionality can be achieved;</td>
</tr>
<tr>
<td>(c) complex or and prestigious projects where the design needs to be under the direct control of the client;</td>
</tr>
<tr>
<td>(d) design changes may need to be made during the currency of the works and the client is prepared to accept the time and cost implications of such changes.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Design-Build System</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) projects where the client requires the overall project period to be as short as possible, but needs to be aware of his financial commitment at tender stage;</td>
</tr>
<tr>
<td>(b) where the client requires one organisation accept responsibility for both design and construction</td>
</tr>
<tr>
<td>(c) where the client is able accurately to define his requirements at tender stage and is unlikely to wish to amend them during the currency of the project;</td>
</tr>
<tr>
<td>(d) when the total project costs needs to be minimised.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Management System</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) when the commencement of the project needs to be accelerated in order to evoke speedy completion;</td>
</tr>
<tr>
<td>(b) where early advice is needed from a management contractor;</td>
</tr>
<tr>
<td>(c) where the client may wish to have flexibility available to incorporate design amendments into the project.</td>
</tr>
<tr>
<td>(d) where certainty of achieving completion dates is required;</td>
</tr>
<tr>
<td>(e) when the project is of high value and complexity and thus requires the use of sophisticated construction and management techniques;</td>
</tr>
<tr>
<td>(f) when the client wishes to be objectively involved in the overall management;</td>
</tr>
<tr>
<td>(g) when the client, in the case of design and manage, wishes to appoint a single organisation to be responsible for the design and management of the works whilst, at the same time, acting in a consultant role.</td>
</tr>
</tbody>
</table>
(iii) Because of uncertainty, ex ante design may need to be modified, which con­sequently engenders the possibility of cost variations. The client needs to be pre­pared to bear the costs of any changes initiated by themselves or by their designer if they adopt this route which does however both allow for such changes and in­corporates procedures (unit prices in bills of quantities) for mitigating the costs of such changes.

2. Design-build system
(i) Partial overlapping of design and construction can lead to shorter project duration with less opportunity cost of time incurred.
(ii) Since design and construction team belong to the same organisation, the responsibility can be easily delineated and communication problem lessened.
(iii) Appropriateness of these systems depends on whether the client is able to clearly specify his requirement ex ante.
(iv) Providing no design changes occurring ex post, applying this system can make the ex ante cost controllable.
(v) This system heavily penalises clients who initiate post-contract changes, and should only be used where the client assesses the likelihood of such changes to be low.

3. Management system
(i) Need for extensive management advice is a reason to adopt this system.
(ii) It is desirable to use this method in carrying out a project with high value and complexity because of sophisticated management techniques involved in the pro­ject.
(iii) The use of this system can shorten the delivery of the project. Maximum over­lapping of design and construction can lead to shortened project duration and thus re­duce opportunity costs of time.

2.1.2 A critical assessment
Most of the existing procurement guidebooks set the scene by comparing the pros and cons of each procurement route and attempt to point out the conditions to which each option may be suitable. These findings are principally based on the perceptions of procurement route implementation performance in the past. Perhaps, the rules of thumb derived by this way sometimes can work, whereas, just as diagnosing diseases wholly on the basis of symptoms, we must be cautious about the danger of taking
stock of the conditions fitting each procurement route purely on the basis of superficial phenomena. This is because the same symptoms are likely to result from different reasons and need different medicine. Treating them indifferently could worsen the situation. And this is why it is important to develop physiology for understanding the fundamental mechanism of diseases. This doesn’t mean to deny the value of reference that the guideline approach can provide, but to stress that, without thorough investigation of causal relations between the observed choice of procurement routes and the underlying determinants behind the observations, the guidelines crystallised out of pure empiricism will be too crude to be scientific. Maybe for this reason, the statistical approach and multi-attribute utility approach emerge to provide a more quantitative tack to improve the client’s decision on the selection of procurement routes.

2.2 The statistical approach

2.2.1 Introduction

While the guideline approach can help transform dispersed knowledge on the implementation of procurement methods in practice into some useful principles, the trade-off relations between different variables stay untouched. To go beyond the crude classification of the circumstances where each procurement route shows advantages, the quality of inference can be improved by employing statistical methods to link up the correlation between project outturn and project attribute variables. For example, Molenaar and Songer (1998) use the regression method to explore the factors affecting the performance of design-build. The basic model is

\[ S = V\beta + \varepsilon \]  

Eq. 2-1

where \( S \) is a 5×1 column matrix, standing for five performance criteria; \( V \) is a 1×n row matrix, denoting n explanatory variables; \( \beta \) is a n×5 coefficient matrix; \( \varepsilon \) is a 5×1 column matrix for random errors.

They propose five criteria for measuring project performance: budget variance, schedule variance, conformance to expectations, administrative burden and overall satisfaction. Then, they set out a hierarchical structure for breaking down a large number of explanatory variables into four categories: project, owner, market, and relationship. By way of the regression analysis, they find that the following variables are
statistically significant for explaining at least one performance variable of Design-Build projects:

(1) Project: scope definition, schedule definition, budget definition, project complexity.

(2) Owner: owner/agency experience, owner/agency staffing, design-builder experience.

(3) Market: design-builder experience

(4) Relationship: design-builder selection, design-builder prequalification.

These variables mainly are measured by semantic differential scale, as listed in Table 2-2.

Table 2-2 The measure of variables in Molenaar and Songer (1998)

<table>
<thead>
<tr>
<th>Type I: Project</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Scope definition:</strong></td>
</tr>
<tr>
<td>(a) What was the level of design (percent complete) at the time of bid or proposal?</td>
</tr>
<tr>
<td>(b) Was the project complete at the end of the design-builder's contract or were additional contract required?</td>
</tr>
<tr>
<td>(c) Were the specifications prescriptive based, performance or a combination thereof?</td>
</tr>
<tr>
<td>(d) Did the solicitation documents (request for proposal or invitation to bid) prevent or encourage additions to scope?</td>
</tr>
<tr>
<td>(e) Did the project have a flexible scope when the design-builder was hired (i.e. a need for the changes during construction was foreseen)?</td>
</tr>
<tr>
<td>(f) What was the level of project scope definition when the design-builder was hired?</td>
</tr>
<tr>
<td><strong>2. Schedule definition:</strong></td>
</tr>
<tr>
<td>(a) At what level of design (percent complete) was the project's construction completion date fixed?</td>
</tr>
<tr>
<td>(b) How crucial were the market factors or outside demands for the project to be delivered?</td>
</tr>
<tr>
<td>(c) To what extent did the agency funding cycle drive the project schedule?</td>
</tr>
<tr>
<td>(d) Did the design-builders have sufficient time to prepare their proposal/bid?</td>
</tr>
<tr>
<td>(e) Did the owner have sufficient time to review the proposals/bids and select the design-builder?</td>
</tr>
<tr>
<td>(f) To what extent was the construction completion date allowed to vary during the selection of the design-build team?</td>
</tr>
<tr>
<td>(g) How critical was staying on schedule to the project's success?</td>
</tr>
</tbody>
</table>

| **3. Budget definition:** |

17
(a) At what level of design (percent complete) was the project budget fixed?
(b) Was the budget conveyed to the design-builder as (i) constraint on scope (ii) maximum allowable (iii) not defined.
(c) What was the owner’s project contingency when the design-builder was hired?
(d) Was the project built with appropriated or non-appropriated funds?
(e) How critical was staying on budget to the project’s success?

### 4. Project complexity:

(a) What was the level of design complexity?
(b) What was the level of construction complexity?
Was the project similar to others which have been built in your locale?

#### Type II: Owner

**1. Owner/Agency experience:**

(a) What was the agency representatives' level of construction sophistication?
(b) What was the owner representatives' level of construction sophistication?
(c) How much experience did the agency representatives have with similar construction projects?
(d) How much experience did the owner representatives have with similar construction projects?

**2. Owner/Agency staffing:**

(a) After the design-builder was hired, was the agency staff adequately sized to answer design-related questions?
(b) After the design-builder was hired, was the owner staff adequately sized to answer design-related questions?
(c) After design-builder was hired, was the agency staff adequately sized to answer construction-related questions?
(d) After design-builder was hired, was the owner staff adequately sized to answer construction-related questions?

**3. Design-Builder experience**

(a) On average, how many prior design-build projects had the agency representative been involved with?
(b) On average, how many prior design-build projects had the owner representatives been involved with?

#### Type III: Market

**1. Design-builder experience**

(a) What was the level of design-builder experience with similar types of projects?

#### Type IV: Market

**1. Design-builder selection**

(a) What type of contract was used with the design-builder?
(b) Was the contract competitively bid or negotiated?
Design-builder prequalification

(a) Was the design-builder selected solely on qualifications, solely on low price or some combination?
(b) Was there any prequalification of proposers/bidders?

On the basis of this research, they develop a computer graphical interface Design-Build Selector (DBS) that consists of two elements:

1. Input the value of explanatory variables of a new project in question into the regression model obtained in their research.
2. Convert five performance indicators into an overall score and benchmark the score of the new project against an average score of the projects surveyed in their study. This comparison will show whether this new project if carried out by design and build will be more likely to outperform the average level of those projects surveyed.

Relative to the guideline approach, this approach has made a stride forward, while there is still much room to improve. This is the topic that we will turn to in the next subsection.

2.2.2 A critical review

We can discuss the weaknesses of the methodology used in Molenaar and Sengers(1998) from three respects. First, the nature of this approach is pure empiricism. It aims to filter the relevant variable from a bunch of possible variables by the aid of regression analysis. The discovery of the explanatory variables wholly depends on whether these variables are significant in the statistical test, for example t-test. But due to lack of a theoretical underpinning, we don’t have a systematic model for thinking about why the results are as they are and why some variables are negligible.

Second, setting the goal of discovering the statistical relations as the only concern in mind is liable to lead to “statistical tyranny”, implying that the alleged high-level-of-confidence variables may be just the consequences of ignored but more fundamental factors. Leaving these variables out of consideration without notice may weaken the soundness of the theoretical foundation of this method.
Last, in application, this approach takes the average performance in the population as benchmark, which will arouse a critical methodology problem in that this point is floating along with the addition and deletion of samples. It doesn’t mean that this benchmark has to be a constant, but it must display stability so that practical users can get a consistent guidance.

2.3 Multi-attribute utility approach

Among voluminous stock of studies related with construction procurement systems since 1970, multi-attribute approach is claimed to be the “foremost technique appropriate for examining the criteria of clients and the preferences of experts’ weights for each method in the most objective way” [Love et al., 1998] and attracts considerable academic attention (e.g., Skitmore and Marsden, 1988; Chan, 1995; Love et. al., 1998; Ambrose and Tucker, 2000), so it is necessary for us to give a more thorough examination on the rationale of this approach.

2.3.1 Fundamentals of multi-attribute utility approach

Expected utility theory (EUT) is a mainstream approach in interpreting decision behaviour in the social sciences. For a decision maker facing a multitude of alternative choices, each choice will lead to a different set of possible consequences with different probability distribution, therefore achieving different degree of satisfaction. The betting game is a good case. EUT dictates that choice $x$ is better than $y$ if and only if the expected monetary utility coming from $x$ is larger than that from $y$. Put in a formal way,

$$x \succ y \text{ if and only if } U(x) > U(y)$$

Eq. 2-2

$$U(x) = \sum_{a \in A} P_x(a)R_x(a)$$

$$U(y) = \sum_{b \in B} P_y(b)R_y(b)$$

where $A$ and $B$ are the sets of consequences brought about by strategies $x$ and $y$; lower case $a$ and $b$ indicate the elements of each set; $P_x(a)$ and $P_y(b)$ are probabilities of each consequence occurring; $R_x(a)$ and $R_y(b)$ are payoffs of each consequence.
However, decision problems in daily life are much more complicated than gambling in the sense that probability and payoffs of consequences are hard to estimate. In dealing with this case, objective measure of probability can be replaced with subjective judgment [Savage, 1954] so that alternative options can be compared on an equal footing.

Sometimes, rather than a single-attribute payoff decision tree (e.g. in the case of gambling, the payoffs are measured by the monetary unit), the problem facing decision makers is the multiple-attribute payoff decision tree (e.g. the selection of nuclear power plant site). To deal with this case, the multi-attribute utility approach is developed, where the expected utility of choice \( j \) is determined by

\[
U_j = \sum_{i=1}^{n} w_i x_i
\]

Eq. 2-3

where \( x_i \) is the value given to the attribute \( i \) of a utility function, decided by the decision maker's subjective evaluation; \( w_i \) are the utility coefficients relating attributes to options. It is to be noted that the consequences of each choice are no longer evaluated by individual events, but against a set of factors \( \{ x_i \} \) that reflects the preference structure of decision makers. Take an example of car purchase. Apart from price, other factors, such as comfort, reliability, efficiency of engine, noise, maximum speed all possibly impinge upon consumers’ preference over cars. Therefore, consumers’ satisfaction from each model of car can be described by a set of the values of these attributes. As applied to policy issues, the attributes affecting public opinion can be used to assess the desirability of each option. An example is the selection of nuclear power plant site [French, 1989]. The construction of this plant is likely to pose direct impacts in different aspects, including financial issues, safety issues, and environmental issues. These are different dimensions for the appraisal of the consequences attendant on the construction of the nuclear power plant. In fact, no matter what issue is under review, the appropriateness of employing multi-attribute approach lies in two links (see Fig. 2-1):

1. Link 1: the attributes of the outcome should be able to fully reflect the decision-maker’s criteria for assessing consequences of each option;
2. Link 2: the effect of each option on these attributes of the outcome should be clearly identified.
2.3.2 Multiple-attribute utility approach for procurement route selection

As applied to construction procurement systems, the multi-attribute approach involves four steps:

(i) Identify the priority variables in the client’s utility function
(ii) Fix the weighting factors relating achievement of priority variables as outcomes to procurement routes.
(iii) Determine the value of priority variables according to the client’s preference
(iv) Sum up the weighted priority variables of each procurement route and choose the one with highest score.

For ease of exposition, we reproduce a simplified version of a procurement route evaluation table used in Love et al. (1998). These four steps will be examined in turn by referring to this table.

2.3.2.1 Identification of priority variables

What are the relevant factors in affecting the client’s choice of procurement systems? In the literature, the factors regarded as crucial are to some degree subject to researchers’ *a priori* judgment. Thus, the list and definition of priority variables in the litera-
ture differs from study to study, as shown in Table 2-3. There seems to have emerged a consensus on the importance of eight factors: (1) speed, (2) price certainty, (3) flexibility, (4) quality standard, (5) complexity, (6) risk allocation, (7) price competition and (8) responsibility. That is, the client is assumed to select procurement systems on the basis of these eight factors. Those factors considered in Love et al. (1998) are shown in the first column (Col.1) of Table 2-4. For simplicity of exposition, the table show all priority variables to have been assigned an equal weight by the client.

### Table 2-3 Documented priority variables affecting the client's decision on procurement systems

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Speed</strong></td>
<td><strong>1. Time</strong></td>
<td><strong>1. Timing</strong></td>
<td><strong>1. Speed</strong></td>
</tr>
<tr>
<td>How important is early completion to the success of your project?</td>
<td>Is early completion required?</td>
<td>How important is early completion to the success of your project?</td>
<td>How important is early completion to the success of your project?</td>
</tr>
<tr>
<td><strong>2. Certainty</strong></td>
<td><strong>2. Cost</strong></td>
<td><strong>2. Price certainty</strong></td>
<td><strong>2. Certainty</strong></td>
</tr>
<tr>
<td>Do you require a firm price and/or a strict completion date for the project before you can commit yourself to proceed with construction?</td>
<td>Is a firm price needed before any commitment to construction is formed?</td>
<td>Do you need to have a firm price for the project construction before you can commit it to proceed?</td>
<td>Does your organisation require a firm price or strict completion time for the project before your organisation can commit to a building project?</td>
</tr>
<tr>
<td><strong>3. Flexibility</strong></td>
<td><strong>3. Controllable variation</strong></td>
<td><strong>3. Flexibility</strong></td>
<td><strong>3. Flexibility</strong></td>
</tr>
<tr>
<td>To what degree do you foresee the need to alter the project in any way once it has begun on site?</td>
<td>Are variations necessary after work has begun on site?</td>
<td>Do you foresee the need to alter the project in any way, for example to update machinery layouts?</td>
<td>During the course of a building project, to what extent does your organisation feel it necessary to alter the project in any way once it has begun on site?</td>
</tr>
<tr>
<td>What level of quality, aesthetic appearance do you require in the design and workmanship?</td>
<td>Is high quality important?</td>
<td>What level of quality do you seek in the design and workmanship?</td>
<td>What level of quality, aesthetic appearance do you require in the design and workmanship?</td>
</tr>
<tr>
<td>Does your building need to be highly specialised, technologically advanced or highly serviced?</td>
<td>Is the building highly specialised, technologically advanced or highly serviced?</td>
<td>Does your building (as distinct from what goes in it) need to be technically advanced or highly serviced?</td>
<td>Does your organisation require a technologically advanced or highly specialized building?</td>
</tr>
<tr>
<td>Is it important for you to choose your construction team by price competition, so increasing the likelihood of a low price?</td>
<td>Is completion on time important? Is completion within budget important?</td>
<td>Do you need to choose your construction team by price competition?</td>
<td>Is it important to select the construction team by competition?</td>
</tr>
<tr>
<td>To what extent do you wish one single organisation to be responsible for the project; or to transfer the risks of cost and time slippage?</td>
<td>Is transfer of responsibility for the consequence of slippages important?</td>
<td>Do you want to pay someone to take the risk of cost and time slippage from you?</td>
<td>Does your organisation want to limit the amount of speculative cost and design liability?</td>
</tr>
</tbody>
</table>
8. Division of responsibility
Is single-point responsibility wanted?
Is direct professional responsibility wanted?

8. Management
Can you manage separate consultancies and contractor, or do you want just one firm to be responsible after the briefing stage?

8 Responsibility
To what extent do you wish one single organisation to be responsible for the project; or to transfer the risks of cost and time slippage?

9. Accountability
Do you want professional accountability to you from the designers and cost consultants?

9. Arbitration and Disputes
To what extent does your organisation wish to avoid disputes and arbitration?

Table 2-4 A simplified version of procurement route evaluation table in Love et al.(1998)

<table>
<thead>
<tr>
<th>Client's priority Variables</th>
<th>Weight assigned by Client (0-1)</th>
<th>Procurement Options</th>
<th>Procurement Options</th>
<th>Procurement Options</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Col.2)</td>
<td>Traditional method</td>
<td>Design and Build</td>
<td>Construction Management</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Utility coefficient (Col.3)</td>
<td>Result (Col.4)</td>
<td>Utility Coefficient (Col.5)</td>
</tr>
<tr>
<td>Speed</td>
<td>0.11</td>
<td>57.3</td>
<td>5.78</td>
<td>76.2</td>
</tr>
<tr>
<td>Certainty</td>
<td>0.11</td>
<td>88.5</td>
<td>9.74</td>
<td>90.3</td>
</tr>
<tr>
<td>Flexibility</td>
<td>0.11</td>
<td>75.6</td>
<td>8.32</td>
<td>59.6</td>
</tr>
<tr>
<td>Quality</td>
<td>0.11</td>
<td>100.0</td>
<td>11.00</td>
<td>60.8</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.11</td>
<td>81.6</td>
<td>8.87</td>
<td>75.6</td>
</tr>
<tr>
<td>Risk Allocation</td>
<td>0.11</td>
<td>80.0</td>
<td>8.80</td>
<td>96.8</td>
</tr>
<tr>
<td>Responsibility</td>
<td>0.11</td>
<td>85.6</td>
<td>9.75</td>
<td>92.5</td>
</tr>
<tr>
<td>Arbitration and disputes</td>
<td>0.11</td>
<td>75.3</td>
<td>8.28</td>
<td>70.8</td>
</tr>
<tr>
<td>Price competition</td>
<td>0.11</td>
<td>94.5</td>
<td>10.40</td>
<td>42</td>
</tr>
<tr>
<td>Total</td>
<td>≈1</td>
<td>80.94</td>
<td>70.09</td>
<td>71.47</td>
</tr>
</tbody>
</table>

2.3.2.2 Correlate priority variables with procurement methods

After priority factors are identified, the next step is to assign a weighting index (called utility coefficient in Table 2-3) (in Love et al. from 10 to 110) to correlate priority factors with procurement methods. As proponents of this approach assert, "[f]or this procedure to be of practical use it is necessary first to fix utility coefficients that relate these attributes to individual procurement methods independent of individual projects" [Love, et al., 1998]. The three shadowed columns (Col.3,5,7) in Table 2-4 display the predetermined utility coefficients linking each priority variable to the expected performance of each route in terms of these variables. According to Love et
al. (1998), one possible way of obtaining the estimate of these utility coefficients is to elicit experts' consensus opinions. Thus, it is not surprising to find that the coefficients provided by different authors vary as the sources of expert opinions.

### 2.3.2.3 Determine the weightings to priority variables

The client has to determine the relative importance of priority variables to the success of his/her procurement. The weighting assigned to each factor signifies the client’s priority to achieve these factors. For example, if the client thinks early delivery is most important goal of the project, other things being equal, the route that can best facilitate the achievement of this goal will stand out, i.e., construction management. A special set of weightings with equal low figures (0.11) to each variable is shown in column 2 (Col.2) in Table 2-4.

### 2.3.2.4 Computation of utility index of procurement methods

Using Eq.2-3, we can compute the utility index that each procurement method would deliver in dealing with the project in question by multiplying client’s weighting to each factor (Col.2) by the utility coefficient of each procurement route with respect to each factor (Col.3,5,7) and totting them up. As shown in the bottom row in Table 2-4, traditional method is the most suitable route in this special case of equal low weightings. Ideally, this index should fully reflect the extent that the client’s requirements can be fulfilled by each procurement option; otherwise, it will mislead the client’s choice. In the following section, we will investigate the weaknesses of this method.

### 2.3.3 Where's the problem?

#### 2.3.3.1 Ignorance of the basic premise of MAUA

The basic idea underlying MAUA is that the client is suggested to select an appropriate procurement route on the basis of priority factors, while it is quite odd to find that to our knowledge, the issue of why these factors are relevant hasn’t been addressed in the literature. The reason may be due to the consensus that these factors actually are the client’s considerations in practice. However, on theoretical grounds, we claim that the set of factors requires careful scrutiny. To understand why the re-examination of

---

1 This arrangement of weights is just for the purpose of illustration in the following section.
these factors is essential, the methodology of MAUA should be reconsidered in the context of construction procurement route selection, as shown in Fig. 2-2.

The soundness of decision that Fig. 2-2 can help make depends on the robustness of Link 1 and Link 2. Link 1 requires the priority variables be (different dimensions of) "ends", not "means", and Link 2 demands an affirmative relation showing the extent that the client's needs can be satisfied by each procurement route in terms of priority factors (procurement routes as means to multiple ends). With a robust relation between procurement systems and their expected performance evaluated in terms of priority variables, the client just needs to give a weight to each criterion, then the option with the highest utility index will be the most desirable one serving the client's preference. However, if either of two links were broken, decision principles formulated on the basis of assigning the weight, \( x_i \), to each factor will not truly reflect the relative desirability of procurement routes.

**Fig. 2-2 The reasoning structure of multi-attribute approach for procurement route selection**

First of all, it is necessary to be cautious about the condition under which the methodology of MAUA can be applied. We have stressed that only when priority factors represent different aspects of the *outcome* of the project implementation can the client's assignment of \( x_i \) be meaningful. However, as the next subsection shows, this point seems not respected in the literature. Second, every cell in columns 3, 5, 7 of Table 2-4 in principle should be able to reveal the desirability of procurement routes with respect to each evaluation criterion (i.e., priority variables). As will be discussed later, their reliability heavily relies upon an unreal assumption of *complete contracting*, implying that *ex ante* promises will be fully and effectively enforced by contract.
On the basis of these two considerations, we will argue in the following subsections on two fronts: (1) the first requirement is to distinguish what types of variables should be used as priority variables; and (2) that even some priority variables of the right type (as will be discussed later, this refers to consequence variables) used in the literature should be reconsidered.

### 2.3.3.2 What should be included as priority variables?

In the previous section, it has been stressed that, in applying MAUA, decision variables should be attributes of consequence of an action (in this case the action in question being the selection of a procurement route). Unfortunately, several variables used in MAUA don’t meet this requirement. In fact, the eight variables\(^2\) with high degree of consensus in the literature belong to three categories of different nature:

1. **Consequence variables (CVs)**
   Consequence variables are criteria against which the achievement of the client’s goal (maximizing expected utility from the outcome) is assessed. According to the doctrine of MAUA, they are the right desiderata in the utility function. The CVs used in Table 2-4 include: speed\(^3\); outturn price, decomposed into two stages, outturn price certainty (that final payment will not exceed bid price or estimate), and price competition (to drive down lowest bidding price); achieving the level of quality specified; and risk avoidance (by transferring risk to the contractor).

2. **Attributes of procurement routes:**
   This category of variable are attributes of procurement routes that may assist clients in achieving their desired CVs for projects of given attributes. Single point of responsibility is a case in point.

3. **Attributes of the project:**
   This category of variables is the factors causing and requiring differentiation from project to project. Flexibility (likelihood of changing requirements) and complexity (of the project) belong to this group.

We are not arguing against the importance of these variables in affecting the client’s procurement route selection, but questioning the appropriateness of the way they are used.

\(^2\) These terms follow the definitions used by the MAUA procurement literature authors (see Table 2-3)

\(^3\) In fact, speed can be classified either as a CV or an attribute of procurement routes.
analysed in the literature. Theoretically, as the equation of MAUA (Eq.2-3) is applied to the case of procurement route selection, all the three categories of variables should be involved. First of all, the CVs should be selected on the basis of the client’s criteria used to evaluate the outturn of the project. Second, utility coefficients ($w_j$) should be jointly determined by the attributes of procurement routes and those of the project. The determination of project attributes itself is a reflection of the ultimate purpose for which the project is built, e.g., housing, office or factory. This context is also a key determinant of the priority that the client will give to each CV. Thus, the framework in Fig. 2-3 describes a correct way to accommodate these three categories of variables into the framework of MAUA. The satisfaction of these three conditions can ensure the robustness of Link 1 and Link 2 and thus the reliability of the suggestions derived from it.

An intuitive understanding of these assertions can be illustrated by the example of car purchase again. As said above, the reasonable CVs in this case include price, comfort, reliability, efficiency of engine, noise and maximum speed. The utility coefficients must be determined by the characteristics of cars (model attributes analogous to procurement route attributes). For example, racing cars should get high score in speed and Cadillacs should be rated higher in comfort. Last, what weighting will the buyer assign to each priority variable? The answer depends in part on what’s the main purpose of buying this car. If the car is used as a commuting vehicle, reliability and comfort will be given higher weighting, while if used for a symbol of status, price may top
the priority list. By analogy, the purpose of car purchase is analogous to the purpose of buying a building. The weightings assigned to each \( x \) should be based on the characteristics of the construction transaction. For example, the first priority of building a hi-tech factory in the boom time is speed, while for a monumental government building, quality may become most important.

We argue that, according to the doctrine of MAUA, the effect of the subjective preference on single point of responsibility, flexibility or complexity on the selection of procurement routes should be ruled out of the utility function. Instead, such project attributes should be taken explicitly into account in arriving at appropriate project specific \( w_j \) (utility coefficients) for each true CV.

For example, Skitmore and Marsden (1988) and Love et al. (1988) conflate an outcome or CV (transfer of risk of cost and time slippage) with a means to an end (single point of responsibility). Inclusion of the latter in the utility function only confuses matters, undermining the usefulness of MAUA.

### 2.3.3.3 What are CVs: the influence of contract incompleteness

Up to this point, it should be clear that only CVs are qualified to be elements in the utility function\(^4\). We also show that the CVs actually considered in the literature are mixed with attributes of procurement routes and of the project. Then, one might ask whether the current application of MAUA can be saved by just restricting the list of priority variables to CVs. This subsection attempts to argue that the answer is no, if the proposition of incomplete contracting is adopted.

Generally speaking, short delivery time, low cost and high quality are upheld as the three principal goals of construction project management. The client’s utility function for procurement route selection should consist of the variables able to reflect the realised outcome of the project implementation with respect to time, cost and quality. The connection between CVs used in the MAUA literature and the three final goals is shown in Fig. 2-4. First, it is necessary to point out an ambiguity in the definitions of

---

\(^4\) In our interpretation we make the simplifying assumption that the client gains no utility from their involvement in the construction process, but only from construction outcomes.
quality used in the MAUA procurement literature. As will be explained later, quality of specification is quite distinct from quality of conformance (Winch et al., 1998). Second, with respect to time and cost, these appear to be linked in the MAUA literature with four documented CVs: speed, certainty (of price and time); price competition; and risk allocation (cost and time slippage risk).

![Fig. 2-4 The relation between consequence variables in MAUA literature and the goals of construction management](image)

However, in a world of contract incompleteness, promise does not assure performance, and we must sharply distinguish between what is contracted and what is performed (actual outcome). The variables used in MAUA mostly relate to the former rather than the latter.

In terms of contract, there are two conditions necessary to the achievement of the client’s goals: (1) agents make an ex ante commitment on terms aligned to the client’s interests; and (2) the commitment can be effectively enforced. Following this logic, the role of including price competition, certainty and risk allocation can be clarified. As shown in Fig. 2-5, at the pre-contract stage, by way of open tendering under traditional procurement, the winner’s tender price can be lowered down to some extent and
in conjunction with the selection of contract form, the client has the option to obtain an *ex ante* commitment on low price. By way of appropriate choice of procurement systems, e.g. design and build, it is likewise possible for the client to get an *ex ante* commitment on the delivery time and upper ceiling of budget. Commitments alone however are not enough. The client also needs to be able to transfer risks likely to occur in the course of transaction to ensure the ex ante commitments can be realised as promised.

**Fig. 2-5** The roles of three consequence variables

<table>
<thead>
<tr>
<th>Pre-contract stage</th>
<th>Post-contract stage</th>
<th>Realisation Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commitment on tender price</td>
<td>Risk transfer</td>
<td>Time</td>
</tr>
<tr>
<td>Commitment on completion time</td>
<td></td>
<td>Cost</td>
</tr>
</tbody>
</table>

The principle to be used for testing the qualification of a CV is whether the efficacy of procurement systems with respect to that CV is *intrinsically* different. In essence, speed and quality are separable and will be discussed as such. The remaining three variables are intrinsically independent. The extent that certainty and risk allocation can be fulfilled are dependent on the choice of contractual arrangements (ie. procurement systems), so we will discuss them together by focusing on the role of contract. However, the main purpose of price competition is to get low tender price. Whether the client can ensure the realization of that price depends on contract, so we have to discuss what the price signal stands for and its relationship with contract completeness.

### 2.3.3.4 Speed of project delivery

Speed is a sensible CV since the performance of procurement systems with respect to this dimension is intrinsically different due to different extent to which it is possible for design and construction to be overlapped. This assertion also gains empirical support. The results of BEDC (1988) in Table 2-5 shows that there is a categorical difference in the performance of procurement systems in this respect.
Table 2-5 Comparison of delivery time through three procurement routes (BEDC, 1988)

<table>
<thead>
<tr>
<th></th>
<th>Percentage projects % (Site time)</th>
<th>Percentage projects % (Total time)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Faster</td>
<td>Average</td>
</tr>
<tr>
<td>Traditional</td>
<td>17</td>
<td>37</td>
</tr>
<tr>
<td>Design and Build</td>
<td>47</td>
<td>11</td>
</tr>
<tr>
<td>Management contracting</td>
<td>50</td>
<td>38</td>
</tr>
</tbody>
</table>

Note: 'Fast' projects are those completed at least 10% faster than the norm. 'Average', completed within 10% of the norm. 'Slow', taking at least 10% longer than the norm.

2.3.3.5 Quality

Quality will be discussed in Chapter 6, and here we just briefly outline the key points. The concept of construction quality contains at least two levels: design quality and conformance quality. The design expressed in the form of drawings and specifications plays the role of yardstick for assessing conformance. Using the integrated route, like design and build, in one respect it will be easier for the client to implement quality enforcement (conformance) since the liability for any defect is clear (single point of responsibility). But on the other hand, in design and build the client lacks a technical specification against which to measure out-turn quality, and must instead rely on a performance specification. This may make it more difficult to detect quality non-performance, since it is intrinsically difficult to write complete and enforceable performance specifications. Relative to the traditional method and management system, it may be harder to maintain design quality by way of design and build. In traditional method, 'separation' effects may make it easier to control design quality, but it may become more difficult to enforce conformance because of divided responsibility. We believe this is a tradeoff, implying no procurement system enjoys absolute advantage in quality delivery.

2.3.3.6 The importance of contract incompleteness

If examined closely, the remaining three CVs, including price competition, certainty and risk allocation, in fact are just proxy variables for the client’s final goals, as shown in Table 2-6. What the client actually wants to achieve is certainty on final
payment and date of completion, while in Table 2-3 certainty in fact just indicate *ex ante* commitment made by the contractor. Similarly, the desire for lower final payment is proxied by price competition, and that for risk avoidance by contractual risk allocation. The legitimacy of taking these contract variables as CVs lies in the hidden assumption of complete contracting. In a world of complete contracting where all the possible contingencies occurring in the course of transaction can be efficiently governed, you can always get what you contract for, ensuring contract variables are equivalent to outturn variables. This reasoning is manifested by what we argue are the “spurious” relations between priority variables and procurement routes (ie. Link 2) in Table 2-4 of Love et al.(1998).

<table>
<thead>
<tr>
<th>Proxy variables</th>
<th>Desired consequences for the client</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ex ante</em> commitment on price and date of completion</td>
<td>Certainty on final payment and date of completion</td>
</tr>
<tr>
<td>Extent of competitive tendering</td>
<td>Low final payment</td>
</tr>
<tr>
<td>Risk transfer by way of contract</td>
<td>Risk avoidance</td>
</tr>
</tbody>
</table>

The contract in construction is often used by the client as an instrument for achieving *ex ante* certainty on price and contractually allocating risks to the contractor. There is a general (implicit or explicit) presumption in much construction economics that, as long as the client requires, the contract will serve his/her demands. It implies that the function of the contract is efficacious. This perspective deserves opposition both from the perspective of theory and practice.

Along with the development of transaction cost economics, we have become aware that the traditional assumption of perfect rationality in economics needs to be relaxed for addressing the problems of organisation. If the costs of using markets were not sometimes prohibitive, markets would exercise their magic power to coordinate all economic activities. However, there are many problems leading to market failure and inhibiting the prevailing of market contracting [Milgrom and Roberts,1992]. First of all, perfect foresight on the future is unachievable in the sense that all contingencies
cannot be foreseen in advance. Second, it is costly to formulate the available information about possibilities that can be foreseen into clauses and express them in written terms. Third, the enforcement of contract will be difficult if its performance is not verifiable by a third party. Taking these three factors into account, we have no reason to place too much optimistic expectation on the efficacy of contract in governing construction transactions. Moreover, excessive reliance on risk transfer by way of contract may make the consequence worse. The reason can be made clear by another numerical example.

Let's start by assigning the equal weight of 0.11 to nine priority factors in Table 2-4. Comparing the summed weighted value of priority variables, we find that the traditional method is the most suitable one for the client indifferent between all priority factors. In dealing with the same project, if the client becomes more risk averse, raising the weight of certainty and risk allocation to 0.3 (so the weightings of other factors are reduced to 0.059), the result will be different, design and build being the optimal solution this time. Does this advice make any sense? The answer depends on whether the attempt of risk avoidance can really reach its goal. We can contrast this case with the decision on the household's investment portfolio. As a rule, bond markets are less volatile than stock markets, so the risk-averse investor is often advised to hold higher proportion of riskless bonds. That is, keeping more bonds can help investors reduce risks and match the holding portfolio with their risk preference. However switching from traditional method to design and build purely on the basis of the client's risk considerations is not quite the same story. A risk-averse client may attempt to insulate himself/herself from risk exposure by using design and build to deal with a high-risk project. How far this intention can reach its goal depends on the effectiveness of contract. As will be clear in Chapter 4 and Chapter 6, this intention in fact is likely to lead to an opposite outcome. A brief preview of the reasoning is given here. When the contract is not able comprehensively to cover all the contingencies, the loopholes may create the opportunity of (economic) rent redistribution by way of ex post renegotiation. If one party is vulnerable to hold up, the incompleteness of contract may induce the exploitation of this opportunity by other parties. This possibility may trigger a bout of rent-seeking behaviour, such as disputed claims, leading to appeal for third-party arbitration and breeding distrust, all of which will be costly. The client's vulnerability to hold up, given the set of project attributes, is limited in the
last resort, by their ability to terminate a contract and switch to a replacement supplier. However, such switching will always be costly. The extent of this cost though will, we argue, depend in part on the procurement route used. Specifically, it will be more costly to replace a D & B contractor. The relevance of project attributes (likelihood of need to make post-contract changes; project complexity) lies, in a TCE approach, in their effect on the severity of the hold up problem. Specifically, for example, the greater the project’s likelihood of need for post-contract changes, the greater is the client’s potential vulnerability to opportunistic post-contract hold up. If a decision supporting method for procurement route selection is not able clearly to identify the potential transaction hazards, the recommendation will be of little value. For this reason, we claim that the sound advice on procurement route selection should be based on what the procurement systems can achieve in a given context, rather than the client’s risk preference itself.

2.3.3.7 The role of price

If the extent of tendering price competition has any importance to the client, it must be the expectation that the client can thereby get a building completed for less money. Providing the keenness of price competition varies as procurement routes, this factor is presumed to influence the client’s utility. The legitimacy of this perspective should be built on the robustness of the relation between the contract price and the outturn price. That is, we are concerned with whether the lower price stipulated in the contract will necessarily turn out to be a lower outturn price, as the literature implies. On several theoretical grounds, we cast doubt on this.

First of all, errors in cost estimation combined with the operation of the “winner’s curse” may lead to the realised costs of the winning bidder systematically exceeding expectation and so increase the contractor’s downside risks. The greater their underestimate, the more likely the contractor is, perhaps, to resort to opportunistic claims to recover their loss. This consequence is likely to nourish the client’s distrust, requiring more resource be devoted to precautionary measures, such as increasing the frequency of quality inspection and hiring more supervisory staff. Obviously this development can be relatively inefficient in terms of total outturn cost of the transaction to the client.
Second, incompleteness of information is likely given the contractor's bounded rationality or incomplete documentation from the client. The contractor is not endowed with perfect foresight on every contingency likely to happen in the course of construction, so the cost estimation must have errors. Besides, the documentation provided by the client is likely to be incomplete due to many reasons.

These two points create the ground for suspicion that a difference in bidding price between bidder A and B may be a poor predictor of difference in outturn price between A and B. To our knowledge, no empirical work exists to demonstrate that there is a significant correlation across the range of projects with given sets of attributes, between lower tender price and lower outturn price. In the absence of such evidence, we believe it is prudent to take an agnostic or sceptical position on this point.

2.3.4 Conclusions of Sec.2.3

After carefully examining the logic of MAUA, it is clear that potential weakness in the application to procurement route selection derive from three respects: (1) selection of priority variables; (2) spurious association of procurement routes with priority variables due to the assumption of complete contracting; and (3) the insensitivity to project attributes of the utility coefficients used to link routes to outcomes.

The possible alternative is founded on the transaction-cost-based approach developed in this study. This paradigm marks it off from MAUA in its stress on the alignment of procurement route with the attributes of construction transaction. That is, normatively this approach suggests the client should rest upon a procurement route most competent to attain his/her goal in the particular project context, rather than make a selection purely on the basis of his/her preference, in particular risk preference. The whole logic will be clear after presenting the rationale of the model in Chapter 4 and Chapter 6. I will go back to this issue in the concluding chapter, chapter 10.
2.4 Multidisciplinary perspectives

In recent years, CIB’s working commission 92 attempts to enrich the research of procurement routes by infusing multifaceted perspectives and extending the scope of analysis to higher level, that is, integrating the procurement route into the company’s whole operation strategy and placing this issue into the institutional context. In other words, the analysis of procurement routes can be reviewed in three levels (see Fig. 2-6). The first level considers the selection of procurement routes on the project-by-project basis, this study being a case. The next level discusses this issue as a part of the company’s integrated procurement strategy, where a bundle of projects replaces the individual project as the unit of analysis. Sometimes, the analysis allows the change of institutional parameters to accommodate the influence of legal systems or culture. The effort of including wider range of factors into consideration itself deserves encouragement, while the complexity of reality and the tractability of theory always exhibit a tradeoff relation. No matter what foundation or perspective is resorted to, the theory must stand up to falsifiability. Unfortunately, according to the current research results shown in collected essays of Rowlinson & McDermott (1999), it is difficult to identify the main cases or a coherent theory that this research programme intends to develop. Moreover, we are not sure whether the actual efforts will be made to make the theory falsifiable. Thus, it is too premature to give any confirmative assessment of the academic contributions of those works.

Fig. 2-6 Procurement route research in three contexts

Level 1
Single Project

Level 2
Company’s Operation Strategy

Level 3
Institutional Environment

---

International Council for Building Research Studies and Documentation (CIB) is one of the largest international organisations for the studies of construction.
In the broader sense, transaction cost approach employed in this study should belong to a part of this research movement. Though its focus is restricted to the project level, the basic framework is extendible and, what is more important, it can live up to the basic requirement for theory assessment used in the mainstream economic methodology: falsifiability.

2.5 Ending remarks of this chapter: where is the niche of this study?
The common goal among the four approaches classified in this chapter is to discover the factors affecting the merits and demerits of procurement systems. The evolving path of procurement routes research is depicted in Fig. 2-7. The most primitive method, the guideline approach, relies upon the observations of the past experiences, from which the positive regularities and the normative advice are derived. We can learn a lot from these guidelines, while I still lack a coherent and systematic way of thinking about the nature of procurement routes and the tradeoff relations between variables. Thus, the best strategy that we should pursue is vague if these crude guidelines are being followed. The pervasive use of this simple method of induction, more often than not, symbolizes the nascent stage of a research topic. The development of this topic moves forward along two branches: the multi-attribute utility approach and statistical approach. The overriding problems of the MAUA arise out of the selection of decision factors. For one thing, the variables are chosen quite arbitrarily. For another, the assumption of complete contracting is firmly held, leading to the fact that the efficacy of contract is overestimated so that the advice based on MAUA may be inappropriate. Another branch tries to apply statistical analysis to look for statistically significant factors in affecting the project performance, and objective relations between variables, like regression equations. Without careful examination of causality, we cannot tell whether the statistical associations are superficial or fundamental. To fill up the lacunae, the new direction-multidisciplinary perspectives try to encompass various aspects of the issues with respect to procurement systems. The attempt to systematically explore these issues is on the right track, while the ambiguity and loose use of terms should be overcome so that the falsifiable hypotheses derived from these perspectives can be tested; otherwise, it will be hard to assess the scientific value of these contributions.
After going through chapter 2 up to this point, some readers may pose a question: now that there have been a multitude of studies on this issue, what can distinguish the present work from previous attempts? The difference mainly stems from reasoning. Since the main goals of the existing studies are either to search for a method, which can guide the client to choose an appropriate procurement route, or to compare different types of procurement routes. In a sense, the purpose of this study is complementary to them. Standing on the established knowledge on procurement route selection, this study attempts to employ economic methodology to build up the causal relations between important variables found in the literature and observed procurement route selection behaviour. By means of the comparative institutional methodology, the theory set out in this study is both tractable and falsifiable and at the same time maintains the consistency and rigor of reasoning. What is more, its implications are likely to open up a new frontier in examining the nature of procurement systems (please see
Sec. 10.4). This is a possible way to upgrade the state of art of procurement system research to a higher scientific standard.
Chaper 3 An Introduction to Transaction Cost Reasoning

Drawing an analogy, Chapter 2 is just like an attempt to tear down an old building. However, destruction itself is not an end but a means by which the real problem can be thoroughly diagnosed. In combination with the valuable elements in the debris, the new building should be able to deliver higher standard service to users. But, without carefully organised plan, rebuilding would be no better than leaving it as it was. As a result, it will be of great help to set out a framework that can surely keep the progress of rebuilding on the right track.

In search of an apposite analytical apparatus for the purpose of this study, the methodology of transaction cost reasoning, or more generally comparative institutional analysis, can provide a well-founded underpinnings upon which further explorations can be built. However, in applying this methodology, there are two ways to go: one, called direct measurement approach (DMA), is to get the direct measurement of transaction costs; another one, called indirect measurement approach (IMA), is to operationalise the transaction and manage to empirically verify the postulated relations between the key determinants and governance selection. Before taking the next step, we have to fully take stock of their strengths and limitations. As will be discussed below, the key point lies in what elements of transaction costs are really relevant. In the ending section, an assessment of IDM and DMA will be made.

3.1 The genesis of transaction cost analysis

3.1.1 The rationale of Coase(1937)

In the traditional theory of value and distribution, the role of firm is no more than a black box in which a bundle of inputs are transformed into outputs. What the firm does is restricted to determine selling price and production quantity given exogenous product price. The key concern is how the magic of market can dole out scare resources to their most valued use. However, the lopsided weight of analysis placed on the importance of market mechanism is obviously incongruent with the fact that a large sphere of economic activities is coordinated within the firm. Inspired by this observation, Coase poses an intuitive question: why the firm exists at all. The answer he provided is that “there is a cost of using the price mechanism” so that in some cases it
is too costly to achieve efficient resource allocation through markets and the intervention of the visible hand by means of internal control, can be used as a remedy.

**Fig. 3-1 The value chain of a production process**

Mode I: buy decision

\[ B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \]

Mode II: make decision

\[ B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \]

Once the firm comes into existence, the firm will expend until where the cost of carrying out an additional transaction within the firm and by way of the market becomes equal. Let’s consider the value chain of a production process in Fig. 3-1. Activity D should be integrated if

\[ ATC_i + APC_i \leq P + ATC_m \]  
**Eq. 3-1**

where \( ATC_i \) is the average cost of organising activity D internally, \( ATC_m \) average cost of using market, \( APC_i \) average cost of internal production, \( P \) prevailing market price. If \( APC_i \) and \( P \) are cancelled out by assuming that D can be organised within the firms as efficiently as it is in a competitive market, the problems of “make or buy”, as Coase claimed, is equivalent to the problems of comparing costs of coordination through the firm or the market.

### 3.1.2 The methodology in Coase (1937)

It is to be noted that, as a matter of fact, Coase (1937) deals with two problems: one is about the existence of the firm and another is related to the boundary of the firm. As will be shown later, the nature of these two issues is quite different, while there is a common methodology underlying Coase’s reasoning: discrete structure analysis [Simon, 1978]. As its name dictates, the focus of analysis now is to compare the relative desirability of limited number of organisation alternatives, rather than to deter-
mine the condition of equilibrium at the margin. That is, the issues concerned here is like “under what circumstances an option will be more preferable”, rather than “at what quantity level the equilibrium conditions can be met”. Let’s see how Coase carried out his analysis.

To fully grasp the reasoning in Coase (1937), there are two points deserving special attention: one is what alternative choices are under comparison and another is the objective function of comparison.

First of all, when speaking of the problem about the existence of the firm, Coase departs from a proposition that production factor providers are specialized autonomous entities. Thus, organizing a production process completely by means of market mechanism means that a factor of production has to make a series of contracts with whom he wants to cooperate. The firm can act as an intermediary role to reduce the number of contracts to one, so the repeated costs of using price mechanism can be saved. This is where the firm displays advantage over the market. This argument can be illustrated graphically. In Fig. 3-2, for simplicity, the core part of production is assumed to include only four factors of production A, B, C, D, each with transaction relationship with other three factors. Their external network of transaction like $EF_{A1} \sim EF_{A4}$ is assumed to be no connection with each other. In this case, the extent of complexity of market coordination that the firm can help alleviate strengthens as the number of core factors increases. Saving repeated costs of using price mechanism then provides the fundamental raison d’etre of the firm. From this example, it is clear that the pair of alternative forms of organisation under discussion is production through the transaction network between production factors versus production through internal management within the firm.

Regarding the boundary of the firm, the starting point of analysis is slightly different. Given a firm that has engaged in a stage of production, the issue of interest here is the best way of governing the production activity to the point where the costs of organizing additional transaction internally and procuring from the market become equal. In the mundane term, this is a choice between make or buy, as shown in Fig. 3-1.
For a firm specializing in production activity E (see Fig. 3-3), the decision on whether it is desirable to integrate upstream activities (D, C, B, A) and how far it should go depends on the interaction two forces: on the one hand, due to limited attention of decision makers, managerial returns to scale is conventionally assumed to be diminishing; on the another, marginal transaction cost is expected to be lower as the volume increases. Coase asserts that the optimal scale of the firm is conditional on the point where the marginal cost curve of management and market transaction are crossed. That is, in Fig. 3-3 the optimal boundary for the firm is B+C+D+E.
A summary table is shown in Table 3-1. In exercising discrete governance analysis, it should be borne in the mind what alternative options are under comparison and what are the conditions that the desirable option will compete out of others; otherwise, the reasoning would be messed up. In fact, a lot of confusion appearing in the works following Coase’s lead results from the ignorance of reexamining the premise of the analysis\(^1\). We will have more words on this point later.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Alternatives under comparison</th>
<th>Objective Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why firms emerge</td>
<td>Production in the presence of firms versus in the absence of firms</td>
<td>Choose governance with lower transaction costs</td>
</tr>
<tr>
<td>Why firms integrate</td>
<td>Production with integration versus without integration</td>
<td></td>
</tr>
</tbody>
</table>

### 3.1.3 Some reflections on Coase(1937)

As far as Coase’s original intention is concerned, his theory is a consequence of compromising tractability and reality of analysis. Its vague exposition notwithstanding, Coase (1937) has generated profound influences:

1. To justify the fact that organisations supersede markets in some respects, he proposed that transaction costs should be added into the cost-benefit analysis of organisation.
2. The existence and boundary of firm is a comparative exercise: among all alternatives, efficient form of organisation will emerge and expand.
3. In the comparative organisation analysis, production cost is implicitly assumed to be constant across different organisation.

We have witnessed that these concepts or principles have generated substantial repercussions on the evolving science of organisation in the last two decades. However, some issues attracting heated debate in academic seem to root in this essay as well. First of all, his attempt to single out some important components according to the stage of contracting (pre-contracting, contracting and post-contracting stage) gave

\(^1\) For example, in explaining why firms integrate, the premise is that firms have existed.
stage of contracting (pre-contracting, contracting and post-contracting stage) gave a
direct impetus to the birth of the contractual perspective on the theory of the firm,
while lack of an operational definition makes this concept less likely to be empirically
falsified, which is often ascribed to the reason why it is charged with tautology, fail­
ing to be accepted as a legitimate explananda by economists before 1970s [William­
son, 1975]. Second, the claim that the selection of organisation can be interpreted in
terms of transaction costs has merited critiques because of two reasons: (1) efficiency
(or transaction cost savings) is not the only relevant criteria in the analysis of organi­
sation and other factor like power should be taken into account; (2) even if efficiency
is viewed as of paramount importance, what should be economised is the sum of pro­
duction costs and transaction costs, instead of the latter only.

3.2 Discrete structural analysis
From Sec.3.1, it is evident that Coase (1937) has made contributions in two respects:
(1) he introduces the concept of transaction costs; (2) he sets out the rationale of com­
parative institutional analysis. With transaction costs included in the comparative in­
stitutional framework gives us a tool for investigating the relative advantages of gov­
ernance structures. The power of this tool was not as impressive as it is recognized
today. The problem in the large part stems from the ambiguous expression of transac­
tion costs, causing the sluggish development of the research methodology for transac­
tion cost reasoning prior to 1970s. In view of these obstacles, the studies following
Coase’s lead, on the one hand, have to clarify what is the operational meaning of
transaction costs and on the other, develop a methodology that can make theory refut­
able. Of course, these two problems are interlinked. This section will present a brief
account of the ways that the transaction cost reasoning can be implemented. The de­
development of transaction costs will be left to the next section.

3.2.1 A general illumination
The overriding themes for organisational studies are to explain the selection of mode
of organisation at the individual level and frequency dominance of organisational
forms at the population level. A rational model based on transaction costs entails
showing there is a clear link between costs of organisation and organisation selection
behaviour. Intuitively, a good reason for explaining why governance structure 1 (GS1)
is chosen instead of GS2 lies in lower costs of using GS1. Thus, whether a reliable comparison of transaction costs between GSs can be made becomes a critical part of the transaction cost theory. To see the possible ways of doing this, a simple mathematical example will be helpful. We are comparing two variables each capturing transaction costs: A and B. Variable A captures the transaction costs associated with GS1, and variable B the TCs of GS2. Their relative magnitude can be demonstrated by: (1) getting the real value of each of the two variables; or (2) assessing the sign of A minus B though logical inference and testing to see if GS1 dominates GS2 in frequency of use wherever the sign of this inferred difference is negative, and vice versa whenever the sign is positive, i.e. Williamson’s approach; or (3) presenting A and B in a common non-monetary numeraire (e.g. ‘number of disputes’) and comparing their respective coefficients.

Certainly, as we change the method to be used, the data requirements will vary. For IMA data is required on measurable transaction attributes and on relative frequencies with which governance structures are used, for transactions with different attributes. For DMA data is required on measurable transaction attributes, the relative sums of all transaction costs for transactions with similar attributes under different governance structures, and the absolute values of the comparatively significant elements of total TCs. Data requirement can be thought of in two dimensions: the number of transactions on which data is required; and the number of dependent and independent variables to be measured per transaction. In terms of the latter, the data requirement of the direct measurement approach (DMA) is, normally, higher. However, the number of transactions on which data is required will depend upon the number of transaction attributes deemed relevant. If, for instance, only one attribute is judged relevant, then both the number of transactions on which data is required and the number of elements of TC needing to be measured per transaction will be less than in an analysis where multiple attributes are judged relevant, potentially varying independently and offsetting one another.

It might intuitively appear that the number of transactions required for a DMA study to yield statistically significant results would be smaller than in the case of an IMA study. However, once it is recognised that some of the more comparatively important TCs (e.g. cost of disputes) have low probability of occurrence but high impact when...
they do occur, on the one hand, and on the other that in IMA it is possible to conduct the analysis of the effect of each attribute independently in terms of *ex ante* subjective probabilities using software such as LIMDEP, it follows that this is not in fact necessarily the case.

One further distinction is that the DMA approach is potentially pragmatic. It is not *essential*, in this approach, to generate *ex ante* predictions of the sign of the difference in TCs between GSs. It is enough to observe such difference *ex post*. Transaction attributes still have to be measured, but simply to ensure that like is being compared with like. Whereas, it is essential, in the IMA, to theorise the *expected* sign of difference on the basis of observed transaction attributes. Moreover, whilst the IMA has both normative (advice-giving) and positive (explanatory) objectives, it is fair to say that the DMA *can* (but need not) have mainly normative (advice-giving) ones. Any resulting explanation of why the difference in TCs has the sign it is found, empirically, to have, can be constructed retrospectively.

The benefit of making this distinction is to provoke thought about how to choose a appropriate research strategy for applying transaction costs. The next subsection will give a brief description of how the reasoning of these two approaches will run.

### 3.2.2 A Methodological Comparison of the DMA and IMA

The first, and intuitively appealing, way of applying transaction cost reasoning is called direct measurement approach, requiring the identification and direct measurement of the magnitude of elements of transaction costs and the linking up of the sum of these costs with the governance structure used (e.g. hierarchy or market). In contrast, the indirect measurement approach is devised to seek the key determinants that *explain* the relative efficacy of governance structures in terms of transaction costs. We can compare these two approaches in a somewhat formal way.

Assume that using a governance structure, say the market, is attendant with *n* categories of transaction costs, labelled as $\text{TC}_i, i=1\ldots n$. According to the DMA, the components of $\text{TC}_i$ need to be identified first. Then, whether those costs are to be used as explanatory variables for organisation selection or simply used in empirical comparison of TCs across governance structures, they need to be quantified. Summing up all
the quantified items of transaction costs can yield the total costs of running a governance structure. Of alternative choices, the one with lower $\Sigma TC_i$ is regarded as more efficient. This means that in using DMA, the absolute level of elements of transaction costs, and the relative or ordinal level of total transaction costs, must be known in order to determine the efficient organisation in most cases. Because there are several, perhaps many, elements of transaction cost, and because the relative magnitude or 'weight' of each element of transaction cost is not known in advance, DMA requires absolute or cardinal measurement of such elements, even though it requires only ordinal ranking of total TCs. At this point it is important to note that, although there may be many categories or elements of TCs (e.g. legal fees for contract writing, court costs, quantity surveyor fees, inspection costs, contractor selection costs, insurances, performance bonds, etc.) mainly these can be traced to just two ultimate sources. Most TCs arise either as costs of collecting information and measuring outcomes or as costs of opportunism. This will be significant later in the argument.

In contrast to DMA, the IMA stresses comparative difference. If there are two alternative governance structures under review, their running costs can be expressed as $TC^1 = \sum_{i=1}^{n} TC^1_i$ and $TC^2 = \sum_{i=1}^{n} TC^2_i$. The transaction costs difference between them is

$$\Delta TC = TC^1 - TC^2 = \sum_{i=1}^{n} TC^1_i - \sum_{i=1}^{n} TC^2_i$$

$$= \sum_{i=1}^{n} \Delta TC_i \quad (\Delta TC_i = TC^1_i - TC^2_i) \quad \text{Eq. 3-2}$$

Now, in the IMA, we are not concerned either with the magnitude of each component of $TC^1$ and $TC^1$ or $\sum_{i=1}^{n} \Delta TC_i$, but with what factors will contribute to their difference and thus to the sign of $\Delta TC$. Put another way, we attempt to find out the following relation:

$$\sum_{i=1}^{n} \Delta TC_i = f(X_1, X_2, \ldots, X_k) \quad \text{Eq. 3-3}$$
The variables $X_i, i=1...k$, are often called transaction attributes in the TCE literature. In this way, we can get over the difficulties in measuring the absolute level of transaction costs, while it is still possible to derive refutable hypotheses. Relevant attributes are generally suggested to include: degree of uncertainty/complexity, degree of asset specificity and frequency.

The procedures for implementing DMA and IMA are summarised in Fig. 3-4. It is clear that the step of identifying categories of transaction costs that may occur in the course of transactions is their common point of departure. After identifying the sources of transaction costs, we have to interpret why transaction costs will change as transactions are organised in a different way, leading to predicted frequency dominance of one GS over another. These two approaches branch out here. If the DMA is adopted, transaction costs have to be broken down into a comprehensive list of elements for which we have no adequate *a priori* reasons to rule out their comparative significance (i.e. reasons to reject in advance the proposition that they may vary significantly between one GS and another). We can make a prediction purely on theoretical grounds, such as: *for the transactions with attributes set A, total TCs will be lower in GS1 than in GS2*. A corollary is that GS1 is predicted to be more widely used than GS2, so GS1 will occupy a larger proportion in the population. We thus have two predictions. If the first hypothesis is to be tested empirically, first of all, all identified elements of transaction costs or their difference between GSs have to be measured in the same unit. If the unit is money, we can obtain the estimate of the real value of costs. If the unit is an arbitrarily chosen numeraire whose ratios with each element of transaction costs can be relatively easily estimated, we can obtain the estimate of costs in a different unit. With these data in hand, it is easy for us to generate statements about the actual sign of $\Sigma \Delta TC$ between GS1 and GS2. Meanwhile, to test the second hypothesis, the actual frequency distribution of GSs for transactions of given attributes has to be investigated. At the last node, theoretical predictions are to be verified by examining whether the dominant GS has lower average transaction costs.
Classify TCs by sources of origin into:
(a) Costs of information collection/measurement
(b) Costs of opportunism

Explore reasons why particular attributes lead to $\Delta TC$ between GSs

Make a comprehensive list of transaction cost elements that may be different across GSs

Provide a priori reasons for identifying relevant transaction attributes

Make predictions of relative TCs of GS1 and GS2 for transaction of given attributes

Verify theoretical predictions by testing whether GS selection is aligned with the transaction attributes.

Generate predictions of frequency dominance for GSs with lower TCs by comparing relative predicted $\Sigma TC$ of GS1 and GS2 for given attributes

Verify whether the frequency of the predicted GS is predominant for given attributes

Observe actual frequency predominance of GSs for given attributes

Observe actual frequency predominance of GSs for given attributes

Measure each type of TCs

For sets of projects of given attributes, verify prediction of sign of difference of TCs between GS1 and GS2

Verify theoretical predictions by testing whether frequency predominant GS displays lower TCs.

<< Empirical investigation at the individual level
<< Empirical investigation at the population level

$\rightarrow$ Theoretical reasoning

$\rightarrow$ Empirical testing
However, if the IMA is adopted, we have to use theory to point out transaction attributes that will make governance structures perform differently and then predict the predominance of a governance structure according to these attributes. Take the standard example of TCE. When hierarchy (GS 1) is compared with market (GS 2), their relative advantage will be changing as the degree of asset specificity. The refutability of this theory is built on the following prediction:

\[
\begin{align*}
\text{If } k > \bar{k} & \quad \text{then } \Delta TC > 0, \text{ so GS2 will be dominant over GS1.} \\
\text{If } k < \bar{k} & \quad \text{then } \Delta TC < 0, \text{ so GS1 will be dominant over GS2.}
\end{align*}
\]

In other words, we focus on the conditions in which hierarchy (or market) is posited to be desirable. To prove the validity of this sort of theoretical inference, transaction attributes and the actual selection of GSs have to be recorded so as to generate statistical regularities. From the direction of attributes acting on the selection of GSs, given the assumption that behaviour is transaction cost minimising (profit maximising), the theoretical prediction of effect of attribute on difference in TCs between one GS and another can be indirectly verified.

It is clear that the success of DMA is dependent on two factors: (1) completeness of the list of elements of transaction costs; (2) accuracy of measurement of each element in that list of transaction costs. First of all, it will be ideal to draw up an exhaustive list of transaction costs for every governance structure. Of course, if it can be justified, the items of minor significance can be ignored. Second, there may appear some practical problems in measurement. Using Eq.3-2 to demonstrate GS 1 is advantageous over GS 2 entails

\[
\Delta TC = TC^1 - TC^2 = \sum_{i=1}^{n} [TC^1_i - TC^2_i] > 0
\]

This inequality can be determined by relative magnitude of aggregate transaction costs of \(TC^1\) and \(TC^2\). In fact, not all categories of transaction costs need to be estimated. Only the categories with a first-order difference need to be computed. The success of DMA is conditional on whether the precise estimate of these costs can be obtained empirically.
In contrast, IMA is, in essence, a deductive method, by which we mean the explained phenomena are rationalised by its logic system, inferring from the first principle – behavioural assumptions. Thus, the validity of theory relies on that of assumptions. According to the mainstream economic methodology [Friedman, 1953], the touchstone for a scientific economic theory is not in the realism of its assumptions, assessed directly, but in the predictive power of the theory incorporating those assumptions. Thus, to what extent the theoretical prediction can fit the data is the main concern.

At the end of this chapter, we will come back to discuss which approach is more suitable for the case of this study. Before that, we first have to know what are the categories of transaction costs that may really matter and why.

### 3.3 What Are Transaction Costs?

#### 3.3.1 Why transaction costs exist at all?

Transaction costs originate from the observation that realising gains from specialisation through markets costs something. However, where are these costs from? Stigler (1967) has given an inspiring answer: it is *the costs of transportation from ignorance to omniscience*. Between ignorance and omniscience lie two gaps [Dosi and Orsenigo, 1988]:

1. **Information gap**

Predicting the future always attracts enormous amount of intellectual efforts. This enduring fascination seems to have something to do with the nature of this task being unattainable. For the fact that what the future will become is beyond our grips at the point of decision-making, the consequences of a commitment made in advance (often in the form of contracts) must involve uncertainty. In the expected utility theory, Knightian risk and uncertainty can be modelled as (objective or subjective) random variables. Collecting information and building up a complete decision tree are the main sources of costs. For Austrian economists, since we are ignorant about mean-and-end structure in advance, they emphasise the importance of discovery process, discovering what is worth searching for [Littlechild, 1986], in contrast to neoclassical economists who stress to search information necessary in the given decision tree. Some radical subjective economists, like Shackle and Lackmann, assert that the future is not only waiting for being discovered, but needed to be created. Indeterminacy
characterises this line of inquiry. Provided the idea of 'ignorance' or 'indeterminacy' is taken seriously, no longer are information problems appropriate to be treated as optimisation within the constraints of information costs, and the existence of information gap should be respected in the analysis.

(2) Competence gap
Were the future to be lack of radical uncertainty just as in the neoclassical world, will all information problems vanish? Not quite. There are two factors affecting human competence in dealing with complex problems [Williamson, 1975]:
(a) physical limits:
More often than not, grave decision problems contain much information than man can receive, store, retrieve and process so that it is costly to draw out complete state-contingent contracts.

(b) language limits
Communicating by the use of words, numbers or graphics possibly exists dissonance in mutual understanding, aggravating the effectiveness of contracts in governing behaviour of trading parties as transactions unfold.

From the above discussion, in principle, it can be said that transaction costs ultimately result from informational complexity and uncertainty [Dietrich,1994]. Nonetheless, we need to know what kind of information is of utmost relevance. Ouchi’s view is pertinent here:

*Transaction costs arise principally when it is difficult to determine the value of the goods or service. [Ouchi,1980]*

It follows that information about the (fair) valuation of the goods or service deserves most attention. Focusing on the transaction per se can be a prolific research strategy and viewing the process of transaction in terms of transaction costs can engender valuable insights on the functioning of institutions, as we have witnessed in the progress of New Institution Economics. This point will be developed further in Chapter 6.
3.3.2 The concept of transaction costs

Transaction costs arise when goods and services are transferred across technologically separable phases of production or distribution [Williamson, 1985]. How costly the process of transferring is, has bearing on the governance structure being used, i.e. the way this process is organised. What reasons make trade incur costs? An intuitive guide is given in Coase (1960):

*In order to carry out a market transaction it is necessary to discover who it is that one wishes to deal with, to inform people that one wishes to deal and on what terms, to conduct negotiations leading up to a bargain, to draw up the contract, to undertake the inspection needed to make sure that the terms of the contract are being observed, and so on.*

Taking Coase’s lead, Dahlman (1979) summarises the key items of transaction costs according to the different phases of transacting as shown in Table 3-2:

<table>
<thead>
<tr>
<th>Stage of Transacting</th>
<th>Pre-contracting Stage</th>
<th>Contracting Stage</th>
<th>Post-contracting stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items of transaction costs</td>
<td>Search and information costs</td>
<td>Bargaining and decision costs</td>
<td>Policing and enforcement costs</td>
</tr>
</tbody>
</table>

According to Dahlman, transaction costs are rooted in lack of information. To get over the initial incompleteness of information, both sides (buyers and sellers) have got to locate suitable trading counterparts, formulate acceptable decision criteria, negotiate for better terms and ensure the realisation of *ex ante* promises. All these efforts involve not only real resource consumption, like telephone bills, information subscription fees, transportation fees, but also the opportunity costs of time spent. In principle, the costs of increasing the available information would turn out to be worthwhile if the quality of decision can be correspondingly improved. And the extent of these costs will affect the efficiency of market relative to hierarchy. This line of definition gives an intuitive understanding to transaction costs caused by, what we can call, the first round of information problems. However, the consequences of incomplete information are more than the necessity of overcoming an information gap, since the strategic manipulation of information by the informed party can be more devastating to the
transaction relation. The later problem, in terms of TCE, results from the interaction between bounded rationality and opportunism. The rationality of decision makers is bounded due to both the incompleteness of the current available information and some extent of ignorance of the future as well as the neurophysiological and language limits. If the extent of incomplete information facing transactors is different, the opportunistic motives for taking advantage of the other's vulnerability can be triggered to create the second round of information problems – the informed party tries *ex post* to reap a larger share of gain from trade than agreed upon *ex ante* by disclosing distorted information to mislead his/her trading party. Provided all the opportunistic strategies can be implemented with no resource cost, the efficiency of carrying out this transaction won't be altered\(^2\), but the distribution of benefits will be. In the general case, the story doesn't wind up here because few people are willing to give up the expected gains without struggling. If opportunistic appropriation demands are made and resisted, the consequence for resource costs will be far-reaching. First of all, renegotiation before the expiry of contract will be a required effect of opportunistic behaviour and how costly it will be depends on the efficacy of dispute resolution mechanism. Second, the lagged effect of a bitter experience will be reflected in the next transaction, leading to two costly outcomes: (1) To reduce the future occurrence of renegotiation, more resources will be devoted to formulating the contract in more detail, preventive actions will be taken, such as bonds, and an effective mechanism for settling disputes will be required to be set up. These all entail costs; (2) A possible welfare loss may happen due to less productive technology being employed to reduce asset-specificity and thus the vulnerability to such hold-up strategies. According to the above analysis, transaction costs should consist of two parts:

\[
TC = TC_I + TC_{II} \quad \text{Eq. 3-4}
\]

The first category or type \((TC_I)\) is caused by the first round of information problems, excluding the intervention of opportunism, while the second category or type \((TC_{II})\) indicates the consequential costs arising from the strategic exercise of information

\(^2\) This is because the total gain from this transaction is still the same, because it is determined by the surplus of value of output over the resource costs of producing and transacting that output. This surplus is the 'gain of trade', and constitutes the economic rent whose appropriation by one party, rather than another, is the object of games of opportunistic strategy.
asymmetry by the informed party. As will be clear in the next subsection, this is a key to understand why there are divergent views on transaction costs in the literature.

3.3.3 A Systematic Presentation Of Second-Round Information Problems

The major information problems in the transaction are described in Fig. 3-5, from which five outcomes may result. First of all, the producer has to make a decision between (1) inefficient but less vulnerable technology (or GS) or (2) efficient but vulnerable technology or GS. If the producer goes for the upper branch at Node I with the probability of $\pi_1$, the well-documented problem of under-investment in efficient technology (or GS) becomes the issue of interest (Grout, 1984; Grossman and Hart, 1986; Hart & Moore, 1988; Hart, 1995). The productivity shortfall arising from the adoption of less resource-efficient technology (or GS) is tantamount to social welfare loss, $\Delta W$. If the vulnerable path is taken, two possibilities may happen at node II: opportunistic behaviour will occur or not, with the probability of $\pi_2$ and $1-\pi_2$. If the vulnerability is not exploited by the trading party, the $\text{TC}_\text{H}$ will vanish. However, if it is exploited, the vulnerable party (at Node III) may either accept the hold-up demand to bear an extra payment $\Delta P$ and/or value loss due to inferior quality $v(\Delta Q)$ relative to the outcome to be expected under perfect competition, or refuse to yield to the threat of hold-up. If the opportunist backs down after a series of disputes and renegotiation costs (at node IV), the additional transaction cost occurring in this process is labeled as $\text{TC}_{c_1}$. But sometimes, the dispute is taken to arbitration or the court. The outcome may be that (1) the held-up party wins the lawsuit, successfully avoiding extra costs due to unfair demand; but (2) in the worst outcome, the vulnerable party may be forced to accept the hold-up demand after struggling, thus incurring $\text{TC}_{c_2}$ (legal fees of both parties plus costs of delay) in combination with extra costs, $\Delta P+v(\Delta Q)^3$.

Therefore, a more complete expression of expected type-II transaction cost ($\text{TC}_{\text{II}}$) consists of three items as follows:

\[ \Delta P + v(\Delta Q)^3 \]

For simplicity, an assumption is made here that $\Delta P + v(\Delta Q)$ is equal to the maximum appropriable quasi-rent. It indicates the upper limit of the loss from conceding to the hold-up demand.
\[ L = \pi_1 \Delta W + (1 - \pi_1) \pi_2 \left[ \pi_3 [\Delta P + \nu(\Delta Q)] + (1 - \pi_3) \right] + (1 - \pi_2) \pi_4 \left[ (1 - \pi_4) (T_{C1} + (1 - \pi_4) (T_{C2} + \Delta P + \nu(\Delta Q))) \right] \]

Welfare loss due to inefficient production  
Loss due to acceptance of hold-up demand 
Loss due to refusal of hold-up demand

This formula delivers a clear message that, in essence, \( T_{C1} \) and \( T_{C2} \) are probabilistic concepts, implying that empirically large samples will be required to get a reasonable estimate of representative values for the population. In this respect, \( T_{C1} \) and \( T_{C2} \) are different in nature.

Fig. 3-5 A structural presentation of information problems in a transaction: possible payoff (outcomes) for a party vulnerable to hold-up

3.3.4 The nature of \( T_{C1} \) and \( T_{C2} \)

From an overall perspective, the TCS that matter are real resource-incurring TCS, which reduce the total economic gain from transaction. However, from the perspective of a single transactor, trying to optimize their profit, behavioural uncertainty (especially opportunism) gives rise to another kind of cost – the negative difference between the promise (on the basis of which the transaction is agreed) and the delivery or
We call this rent-transferring TCs. This is admittedly an appropriation or transfer of economic rent from one transactor to another, rather than a resource-cost. This type of TCs may either be anticipated or not. Where they are anticipated, they result in loss from refusal to engage in a potentially efficient transaction. This loss (the opportunity cost) is measured by the difference between the total economic net benefit that would have accrued from the aborted transaction, in the absence of opportunism, and the economic net benefit of the best alternative arrangement or transaction chosen instead. Where they are not anticipated, they result in a loss of expected return to the transactor, which may reduce return below opportunity costs.

We classify the elements of TC₁ and TC₁I according to the nature of those costs, resource-incurring or rent-transferring, in Table 3-3. By definition, TC₁ is caused by the first-round information problem in the absence of opportunism. All the costs used to fill up the information gap, as listed in Table 3-2 are of resource-incurring type. However, this is not the case for TC₁I, which meanwhile contains resource-incurring TCs, such as legal fees and opportunity cost of delay, and rent-transferring TCs, such as extra payment (ΔP) and loss due to inferior quality (v(ΔQ)). The purpose of introducing this classification system is to highlight that the fundamental distinction of TC₁ and TC₁I lies in differential measurement difficulties. This is the key point that we will make for distinguishing the applicable scope of the IMA and DMA.

Table 3-3 A classification system of transaction costs

<table>
<thead>
<tr>
<th></th>
<th>TC₁</th>
<th>TC₁I</th>
</tr>
</thead>
<tbody>
<tr>
<td>**Resource-</td>
<td>1. cost of information</td>
<td>1. legal fees and opportunity cost of delay (TC₂)</td>
</tr>
<tr>
<td>incurring TCs</td>
<td>collection and search costs</td>
<td>2. welfare loss due to adoption of inefficient technology or GS (ΔW)</td>
</tr>
<tr>
<td></td>
<td>2. cost of bargaining</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. cost of measuring</td>
<td></td>
</tr>
<tr>
<td></td>
<td>performance of agents or</td>
<td></td>
</tr>
<tr>
<td></td>
<td>quality of products</td>
<td></td>
</tr>
<tr>
<td>**Rent-</td>
<td>extra payment (ΔP) or loss</td>
<td></td>
</tr>
<tr>
<td>transferring</td>
<td>due to inferior quality (v(ΔQ)</td>
<td></td>
</tr>
<tr>
<td>TCs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.3.5 The ways of modelling transaction costs

Outside of the comparative studies of economic organisation, the most common way of theorizing transaction costs is to treat them as another category of cost, that can be "formalized in a transaction cost function analogous to a production function" (Niehans, 1987). In the simplest form, the transaction function is often treated as in proportion to the number or volume of transactions. The transaction cost simply plays the same role as a tax, taking away a part of revenue from the receiver (Allen, 1991). This setting can be frequently seen in the works of financial economics and monetary economics, but rarely found in the studies of comparative economic organisation and economic history. Why? The key point lies in whether the trading parties are able to manipulate private information to change the counterparty's evaluation of this transaction to their own benefit. If the price cannot be determined independent of the trader's actions, the information problem will not terminate in the first round and its side-effect should not be ignored. Put another way, if the value of the underlying object of transaction is a stochastic variable, determined by acts of God, \( TC_1 \) is the principal component of transaction costs and its magnitude will generally vary as the number or volume of transactions. Buying shares is a case. The expected return of this investment is not affected by the individual buyer or the seller, so the costs of this transaction mainly belong to the category of \( TC_1 \), increasing as the volume of trading. Buying shampoo in the neighbouring store of Boots, Drugstore or Tesco is another example. To find the best offer, the time and travel cost spent in every search is quite similar. The transaction cost in this case is proportional to the number of searches. As a result, it is sensible to treat \( TC_1 \) as a tax-like cost in modelling. However, this is not the case when the second round of problems may be triggered by moral hazard and asset specificity. Difficulties in enforcing the performance of the agent as well as the irreversibility and lock-in effect of lump-sum investments are two contributing factors to transaction problems. The transactions subject to these two problems are characterized by (1) the dependence on the agent's efforts to produce goods or service; and (2) a longer period of transacting time than spot exchange. Because the outcome value of the transaction object, such as quality, can be changed by the trading parties, its value is not only affected by uncontrollable uncertainty (acts of God), but by behavioural uncertainty due to the manipulation of asymmetrical information. Meanwhile, if the transaction requires a period of time rather than being completed on the spot, the influence of post-contractual interaction between trading parties will be growing.
weighty, making $\text{TC}_\text{II}$ dominate over $\text{TC}_1$. When the different types of transaction costs become dominant, a principal effect will be on the appropriate mode of their modelling. For the category of $\text{TC}_1$ - dominant transactions, transaction costs can be approximately depicted as a tax. However, for the category of $\text{TC}_\text{II}$ -dominant transactions, transaction costs are not the "fixed costs" per transaction, but are governance-dependent. That is, the right way of approaching transaction costs in this case is not from the number or volume of transactions but from the way that the transactions are organised, in that the consequential costs arising from behavioural uncertainty are a joint function of attributes of transaction and attributes of governance structures. Only when the chosen governance structure is best suited to the characteristics of the project can the transaction costs be minimized relative to other alternatives. In fact, the theoretical role of transaction cost of the first category is quite different from that of the second category. More often than not, the theorists are forced to consider the effect of transaction costs just because of explanatory failures of the frictionless model. Thus, transaction costs are treated as another category of cost like production cost in the marginal analysis. In contrast, the transactions of the second category are generally associated with the organisational issues where involve coordination of trading parties in a certain period of time.

Table 3-4 A Summary table for the comparison of $\text{TC}_1$ and $\text{TC}_\text{II}$

<table>
<thead>
<tr>
<th>Characterization of transaction</th>
<th>$\text{TC}_1$</th>
<th>$\text{TC}_\text{II}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is large lump-sum specific investment involved?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Alterability of the value of outcomes by strategic behaviour</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Sources of Uncertainty**

- $\text{TC}_1$: Parametric uncertainty
- $\text{TC}_\text{II}$: Parametric uncertainty, Behavioural uncertainty

**Way of modelling**

- $\text{TC}_1$: Transaction cost as a tax
- $\text{TC}_\text{II}$: Transaction cost as the cost of governance

**Standard Case**

- $\text{TC}_1$: Purchase of shampoo from the neighbouring stores
- $\text{TC}_\text{II}$: Comparative institutional analysis, such as choice of procurement route
The points of difference between these two approaches are summarised in Table 3-4. Whether there is the lump-sum specific investment and alterability of the outcome value of the transaction object by trading parties are two important characterisations of transactions that can be used to distinguish the condition under which each approach is analytically advantageous. This factor has direct bearing on the sources of uncertainty prevailing in the transaction. Parametric uncertainty is uncertainty facing all parties and itself will not be affected by an individual party's strategy. The expected return of shares in the perfect competition financial market is a case. In contrast, uncertainty may be caused by the possibility that the strategies taken by the traders may drift out of contract stipulations. Actually, this is a manifestation of the incomplete contract.

Finally, from the above analysis, the best research strategy for modelling transaction costs should be determined by the composition of $TC_I$ and $TC_H$ of the transaction at issue. Transaction cost as a tax is a sensible simplified view of examining the effect of transaction costs if $TC_I >> TC_H$, while we have to put the analysis of transaction costs into the context of governance structure if $TC_I << TC_H$. This provides a basic reason why these two views coexist in the literature.

### 3.3.6 Which Way To Go: DMA V.S. IMA

With reference to Table 3-5, the DMA and IMA are distinguished from the perspectives of (1) reality of assumptions, and (2) measurability requirements. The DMA has advantage in requiring less demanding behavioural assumptions, that is, whether the agent is able to find the optimal solution is not critical to the verification of the theory in that transaction cost itself has to be measured. In this respect, the IMA seems more vulnerable to critiques from economic methodologists. In a pure deductive model, the result of analysis can be derived from the basic premises, the proof of mathematical theorems being the case at issue. This has been the ultimate goal for natural or social scientists. But the problem is that the reality is so complicated so that we can hardly find a premise that is both realistic and tractable. It is obvious that most researchers appreciate the latter more than the former. This view leads to the pervasive acceptance of assessing the theory by way of its predictive power instead of realism of assumptions. Following this mainstream perspective, the question asked of the IMA is
whether the more or less unrealistic assumption will bring too many errors so that the validity of the theory is severely affected. Moreover, if unexpected results are obtained, the interpretation may be ascribed either to (1) the mistaken selection of variables, (2) inappropriate measurement of variables or (3) the wrong behavioural assumption(s). However, the test of hypotheses by empirical investigation is actually sensitive to (1), (2) and (3). That is, only when all three premises are valid can we obtain statistically significant results.

Table 3-5 Comparison of advantage and disadvantage of the DMA and IMA

<table>
<thead>
<tr>
<th>Dimension for Comparison</th>
<th>Direct Measurement Approach</th>
<th>Indirect Measurement Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Advantage</td>
<td>Disadvantage</td>
</tr>
<tr>
<td>Reality of assumptions</td>
<td>No need to assume <em>ex ante</em> efficiency.</td>
<td></td>
</tr>
<tr>
<td>Measurability of require-</td>
<td>Need to obtain reliable measures for transaction cost elements of first-order importance</td>
<td>The actual choice can be easily observed and transaction attributes can be measured.</td>
</tr>
<tr>
<td>ments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Regarding measurability requirements, there is no way out of context to make a judgment about which approach is better. Any reasonable comparison has to be made on the basis of the decision context. As enunciated before, modelling transaction cost as a tax is a sensible way to simplify the problem for dealing with the transaction with parametric uncertainty only. This is also where the DMA is relatively more feasible in that the costs in proportion to the number or volume of transactions can more easily be estimated reliably. In contrast, if the main sources of transaction costs stem from behavioural uncertainty, the potential costs incurred will go beyond resource-incurred transaction costs and include rent-transferring transaction costs such as extra prices paid or downgraded quality accepted. To reveal the complete picture of the effects of behavioural uncertainty, both of these two components have to be taken into account. But, this will pose a tremendous obstacle to those who want to undertake the
actual measurement. Third-party measurement of quality shortfall is intrinsically difficult, for obvious reasons. Moreover, whilst it may be feasible to measure the difference between renegotiated or outturn price and contract price, it is not legitimate to assume a priori that all of this difference is necessarily a rent-transferring transaction cost arising from opportunism.

Regarding $TC_c$, the difficulty is even greater. $TC_c$ has to include the opportunity cost of delay, legal bills and so on. Obviously, the magnitude of these costs varies considerably among transactions. Now the problem facing us is how to infer a representative estimate for the population from the limited samples. In other words, what we want to explore is not whether some particular transaction using GS1 has lower transaction costs than those using GS2, but rather the proportion of GS1 transactions with lower costs in the population. However, due to the stochastic nature of $TC_{II}$, we need to have substantially large samples to eliminate the biases caused by its high variability. Therefore, for most research topics where $TC_{II}$ is dominant, the costs of meeting this requirement will be too substantial to be practical. And this is where the DMA is least likely to be applicable.

3.4 Concluding remarks

To sum up, since the potential measurement difficulties vary as the nature of transaction cost types, the relative desirability of the DMA and IMA is naturally dependent on what are the principal elements of transaction costs with comparative significance, which in turn is affected by what are the alternative GSs under comparison. Accordingly, to give a fair assessment of which one of the IMA and DMA is more suitable to the analysis of construction procurement routes, we have to deepen our understanding about the main sources of transaction costs that may occur in the construction project. This is one of the issue that will be dealt with in Chapter 6.
Chapter 4  The Doctrine Of Transaction Cost Economics

It is clear in Chapter 3 that the IMA is comparatively more suitable to the organisational analysis. This is evidenced by the phenomenal revival of institutional studies since 1970s, often called new institutional economics. The reason of propelling this trend lies in the intention to seek an approach that can avoid the charge of ad hoc specifications of the DMA but still can make theory refutable. Two seminal contributions made by Alchian and Demsetz (1972) (A&D hereafter) and Williamson (1975) broke up a new frontier for economic organisational analysis. They are branched out into two streams: asset specificity branch and measurement branch. This chapter will reorganize the Williamson’s and A&D’s methodology around a common theme of comparative institutional analysis: operationalisation of transaction.

4.1 Williamson's three-stage comparative institutional analysis

Transaction cost economics is almost single-handedly established by Williamson. Tracing the development path of his theory in Williamson (1975,1985,1996), we can find that his concerted efforts are devoted to the construction of a workable methodology for comparative institutional analysis. This approach is clearly enunciated in Williamson (1988) where he sets out a three-stage scheme for operationalising transaction costs. ‘The first stage entails comparative contractual analysis of a microanalytic kind’. ‘The second stage entails a concerted effort to develop refutable implications based on the logic of discriminating alignment’. And the third stage is ‘an examination of intertemporal process differences among modes’.

Developing in parallel with Williamson’ transaction cost theory, Alchian and Demsetz (1972) and Barzel (1982) lay a foundation for viewing organisational issues in terms of measurement problems. Their purpose is also to pursue the operationalisation of theory and make theory testable. In other words, the common question they addressed is to discover the key determinants that can best differentiate the relative efficacy of GSs. In the 1980s, the focus of analysis is mainly directed to the operationalisation of transaction attributes. However, we gradually come to recognize that the symmetrical issue, the dimensionalisation of GSs, has to be worked out so that the alignment of transaction attributes with GSs can be systematically analyzed [Williamson-
son, 1991]. Thus, to put all the elements into an integrated framework, the analysis scheme set out in Williamson (1988) has to be extended. First, the three stages for operationalising transaction costs are condensed into a single stage. Second, the dimensionalisation of GSs is treated as an independent part.

This modified trilogy represents a complete explanatory structure of transaction costs reasoning, as shown in Fig. 4-1. From the first stage, we can uncover the key attributes with respect to which governance structures differ and realise how transaction costs will change in dealing with transactions of different attributes. From the second stage, by way of examining the contractual and informational nature of GSs, we likewise hope to point out the major dimensions as to which the comparative advantage of GSs on transaction costs can be revealed. Combining the analysis results obtained at the first and second stage, the last step aims to align the attributes of transaction with that of GS in a discriminating way. By discriminating what we mean is the extent that GSs can satisfy the objective function. Minimising transaction costs is assumed to be a reasonable approximation to the “actual” goal of organisation selec-
tion. It is worth noting that the meaning of the operator "minimising" here is different from that in the marginal analysis. In contrast to production cost minimisation that requires the ratio of factors of production, say labour and capital, equal to the ratio of their relative prices, i.e., wage and rent, at the margin, transaction costs can be said to have been minimised as long as there is no better GS available.

As mentioned at the start of this section, there are two lines of transaction cost reasoning with different point of emphasis. One line ascribes the main transaction hazards to the presence of asset specificity; another line upholds measurement difficulties as the explanatory variable. Which one is weighty is largely conditional on what are the alternative governance structures under comparison. We will discuss more about when the balance will tip in Sec.4.4.

4.2 The asset specificity approach
Williamson labels his approach as governance branch of transaction cost theory. The governance structure, by definition, is "the institutional matrix in which the integrity of a transaction is decided" [Williamson, 1996]. The way a transaction is governed will decidedly affect how costly it will be. However, to get around the obstacle to direct measurement of transaction costs, a detour can be taken by producing the predictions on the basis of transaction attributes instead of transaction cost itself. This section will expand on how Williamson achieved it.

4.2.1 The unit of analysis: 'transaction'
Coase (1937) takes market transactions as the focus of analysis. Coincidentally, Commons (1932) put forth a similar perspective that a transaction is an informative unit of analysis, despite out of different consideration. Coase's research strategy can be said to come from the positive motive in the sense that the exploration of market transactions is carried out for rationalising a fact: the existence of the firm. In contrast, Commons employs a somewhat normative way to "seek" an appropriate answer. He proposes three key components needed to be considered: conflict, mutuality and order. Rephrased in modern terms, they correspond to sub-goal pursuit, interdependence and continuation. Using transaction as the unit of analysis, he reckons, can capture all these aspects. As a matter of fact, this view should be thought of as a counteraction to the neglect of the relations between man and man in the classical and neoclassical
economics, where the only origin of uncertainty faced by us is nature [Medema, 1992]. Founded on this authority, Williamson further attempts to integrate the man-man and man-nature relation into a unified framework.

There is no denying that, to get over some analytic difficulties, Williamson trims off a few elements originally pinpointed by Commons, such as ongoing power relationships, trust, habit, custom and so forth. Instead, he defines a transaction as the transferring of a good or service across a technologically separable interface [Williamson, 1985]. This definition evidently has a bearing on his proclaimed focal problem – vertical integration. The choice of the unit of analysis can be different varying as the nature of the problem, whereas ‘the critical dimensions with respect to which that unit of analysis differs need to be identified [:] otherwise the unit will remain non-operational’ [Williamson, 1993] and non-refutable. This is the essence of his demand that the symmetry of examination should be done in assessing the rival theories. Only if refutable predictions can be derived from the theory, it is possible to tell the chaffs from the wheat. This criterion for theory appraisal will be discussed in Chapter 5.

4.2.2 A four-step scheme for operationalising transaction costs

If we examine Williamson’s logic in scrutiny, the process of identifying transaction costs actually includes four steps (included in the first stage in Fig. 4-1):

(1) Behavioural assumptions:
In the tradition of methodological individualism, human behaviour must be interpreted in terms of individual behaviour, which is guided by some principle, often in the form of behavioural assumptions. The purpose of these assumptions is to specify the goal of individual decisions. For example, in pursuit of utility maximisation is the postulated behaviour criterion for the rational consumers.

(2) Interaction structure:
Economists get accustomed to examine the interaction relationships between people through the lens of incentive structure. Game theory is a standard case, according to which payoff structure in combination with the assumption of optimisation constitutes the decision situation and decision makers will choose a course of actions so as to maximise the expected gains. Unlike Dahlman (1979), Williamson conceives of
transaction costs as a function of the strategy that trading parties may take, and explores its implications.

(3) Itemisation:
Before going into operationalisation, Williamson also gets through the task of itemising the crucial components of transaction costs that are likely to happen in a series of related transactions. In particular, he highlights the importance of process analysis, which marks another feature of his analysis.

(4) Operationalisation:
Distinct from that in North (1986), the purpose of operationalisation in Williamson’s framework is to ‘relate the relative merits of alternative institutional arrangement to observable attributes of transactions’ so as to ‘[generate] an array of refutable implications regarding the institutions and organisational forms used to govern production and exchange’ [Masten, 1995].

How this four steps work in transaction cost economics (TCE) will be spelt out in the following four subsections.

4.2.3 Behavioural assumptions interaction structure in TCE
The concept of bounded rationality, despite its intuitive appeal, has been dismissed as unworkable for a long period of time since it was first espoused in Simon (1957). One may wonder why this notion is elevated to such a prominent status in TCE. A direct answer is that this assumption can endow “organisation” with a role to play in economic theory – as an apparatus for coping with real uncertainty. This assertion needs a bit more expositions.

According to Simon, an economic man is assumed to be “intentionally rational, but only limitedly so” [Simon, 1961]. This intention to be rational reflects an implicit premise that the pursuit of economising is the ultimate goal, but thanks to limited capability and information, it must be admitted that the conditions of unbounded rationality sometimes are unachievable. At this point, we are facing an intriguing problem. Because bounded rationality is defined by way of another term, perfect rationality, it is natural for us to inquire about what perfect rationality is. In the neoclassical eco-
omics, rationality is not defined by itself, but by means of a process for achieving it. Optimising an objection function is the standard setting-up for modelling “rationality”. Taken to the extreme, it leads to Arrow-Debreu general equilibrium theory in which comprehensive contracting is possible by assuming that contracts for all goods and service across all future contingencies have existed at the point of making a decision. Living in this imaginary world, all agents won’t expect any surprise because every contingency has been stipulated in the contract. It follows that transactions through markets are always efficient, making economic organisations redundant. Accordingly, Williamson has to relax the assumption of rationality to address the problem of incomplete contracting, while at the same time, he still wants to stick to some powerful analysis concepts, such as economising and equilibrium. This ‘eclectic’ attitude has been doubted because, now that rationality is bounded, how is it possible to achieve an optimal outcome. In the face of this challenge, Williamson shifts the focus of this controversy to higher level of philosophical debate. We will come back to this issue in Sec.6.2.3.

4.2.4 Interaction structure in TCE

In comparison to bounded rationality that is pervasively accepted by heterodox economists, opportunism is a peculiarity of TCE. Williamson extends the classic behavioural assumption of self-interest seeking to encompass strategic exercises of advantages in information. An opportunistic trader is assumed to have the proclivity to disclose the incomplete or distorted information so as to “mislead, distort, disguise, obfuscate, or otherwise confuse” his/her trading partners [Williamson, 1985]. Whereas opportunism is treated as one of two pillars of TCE’s behavioural assumptions, its function is to signify the interaction structure of trading parties. Let’s make a brief comparison. The assumption of bound rationality states that the rationality of the decision-makers must be bounded, while opportunism just indicates that opportunistic behaviour is apt to happen, particularly under certain circumstances. Have a closer look at the standard definition of opportunism – self-interest seeking with guile. Same as bounded rationality, the first part is an affirmative assumption that an economic agent must be going to act on the behalf of his/her own interest. However, the component of guile conveys a distinctly probabilistic concept. That is, the vulnerability of a party may be taken advantage and ex ante promises are not always binding.
Plainly, opportunism takes into considerations the various possibilities of the strategies that may be adopted. The conjuncture of incomplete contracting and opportunism characterises the property of open-endedness in TCE, which is the reason why it is grouped into the open-system organisation theory [Scott, 1987] and why there is an ontological difference between TCE and Alchian & Demsetz (1972) [Foss, 1994]. We'll go back to this issue in Sec.4.4.1.

4.2.5 Itemisation of transaction costs: a process perspective

Taking as the cutting line whether the contract is signed, Williamson (1985) distinguishes two kinds of transaction costs: ex ante and ex post transaction costs. The former refers to ‘the costs of drafting, negotiating, and safeguarding an agreement’. The latter includes (1) maladaptation costs, (2) haggling costs, (3) costs of dispute resolution and (4) bonding costs. In contrast to Dahlman (1979), Williamson downplays the significance of search costs, maintains the costs of negotiation and transforming costs of enforcement into that of adaptation in this framework. The picture he intends to depict is a livelier interaction process during the transaction: if the governance structure isn’t suitable to deal with uncertain future, maladaptation will occur, renegotiations for better terms may be necessary, appeal to third-party arbitration will ensue provided disputes cannot be settled, and extra bonds to prevent it from happening again will be required, all of which incur costs. For the single transaction analysis, the story will end here. But, for Williamson, it is the beginning of a series of events. He insists “farsighted contracting” and considers the transactions “in its entirety” [Williamson, 1996]. This claim entails slow digestion.

First, Williamson is aware that delayed effects of the past experience in carrying out a type of transaction will be passed on to the next decision [Williamson, 1993]. To realise this mechanism, it's right time to put the arguments in these three steps together. Under the assumption of bounded rationality, the completeness of contract is ruled out, so certain trading party is likely to fall victim to vulnerable positions caused by uncertainty. If this does happen, both parties' suffering from ex post transaction costs can be expected. And this memory will reinforce the victim’s incentive to take precautionary actions next time, increasing ex ante costs. This is a vicious cycle. To the contrary, if the transaction process always goes smoothly, the preventive measures will be seldom prepared, reducing ex ante costs. This is a virtuous cycle. It is to be noted that
the possibility of opportunistic behaviour is a function of vulnerability, which in turn is a function of uncertainty. Uncertainty represents the fact that we cannot foretell what will turn up and when. Thus, one time of safe experience doesn’t mean that an efficient governance structure is being employed, that may be just lucky. For avoiding this bias, we ought to include a series of related sequential transactions into the horizon of decision-making. Of course, this is a normative suggestion. Its rationale is founded on positive evidence. The transformation from a positive theory to normative guidelines lacks a link: weak-form selection. Simply speaking, through weak-from selection, only can the organisations without remediable gains survive. If the hypotheses derived from this proposition line up with the reality, it is corroborated that the governance structures selected by far-sighted contracting are efficient. If you want to make a right choice as well, far-sighted contracting is what you ought to follow. Failure to make this point explicit would obscure the underlying interpretation that TCE opts for: intentional or functional explanation [Knudsen, 1993].

4.2.6 Operationalisation

It is the task of operationalisation to point out ‘the factors responsible for transaction cost differences’ [Williamson, 1985]. Three factors are demonstrated to be relevant: asset-specificity, uncertainty and frequency. However, they are not with equal footing. The last two factors just serve as the subordinate roles. Only when the condition of asset-specificity holds, neither of them matters. The following subsections will examine these factors in turn.

4.2.6.1 Asset specificity

Asset specificity arises when an investment is idiosyncratic to a transaction, making its value greatly reduced in the second best use. Once the investment becomes irreversible, the dedicated resources will be locked in to this relation. The reduction in ex post opportunity costs of the investment leads to the presence of quasi rent. This concept can be illustrated with the aid of a paradigmatic case- vertical integration. With reference to Fig.3-1, the condition that can induce a producer to engage in production activity E is

\[
\text{Seller's rent} = pQ - TVC - I > 0
\]

Eq. 4-1
where $p$ is the price per unit, $Q$ quantities produced, TVC total variable costs, $I$ the ex ante opportunity cost of the investment. After the investment is sunk (e.g. the plant is built and machinery is installed), the value of the original investment can be different, varying as the degree of specificity. The condition that makes the seller stick to the relation can be expressed as

\[
\text{Seller's quasi rent} = pQ - TVC - S > 0 \quad \text{Eq. 4-2}
\]

where $S$ is the ex post opportunity cost of the investment. Consequently, a quasi rent represents the excess above the return necessary to maintain this production flow (TVC+S) [Alchian and Woodward, 1987]. Assuming that at the onset, large number of bidders measure up to perfect competition, i.e., economic rent being squeezed to zero, we can get $p = \frac{TVC + S}{Q}$. Substituting $p$ into Eq.4-2, we can find that the quasi rent results from the difference between ex ante and ex post opportunity cost of the investment, $I - S$. It can be interpreted as \textit{costs of illiquidity} in the sense that the physical capital is less liquid than money.

Since the producer is not easy to switch elsewhere without costs, the seller’s quasi rent appears approbiable by the buyer. But this can be realised only when unanticipated events are present or more exactly when the situations are not stipulated in the contract and/or cannot be arbitrated by the third party. Simply speaking, this concept tries to describe a scenario: when the opportunity of snapping up extra gains comes out and the possibility of success of this attempt is high, this situation turns out troublesome. How troublesome it is depends on how large the targeted pie is, which is proportional to the degree of specificity given a constant investment scale.

After introducing the basic idea of asset specificity, we needs to link up this factor with the costs itemised in Step 3. As is explained in the last paragraph, when quasi rent appears easy to catch, the incentive for renegotiating a better term intensifies. This tug-of-war is likely to go through bargaining, arbitration, litigation, all of which are just like ex post transaction problems enumerated in Sec4.2.5. This experience certainly will give a negative feedback to the next decision. Once the vicious cycle is triggered, the trouble of transacting through markets is going to be getting worse, increasing the incentive to shift from one governance structure to another. Since asset
specificity can suitably reflect the sources of transaction difficulties, it is reasonable to view transaction costs as a function of this factor.

4.2.6.2 Uncertainty

In the most economic analysis, the future is often treated as the probability distribution of risk or uncertainty. However, the real uncertainty must involve the possibility of surprise. Not confined to “the God”, the sources of surprise may come from the opportunistic behaviour, termed behavioural uncertainty. No matter how uncertain is the environment, as long as the market is teeming with alternative trading partners that can be switched to without much cost, real uncertainty will do no harm to the transaction relation. The presence of asset specificity is a necessary condition that can make uncertainty have effects on transaction costs. This inference is well corroborated in most empirical studies [Shelanski and Klein, 1995].

4.2.6.3 Frequency

Why frequency is relevant? Williamson (1985) sheds a light on this problem: ‘[t]he cost of specialised governance structures will be easier to recover for large transaction of a recurring kind; [H]ence the frequency of transactions is a relevant dimension’. The cost of specialised governance referred here plainly includes more than transaction costs (listed in Sec. 4.2.5). The pendulum of focus has swung to production costs, in particular set-up costs of the investment. Ceteris paribus, the more frequent is the commodity or service needed to the whole operation, the more likely for the buyer to be willing to bear a large number of installation costs (plant and machinery), and the more likely to take an integration strategy. This argument is based on the same level of asset specificity. While the transaction is not idiosyncratic, even the most frequent demand can be met by means of price mechanism without much trouble. Consequently, similar to uncertainty, the importance of frequency is not on its own but on its interaction with asset specificity.

4.3 The measurement approach

4.3.1 Alchian and Demsetz’s theory of team

The theory of team espoused by Alchian and Demsetz (1972) starts off with the premise that technology can be non-separable, so organising the production activities in the form of a team is necessary. Perhaps the output of a team may be easily measured, but
it is not the case for the output of individual workers. Because of the measurement difficulties, the wage is often not directly linked to marginal productivity, leading to free-rider problems, which means that someone may choose to shirk at the expense of total output. An agent in charge of monitoring can curb this kind of intention. To avoid shirking, the reward scheme of the monitor should be correlated to the residual product above the prescribed level. Besides, a metering system is to be effective if the monitor has the right to ‘revise the contract terms and incentives of individual members without having to terminate or alter every other input’s contract’ [Alchian and Demsetz (1972)]. Put all together, the ownership of the capitalist firm is characterised as ‘a bundle of rights (1) to be a residual claimant; (2) to observe input behaviour; (3) to be the central party common to all contracts will inputs; (4) to alter the membership of the team; and (5) to sell these rights’.

Not only does the significance of this work lie in its intuitively appealing, but also its manifestation of investigating the condition that can discriminate the relative efficacy of the market and the firm. The finding that the nature of the output of each production stage (whether easily being measured) determines which type of organisation is more efficient brings about the further development along this line of inquiry, which is called measurement branch of transaction cost theory. This is the topic to which the next subsection will turn.

4.3.2 Measurement costs and organisations

According to Lancaster’s consumer theory, the good is a mixture of different characteristics, which are the ultimate sources of utility [Lancaster, 1966]. The force of competition determines the market value of the commodity with given attributes. In exchange, those attributes with ambiguous properties will be in the public domain, from which potential rent comes. For the buyer, inspecting quality incurs costs and for the seller, reducing quality risks losses if detected. The rent-seeking behaviour at least will waste a portion of the potential rent, whence the arrangements, such as share contracts, warranties, trust, brand name or repeat purchases, exist to lower these losses [Brazel, 1982]. Among others, vertical integration has the benefits of saving measurement costs as well. Referring to Fig.3-1, if the quality of product D is very difficult to be determined at the point of transaction, its buyer may need to monitor the whole production so as to know the value of inputs as well as added value from activity D.
itself. Thus, ‘when inputs have to be measured at two successive junctures, a rationale for an integrated firm emerges’. The trade-off behind this inference is between the costs of measurement required in the exchange as well as the costs of inspecting the quality of input and production process. It is to be noted that market price and production cost are not mentioned in this framework.

Reliance on measurement costs as the only explanatory variable for vertical integration has aroused head-on confrontation with asset-specificity explanation. What weight should be given to each has been a topic of keen debate. Why their focuses are so different? What are the conditions for determining the weight that should be placed on each factor? These questions will be explored in Sec. 4.4.

4.4 Comparison of asset-specificity and measurement approach

4.4.1 Ontological differences
Before proceeding to the comparison of asset-specificity and measurement approach, it can be helpful to give a brief review of their tenets. TCE claims that incomplete contracting, arising out of bounded rationality and uncertainty, can give rise to vulnerable positions in the presence of asset specificity, which is subject to the possibility of being exploited by trading partners. The motivation of reaping additional rents is the site of transaction problems (with reference to Fig. 4-2).

Fig. 4-2 Logic of asset-specificity branch of transaction cost theory

Bounded rationality

Uncertainty

Incomplete contracting

Asset specificity

Transaction costs

Opportunism

The measurement approach holds the view that the ambiguity of quality requirement affects the rent distribution between transactors, so the seller is likely to conceal or
manipulate the asymmetrical (true) information about quality for the realisation of additional rent from lower quality. The fear of quality certainty may evoke the buyer to reinforce the intensity of inspecting the *assembled products* in the exchange until inspecting internally the *production process* becomes less costly (with reference to Fig. 4-3).

After cosmetic comparisons, some ontological difference is surfacing in sight. TCE is dealing with the triangle relations of a decision-maker, his/her trading partner and real uncertainty, while in the measurement approach, real uncertainty is reduced to the nature of the commodity. It follows that the transaction problems come from what is being transacted only, regardless of how the external environment is changing. It is this close-endedness that distinguishes two approached in terms of ontology [Foss, 1994].

**Fig. 4-3 Logic of measurement branch of transaction cost theory**

Ambiguous attributes of goods or service

Uncertain intrinsic value of goods or service

Opportunism ➔ Measurement costs

4.4.2 Why measurement costs are downplayed in TCE

4.4.2.1 The role of measurement costs in TCE

In Williamson (1985), measurement cost is not totally ignored, but treated case by case. His standard assertion is as follows:

*Although the measurement branch plays a less substantial role in the analysis of problems dealt with in this book (partly because the problems selected for study have a bilateral trading quality), it is nevertheless important.* [p.112]
But under what circumstances will measurement costs seem non-negligible relative to the hold-up problem? Williamson doesn't address this question, not to mention answer it. Nonetheless, he does seriously contemplate the importance of measurement costs in certain contexts. The first appears in the interpretation of manufacturing firms extending their operations into distribution [p.112]. Owing to the spillover effect of the individual distributor's efforts in promotion, the manufacturer cannot meter their performance by their volume of sales only. Difficult measurement problems obviate the feasibility of an output-based reward scheme, so sub-optimal level of effort prevails. A possible remedy for this is vertical integration.

The second case is related to the organisation of work [p.212]. In talking of internal organisation, Williamson is concerned with principal-agency relationship on the similar ground to Alchian and Demsetz (1972). To check the employee's opportunistic behaviour, two alternatives can be effective: (1) reducing the incentive by way of different contractual arrangements, e.g. changing from output-based contract to wage-based one to curb the proclivity to increase productivity at the expense of quality; (2) decreasing the measurement costs by means of product redesign and task reorganisation.

Whereas the use of measurement costs disperses across several topics, there is something in common – difficulties in measuring the performance of the agent. This difficulty is often considered as a function of the quality characteristics of the goods or service. However, in construction it is the joint function of project characteristics and procurement routes. We will return to this issue in Chapter 6.

4.4.2.2 A preliminary analysis

In discussing the paradigmatic problem of make-or-buy decision, Williamson justifies the omission of measurement costs by arguing that 'the problems selected for study have a bilateral trading quality' [Williamson, 1985,p.112]. Regarding search costs, no apology is expressed, as if they didn't exist [Alchian and Woodward, 1988]. To delve into the reason, we need to focus on how measurement costs and search costs are changed after integration.
For the following discussion, the principle of cost transferability requires discussion first. Under the circumstance of large-number bidding, the producer's rent will be squeezed to zero \textit{ex ante}, that is

\[ \text{Seller's rent} = pQ - TVC - TTC - I \quad \text{Eq. 4-3} \]

where \( p \) is the price per unit, \( Q \) quantity produced, \( TVC \) total variable production costs, \( TTC \) total transaction costs, \( I \) the ex ante opportunity cost of the investment. As a consequence,

\[ p = \frac{TVC + TTC + I}{Q} \quad \text{Eq. 4-4} \]

This equation indicates that the selling price contains average (seller's) transaction costs, implying that the costs of co-ordinating internal production will be transferred to buyers through the market in the form of increased price. Though the transferred proportion of transaction cost varies as the condition of competition, it does not make sense to preclude the possibility that this cost will not be absorbed by its direct bearer. One reason for this is that competition is always changing; another is that ignoring the possibility of cost transfer is liable to camouflage the potential hazards of contracting. To make a sound appraisal of organisation forms, the redistribution of transaction costs through contracting cannot be ignored.

\[ \text{Fig. 4-4 Buy or make decision} \]

\[ \text{Mode I: buy decision} \quad A \rightarrow B \rightarrow p \rightarrow \text{Core production activity} \quad C \]

\[ \text{Mode II: make decision} \quad A \rightarrow B \rightarrow \text{Intermediate goods} \quad b \rightarrow \text{Activity C produced by producer C} \rightarrow \text{Output} \quad c \]
We can examine the significance of search costs and measurement costs by considering the paradigmatic case - vertical integration. In Fig. 4-4, when B and C are independent firms (mode I), firm C purchases product $b$ from the market with unit price $p$ and incurs average transaction costs $ATC_m$. $ATC_m$ covers search costs for finding producer B and measurement costs for checking the quality of goods $b$. To constrain employees' shirking and assure quality of production, firm B needs to bear costs for setting up an internal inspection system. Meanwhile, searching for producer A and buyer C as well as measuring the quality of product $a$ are also transaction costs that need to be considered.

| Table 4-1 A comparison of search costs and measurement cost influenced by integration |
|---|---|
| **Mode I** | **Firm C** | **Firm B** |
| | $TC^C_{i(1)}$: search costs for finding and bargaining costs with producer B | $TC^B_{i(1)}$: search costs for finding and bargaining costs with producer A |
| | $TC^C_{i(2)}$: measurement cost for assuring the quality of product $b$ | $TC^B_{i(2)}$: measurement cost for assuring quality of product $a$ |
| | $TC^C_{i(3)}$: search costs for finding and bargaining costs with buyer C | $TC^B_{i(3)}$: measurement costs of internally controlling product $b$ |
| | $TC^C_{i(4)}$: measurement costs of internally controlling the quality of product $b$ | $TC^B_{i(4)}$: search costs for finding and bargaining costs with buyer C |

After production activity B is internalised (mode II), firm C is facing internal costs of quality control and search costs for finding supplier A and measurement costs for checking product $a$. All types of these costs are listed in Table 4-1. Now we need to address a question: whether measurement costs and search costs matter (i.e. differ) under different governance structures, market or hierarchy? First, thanks to integration, apparently the total (search and measurement) costs that firm C needs to bear change from $TC^C_{i(1)} + TC^C_{i(2)}$ to $TC^C_{i(1)} + TC^C_{i(2)} + TC^C_{i(3)} + TC^C_{i(4)}$. But we should not omit the fact that in mode-I transaction costs originally incurred by firm B will be transferred to firm C by way of selling price of product $b$. That is, from the perspective of firm C, we are
interested in the question of whether $TC^C_{i(1)} + TC^C_{i(2)} + TC^B_{i(1)} + TC^B_{i(2)} + TC^B_{i(4)} + TC^B_{i(5)}$ will be much different from $TC^C_{i(1)} + TC^C_{i(2)} + TC^C_{i(4)}$.

Next we introduce the assumption that firm C and B face similar costs whichever of them undertakes the activities of dealing with A and internally controlling the quality of product $b$. This makes $TC^B_{i(1)} = TC^C_{i(1)}$, $TC^B_{i(2)} = TC^C_{i(2)}$ and $TC^B_{i(3)} = TC^C_{i(3)}$. When the aforementioned three items are cancelled out, the three items left are the costs of using the market for product $b$. The benefit of using hierarchy for saving these costs was first systematically introduced into economic analysis in Coase (1937). This becomes one of the most fundamental reasons why the firm exists. Thus, an interesting question is why these costs don't merit equal emphasis in Williamson's TCE. A possible justification is to take the assumption of $TC_I \ll TC_{II}$, meaning that, compared with opportunism-derived costs, resources used for collecting and verifying information are relatively trivial so that they can be ignored as items of second-order significance. Whether this assumption is sensible depends on the actual relative magnitude of these types of costs. Obviously, this is an empirical problem. However, as explained before, in principle, we can make a judgement according to the severity of opportunism, which is affected by how much quasi-rent will be present [Klein et al., 1978]. The larger is the lump-sum specific investment in the first place, the more substantial is $TC_{II}$ and the less is the error resulting from the omission of $TC_I$. In terms of empirical evidence as documented in Shelanski and Klein (1995), the predictive power of TCE seems not to be blunted by exclusion of $TC_I$. It seems safe to say that, for most vertical integration cases, this assumption can be maintained. However, whether it still hold in the case of construction. This is the issue we want to explore in Sec.6.5.1.1.

4.5 Dimensionalisation of governance structures

There are two key dimensions singled out by Williamson as relevant to distinguish market and hierarchy: adaptation and instruments [Williamson, 1991]. In the economic literature, adaptation connotes two kinds of meanings:

(1) Hayek (1945) perceptively pinpoints that “the marvel of the market” comes from the participants’ ability to take the timely action to respond to changing en-
vironment. Since economic information disperses everywhere, the most efficient way of running an economy is to allow everyone to make autonomous decisions according to perceived price signal.

(2) In contrast, Barnard (1938) emphasises "the marvel of internal organisation", which is the manifestation of "kind of cooperation among men that is conscious, deliberate, purposeful." Through cooperation, it is less costly to arrive at coordination.

The adaptability of governance structures is chiefly conditional on the instrument that can be employed within them. Instruments to which Williamson refers include (1) differential incentive intensity and (2) differential administrative controls. Incentive intensity represents the degree of the actor's incentive to reduce costs and adapt efficiently. Market is of high-powered incentives in that one's own efforts are closely linked to rewards or losses. In contrast, 'internal organisation degrades incentive intensity and added bureaucratic costs result' [Williamson, 1985]

Administrative controls mean resources are allocated by fiat rather than by competition. Hierarchy can impose tighter controls on the actions of employees by monitoring or career rewards or penalties. Therefore, it is more suitable for sequential decision making in that repeated bargaining between members within the firm can be reduced to a minimum by administrative orders. By dint of exercising the instrument in the right situation, the transaction can be governed in a more efficient way. Then why does the feature of these instruments decisively affect the magnitude of transaction costs? As regards this question, a more detailed account is needed.

A governance structure can be said to be more adaptable if it is less costly to accommodate new information into the new economic decision. In the perfect competition market, equilibrium is restored instantaneously in the presence of disturbances, so economic actors can modify the decision in response to the new prices. However, it is not the case for the transaction of bilateral dependency. When the specific investment is sunk, the producer will have little trading partners to choose. Once unanticipated events intrude, the gap of incomplete contract will be widening, which is likely to

1 In the sense of Taylor's expansion, the terms of higher orders can be ignored.
trigger a bout of renegotiations on prices and quality. In the process of attempting to
effect realignment, determining new terms through bargaining seems unavoidable if
both parties are autonomous entities. ‘Such bargaining is itself costly. The main costs ..
are that transaction are maladapted to the environment during the bargaining inter­
val. Also the prospect of ex post bargaining invites ex ante prepositioning of an ineffi­
cient kind.’ [Williamson,1991] In this case, hierarchy is a better alternative.

Hierarchy has advantages over market in coping with adaptability problems in the
situation of high specificity. For instance, when the headquarter of a company per­
ceives some modifications for the product is needed for meeting the market demands,
this requirement can be more easily implemented if the product is manufactured inter­
nally as a result that ‘(1) proposals to adapt requires less documentation, (2) resolv­
ing internal disputes by fiat rather than arbitration savers sources and facilitate
timely adaptation, (3) information that is deeply impacted can more easily be ac­
cessed and more accurately assessed, (4) internal dispute resolution enjoys the sup­
port of informal organisation, and (5) internal organisation has access to additional
incentive instruments – including especially career reward and joint profit sharing –
that promote a team orientation’ [Williamson,1991]. Otherwise, this company is
highly possible to be overcharged because of limited supply channels.

In summary, transacting through market can fully reap the benefits of specialisation
and with the pressure of competition, the incentive to lower down production costs
will prevail in the market. However, as interdependency is deepening, its awkward
ability to make sequential adjustments through market will offset its advantage of
high-powered incentives. Accordingly, the problems becomes what condition will
make market or hierarchy dominate. This is the task of alignment in the following
section.

4.6 Alignment of governance structures with transaction
dimensions

A graphic illustration is an informative way to grasp the alignment of governance
structures with transaction features. With reference to Fig. 4-4, let’s consider a deci­
sion scenario in which firm C decides to buy or to make intermediate goods. Hierar­
chy will be advantageous if
TC_i + PC_i ≤ PQ + TC_m \quad \text{Eq. 4-5}

where \( TC_i \) is cost of organising activity B internally, \( TC_m \) cost of using market, \( PC_i \) cost of production internally, \( P \) prevailing price in the market. Eq.4-5 can be expressed as difference equation of asset specificity \( k \)

\[ \Delta G(k) + \Delta C(k) ≤ 0 \quad \text{Eq. 4-6} \]

By definition, \( \Delta G \) is the difference between costs of bureaucracy and costs of using market, and \( \Delta C \) is the difference between the cost of producing internally and the price of procuring from the market. As shown in Fig. 4-5, \( \Delta G \) is a decreasing function of \( k \left( \frac{\partial G}{\partial k} < 0 \right) \) with positive intercept with y-axis, \( \beta_0 \). This setting indicates that when \( k < \tilde{k} \) (the investment is relatively redeployable), transaction costs through the market mechanism is less costly than costs of internal management, while once \( k > \tilde{k} \), the governance costs of market will exceed that of internal organisation in that coordination through authority can more effectively respond to changing environment and more efficiently allocate quasi-rent as mutual dependency deepens.

Likewise, \( \Delta C \) is a decreasing function of \( k \left( \frac{\partial C}{\partial k} < 0 \right) \) as well, meaning the adverse effects of internal production on scale of economy and/or scope of economy will appear trivialised as asset specificity is getting large and the market is getting thin. Add up the curves of \( \Delta C \) and \( \Delta G \) vertically to get \( \Delta C + \Delta G \) (see Fig. 4-5) with intersection point with horizontal axis at \( k^* \). \( k^* \) is the critical value that can distinguish the relative efficacy of market and hierarchy. By solving Eq.4-7, we can find the algebraic solution of \( k^* \):

\[ \Delta C(k^*) + \Delta G(k^*) = 0 \quad \text{Eq. 4-7} \]

Only when \( k > k^* \), hierarchy has advantage over market; on the contrary, when \( k < k^* \), market becomes dominant.

From the above analysis, a complete transaction cost reasoning must consider both transaction costs and production costs. Sometimes keeping production costs constant can be viewed as a sensible approximation, but this assumption needs to be justified. Besides, theoretically, we can tell comparative advantage of hierarchy from market
according to the degree of asset specificity. But we cannot say that hierarchy is advantageous as long as asset specificity is present. This is a comparative undertaking!

Fig. 4-5 Comparative Production and Governance Costs
[Williamson, 1985]

\[ \Delta C = \text{production cost of producing internally} - \text{production costs of the suppliers in the market} \]

\[ \Delta G = \text{transaction costs of internal production} - \text{transaction costs of using the market} \]
Chapter 4 outlines a structured account of the rationale of TCE. To ensure the successful application of this approach, it is fundamental to compare the nature of problems at issue to those of the paradigmatic cases out of which TCE has developed. In principle, the more clearly we can identify the underlying assumptions upon which TCE rests, the more closely the application can be fitted to the nature of construction procurement systems, and the more rigorous can be the reasoning. Thus, it is worth devoting a chapter to exploring this issue.

As will be explained in Sec.5.1-5.4, there are five premises in TCE relevant to this study: (1) invariability of revenue across GSs; (2A) irreversibility of lump-sum investment and (2B) continuous production; (3) exogenous design; (4) invariability of production costs across GSs. These assumptions have not always been made explicit. In this chapter, we would like to take a step further to canvass the plausibility of the hidden assumptions in the standard case of manufacturing vertical integration, and then investigate these issues in the context of construction procurement in Sec.5.5.

5.1 Assumption 1: Invariability of revenue across GSs
The first item ignored in the standard TCE objective function is the revenue from production. Referring to Fig. 5-1, now we have two alternative ways of organizing production activity $D$. One mode is to purchase intermediate goods $d$ from the producer $D$ in the market; another mode is to produce goods $d$ internally, perhaps in one of the divisions of company $E$. Suppose company $E$ sells goods $e$ directly to consumers. Then the revenue obtains from the sale of goods $d$ is determined by the competition condition of market for goods $e$. This means that the revenue from goods $d$ to company $E$ is independent of the way that goods $e$ are produced, market or hierarchy.

For the more complicated case where activity $E$ is to assemble a number of inputs, including goods $d$ to produce goods $e$, the same conclusion can be made by assuming that, in analysing the mode I and II,

1. the composition of quantities of inputs required per unit output of goods $e$ is kept constant;
(2) the price of other factors of production and intermediate goods are exogenously
determined by their individual markets;
(3) the competition condition of market for goods \( e \) is exogenous.

**Fig. 5-1 The Production Chain of the final product E**

Mode I: buy decision

\[ \text{B} \rightarrow \text{C} \rightarrow \text{D} \rightarrow \text{E} \rightarrow \text{Market for goods } e \]

Mode II: make decision

\[ \text{B} \rightarrow \text{C} \rightarrow \text{D} \rightarrow \text{E} \rightarrow \text{Market for goods } e \]

Under these premises, the assumption of revenue invariability can hold in dealing
with the case of vertical integration. But, the reasons for justifying this assumption are
not applicable to the analysis of construction procurement systems, so some modifica­tions has to be made. We will turn to this issue in Sec.5.5.

**5.2 Assumption 2: irreversibility of investment and continuous production**

The transaction attribute of foremost importance in TCE is asset specificity, indicative
of the extent that the value losses of the lump-sum investment may be reduced if it is
switched to alternative use. The losses, in fact, results from two reasons:

(A) Irreversibility of investment:
By irreversibility what we mean is that once a lump-sum investment is made, the de­cision cannot be reversed to the original point without losing a large part of value of
that investment. This condition may be due to the immovability\(^1\) of a physical investment (e.g. factory construction and machine installation) or time spent (e.g. training) or other reasons. In the context where the decision maker cannot reverse without a cost, what may happen due to this decision in the transaction process needs to be considered in the first place. This is one of the main reasons why TCE argues for using the process perspective to scrutinise the mechanism of GSs in detail\(^2\).

(B) Continuous production:
According to the discounted cash flow (DCF) valuation model, the value of the investment is equivalent to the discounted value of its future revenue stream. Before the investment is made, money is money without any illiquidity problems. But after the investment is sunk, money is transformed into production capital, the value of which will be varying as the profits that this investment can bring in.

Assume that a producer decides to make an investment in physical capital goods with the economic life of \(n\) years. The cash flow of this investment can be visualised in Fig. 5-2. For the purpose of illustration, the expected profit is assumed to be driven down to zero by keen competition, i.e.

\[
0 = \sum_{t=1}^{n} [p_t Q_t - TVC_t - I_t] \delta_t, \tag{5-1}
\]

where \(p_t\) is the selling price per unit at time \(t\), \(Q_t\) is the total quantity sold at that time, \(TVC_t\) is total variable costs, \(I_t\) is the ex ante projected opportunity cost of the capital goods at time \(t\) and \(\delta_t\) is the discounting factor. It is to be stressed that it is a series of cost estimates in different periods on the basis of ex ante costs of capital without considering the lock-in effect.

After the investment is made, the residual value of the investment at each period may not be as expected. The value of this investment at this time is

\(^1\) Immovability of the machinery/equipment is defined by the cost of moving them. In fact, factory or installed machine can be movable if they are dismantled. But it is costly, so it can be said to be immovable.

\(^2\) Further discussion please see Sec.6.2.
\[ v = \sum_{t=1}^{n} [p_t Q_t - TVC_t - I_t] \delta_t, \]  

Eq. 5-2

where \( I_t \) is the ex post realised opportunity cost of the capital goods at time \( t \). If \( I < S \), it means that \textit{ex ante} the alternative use of the capital goods is not as valuable as it was \textit{ex ante}. Apparently this can make the profit swell, whereas this swollen profit is not a real profit, but an appropriable rent in that the cost of a lump-sum investment is completely sunk at \( t=0 \). The vulnerable party actually may be willing to accept a demand on the redistribution of quasi rent \textit{ex post} in return for getting back the money stuck in this investment by maintaining the positive revenue stream \((p_t Q_t - TVC_t > 0, t = 1, \ldots, n)\). The calculation of gaining more in the future by sacrificing now is the ultimate reason making asset specificity relevant. That is, if production flow generated by the capital goods is not continuous, \( p_t Q_t - TVC_t \) will not necessarily provide enough incentive to induce the vulnerable party to accept the overcharge of \( \Sigma[I_t - S_t] \delta_t \). As a result, whether the production flow can last as long as the economic life of the capital goods will fundamentally influence the severity of hold-up problem.

**Fig. 5-2 The benefit and cost flow of a capital good investment**

This clarification appears superfluous in most applications of TCE. Continuous production is a sensible condition for the manufacturing firms because the production line often operates continuously to make products. The manufacturing production process is often modelled as a continuous function of mapping inputs into outputs. This continuity of mapping is only an approximation only when the same capital goods can produce the same products with different intensity (quantity per unit time). And only when the recovery of the investment costs depends on maintaining the
transaction relation (producing the same products), the appropriability of quasi rent poses the threat to the vulnerable party. Though the quasi rent, a portion of gains from trade that can be redistributed, is an inducement to opportunism, the ultimate reason of one party being held up is his/her being afraid of losing the value of the investment due to the disruption of the production flow.

But, this is not the case in construction. Specific location of construction site of an individual project and unique design make it impossible to set up a fixed production line, so the importance of sunk capital investment, like equipment and machinery, is not as critical as it is to the manufacturing firms in affecting organisational choices. Construction doesn’t meet this (too-basic-to-be-mentioned) proposition. For this reason, it is imperative for us to meticulously examine the robustness of this factor in applying to construction (see Chapter 5.5).

5.3 Assumption 3: Exogenous design

The third implicit assumption of TCE is exogenous design, namely that the product design is common knowledge between traders. The producer knows exactly what to make and the buyer clearly knows what to purchase. The quality is assessed against the predetermined standard. The measurement problems mainly depend on how costly it is to check the extent that received quality deviates from the yardstick. If this mutually accepted quality standard becomes the norm of the market, the quality disputes will relatively easily be resolved by a third party. Under this assumption, how design is decided and the possible impact of design on transaction process will not be of concern. This assumption is most likely to hold for the case of intermediate goods where the attributes of goods have been standardized or/and modularized so that industry standards can be easily followed. In contrast, for the tailor-made production, due to uncertain design, the quality is not as easily verified as it is in the case of standardized products. The measurement problem is not only in quality variations relative to given specifications, but also the assurance that the product quality can fully satisfy the buyer’s demands. Owing to the inclusion of design factors, the efficacy of court-ordering or other arbitration mechanisms will be weakened, so the transaction hazards arising from design problem can be more severe under some circumstances.
Obviously, design plays a pivotal role in construction and so it is necessary to take design into account in the construction procurement system analysis. In the concluding section, we will take a step further to discuss how we should do this.

5.4 Assumption 4: Invariability of production costs

Whilst the objective function of TCE is claimed to be minimisation of transaction costs and production costs [Williamson, 1985], the actual attention is paid to transaction costs by assuming that production costs are invariant to governance structures (hereforth invariance assumption). In the literature, this assumption has been attacked on three grounds – (1) inseparability of production cost and transaction cost, (2) economies of scale, and (3) tacit knowledge. The same concern has been expressed in applying TCE to construction (e.g., Winch, 1996; Walker and Wing, 1999). As a result, the ignorance of production cost cannot be taken for granted in the analysis of construction organisation.

The invariance assumption in fact involves two different versions. The strong form entails production costs are equivalent across governance structures, while the weak form just requires the difference in production costs between governance structures is trivial relative to the differences in other items of costs or revenue so that the ignorance of production costs in the objective function won’t degenerate the prediction power of theory. The testing of the appropriateness of this assumption may either rely on real data or theoretical analysis. This paper prefers the later to the former in that, apart from the formidable costs of data collection when the former approach is used, it is extremely difficult to find identical projects carried out by different procurement systems. Without real data, it is impossible to test the strong form of invariance assumption. Owing to this reason, the objective of this paper is to examine the weak-form invariance assumption in a specific construction context along the existing three lines of critiques. Through the analysis, an attempt is made to demonstrate it is a reasonable approximation to consider production costs invariant to procurement systems.
5.4.1 Three lines of critiques

Critique I: Inseparability of production costs and transaction costs
The first critique is based on a premise that transaction costs and production costs are hard to separate, if not inseparable (Demsetz, 1988; Goldberg, 1985). Lack of a neat delineation between production costs and transaction costs naturally will pose a barrier to properly measure transaction costs and hence the reasoning on the basis of transaction costs cannot be falsified.

Critique II: Ignorance of economies of scale
The second objection arises from the ignorance of the effect of organisation form on the scale of production. As the case made in Milgrom and Roberts (1992), it is a general practice for the patent-controlling electronic company to give another competitive company license to produce the same product. In doing so, his/her potential users’ fear of being held up can be alleviated, which is conducive to the acceptance of this new product. However, this benefit is gained against the loss of production scale. In a industry requiring the large initial capital expenditure, the effect of economies of scale can be a critical factor and the ignorance of this factor may make the analysis go awry.

Critique III: Unreal assumption of constant competence
The third line of attack stresses that the producer’s competence is tacit knowledge that cannot be conveyed by way of audio and visual media, such as language, graph, so, unlike information, it is not tradable and should be accumulated by way of a long period of learning process (e.g., Langlois, 1995; Langlois and Foss, 1996). Under this proposition, the identity of producers will change production costs, making the invariance assumption less likely to hold.

5.4.2 The Analysis
5.4.2.1 Methodology
In economics, production costs are measured by opportunity costs of inputs that are consumed in a production process. The inputs may be devoted to three levels of activities: production activity level, production management level and business man-
agement level. In Fig. 5-3, a company, for the purpose of illustration, is assumed to consist of two production units (e.g. factory), both of which are made up of two production activities (e.g. two production stages in a assembly line). And all of the resource consumption are broken down into two basic types of factors of production: labour and capital\(^3\). In this example, what should be classified as production costs? The key to this question lies in what can be regarded as production activities.

**Fig. 5-3 Three types of production costs**

First of all, few people will disagree that the resources spent at the production activity level (e.g. wages paid to workers and costs of capital employed) are production costs. Second, to reduce the shirking of workers, ensure the quality of outputs and coordinate different production activities, some supervisory and administrative personnel need to be hired. It is generally accepted that the cost of monitoring are transaction costs. For example, North and Wallis (1994) take a position that "hiring a foreman to supervise workers is a transaction cost, since it changes the property rights attached to the labour services by transferring the right to direct labour from the worker to the foreman." Last, for the functioning of the whole company, at business management level, specialised divisions, including accounting, procurement, marketing and so on, may be required. The classification of these costs may be disputable.

\(^3\) The third basic type of factor of production, land, is not directly relevant in the following analysis, so it is ignored.
To address three lines of critiques in a single framework, the strategy of analysis is to restrict the scope of production costs to a smaller set containing items generally accepted as production costs. Those items with unclear classification are grouped as grey costs (GC).

In the literature, production costs per unit output can be affected by two factors: one is the producer's competence (mainly in technology); another is the volume of production. That is, production costs can be expressed as

\[ PC = F(Q,K) \]  

\[ \text{Eq. 5-3} \]

where \( Q \) is the quantity of production and \( K \) the producer's competence. Based on this framework, we want to explore: under what circumstances the aforementioned three lines of attack are effective? The answer lies in what are the alternative governance structures under comparison. Thus, in the following, the invariance assumption of production costs will be examined in different pairs of organisational choices.

5.4.2.2 Comparison of two intraorganisational forms

The first case to be placed under review is the comparison of two internal organisation forms, e.g. labour-owned firm versus capitalist firm. Since these two organizational forms are to be implemented in the same firm, production scale, which is dependent on market conditions, or producer's competence can be held constant. However, the problem as to inseparability of production costs and transaction costs will matter if the cost items in the grey area are quite different. Most of the disputable items of costs come from management activities associated with production, such as administrative outlays of procurement division in the headquarter of a company. Majority of the costs of this sort are of the similar level in different governance structures. For example, the costs of retaining staff in charge of arranging material delivery are unavoidable expenditure, no matter what the firm is labour-managed or capitalist. As long as the classification of these costs is consistently applied, the difference of GC between two intraorganisational forms will be trivial. Thus, the doubt about vague boundary between production cost and transaction cost is not as serious as it seems.
5.4.2.3 Comparison between intraorganisational and interorganisational forms

If we shift the focus of analysis to mundane decision on “make or buy”, there are two production processes for comparison: one is to produce internally; another is to buy goods from the upstream providers. In this case, economies of scale would be an important factor in affecting production costs, in particular for the capital-intensive industries, such as IC manufacturers. If the scale of internal production is much smaller than the production scale that the upstream providers can achieve, it will be disadvantageous to produce internally, making the assumption of constant production costs inappropriate. This point can be illustrated by Fig. 5-4, where $\Delta C$ and $\Delta G$ stand for production cost and transaction cost difference between hierarchy and market (hence $\Delta C + \Delta G$ is the difference in total cost). The magnitude of $|\partial \Delta C / \partial k|$ is an indication of the degree of economies of scale. When $|\partial \Delta C / \partial k|$ increases (from $\Delta C_1$ to $\Delta C_2$), the line of $\Delta C + \Delta G$ will shift to the right, making the threshold value of asset specificity move from $k_1$ to $k_2$. In comparing two alternative organisation choices, if the real cost difference $\Delta C_1$ is inappropriately assumed to be $\Delta C_2$, i.e. the effect of economies of scale is ignored, the transaction with asset specificity in the range of $k_1$ to $k_2$ will be wrongly classified as the transactions suitable for the possible error of prediction will be

$$\varepsilon = \frac{k_2 - k_1}{k_1} \times 100\%$$

Therefore, the quantity of production matters in deciding vertical integration if economies of scale are significant.

Second, owing to the similar reason, the differential production competence will have an impact upon the relative efficacy of alternative governance structures in that the producer will be different between the choice of make or buy.

Last, since part of the costs of internal production is hard to neatly classify, we can separate this disputable items of costs into grey area. However, in this case, the clear criteria for cost classification do not need to be known, since, as long as the criteria are consistently applied to both firms, the difference in the grey area will become trivial.
5.5 An examination of the basic premises of TCE in the context of construction procurement

5.5.1 Three stereotypical categories of procurement systems

It has been stressed that the plausibility of the five premises depends on the nature of procurement systems under comparison, so in the first place it is essential for us to realise the implementation procedure for different procurement systems. With reference to Frank (1998), the process charts of the three stereotypical types of procurement systems - traditional system, design-and-build system and management system are exhibited in Fig.5-5.

The first type is named the traditional system after its popular use in Britain since 18th century. In this system, after perceiving its construction demands, the client has to formulate a brief and retain a design team to crystallize these ideas into a complete set of drawings and specifications. In principle, with complete design, the quantity surveyor can then give a reliable *ex ante* estimate of the construction cost on the basis of bill of quantities, from which the tendering price be objectively assessed. Secondly, because bill of quantity reduces cost of tendering to each bidder (they share the bill of quantity produced by the client), it facilitates large number bidding.
Thirdly, because in principle all bidders are undertaking to constant design and bill of quantity, it encourages use of lowest price as bid evaluation criterion. This is why the traditional system is often characterized by its ability to drive down construction cost by dint of price competition. In the most cases, the project is awarded to the contractors with lowest price. The designer then takes up the responsibility of supervision, checking whether construction work is done as instructed.
If D & B is the procurement system used, the job of the client is to prepare a brief account of his/her requirements, then elicit design and construction proposals from qualified contractors and then award the project to the contractor submitting the most satisfactory plan. The client meanwhile has to organize a team for reviewing the detail design and to supervise the outcome of construction along with the unfolding of the project.

In contrast to D & B where responsibilities are concentrated on the single party, management system tries to separate the project into a reasonable number of parts and introduce a professional organization to coordinate the sophisticated contractual relationships between trade subcontractors and the client. In this system, the client still needs to find a designer first and ask for consultation from the quantity surveyor to get the preliminary estimate of construction costs. In the first three steps, management system is quite similar to but less complete than the traditional system. A key difference is that design is incomplete when the contractor is appointed. However, this time the information available up to this step is used to select management contractor (MC) instead of main contractor. MC is in charge of subcontracting work packages and coordinating different trade contractors involved in the project.

Fig. 5-6 Contractual structures within three procurement routes

(A) Design and Build

(B) Traditional Method

(C) Construction Management

With the aid of these process charts, the contractual structures within these three procurement routes can be exhibited in Fig. 5-6, in which for convenience of exposition the project contains two work packages only. From the contractual perspective, there
is no difference in the total number of contracts used in the three procurement routes. What differs is the boundary of the dotted rectangle and its accompanying effects on the efficacy of procurement routes. Each contract can be said to be a transaction, thereby involving transaction costs. With the dotted area shifting, the bearers of each item of transaction costs change as well. This is an essential point to the analysis of transaction costs in construction.

5.5.2 The assumption of revenue invariability

The value of a project is often measured by the discounted revenue stream generated from it, i.e.,

\[
\text{NPV of revenue} = \sum_{t=1}^{n} \frac{R_t - C_t - C_t^0}{(1 + r)^t} + \frac{RV}{(1 + r)^n}
\]

where \(R_t\) is the revenue stream directly generated from the completed project at time \(t\), \(C_t^0\) is the operating cost at time \(t\), \(C_t^1\) is the cost outlay for land purchase at time \(t\), \(RV\) is residual value of the project after \(n\) period and \(r\) is the discount rate.

In the equation, \(R_t - C_t\) describes the net revenue stream. In principle, the earlier it turns into surplus, other things equal, the more profitable is the project. Thus, the date of completion is a critical variable in the client’s utility function.

Since the procurement systems in fact play the role of instruments with differential ability in achieving the schedule goal\(^4\), so the revenue from the project will be different due to the selection of procurement systems. This means that revenue, an element of the utility function omitted in TCE has to be reconsidered in the analysis of construction procurement.

5.5.3 The assumptions of irreversibility of lump-sum investments and continuous production

The underpinnings of asset specificity are based on two premises: the irreversibility of investments and the continuous revenue stream from the investment. The more costly

\(^4\) Please refer to Sec.6.9 for details.
it is to reverse the investment decision and the more important to maintain a continuous revenue stream, the more serious is the hold-up problem. In other words, the acid test for the explanatory power of asset specificity in construction is to examine how far is construction deviant from the standard case in TCE.

First of all, construction is characterized by project-dependent construction site, project-specific design and hence wide varieties of technology composition across a firm’s portfolio of projects, so construction is a sort of batch production, namely a whole project as the unit of production. With no possibility of setting up a mass production line, most construction equipment and machinery are movable and devised for general purposes. Accordingly, from the contractor’s perspective, the physical capital invested for undertaking the project is not generally specific to the project.

In addition, the postulated continuous production stream resorting from the lump-sum investment is much weaker in construction because the economic life of the construction machinery and equipment is, more often than not, much longer than the duration of the project. To make it more explicit, assume that the contractor has to purchase a precious construction machine with economic life of \( n \) years for undertaking the project with the duration of \( m \) years. If the machine is continuously employed, the NPV of its quasi-rent (QR) is

\[
QR = \sum_{t=1}^{n} (R_t - TVC_t - S_t)\delta_t
\]

where \( R_t \) is the value of production at the time \( t \), \( TVC_t \) is the operating costs of the machine at the time \( t \), and \( S_t \) is the opportunity costs of the machine, often evaluated by rental price. If the project duration is longer than the economic life of the machine, i.e., \( m > n \), then the total QR can be appropriated by the client, just like the case of manufacturing firms. However, in the majority of construction projects, \( m < n \), the appropriable QR in a project becomes

\[
QR = \sum_{t=1}^{m} (R_t - TVC_t - S_t)\delta_t
\]

That is, a part of QR,

\[
QR_t = \sum_{t=m+1}^{n} (R_t - TVC_t - S_t)\delta_t
\]
is not in the grip of the client, so the degree of asset specificity in this case will be much alleviated in comparison to the case of \( m > n \).

Moreover, the significance of asset specificity bears some positive relation with the ratio of lump-sum investment cost \( v_I \) to transaction value \( v_T \).

\[
r_{ii} = \frac{v_I}{v_T}
\]

It implies that the more capital intensive is the production, the more serious are the asset specificity problems, were the lump-sum investment to be locked in the transaction.

Putting two factors together, we can construct an indicator for showing the severity of standard types of asset specificity of physical investments in the mainstream TCE:

\[
w = (1 - \frac{QR}{QR}) \frac{v_I}{v_T} \quad 0 \leq S \leq 1
\]

In the extreme case where \( w \) approaches 1, namely that the economic value of the investment can be fully exploited during the period of a single transaction and the costs of making this investment is nearly equal to the value of the transactions, the specificity of the investment will have the greatest influence on the extent of the potential transaction hazards. Conversely, where the duration of transaction is relatively shorter than the economic life of the machine and its value is only a small part of the transaction value, the importance of asset specificity of this machine in affecting the transaction relationships will be much weaker. This is the case in construction. On the one hand, the value of most construction machinery and/or equipment relative to the value scale of the project is insubstantial; on the other, the duration of a project is often much shorter than the economic life of the equipment. Thus, asset specificity of standard types in TCE will lose much of its power in predicting construction procurement system selection.

5.5.4 The assumption of exogenous design

It is typical in construction that design varies as project. What is required in design will determine what is to be constructed on site. We assume quality can be represented as a proportion of 100% delivery of the design as specified (as \( q = 1 \)). Hence the
value scale of the project, \( v \), is a function of both design \( D \) and quality index \( q \), \( 0 \leq q \leq 1 \).

\[
v = F(q, D)
\]

In construction, there are two types of decision sequences as to design and construction, as simplified in Fig. 5-7. In the first mode (corresponding to traditional system and management system) the client has to decide a preferred design at the first stage, then the contractor needs to choose a conformance quality level of the project according to that design, and finally the client (or his/her representative) has to inspect the quality level \( q \) against the chosen design requirements. In the second mode (corresponding to design and build system), initially, the client has to formulate a brief to instruct the design and build contractor. According to the client’s requirements in the brief, the contractor chooses detailed design and delivered quality standard \( q \).

The first mode is analogous to the standard case in TCE in the sense that design is independent of the production decision. The amount of added value created in the production process is the function of delivered quality standard. Nonetheless, the second mode is quite a different case. When design and construction are under the control of the same agent, i.e., D&B contractor, the problem is not just how to slice a pie of fixed size, but how to slice a pie of adjustable size. As a result, searching for a good
way to accommodate this factor into the analysis is the issue to be tackled. We will return to its analytic implications Sec.6.6.2.

5.5.5 The invariance assumption of production costs
Now we turn to the issue of whether the difference of production costs across procurement systems can be ignored. In answering this question, let’s examine the three lines of critiques (hard separation of TCs and PCs, economies of scale and tacit knowledge) on the assumption of production cost invariability in turn. In the following analysis, reader’s basic understanding of construction procurement systems, such as traditional method, design-build and management system is assumed.

5.5.5.1 Vague boundary between TCs and PCs
In comparing transaction costs of two governance structures, the unclassifiable costs in the grey area (GC) would create a problem if it is probable that the relative proportion of or mix of transaction costs and production costs in the grey area are significantly different between governance structures under comparison. To explore this issue, the analysis start with a narrow definition of production cost (PC), so that residual is GC+TC. Then apply this method to a construction project procurement setting.

Production costs in construction
The definition of production costs in the construction process would be a bit complicated. For the client, the money paid to his/her agents, such as designer and contractor, may be considered as production costs, while the costs incurred for effecting the project may be treated as transaction costs. However, the payment the contractor receives can still be decomposed into production costs that the contractor pays to the subcontractors and transaction costs that are used to bid and manage a construction project. Therefore, defining the production costs according to where the payments go is not helpful in the different context of procurement systems. A better way to explore the relation between production costs and transaction costs in construction should go back to its original definition, i.e., the opportunity cost of labour and capital which is directly connected to design and construction. Any managerial expenditure, no matter who incurs it (designer, main contractor, or subcontractor) should be treated as transaction costs in terms of the whole system. This classification system can help sharpen up the concept of production costs.
Given this narrower definition of production costs, the justification of production cost invariability with respect to the critique of inseparability of transaction cost and production cost entails demonstrating that (1) PC will not vary as procurement systems and (2) to which groups of costs GCS are assigned will not change the result of transaction cost analysis of procurement systems as long as the rule is consistently applied.

**Analysis of PC:**

Production comprises two main processes – design and construction. Each uses labour and capital inputs to produce output. Production costs are the opportunity costs of these inputs. Assume that the opportunity costs of inputs (therefore their shadow prices) are the same across routes. If this holds, then even if market prices depart from shadow prices, we can still compare PCs across GSs by comparing quantities of inputs multiplied by market prices of inputs, since any deviation will be of similar proportion in each GS. Now we need to show:

1. all routes exist within same factor markets, and thus have common factor prices;
2. that there is not significant difference in input quantities to produce same output, between GSs.

The first point seems highly plausible. What about the second point? We need not to define what mean by 'same output'? Because design and construction are integrated in one of the GSs, the analysis needs to consider the combined PCs of design and construction (unless we say that relative production costs of design are trivial in relative to production costs of construction, and thus can be neglected; in that case, we need to show only invariance of production costs of construction). We assume that all clients proceed first by writing a performance or functional output specification; and second by applying value engineering process to design prior to approving them (if this remain a client role, i.e., if a non-integrated route is chosen). This value engineering for the client we assume produces parallel convergence of finally adopted design on the production cost minimizing design to that achieved by an integrated producer applying their own value engineering or buildability heuristics.

**Analysis of GC**

As mentioned before, we may not have consensus over the categorisation of some cost items, GCS. If the invariance assumption has to hold, what needs to be shown is
that, no matter what types of these costs are classified, this part of costs will be can­
celled out when alternative procurement routes are compared and won’t affect trans­
action cost reasoning. We have a reason to claim that it is the case because the total GCs varies as the total number of procurement activities in a construction process, which can be considered as nearly constant for a given design.

5.5.5.2 Economies of scale
Second, in the industry with homogenous products, the orders from different buyers (say A, B, C) can be added together to generate the effect of economies of scale, as shown by the bold arrow in Fig.5-8. However, construction is a sort of project-oriented production, so there are no obvious spillover benefits when undertaking many projects at the same time. Put it more formally, the totality of the production costs of carrying out the project X and Y sequentially won’t make too much difference to that of carrying out X and Y at the same time, i.e.

\[ C(X) + C(Y) \approx C(X,Y) \]

Consequently, this critique can be dismissed fairly quickly.

Fig. 5-8 The effect of economies of scale

5.5.5.3 Tacit knowledge
The challenge that tacit knowledge is non-transferable so that production costs cannot be equal between competing producers is important in construction process as a whole. But this concern is not relevant in the analysis of construction procurement system selection because most of that tacit knowledge is held by specialist subcontractors, not
by coordinators. The procurement setting is about appointment of project coordinators. But however the appointed coordinator will draw upon the same pool of specialist subcontractors, and thus the same pool of tacit knowledge in this respect.

5.5.5.4 Summary: production cost invariability in three contexts
The results of the analysis in Sec.5.4 and Sec.5.5 are summarized in Table 5-1, from which we conclude that the assumption of production costs invariability is most likely to hold in dealing with the problems of procurement system selection since for this case, there are plausible defence against each of the three lines of attack.

Table 5-1 Relevance of three lines of attack to the three sets of alternative governance structures

<table>
<thead>
<tr>
<th>Points of attack</th>
<th>Governance Structures under Comparison</th>
<th>Vague boundary between transaction costs from production costs</th>
<th>Scale economies</th>
<th>Tacit knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labour-owned firm vs. capitalist firm</td>
<td>Irrelevant</td>
<td>Irrelevant</td>
<td>Irrelevant</td>
<td>Irrelevant</td>
</tr>
<tr>
<td>Producing internally vs. Outsourcing</td>
<td>Irrelevant</td>
<td>Relevant</td>
<td>Relevant</td>
<td>Relevant</td>
</tr>
<tr>
<td>Traditional method, design and build, management system</td>
<td>Irrelevant</td>
<td>Irrelevant</td>
<td>Irrelevant</td>
<td>Irrelevant</td>
</tr>
</tbody>
</table>

5.6 A summary of chapter five
Through examining the five premises in TCE in the context of procurement systems, some conclusions can be made, as summarized in Table 5-2.

(1) The NPV of the revenue from the project can be quite different if different procurement systems are used, so the revenue cannot be assumed to be invariant to procurement systems.

(2) On the one hand, construction equipment and machinery are more or less removable, so the degree of irreversibility of the investment is much slighter than that in the manufacturing industry. On the other, even for projects which require expensive specific machinery, the duration of the project is often much shorter than the economic life of that machinery, so the threat of expropriating quasi-rent is not as severe as the general case of continuous production. Putting these factors together
leads us to doubt whether the standard types of asset specificity provide the same explanatory power to the construction case.

(3) Exogenous design obviously is not a reasonable premise for tailor-made production, like construction.

(4) Fixing production costs in carrying out the benefit-cost analysis of procurement systems can be regarded to be a sensible approximation from the comparative perspective.

Table 5-2 Comparison of five assumptions in TCE as applied to construction procurement system selection

<table>
<thead>
<tr>
<th>Assumptions in TCE</th>
<th>Assumptions (Hold or not)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Invariability of revenue from production</td>
<td>No</td>
</tr>
<tr>
<td>2.A Irreversibility of lump-sum investment</td>
<td>slight</td>
</tr>
<tr>
<td></td>
<td>Most construction equipment/machinery are movable and for general purpose</td>
</tr>
<tr>
<td>2.B Continuous production</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Project duration is much shorter than the economic life of the specific machinery</td>
</tr>
<tr>
<td>3. Exogenous design</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Design is not given.</td>
</tr>
<tr>
<td>4. Invariability of production costs</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>The difference of production costs across procurement systems is relatively minor to that of transaction costs, so production costs can be assumed to be invariant to organization.</td>
</tr>
</tbody>
</table>

To remove the barriers on our way to successfully apply transaction cost reasoning to the problem of construction procurement system selection, we have to modify the original TCE framework in the following respects (see Table 5-3).

First of all, revenue will change as procurement system choice, so the postulated objective should be changed from minimization of transaction costs to maximization of profits. Second, the weak power of standard types of asset specificity in differentiating the relative efficacy of procurement systems, together with the observed importance of hold-up problems in construction, gives rise to the need to reexamine the origin of the transaction hazards and discover the new factor that can replace the func-
tion of asset specificity in TCE. Last, the economic implications of including design into the model have to be fully explored. All three points will be the focus of analysis in the next chapter.

### Table 5-3 The modifications of TCE for the analysis of construction procurement systems

<table>
<thead>
<tr>
<th>Assumption not valid in the case of construction</th>
<th>Modification to be made in Chapter 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invariability of revenue from production</td>
<td>Add revenue into the objective function</td>
</tr>
<tr>
<td>Irreversibility of fixed investment + continuous production</td>
<td>Identify process specificity as the origin of quasi-rent in the construction context</td>
</tr>
<tr>
<td>Exogenous design</td>
<td>Identify the source of rent generated from exogenous design</td>
</tr>
</tbody>
</table>
6.1 An overview of the theory

6.1.1 The basic logic

The construction procurement process can be regarded as a special form of economic organisation, so the basic principles of transaction cost economics enunciated in Chapter 4 can be applied. Nonetheless, to incorporate the unique features of construction (see Sec.5.5.2-5.5.5), a significant and substantial modifications and extensions should be made.

Fig.6-1 The trilateral relation between procurement routes, decision function and observable factors

As TCE’s logic is applied to construction procurement, we in fact are dealing with the problem of discovering the trilateral relation between procurement route selection, decision function and observable factors (see Fig.6-1). That is, we need explicitly to point out the observable conditions under which one procurement route becomes more likely to achieve the client’s objective function than others. If procurement routes are conceived of as instruments for achieving the client’s investment objective, those factors affecting the effectiveness of these instruments will be our concerns. To interpret the regularities underlying the client’s choices, some observable factors need to be found for connecting up the causal link between the revealed outcomes of the client’s choice and the postulated objective function. Three observable variables will be proposed in this chapter as relevant: (1) the client’s time value, (2) process specificity and (3) degree of specialisation (between subcontractors or trade
contractors). Different combinations of these three factors constitute the set of conditions under which each procurement route is most desirable. We can use one of the conclusions in Sec.6.7 as an example: from the theoretical perspective, it can be inferred that high cost of time, high process specificity and low degree of specialisation will be the conditions making management system most likely to achieve the client's goal of profit maximisation. Lying behind this kind of conjecture, there should be an explanatory model justifying the implied causal relations. The methodology of TCE presented in Fig.4-1 can serve as a blueprint for building up this model.

6.1.2 A brief overview of the model
Since the construction client is assumed to pursue the goal of profit maximization instead of transaction cost minimization, an additional step - operationalisation of revenue function - should be added into the framework in Fig.4-1. A modified three-stage scheme for the analysis of construction procurement route is exhibited in Fig.6-2.

![Fig. 6-2 A modified trilogy for the analysis of construction procurement routes](image-url)
In addition to the objective of transaction costs minimization being replaced by profit maximization, the second difference of this model from mainstream TCE lies in its stress on the role of measurement costs. In the tradition of Klein et al. (1978), it is claimed that the presence of quasi rent is the origin of major contractual problems. Another source of rent - quality rent - underlying the foundation of measurement branch of TCE is in large part downplayed in Williamson’s theory. This study attempts to restore its significance out of two considerations. The first stems from the observation that the relative importance between quasi rent and quality rent will tip towards the latter in the service-intensive transaction where smaller initial fixed costs are sunk and the complete description of mutual agreed quality standard is hard to specify in the written form ex ante. Another is the presence of design rent, a special form of quality rent, due to inclusion of design in the analysis. For these two reasons, measurement costs are elevated to be the third pillar (other two being the client’s time value and process specificity) for interpreting construction procurement behaviour.

The third point distinguishing the methodology of this study from TCE is the introduction of a new type of asset specificity: process specificity. As will be explained further in Sec.6.5.4, the ultimate reason for explaining the occurrence of process specificity lies in a unique feature of construction production: product is fixed on a site provided by the buyer. Because the loss of scraping the partially completed project is much dearly than that of finding a replacement contractor, the client will choose to the latter option rather than the former. That is, the feature of on-site assembly is a major reason of making the client vulnerable. This source of specificity only takes place in construction, so process specificity is a special type of asset specificity that have never been identified in the literature.

6.2 The process perspective in the context of construction
The demand for construction services is derived from the final demand for constructed products, such as buildings, roads and plants. To make the project be completed on time, within budget and with required quality, the initiator (the client) has to bring together related trades and professionals (architect, contractor, subcontractor) to form a project coalition to take part in the process of project delivery. This process can be examined in three contexts (see Fig.6-3).
The first context is a standard economic decision problem, in which decision is made by looking forward, i.e., taking the expectations of decision variables and discounting them to the present. The second context attempts to capture the interaction between trading parties in the course of transaction. In the last context, the decision maker has to assess the decision made in the first context with reference to the realized outcome and explore the implications for the decision in the first context. This framework of analysis is a manifestation of the thorough implementation of Williamson’s *farsighted contracting in its entirety* [Williamson, 1996].

### 6.2.1 Context I: looking forward

According to the basic economic principle, the condition for the client to launch a new project is the presence of non-negative economic rent\(^1\), i.e.,

\[
V - \sum P_i \geq 0 \quad \text{Eq. 6-1}
\]

where \(V\) is the value of the project and \(P_i\) is the payment to agent \(i\). On the same ground, the incentive for all the agents to participate the project coalition is their non-negative rent, i.e.,

\[
\forall i \quad P_i - C_i \geq 0 \quad \text{Eq.6-2}
\]

---

\(^1\) Economic rents are defined as “the portion of earnings in excess of the minimum amount needed to attract a worker to accept a particular job or a firm to enter a particular industry.” [Milgrom and Roberts, 1992]
where $C_i$ is the cost incurred by agent $i$. It follows that

$$\Sigma (P_i - C_i) \geq 0$$  Eq.6-3

Combining Eq.6-1 and Eq.6-3 gives the condition for the existence of a project coalition,

$$V - \Sigma C_i \geq 0$$  Eq. 6-4

This means that, when the initiation of a project is profitable to every party involved, this transaction will take place. In contrast, such will not be the case, if

$$V - \Sigma P_i \geq V - \Sigma C_i \geq 0$$  Eq. 6-5

This signifies that, whilst from the point of view of the client (the first sign of inequality) and the whole system (the second sign of inequality), the project coalition will emerge, some agents will suffer losses in engaging in this transaction due to $\Sigma P_i - C_i \leq 0$. However, why would some agents voluntarily enter an agreement to lose money? According to observations, some contractors in the period of recession may undertake projects purely for the reason of cash flow. These cases are relatively infrequent and should not be of main concerns. The real problems lie in the divergence between *ex ante* commitments and *ex post* realizations. Before forming the project coalition, all the participants have to estimate $C_i$ according to available information, upon which the formulation of pricing strategy is based. What the client is concerned with is not the estimate of $C_i$, but the price agreed upon by the agents. If the offer $P_i$ is attractive enough, the client will still carry on the project even though some agent $i$ may suffer losses (i.e., $P_i - C_i \leq 0$). If $P_i$ is deliberately set below than the expected costs $C_i$ *ex ante* and agent $i$ seeks compensations by downgrading quality or reducing quantity *ex post*, then the apparent cost saving of $|P_i - C_i|$ will backfire to the client. Accordingly, the positive economic rent for both the client and his/her individual agents, i.e., requiring Eq.6-1 and Eq.6-2 to hold simultaneously, is a safe condition to avoid dissolution of a project coalition *ex post*.

### 6.2.2 Context II: dynamic process of construction

The construction procurement route in fact is a process of organising project participants. After the agreement is entered into, the estimated costs will be realised gradually. A complete description of the dynamic process of a construction transaction contains two parts: one is accounting-based cost-benefit analysis and another is opportunity-cost-based cost-benefit analysis. To play a role in a project
coalition, all participants need to spend “actual” resources in exchange for the expected profits, while at the same time they also lose the opportunity to switch the dedicated resources to the alternative use without costs. These two systems are complementary, one showing how much money has been spent and earned and another indicating how wisely you have exploited the economic value of the resources that you manage. In principle, opportunity costs play the role of choosing the right means to achieve the goal of increasing accounting surplus. And the dynamic changing of opportunity costs and accounting costs will affect the interaction relation between trading parties in the course of transaction.

Different from the paradigmatic case of vertical integration explicated in Chapter 4, the vulnerability in the construction process tilts to the client. Analytically, as discussed above, the condition under which the client is willing to initiate a project is the presence of non-negative economic rent at $t=0$, i.e.,

$$\text{Client's rent } CR = V - P \geq 0$$  \hspace{1cm} \text{Eq. 6-6}

where $V$ is the value of the project to the client and $P$ is the total payment to the agents involved for covering production costs incurred by them. We are concerned with the hold-up problem facing the client, were the project to be disrupted. This problem can be discussed in two scenarios: design process and construction process.

### 6.2.2.1 Hold-up problem in the design process

This specifically refers to the interaction between the client and the designer in the traditional method. In the first place, the initiation of the design process must be based on the client’s non-negative economic rent $CR_d$, i.e.,

$$CR_d = V_d - P_d \geq 0$$  \hspace{1cm} \text{Eq. 6-7}

where $V_d$ is the client’s evaluation of design work; $P_d$ is the total payment to the designer. Likewise, we can express the designer’s rent as

$$DR_d = P_d - C_d$$  \hspace{1cm} \text{Eq. 6-8}

where $C_d$ is production cost incurred by the designer. Since the nature of hold-up problem lies in the opportunity cost and the sunk cost of the vulnerable party, we need to address the problem of quasi-rent from the standing point of both parties.
Client’s quasi-rent during the design process

At a point of time, say t=1, in the post-contract stage, the client’s valuation on this design is made up of two parts: evaluation on the finished part of design work, $V_{DF}$ and that on the unfinished part, $V_{Du}$ (i.e., $V = V_{DF} + V_{Du}$). In terms of cost, if we denote the design fee to be paid for the remaining part of design $P_{Du}$, the payment gone to the designer for the completed part of design $(P_{D} - P_{Du})$ is sunk and thus should be expressed as the value of the completed part of design in the alternative best use, $V_f$. As explained in Sec.4.2.6.1, quasi rent shows the condition under which transactors are willing to stick to the original contractual agreement. In this case, the greater value of the uncompleted part of design than its cost is a condition inducing the client to carry on the original contractual relation with the designer. Formally, this can be shown that

$$CQR_D = (V - V_f) - P_{Du}$$

$$= (V_{DF} + V_{Du} - V_f) - P_{Du} \geq 0$$  \hspace{1cm} (Eq. 6-9)

To simplify the following discussion, we assume that, at t=1, the client evaluates the uncompleted part of design in the same way he/she did ex ante, i.e., $V_{Du} = P_{Du}$. Under this assumption, Eq.6-9 can be reduced to

$$CQR_D = V_{DF} - V_f$$  \hspace{1cm} (Eq. 6-10)

This equation means that the client’s quasi rent comes from the difference between the client’s $ex\ ante$ evaluation on the completed part of design and its $ex\ post$ value. For ease of giving this statement an intuitive interpretation, a further assumption is made that the client’s $ex\ ante$ rent is zero\(^2\), i.e., $V = P_{D}$. It follows that

$$CQR_D = (P_{D} - P_{Du}) - V_f$$  \hspace{1cm} (Eq. 6-11)

On the right side of this equation, the first term is the client’s $ex\ ante$ willingness to pay (for the completed part of design) and the second term is the value of that partly-completed design in the alternative best use. It means that the origin of this difference lies in the fact that $the\ combined\ value\ priced\ according\ to\ individual\ constituent\ parts\ of\ a\ design\ work\ is\ less\ than\ the\ value\ of\ a\ design\ work\ priced\ as\ a\ whole.$

A numerical example will be helpful to understand the economic meaning of Eq.6-11.

\(^2\) The following argument can run and the same conclusion be reached by relaxing this assumption, but the discussion would be more complicated.
Assume the client’s willingness to pay for the design work is £100,000 and the fee will be paid according to the percentage of completion. At the point of 30% completion, the client should have paid the designer £30,000. This amount of money also indicates the client’s evaluation on this part of work under the precondition that the design will be completed. However, in terms of the opportunity cost of this part of design, 30% of drawings and specification is of little value, say £5,000, if the design is not to be continued. When the hold up demand from the designer occurs, the client is in a weak position to bargain with the designer since the value of the partially-completed design in the alternative use is low. This means that CQRd (in this case £25,000) is an appropriable rent.

**Designer’s quasi-rent during the design process**

Based on our above discussion, at t=1, the designer’s quasi rent can be expressed as

\[ DQR_d = P_{du} - C_{du} \]  

Eq. 6-12

where \( P_{du} \) is the payment to be received; \( C_{du} \) are the cost to be incurred by the designer. If we maintain the two assumptions that (1) the *ex ante* rent of the client is zero and (2) the cost expectation, at t=1, of the uncompleted part of design (\( C_{du} \)) is consistent with *ex ante* expectation, i.e., \( P_{du} = C_{du} \), then designer’s quasi rent at t=1 is zero. This means there is no lock-in effect on the designer’s side. This result will hold even if the *ex ante* economic rent is positive. For the case where the expected profit rate is \( \pi \), i.e., \( P_D = (1+\pi) C_D \), designer’s quasi rent becomes \( \pi C_{dc} \). The rent is not appropriable since the rent is a part of realised profit margin. The client is not able to appropriate this rent since the designer’s sunk cost has been fully reimbursed.

**Who is more vulnerable between the client and the designer?**

According to the above discussion, the client seems much vulnerable than the designer. However, it is to be noted that the principal reason making the designer free from hold-up problem relies on a hidden premise that payment and cost are synchronised. In practice, the interim payments are often held until the client is satisfied with the progress and quality of design work. This practice will substantially weaken the designer’s position to refuse to modify design as the client demands. If \( P_D - P_{du} \) is only a fraction of the money that should have been paid, CQRD will be
lower, meaning the client is not as vulnerable as it appears under the assumption of synchronisation of payment and cost.

Besides, to get over the problem of information asymmetry prevailing in the construction design market, the client heavily relies on the track record to prequalify the competence of designers. As a result, the reputation effect in the construction design market is also a critical instrument to discipline the designer. For these two reasons, the hold-up problem in the relation between the client and the designer is not severe.

6.2.2.2 **Hold-up problem in the construction process**

**Client's quasi-rent in the construction process**

We now turn to the hold-up problem that may happen after construction starts on site. The first scenario is that, were the project to be discontinued, the client as last resource is to alter the final purpose of the partly-completed project. In this case, the client's quasi rent will be

\[ CQR_{C1} = (V - V_c) - P_{cu} \quad \text{Eq. 6-13} \]

where \( V_c \) is the return of the completed part of the project in its alternative best use and \( P_{cu} \) is the payment for the uncompleted part of construction work. Thus, Eq.6-13 stands for the cost-benefit appraisal of the remaining part of the project.

Correspondingly at this point (i.e., \( t=1 \)), the original payment scheme can be divided into two parts: (1) payment made \( (P_{Cr}) \) and (2) payment to be made \( (P_{Cu}) \). These can be regarded as the client’s evaluation of the project in terms of costs. For simplicity, we assume that the ex ante economic rent is zero due to competition in the market for which the project is built, i.e.,

\[ V = P_{Cr} + P_{Cu} \quad \text{Eq.6-14} \]

Substituting Eq.6-14 into Eq.6-13, we can get

\[ CQR_{C1} = P_{Cr} - V_c \quad \text{Eq.6-15} \]

Normally, the value of a partly-completed project is appreciably lower if the project is not completed to serve the planned function, i.e., \( P_{Cr} \gg V_c \). As long as \( CQR_{C1} > 0 \), the client will stick to the original contract to carry on the project. Thus, \( P_{Cr} - V_c \) is a source of the appropriable rent.
In the second scenario, the client may choose another option, that is, finding a replacement contractor to take over from the original contractor. In this case, the client’s quasi rent $CQR_{c2}$ is equal to the cost of switching$^3$.

\[
\text{Fig. 6-4 Client’s quasi-rent at one point of time during construction process}
\]

Thus, a measure for the lock-in effect on the client’s side can be defined as

\[ CQR = \min \{CQR_{c1}, CQR_{c2}\}. \]

The graphical presentation of this function is shown in Fig.6-4. $CQR_{c1}$ is described as a vertical line, i.e., a constant because $P_{cr} \gg V_c$ and, at a point of time, $P_{cr}$ is given. In contrast, $CQR_{c2}$ will change as design/technology characteristics of the project, so it is a straight line going through the point of origin. When $CQR_{c1} \geq CQR_{c2}$, $CQR$ is varying from 0 to $CQR_{c1}$; when $CQR_{c1} \leq CQR_{c2}$, $CQR = CQR_{c1}$. It means that the value of the partially-completed project in the alternative use is given, dependent on the final use of the project, but the cost of switching is endogenous, changing as design or technology adopted. This definition tries to capture the spirit of opportunity costs. More explicitly, the alternative option that the client has will always impinge upon his/her bargaining power in the current transaction, which will in turn affect the transaction hazards and the severity of the accompanying rent-seeking problems. We assume that the original function and scope of the project is chosen by way of sound

---

$^3$ We will scrutinise the economic meaning of this term Sec.6.5.4.
feasibility study, implying that scrapping a project halfway will be much more costly than replacing the contractor. Thus, in this study, CQRc is treated as the binding condition.

**Contractor’s quasi-rent in the construction process**

The contractor’s hold-up problem involves fixed investment, which makes the analysis more complicated. For convenience of the following analysis, a simplified form of the contractor’s (expected) rent at t=1 is expressed as

\[ \text{BR}_C = P_c - \sum_{t=1}^{t_w} (VC_{Ct} + I_{Ct}) \]

where \( P_c \) is the total payment for undertaking the construction work; \( VC_{Ct} \) is the variable cost, such as project overhead, or wages paid to supervisory personnel and payments to subcontractors; \( I_{Ct} \) refers to the lump-sum investment for machinery and equipment. The subscript \( t \) indicate the cost incurred at the period of \( t \); \( t_w \) is the whole duration of the project.

At the post-contract stage, say \( t_1 \), the contractor’s quasi-rent becomes

\[ \text{BQR}_C = P_{Cu} - \sum_{t=t_1}^{t_w} (VC_{Ct} + S_{Ct}) \]  
Eq. 6-16

where \( P_{Cu} \) is the payment to be received; \( VC_{Ct} \) variable cost; \( S_{Ct} \) ex post opportunity cost of the lump-sum investment. Assume the ex ante profit is zero, so with reference to the expectation made at \( t_1 \),

\[ P_{Cu} = \sum_{t=t_1}^{t_w} (VC_{Ct} + I_{Ct}) \]  
Eq. 6-17

Substituting Eq.6-17 into Eq.6-16 gives

\[ \text{BQR}_C = \sum_{t=t_1}^{t_w} (I_{Ct} - S_{Ct}) \]  
Eq. 6-18

From Eq.6-18, we can realise contractor’s quasi rent comes from the same source as the case discussed in Chapter 4 – the difference between *ex ante* and *ex post* opportunity cost of lump-sum investment. Take a numerical example for illustration.

If we have the following information,

---

4 We don’t consider discounting factors in this formula.
(1) a contractor spent £1,000,000 for purchasing a machine at the outset of a project;
(2) the prevailing interest rate is 8% per annum;
(3) the rental income of contractor’s newly purchased machine is 3% of its purchasing price per annum;
(4) the total duration of the project is three years and \( t_1 \) is at the end of the first year.

Before the machine is purchased, the contractor can earn interest by depositing one million pound in the bank. However, after this investment is sunk, the alternative best use of this machine is assumed to be used for rent. For simplicity, we further assume that discounting is ignored and all income payment is made at the end of each year. Then at \( t_1 \), the contractor’s quasi rent is equal to

\[
BQR_C = 1,000,000 \times \sum_{t=2}^{3} (0.08 - 0.03) = 100,000
\]

This means that, if at \( t_1 \), the maximum rent can be expropriated by the client is £100,000. Beyond this amount of money, the contractor would rather stop the project and rent the machine to other contractors. Nonetheless, generally speaking, most high-valued equipment and machinery in construction are removable and not specific to a particular project, so relative to the scale of the whole project, \( I_e - S_e \) is not expected to be large.

**Who is more vulnerable in the construction process?**

From the above discussion, we come to realise that the client appears to be more vulnerable than the contractor at the post-contract stage since it is more costly for the client to find a replacement contract than for the contractor to switch the machinery or equipment to an alternative use. As a result, in modeling the interaction between the contractor and the client, it seems reasonable to treat the contractor as net rent advantage party (NRA party).

One may be interested in the question as to whether the contractor’s advantage on relatively slighter lock-in will be weakened by retention or postponement of payment as in the analysis of the designer. The answer is yes, but to a much lesser extent. There are three reasons. Firstly, the contractor can mitigate the negative effect of payment retention on cash flow by taking the strategy of ‘pay when paid’ to subcontractors.
Take an analogy of reservoir. This is just like controlling outflow in response to inflow so as to keep the reserve of the reservoir stable. If the contractor can easily transfer the pressure of keeping positive cash flow, the function of retention in disciplining the contractor will be weakened.

Secondly, to achieve the purpose of positive cash flow earlier, the contractor is liable to 'frontload' the payment by setting the unit prices of cost items incurred at the early stage higher than it would be and evening out the unbalanced payment by lowering down the unit price of items incurred at the later stage. This strategy makes the contractor’s payment ahead of cost, implying under this condition retention is not as effective as it appears to be.

Last, the strategy of payment retention is at best a preemptive instrument, and is not able totally to curb the potential opportunism from the contractor. Our reasoning is as follows. Payment retention is not the money that the client is entitled to determine whether to pay unilaterally. When disputes between both parties arise, the final resolution of disputes has to go through a pre-agreed-upon mechanism, such as third-party arbitration or the court. The key point is still whether the client can afford to the loss arising from going through this process. If no, the client’s bargaining power coming from holding a part of payment due will lose a lot of its preemptive function.

In this section, we try to argue that payment retention that can be used as a powerful instrument for disciplining the designer is not effective enough to counterbalance the client’s weak position. This conclusion leads this study to explore how the client can mitigate his/her vulnerability by choosing an appropriate procurement system in Sec.6.6.

6.2.3 Context III: Equilibrium process
When a transaction goes to the final point of \( t_2 \), the client has to assess the performance of a chosen procurement system. Profit is a standard indicator. The positive profit signifies that the right GS has been made. The same decision will be followed next time. On the contrary, if the return is not large enough to compensate the revenue loss resulting from giving up the opportunities of using the resources in
the alternative best use, the principle of decision needs to be modified to improve the outcome of the next decision. In the organisation analysis, the reasonable response to the bad performance of the currently used GS is to switch to alternative GS next time. In equilibrium, the transaction can be dealt with by the most appropriate GS so as to maximise the gains from the transaction. This process must be very complicated, involving a multitude of factors acting on it. In the neoclassical economics, the achievement of equilibrium is assumed, rather than explored explicitly. This assumption makes the inference of theory determinate and parsimonious, but it also attracts critiques on different grounds, unrealism being the most prominent. There is no denying that those attacks have made some mark. But, most refutable hypotheses derived from equilibrium analysis have been proven to stand up to empirical testing (in particular those derived from TCE), so this assumption still firmly stands in the economic literature. This study follows this tradition and assumes that in equilibrium the client is able to make an appropriate choice of procurement routes.

6.2.4 Why process perspective?
The three contexts - looking forward, dynamic process of transaction and equilibrium process, corresponds to the key issues emphasised in the three streams of contemporary economic theories of organization - agency theory, transaction cost theory and evolutionary theory. Agency theory concentrates on the role of contract in providing incentive and allocating risks. The contract is assumed to be a complete framework under which every possible contingency is well stipulated by written clauses so that trading relationships can be effectively governed. In this setting, the second context is redundant because potential transaction problems have been assumed away. In contrast, transaction costs theorists try to unravel the black box of the transaction process, placing the main attention on the relation between process dynamics and efficiency of GSs. Both agency theory and transaction cost theory rely heavily upon the hidden premise that the force of natural selection will weed out the inefficient type of organisation and leave the efficient one to survive. Thus, in equilibrium, the observed choices of organisation are all optimal as if they follow the theoretical principles to make the decisions.
This assumption has been criticized by the proponents of evolutionary theorists [Hodgson, 1988]. As its name shows, this method is characterized by stressing the evolutionary process and selection mechanism, not just focusing on the outcome of evolution. However, up to the present, an unified framework is still awaited.

Though, of the three approaches, no approach enjoys absolute dominance, transaction cost theory seems to have shown its advantages in two respects: (1) more realistic assumptions relative to agency theory and (2) its ability to develop a refutable theory for explaining new issues, such as the boundary of the firm. The logic underlying Williamson’s theory is the comparative institutional analysis, by which alternative institutional arrangements are assessed against each other in the context of transaction process. We have witnessed the impact of this approach on the economic analysis of organisation. Along with the unpacking of the “black box” of internal organisation, some interesting issues, like the boundary of the firm, can be addressed. Our concern is not restricted to the function of contract in risk allocation, but furthermore to explore the efficacy of organisation in governing the whole transaction. Let’s compare the fundamental difference between contract paradigm and governance paradigm (see Table 6-1).

<table>
<thead>
<tr>
<th>Points of difference</th>
<th>Process perspective</th>
<th>Contract perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>The effectiveness of contract</td>
<td>Incomplete</td>
<td>Complete</td>
</tr>
<tr>
<td>Nature of procurement systems</td>
<td>A type of governance structure</td>
<td>A form of contract</td>
</tr>
<tr>
<td>Main issue of analysis</td>
<td>Alignment of governance structure with transaction attributes</td>
<td>Incentive compatibility</td>
</tr>
<tr>
<td>Focus of analysis with respect to time</td>
<td>The process in its entirety</td>
<td>A point in time</td>
</tr>
</tbody>
</table>

To model the transaction behaviour, the trading rules have to be specified. Contract in the standard principal-agent theory is assumed to be a perfect instrument for enforcing the predetermined division of economic rent. The incentive structure is embedded in
the form of contract, under which the range of actions that all parties can choose in response to different contingencies is well defined. Hence, trading parties just need to negotiate for mutually beneficial terms of transaction with reference to their risk preference, and the efficient outcomes will be assuredly obtained, i.e., the total utility will be maximized. In this setting, the whole transaction process is treated just as a point in time. All of the transaction problems originate from whether the incentive clauses are properly specified in the contract, ruling out the potential hazards likely to arise in the course of transaction.

In contrast, the proponents of TCE start from an opposite proposition that contract must be incomplete under the constraint of bounded rationality. The relaxation of unbounded rationality makes TCE shift the focus of analysis to the implications of incomplete contracting on the efficacy of organization. Under this paradigm, the presence of loopholes in contract may lead to inefficient redistribution of quasi-rent. This study attempts to demonstrate the implication of this process perspective as potentially able to generate the similar impact on the analysis of construction procurement systems.

6.3 The unit of analysis: construction transaction
6.3.1 The nature of construction transaction
In the analysis of construction procurement system selection, construction transaction is taken as the unit of analysis. It consists of the process of designing and constructing a construction project. That is, what to be exchanged in this transaction are design and construction of a project. It is desirable for the client to ensure the process of this transaction is well organised so as to get the project delivered against as the targets set ex ante. This process can be organized in various ways. The way this process is organised can be viewed as a governance structure by which the relations between the client and his/her agents are specified normally by way of contractual arrangements. The client’s concern is with design and construction delivered by the agents who bear direct contractual obligations to the client. Varying as procurement systems, the project governance structures involve different agency relationships as a result. Both in the traditional system and management system, there are multi-agents directly responsible for the single principal, the client. In contrast, in the design-build system,
the client has a bilateral trading relations with D&B contractor. Whilst the relationship between the client and his/her agents change as procurement systems (bilateral transaction or multiple agency relationship), the subject matter of analysis is still the same, i.e., how costly it would be for the client to achieve the managerial goal (ex. profit maximisation) by way of different contractual arrangements. As a result, the essence of taking transaction as a unit of analysis is about the "content" of a transaction and explore why the efficacy of different governance structures will display differential competence for different types of transactions, but not about governance structure itself.

6.3.2 Efficiency of procurement systems
One of the most important functions of an economic organisation is effectively to coordinate all trading parties involved so as to complete a value-generating process. The value added by the producer stems from the surplus generated from higher selling price of exchanged goods or service than opportunity costs expended in production and transaction. Due to the intermediation of economic organisation, a great deal of profitable business ventures can possibly be realized.

When it comes to efficiency of running an organisation, the principle of wealth maximization is often applied [North, 1990]. In a world where uncertainty prevails, this principle is not as straightforward as it seems in Sec.6.2 and entails detailed scrutiny.

First of all, the organizational efficiency can be assessed on the basis of wealth maximization only when three assumptions can hold: (1) risk neutrality, (2) no financial constraint, (3) common belief. If trading parties share the same belief about the future and are neutral to risk allocation, the total risk of the transaction should be shared evenly by both parties through bargaining. With the assumption of perfect capital market, financing availability will be unlimited, i.e. any transaction can get external finance. Then, the change of wealth of every agent will fully reflect the change of utility. But we wonder why wealth maximization implies efficiency? According to the standard criterion of efficiency – Pareto optimality - the situation can be said to be optimal when no one can be made better off without making other
people worse off. If the measure for “better off” or “worse off” is in monetary unit, wealth-maximizing organisational arrangement indicates that the gains from transaction are on the attainable frontier. One party’s extra gain will signify the other party’s loss. Thus, wealth maximization is an equivalent goal of efficiency. It is to be noted that when **wealth effect** is present, the situation will be more complicated. Wealth effect means that the marginal utility of one unit increase in wealth is lower for the rich than for the poor. Thus, it is possible to find non-wealth-maximizing arrangement is efficient as well. For example, the organization is made up of two members, high wealth agent (HWA) and low wealth agent (LWA), with different initial wealth $W^0_{HW}$ and $W^0_{LW}$. If GS1 can increase the wealth of two agents by the amount of $W^1_{HW}$ and $W^1_{LW}$ (both of which can be positive or negative), so the total utility $U^T$ is

$$U^T = U(W^0_{HW} + W^1_{HW}) + U(W^0_{LW} + W^1_{LW})$$

When the wealth effect is present, i.e.,

$$U'(W^0_{HW}) < U'(W^0_{LW})$$

It is possible for us to find a pair of $X$ and $Y$ so that

$$U^T = U(W^0_{HW} + W^1_{HW} - X) + U(W^0_{LW} + W^1_{LW} + Y)$$

Keeping the total utility at the same level, the wealth effect gives

$$X > Y$$

If a GS2 can deliver this performance, the total wealth of HWA and LWA is

$$W^0_{HW} + W^1_{HW} - X + W^0_{LW} + W^1_{LW} + Y < W^0_{HW} + W^1_{HW} + W^1_{LW}$$

This means that both GS1 and GS2 are efficient GSs, but only GS2 is the wealth-maximizing one. Thus, when the marginal utility of wealth varies as the initial wealth, the efficient GS is not necessarily the wealth-maximising one. In determining the relative desirability of alternative GSs, the utility function of the trading parties should be taken into account. And this will make the analysis much more complicated. Assuming away the wealth effect can greatly simplify the analysis of alternative GSs by comparing their ability to create wealth.

The wealth effect often takes the form of financial constraints. The more stringent is the financial status, the higher is the marginal utility of a loss. A standard assumption taken in organisational studies is that the constituent parties have similar financial constraints. For example, in analysing the selection of hierarchy and market, the
objects of analysis are two firms, so the assumption of no wealth effect is more likely to stand than the case of analysing the employment relationship where employer and employee obviously face different financing conditions. In this study, the key concern is the influence of the interaction between the client, the designer and the contractor on the efficiency of procurement systems. Perhaps, generally speaking, the client is relatively easier to get external funding than the contractor and the designer. However, if we consider the fact that the client bears much heavier financial commitment in the construction process, the different level of financial constraints facing each party in a project will not be a serious problem. As a result, the wealth effect is assumed way in the following analysis.

6.4 Operationalisation of construction transaction(I): revenue

The value of a project depends on its final use. For the commercial building project, such as office and speculative housing, the ownership of the buildings is divided into small units for sale or rent. For the industrial building project, the revenue stream is generated from the operation of the project. Regardless of the source of revenue, the value of the project can be evaluated in terms of discounted revenue. For a project with the delivery date of \( t_a \) and the economic life of \( n \), the net present value (NPV) of the net revenue (NR) of the project can be expressed as

\[
NR(t_a) = \sum_{t=1}^{n} \frac{R_t - C^o_t - C^l_t - C^f_t}{(1+r)^{t_a+t}} + \frac{RV}{(1+r)^{t_a+n}} \quad \text{Eq. 6-19}
\]

where \( R_t \) is the revenue stream directly generated from the completed project, \( C^o_t \) is the operating costs of the project, \( C^l_t \) is the cost of purchasing land, \( C^f_t \) is the cost of financing for construction and for purchasing land (so \( R_t C^o_t \) is the expected net income from the completed project at the \( i^{th} \) period), \( RV \) is residual value of the project after \( n \) period and \( r \) is the discount rate.

This formula is an exact form for the evaluation of revenue from the project, from which it is evident that the timing of the project delivery is a determinant of NPV. We want to explore what kind of transaction attributes will make this differ between procurement routes. The answer lies in the client’s time value, which can be measured
by the average opportunity cost of time (OPPCT) between two delivery dates. For ease of notation, we will simplify Eq. 6-19 as a linear function,

\[ NR(t_a) = -a \times v + b \]

where \( a \) and \( b \) are constant, standing for the slope and intersection of the line. In Fig.6-5, the bold line indicates the meaningful range of delivery time. \( t_{ae} \) and \( t_{al} \) are the earliest and latest delivery date. The slope is average opportunity cost of time (of the project).

In principle, the larger is OPPCT, the more important it is to get the project delivered earlier. This also implies that under this condition, the fast-track strategy will appear to be more desirable. As a result, it is a factor responsible for differential efficacy between procurement routes with respect to discounted (net) revenue.

![Fig. 6-5 A graphical illustration of the client's time value](image)

6.5 Operationalisation of construction transaction (II): transaction cost

From this section on, we will implement Williamson’s three-stage transaction cost reasoning shown in Fig.4-1 to the context of construction procurement. Regarding the first stage, the assumption of bounded rationality and opportunism is taken, so we start the analysis by exploring the important items of transaction costs with comparative significance and move on to the issue of discovering significant construction transaction attributes by an illustrative model.
6.5.1 Identification of transaction costs in construction

6.5.1.1 Itemisation of transaction costs in construction

Which items of transaction costs in construction have comparative significance between different procurement routes? We can approach the answer by putting together all types of transaction costs that have been pinpointed in the literature, then proceeding to comparative analysis.

Gruneberg and Ive (2000) proposes a list of the important items of transaction costs in construction, including:

1. Search costs of finding out information about who is offering what products or services and at what prices.
2. Product or service specification costs
3. Contract selection, contract design and negotiation costs
4. Supplier selection costs
5. Contract performance monitoring costs
6. Contract enforcement costs of legal bills and delays

Compared with Dahlman (1979) and Williamson (1985), this list is by and large based on the former without much attention to the latter except for the last point. Similarly, Ismail (1997) also asserts that the quantifiable items of transaction costs in construction include tendering costs, simultaneous management costs and costs of dispute resolution. After surveying these costs occurring in three procurement routes - traditional system, integrated design-build and integrated design-build with subsidiary contractor - a conclusion is made that “transaction costs can be quantified in procurement routes” [p.39]. In this section, we claim that it is likely to prove futile to attempt to provide explanatory foundation for procurement behaviour by quantitatively measuring each important item of transaction costs in the construction process, since a majority of costs with comparative importance are fairly difficult, if not impossible, to estimate.

---

5 This cost allegedly comes from the overlapped responsibility of in-house project manager, architect and quantity surveyor in monitoring the process.
Table 6-2 A comparison of measurement cost in three procurement systems

<table>
<thead>
<tr>
<th></th>
<th>The client</th>
<th>The contractor and designer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traditional method</strong></td>
<td>$\text{TC}^\text{TM}(1)$: tendering costs for finding designer</td>
<td>$\text{TC}^\text{TM}(1)$: tendering costs for finding subcontractors</td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{TM}(2)$: measurement cost for assuring the performance of designer</td>
<td>$\text{TC}^\text{TM}(2)$: measurement cost for monitoring the performance of subcontractors</td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{TM}(3)$: tendering costs for finding main contractor</td>
<td>$\text{TC}^\text{TM}(3)$: main contractor's tendering costs for bidding for the project</td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{TM}(4)$: measurement cost for assuring the performance of main contractor</td>
<td>$\text{TC}^\text{TM}(4)$: designer's tendering costs for bidding for the project</td>
</tr>
<tr>
<td><strong>Design and Build</strong></td>
<td>$\text{TC}^\text{DB}(1)$: tendering costs for finding main contractor</td>
<td>$\text{TC}^\text{DB}(1)$: tendering costs for finding subcontractors</td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{DB}(2)$: measurement cost for monitoring the performance of main contractor</td>
<td>$\text{TC}^\text{DB}(2)$: measurement cost for monitoring the performance of subcontractors</td>
</tr>
<tr>
<td><strong>Management System</strong></td>
<td>$\text{TC}^\text{MS}(1)$: tendering costs for finding designer</td>
<td>$\text{TC}^\text{MS}(1)$: construction manager's tendering costs for bidding for the project</td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{MS}(2)$: measurement cost for assuring the performance of designer</td>
<td>$\text{TC}^\text{MS}(2)$: designer's tendering costs for bidding for the project</td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{MS}(3)$: tendering costs for finding construction manager</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{MS}(4)$: measurement cost for assuring the performance of construction manager</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\text{TC}^\text{MS}(5)$: construction manager's fees covering tendering costs for finding qualified package contractors and monitoring their performance</td>
<td></td>
</tr>
</tbody>
</table>

With reference to Table 4-1, we can attempt to build up a similar table, Table 6-2, to compare the significance of different items of transaction costs. From the standing point of the client, construction management carries the highest direct transaction costs, next is traditional system, and the lowest is integrated system. However, from the perspective of the whole system, including the client and his/her retained agents, these may be no noticeable difference in the sum of type-I transaction costs. For the client, the change in identity of the bearer of transaction costs doesn't mean they disappear. Much depends on whether the principle of cost transferability applies. Maybe, sometimes in some special cases, it will not apply. But, peculiarities should not be mixed up with the economic principle. To reveal the potential transaction

---

6 Assume that designers are all directly employed by the main contractor.

7 It is to be noted that our focus here is only on type-I transaction costs. Discussions about type-II ones are left to the next subsection. (refer to definitions of these types in section 3.3)
hazards, we propose that the principle of cost transferability should be considered. Keeping this principle in mind will help us undertake the following analysis.

Let’s go back to Ismail (1997) to examine the comparative importance of the quantifiable costs. First of all, according to observations, the client’s direct costs of tendering for a contractor in the three procurement routes - $ TC^{C_TM(3)}$, $ TC^{C_{DB(1)}}$, and $ TC^{C_{MS(3)}}$ - are different. $ TC^{C_{DB(1)}}$ tends to be higher than the other two. However, the client’s tendering costs actually encompass $ TC^{C_{TM(3)}}, TC^{C_{DB(1)}}$, and $ TC^{C_{MS(3)}}$. Taking $ TC^{C_{TM(1)}}$ and $ TC^{C_{MS(1)}}$ into account will greatly weaken the relative disadvantage of design-build in tendering costs.

Second, regarding simultaneous management cost, Ismail claims that the quantity surveyor in the traditional system is an extra monitor of construction process relative to design-build. This seems to imply that the fee paid to quantity surveyor is an additional transaction cost to the client if traditional system is used. What’s the problem here? It is a fact that design-build can provide the client with an *ex ante* commitment to cost ceiling. Nonetheless, the tasks of assuring value for money are still to be done. Without the aid of a quantity surveyor tracking the cost record, the client needs to appoint a project monitor periodically to review the construction plan submitted by the main contractor and keep an eye on the quality of what has been done. Accordingly, the mission of cost monitoring is just being transformed into the form of intensified inspection of quality. How wide would this difference in transaction cost be? It is not as large as it first appears.

Last, there is no estimate of costs of dispute resolution shown in Ismail (1997) because no interviewed clients have experienced arbitration. But it does not necessarily follow that there is no cost incurred due to opportunism. The loss from opportunism may take the more subtle form of concession to hold-up demand (and thus cost overrun), delay, and extra costs of intensifying quality inspection. Unfortunately, these costs are very hard to quantify. We need to find other way out.

Sometimes the fees for construction manager are deemed as the peculiar item of transaction costs in using construction management. As emphasised in the principle of
cost transferability, the cost of co-ordinating subcontractors or trade contractors appears in all routes. In the traditional system and design-build, the cost is included in the awarded contract price; in the construction management method, it is paid as the agent’s fees. They may not be absolutely equal, but their difference in comparison with the costs attendant with the second-round information problems is not paramount.

As a result, from the above discussion, it seems safe to say that the differences of TC\textsubscript{1} across procurement routes are not significant. But what are the relevant items of transaction costs? This question will be answered in the following subsection.

6.5.1.2 What are the relevant items of transaction costs to the selection of procurement routes?

The previous subsection demonstrates that measurement and search costs are of little comparative significance in comparing construction procurement routes. This finding is understandable. To complete a transaction, there is information that you need to know, but you don't know. The costs of obtaining this kind of information won't be much different across governance structures. What may be different is the bearers of costs. In contrast, Williamson places slight attention on these can-be-expected type-I transaction costs, but concentrates on unexpected type-II ones. These costs are the outcome of interaction between an uncertain future and possible strategic reactions of other parties. When the contingency falls outside of the stipulation of the contract clauses, renegotiating a term is inevitable. For example, when ground conditions force the original design to be modified, the client is vulnerable to be overcharged for the additional or modified works resulting from requirement changes. If the client is determined to defy the unfair demand, a bout of haggling occurs with consequent costs of delay to the client. This story may end in the client’s yielding to taking the extra costs, or appealing to third-party arbitration or, even worse, litigation. How painful is the process is conditional on how efficient is the procurement route in dealing with the re-distribution of ex post quasi-rent. Because the adaptability of every procurement route to unanticipated events is distinct (see Sec.6.6), their differentials in the magnitude of the type-II transaction costs must be significant. Compared to TC\textsubscript{1}, the possible losses caused by delay, dispute resolution and
litigation may be very huge. This is why we claim type-I costs should be relegated to
the minor role. It bears emphasising that asset specificity is not the only important
factor, in that measurement problems associated with opportunism and rent-seeking
behaviour cannot be assumed to be constant under different procurement routes. A
complete explanation of this difference will be put forward in Sec.6.6. Suffice it to
say here that type-II transaction costs in construction are not only affected by the
characteristics of the project, but also by the way designer and contractor are
organised. For example, the design-build route is more plastic\(^8\) than the other two in
the sense that the main contractor has a wide range of legitimate decisions within
which he/she can choose. As a consequence, a complicated project executed by way
of this route may consign to the client severe quality control problems. This may
make the rent-seeking problem even worse. As a result, we take the same position as
Williamson that TC\(_{II}\) will be dominant over TC\(_I\) and the focus of analysis should be
placed on how TC\(_{II}\) will vary between different procurement routes under different
conditions.

6.5.2 The basic reasoning of operationalising transaction

The key point of making TCE refutable lies in the operationalisation of transaction.
Most empirical works just rely upon those factors pointed out in TCE as the
explanatory foundation for the organisational issues of interest, while, for the case
failing to meet the basis propositions of TCE, like construction, it is imperative to
re-examine the way of discovering the factors responsible for transaction cost
differences between procurement systems.

As introduced in Chapter 4, Williamson's comparative institutional analysis is made
up of three parts: operationalisation of transaction, dimensionalisation of GSs and
alignment of transaction attributes with GSs (Williamson, 1996). This three-step
scheme indicates a complete procedure for implementing transaction cost reasoning,
while it is not explicit about how to achieve each step. This is not a problem for the
case where the basic premises of TCE, like continuous production and exogenous

\(^8\) The definition of plasticity is taken from Alchian and Woodward (1987). More details please refer to
Sec.6.6.
design, can be met. However, as explained in Chapter 5, construction procurement route selection is not such a case. Thus, unlike some TCE empirical works which can directly use the key variables, such as different types of asset specificity, pointed out as theoretically relevant, we have to investigate the process of operationalisation.

Behind Williamson’s analytical framework, in fact, lies an rent-based reasoning set out by Klein, Crawford and Alchian (1978). The basic assumption is that the transaction costs derived from the process of tussling for redistribution of quasi-rent is a function of approducible quasi-rent (AQR). It follows that the larger is AQR, the more lucrative the opportunities to exploit the other party’s vulnerability and the more resources will be incurred in the non-productive activities. With this basic principle, Williamson’s three-step scheme can be reformulated in terms of AQR. First of all, we have to discover the conditions under which the magnitude of AQR is different; second, explore why AQR can be affected under different organizational setting; third, align GSs with transaction attributes so as to efficiently govern the transaction hazards arising from AQR. The purpose of this reformulation is to show that AQR-based argument provides a helpful guidance on carrying out operationalisation in dealing with the new case that hasn’t been discussed in TCE. In the following two sections, we will investigate the two main sources of AQR in the literature [Williamsons, 1985; Alchian and Demsetz, 1972].

6.5.3 A transaction cost theory of exchange

In this section, we try to illuminate two key concepts of this chapter, quality rent and quasi-rent, by using a simple model. For ease of discussion, firstly we will focus on the ex post measurement problem by putting the hold-up problem aside for a while. We then shift the focus to hold-up problem by holding measurement problem constant. Finally, an attempt is made to explore the possible consequences of the interaction between these two problems.

6.5.3.1 Quality rent: definition

In order to give quality rent a clearer definition, we will set up a simple model in which a buyer purchases a type of intermediate goods from a large-number competition market. The development of decision scenario is shown in Fig.6-6. At t=1,
the buyer determines the optimal quantity $q^*$ for given ex ante specified quality standard $s (=1)$ by reference to the market equilibrium price $p$. Then, the buyer contracts with one of the suppliers at $t = 2$. To ensure the delivered quality, the buyer has to determine the level of quality inspection at $t = 3$.

The buyer chooses a surplus-maximising quantity $q^*$ for given ex ante specified quality standard $s (=1)$.

The buyer goes into the contractual relationship with the seller.

The buyer determines the optimal control level of quality $s^*$.

The buyer is a monopsonist who can freely choose a surplus-maximising quantity $q^*$ for given ex ante specified quality $s (=1)$. For ease of discussion, $q^*$ is set to be one. The condition inducing the buyer to engage this transaction comes from the excess of his valuation on the transacted goods over their purchasing prices. Let the buyer’s valuation function $V$ varies as quality index, $s$, $0 < s < 1$, so we can express the buyer’s rent as

$$R_b(s) = V(s) - p \geq 0$$

Likewise, if the producer’s average production cost $AC$ varies as quality produced, we can express the producer’s rent as

$$R_p(s) = p - AC(s) \geq 0 \quad \text{Eq. 6-20}$$

To find more specific results, we have to make some assumptions that $R'(s) > 0$, $R''(s) < 0$, $AC'(s) > 0$, $AC''(s) > 0$. In this setting, total rent is $V(s) - AC(s)$, varying as quality level. Since price is usually determined according to a given quality requirement, we can interpret quality index as the extent of this requirement being fulfilled. As shown in Fig. 6-7, for the buyer, the minimum acceptable level of quality is $s_1$, where the buyer’s rent becomes zero. Along with the quality received changing from $s = 1$ to $s = s_1$, the buyer’s ex post rent is shrinking from $V(1) - p$ to 0. At the same time,
producer’s rent increases from minimum $p \cdot AC(l)$ (equal to zero since ex ante large-number competition between sellers is assumed) to maximum $p \cdot AC(s_i)$. In principle, the buyer can be aware of this situation and take some measures to ensure the quality of the delivered goods as promised. On the contrary, the producer has an incentive to reduce the quality standard so as to enlarge his rent ex post. This misalignment of incentive in quality is a potential seat of transaction problems.

Fig. 6-7 A graphic illustration of buyer’s and producer’s rent

Next, we introduce the costs of quality inspection, $C_i(s)$ ($C_i' > 0$ and $C_i'' > 0$), indicating the minimum costs that the buyer needs to bear for assuring the quality standard of $s$. It is assumed to be a concave function of $s$, namely that improvement from high quality costs more than that from low quality. Where opportunism and complexity are absent, this curve is almost a horizontal line with much smaller positive intercept with x-axis (the implications is that checking small samples suffices safely to infer the quality of the whole population). Once we verify the small part of the product is fine, the rest of them can be expected to be fine (except for random variations) in that opportunism and complexity are excluded. This belongs to pure cost of cognition in Matthews’s sense [Matthew, 1986]. What is important here is the quality gap between

---

9 Barzel(2000) makes the similar assertion that “[w]ere the specimens of each commodity strictly uniform, full information would be obtained by counting the number of specimens and measuring just one.” (p.9)
the transmission of the buyer’s quality requirement and the delivered quality of the final product.

After adding $C_i(s)$ into framework, we can find that the buyer’s optimal level of quality is determined by the marginal condition of $R'(s)=C'_i(s)$ (see Fig. 6-8). Thus setting the intensity of quality inspection for assuring the delivery of $s_2$ is the optimal choice to the buyer. When the seller takes costs of quality inspection into account, $AC(1)-AC(s_2)$, is an **appropriable quality rent** ($r_q$). However, in practice if $s_2$ is unknown to the seller and the buyer’s controlled level of $s$ exceeds the seller’s provided level, the seller runs the risk of being detected. Thus, to the seller, quality rent would be a probability distribution. For simplicity, we can assume the expected value of quality rent is $AC(1)-AC(s_2)$. Whenever the magnitude of this rent is not trivial, the producer’s motive for rent-seeking can be prompted by, for example, providing inferior quality goods.

**Fig. 6-8 A graphic illustration of buyer’s and producer’s rent**

Since it is costly to evaluate the true value of the goods with ambiguous bundle of attributes, $C_i(s)$ will increase much faster as the required quality is approaching the perfection level ($s=1$). Along with $C_i(s)$ shifting up to $C^d_i(s)$ ($C^d_i(s)$ denotes the costs of inspection with complexity), the optimal quality level will become $s_3$, increasing the quality rent to $AC(1)-AC(s_2)$ (see Fig.6-9). As the quality rent is enlarging, the seller’s incentive of taking advantage of asymmetric information on the product quality (i.e. providing lower-than-promised quality) will intensify, likely to result in
the dissipation of rent. According to Barzel(1982), this is an explanation of why integration occurs. However, this is only half of the story. Another part lies in the consideration of correction costs in the event that opportunistic intent gets caught. It seems sensible to speculate that the producer will try to reap the quality rent only when the expected value of quality rent exceeds that of correction costs. In fact, what are considered as remedies for quality problems in Barzel(1982), like brand name, repeated purchase, all can give rise to the effect of increasing the (discounted) correction costs (in the broader sense) in that today’s record of inferior quality will occasion tomorrow’s distrust from the buyer, resulting in more costs for saving the bad reputation. As a result, in some measure the costs of correction would considerably reduce the net quality rent (quality rent minus correction costs), thereby thwarting the motive of opportunism. It seems safe to contend that measurement problems matter in the interpretation of organisation selection only when the net quality rent instead of just quality rent appears substantial.

Fig. 6-9 A graphic illustration of buyer’s and producer’s rent

\[ R(s) \]
\[ R(s) - C_f(s) \]
\[ AC(s) - AC(t) \]
\[ C_f(s) \]
\[ C_f'(s) \]

- \( C_f(s) \): costs of inspecting in absence of complexity
- \( C_f'(s) \): costs of inspecting with complexity

6.5.3.2 Quasi rent
The quasi-rent-based argument, developed along the line of Williamson (1971) and Klein et al. (1978), commits itself to relating the presence of vulnerability to
contractual remedies. The reduced ex post opportunity cost of the irreversible investment is claimed to be disadvantageous to the producer. When the specific investment is reduced from $I$ to $I_1$, $AC(s)$ will be replaced by $AC_1(s)$ (see Fig.6-10). This movement of the curve doesn’t affect the buyer’s optimal quality inspection level, but the distribution of rent. The producer’s advantage resulting from quality rent $AC(1)-AC(s_2)$ will be offset to some extent by quasi rent, $AC_1(s)-AC(s)$. Net balance of quality rent and quasi rent will determine which party is more vulnerable.

![Fig. 6-10 The effect of idiosyncratic lump-sum investment](image)

6.5.3.3 The interaction between quality rent and design rent

Here we have an intriguing question to address: when two types of rents are balanced out, whether will the rent-seeking behaviour be contained? The answer is “not entirely”. In a system where counterbalancing forces are acting at the same time, the equilibrium can possibly be maintained at the outset, but it is not necessarily stable. In an open system, the equilibrium will be readily disturbed as external factors intervene. The concept of incomplete contracting is characterised by the possibility that contract may drift out of incentive alignment owing to unforeseen contingencies. These triggering events would stretch the tension of unstable equilibrium beyond the critical point since the party with net rent advantage (henceforth NRA party) may find rent seeking strategies lucrative due to higher probability of success under certain circumstances.

However, it is to be noted that in essence quality rent differs from quasi rent in how to
detect the occurrence of rent being appropriated. The producer’s intention to snap up quality rent cannot be known unless inspection efforts are made (so costs are incurred). In contrast, the buyer’s quasi rent seeking strategies often take the form of ex post renegotiation for terms disadvantageous to the vulnerable party. This distinction would help explain their interaction.

Let’s consider a case where buyer and producer are considering the choice of opportunistic or non-opportunistic strategy. In the following analysis, we take the standard game theory assumption of common knowledge, which means that if information is known to all player, each player knows that all of them know it, each of the them knows that all of them know that all of them know it, and so forth [Rasmusen, 1989]. Assume that both parties are able to estimate the magnitude of quasi rent and quality rent ex ante, thereby calculating the payoff matrix of each pair of strategies\(^\text{10}\). The buyer is a NRA party with expected appropriable quasi rent of 4 units and the effective level of inspection cost is 1 unit given the producer acts non-opportunistically. The producer is assumed to have 2 units of (net) quality rent to his advantage. On the basis of the above basic information, without considering the rippling effects of these strategies (i.e. TC\(_\text{II}\)), the payoff tree is described in Fig.6-11.

**Fig. 6-11** A payoff tree

<table>
<thead>
<tr>
<th>Buyer decides whether to appropriate quasi-rent</th>
<th>Seller decides whether to appropriate quality-rent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(buyer’s payoff, seller’s payoff)

- (4-1-2, -4+2) = (1,-2)
- (4-1, -4) = (3,-4)
- (-2-1, 2) = (-3,2)
- (-1, 0) = (-1,0)

The payoff structure in Fig.6-11 can be summarised in Table 6-3. For instance, in the

\(^\text{10}\) It is to be noted that we consider the payoff of these strategies in isolation, namely that the payoff refers to the extra revenue coming from this strategy.
right upper cell, the buyer takes an opportunistic strategy with expectation to receive 4 units of quasi rent subject to non-opportunistic response from the producer, so the net return is 3 units (net of quasi rent (4) and inspection costs (1)). For the producer, 4 units of quasi rent being exploited are equivalent to 4 units of loss ex post. The rest of cells are computed in the same way.

<table>
<thead>
<tr>
<th></th>
<th>Producer’s strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Buyer’s strategy</strong></td>
<td>Opportunistic</td>
</tr>
<tr>
<td>Opportunistic</td>
<td>1,-2</td>
</tr>
<tr>
<td>Non-opportunistic</td>
<td>-3,2</td>
</tr>
</tbody>
</table>

Given this payoff structure, the strategy mix of (opportunistic, opportunistic) will be dominant. In the face of this situation, precautionary actions are much more likely to be adopted than collaboration due to the nature of information asymmetry. In other words, the buyer’s goodwill to give up the opportunity to take advantage of producer’s weak positions will not necessarily be appreciated by producer, and vice versa. It follows that, for fear of the producer’s adoption of opportunistic strategies such as quality downgradation due to uncertainty as to whether the producer will exploit quasi rent, the buyer may impose more stringent quality inspection above the optimal level (because the optimal level is unknown in practice), leading to more costs incurred but with less effectiveness. As a result, generally speaking, quality inspection problem will deteriorate the severity of quasi-rent-seeking behaviour.

6.5.3.4 A summary of Sec.6.5.3
In Sec.6.5.3.1~ Sec.6.5.3.3, we intend to demonstrate how the possibility of rent redistribution is prompted. From the doctrine of governance branch and measurement branch of transaction cost theory, rent is conceivably generated from two sources: (1) quasi rent exists due to the non-redeployability of the durable investment; (2) quality rent is generated from the provision of inferior quality by saving production costs.
Certainly, they are interrelated. The possibilities of appropriating quasi rent are on the buyer’s side and can be expropriated by forcing the producer’s to lower the price. Every time the buyer succeeds in reducing the procurement price, the seller’s rent begins being nibbled.

Quality rents arise from the fact that the complete attributes of the goods are ambiguous and can be reaped regardless of whether asset specificity is present. However, its significance should be assessed by the magnitude of net quality rent rather than quality rent per se.

### 6.5.4 Process specificity in construction

#### 6.5.4.1 Types of asset specificity in the literature

In TCE, asset specificity, a term used to describe the degree of interdependence between transactors, can be measured by the decrease in value if the original transaction is not carried on. The party who suffers higher losses due to the disruption of the transaction will be more vulnerable to hold up from other parties. As explained in Williamson (1983), lock-in effect may be caused by different types of asset specificity:

1. **site specificity:** A “cheek-by-jowl” relationship between the buyer and seller is often based on the consideration of saving transportation and inventory costs. An example is the construction of a steel plant near an ore mine. When the factory starts operating, resorting to other mine supply will become much more costly and this will subject the plant to unreasonable over charges from the mine owner.

2. **physical asset specificity:** For the case where a transaction involves 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 a transaction entailing an investment in relationship-specific human capital through the learning-by-doing process, this investment will be of no value when the personnel with special expertise are transferred to other transactions.

4. **dedicated asset specificity:** When an investment is made for tapering economies of scale, the premature cancellation of orders can leave the producer with excess
According to Winch (1989), none of them is relevant to construction. However an item newly added in Williamson (1996), temporal specificity, infuses new blood to this concept. By way of examining the distinctive features of construction operations, Masten et al. (1991) found that 'delays become a potentially effective strategy for exacting price concessions' when 'the timely performance [of material suppliers] is critical'. Put another way, the presence of hold-up problems doesn't necessarily involve the specificity of the durable investment. 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 have an advantageous position to increase the price to seize the rents generated by temporal specificity. Pirrong (1993) applies a similar concept to the analysis of the bulk shipping industry. He claims that 'time and space factors in shipping markets may create "temporal specificity" that encourage costly haggling between shippers and carriers over quasi rents if they rely on spot markets.' Compared with Masten et al. (1991), he finds that “space” factors also lead to temporal specificity. In his case, only a vessel with special apparatus can transport some commodities. If there is only one suitable ship in the neighbouring harbour, the sender may be subject to temporal hold-up problem, the severity of which depends on the transportation costs and time delay in getting another vessel available in the same harbour. Hubbard (1999) conducts a quantitative test of the relation between temporal specificity and contractual arrangement in the trucking industry, concluding 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 finding out other possible sources of asset specificity.

6.5.4.2 A new type of asset specificity: process specificity

The process of a construction project is a kind of production flow. Once it commences, keeping the flow running without stop until the end will be beneficial to all parties involved since the realisation of their expected economic rent depends on whether the project unfolds as expected. In the event that the continuity is disrupted, all parties will incur costs of delay or even worse, costs of switching to other projects. On the
client's side, this cost can be defined as the additional cost for restoring the production flow. More exactly, the magnitude of process specificity can be expressed as:

\[ PS = v(t_1 + t_2) + C_e + C_{rp} + C_d \]  

Eq. 6-21

where (1) \( t_1 \) and \( t_2 \) stand for the time for finding a replacement agent and the time for the replacement agent to carry on the unfinished project. \( v \) is the average cost of time. Thus, the first term indicates the potential loss arising from the opportunity costs of time, were the project to be disrupted. \( C_e \) is the extra costs due to repeated set-up on the construction site\(^{11}\). \( C_{rp} \) is used to indicate the risk premium asked by the replacement contractor for uncertain quality of work done by the first contractor. The last term, \( C_d \) is indicative of the cost incurred due to difficulties in identifying liabilities between original and replacement contractor, were defects found in the interface between them.

In Eq.6-21, \( v \) is an attribute of the project (see Sec.6.4). Other elements, including \( t_1 \), \( t_2 \), \( C_{rp} \), \( C_d \) and \( C_e \) change as both design and/or technology and procurement systems used in the project.

To distinguish from the temporal specificity arising from lack of alternative choices ex ante due to time or space factors, the specificity resulting from the cost of disruption can be called **process specificity** to indicate the interdependent relation between parties in the construction process. On the side of contractor and designer, the personnel devoted to this project will be redundant if this project stops and they cannot be transferred to other projects at once. As a result, during the construction process, the client and his/her agents are in the condition of bilateral monopoly. But the degree of exposure to hold up is not symmetrical, so the exploitation of other party's weakness is likely to occur\(^{12}\). As explained in Sec.6.2.2, the value reduction of the partly-completed project in the course of construction makes the client face a weaker position than his/her agents. Hence, in the analysis of procurement route selection, we will focus on the problems of process specificity on the client's side.

---

11 This cost includes costs for installing construction machinery, administrative facilities (say temporary office), management hardware (say computers), and so on.
12 Otherwise, the bilateral monopoly would be in the balance of tension since nobody will get the net gain under the condition of symmetrical degree of specificity.
6.5.4.3 A Comparison of process specificity with other types of asset specificity

In this subsection, the concern will be whether process specificity is intrinsically different from those identified in the TCE literature and only applied to construction.

Seven types of asset specificity are put under scrutiny in Table 6-4, from which the sources of quasi rent in fact can be understood in terms of the option-like model introduced in Sec.6.2.2.2. Put explicitly, the vulnerable party has two options in response to hold-up demands: one is to give up the original investment and switch the invested resources to other uses with the loss \( L_1 \); another is to find a replacement trading partner to complete the original transaction with the loss \( L_2 \). For a rational decision maker, the option with lower loss will be chosen. The magnitude of this loss can be defined as a measure of degree of asset specificity (DAS). That is,

\[
DAS = \min \{ L_1, L_2 \} \quad \text{Eq. 6-22}
\]

When seven types of asset specificity are classified according to Eq. 6-22, two types of asset specificity can be identified:

**Type I: \( L_1 > L_2 \)**

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

**Type II: \( L_1 < L_2 \)**

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 those resulting from finding an alternative trading partner \( L_2 \) because there is no other parties who want to employ the resources in the same way and thus \( L_2 \) can be considered as infinity. For example, in the case of General Motor versus Fisher Body, Fisher Body needs to make an investment on dies and stamping machines specific to GM’s requirements (i.e. physical asset specificity), so no other car manufacturer will buy this type of car body. Another example is human asset specificity. For a marketing company, if the on-job training of its salesmen is designed only for the products this company is promoting, this knowledge will be of
no value for its salesmen in undertaking other jobs and the knowledge of the new salesman hired from outside is also of little value to this company. For these two types of asset specificity, the termination of transaction means only the salvage value of the sunk investment can be recovered.

Table 6-4 Sources of quasi rent arising from seven types of asset specificity

<table>
<thead>
<tr>
<th>Types of asset specificity</th>
<th>Example</th>
<th>Vulnerable party</th>
<th>Sources of quasi rent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical asset specificity</td>
<td>Car assembler and car body-making company</td>
<td>Car body-making company</td>
<td>Value depreciation caused by switching the specific investment, including stamping machines and dies, to other uses</td>
</tr>
<tr>
<td>Site asset specificity</td>
<td>Ore mining company and steel plant</td>
<td>Steel plant</td>
<td>Increased costs in transportation and inventory when switching to alternative suppliers</td>
</tr>
<tr>
<td>Dedicated asset specificity</td>
<td></td>
<td></td>
<td>Costs of maintaining excess capacity when the original transaction doesn’t carry on</td>
</tr>
<tr>
<td>Human asset specificity</td>
<td>Salesperson and marketing company</td>
<td>Marketing company</td>
<td>Costs of training salespersons with company-specific knowledge if the experienced salespersons leave</td>
</tr>
<tr>
<td>Temporal specificity – timing</td>
<td>Contractor and material suppliers</td>
<td>Contractor</td>
<td>Opportunity costs of delay caused by postponed delivery of materials</td>
</tr>
<tr>
<td>Temporal specificity – space</td>
<td>Cargo sender and cargo carrier</td>
<td>Cargo sender</td>
<td>Opportunity cost of delay caused by finding a replacement cargo carrier in the neighbouring harbour</td>
</tr>
<tr>
<td>Process specificity</td>
<td>Construction client and contractor</td>
<td>Contractor</td>
<td>Increased costs of switching to a replacement contractor for resuming the disrupted project</td>
</tr>
</tbody>
</table>

Now we are wondering 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 the client and the designer/contractor. The losses resulting from resuming a disrupted project is an ultimate source of process specificity. The costs of switching are a most general term for these kind of losses. In fact, all types of asset specificity listed in Table 6-4 occur because of this reason. The difference between them is what reasons cause these losses. For the type-II asset specificity, the value depreciation of the sunk investment makes the carrying-on of the transaction critical. For the type-I asset specificity, the relevant option is to find a replacement trading partner to consummate the transaction because it is relatively more costly to give up the original use of the sunk investment, i.e., $L_1 > L_2$. But in going further to explore why $L_1$ is so high for the type II asset specificity, various reasons can be identified:

1. Sunk physical investment: For site specificity and dedicated asset specificity, the consideration of recovering the sunk lump-sum physical investment makes it worthwhile to bear extra costs (e.g., transportation cost and inventory cost 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 one demands too much extra charge, because the ongoing transaction in which the held-up transaction is embedded is much 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 value of the cargo. For a contractor, the value of the whole construction project normally greatly outweighs that of an individual building material. As a result, being embedded in a large transaction is the ultimate reason for making $L_1 > L_2$ in the case of temporal specificity.

3. Built on a fixed site: In the case of process specificity, the client will choose to resort to an alternative contractor, were the project disrupted, because scraping the project that has been partly fixed on site is much more costly.

This comparison clearly demonstrates that the origin of process specificity is intrinsically different from other types of specificity because the product is fixed on the land provided by the client. The interim payment to the contractor for this part of
the project amounts to an irreversible cost to the client. This is the ultimate reason why the client is in a vulnerable position.

6.5.5 Quality rent and design rent in construction

According to the measurement branch of transaction costs theory, the measurement problems are varying as the characteristics of the goods or service. Their focus is on the difficulties in assuring value for money. In this case, the costs of inspection are the main concerns. To discuss this problem in the context of construction, an effort is made here to extend its scope to cover three aspects: (1) inspection of design quality; (2) inspection of construction quality (conformance quality or fitness for purpose) and (3) costs of making project defects corrected. Distinct from the analysis of manufacturing firms, the feature of this distinction lies in the inclusion of design factors and costs of defect correction. Let's discuss them in turn.

Leaving design out of the framework is a regret of transaction cost economics [Williamson1988]. Construction is a right case to take it into consideration. Having design factors in mind, we want to explore furthermore: is there any effect of design on the creation of rent? The answer depends on whether designer and contractor is the same agent. When design and construction are undertaken by different agents, the responsibility of the contractor is to produce a project in accordance with design and specification controlled by the client. In this case, standard quality rent arises from contractor’s intention to downgrade the required quality. Nonetheless, when in some measure the producer has the right to determine the basic design of the project, the rent can appear by changing the way the project will be built. This point can be illustrated by Fig. 6-12. Let $D_1$ be the original design that the client expects ex ante, and $D_2$ be the design plan the contractor submits. For simplicity, it is assumed that (1) different design doesn’t alter the cost of monitoring, $C_i'(s)$ and (2) the average cost can be reduced if $D_2$ is used, that is $AC(s;D_1)$ shifting down to $AC(s;D_2)$. At this time, two types of rents arise from measurement problems. The first part, $AC(1;D_1)-AC(1;D_2)$, comes from the change of design irrelevant to controlled degree of quality ($s=1$), while the second part, $AC(1;D_2)-AC(1;D_1)$ results from lower quality standard for given design $D_2$. To distinguish it from quality rent, the former is called design rent. Where both rents are present simultaneously, the severity of rent-seeking
behaviour will be ameliorated.

Fig. 6-12 Quality rent and design rent in construction

Another distinctive aspect of measurement costs considered here are the costs of correcting defects. Why should this cost ignored in case of manufacturing industry be elevated to such a prominent place? The reason is simple. In the general commercial transactions, the sellers bear the responsibility to assure the quality they provide and correct defects according to contract or law. However, in the construction process, there is something different. As the description in Eccles(1981), ‘[a]t any point in time a number of these specialities will be simultaneously involved on the project and often the work of one cannot proceed until a phase of work has been completed by several others’. The production with this characteristic is likely to pass on the error of the previous stage to the next one, making the identification of liabilities more difficult. According to the current practice, most trade contractors (or subcontractors) are autonomous entities. If the responsibilities cannot be clearly identified, it is impossible to claim back the loss. How costly it will be is one of our concerns. Another concern is related with subcontractors’ ability to bear the compensation for the loss that they cause. Generally speaking, the capitalisation of subcontractors is much smaller than that of main contractors so that their capability of absorbing risks would be relatively limited, implying it is more likely for the client to encounter the situation where the loss cannot be reimbursed. These two considerations generate another source of quality rent. Following the same logic in Sec.6.5.3.1, if it is very
costly to identify the bearer of responsibility and get the loss compensated, the client will not necessarily intend to control quality to no-defect standard, whereby quality rent appears. It is to be noted that quality rent and design rent have the common origin in measurement problems, whereas their effects on rent-seeking behaviour are different, which is the subject the next subsection will turn to.

6.5.6 The interaction effect of quasi rent, quality rent and design rent

In contrast to the case of vertical integration where opportunity to appropriate quasi rent and quality rent is in the hands of the buyer and the seller respectively, in construction, the client will be facing harder situations since all the three rents are controlled by the contractor\(^{13}\). First of all, the rent from process specificity can be exploited in different ways. The first possibility is to require the client to pay higher price with some alleged excuse, like hiking costs of building materials. The second possibility is to exploit the client’s weakness once a change order is initiated. The third possibility is by way of submitting a design with less construction costs, but with lower client’s satisfaction.

The strategy of exploiting design rent becomes feasible only under the integrated route. As a matter of fact, the presence of design rent \textit{per se} will do no harm to the client if process specificity vanishes in that the client can then reject any proposal he/she dislikes with no fear of switching to alternative contractors. In fact, design rent is a special form of quality rent arising from integration of design and construction. Its magnitude depends on the controlled quality standard set by the client and complexity of the project (which affects \(C_i(s)\)).

From the perspective of transaction cost economising, which procurement route is more efficient depends on the trade-off between transaction costs arising from seeking of quasi rent, quality rent and design rent. Only when the conditions that affect the severity of rent-seeking behaviour are pointed out can the refutable hypotheses be derived. And this entails further empirical investigations.

\(^{13}\) It is assumed that the client’s costs of disruption are higher those of the contractor.
Table 6-5 Comparison of procurement systems in terms of contractual characteristics

<table>
<thead>
<tr>
<th>Degree of fragmentation</th>
<th>Degree to which design and construction can be overlapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>Management systems</td>
</tr>
<tr>
<td>M</td>
<td>Traditional method</td>
</tr>
<tr>
<td>L</td>
<td>Design and Build</td>
</tr>
</tbody>
</table>

6.6 Dimensionalisation of procurement routes

This step aims to answer the question: why will three main types of procurement systems display differential competence in dealing with projects with different transaction attributes? The explanation lies in their different contractual arrangement.

In principle, the procurement systems can be characterised by two dimensions (see Table 6-5): (1) the extent to which design and construction can be overlapped and thus faster completion be achieved and (2) degree of fragmentation. Fragmentation is defined as the inverse of concentration of responsibility to the client. According to the arrangements of the three main types of procurement systems to be examined in this study, fragmentation can be achieved by two ways: one is to delegate design and construction to different agents (i.e., designer and main contractor); another is to split the whole construction work into small parts and assign them to different work package contractors. As will be explained later, the fragmentation of procurement systems will give rise to two opposite forces at work (see Fig.6-13). More exactly, as the degree of fragmentation increases, adaptability to change will be better, while the efficacy of ensuring conformance quality becomes weaker.

This section attempts to explain why different contractual arrangement is an intrinsically decisive determinant affecting the efficacy of procurement systems with respect to three dimensions: time value, process specificity and measurement difficulties. An overview of the expository framework is exhibited in Fig. 6-14. The analysis is made up of two steps:
Fig. 6-13 The effect of fragmentation of procurement routes

Fig. 6-14 The matching of characteristics of procurement routes with objective

<table>
<thead>
<tr>
<th>Characterisations of procurement systems</th>
<th>Dimensions of procurement systems</th>
<th>Effects of each dimension on the objective function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree to which design and construction can be overlapped</td>
<td>Time of project delivery</td>
<td>NPV of revenue from the project</td>
</tr>
<tr>
<td>Degree of Fragmentation</td>
<td>Adaptable to design changes</td>
<td>Transaction costs from process specificity</td>
</tr>
<tr>
<td></td>
<td>How easily the liability of agents can be identified</td>
<td>Transaction costs from measurement difficulties</td>
</tr>
</tbody>
</table>

(1) The first step aims to wrestle with the problem as to the relation between the characteristics of contractual arrangement across procurement routes and their effects on the differential efficacy of procurement routes;

(2) The second step is to fill up the gap between the dimensions of procurement routes and their effects on the elements of the objective function, i.e. profit function (revenue net of transaction costs for fixed production cost). Transaction
costs have two origins: one type arises due to process specificity; another arises because of measurement difficulties.

In the following, how these two steps are linked with respect to the three elements of objective function will be explained.

6.6.1 Delivery time

Different contractual arrangement of the project coalition can give rise to different delivery time. The traditional method follows the sequence of design, review of design, tender and construction, while fast track strategy tries to overlap the task of different stages, so the total time can be compressed (an example is shown in Fig.6-15). However, what kind of arrangement can render fast track strategy more feasible? Let's examine three procurement routes in turn.

The traditional method is characterized by the sequence of "design then tender". This makes it harder to use fast tracking strategy. As a matter of fact, it is possible to use accelerated traditional method to speed up the completion of the project. However, it will be subject to the costs of sequential bargaining on price with contractor, so the traditional method can be said to be least suitable for fast tracking.

![Fig. 6-15 Comparison of traditional method and fast-track method](image)

In contrast, this strategy can be more easily implemented by design-build and management system. Nonetheless, to help the bidding D&B contractors formulate a
proposal based on sufficient information, it takes more time for the client to prepare preliminary design in advance. As a result, the benefits of fast tracking in shortening delivery time can be more fully crystallised by management system. The above inference is consistent with empirical evidence. According to Building Economic Development Council’s *Faster Building for Commerce* (1988) (see Table 6-6), regardless of site time or total time, management contracting is most likely to deliver the earliest date of completion. Design-build comes second and the traditional method is the slowest route.

Yet, why do we need to be concerned with the time of delivery? That’s due to its effect on the discounted revenue from the project. In principle, for the commercial or industrial clients, time is money. It seems safe to assume that the discounted revenue is a decreasing function of delivery time. The order of delivery time by three procurement routes is $t_{MS} < t_{DB} < t_{TM}$, so we can infer that the order of the corresponding discounted revenue will be $NR(t_{MS}) > NR(t_{DB}) > NR(t_{TM})$. This is an effect of a procurement route attribute on revenue from the project.

### Table 6-6 Comparison of delivery time through three procurement routes (BEDC, 1988)

<table>
<thead>
<tr>
<th></th>
<th>Percentage projects % (Site time)</th>
<th>Percentage projects % (Total time)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Faster</td>
<td>Average</td>
</tr>
<tr>
<td>Traditional</td>
<td>17</td>
<td>37</td>
</tr>
<tr>
<td>Design and Build</td>
<td>47</td>
<td>11</td>
</tr>
<tr>
<td>Management system</td>
<td>50</td>
<td>38</td>
</tr>
</tbody>
</table>

Note: 'Fast' projects are those completed at least 10% faster than the norm.

'Average', completed within 10% of the norm.

'Slow', taking at least 10% longer than the norm.

### 6.6.2 Adaptability to client-originated design changes

The second dimension of procurement systems is the differential flexibility in accommodating design changes during construction. An indicator for the degree of flexibility is the cost of implementing change orders. The key issue here is how this cost varies as the contractual arrangement of procurement systems. We want to argue from two respects (see Fig.6-16): contractual arrangement of procurement systems
will, on the one hand, alter the extent of process specificity and, on the other, the
efficacy of court ordering. Both of them have bearing on the cost of making changes
to the client’s requirement. In the following two subsections, we will try to build up
these two links.

Fig. 6-16 The link between procurement systems and adaptability to design changes

![Diagram showing the relationship between contractual arrangement of procurement systems, extent of process specificity, and cost of change orders.]

The key issue here is how this cost varies as the contractual arrangement of
procurement systems. We want to argue from two respects: contractual arrangement
of procurement systems will, on the one hand, alter the extent of process specificity
and, on the other, the efficacy of court ordering. Both of them have bearing on the
cost of making changes to the client’s requirement. In the following two subsections,
we will try to build up these two links.

6.6.2.1 Procurement systems and process specificity

Let’s discuss this problem from two scenarios. At the pre-contract stage, the client can
elicit market prices of any constructed product with given design from all the
potential bidders. However, at the post-contract stage, the price of a design change is
determined by the outcome of negotiation between the client and the winner
contractor. The difference is, in the latter case, the loss of options for switching to
alternative contractor without a cost. The issue of interest is why the degree of process
specificity will change as contractual arrangements.

In principle, as the route becomes more fragmented, the process specificity becomes
less severe. In other words, the time spent in finding a replacement contractor and the
time the replacement contractor needs to ‘warm up’ will decrease as degree of
fragmentation. When the responsibility is more concentrated on one party (like design and build), more of the tacit knowledge obtained in carrying out this project (in particular defects in design and construction) cannot be transferred, and the more difficult it is for other agents to continue the unfinished tasks (longer $t_2$), the fewer contractors that are willing to take over this task, so the more time is needed to find the replacement contractor (longer $t_1$). Moreover, using an integrated procurement system will increase the danger of quality uncertainty since the problem of information asymmetry as to quality will be more severe. This will increase the risk premium required by the replacement contractor $C_{rp}$ or/and the difficulties in handling the interface problem, i.e., higher $C_d$. In respect of $C_e$, if design-build or traditional method is used, temporary works and construction machinery or equipment has to be set up by the main contractor. When the original main contractor is replaced, all these set-ups need to be removed. In the management system, thought the major part of temporary works are arranged by the management contractor, the individual trade contractor is responsible for installing machinery for his own use. Thus, $C_e$, is much smaller in the management system than in other systems. The above discussion is summarised in Table.6-7. It evidently shows that it is most costly to replace the contractor in the design-build, implying this route puts the client in the weakest position to bargain with the contractor for the pricing of ex post requirement change. This means that design-build is least appropriate to deal with these projects with high possibility of changing the original requirements in the brief. For such projects, traditional method comes next and management system is the most desirable one.

<table>
<thead>
<tr>
<th></th>
<th>Traditional method</th>
<th>Design-build</th>
<th>Management system</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>Medium</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>$t_2$</td>
<td>Medium</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>$C_{rp}$</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>$C_d$</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>$C_e$</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 6-7 The effect of contractual arrangement on process specificity
6.6.2.2 Procurement systems and efficacy of court ordering

The presence of quasi-rent is the ultimate reason inducing rent-seeking behaviour. The severity of this problem is dependent on both the magnitude of the rent and the likelihood of successfully appropriating it. The latter is considerably affected by the differential efficacy of court-ordering under different circumstances. The origin of this differentiation mainly lies in the plasticity of the decision at issue. Plasticity involves two aspects: (1) the range of legitimate options that can be chosen by decision-makers, and (2) the costs of monitoring [Alchian & Woodward, 1987]. Others things being equal, the more choices the agent can take, the more plastic is the case. However, the number of choices just reflects half of the story, and the other half depends on whether reliable and less costly monitoring instruments are available. If it is easy for the principal to check the performance of his/her agents, how many options are open to the agent to choose won’t be of concern to the principal. On the contrary, if monitoring cost is prohibitive, even though options were few, the principal’s interest is still not secure.

When the procurement route becomes plastic, the efficacy of court ordering will be weakening because it is relatively more difficult for the third-party to make a judgment on the basis of an ambiguous yardstick. Of the three main procurement routes, design-build is most plastic since the detailed design is open at the time when the contract is placed and the contractor is given more latitude in determining undecided design as well as selecting technology and construction procedures [Ndekugri and Turner, 1994; Rowlinson, 1987]. In the management system design is also not finished at the stage of appointment of the management contractor, but what differs is that drawing and specification are offered by an independent agent of the client, the architect. In contrast, in the traditional method design has completed before tendering, while, as far as the plasticity of the contractor’s decision is concerned, these two routes are of the similar level.

6.6.2.3 A conclusion

It is claimed that adaptability to design changes of a procurement system is conditional on both its effect on process specificity and differential efficacy of court-ordering under different contractual arrangements. The results of the above
discussion can be summarised in Table 6-8. It manifestly shows that, in terms of both of the two dimensions, design-build is the least efficacious route to accommodate design changes *ex post*. In contrast, management system gives rise to slightest process-specificity effect for given attributes of the project and the similar level of court-ordering effect as traditional method, so it follows that management system would be the most desirable route to carry out these projects with high process specificity.

| Table 6-8 A comparison of adaptability to design changes of procurement systems |
|----------------------------------------|-----------------|-----------------|-----------------|
|                                       | Traditional method | Design-build | Management system |
| Extent of process specificity          | Medium            | High           | Low             |
| Efficacy of court ordering on design quality | High            | Low            | High            |

The relative efficacy of procurement systems in dealing with process specificity: a mathematical representation

For ease of notation in Sec.6.7, transaction costs are expressed as a function of three variables:

(1) \( v \): the client’s time value

(2) \( m \): technology/design related process specificity

The degree of process specificity is affected by two attributes of the project: opportunity cost of time \( (v) \) and technology characteristics. To separate the influence of \( v \) on degree of process specificity, in the following discussion, process specificity is referred to technology-related process specificity by holding \( v \) constant.

(3) \( k \): degree of measurement difficulties in dealing with interface problems between trade contractors or subcontractors.

That is,

\[ TC = TC(v,m,k) \]

To distinguish the variable in focus and the variables kept constant, alphabets in italic stand for given values of variables. For example, \( v \) is a variable, but \( v \) is the value of
With the help of this notation system, differential capability of three procurement routes in dealing with process specificity can be expressed in a mathematical form:

$$TC'_{db}[v,m,k] > TC'_{TM}[v,m,k] > TC'_{ms}[v,m,k]$$  

Eq.6-23

This order implies that as the degree of process specificity is deepening, the marginal transaction costs with respect to process specificity, $m$, will be rising relatively more quickly if an integrated route, like design-build, is used.

### 6.6.3 Degree of efficacy in assuring construction quality

Apart from the cost of monitoring design discussed in the previous section, measurement costs are made up of two other components: (1) costs of inspecting the completed parts of a project; and (2) costs of getting defects corrected. To make these costs comparable under the three procurement routes, the following analysis proceeds on the basis of the same design.

As analysed in Sec.6.5.1, the costs of inspecting should be of the similar level across procurement routes. Thus, we will focus on the difficulties in getting defects corrected. Single point of responsibility is more efficient in dealing with this problem. Design-build contractor has no scapegoat to get rid of responsibility once something has gone wrong. In the traditional system, the grey area of liability between designer and contractor may give one party the opportunity to impose charge upon another. The arrangement of management system would make the situation even worse since in this setting, designer, management contractor, trade contractors are all autonomous entities who are responsible for their own faults. For the client, this would lead to a big trouble in clearly identifying the bearer of liabilities. On the basis of this logic, it implies that marginal transaction cost arising from measurement difficulties will increase most substantially by using management system.

As a result, the relative competence of three procurement systems in dealing interface problems between trade contractors or subcontractors can be presented in the following relation,

$$TC'_{ms}[v,m,k] > TC'_{TM}[v,m,k] > TC'_{db}[v,m,k]$$  

Eq.6-24

159
6.6.4 Implications

6.6.4.1 An inconsistent trinity

The overriding task of procurement systems dimensionalisation is to discover the attributes of procurement systems that have direct bearing on the profit function. Contrasting Table 6-9 with Table 6-10, we can find that the intrinsic distinctions in terms of the extent to which design and construction can be overlapped and of fragmentation manifest themselves in the differential competence of procurement systems in the three dimensions: delivery time, adaptability to changes and degree of measurement difficulties.

Table 6-9 Micro-economic characteristics of three main procurement routes

<table>
<thead>
<tr>
<th>Characterizations of procurement systems</th>
<th>Traditional Method</th>
<th>Design &amp; Build</th>
<th>Management System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent to which design and construction can possibly be overlapped</td>
<td>L</td>
<td>M</td>
<td>H</td>
</tr>
<tr>
<td>Fragmentation of procurement routes</td>
<td>M</td>
<td>L</td>
<td>H</td>
</tr>
</tbody>
</table>

Table 6-10 Micro-economic features of main procurement routes

<table>
<thead>
<tr>
<th></th>
<th>Traditional Method</th>
<th>Design and Build</th>
<th>Management system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery time</td>
<td>Slow</td>
<td>Fast</td>
<td>Very fast</td>
</tr>
<tr>
<td>Adaptability to design changes</td>
<td>Weak</td>
<td>Very Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Ease of identifying liabilities of agents</td>
<td>Less difficult</td>
<td>Easy</td>
<td>Difficult</td>
</tr>
</tbody>
</table>

As exhibited in Fig.6-14, fast delivery is mapping to higher net present value of revenue stream from the project, lower degree of fragmentation (or single point of responsibility) reflects less severe burden on the measurement of performance (of each participant) and high adaptability to design changes gives contractor less
opportunity to take advantage of the client’s quasi-rent. As a consequence, to maximise profit function, a procurement route with the feature of fast delivery, single point of responsibility and high adaptability would be the ideal option. Unfortunately, this is unobtainable. The dilemma facing the client in selecting the procurement route is similar to that facing a monetary authority in selecting exchange rate policy. 1999 Nobel laureate Robert Mundell put forth a very famous “impossible trinity”: a country cannot pick free capital movement, a fixed exchange rate and an effective monetary policy at the same time. Likewise, it is impossible for the client to choose a procurement route meeting these three conditions simultaneously. This inconsistent trinity connotes the necessity of seeking balance between counteracting factors.

6.6.4.2 A formal presentation

Assume the same project is carried out by three procurement routes; the profit function corresponding to each route can be expressed as

\[ \pi_{TM} = NR(t_{TM}) - TC_{TM} - PC \]  
\[ \pi_{DB} = NR(t_{DB}) - TC_{DB} - PC \]  
\[ \pi_{MC} = NR(t_{MC}) - TC_{MC} - PC \]

where discounted revenue stream \( NR(t) \) is a function of total time of design and construction; \( TC_{TM}, TC_{DB} \) and \( TC_{TM} \) are transaction costs of three routes; \( PC \) is production cost, kept constant across procurement routes by following the standard assumption of TCE.

In the discrete structural analysis, the determination of the efficient governance structure is a comparative undertaking. Thus, we subtract Eq.6-25–Eq.6-27 permutively by each other, i.e.,

\[ \pi_{DB} - \pi_{TM} = [R(t_{DB}) - R(t_{TM})] - [TC_{DB} - TC_{TM}] \]  
\[ \pi_{MS} - \pi_{TM} = [R(t_{MS}) - R(t_{TM})] - [TC_{MS} - TC_{TM}] \]  
\[ \pi_{MS} - \pi_{DB} = [R(t_{MS}) - R(t_{TM})] - [TC_{MS} - TC_{TM}] \]

For ease of notation, the average cost of time, \( v \), is introduced to express the difference in revenue of two different delivery times. For example,

\[ R(t_{DB}) - R(t_{TM}) = v \times (t_{TM} - t_{DB}) \]

Besides, from Eq.6-21, we can find two categories of project attributes affect process
specificity: average cost of time and design/technology. If we denote design/technology derived process specificity as $s$ and assume a linear relation between delivery time and revenue, Eq.6-28~Eq.6-30 can be transformed into

$$\pi_{DB}-\pi_{TM} = v \times (t_{TM}-t_{DB}) - [TC_{DB}(v,m,k)-TC_{TM}(v,m,k)]$$  Eq.6-31

$$\pi_{MS}-\pi_{TM} = v \times (t_{TM}-t_{MS}) - [TC_{MS}(v,m,k)-TC_{TM}(v,m,k)]$$  Eq.6-32

$$\pi_{MS}-\pi_{DB} = v \times (t_{DB}-t_{MS}) - [TC_{MS}(v,m,k)-TC_{DB}(v,m,k)]$$  Eq.6-33

To get a more specific finding, we assume there is a fixed relation between $t_{TM}$, $t_{DB}$, and $t_{MC}$. For simplicity, $t_{MC}$ and $t_{DB}$ are assumed to be proportional to $t_{TM}$ with a given ratio of $\alpha$ and $\beta$ respectively ($1 > \alpha, \beta > 0$). For a project, the expected delivery time by traditional method can be estimated ex ante. This is, $t_{TM}$ is given. Then, Eq.6-31~6-33 are simultaneous equations, consisting of three equations and three unknown variables. The problem of interest is whether there is a condition under which the desirability of the three routes is indifferent to the client. Put it in a mathematical way, this problem requires to find the solution of Eq.6-31~6-33 by setting $\pi_{DB} - \pi_{TM} = \pi_{MC} - \pi_{TM} = \pi_{MC} - \pi_{DB} = 0$. Its solution $(v^*, m^*, k^*)$, if any, stands for a set of transaction attributes where the three routes will perform equally well (see Fig.6-17). In other words, in dealing with the project with the attributes of $(v^*, m^*, k^*)$, the client will be indifferent over using any of the three procurement systems. Except for this type of project, the choice of procurement routes will influence the efficiency of project implementation.

**Fig. 6-17 Point of indifference in procurement route selection**
6.7 Determination of the efficient procurement route

6.7.1 Opportunity cost of time

Since in the most commercial and industrial projects, the earlier completion often brings in higher discounted value of revenue. For convenience of the following discussion, discounted revenue is assumed to be the function of delivery time \( t \) (an attribute of procurement routes) and opportunity cost of time (OPPCT; \( v \)) (an attribute of the project), i.e., \( R = R(t,v) \). However, as OPPCT becomes larger, process specificity also increases as indicated in Eq.6-21. Thus, it is essential for us to distinguish the origin of process specificity. On the one hand, the degree of specificity depends on the availability of construction technology employed in the project, labelled as \( m \). On the other, OPPCT can alter the opportunity cost of time in delay, were the contractor to be replaced. Thus, asset specificity is a joint function of \( v \) and \( m \).

After these factors are taken into account, a complete expression of Eq.6-25~6-27 is

\[
\begin{align*}
\pi_{TM}(v) &= NR(v, t_{TM}) - TC_{TM}(v,m,k) - PC \quad \text{Eq.6-34}\\
\pi_{DB}(v) &= NR(v, t_{DB}) - TC_{DB}(v,m,k) - PC \quad \text{Eq.6-35}\\
\pi_{MS}(v) &= NR(v, t_{MS}) - TC_{MS}(v,m,k) - PC \quad \text{Eq.6-36}
\end{align*}
\]

In this set of equations, technology-related process specificity \( m \) and measurement difficulty \( k \) are held constant and the delivery time of three routes \( (t_{TM}, t_{DB}, t_{MS}) \) are given. All the exogenous variables are expressed in italic font. Thus, the only variable here is \( v \). The relative location of Eq.6-34-Eq.6-36 in a graph is determined by their slope and interception with y-axis of each curve.

First of all, regarding the relative slope of three curves, we have to do the first differentiation with respect to \( v \) to Eq.6-34-Eq.6-36, giving

\[
\begin{align*}
\pi'_{TM}(v) &= NR'(v, t_{TM}) - TC'_{TM}(v,m,k) \quad \text{Eq.6-37}\\
\pi'_{DB}(v) &= NR'(v, t_{DB}) - TC'_{DB}(v,m,k) \quad \text{Eq.6-38}\\
\pi'_{MS}(v) &= NR'(v, t_{MS}) - TC'_{MS}(v,m,k) \quad \text{Eq.6-39}
\end{align*}
\]

Because design-build is the most inappropriate route to deal with the project with high process specificity, we can get the order of

\[
TC'_{DB}(v,m,k) > TC'_{TM}(v,m,k) > TC'_{MS}(v,m,k). \quad \text{Eq. 6-40}
\]

Concerning the ability to implement fast-track strategy, management system > design
and build > traditional method. In a mathematical term, we can get

\[ NR'(v, t_{MS}) > NR'(v, t_{DB}) > NR'(v, t_{TM}) \]  
Eq. 6-41

Combining Eq.6-40 and 6-41, we can know the slope of \( \pi_{MS} \) is the steepest, but the relation between \( \pi'_{TM} \) and \( \pi'_{DB} \) cannot be determined, ie.,

\[ \pi'_{MS} > \pi'_{TM}, \pi'_{DB} \]  
Eq. 6-42

For a maximisation problem, there are two possibilities: (1) \( \pi'_{MS} > \pi'_{TM}, \pi'_{DB} > 0 \) and \( \pi''_{MS}, \pi''_{TM}, \pi''_{DB} < 0 \); or (2) \( 0 > \pi'_{MS} > \pi'_{TM}, \pi'_{DB} \) and \( \pi''_{MS}, \pi''_{TM}, \pi''_{DB} > 0 \). These two cases will be discussed in turn.

To locate three curves, apart from the relative slope of three curves, we have to know the interaction of three curves with y-axis. Without further assumptions, from Eq.6-37~6-39, we cannot decide the relative magnitude of \( \pi'_{TM}(0), \pi'_{DB}(0), \pi'_{MS}(0) \). However, with reference to the fact that all three routes appears along the continuous line of \( v \), if we can rule out the cases where there is any one route always dominant\(^{14}\), the location of three curves can be easily identified.

**Fig. 6-18 The determination of the efficient route with respect to time value:**

**Case I**

(a) \( \pi'_{MS} > \pi'_{TM} > \pi'_{DB} > 0 \)

---

\(^{14}\) My only defence for this practice relies on Williamson’s authority (Williamson, 1991). If \( H(k) \) and \( X(k) \) denote governance costs of hierarchy and hybrid organisation as a function of asset specificity, \( k \). Williamson assumes \( X(0) < H(0) \) “since otherwise the hybrid mode could be dominated throughout by the least-cost choice of either market or hierarchy which may occur for certain classes of transaction”.
Case I: $\pi'_{MS} > \pi'_{TM}, \pi'_{DB} > 0$

For this case, there are two possibilities of the location of three curves, as shown in Fig.6-18(a) and (b). The bold curves mark the attainable efficient frontier (AEF) of two cases. In both case, when $v > v_f$, $\pi_{MS} > \pi_{DB}, \pi_{TM}$. Thus, management system is the most desirable procurement system under this condition. However, when $v_f > v > v_2$, the most desirable route can be traditional method in Fig.6-18(a) and design-build in Fig.6-18(b). That is, the relative desirability of design-build and traditional method with respect to time value is indeterminate.

Fig. 6-19 The determination of the efficient route with respect to time value of the project: Case II

(a) $0 > \pi'_{MS} > \pi'_{TM} > \pi'_{DB}$

(b) $\pi'_{MS} > \pi'_{DB} > \pi'_{TM} > 0$
Case II: $0 > \pi'_{MS} > \pi'_{TM} > \pi'_{DB}$

In this case, three curves are negatively sloped. Fig. 6-19 shows that there are two possibilities. From Fig. 6-19(a), we know that,

1. when $V > v_f$, $\pi_{MS} > \pi_{DB} > \pi_{TM}$. Thus, management system is the most desirable procurement system under this condition.

2. when $v_f > v > v_2$, $\pi_{TM} > \pi_{DB}$, Thus, traditional method is the most desirable procurement system under this condition.

3. when $v < v_2$, $\pi_{DB} > \pi_{TM} > \pi_{MS}$. Thus, design and build is the most desirable procurement system under this condition.

In Fig. 6-19(b), the reverse order between $\pi_{DB}$ and $\pi_{TM}$ gives a different result on the condition under which traditional method and design-build is applicable, namely the former is more desirable for the project with lower OPPCT and the latter for the project with medium OPPCT. This means that we cannot differentiate TM and DB a priori in terms of OPPCT.

6.7.2 Degree of process specificity

As explained above, process specificity is affected by $v$ and $u$ simultaneously. In this subsection, we focus on technology-related process specificity, by holding the effect of time value, $v$, constant.

In terms of transaction costs, we can unambiguously identify the range of each procurement system being more advantageous with respect to process specificity. However, if we taken revenue term into consideration, the situation will be more
complicated. In fact, what really concerns us is the comparison of the profit that can be generated by each procurement system, i.e.,

\[
\pi_{TM}(m) = NR(v, \tau_{TM}) - TC_{TM}(v, m, k) - PC \quad \text{Eq. 6-43}
\]

\[
\pi_{DB}(m) = NR(v, \tau_{DB}) - TC_{DB}(v, m, k) - PC \quad \text{Eq. 6-44}
\]

\[
\pi_{MS}(m) = NR(v, \tau_{MS}) - TC_{MS}(v, m, k) - PC \quad \text{Eq. 6-45}
\]

When we try to plot \( \pi_{TM}(m) \), \( \pi_{DB}(m) \) and \( \pi_{MS}(m) \) vs. \( m \) on the same diagram, we have to know the slope and y-axis intersection of each curve. First of all, we need to compare the first order differentiation of Eq.6-43–6-45 with respect to \( m \). From our discussion on process specificity, we can infer

\[
TC'_{DB}(v, m, k) > TC'_{TM}(v, m, k) > TC'_{MS}(v, m, k)
\]

It follows

\[
|\pi'_{DB}| > |\pi'_{TM}| > |\pi'_{MS}|
\quad \text{Eq. 6-46}
\]

Likewise, if we preclude the case where any route is always dominant over others. The only possibility is shown in Fig.6-20, from which we can find that

1. if \( m > m_2 \), management system is the most desirable one;
2. if \( m_2 > m > m_1 \), traditional method is the most efficient one;
3. if \( m < m_1 \), design-build will stand out.

![Fig. 6-20 The influence of process specificity on transaction costs](image)

**Fig. 6-20** The influence of process specificity on transaction costs

### 6.7.3 Measurement difficulties

The effect of measurement difficulties can be examined by holding other factors equal. The set of objective functions to be compared is
\[ \pi_{TM}(k) = NR(v, f_{TM}) - TC_{TM}(v, m, k) - PC \]  
Eq. 6-47

\[ \pi_{DB}(k) = NR(v, f_{DB}) - TC_{DB}(v, m, k) - PC \]  
Eq. 6-48

\[ \pi_{MS}(k) = NR(v, f_{MS}) - TC_{MS}(v, m, k) - PC \]  
Eq. 6-49

To determine the location of these curves, we have to make use of the analysis results of

\[ TC'_{MS}(v, s, k) > TC'_{CM}(v, s, k) > TC'_{DB}(v, s, k) \]

, it follows that

\[ |\pi'_{MS}| > |\pi'_{TM}| > |\pi'_{DB}| \]  
Eq. 6-50

Considering these conditions and excluding the cases where any route is always dominant over others, the only possibility is shown in Fig. 6-21, from which we can find:

1. When \( k > k_2 \), design-build is the most efficient procurement route;
2. When \( k_1 < k < k_2 \), traditional method stands out;
3. When \( k < k_1 \), management system becomes the most efficient route.

As a result, we can be sure that when serious measurement difficulties make design and build more advantageous and slight measurement difficulties can make management system more desirable. Moreover, the possible condition suitable for traditional method will be between two ends.

**Fig. 6-21 The influence of measurement difficulties on transaction costs**

![Fig. 6-21 The influence of measurement difficulties on transaction costs](image)

168
6.7.4 Concluding remarks of this section

6.7.4.1 The logic of alignment

The way that procurement systems are aligned with transaction attributes is illustrated in Fig.6-22. In the middle lie the constituent elements of the objective function: revenue, transaction costs caused by asset specificity and by measurement difficulties. To achieve the goal of profit maximisation, we search the conditions (a certain set of transaction attributes) under which each procurement route becomes most desirable.

On the right hand side, three attributes of the project are the key determinant in affecting the objective function: high client’s opportunity cost of time makes the earlier completion matter, high degree of process specificity and complexity of specialisation between trade or subcontractors makes the rent-seeking problems more serious. But, why the competence of procurement systems is different with respect to these three dimensions? The answer lies in the differential capability of procurement systems in three respects: the extent to which design and construction can be overlapped, adaptability to design changes and degree of responsibility concentration. These three dimensions are mapping to three elements of the objective function, namely that high degree of design and construction being overlapped can deliver higher net present value of revenue, high extent of responsibility concentration can help reduce transaction cost caused by measurement difficulties and high adaptability.
to design changes can alleviate the rent-seeking behaviour caused by process specificity. This framework enables us to choose the route best able to handle projects with different attributes.

6.7.4.2 Conditions under which three routes are most suitable

Regarding the properties of procurement systems, conclusions can be drawn as follows. Client’s opportunity cost of time matters in comparing the present value of revenue that can be generated by different procurement systems. If the delivery time of the project makes the revenue stream different, the fast tracking route will be favoured. From the theoretical analysis, we can only identify the condition under which high OPPCT project is most suitable for management system. The relative desirability of traditional method as well as design-build is indeterminate.

The second critical factor, process specificity will impinge upon transaction costs due to the differential capability of procurement routes in accommodating unforeseen changes. Management system is best for the condition where process specificity is very high, traditional method may be suitable for the project with medium process specificity and design and build is most competent to deal with the projects with low process specificity.

Fig. 6-23 Alignment of procurement routes with transaction attributes: results
The last factor concerned in this chapter is measurement difficulties in identifying the liability of contractors involved in the project. The dispersion of responsibility makes it hard for the client to find the right agent to blame. This problem can be overcome by devolving the project to a single party. Thus, when the interface problem between participating parties are serious, design-build becomes the most suitable choice. As the severity of this problems is alleviated, more fragmented routes can be applied. That is, traditional method and management system can be used to deal with the project with medium and low measurement difficulties.

The above discussion is summarised in Fig. 6-23. It indicates the combinations of three factors under which each route is most desirable. Apart from these three conditions, which route is better depends on the relative strength of these factors.

These factors are upheld as the theoretical underpinnings of the model of construction procurement behaviour due to their theoretical significance in operationalising the construction transaction. Moreover, the importance of these factors is on their own, rather than depends on the presence of other factors. For example, as explained in Chapter 4, the effect of uncertainty on transaction hazards is conditional on the degree of asset specificity. Thus, this kind of variable is considered as of secondary importance.

6.8 Conclusion
It is found that the current literature is unable to demonstrate causality behind empirical facts. These works perhaps are sensible starting points, but definitely not the end. Developing a model by using the deductive method, like economic approach can deepen our understanding to the nature of procurement route selection. In fact, without a theoretical (economic) analysis, it is hard to check whether the pure empiricism would go wrong. In this sense, transaction cost based model can be a complement approach, able to fill up the vacuum of theoretical reasoning and enrich our understanding of the interaction between the participating parties of a project. Because the postulated objective function is profit maximisation, the model is most suited to explain procurement system selection for commercial and industrial buildings. However, it can be extended to other types of projects by including more relevant factors into the objective function.
Chapter 7 Empirical Investigations Of Construction Procurement Behaviour

Chapter 6 arrives at an essential theoretical conclusion that in selecting an appropriate construction procurement system, the client in fact is facing an inconsistent trinity since there is no procurement system can achieve fast delivery, low process specificity and low measurement difficulties simultaneously. To test the validity of this proposition empirically, we need to seek suitable proxy variables to capture the influence of client’s time value, process specificity and measurement difficulties on the profit function. As a result, this chapter aims to: (1) describe the process of searching observable variables that have direct bearing on the three key components of profit function; (2) illuminate the empirical methodology – reduced-form analysis; and (3) develop hypotheses.

**Fig.7-1 Three stages of operationalising transaction costs in construction**

<table>
<thead>
<tr>
<th>Operationalisation</th>
<th>Stage I</th>
<th>Stage II</th>
<th>Stage III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue function</td>
<td></td>
<td>OPPCT</td>
<td>Measurement of variables</td>
</tr>
<tr>
<td>Profit function</td>
<td>Transaction costs from process specificity</td>
<td>UNCER</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DEGPS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PARTR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SCALC</td>
<td></td>
</tr>
<tr>
<td>Transaction costs from measurement</td>
<td>COMSP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7.1 Preview of empirical investigations

Theoretical analysis of governance structures in TCE often stops where Chapter 6 has done, while its empirical testing starts from there. To make three key factors identified in Chapter 5 more operational, the step of operationalisation in Fig.4.1 ought to be extended into three stages as shown in Fig.7-1. Two follow-on stages are devised to conjure up the variables that have direct impact upon the revenue function, transac-
tion cost from process specificity and that from measurement difficulties (stage II), and work out the problems as to how these variables can be measured in a feasible way (stage III). The proposed procedure has been pervasively employed in the TCE empirical studies in the 1980s and 1990s. Masten (1995) compiles a good record of literature along this line.

In Fig. 7-1, the output of chapter 6, are taken as input to chapter 7 and an attempt is made to further explore the relevant variables directly affecting the three components underlying the profit function. Then the focus shifts to problems of measurement of these variables in chapter 8.

### 7.2 In search of observable variables

#### 7.2.1 The principle of search

Fig. 7-2 displays the link for searching for relevant factors that can most effectively distinguish the relative desirability of procurement systems in terms of three components of the objective function: revenue, transaction cost from process specificity and that from measurement difficulties. The differential capability of two governance structures in generating profit can be represented by two categories of variables: the client attributes and the project attributes. The factors that considerably affect the magnitude of difference in revenue and transaction costs across procurement systems are those for which we are searching. The following subsections aim to elaborate on the process of finding these relevant factors.

**Fig. 7-2 In search of observable variables**

<table>
<thead>
<tr>
<th>Stage I</th>
<th>Stage II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Revenue</strong></td>
<td><strong>Linkage</strong></td>
</tr>
<tr>
<td></td>
<td>Difference in NPV of revenue for different completion dates</td>
</tr>
<tr>
<td><strong>Transaction costs from process specificity</strong></td>
<td><strong>Influenced variables</strong></td>
</tr>
<tr>
<td></td>
<td>Magnitude of specificity</td>
</tr>
<tr>
<td></td>
<td>Severity of rent-seeking behaviour</td>
</tr>
<tr>
<td><strong>Transaction costs from measurement difficulties</strong></td>
<td>Costs of defects correction</td>
</tr>
</tbody>
</table>

- OP CCT
- DEG PS
- UNCE R
- PAR TR
- SCAL C
- COP MSP

173
7.2.2 An example in the literature
Before we start searching for relevant critical factors, it will be helpful to give an example of how this is done in the TCE literature.

Most TCE empirical works depart from the hypotheses derived from the theoretical analysis, as shown in Chapter 4. Masten et. al. (1991) is the first work to use the concept of temporal specificity. Pirrong (1993) follows this line to explore the contracting practices in shipping markets. He argues that "forward contracts economise on transaction costs in markets where temporal specificity is important. 'To determine whether actual contracting practices conform to this pattern, one must identify conditions that affect the magnitude of the temporal specificities.' Three factors are pointed out as relevant: (1) the availability of geographically dispersed alternative cargo sources for the shipper, (2) thickness of the market, (3) specificity of vessels. He then attempts qualitatively to compare the contractual form in practice with theoretical predictions.

By means of this brief account of the explanation structure of TCE empirical studies, it is evident that the focus of this chapter should be directed to ferret out the key determinants that will impinge upon discounted revenue, process specificity and measurement difficulties of the project.

7.2.3 Factors affecting revenue difference
First, as far as $\Delta R$ is concerned, we want to address the question: what is the factor that may widen the gap in revenue if the project is delivered at different dates? The answer has been mentioned in chapter 5: opportunity cost of time (OPPCT), which can be understood as a cost change in financial units of a one-period increase in time taken. Suppose procurement system 1 and 2 can deliver the project in $t_1$ and $t_2$ days ($t_1 > t_2$). For the client, the differential revenue defined in Sec.6.5 will be

$$\Delta R = \underbrace{\text{OPPCT}}_{\text{Attribute of the project}} \times \underbrace{(t_1 - t_2)}_{\text{Attribute of the procurement systems}}$$
If the hidden costs of late completion are high, AR will go larger and the date of completion becomes more critical to the success of the project. For the project with this attribute, the advantage of fast-tracking strategy will be more obvious to the client. That is, the procurement system with the attribute of shorter delivery time of $t_6 (< t_1)$ is more suitable to the deal with this type of project. Other things being equal, the procurement routes that can serve the client's demand for fast delivery will be more likely to be chosen.

However, as stressed in Sec.6.7.1, as OPPCT increases, the process specificity increases as well, making design-build less desirable. Thus, only management system enjoys the relative advantage in this dimension.

7.2.4 Factors affecting transaction costs from process specificity (I) - DEGPS

Next, we want to explore what determinants have bearing on the magnitude of transaction costs arising from process specificity. This question is a bit more complicated. From Chapter 4 and Chapter 6, it can be seen that the troubles caused by process specificity are not only dependent on its extent per se but also on its interaction with other variables, like uncertainty. Thus in fact this main question comprises two parts: one is, what factors are responsible for the degree of process specificity? And another is, what are the factors that may aggravate the severity of process specificity? This subsection will focus on the first question.

Process specificity is indicative of the degree of interdependence between trading parties in the construction process, so the degree of process specificity (DEGPS) is a vital factor in affecting the balance of bargaining power between the client and the contractor. A good indicator for this is the cost of switching, defined by the extra costs of changing contractor post-contract. The sources of these costs include opportunity cost of delay, repeated costs for temporary facilities, risk premium charged by replacement contractor for uncertain quality, and costs in resolving the disputes of quality problems arising from the grey area of liabilities involving the original and replacement contractor. The totality of these costs signifies the importance of production continuity to the client and affects the bargaining power of the client relative to his/her agents. When the potential loss from disruption is substantial, the client will be in a weak position to prevent the post-contractual redistribution of economic rent. The higher is the
appropriable quasi-rent, the more attractive is the opportunistic strategy, and the more resources will be devoted to unproductive activities. Thus, DEGPS is a variable that we should consider.

7.2.5 Factors affecting transaction costs from process specificity (Il-I) — uncertainty (UNCER)

Before answering the second question, it is worth repeating that transaction costs from process specificity are rooted in the occurrence of opportunistic intentions. Thus, the factors that can mitigate or aggravate the proclivity to exercise opportunistic strategies should be encompassed in the list. With reference to TCE, the first factor coming to mind is uncertainty. As the occurrence of unanticipated events increases, the cooperation relation between the client and he/her agents may deteriorate due to rent-seeking behaviour. The incompleteness of contract makes uncertainty appear threatening to the client’s interest under some circumstances. This is a factor on the negative side.

On the positive side, we have to explore what factors might mitigate the likelihood of rent-seeking behaviour. A direct answer is trust. However, owing to its multifaceted connotations, an effort should be made to clear away the confusion that the term may cause, to which the next subsection will turn.

7.2.6 A digression: the role of trust

The basic premise of the theory of human behaviour (or action) seems to be fairly different in sociology and economics. Economists are used to take the postulate that self-interest seeking is the motivation of paramount importance in guiding human behaviour. In contrast, sociologists show great uneasiness with this assertion. This is evidenced by the majority of fierce attacks on Williamson’s assumption of opportunism coming from sociologists. Are self-interest seeking and trust really completely inconsistent? This question entails careful scrutiny.

In TCE, transaction cost is regarded as an equivalence of friction in the physical world. By the same analogy, trust acts like the lubrication oil that facilitates the smooth working of the economy. In the real world, the main barrier to transaction lies in the incompleteness or asymmetry of information. The presence of trust can help
enhance the predictability of trading parties’ behaviour. As a much-quoted definition of trust in economics shows,

*When we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial... is high enough for us to consider in engaging in some form of cooperation with him. Correspondingly, when we say that someone is untrustworthy, we imply that that probability is low enough for us to refrain from doing so.* [Gambetta, 1988]

The predictability of behaviour comes from different origins. In the literature, many different classification systems have been proposed. For example, Sako (1992) in her work on the comparison of inter-firm commercial relations between Britain and Japan classified trust into three types:

1. **Contractual trust**, which means entrusting a person to endeavour to fulfill their contractual agreement;
2. **Competence trust**, which means entrusting a person on the basis of his/her ability to carry out the task;
3. **Goodwill trust**, which is based on the expectation of one party’s willingness to do more than contractually stipulated.

Another important work by Shapiro, Sheppard, and Cheraskin (1992) points out three reasons that can create trust. First, trust can be said to exist if the punishment is serious enough to discourage the proclivity to opportunism. This kind of trust, labeled *deterrence-based trust*, can be maintained due to the positive benefit of abiding by the original promises. Another type of trust can be built owing to informational advantage on the decision style of trading parties. This is called *knowledge-based trust*. The third type of trust, *identification-based trust*, comes from the deliberate design of incentive compatible terms and conditions of transaction so that both parties’ preferences can be fully aligned in the contract.

We can try to interpret the above six types of trust for purposes of this study. First of all, the nurturing of trust may rely on the efficacy of contract by fine-tuning the way that risk is allocated or resort is made to court ordering. Second, trust can be culti-
vated by collecting the critical information about the trading partners, such as their competence or decision style. Third, trust is a natural consequence of repeated transactions in that what you do today will determine what you will get tomorrow, sanction or reward, and will also be influenced by what the transactor did yesterday. Folk theorem in game theory gives a good example of why prisoner's dilemma can be sorted out when transactions are likely repeatedly to occur.

Because the basic proposition of TCE is risk-neutrality, the identification-based trust is not directly relevant in this setting. The second category of information-related trust, like competence trust or knowledge-based trust, is important in the analysis of organisations. Whether the contractual arrangement of the procurement system can improve incompleteness or asymmetry of information will change the likelihood of vulnerable positions being exploited and alter the desirability of the 'vulnerable' procurement route, like design-build. The deterrence category of trust can be understood as follows. If we extend the horizon of decision into the inter-temporal setting, one important kind of deterrence that can discourage contractor's opportunistic behaviour is the loss of opportunity to get the future works. Corresponding to these two categories of trust, we can find two factors relevant: whether partnering is used and the total demand of the client.

7.2.7 Factors affecting transaction costs from process specificity (II-2): partnering agreement
Use of a partnering agreement between construction clients and their agents, like contractors or designers, has gained attraction in construction practice in recent years. In response to the long-term complaints from the client side about the confrontational atmosphere in construction, the infusion of trust element is popularly accorded recognition as an encouraging way forward to soothe the tension. Two of the principal characteristics of long-term partnering are open book costing and the continuity of workload.

Open book costing
A direct benefit of transparent cost record is to alleviate distrust owing to asymmetry of information. When the information affecting the distribution of quasi-rent is public to all stakeholders, the ex ante commitment can relatively easily be enforced and the
potential disputes from unforeseen contingencies can be resolved with less cost. This will discourage the motivation for exercising opportunistic strategies, reducing the unpredictability of behaviour arising from tussling for quasi rent. Thus, trust can be growing.

**Continuous workload**

Another important spirit of partnering is providing the possibility of continuous workload. As analysed in Sec.6.7, in the course of construction, the contractor and the client are playing a game with loss-loss result. To avoid going into the vicious cycle, performance-dependent provision of future workload as an inducement to make the contractor give up opportunistic strategies can serve the function of building deterrence-based trust, which implies behaving in a cooperative way, otherwise losing certainty of continuous workload. This result has been theoretically demonstrated in the game theory literature.

In Kreps (1990), the effect of trust is captured by a game where player 1 chooses to Trust or Not Trust player 2 and player 2 then decides to honour or betray player 1's trust. The extensive form of this game is shown in Fig.7-3. The game assumes that, if player 1 offers trust first, he/she will possibly get the return of S if player 2 honours it or suffer a loss of L if player 2 betrays it. If player 1 doesn't take the strategy of Trust, the game terminates, meaning both players leave with payoff of zero.

![Fig. 7-3 The extensive form of a trust game](image)

The game can be solved by backwards induction, i.e., inducing backwards one node at a time along the decision tree. First, when player 2 chooses the strategy between Be-
tray and Honour, the former will be favoured because its payoff, $S+H$ is higher than that of the later ($S$). With this knowledge, player 1 would rather stop this game in the first place with no gain than trust player 2 and suffer the loss of $L$. It is obvious that this result is socially improvable since the strategy set of (Trust, Honour) will make both players better off. This dilemma can possibly be overcome by repeating the game.

Suppose player 2 is facing a two-period decision: (1) honour player 1’s trust offer and gain the player 1’s trust in the second period or (2) betray player 1’s trust offer and get nothing in the next period. The condition for inducing player 2 to take a cooperative strategy is

$$(1 + \frac{1}{r})S > (S + H)$$

where $r$ is the discounting rate of player 2. That is, when player 2’s time value is less than $S/H$ ($r < S/H$), it is more desirable for player 2 to pursue long-term benefits (i.e., honouring player 1’s trust offer) rather than realise the short-term gain (i.e., betray player 1’s trust offer). Thus, the new equilibrium will be (Trust, Honour). This result signifies the effect of repeated transactions in repressing the occurrence of opportunistic behaviour out of short-termism. From this result, it seems reasonable to infer that the use of partnering agreement can reduce the likelihood of opportunism happening induced by process specificity when the vulnerable route, like design-build is used. And this factor will possibly affect the client’s selection of procurement systems.

### 7.2.8 Factors affecting transaction costs from process specificity (Il-3): prospect of future works

Besides the use of partnering agreement, another factor playing the similar role is the size of the client, measured by the total scale of projects that the client commissions every year. The repeat client enjoys more bargaining power with his/her agent by giving preference in the award of the next project in exchange for the cooperation in this project. Therefore, the amount of construction demand that a client can provide indicates the extent of deterrence that a client can exercise. That is, given the same project attributes, the vulnerability of design-build will not be of the same level for the large client as for the small client. The large client will enjoy the benefit of owning the
“hidden” deterrence power and face lower risk in making himself exposed to the vulnerable positions.

7.2.9 Factors affecting transaction costs from measurement difficulties: COMSP
Lastly, as claimed in Sec.5.8.2, measurement problems in the analysis of procurement systems come from the costs of defect correction. What factors will make this cost become crucial in distinguishing the efficacy of procurement routes? The answer lies in the complexity of specialisation (COMSP). In principle, as the types and number of subcontractors or trade contractors increase, the more serious are the interface problems between them, the more difficult it becomes to clearly identify the bearer of liability. This situation may induce the subcontractors to exploit quality rent, and the second round of information problems may ensue. A single point of responsibility can help alleviate this problem, that is, an integrated procurement route will be favoured.

7.3 The empirical methodology: reduced-form analysis
In Williamson’s TCE, to get around the problem of transaction costs unmeasurability, the ingenuous solution of comparing alternative GSs by operationalising transaction attributes is introduced. In the reduced form, the difference in the cost of governance structure of GS_i and GS_j, TC_i and TC_j, can be expressed as

\[ TC_i - TC_j = aX + \varepsilon \]

where \( a \) and \( X \) stand for coefficient vector and independent variable vector, respectively and \( \varepsilon \) random disturbances. As a result, the absolute magnitude of transaction cost is of no significance in this discrete structural analysis [Williamson, 1996]. What are relevant are the key determinants that will impinge on the differential cost of two governance structures.

According to typical marginal analysis, efficiency requires that prices and quantities be adjusted to the margin. But in dealing with organisational problems, like how to determine the boundary of a firm, this sort of second-order economising (adjusting the margin) should give way to first-order economising (getting the choice of governance structures right) because of the discreteness of the choice set.
As is explained in chapter 5, the criterion by which the client selects the procurement route is to maximise the discounted profit function, net of revenue and transaction costs. We follow a core proposition in TCE that transaction costs are very difficult directly to measure, if not directly unmeasurable, whereas by means of discovering the determinants able to distinguish the differential competence in coping with various transaction relations, it is possible to predict the choice of governance structures from some observable factors. The keynote of this approach tells us that factors matter only in the comparative institutional sense. Therefore, the question needed to be brought to the fore, is what factors would impinge upon the difference in revenues and transaction costs under different procurement routes?

### 7.4 Conditions to which the three procurement systems are more desirable

#### 7.4.1 The analysis

Following the analysis in Sec.6.7, we can express the difference in profits generated from three routes as

\[ \pi_{DB} - \pi_{TM} = v \times (t_{TM} - t_{DB}) - \sum_{i=1}^{5} \Delta TC_{DB-TM}^i (x_i) \]  
\[ \text{Eq.7-1} \]

\[ \pi_{MS} - \pi_{TM} = v \times (t_{TM} - t_{MS}) - \sum_{i=1}^{5} \Delta TC_{MS-TM}^i (x_i) \]  
\[ \text{Eq.7-2} \]

\[ \pi_{MS} - \pi_{DB} = v \times (t_{DB} - t_{MS}) - \sum_{i=1}^{5} \Delta TC_{MS-DB}^i (x_i) \]  
\[ \text{Eq.7-3} \]

where \( v \) is the opportunity cost of time; \( t_{TM}, t_{DB}, t_{MS} \) are delivery time that can be achieved by three routes; \( x_i \) are the transaction attribute variables. To find the conditions to which each of three procurement systems are more desirable, let's examine six variables contained in the equation set of Eq.7-1 ～ 7-3 in turn by holding other variables constant.

Compared with Sec.6.7, there are three factors in common: OPPCT, DEGPS and COMSP, so in the following subsections, we concentrate on the other three factors: UNCER, PARTR, SCALC.

#### 7.4.2 Uncertainty

Uncertainty is one of the most critical features of economic decision-making. Its significance should not only be understood as a stochastic factor in the ex ante sense but, what is more important, its interactions with other factors ex post. From the
what is more important, its interactions with other factors ex post. From the viewpoint of TCE, uncertainty becomes relevant not in its own right, but due to its role in triggering the vicious cycle of rent-seeking behaviour when the transactors are closely interdependent. The basic logic can be described in Fig. 7-4, which shows two categories of factors at work: uncertainty and process specificity. The outcome of their interactions is what we want to explore. We can elaborate upon the story this way.

(1) The likelihood of renegotiation for dividing up quasi-rent ex post will be increasing as uncertainty facing the transactors increases, so uncertainty has a positive effect on the likelihood of renegotiation, labeled as +. The direct consequence of increasing the opportunities for renegotiation is to give the advantageous party the motive to bargain for extra share of trading gains ex post.

(2) Process specificity indicates how costly is it for the client to switch to alternative contractor to resume the disrupted project, so high process specificity naturally foretells the presence of large appropriable quasi-rent.

(3) Two factors jointly affect the site of transaction problems in rent-seeking behaviour. It is seems reasonable to model the severity of rent-seeking behaviour as the function of two variables: one is how lucrative is the opportunistic attempt; another is how likely it is to succeed. The former is determined by the degree of process specificity of the project, i.e., how large the appropriable quasi-rent would appear ex post. The latter is, to a large part, affected by the uncertainty of the project. The presence of quasi-rent is the ultimate reason for opportunism and uncertainty is only the trigger for making it happen.

![Fig. 7-4 The interaction of uncertainty with process specificity](image-url)
From the above analysis, the transaction cost arising from uncertainly can best be mitigated by the same means that will mitigate process specificity. In other words, ceteris parabus, the higher the high uncertainty of a project, the greater relative efficiency advantage of the management system.

7.4.3 Partnering agreement
As illuminated in Sec.7.2.7, the effect of partnering agreement on transaction costs relied upon its ability to mitigate the hazards from opportunism. Unlike the other five variables in Fig. 7-2, this is a binary variable, labeled as 1 for the case of the partnering agreement being used and otherwise 0. If this factor is taken into account by the client in choosing procurement systems, its effect would be to enlarge the range for which design-build can be desirable, that is, increasing the possibility of design-build being chosen as a whole. The reason lies in the nature of partnering agreement.

For most proponents of construction partnering, this method should stand in its own right. However, for convenience of the following analysis, we would like to view it as a supporting contractual arrangement for solving the mounting transaction problems appearing in the functioning of each procurement route. In using the traditional method and design-build, the main contractor is the principal party with whom the client intends to form the close collaboration relation by entering into the partnering agreement. In using management system, the designer or sometime construction management firm is the most frequent party with whom the client can effectively improve cooperation atmosphere by partnering. Now the problem is how much the original problems can be alleviated after partnering is introduced.

In the contractual arrangement of management system, the responsibilities of construction are so dispersed to several trade contractors that the efficacy of partnering agreement in constraining opportunism arising from measurement difficulties will be more blunted. In principle, the more concentrated the responsibilities in the procurement system, the more efficacious is the partnering agreement. Thus, we have a priori ground for claiming that the use of partnering will have much more obvious effect on design-build than on other two routes.
7.4.4 The scale of the client
As explained in Sec.7.2.8, the role of the client’s scale is an important factor in generating deterrence for opportunism, so the possible effect of this factor is similar to the use of partnering agreement. In principle, the more the expected construction expenditure in the future, the more able is the client to bargain with his/her agent to give up the present Appropriable rent in exchange for the future awarding of workload. The effect of alleviating short-termism by giving future works intensifies along with the scale of the client. As a result, the postulated order of relative desirability of procurement systems as this variable increase is: design-build > traditional method > management system.

7.4.5 Summary
To sum up, using Eq.7-1~7-3, in conjunction with the analysis in Chapter 6 and Sec.7.4, we can construct the conditions under which each procurement system is most suitable, and therefore most likely to be chosen by rational clients. The sets of conditions are as graphically presented in Fig.7-5.

**Fig. 7-5 Alignment of procurement routes – empirical predictions**

1. When (a) time value is very high, (b) uncertainty of the project is medium, (c) degree of process specificity is medium, (d) the scale of the client is medium, (e) the interface problems between subcontractors or trade contractors are not serious, management system is the most suitable procurement strategy and so more likely to be chosen.
2. When (a) time value is very low, (b) uncertainty of the project planning is low, (c) the degree of process specificity is low, (d) the scale of the client is medium, (e) the interface problems between subcontractors and trade contractors are average, **traditional method** is the most suitable procurement strategy and most likely to be chosen.

3. When (a) time value is high, (b) uncertainty of the project planning is very low, (c) the degree of process specificity is low, (d) the scale of the client is large, (e) the interfaces problems between subcontractors and trade contractors are very serious, **design-build** is the most suitable procurement strategy and so most likely to be chosen.

Apart from the five variables, partnering agreement can more effectively improve the relative desirability of design-build than traditional method and management system.

7.5 Development of Hypothesis

The section tries to transform the results concluded in Sec.7.4.5 into hypotheses. Combined with the variable of PARTR (use of partnering), six hypotheses can be derived. The effect of each variable on the selection of procurement systems is formulated in the form that "as X variable is higher, procurement option R will be more likely to be chosen relative to option Q, other factors being equal". The relation of one route being more likely to be chosen than another one is expressed as \( R > Q \).

7.5.1 Opportunity costs of time (OPPCT)

Opportunity costs of late delivery relate to the client’s sensitivity to the timing of completion. Higher OPPCT will enlarge the advantage of fast-track route in increasing (discounted) revenue but at the same time by increasing process specificity enhance the disadvantage of an integrated route. Thus, a proposition can be stated as follows:

**Hypothesis 1**

The higher is the opportunity cost of time, the more desirable and therefore the more likely for the client to employ a fast-track strategy for shortening the total time. However, higher opportunity cost of time also leads to more serious process specificity problems, making an integrated route less desirable. The joint effect of these two forces leads to the order of management system > design-build, traditional method.
7.5.2 Uncertainty (UNCER)
This variable is intended to encompass the exogenous factors that may cause the ne-
cessity of renegotiating the price and quality standards ex post. Provided the execu-
tion of the project is full of uncertainty, the client possibly will be forced to change
requirements and design. This kind of variation is initiated by the client himself,
rather than due to designer’s error. Whose fault is relevant here. If the liability is on
the designer side, the client can stave off this risk by choosing design-build. If not, the
client will face the problem of re-negotiation with contractors during construction.
This will aggravate the hazard of process specificity. In this case, an integrated route
like design-build may make the client vulnerable to be held up, increasing transaction
costs. According to the analysis in Sec.6.6, management system is more flexible than
traditional method and design-build. Thus, a hypothesis can be formulated as follows:

Hypothesis 2
The more uncertain factors are involved during design and construction, the more
likely for the client to need to renegotiate with contractors after contract is signed, and
the more likely for him/her to employ a flexible strategy. Thus, management system
> traditional method > design-build.

7.5.3 Degree of process specificity (DEGPS)
The degree of process specificity can be measured by the cost of switching. If an
alternative contractor cannot easily be found, the client would be in a weak position to
bargain with the current one. This situation would be worse, if the integrated route is
used. Accordingly, we can get the following proposition:

Hypothesis 3
The more costly to switch to a replacement contractor, were the project to be dis-
rupted, the more desirable and therefore the more likely for the client to adopt a frag-
mented route. Thus, management system > traditional method > design-build

7.5.4 Partnering agreement (PARTR)
As explained in Sec.7.4.3, the use of partnering agreement can more effectively miti-
gate the transaction hazards arising from rent-seeking for the project using design-
build than that using management system. Thus, the hypothesis will be:
Hypothesis 4
When a partnering agreement is used in the project, the more desirable, and therefore
the more likely for the client to adopt design-build. That is, design-build > traditional
method > management system

7.5.5 The scale of the client (SCALC)
As claimed in Sec. 7.4.4, the scale of the client, measured by the total construction
outlays in a period of time, can reduce the occurring likelihood of rent-seeking behav­
iour, so the hypothesis can be stated as follows:

Hypothesis 5
The more future works that the client can provide, the less threatening is the rent­
seeking behaviour and the more desirable is the integrated procurement system. That
is, design-build > traditional method > management system.

7.5.6 Complexity of specialisation (COMSP)
This variable is associated with how many number and types of subcontractors will be
involved in the project. The increase in the number of independent subcontractors will
increase the difficulty of monitoring quality in that there are many grey areas existing
between the responsibilities of subcontractors. When COMSP becomes severe, the
disadvantage of using management system is evident. Thus, we can infer that

Hypothesis 6
When the degree of ambiguity of interface between autonomous subcontractors be­
comes severe, a more integrated procurement route will be more desirable because it
is able to reduce measurement costs in identifying the liability of defects of the project.
Thus, design-build > traditional method > management system.

7.6 Testing of hypotheses
7.6.1 Econometric model
Among econometric models, qualitative response model is designed for the analysis
of discrete choices. A popular technique is the logit model [Greene, 1993]. In this
study, the client has three alternatives: traditional method, design-build, and manage­
ment system. Six main variables are identified as critical in affecting the client’s
choice: OPPCT, UNCER, DEGPS, PARTR, SCALC, COMSP. The logistic response
function used to represent the impact of these six effects is expressed below.
If we use the logit specification, the probability of each procurement route being chosen is as follows:

\[ P_{DB} = P(DB) = \frac{e^{\beta_1 X}}{1 + \sum_{i=1}^{2} e^{\beta_i X}} \]  
**Eq. 7-4**

\[ P_{MS} = P(MS) = \frac{e^{\beta_2 X}}{1 + \sum_{i=1}^{2} e^{\beta_i X}} \]  
**Eq. 7-5**

\[ P_{TM} = P(TM) = \frac{1}{1 + \sum_{i=1}^{2} e^{\beta_i X}} \]  
**Eq. 7-6**

where \( X \) is row matrix of independent variables, including OPPCT, UNCER, DEGPS, PARTR, SCALC, COMSP and a constant; \( \beta_1 \) and \( \beta_2 \) are column matrix of the estimated coefficients. Eq.7-4 ~ Eq.7-6 represent the probability that the client chooses design-build. Taking the natural logarithm of both sides in Eq.7-4 ~ Eq.7-6 yields the relation between the factors and the log odds ratio:

\[ \ln \left( \frac{P_{DB}}{P_{TM}} \right) = X\beta_1 \]  
**Eq. 7-7**

\[ \ln \left( \frac{P_{MS}}{P_{TM}} \right) = X\beta_2 \]  
**Eq. 7-8**

\[ \ln \left( \frac{P_{MS}}{P_{DB}} \right) = X(\beta_2 - \beta_1) \]  
**Eq. 7-9**

For example, Eq.7-7 indicates the relative probability of design-build being selected compared with traditional method. This probability is affected by six variables, so the complete form of Eq.7-7 is

\[ \ln \left( \frac{P_{DB}}{P_{TM}} \right) = \beta_1^0 + \beta_1^{OPPCT} + \beta_1^{UNCER} + \beta_1^{DEGPS} + \beta_1^{PARTR} + \beta_1^{SCALC} + \beta_1^{COMSP} \]  
**Eq. 7-10**

By the same way, we can get,
\[
\ln \left[ \frac{P_{DB}}{P_{RM}} \right] = \beta_2^0 + \beta_2^1 \text{OPPCT} + \beta_2^3 \text{UNCER} + \beta_2^4 \text{DEGPS} + \beta_2^4 \text{PARTR} + \beta_2^4 \text{SCALC} + \beta_2^4 \text{COMSP}
\]

\text{Eq. 7-11}

\[
\ln \left[ \frac{P_{DS}}{P_{DB}} \right] = (\beta_2^0 - \beta_1^0) + (\beta_2^1 - \beta_1^1) \text{OPPCT} + (\beta_2^3 - \beta_1^3) \text{UNCER} + (\beta_2^4 - \beta_1^4) \text{DEGPS}
\]

\[
(\beta_2^4 - \beta_1^4) \text{PARTR} + (\beta_2^4 - \beta_1^4) \text{SCALC} + (\beta_2^6 - \beta_1^6) \text{COMSP}
\]

\text{Eq. 7-12}

The coefficients can be estimated by maximizing the likelihood function:

\[
L(D_j | X_j; \beta) = \prod_{j=1}^{N} P(D_j = 1 | X_j)^{D_j} (1 - P(D_j = 1 | X_j))^{1-D_j}
\]

where \( N \) is the number of samples collected by survey.

### 7.6.2 A brief account of logit model

In the social science, lots of issues of interest involve limited discrete choices. Due to discontinuity of dependent variable, the traditional regression analysis is no longer applicable since the estimation of coefficients is not efficient. To see the reason, a simple example is given for illumination.

For example, we might be interested in the condition that makes a household own a car. In this case, the qualitative dependent variable can be defined as follows:

\[
D_i = 1 \quad \text{if ith household owns a car}
\]

\[
D_i = 0 \quad \text{if ith household does not own a car}
\]

For ease of exposition, suppose the income level is taken as the only explanatory variable. Thus, the linear model can be specified as

\[
\forall i \quad D_i = \beta_0 + \beta_1 I_i + \epsilon_i
\]

\text{Eq. 7-13}

Assuming \( E(\epsilon_i) = 0 \), we then have

\[
\forall i \quad E(D_i) = \beta_0 + \beta_1 I_i
\]

\text{Eq. 7-14}

The expected value of \( D_i \) has an intuitive interpretation. According to the definition,

\[
E(D_i) = 1 \times \Pr(D_i = 1) + 0 \times \Pr(D_i = 0) = 1 \times P_i + 0 \times (1 - P_i) = P_i
\]
That is, \( E(D_i) \) is simply the probability that household \( i \) owns a car. Thus, we can re-write Eq.7-14 as

\[
\forall i \quad P_i = \beta_0 + \beta_1 I_i
\]

Eq. 7-15

This linear probability model indicates that the probability of the \( i^{th} \) household buying a car is determined by substituting the income level into the regression equation. The estimation of Eq.7-14 can be done by OLS. However, there arise problems because the error term \( \varepsilon_i \) is binomial distribution rather than normal distribution. Using Eq.7-13, we can get

\[
\varepsilon_i = 1 - P_0 - P_{ii} \quad \text{if} \quad D_i = 1
\]

\[
\varepsilon_i = -P_0 - P_{ii} \quad \text{if} \quad D_i = 0
\]

Thus,

\[
Var(\varepsilon_i) = E(\varepsilon_i^2) = P_i \times \text{(value of } \varepsilon_i \text{ when } D_i = 1) + (1 - P_i) \times \text{(value of } \varepsilon_i \text{ when } D_i = 0)
\]

\[
= P_i \times (1 - \beta_0 - \beta_1 I_i)^2 + (1 - P_i) \times (-\beta_0 - \beta_1 I_i)^2
\]

\[
= P_i \times (1 - P_i)^2 + P_i \times (1 - P_i)
\]

\[
= P_i \times (1 - P_i)
\]

Obviously, the variance of error terms varies from household to household. This means that Eq.7-15 is heteroskedastic, violating the basic assumption of classical regression analysis that \( Var(\varepsilon_i) = \text{constant} \). Though the estimators themselves are still unbiased, the estimated standard errors will become unreliable.

A possible way to sort it out is to introduce a latent variable \( L \), defined by

\[
L^* = \beta_0 + \beta_1 I + \varepsilon_i
\]

In our example, \( L^* \) might be explained as an unobservable index of the willingness to own a car. Its expected value is

\[
E(L^*) = \beta_0 + \beta_1 I
\]

To link the unobservable \( L^* \) with observable outcomes, we can simply specify

\[
D_i = 1 \quad \text{if} \quad L^* > 0
\]

\[
D_i = 0 \quad \text{if} \quad L^* < 0
\]

To obtain the expression of \( P_i \), note that

\[
P_i = \Pr(D_i = 1) = \Pr(L^* > 0) = \Pr(\varepsilon_i > -(\beta_0 + \beta_1 I_i))
\]

Therefore, \( P_i \) depends on the way in which \( \varepsilon_i \) is distributed. A popular model is called logit model. The year 2000's Nobel laureate Professor MacFadden is credited as mak-
ing substantial contributions in developing the multinomial logit model, which is used in my study.

The basic form of logistic function is

\[ P_i = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 I + \varepsilon_i)]} \]

It is easy to see that

\[ 1 - P_i = \frac{\exp[-(\beta_0 + \beta_1 I + \varepsilon_i)]}{1 + \exp[-(\beta_0 + \beta_1 I + \varepsilon_i)]} \]

so that

\[ R_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 I + \varepsilon_i \] \hspace{1cm} \text{Eq. 7-16}

This is called odd ratio and its natural logarithm as the logit. If we put \( D_i \) directly into the logit \( R_i \), we get:

\[ R_i = \ln\left(\frac{1}{0}\right) \text{ if a family owns a car} \]

\[ R_i = \ln\left(\frac{0}{1}\right) \text{ if a family doesn’t own a car} \]

Obviously, these expressions are meaningless. However, matters change when we switch to predicting the behaviour of \( N \) actions. To give the odd ratio an intuitive interpretation, we can use the hypothetical example in Table 7-1 for explanation.

<table>
<thead>
<tr>
<th>( I_i )</th>
<th>( N_i ) (number of families at income ( X_i ))</th>
<th>( n_i ) (number of families owning a car)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>40</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>60</td>
<td>18</td>
</tr>
<tr>
<td>13</td>
<td>80</td>
<td>28</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>45</td>
</tr>
<tr>
<td>20</td>
<td>70</td>
<td>36</td>
</tr>
<tr>
<td>25</td>
<td>65</td>
<td>39</td>
</tr>
</tbody>
</table>
As this table shows, corresponding to each income level $X_i$ there are $N_i$ families, $n_i$ among whom own a house. Therefore, if we now compute the relative frequency

$$\hat{P}_i = \frac{n_i}{N_i},$$

which is an estimate of the true $P_i$ corresponding to each $N_i$. Using the estimated $P_i$, we can obtain the estimated logit as:

$$R_i = \ln\left(\frac{\hat{P}_i}{1-\hat{P}_i}\right) = \hat{\beta}_1 I_i + \epsilon_i$$  \hspace{1cm} \text{Eq. 7-17}$$

In short, given the grouped or replicated data, one can obtain the dependent variable, the logits, to estimate Eq. 7-17.

Conversely, if we find a household with the income level of $I_i$, the probability of this household having a car can be computed by substituting $I_i$ into Eq. 7-17. Suppose we get a result, say, 0.6, then what does this prediction mean? It means, if ten families have the income level of $I_i$, we expect to find six families owning a car.

Table 7-2 The actual observations and predicted results of a hypothetical example

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Families own cars</th>
<th>Families don’t own cars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Families own cars</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Families don’t own cars</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>35%</td>
<td>65%</td>
<td></td>
</tr>
</tbody>
</table>

We can hypothetically tabulate the actual and predicted distribution of whether families own cars in Table 7-2. There are 11 $(4+7)$ out of 20 successfully classified. What does this success rate mean? That shows that if we collected 100 families with the income level of $I_i$, the outcome (owning a car or not) could be correctly predicted by the model for 55 of them. In reality, what we observe is a single outcome, owning a car or not, at a first glance, probability prediction seems meaningless. However, we
shouldn't forget that the prediction of logit models is based on grouped data or hypothetical repeated observations. That is, if we had another 100 samples grouped together according to the same income level, the logit model can help us correctly identify 55% of the samples in the ex ante sense.

7.6.3 How to verify hypotheses by way of logistic model
This subsection aims to explain how to test hypotheses listed in Sec.7.5 by logit model. First of all, the hypotheses state a logic that for a given attribute, a procurement route is desirable relative to others due to its superior efficacy in dealing with this kind of project so that the client's objective can be best satisfied. What the logit model can do is corroborate the statistical relation between the probability of each procurement route being used and the attributes of transaction. Relative higher frequency of a procurement route being selected means that that route is preferable to the client for the given conditions. Under the assumption that the client is rational, the frequency of a procurement route being used is tantamount to showing the extent that this procurement route can satisfy the client's objective for the given attributes. Thus, the higher frequency of a procurement route relative to others means that it is the most efficient procurement route under that set of conditions. The logic can be summarised as follows: (1) theoretical prediction that higher factor A makes procurement route $PR_i$ relatively desirable, and (2) the client is assumed to choose the most efficient route, so (3) higher factor A will increase the likelihood of procurement route $PR_i$ being chosen, meaning (4) the coefficient of factor A will be positive in the log odd ratio of $PR_i$ relative to other routes. This is, if the higher value of a variable leads to the increased likelihood of traditional method being chosen relative to design-build, the sign of coefficient of this variable will be predicted to be positive. In contrast, if the higher value of this variable causes the log-odd ratio to reduce, the sign of coefficient of this variable will be expected to be negative. Following this logic, the logit model of Eq.7-10 ~ 7-12 can be employed to test the six hypotheses listed in Sec.7.5.

For example, from hypothesis 1 ~ 6, we can know that design-build is more favourable than traditional method for the project of lower uncertainty, lower process specificity and higher measurement difficulties and when the partnering agreement is used and the scale of the client is larger. Transforming this set of hypotheses into the predicted sign of coefficients of variables in Eq.7-10 where design-build is compared
with traditional method gives five predictions as shown in Table 7-3. On the basis of the similar reasoning, the predicted sign of coefficients of each variable in Eq.7-11 and Eq.7-12 can be found, as shown in Table 7-4 and Table 7-5.

Table 7-3 Expected results of logit estimation: Eq.7-10

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Expected sign</th>
<th>Hypothesis tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>+/-</td>
<td>Hypothesis 1</td>
</tr>
<tr>
<td>OPPCT</td>
<td>-</td>
<td>Hypothesis 2</td>
</tr>
<tr>
<td>UNCEP</td>
<td>-</td>
<td>Hypothesis 3</td>
</tr>
<tr>
<td>DEGPS</td>
<td>+</td>
<td>Hypothesis 4</td>
</tr>
<tr>
<td>PARTR</td>
<td>-</td>
<td>Hypothesis 5</td>
</tr>
<tr>
<td>SCALS</td>
<td>-</td>
<td>Hypothesis 6</td>
</tr>
<tr>
<td>COMSP</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

Table 7-4 Expected results of logit estimation: Eq.7-11

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Expected sign</th>
<th>Hypothesis tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>+</td>
<td>Hypothesis 1</td>
</tr>
<tr>
<td>OPPCT</td>
<td>+</td>
<td>Hypothesis 2</td>
</tr>
<tr>
<td>UNCEP</td>
<td>+</td>
<td>Hypothesis 3</td>
</tr>
<tr>
<td>DEGPS</td>
<td>+</td>
<td>Hypothesis 4</td>
</tr>
<tr>
<td>PARTR</td>
<td>-</td>
<td>Hypothesis 5</td>
</tr>
<tr>
<td>SCALS</td>
<td>+</td>
<td>Hypothesis 6</td>
</tr>
<tr>
<td>COMSP</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table 7-5 Expected results of logit estimation: Eq.7-12

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Expected sign</th>
<th>Hypothesis tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>+</td>
<td>Hypothesis 1</td>
</tr>
<tr>
<td>OPPCT</td>
<td>+</td>
<td>Hypothesis 2</td>
</tr>
<tr>
<td>UNCEP</td>
<td>+</td>
<td>Hypothesis 3</td>
</tr>
<tr>
<td>DEGPS</td>
<td>-</td>
<td>Hypothesis 4</td>
</tr>
<tr>
<td>PARTR</td>
<td>-</td>
<td>Hypothesis 5</td>
</tr>
<tr>
<td>SCALS</td>
<td>-</td>
<td>Hypothesis 6</td>
</tr>
<tr>
<td>COMSP</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
8 Development Of Questionnaire And Survey Design

Chapter 7 concludes the stage for searching for relevant variables that are most likely to affect discounted revenue, transaction costs from process specificity and measurement difficulties of the project, i.e., going through stage I to stage II in Fig.7-1. This chapter continue the development from stage II to stage III for dealing with measurement of variables.

8.1 Development of measurement

8.1.1 The process of questionnaire design

Using a semantic scale to ask for the respondent’s subjective evaluation can easily lead to errors without careful design. Every effort has to be made to minimise the possible sources of errors. Churchill (1995) suggests a sequence of steps that can be followed to develop valid measures (as shown in Fig. 8-1). Corresponding to the purpose of this study, we can formulate the design procedure as follows:

Fig. 8-1 A suggested procedure for developing measures [Churchill, 1995]

Step 1: Specify domain of the measure

Step 2: Generate sample of items

Step 3: Collect data

Step 4: Purify measure

Step 5: Assess validity

Step 1: Review literature to find the multifaceted dimensions of the six variables used in the hypotheses.

Step 2: Attempt to develop the first version of questionnaire so as to capture the multifaceted dimensions of the six variables.

Step 3: Pilot that questionnaire to test the answerability of the questions and modify it to generate the second version of questionnaire with reference to feedback opinions.
Step 4: Conduct the survey of the second version of questionnaire and analyse the result with an aim to improve the validity of measurement of variables by testing the reliability of set questions that are devised to measure the different aspects of a variable and eliminate the questions with low correlation with others.

In implementing these four steps, this study went through a three-stage process for questionnaire design and survey design (see Fig. 8-2). At the first stage, the questions were sent to five clients that have close connection with my tutors at UCL with the aim of finding two things: (1) what kind of questions can be answered and (2) whether the wording of questions are clear enough to avoid confusion. This step took one month to get back questionnaires and complete the analysis. After communicating with the respondents, the phrasing of the questions was modified and some questions are deleted that were hard to answer according to the received comments.

Fig. 8-2 The flow chart of questionnaire and survey

Then, at the second stage, 100 clients were selected randomly from Glenigan database to test the content validity of the questionnaire. Eliciting responses from larger samples can provide additional information. Of the 35 returned questionnaires (i.e. return rate is 35%), 15 projects (43%) used traditional method, 14 projects (40%) used design-build, and the rest 6 projects (17%) used management system. Though it is found that categorising procurement systems only into three types is too crude, we still can

---

1 It should be noted that this return rate is obtained after follow-up actions, like phone calls.
get some basic idea about the approximate proportion of procurement systems currently used. At the final stage, the full-scale survey was conducted. Three months is spent for preparing and administering the survey. The results of the survey are discussed in Chapter 9. This chapter just focuses on the process of finding sensible means of measurement of the variables.

8.1.2 Demands for new microeconomic data in TCE
Operationalisation is a critical step in TCE to derive refutable hypotheses. By this term we mean the process of "relating the relative merits of alternative institutional arrangements to observable attributes of transactions" [Masten, 1995]. According to this definition, discovering the factors that are intimately linked with the relative performance of governance and that can be measured practically will be our goal. If we successfully achieve it, a set of variables can be obtained. The problem then is how to measure these variables. As will be seen, these variables mainly fall into the category of qualitative variables, distinct from quantitative variables, such as price, interest rate often used in the economic empirical studies. Because this type of variable is liable to be criticised as lacking objectivity due to the errors resulting from personal judgement, many social science disciplines, such as psychology, marketing and sociology have been struggling to cope with them, and generated a vast stock of useful techniques. A sensible starting point is to review psychometric theory of measurement errors [Nunnally & Bernstein, 1994] and attempt to employ some guidelines for questionnaire design prevalently accepted in management studies, such as Churchill (1995) and Kinnear & Taylor (1991).

As aforementioned, TCE attempts to develop a refutable theory by analysing (and measuring) microeconomic features of economic organisation. In contrast to other branches of economics, there are few quantifiable variables readily available to the study of organisation. In face of this situation, Simon's perceptive analogy between physical science and economics provides a good starting point of thinking:

*In the physical sciences, when errors of measurement and other noise are found to be of the same order of magnitude as the phenomena under study, the response is not to try to squeeze more information out of the data by statistical means; it is instead to find techniques for observing the phenomena at a higher level of reso*
lution. The corresponding strategy for economics is obvious: to secure new kinds of data at the micro level. [Simon, 1984]

For researchers who are interested in organisation, the difficulties in variable measurement must be somehow sorted out. After over two decades of development, a phenomenal burgeoning of academic interest springs out of this issue. After examining nearly 200 pieces of TCE empirical work with topics ranging from vertical integration, hybrid contracting modes, long-term commercial contracts, informal agreements, and franchise contracting, Shelanski and Kline (1995) observe that

[t]ypically, these characteristics (asset specificity, uncertainty, frequency) are estimated based on surveys or interviews: for example, a manager might be asked to rate on a Liker-type scale of 1 to 7 the degree to which an investment has value in outside uses.

Whilst there is no denying that 'such measurements are of course subject to the general limits of survey data', the potential errors can be greatly reduced through meticulous design of survey questions. Generally speaking, this is often relied upon as the last resort for the cases where the variables of interest are too soft to be quantified or a way of quantification is not available yet. The situation facing this study is the latter. While most of the variables used in this study, such as uncertainty, process specificity and complexity, are quite soft, we don’t need to be too pessimistic about the possibility of quantifying these variables. As far as the current states of art are concerned, many attempts have been made to measure this kind of variables (see Sec.8.1.3). Certainly, more efforts should be constantly directed to improve the quality of measurement. However, this issue is beyond the scope that this study can cover\(^2\). At this point, it seems sensible to say that, to find a practical solution for variable measurement, a compromise has to be made between practicality and precision. This is the main reason for this study to adopt the discrete (semantic differential) scale in the questionnaire design.

8.1.3 The established practice in TCE

An effort is made here to summarise the way that the variables associated with this study are measured in the TCE literature. The mainstream approach to the design of questionnaire mainly follows the line of scale rating or classification, such as Monte-

\(^2\) A further exposition of variable measurement please refers to Sec.10.5.2.

**Table 8-1 The measurement of variables in TCE**

<table>
<thead>
<tr>
<th>Main issues of this study</th>
<th>Key variables in TCE</th>
<th>Proxy variables</th>
<th>Measurement</th>
<th>Source of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monteverde &amp; Teece (1982)</td>
<td>Buy-or-make decisions for car components</td>
<td>Human asset specificity</td>
<td>Engineering effort</td>
<td>10-point scale rating from &quot;none&quot; to &quot;a lot&quot; of engineering investment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Component specificity</td>
<td>Dummy variable (if the component is specific to a single assembler)</td>
<td>A replacement parts wholesaler</td>
</tr>
<tr>
<td>Masten (1984)</td>
<td>Buy-or-make decisions for aeroplane components</td>
<td>Asset Specificity</td>
<td>Dummy variable (if the item is highly specialised)</td>
<td>Rating system used internally by the company</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uncertainty</td>
<td>Complexity</td>
<td>Dummy variable (if the item is rated as complex)</td>
</tr>
<tr>
<td>Masten, Meehan &amp; Snyder (1991)</td>
<td>Make-or-buy decision in the naval shipbuilding industry</td>
<td>Temporary specificity</td>
<td>Schedule</td>
<td>A ranking of the importance of having the component or performing the task on schedule (1 to 10)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human asset specificity</td>
<td>Worker's knowledge</td>
<td>the degree to which skills, knowledge or experience of workers are specific to this application</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Engineering effort</td>
<td>A ranking of the amount of engineering effort involved in developing the component</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dedicated specificity</td>
<td>The degree to which facilities and equipment used in the production process are specific to this application</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uncertainty</td>
<td>Complexity</td>
<td>A ranking of the complexity of the component or task</td>
</tr>
<tr>
<td>Anderson &amp; Schmittlein (1984)</td>
<td>Choice between independent and integrated sales forces</td>
<td>Human asset specificity</td>
<td>Nature of the company</td>
<td>Measured as a 7-point scale from &quot;disagree&quot; to &quot;agree&quot; in response to the statement &quot;It's difficult to learn all the ins and outs of our company that a salesperson needs to know to be effective&quot;.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Nature of the product</td>
<td>Extra number of weeks of training that a new salesperson needs to familiarise the products</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Confidential information</td>
<td>Measured as a 7-point scale from &quot;disagree&quot; to &quot;agree&quot; in response to the statement &quot;An experienced salesperson's inside information could do us a lot of damage if it got out.&quot;</td>
</tr>
</tbody>
</table>

200
<table>
<thead>
<tr>
<th>Nature of the customer</th>
<th>Measured as a 7-point scale from “disagree” to “agree” in response to the statement “to be effective, the salesperson has to take a lot of time to get to know our accounts.”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer loyalty</td>
<td>Measured as a 7-point scale from “disagree” to “agree” in response to the statement “Personal relationships between our salespeople and accounts have little influence on sales of our product line.”</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>Expected deviation between forecast and actual sales in the next year in percentage</td>
</tr>
<tr>
<td>Environmental unpredictability</td>
<td>Measured as a 7-point scale from “disagree” to “agree” in response to the statement “It is very difficult to measure equitably the results of individual salespeople.”</td>
</tr>
<tr>
<td>Human asset specificity</td>
<td>How much time is required for a newly hired salesperson with experience in the industry to become adequately familiar with your product and customers? (months)</td>
</tr>
<tr>
<td>Choice between direct or indirect marketing channel</td>
<td>Measured as a 3-point scale in response to “How would you describe these products compared to other products in general?” in terms of (1) stable v.s. volatile market shares (2) easy to monitor trends v.s. difficult to monitor trends (3) stable industry volume v.s. volatile industry volume (4) sales forecasts are quite accurate v.s. sales forecasts are quite inaccurate (5) predictable v.s. unpredictable</td>
</tr>
<tr>
<td>Sales Manager</td>
<td>What is the typical time between an initial contact concerning the product and the ultimate placement of an order? (months)</td>
</tr>
</tbody>
</table>

From Table 8-1, it is clear to find that the measurement of asset specificity or uncertainty mainly relies on the evaluator’s perceptions instead of objective observations. To successfully apply this practice, how to minimise measurement errors is the first
and foremost objective in the design of survey. To explore this issue, in the following
subsections, we will examine some methodological issues.

8.1.4 Selection of an appropriate measure scale

Psychometric theorists have developed measurement techniques that can minimise
errors in measuring uncountable variables. They claim that measurement errors are
inevitable, but can be reduced by devising a composite measure of the variable.
[Nunnally & Berstein, 1994]. According to the classic measurement theory, "if we
take not one but a set of items all designed to measure the same underlying attribute
or continuum in a variety of different ways, and if we score these items and add up the
scores for each respondent, then this total score will be a better, more reliable meas­
ure," [Oppenheim, 1992] in that random errors can be cancelled out. As a result, ap­
proaching an abstract concept from different aspects can help reduce errors.

A good composite measurement can be “reliable” if all items of the measure can
maintain internal consistency. Typically, reliability is assessed by computing
Cronbach’s alpha, an indicator ranging from 0 to 1. By convention, values over 0.7
are widely considered adequate for basic study. If the alpha is lower than that
threshold, the next step that should be taken is to carry out the factor analysis, with
which the “garbage items” can be identified [Churchill, 1979]. By resorting to this
theory as theoretical foundation, scaling method is widely used in the social sciences.
The selection of scale mainly depends on the statistical method used in the analysis.
In principle, for research only focusing on the association of variables, 5-point,
6-point and 7-point scales are often applied. However, if the study uses econometric
tools, like logit, probit model, higher-resolution scales are required. Thus, many of
important TCE empirical studies (as listed in Table 8-1), like Monteverde and Teece
(1982), Masten et al. (1991) take this method. Following this tradition, this study also
adopts a 10-point scale to measure the variables.

8.2 Measurement of variables

Just as most TCE empirical literature, a majority of the variables in this study are not
easy to obtain objective estimates. In determining the way these variables should be
measured, we encounter a trade-off between reliability and availability of the data. In
principle, if there are lots of alternative choices of measures, objective variables are the first priority, subjective direct measure is the next, and the last one is subjective indirect measure.

To demonstrate the practical importance of these measures, the present study heavily relies on the survey information from “Faster Building for Commerce” conducted by Building Economic Development Council (1988), in which information is collected by the case study of 60 commercial projects and survey from 260 projects, covering a wide range of buildings, such as speculative office, sports centre, warehouse, supermarket, shopping mall, hotel, restaurant. By reference to that report, it can be shown that some assumptions as to matters of fact taken in this study are representative of results from the actual observations in construction practice.

The final version of the questionnaire is exhibited in Appendix. In the following six subsections, we will discuss why the questions are formulated as they are.

8.2.1 Opportunity costs of time (OPPCT)

8.2.1.1 Key dimensions
Although conceptually OPPCT can be expressed in a precise formula (see Eq.6-19), the exact number is hard to obtain from the client because in most cases, it is either not estimated or kept as confidential information. However, the client does know how critical it is to set the target of completion date and keep the programme of the project on schedule. It is manifested in two respects: one is to require completion before a targeted time, e.g. opening before Christmas; another is to require project delivery as early as possible, such as factories. Both cases will make difference in the date of completion become the client’s serious concern.

It is generally observed that the client for commercial buildings tends to put four factors as top priorities in assessing the urgency of the project [BEDC, 1988]:
(1) longer periods with capital tied up, incurring extra bank borrowing charges or loss of return;
(2) loss of business;
greater uncertainty in managing business assets when opening dates or letting
dates cannot be fixed;
(4) opportunities for the competition to get ahead.

These four points are the key determinants affecting the client’s perception on time
value in executing the project. Thus, in the pilot study, an attempt was made to cap-
ture these dimensions by asking the respondents to answer the following questions
with reference to the scale of 1 to 10:

To what extent do you agree with the following statement:
1. To what extent will the revenue of your business be negatively influenced by later completion (than your targeted date)?
2. To what extent will your competitive position be influenced by later completion (than your targeted date)?
3. To what extent will the prolonging of uncertainty in asset management caused by later completion (than your targeted date) influence your company.
4. Please indicate to what extent will the capital tied up in this project affect your organisation if the project cannot be commissioned at your targeted date.

From the responses, we found that these questions encounter a common difficulty that
the importance of these four dimensions of the time value available have to be meas­
ured against a baseline – the targeted date of completion. This is the site of the prob­
lem. When the early project delivery would lead to much higher discounted revenue
stream, delay must be something that the client is keen to avoid. And the intensity of
the client’s aversion to delay should be higher as the four dimensions become non­
trivial. However, the result is not quite as expected. This is because the selection of
target date is largely based on both the operational requirement of the project and the
procurement system used. Once the demand for earlier completion has been factored
into the targeted date, the severity of delay won’t necessarily be able to correctly re­
fect the client’s true valuation of time of the project in question. That is, what really
needs to be explored is why the targeted date is determined as it is. But how? We need
to recognise one thing: there is a wide range of clients with different background and
needs³, so we have to find a way of measurement that can be easily understood and
answered by the respondents and that can fully capture the theoretical meaning of
OPPCT. Finally, it was decided to use the concept of “the client’s willingness to pay

³ Commercial and industrial buildings include office, shop, supermarket, club house, hotel, food retail
warehouse, shopping centre, restaurant, factory, and so on.
in exchange for earlier completion” as the proxy for OPPCT. Actually, it has been
tried at the first stage of pilot study to formulate the questions in the quantitative way
by asking, like

*Please indicate the amount (as % of contract price) by which, at pre-contract stage, a
bid would have had to be reduced in order to compensate for it being for a contract
period one month greater. (£/day)*

Unfortunately, the comments returned on the practicability of this question were quite
negative, mainly because the answer is at best a guess, potentially involving substan-
tial extent of error. Thus, this kind of question is discarded. In the full-scale survey,
we instead used semantic differential scale so that the client’s willingness-to-pay can
be measured by the following questions:

**To what extent do you agree with the following statement:**

1. You are willing to pay a higher price in exchange for earlier completion of this
project.
2. Getting the project completed as early as possible, even with higher costs, would
do much good for the success of this project.
3. Earlier delivery of the project will greatly increase the positive benefits of this
project to your organisation.
4. The benefits of getting the project completed earlier would by far outweigh the
extra costs of achieving it.

**8.2.2 Uncertainty (UNCER)**

Looking up construction risk management textbooks, we can find that lots of kinds of
uncertainties may appear in the construction process. However, only those that will
worsen rent-seeking behaviour should be considered. According to the doctrine of
TCE, uncertainty causes much trouble due to its interaction with asset specificity. In
chapter 6, it is shown that the presence of design rent and quality rent would give rise
to substantial transaction costs. Rent-seeking behaviour on the contractor side may
take the form of asking for increased payment or forcing the client to accept lower
quality level. Then this problem is, under what conditions will these situations be
more likely to happen? The main reason lies in the possibility of variations or ex post
design changes.

In general, variations refer to ‘works which are not expressly or impliedly included in
the original contract and therefore are not included in the contract price. whether they represent a change or alteration of the original work or simply an addition to or omission of it' [Wallace, 1994]. The occurrence of variations often renders the economic rent of all participants liable to be redistributed. If adjustments to deal with the new situations are inefficient, the relation between the client and his/her agents will be more confrontational. According to BEDC's investigation, 'on several projects apparently superficial changes requested by the customers caused inordinate upheaval and extensions to programme' [BEDC(1988), p.76]. In fact, it has been well documented that in practice a large part of disputes arise from variations [Chan and Yeong, 1995 and references there]. As a result, a tentative conclusion can be made that high expectation or possibility of variations happening in the course of construction will make the flexible route (like management contracting) more desirable. That is, the differential capability between procurement routes varies as uncertainty. The more uncertain is the project at the moment procurement route is selected, the more desirable to use management system.

Then, what sources of uncertainty can be considered as contributing factors that make the client change his/her original requirements? There are six categories of factors to be pointed out in the following subsections: (1) speed of obsolescence of original design, (2) uncertain site conditions, (3) volatility of expected revenue from project, (4) regulatory uncertainty, (5) change of budget, and (6) design complexity of the project.

8.2.2.1 Speed of obsolescence of original design

'Commercial [and industrial] building operates in a fast moving and competitive environment. Developments in process technology, information technology, fashion, etc, may cause building users to change their requirements for accommodation and facilities quite suddenly. As a consequence, many customers consider it necessary to make changes to the design ... even late into construction.' [BEDC(1988), p.101] For the client falling into this category, the design of the project should be carried out in a sequential way, namely that the procurement route should make allowance for the flexibility in accommodating the update requirements from the market. To elicit the client's perception of the need to alter design due to the changing market demand, the question is asked in the following way:
To what extent do you agree with the following statement:

*It is highly likely that the original brief of this project may need to be changed to cope with changing external market requirements after the engagement of the detailed designer.*

### 8.2.2.2 Uncertain site condition

The high frequency of unforeseen ground conditions appearing during construction may pose serious challenge to construction management. One in two projects surveyed in BEDC(1988) are adversely affected by this kind of uncertainty. Two principal reasons are underlined: (1) 'a site survey had failed to detect the problems’ and (2) 'the close follow-on of redevelopment after demolition did not have a thorough investigation of the site' [p.77]. Another two sources of disturbance from site are interference with utility lines [Kumaraswamy, 1997] and restricted access to the site [Semple et al., 1994]. In most cases, these problems just hamper construction and won't be serious enough to require the design modified. These two aspects can be comfortably ignored.

In the event that site conditions are Shacklean surprises, there is no way for the client to foretell them in advance. However, the client would be able to take stock of the ground uncertainty by the information that he/she possesses. In principle, an extensive ground survey can reduce the likelihood of altering foundation design. Thus, the following question is asked to elicit the client’s awareness of this type of uncertainty:

To what extent do you agree with the following statement:

*It is highly likely that this project may encounter ground conditions that may change the original brief after the engagement of the detailed designer.*

### 8.2.2.3 Volatility of revenue from project

The scale of the project is chiefly dependent on the expected revenue from it. If the revenue is too volatile to be foresee in advance, it seems imperative to reserve more or less flexibility in scaling up and down the project during design and construction. Thus, this factor can be measured in the following way:
To what extent do you agree with the following statement:

The expected revenue stream from the project, compared with other commercial or industrial buildings in general, is highly likely to be revised during the construction period.

8.2.2.4 Regulatory uncertainty

In Britain, there are three main types of statutory regulations on commercial and industrial buildings: planning and building regulation approvals and fire certificates. Let's consider them in turn.

The purpose of imposing planning control is to assure harmonisation of the new building with its neighbouring environment. In particular, in the central metropolitan city, the supply of easy-to-build-on sites falls short of demand. Many projects need to overcome some problems with regard to meeting civic concerns, such as noise, traffic control caused by construction reasons. Strong possibility of opposition would increase the possibility of changing the plan ex post. 'Developers were usually aware of the difficulties and risks attached to particular sites' [BEDC(1988), p.96]. This awareness will make integrated route less desirable in that a flexible route can modify design more easily.

Building control may result in uncertainty to the client as well. Most of the projects won't face difficulties in getting approval, while the projects with uncommon and innovative components would be different. 'Problems arose mostly in connection with new designs, such as atria, or with the use of materials that could be associated with known failures of fire and health risks – notably steel frames, claddings, glass, plastics – where building officers proved difficult to convince, even in the face of well substantiated arguments' [BEDC(1988), p.97].

Lastly, the different interpretation of fire control regulations would pose uncertainty on the acceptability of the project design. For example, stricter fire safety standard required by the officer would lead to 'redesign of foundations and changes in the configuration of the frame, amendments to roof supports and, more frequently, increases in the space needed to accommodate plant and services' [BEDC(1988), p.77].
In the pilot study, we separated the uncertainty arising from statutory regulation into three parts, but afterwards we found that doing so places too much weight on this factor (3 out of 9 questions). Thus, the three questions are cut down to a single general question:

**To what extent do you agree with the following statement:**

*It is highly likely that the enforcement of statutory regulations (such as planning, building and fire regulations) in the region of your site may require the original brief to be modified after the engagement of the detailed designer.*

### 8.2.2.5 Change of budget

Budget is one of the most common constraints in carrying out the project. The exact amount of money can be fixed only when the detailed design of the project is done. Some clients have flexible budget for the construction costs, while some set a fixed upper limit. In general, the budget control of the project varies project by project. Provided the budget can be flexibly adjusted, then there should be no problem. However, if the budget limit becomes more stringent during construction, the need to change the requirements in the brief will result and the vicious cycle of rent-seeking behaviour will be triggered. Thus, the factor has to be considered. The question is phrased as follows:

**To what extent do you agree with the following statement:**

*It is highly likely that you may face a more stringent budget constraint than originally foreseen that may cause you to change the original brief after the engagement of the detailed designer.*

### 8.2.2.6 Complexity of design

Design-inspired variations can yield impacts on the project in two aspects: (1) deficiencies in the completeness and coherence of design information; (2) buildability of the design [BEDC (1988), p.66]. Both of them may raise the possibility of change order during construction. Contractors often show grave complaints on poor design, such as ‘fiddly details, awkward positioning, impractical material combination, difficult injunctions’ and so on [p.69]. The problem seems to stem from that fact that ‘[designers] appeared not to be sufficiently aware of the likely dimensional tolerances and variations in appearance of the specified products’ [p.69]. For instance, ‘[the] shape,
size and colour of hand-made and stock bricks, required for decorative uses, proved less controllable... than the qualities of machine-made bricks and they proved difficult to accommodate within the confines of the frames and openings' [p.69]. This is the case with the problem arising from special visual requirements.

These observations in practice can help us think out what factors may make design more likely to be unbecoming ex post? Generally speaking, when some part of the design is an innovative attempt, its quality (including buildability) should be appraised not only by design drawings, but also by the actual outcome of the project. However, the outcome is the joint function of designer’s and contractor’s efforts. Once something goes wrong, who should take up the responsibility would arouse disputes.

From literature, several factors are mentioned as conceivable sources that may increase design variations. Time for design [BEDC,1988], advanced construction technology [Molenaar and Songer,1998], demanding visual requirement of building appearance [Turner,1997]. For these reasons, this factor is divided into four questions:

**To what extent do you agree with the following statement:**

1. It is highly likely that design variations would happen during construction in terms of the general characteristics of your project.
2. The design of this project was expected to be very complex.
3. The construction technology to be employed in your project was anticipated to be relatively advanced.
4. Roughly estimate what (value) percent of the project will contain the innovative elements that haven’t been built ever?

**8.2.3 Degree of process specificity (DEGPS)**

Costs of switching represent the potential additional costs that the client may incur when construction process is disrupted. The magnitude of these costs is affected by the degree of difficulties in replacing another competent designer and contractor in the event that the contractual relation breaks up. For most of the respondents may not have experiences in searching for replacement agents, the question should be designed in a more indirect way. It seems sensible to assume that the more easily for the client to find agents before tendering, the more easily to find replacement agents if necessary. It is implied that the thickness of the market and specificity of technology would
be two determinants. There are four questions devised to capture these two dimensions.

**To what extent do you agree with the following statement:**

1. *It is difficult to find a designer competent to undertake the detailed design for the project.*
2. *It might be difficult to find a contractor competent to undertake your project.*
3. *The contractor will need to have specific construction technology to undertake your project.*
4. *The construction technology that your project was perceived to need is not easily available from the UK construction industry.*

### 8.2.4 Partnering

The measurement of this variable is straightforward, asking for the client to tick whether the project in question has a partnering agreement at work. Hence, the question is as follows:

*Did you use a partnering agreement in this project?*

- [ ] No
- [ ] Yes, informal agreement
- [ ] Yes, formal agreement

### 8.2.5 Scale of the client

This variable tries to show the total construction expenditure that the client spent in the last five years, assuming that this factor is a good proxy for predicting the client’s future construction demand. A quantitative measure is attempted as follows:

*Your organisation’s total expenditure on construction projects in the last five years? (please fill in approx. number) __________ (£m)*

### 8.2.6 Complexity of specialisation (COMSP)

Before tackling the question as to what will be the appropriate proxy variables for COMSP, let’s briefly review what it is that this variable tries to capture. As claimed in chapter 5, measurement costs are one of the tri-pillars in the explanatory framework. From chapter 6, we can realise that COMSP serves as the factor for distinguishing the relative competence of the three procurement routes in dealing with measurement problems. Put another way, when COMSP becomes trivial, there is no difference in measurement costs under three routes. This is the essence of COMSP. However, to get the practical measures, it is essential for us to answer what conditions will make
the output of the project equally easy to be inspected regardless of the way designer, main contractor or subcontractors are organised. The approximate answers are simplicity and small scale. A project is said to be simple in the sense that the design of the building is commonplace, namely that the similar buildings have been constructed many times so that it will be considerably easy to find a qualified designer and contractor to undertake this project. Small-scale projects refer to projects with low construction costs. When the project is both simple and small-scale, the quality of design and construction is relatively easy to be defined in that there are lots of similar cases for reference. Thus, the benefits of single point of responsibility (i.e. design and build route) in reducing measurement costs will turn out less vital. When the scale of the project becomes larger, the more labour-only subcontractors and specialist trade contractors will involve during construction, the larger is the grey areas between the individual output of trade contractors or subcontractors, and the more difficult to delineate the liability of designer, contractor and subcontractors. This inference can be demonstrated from another perspective. As observed in BEDC(1988): ‘the smaller the number of subcontractors, the easier it was to achieve good performance’ [p.8]. This implies that small number of autonomous participants can alleviate the degree of interference from other parties. Stronger control on one’s own programme without need to co-ordinate with other subcontractors can enhance performance on the one hand, and reduce costs of defect correction for the client on the other. This perspective is consistent with the general observation that ‘the increasing sophistication of buildings has led to a more complex set of building requirements and often to many subcontractors each having some design responsibilities, thus increasing problems of project coordination and the task of determining liability should anything go awry’ [BEDC(1988), p.19].

In principle, this variable can be measured in two ways. The first idea is that the unit scale of subcontractors is an indicator for showing the complexity of specialisation between subcontractors, so we want to know how many types and numbers of traders are expected to involve in the project and divide it by the scale of the project. The questions are formulated as follows
1. On the basis of your preliminary design, how many types of principal trade contractors or subcontractors did you think would be involved in the production process according to the prevailing construction practice? ____________________________
(please fill in number)

2. On the basis of your preliminary design, what was your estimate of the total number of different principal trade contractors or subcontractors that would be involved in the production process? ________________
(please fill in number)

However, it is found in the pilot study that this measure is subject to two problems: first, the number of subcontractors or trade contractors is not an intrinsic characteristics of the project, but determined by managerial considerations, so it may not be able to fully reflect the theoretical meaning of this variable; second, it is relatively hard for the client to provide an accurate estimate of this variable, which is evidenced by the high proportion of incomplete questionnaires due to this question. As a result, in the full-scale survey, the attempt of quantitative measure is given up and semantic-differential questions are instead used by asking:

**To what extent do you agree with the following statement:**

1. It might be difficult to clearly identify the liability of subcontractors (or trade contractors) involved in this project due to complicated interfaces between their works.

2. It could be expected that the overlapped area of responsibility between subcontractors (or trade contractors) might lead to serious difficulties in ensuring the quality standard of this project.

3. This project involves the instalment of complicated facility equipment, such as communication systems so that this project might be more likely to encounter difficulty in getting over the interface problems among (subcontractors or trade contractors).

It is to be noted that, according to the finding in the pilot study, some respondents are likely to confuse interface problems themselves with the responsibility or risk that they bear under the procurement route used in their projects. For example, if design-build is used, the coordination of subcontractors contractually has nothing to do with the client since D&B contractor has to bear the legal liability of delivering the project with promised quality. In this case, the client tends to answer the questions on the basis of whether the interface problem will influence his/her interests rather than the severity of that problem itself. Thus, in the full-scale survey, a detailed account of what this variable aims for is given in front of the question set.
Question 2-28–2-32 are concerned with difficulties in clearly identifying the liabilities of trade or subcontractors. Would you please answer the following questions independent of the procurement routes used and on the basis of the technological characteristics of your project only. That is, we are not asking whether it is your responsibility to clearly identify the liability of traders, but how serious the problems are.

8.4 Summary of this chapter

The discussions in Sec.8.3 can be summarised in Fig.8-3. In designing the measure of variables, we try to find a practical way to rightly capture the theoretical implications of the six explanatory variables.

First of all, regarding OPPCT, we use four dimensions to measure the same thing so as to minimise the possible errors, including: (1) willingness to pay for earlier completion, (2) strategic importance of earlier delivery, (3) benefit of earlier delivery, (4) comparison of cost and benefit of earlier completion.

Second, for purposes of encompassing all the important sources of uncertainty, uncertainty should be viewed from six angles: (1) speed of obsolescence of the original design, (2) uncertain ground condition, (3) volatility of revenue from the project, (4) likelihood of budget being more stringent, (5) regulatory uncertainty, (6) complexity of design.

Third, degree of process specificity are assumed to be conditional on the four factors: (1) availability of contractors, (2) availability of designers, (3) specificity of technology and (4) availability of construction technology.

Forth, PARTR is a binary variable, labeled as 0 for the project without using partnering agreement and 1 for those using partnering agreement.

Fifth, the scale of the client is measured by the total expenditure in construction projects in the last five years.

Last, complexity of specialization can be assessed in terms of different dimensions: (1) degree of complexity of interface problems, (2) difficulties in assuring quality due to overlapped responsibility and (3) involvement of complex facilities in the building.
Fig. 8-3 Development of variable measurement

Stage II

OPPCT

Sec.8.3.1

- Willingness to pay for earlier completion
- Strategic importance of earlier delivery
- Benefit of earlier delivery
- Comparison of costs and benefits of earlier completion

Stage III

- Complexity of design
- Speed of obsolescence of original design
- Uncertain ground condition
- Stringent budget
- Volatility of revenue from project
- Regulatory uncertainty

UNCER

Sec.8.3.2

- Availability of contractors
- Availability of designers
- Specificity of technology
- Availability of construction technology

DEGPS

Sec.8.3.3

- Whether partnering agreement is used

PARTR

Sec.8.3.4

- The total expenditure on construction in 5 years

SCALC

Sec.8.3.5

- Complicated interface problems

COPMSP

Sec.8.3.6

- Difficulties in assuring quality due to overlapped responsibility
- Involvement of complex facilities in the building
9.1 Data collection
9.1.1 Source of data
Just as some of the empirical works in TCE reviewed in Chapter 7, the data required in this study needs to be collected by survey. The targeted respondents are the client's in-house project managers who have been responsible for carrying out a new or refurbishment project for commercial or industrial buildings in the year prior to the date of survey. The set of such projects and clients was obtained from Glenigan's database. This is one of the largest construction information companies in Europe. Last year about seventy percent of the British construction projects put out for tendering were recorded through Glenigan. In their database, we can find the brief description of the project, and correspondence address of the promoter (i.e., the client), architect, and quantity surveyor. Hence, we can choose the most appropriate persons to take part in the survey.

9.1.2 Basic information of the survey
The purpose of this chapter is to present the result of testing the fitness of transaction cost approach in explaining construction procurement behaviour with survey data. This survey collected 207 observations out of a sample population of 851, that is, a response rate of around 24.3%. Among them, 15 were incompletely answered and thus discarded.

Of the complete observations, the proportions of the three main categories of procurement routes are shown in Table 9-1. The traditional method still tops the ranking, having the lion's share of nearly 50%. The second largest share goes to design-build with the proportion of 39%. In contrast, management system appears out of favour, being used in as few as 12% of the surveyed projects. Within each category, traditional method, design-build and construction management system are polar forms that interest this study. Without surprise, there are different variants existing in practice. Except for accelerated tradition method (ATM), by which design and construction can proceed in parallel, other variants, including influenced design-build (IDB) (in which the client provides more than concept design) and design-manage (DM) (in which the designer also act as the role of construction manager in charge of managing work package contractors), have fairly low frequency of being chosen. For increasing the
number of MC projects to run the logit model, DM, MC and CM are grouped together. But why not classify other hybrid routes into stereotypical forms as well? This question will be answered in the succeeding subsection.

### Table 9-1 Proportion of procurement routes used in the sample projects

<table>
<thead>
<tr>
<th>Main category</th>
<th>Subcategory</th>
<th>Number of response</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional system</td>
<td>TM</td>
<td>73</td>
<td>38.0%</td>
</tr>
<tr>
<td></td>
<td>ATM</td>
<td>19</td>
<td>9.9%</td>
</tr>
<tr>
<td>D &amp; B system</td>
<td>DB</td>
<td>69</td>
<td>35.9%</td>
</tr>
<tr>
<td></td>
<td>IDB</td>
<td>5</td>
<td>2.6%</td>
</tr>
<tr>
<td>Management System</td>
<td>DM</td>
<td>5</td>
<td>2.6%</td>
</tr>
<tr>
<td></td>
<td>CM</td>
<td>18</td>
<td>9.4%</td>
</tr>
<tr>
<td></td>
<td>MC</td>
<td>3</td>
<td>1.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>192</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Legend:**
- TM: Traditional Method
- ATM: Accelerated Traditional Method
- DB: Design-Build
- IDB: Influenced Design-Build
- DM: Design-Management
- CM: Construction Management
- MC: Management Contracting

#### 9.1.3 The classification of procurement systems

This study assumes that there are three subgoals that the client pursues: earliest delivery of the project; low transaction costs from process specificity; and low measurement difficulty. To get the project completed quickly, the overlapping part of design and construction has to increase. To alleviate the severity of rent-seeing behaviour arising from process specificity, reducing the degree of responsibility concentration will help. To the contrary, to lessen the difficulty in identifying liabilities of the client’s agents, the extent of responsibility concentration has to be raised. These three dimensions can be used as a tripod for formulating the construction organisation strategy, as summarized in Table 9-2.

As a result, as long as the hybrid systems are not different from the polar systems with respect to two dimensions – the extent to which design and construction can be overlapped and degree of responsibility concentration, the aggregation of hybrid and polar systems will not degenerate the prediction power of the model set out in Chapter 6.
Table 9-2 A basic framework for construction organisation strategy

<table>
<thead>
<tr>
<th>The client’s subgoals</th>
<th>Organization Strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Earlier delivery of the project</td>
<td>1. Higher proportion of design and construction being overlapped</td>
</tr>
<tr>
<td>2. Low process specificity</td>
<td>2. Lower degree of responsibility concentration</td>
</tr>
<tr>
<td>3. Low measurement difficulties</td>
<td>3. Higher degree of responsibility concentration</td>
</tr>
</tbody>
</table>

First of all, accelerated traditional method differs from traditional method in its ability to overlap design and construction. Because the primary interest of this study is to explore the selection of polar types of procurement systems, this system is excluded from the analysis. However, as far as the frequency of this method being used, it deserves some attention in the following studies.

Second, the influenced design-build is distinct from the pure form in respect of more design input from the client’s side. It was ruled out from samples owing to two considerations. One the one hand, this study is not able to distinguish the extent to which the client’s effort has been made to design and thus its effect on the degree of responsibility concentration; on the other, its frequency of being used is relatively trivial.

Finally, three variants of management system can be grouped together because the only difference between design-manage and construction management (or management contracting) lies in who plays the role of management contractor. For design-manage, the designer takes this role, while for construction management, a professional construction management firm is retained. This doesn’t affect the degree of responsibility concentration[1]. Likewise, construction management and management contracting differ only in who has the direct contractual relationships with work package contractors – the client or the management contractor [2]. This change doesn’t affect who are responsible for construction works (work package contractors are final undertakers of responsibility in both cases). As a result, in terms of three dimensions in Table 9-2, three variants of management system don’t display obviously different characteristics, so it seems plausible to classify them into the same group.
9.1.4 Measures for reducing errors

9.1.4.1 Experience of the respondents

Though n-point rating scale is a popular way of measuring variables in most "soft" disciplines in the social sciences, such as sociology, politics and marketing, it is still subject to judgmental errors from the respondents. When you are asked to answer a question like "to what extent do you agree ...... ", you have to start exercising your brain to search a benchmark from your knowledge base, mainly build up from your past experience. Therefore the background information about the experience of the respondents is important to the success of a survey. According to Table 9-3, the people in charge of management system projects are generally more experienced than D&B projects and traditional method projects in terms of minimum, maximum, average and mode. For the whole samples, the average is about 17.7 years and the mode is 12 years. From the distribution diagrams in Fig.9-3, we can find that most samples fall into the interval of 5 to 20 years.

<table>
<thead>
<tr>
<th>Years</th>
<th>Traditional method</th>
<th>Design and Build</th>
<th>Management System</th>
<th>Whole samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>40</td>
<td>34</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Average</td>
<td>16.5</td>
<td>17.4</td>
<td>22.6</td>
<td>17.7</td>
</tr>
<tr>
<td>Mode</td>
<td>10</td>
<td>12</td>
<td>20</td>
<td>12</td>
</tr>
</tbody>
</table>

If all respondents can answer the questions on the basis of the same benchmark point, that is, ranking relative to the same scale, the answers can be used for inter-project comparisons. However, this ideal condition is very hard to attain. To reduce the bias resulting from different benchmarks, this study tries to screen out the samples according to the respondents’ years of experience in construction practice. But where is the cutoff point? A sensible point can be chosen with reference to the years of experience required in professional qualification in construction, like project manager, architect or quantity surveyor to be chartered professionals. Generally speaking, three years are thought of as the minimum limit for novel practitioners to get a fuller picture of what is going on in practice. Another reason for supporting this discretion is that a high proportion (72%) of incomplete questionnaires come from the respondents with ex-
periences of less than 3 years, implying practice experience does have something to do with the ability to complete questionnaires. It seems sensible to assume that this is a crucial factor in affecting the quality of the answers. Thus, from these two reasons, taking three years of experience as a cut-off point for controlling survey quality seems reasonable. Finally, therefore we take off ten samples, seven from traditional method projects and three from design-build projects.

**Fig. 9-1 Distribution of years of experience of the respondents**

(a) For traditional method projects

(b) For design-build projects

(c) For management system projects
9.1.4.2 Check whether the procurement route is chosen project by project

Another source of errors comes from the client's organisational inertia in choosing procurement routes. Some clients may have irrational preference for a certain type of procurement route. This kind of case must be excluded from our samples. We found that there are six respondents giving the answer that they always use certain type of procurement route. Of them, five use traditional method and the remaining one uses design-build.

After samples from inexperienced respondents and irrational choice are taken off, the distribution of samples is shown in Table 9-4.

<table>
<thead>
<tr>
<th>Main category</th>
<th>Number of response</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional system</td>
<td>61</td>
<td>40.1%</td>
</tr>
<tr>
<td>Design-build</td>
<td>65</td>
<td>42.8%</td>
</tr>
<tr>
<td>Management System</td>
<td>26</td>
<td>17.1%</td>
</tr>
<tr>
<td></td>
<td>152</td>
<td>100%</td>
</tr>
</tbody>
</table>

9.2 The meaning, measure and descriptive statistics of explanatory variables

The empirical model of this study contains six explanatory variables that are mainly derived from transaction cost reasoning. The description and measure of each variable will be stated in the following subsections. Most variables are measured by asking respondents to answer questions on the following 10-point semantic differential scale:

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

Moreover, all variables are normalised to standard normal deviates with mean 0 and variance 1 so that they can be compared in the same metric. For example, if ten samples of X variable measured by 10-point scale are collected as listed in Table 9-5, the
sample mean \((X_m)\) and standard deviation \((X_{sd})\) are 5.4 and 2.5, respectively. The normalized variable \(X_n\) can be computed by the formula

\[
X_n = \frac{(X - X_m)}{X_{sd}}
\]

as shown in the third column of Table 9-5. Through this transformation, normal deviates with zero mean and unit variance are derived.

### 9.2.1 OPPCT

This variable indicates the client’s willingness to tradeoff between earlier completion and higher cost, which is an indirect measure of opportunity cost of time between two completion dates. Questions 2.1-2.4 are used to measure this variable. Taking the average of these four questions can help reduce measurement errors only if there exists consistency between them. This can be tested by Cronback’s alpha. Thus, the next two subsections will first give a brief account of what this index is and present the results.

#### Table 9-5 An example of a normalised variable

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Normalised (X_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>-0.94</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>-0.16</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>0.24</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-0.94</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>1.02</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>1.02</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>-1.73</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>-0.16</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>0.24</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>1.41</td>
</tr>
</tbody>
</table>

\[X_m\] 5.4 \[X_{sd}\] 2.5

#### 9.2.1.1 A digression on Cronbach’s \(\alpha\)

To test whether each dimension of a measure is closely correlated with the “true value”, we need to do the reliability test. Cronbach’s \(\alpha\) is often used in empirical practice. With reference to Table 9-6, it can be calculated as
\[
\alpha = \left( \frac{k}{k-1} \right) \left( 1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma^2} \right) \tag{Eq. 9-1}
\]

where
- \(k\) = the number of components in the scale
- \(\sigma_i^2\) = variance of scores on item \(i\) across samples
- \(\sigma^2\) = variance of total scores across samples where the total score for each respondent represents the sum of the individual component scores

Table 9-6 An illustration of the calculation of Cronbach’s alpha

<table>
<thead>
<tr>
<th>OPPCT</th>
<th>Client1</th>
<th>Client2</th>
<th>………</th>
<th>Client100</th>
<th>Row variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>A1,1</td>
<td>A1,2</td>
<td>………</td>
<td>A1,100</td>
<td>(\sigma_1^2)</td>
</tr>
<tr>
<td>Q2</td>
<td>A2,1</td>
<td>A2,2</td>
<td>………</td>
<td>A2,100</td>
<td>(\sigma_2^2)</td>
</tr>
<tr>
<td>Q3</td>
<td>A3,1</td>
<td>A3,2</td>
<td>………</td>
<td>A3,100</td>
<td>(\sigma_3^2)</td>
</tr>
<tr>
<td>Q4</td>
<td>A4,1</td>
<td>A4,2</td>
<td>………</td>
<td>A4,100</td>
<td>(\sigma_4^2)</td>
</tr>
<tr>
<td>Column sum</td>
<td>C1</td>
<td>C2</td>
<td>………</td>
<td>C100</td>
<td>(\sigma_t^2)</td>
</tr>
</tbody>
</table>

Low alpha means that some components do not share equally in the common core and should be eliminated. The easiest way to find them is calculate the correlation of each component with the total score and to plot these correlations by decreasing order of magnitude. Items with correlation near zero would be eliminated. Further, items that produce a substantial or sudden drop in the component-to-total correlation would also be deleted [Churchill, 1995]. To make Eq.9-1 more explicit, let’s take the example of OPPCT in Table 9-5, where four questions are asked to capture the implications of OPPCT. Hence, Eq.9-1 can be transformed to

\[
\alpha = \left( \frac{4}{3} \right) \left( 1 - \frac{\sum_{i=1}^{4} \sigma_i^2}{\sigma^2} \right) \tag{Eq. 9-2}
\]

When alpha is greater than 0.7, the combination of four questions can be said to be a reliable factor to measure OPPCT [Churchill, 1995].
9.2.1.2 Analysis of the variable - OPPCT

Four questions display a lower alpha of 0.48 than the generally acceptable level of 0.7. After question 2.3\(^1\) is excluded, the alpha rises to 0.78, so this variable is measured by the average of 2.1, 2.2 and 2.4. Referring to Table 9-7, the maximum and minimum value of OPPCT is 1 and 10. The mean value is 4.1.

Table 9-7 Basic statistics and distribution of dependent variables

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Average</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPPCT</td>
<td>1</td>
<td>10</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>UNCER</td>
<td>1</td>
<td>8.6</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>DEGPS</td>
<td>1</td>
<td>8.7</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>PARTR</td>
<td>0</td>
<td>1</td>
<td>0.28</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) Please refer to questionnaire in the appendix.
9.2.2 UNCER
This variable aims to capture the various sources of uncertainty through questions 2.5–2.13. Nine aspects of uncertainty are taken into account by this variable, including possibility of changing the original brief, uncertain ground conditions, volatility of revenue stream from the project, the influence of statutory regulation, budget constraints, design variations, complexity of design, complexity of construction technology and percentage of innovative elements of the project. The average answer of nine questions is used as a proxy of the project uncertainty. It is to be noted that UNCER is a variable different from OPPCT in the sense that the nine questions are indicative of nine different sources of UNCER instead of measuring the same thing from different ways, like OPPCT, so the consistency of measures is not concerned here.

Of the samples, the value of UNCER ranges from 1 to 8.6 with mean value of 4.2. The value that projects most frequently take on is 5.

9.2.3 DEGPS
This is a standard variable in the empirical research following transaction cost logic, standing for degree of production technology availability. It is measured by the average of questions 2.14–2.17. Higher value indicates that construction technology employed in the project under review is less easily obtained from the industry.
This variable has a distribution in the interval of 1 and 8 with mean value of 3.3. The value of 3 has the highest proportion of projects. Meanwhile, the four questions show high extent of consistency (Cronbach’s alpha = 0.81).

9.2.4 PARTR
PARTR represents whether the client uses a (formal or informal) partnering agreement in the project under survey. This is a dummy variable, 1 for partnering and 0 for non-partnering. The average of this variable is 0.28, i.e., 28% of projects used partnering.

9.2.5 SCALC
This variable is measured in the unit of million pounds. For the basic statistics and distribution please refer to Table 9-7. The total construction expenditure in the last five years of the client’s organization surveyed has a large variation, ranging from 0 to £1500 million. The average size of the client is about £160 million.

9.2.6 COMSP
This variable is concerned with the difficulties attendant with the growing number of subcontractors involving in the project in quality control and liabilities identification. The average of Questions 2.29–2.31 is used to represent this variable.

The value of this variable ranges from 1 to 8.7 with the mean of 4.1. It shows a somewhat symmetrical distribution around the centre value of 4. Moreover, there is acceptable consistency between three questions (Cronbach’s alpha = 0.75)

9.3 The analysis of empirical results
9.3.1 The results
For convenience of interpreting the estimated coefficients, all variables are normalized to the variate with mean of 0 and variance of 1. The data from questionnaires are dealt with by Microsoft Excel first, and then an econometric software, LIMDEP is applied to run the Logit model. The estimation of variables is exhibited in Table 9-8. In any one analysis there are two equations being run; the identity of these equations depending upon the benchmark route being used in the analysis. Thus, there are two sets of equations being run:
I. Traditional method as the benchmark route

\[
\ln \left( \frac{P_{DB}}{P_{TM}} \right) = \beta_1 + \beta_1^{\text{PARTR}} + \beta_1^{\text{OPPCT}} + \beta_1^{\text{UNCER}} + \beta_1^{\text{DEGPS}} + \beta_1^{\text{SCALC}} + \beta_1^{\text{COMSP}} \quad \text{Eq. 9-3}
\]

\[
\ln \left( \frac{P_{MS}}{P_{TM}} \right) = \beta_2 + \beta_2^{\text{PARTR}} + \beta_2^{\text{OPPCT}} + \beta_2^{\text{UNCER}} + \beta_2^{\text{DEGPS}} + \beta_2^{\text{SCALC}} + \beta_2^{\text{COMSP}} \quad \text{Eq. 9-4}
\]

II. Design-build as the benchmark route

\[
\ln \left( \frac{P_{TM}}{P_{DB}} \right) = \beta_3 + \beta_3^{\text{PARTR}} + \beta_3^{\text{OPPCT}} + \beta_3^{\text{UNCER}} + \beta_3^{\text{DEGPS}} + \beta_3^{\text{SCALC}} + \beta_3^{\text{COMSP}} \quad \text{Eq. 9-5}
\]

\[
\ln \left( \frac{P_{MS}}{P_{DB}} \right) = \beta_4 + \beta_4^{\text{PARTR}} + \beta_4^{\text{OPPCT}} + \beta_4^{\text{UNCER}} + \beta_4^{\text{DEGPS}} + \beta_4^{\text{SCALC}} + \beta_4^{\text{COMSP}} \quad \text{Eq. 9-6}
\]

The first two of these equations represent a logarithmic value of the ratio of probability of design-build (DB) and management system (MS) being chosen relative to that of traditional method being chosen, respectively. The sign of the coefficients indicates the influence of each variable on the relative probability of each pair of procurement routes being chosen. A positive coefficient means the ratio on the left side of the equations will change in the same direction as the variables. For example, if \( \beta_i^{\text{PARTR}} \) is positive, the more serious is the interface problems between sub- or trade contractors, the more likely is DB to be chosen. Following this intuitive meaning of the coefficients, we are ready to analyse the results in Table 9-8.

Table 9-8 contains two sets of results on the basis of different reference or benchmark procurement systems—traditional method and design-build are used in the left and right columns respectively. First, we examine the comparison of traditional method relative to design-build. Apart from OPPCT, other five explanatory variables are statistically significant. The coefficients of PARTR, SCALC and COMSP are positive, meaning that the use of partnering agreement, high expenditure on construction project and more serious interface problems make design-build more preferable to tradi-
tional method. Conversely, uncertainty and process specificity are disadvantageous factors to design-build.

Second, as the result shows, management system tends to be more preferable compared to traditional method when there is high OPPCT, uncertainty and process specificity, while its advantages in these aspects are offset by interface problems.
Third, we turn attention to the second column. The first pair of comparison, traditional method vs. design-build, reports the same results as the first comparison in the left column. The result of second pair, management system vs. design-build, shows that time value of the client, uncertainty and process specificity are three important factors making management system desirable relative to design-build. However, partnering and interface problems can raise the likelihood of design-build being chosen.

9.3.2 A summary
The logit model can demonstrate the statistical relation between the relative probability of each procurement route being chosen and the six explanatory variables. The positive coefficient means the higher value of that variable will lead to the higher likelihood of numerator procurement route being used. In conjunction with the basic premise that all the clients are assumed to rational decision makers, the higher frequency of a procurement system being used signifies its advantage in efficiency under a given set of conditions. Thus, the effect of dependent variables on the relative likelihood of each procurement system being chosen, as the sign of coefficients shows, can be interpreted as the relative desirability of procurement systems in the form used in hypotheses. As a result, we can convert the results in Table 9-8, into the order of desirability of procurement systems, as shown in Table 9-9.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Empirical corroborated order of procurement system desirability</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPPCT</td>
<td>management contracting &gt; design-build, traditional method.</td>
</tr>
<tr>
<td>UNCER</td>
<td>management contracting &gt; traditional method &gt; design-build</td>
</tr>
<tr>
<td>DEGPS</td>
<td>management contracting &gt; traditional method &gt; design-build</td>
</tr>
<tr>
<td>PARTR</td>
<td>design-build &gt; traditional method, management contracting</td>
</tr>
<tr>
<td>SCALC</td>
<td>design-build &gt; traditional method, management contracting</td>
</tr>
<tr>
<td>COMSP</td>
<td>design-build &gt; traditional method &gt; management contracting</td>
</tr>
</tbody>
</table>

9.4 The fitness of model specification
This section is to examine the overall performance of Eq.9-3~Eq.9-6. There are two ways to do this. The first one is to test the equation against the null hypothesis that all
coefficients are zero. According to chi-square (166.7), the significance level is higher than 99.9%, namely that we can comfortably reject the null hypothesis, accepting the specification of model used in this study.

The second way is to compare the degree to which projects actually using each procurement route are successfully classified. From Table 9-10, we can know that the percentage of project actually using TM, DB and CM being correctly predicted are 71.62%, 71.43% and 66.7%. The aggregate rate of successful classification is around 70.8%. This result, compared with other researches published in the applied economic journals is quite acceptable.

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>71.62%</td>
</tr>
<tr>
<td>1</td>
<td>20.63%</td>
</tr>
<tr>
<td>2</td>
<td>16.67%</td>
</tr>
</tbody>
</table>

Rate of correct classification = 70.8%

9.5 A discussion
This chapter presents the empirical results of hypotheses testing. As a whole, the variables out of transaction cost consideration, DEGPS and COMSP are successfully demonstrated as the key factors for differentiating the relative efficacy of three pure forms of procurement systems. However, OPPCT only provides the explanatory power in separating management system and other two types of routes. These three variables alone can classify 61% of projects into the right category. If the other three variables are included, the rate can be raised to nearly 71%. When we compare the prediction performance of any three variables in terms of rate of correct classification, we find that OPPCT, DEGPS and COMSP are the best performers. This provides a justification for using these three factors as the mainstay of the theoretical model in chapter 6.
Chapter 10 Conclusions And Suggestions

10.1 Recapitulation of the theoretical innovation:

10.1.1 Conclusions from the theoretical model in Chapter 6
This study provides a rational foundation for interpreting the selection of construction procurement systems by means of an extended transaction cost theory, in which, with revenue added into consideration, the traditional objective function of transaction costs minimising is modified to be profit maximization. Procurement systems are regarded as instruments for achieving the client’s postulated goal of maximising a profit function, revenue net of transaction costs given fixed production costs. A successful transaction cost explanation ought to consist of four parts: (1) identification and operationalisation of relevant transaction attributes, (2) dimensionalisation of governance structures, (3) alignment of governance structures with transaction attributes in an economising way and (4) empirical testing of predictions in (3).

10.1.1.1 Refining the concepts of construction transaction attributes
The theoretical foundation of transaction operationalisation is built on the rent-based argument, the tenet of which pinpoints that the presence of quasi-rent in the course of transaction is the motive for rent-seeking behaviour, and the real resource consumption attendant on unproductive actions for larger share of gains from trade is the most consequential source of transaction costs with comparative significance. It follows that the discovery of the association of sources of quasi-rent with transaction attributes can serve as a way of operationalising the concept of transaction cost. After careful examination of the nature of construction procurement systems, we obtain two important findings:

(1) Transaction costs resulting from traditional categories of asset specificity, including site specificity, physical asset specificity and dedicated specificity, are less severe than in the case of manufacturing supply chains since most construction equipment is designed for general purposes and movable, making the magnitude of quasi-rent much lower. The principal source of quasi-rent in construction arises instead from potential losses from production disruption incurred by the client and the contractor. Generally speaking, the client is more vulnerable than the contractor if locked in the transaction since he/she has larger stakes in the project. In con-
Contrast to producer’s quasi-rent caused by irreversibility of the sunk investment, the client would suffer the hold-up problem due to (a) postponed realization of expected profits from the project and (b) value depreciation of the partially completed project which is claimed by a replacement contractor as risk premium for making up for uncertainty about the construction quality of the previous contractor. Relative to unexpected halt in or delay to construction work, keeping production progressing as scheduled will serve the interests of both the client and the contractor. When the potential losses of both sides are unevenly distributed, however the party with larger loss will be in the vulnerable position. In other words, the avoidance of disruption is more important for that party. The difference of losses from two parties plays the same role as quasi-rent. The new specificity resulting from this reason is called \textit{process specificity} with the emphasis that the sources of specificity are the results of mutual reliance in the production process.

(2) The role of quality rent, which often manifests its importance in service transactions, is demonstrated to be much weightier than that it is traditionally analysed to be in transaction cost economics because the inclusion of project design into governance structure will considerably alter the nature of product quality. For the case where product quality is checked against the pre-specified criteria, costs of inspection in exchange is the key concern. For cases where the producer is authorized to have discretion to change any quality attribute of the product, like shape, material, or colour, quality is not determined against a given but a moving yardstick, as product quality varies as design quality and design quality is affected by the degree of details on product requirements that the buyer can provide before production. The rent generated by inferior quality is called \textit{quality rent}. The inferiority of quality may stem from (1) failure to meet product requirements (production quality) or (2) unsatisfied product design (design quality). As a result it will be useful to make a distinction of two types of product quality: design quality and production quality.

Once the general perceived quality is not a sufficient index for reflecting the buyer’s satisfaction on quality, there exist grey areas where the dissonance of buyers and producers over quality cannot be verifiable by a third party, as the court will judge design quality by comparing the effort of designer with the general standard instead of assessing design itself. This practice will make the
mechanism for fairly settling disputes over design quality less efficacious. Combined with the general observation that cost is in proportion to quality, it would be tempting for the producer to downgrade quality so as to enlarge the expected profit. The likelihood of success in expropriating quality rent is conditional on two factors: (1) efficacy of court ordering, (2) severity of asset specificity. Providing the court can assess design quality with low cost, the attempt to expropriate quality rent by producing goods with low-production-cost design will be infeasible, meaning quality rent is no more than measurement cost. If this is not the case, quality rent may take the form of design rent, which will complicate the transaction where the producer plays the role of the designer as well.

Due to the presence of both process rent and quality rent in the construction transaction, the appraisal of its efficiency should include transaction costs caused by these two sources, and the factors affecting their magnitude are the determinants that can differentiate the strong and weak points of procurement systems. We call this operationalisation of TCE, so the degree of process specificity and of quality measurement difficulties are just what we are searching for.

Table 10-1 An assessment of three stereotypical procurement systems

<table>
<thead>
<tr>
<th>Elements of objective function</th>
<th>Traditional Method</th>
<th>Design and Build</th>
<th>Management System</th>
</tr>
</thead>
<tbody>
<tr>
<td>High NPV of expected revenue</td>
<td>-</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>Low transaction costs due to process specificity</td>
<td>+</td>
<td>-</td>
<td>++</td>
</tr>
<tr>
<td>Low transaction costs due to measurement difficulties</td>
<td>+</td>
<td>++</td>
<td>-</td>
</tr>
</tbody>
</table>

++ The procurement system showing best performance in that criterion
+ The procurement system showing second best performance in a criteria
- The procurement system showing worst performance in that criterion
In addition to costs, revenue should also be counted as a determinant of efficiency since different procurement systems will have different delivery time, affecting the present value of the completed project to the client. Putting them together, we can formulate the efficiency of construction systems as maximisation of revenue net of transaction costs (i.e. revenue – transaction costs due to process specificity – transaction costs due to measurement difficulties). We must now consider in what way procurement systems can serve this objective. Table 10-1 shows an examination of the performance of three stereotypical procurement systems in terms of three components of objective function. It is evident that there is no perfect procurement system that performs equally well in all three aspects. One system that is advantageous in one criterion at the same time has weaknesses in another. The trade-off relation between these three criteria that each procurement system exhibits leads to one of the most critical theoretical conclusions of this study, inconsistent trinity: the client faces an inconsistent trinity in choosing procurement systems, implying there is no best procurement system for all cases.

10.1.1.2 Dimensionalisation of procurement systems
The second stage of transaction cost explanation framework is to explore why a different mode of governance structure can generate outcomes of differential efficiency. To unravel the reasons, the first thing is to point out what are the most important dimensions characterising procurement systems. The fundamental difference of procurement systems can be viewed from three dimensions: degree to which design and construction can be overlapped; degree of responsibility concentration; and adaptability to design changes. A summary table is shown below, presenting an assessment of the relative advantage of procurement systems.

To interpret the effects of dimensions of procurement systems on the objective function, a mapping relation between them needs to be built. From Fig.10-1, we can see that degree to which design and construction can be overlapped influences how far the fast-tracking strategy can shorten the total duration of the project and thus increase the NPV of net revenues. That is, the larger the proportion of design and construction phases proceeding in parallel, the earlier the completion of the project and higher net present value of revenue from the project.
Table 10-2 A comparison of dimensions of three procurement systems

<table>
<thead>
<tr>
<th>Dimensions of procurement systems</th>
<th>Types of procurement systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management System</td>
<td>Traditional Method</td>
</tr>
<tr>
<td>Degree to which design and construction can be overlapped</td>
<td>High</td>
</tr>
<tr>
<td>Degree of responsibility concentration</td>
<td>Low</td>
</tr>
<tr>
<td>Adaptability to design changes</td>
<td>High</td>
</tr>
</tbody>
</table>

The extent of responsibility concentration affects whether it is easy to find the right one to blame if something goes wrong. If the price paid and quality of output can by this means be closely linked, the producer will take much care of production process and strategically manipulating his information advantage on quality will be futile. Thus, transaction costs arising from measurement difficulties will be much lower.

Fig. 10-1 The linkage between attributes of procurement systems and elements of the objective function
Adaptability to change orders is a decisive factor in affecting the extent that the client should keep "the option to change his mind" open. The cost of keeping options open is to take up the ensuing uncertainty about the final costs, while making up one's mind at an early time may reduce ability to meet the need to change the original requirements at later stage. Whether a procurement system is flexible enough to cope with change orders will alter the relative desirability of keeping options open because the severity of transaction costs arising from process specificity will vary with adaptability to change orders of procurement systems.

Fig. 10-2 Alignment of procurement systems with transaction attributes

10.1.1.3 Alignment of governance with procurement systems
At the last stage, the major task is to align the attributes of construction transactions with dimensions of procurement systems in an economising way. Put another way, the objective function plays the role of linking up the dimensions of procurement systems and attributes of construction transactions, as shown in Fig.10-2. The mapping needs to be interpreted in two ways. In pursuit of efficiency, higher net present value of revenue from the project, and lower transaction costs from both process specificity and measurement difficulties are the goals to be achieved. To increase NPV of reve-
nue, the phases of design and construction have to proceed in parallel to some extent to make the project delivered earlier (link on the left hand side). For the projects with high opportunity cost of time (OPPCT), a slight delay will cause a large difference in NPV of revenue (link on the right hand side). Combining them together, we can derive a conclusion that for a project where OPPCT is high, it is suitable to select the procurement system that can overlap the phases of design and construction. However, it is to be noted that high OPPCT also reduces the advantage of an integrated route because it increases client vulnerability to hold up. This means that, relative to traditional method, design-build’s advantage in generating higher NPV of revenues will be offset by its weak adaptability to design changes and make its relative desirability with respect to OPPCT indeterminate \textit{a priori}. The same reasoning can be applied to the other two factors. Therefore, we can obtain the conditions in Fig.10-3 to which three main types of procurement systems are best suited.

\textbf{Fig. 10-3 Alignment of procurement systems with transaction attributes: results}

\begin{center}
\begin{tabular}{c|c|c|c}
\hline
OPPCT & H & M & L \\
\hline
Process specificity & & & \\
\hline
Degree of measurement difficulties & & & \\
\hline
Traditional method & & & \\
Design and Build & & & \\
Management System & & & \\
\hline
\end{tabular}
\end{center}

\textbf{10.1.2 Conclusions from empirical investigation}

\textbf{10.1.2.1 Formulation of hypotheses}

Williamson’s theoretical analysis is to derive refutable hypotheses for empirical verification. This is has been done in principle in Fig.10-3. However, these theoretical hypotheses are still not operational enough to do the testing. We still need to seek out a complete list of critical variables that may affect the magnitude of transaction costs.
The hypotheses are stated in the form that “as X variable is higher, procurement option R will be more likely to be chosen than option Q, other factors being equal”. This relation is expressed as \( R > Q \).

**Hypothesis 1: Opportunity costs of time (OPPCT)**

The higher is the opportunity cost of time, the more desirable and therefore the more likely for the client to employ a fast-track strategy for shortening the total time. However, higher opportunity cost of time also leads to more serious process specificity problems, making an integrated route less desirable. The joint effect of these two forces leads to the order of management system > design and build, traditional method.

**Hypothesis 2: Uncertainty (UNCER)**

The more uncertain factors are involved during design and construction, the more likely for the client to need to renegotiate with contractors after contract is signed, and the more likely for him/her to employ a flexible strategy. Thus, management system > traditional method > design and build.

**Hypothesis 3: Degree of process specificity (DEGPS)**

The more costly to switch to a replacement contractor, were the project disrupted, the more desirable and therefore the more likely for the client to adopt a fragmented route. Thus, management system > traditional method > design and build.

**Hypothesis 4: Partnering agreement (PARTR)**

When a partnering agreement is used in the project, the more desirable, and therefore the more ideally for the client to adopt design and build. That is, design and build > traditional method > management system.

**Hypothesis 5: Scale of the client (SCALC)**

The more future works that the client can provide, the less threatening is the (quasi) rent-seeking behaviour and the more desirable is design-build. That is, design and build > traditional method > management system.
Hypothesis 6: Complexity of specialisation (COMSP)
When the degree of ambiguity of interface between autonomous subcontractors becomes severe, an more integrated procurement route will be more desirable because it is able to reduce measurement costs in identifying the liability of defects of the project. Thus, design and build > traditional method > management system.

10.1.2.2 Empirical findings
Through empirical investigation, the following findings are obtained.

(1) The client’s willingness to pay for earlier completion is demonstrated as an important consideration for the selection between fast-track and non-fast-track routes. There is a high frequency of fast-track routes being used in projects with high OPPCT. If the client is rational, it can be inferred that fast-track routes are the more efficient ones to deal with the projects with high OPPCT. However, higher OPPCT also leads to higher process specificity, weakening the desirability of design-build relative to traditional method. The empirical result bears out this hypothesis, i.e. management system > design and build , traditional method.

(2) The combined effects of five principal sources of uncertainty of the project, including (1) speed of obsolescence of original design, (2) uncertain site condition, (3) volatility of revenue from project, (4) regulatory uncertainty and (5) design complexity of the project, are proved to be the key determinants to differentiate the relative desirability of procurement systems. We found that the higher is the uncertainty, the less likely it is for the client to choose integrated procurement system. It follows that management system is most efficacious one to carry out the projects with high uncertainty, traditional method next, and design-build comes last. Thus, hypothesis 2 is proven to be valid.

(3) The costs of switching to alternative contractor shows its statistic significance in affecting the relative probability of procurement routes being chosen. For the case where it is particularly costly to disrupt the project and resume it by other contractor, management system can provide the most efficient way to mitigate the contractor's motive to take advantage of the client’s vulnerability, and hence reduce transaction costs. As this kind of cost become non-trivial, the benefit of using management system will be more obvious. On the same reasoning, traditional
method is more desirable than design and build. As the empirical results show, the conjecture of hypothesis 3 is completely corroborated.

4) The use of partnering agreement can make the client more likely to use design and build. It means that partnering can help improve transaction atmosphere and cultivate trust by sharing information that may affect mutual interests. The data show that the projects with partnering agreement at work are more likely to use design and build than non-partnering projects. Though theoretically (hypothesis 4), partnering should more effectively alleviate transational hazards in the context of traditional method than management system, we fail to find this relation empirically. The right order appears to be: design and build> traditional method, management system.

5) The scale of the client, in terms of construction expenditure, is only found to be a significant factor in interpreting the likelihood of design-build being chosen relative to traditional method and management system. That is, only the first part of hypothesis 5 is empirically valid. This means that frequent construction clients find it more desirable to use design and build than the infrequent client, other things being equal.

6) The complexity of specialisation of sub- or trade contractors is shown to be a critical factor in explaining the condition under which disintegrating of a project into small contracts is not efficient. The results shows that, as the interface problems appear more serious, it will be more beneficial to use the integrated route. Thus, the order of desirability in hypothesis 6 is empirically robust.

10.2 The contributions of this study for TCE theorists

Williamson’s transaction cost theory mainly is devised to deal with the problem where the lump-sum investments provide continuous flow of output, physical product or service with given quality. When the theory is applied to project-oriented mode of production like construction, the power of this theory is blunted due to the key explanatory variable – asset specificity - failing to work well and the design factor setting in. Both of them cannot be squared with the basic premises of mainstream TCE, posing obstacles to direct application of TCE to construction. This difficulty can be overcome by proposing two three concepts: process specificity, design rent and quality rent.
10.2.1 Process specificity

In the standard analysis of TCE, asset-specific investment can pose transaction hazards because it is costly to switch the sunk investments to alternative uses. When the investment, in particular where it entails huge amount of capital expenditures for physical facilities, equipment and machinery, is designed for specific purposes, the potential value loss arising from change of production plan becomes the main source of appropriable quasi rent, an inducement for rent-seeking behaviour. Thus, for the types of specificity, like site, physical asset or dedicated specificity, the irreversibility and lock-in effects of lump-sum investments play a key role. According to this interpretation of asset specificity, quasi-rent will not appear in the case where there is no lump-sum physical or human capital investment at the outset. However, in examining the production process of shipbuilding industry, Masten (1991) found that supplier’s strategic delay in delivering or installing intermediate goods would cost shipbuilding companies a lot. In this case, how harmful is the supplier’s strategic delay depends on the extent of losses due to disruption of production. If delay may cause loss of revenue and market share as well as damage to reputation and competitiveness, the accompanying opportunity costs will be high. In some cases such as integrated circuit production line, once production stops, all partially finished products will be wasted. This gives another source of quasi-rent. The supplier’s strategic action or move is to threaten delay to the progress of the works and thus to completion of the project unless the client agrees an extra payment. The client’s ultimate last-resort strategic response is to threaten to determine the contract and replace the original supplier by another firm, to complete the works. Dispute-resolution will normally be possible, so that this client sanction is rarely used. Indeed it is only a credible threat where the amount of extra payment demanded is relatively large.

When the focus of analysis is shifted to construction process, the reason that the main contractor can hold up the client is similar to Masten’s temporal specificity, namely that continuing construction is better than disruption. However, the sources of quasi rent are slightly different. Although opportunity costs of delay may be the same, the client may be required to pay for quality uncertainty at handover to new contractor at a risk premium, and the difficulty in identifying liabilities of the old and new contractors may also be likely to arise. The extended meaning of asset specificity can make
us rightly examine the potential transaction problems in construction and avoid making a conclusion that all construction transactions are characterised by low specificity and so low transaction costs.

<table>
<thead>
<tr>
<th>Categories of specificity</th>
<th>Example</th>
<th>Two alternatives under comparison</th>
<th>Main sources of quasi-rent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site specificity, physical asset specificity and dedicated specificity (Williamson)</td>
<td>The investment stays in the original use vs. The investment switches to the second-best use</td>
<td>difference of <em>ex ante</em> and <em>ex post</em> opportunity costs of the investment</td>
<td></td>
</tr>
<tr>
<td>Temporal specificity</td>
<td>keep production flow going without disruption vs. Find the replacement supplier to take over the production</td>
<td>1. opportunity costs of delay 2. losses from discarded partially-finished products</td>
<td></td>
</tr>
</tbody>
</table>

10.2.2 Design rent

In the standard setting in TCE, the product design is exogenous to the model, that is, trading parties are clearly aware of what is to be transacted. Quality at most is a problem of examining conformity to specification, and quality rent is indicative of cost savings from producing inferior products at lower costs than otherwise would arise. The cases in construction may be a bit more complicated due to different contractual arrangements. Conceptually, there are two modes of specialization between designer and contractor (see Fig10-4):

(1) Mode I: separation of design and construction

The first mode is to devolve design and construction to two independent parties. The client first gives briefs to the designer for producing drawings and specifications and then requires the contractor to carry out construction on the basis of this blueprint.
(2) Mode II: integration of design and construction
The second mode is to choose a single organisation for design and construction. The client first conveys his/her requirements to DB contractor and DB contractor then is required to put forth design and construction proposals that have to be reviewed and approved before being implemented on site.

**Fig. 10-4 Two modes of specialisation between designer and contractor**

(a) **Mode I**

Client → Briefs → Designer → Contractors → Physical Facilities

Inspection for conformance with design specification

(b) **Mode II**

Client → Briefing → Proposal → Review Result of → Design-build Contractor → Physical Facilities

Inspection for fitness for purpose on the basis of the approved design

Along with the integrated arrangement of design and construction, the incentive structures in the construction team will also be different. In mode-I, the client finds it relatively easier to ensure design quality that in mode-II on the ground of two reasons. First, though the net rent advantage is on the designer’s side, the absolute value is not so large and this disadvantage can be counterbalanced by using the strategy of payment retention and postponement. Second, rather than price, construction design market is a market competing for quality or more exactly, the caliber for design. For this kind of plastic transaction, the client has to rely upon the reputation effect to some extent. A successful project will give the designer a credit, which is beneficial to enhance the market evaluation of the designer’s competence. Thus, not many designers tend to run the risk of getting records damaged thanks to refusal to make further modification for the client’s satisfaction. After all, the costs of resolving disputes in
court are not only legal fees but the bad name for favouring confrontation. Also, the payment mechanism and forms of liability of designer give designer strong incentives to attempt to satisfy the client. Therefore, it seems quite sensible to say that the designer market in mode I can be disciplined to meet the client's demands by reputation effect. Meanwhile, with full documents of design, the client can monitor the contractor by checking the degree of conformity of the completed project with his/her documented requirements. This is also the principal criterion for the court to resolve disagreements between the client and the contractor.

Table 10-4 Comparison of two modes of specialisation between design and construction

<table>
<thead>
<tr>
<th>Mode</th>
<th>Mechanism for disciplining the designer</th>
<th>Mechanism for disciplining the contractor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode-I</td>
<td>(1) Reputation effect, rarely resort to legal system</td>
<td>Assessing conformity to requirements by designer's inspection, with arbitration or legal system to resolve disputes.</td>
</tr>
<tr>
<td></td>
<td>(2) Designer's net rent advantage is counterbalanced by payment retention or postponement.</td>
<td></td>
</tr>
<tr>
<td>Mode-II</td>
<td>Judging the quality of design with reference to fitness for purpose</td>
<td>Assessing conformity to approved design plan by arbitration or legal system</td>
</tr>
</tbody>
</table>

In contrast, in mode-II, design and construction are carried out by the same party, so the DB contractor has much wider range of options for manipulation. For instance, the contractor may propose a cost-saving design plan that can generally meet the functionality but lack aesthetic attraction and may shorten the economic life of the project. Contractually, the client has the right to disapprove any proposal that he/she is not happy with, whereas the legal system is not always on the client's side since design quality in this case lacks an objective yardstick, forcing the court to make a judgment by taking the standard of fitness for purpose as benchmark. This legal practice will leave a large gray area that can be exploited by the contractor.

From the above discussion, a simple conclusion can be made that the presence of design rent is due to the differential efficacy of court ordering in resolving design disputes in mode-I and mode-II. Where the court cannot be fallen back on as an effective
A preemptive mechanism for dealing with opportunism and where reputation effect is not able to function well, the client will be vulnerable to the contractor's strategic manipulation of his/her active right to propose design plan. This is not going to be a problem for the transaction where mutual trading quality has been standardized like manufacturing industries. In construction, it is possible for design and construction to be put under the single organisation, so its ensuing transaction problem, the presence of design rent, need to be taken into consideration. This is an element that is not discussed in mainstream TCE.

10.2.3 Quality rent
Both design and quality rent are originated from measurement problems. The former is related with design quality, while the latter is associated with construction conformance quality. When design-build is used, quality rent, in conjunction with process specificity, will put the client in a particularly vulnerable position. But why is the use of this route so pervasive? The reason lies in its better ability to deal with conformance quality. There are two aspects of this notion: (i) inspection and (ii) identification of project defects. It is assumed that the task of inspection of the completed project doesn't change as procurement systems. The key difference results from the differential degree of difficulties in identifying the liability of project defects. If a project is divided and undertaken by different work package contractors, it would be a problem for the client to enforce conformance quality. This possibility is likely to be taken advantage by work package contractors to cause type-II transaction costs. To deal with this case, single point of responsibility will be an effective measure to preempt this kind of opportunism.

10.3 The contributions of this study for construction management theorists
10.3.1 Set out a complete framework for the application of TCE to construction
This study claims to be the first complete application of transaction cost reasoning to the analysis of construction procurement systems. The only previous study following this line, Winch (1989) failed to examine the applicability of basic premises in TCE. It attempts directly to apply Williamson’s analysis and its results of operationalising transactions without making efforts to examine the basic propositions that would affect the applicability of TCE to construction. The premises that don’t hold in applying
to construction firms include (a) governance-independent revenue (b) irreversibility of lump-sum investment and continuous production, (c) exogenous design, and (d) fixed production costs\(^1\). Factoring the characteristics of construction into TCE framework is a very critical step to the successful application of TCE approach. This state of affairs makes it necessary to go back to the fundamentals to seek the right modifications. The author claims that this study has met this call and successfully found the way forward.

In principle, this reasoning framework can be extended to deal with other pivotal organisational issues in construction, such as subcontracting and joint venture. The potential usefulness of this approach in the analysis of subcontracting has been demonstrated in Chau and Walker (1994). Developing construction organization theory along this line of inquiry will be a fruitful undertaking.

### 10.3.2 Place the role of the client at the central stage of analysis

As far as the current development of construction management is concerned, the client is located at the periphery of analysis. The role of the client is just like shadow, the recognizance of its existence relying on the presence of sunlight instead of its own importance. Construction industry is dealt with as a service provider, trying to satisfy the client's needs that have existed in his mind. This one-way thinking, neglecting the necessity of taking account of interaction between the client and construction industry, is liable to lead to conclusions with no awareness that other important forces may be left out of analysis. This problem can be found in Winch (1989), where the contractor's boundary is determined mainly on the basis of contractor's interests, ignoring the fact that the client has the initiative to choose the procurement route. A complete framework should contain a coevolutionary process in which two adjustment processes proceed in parallel, as shown in Fig. 1-1.

1. The client's adjustment process: selection of procurement routes (Fig.10-5)

By any rule, the client, at the outset, has to choose a procurement system (or even invent a procurement system) to carry out an ongoing project. The appropriateness of

---

\(^{1}\) Noted that this assumption was also required in our analysis of procurement systems. However, in the analysis of construction firms, whether it is desirable for the contractor to maintain an in-house design team involves the sharing of production costs (wage paid to hired designers) between projects. As a result, production costs will vary between market and hierarchy.
this procurement system in handling the project at issue considerately depends on the attributes of the project. The final performance of the project is an indicator for testing whether the original pick is right or wrong. The experience of carrying out this project can be accumulated by the client's organisation, his consultants or other participants. The accumulation of this experience can gradually shape up principles for improving decision quality on the procurement system selection. Applying these principles, the client is able to select an appropriate procurement system in the context of different projects. Through loops of principle correction, the proportion of chosen procurement systems being consistent with transaction attributes will increase over time until equilibrium is approached.

Fig. 10-5 The client's learning mechanism for procurement system selection

![Diagram]

(2) The contractor's learning process: selection of organisation form (Fig.10-6)

On the contractor side, the main organization issues are (A) whether to maintain a design department and (B) whether to retain in-house labour force or specialist trade workers (e.g., structural steelwork, concrete, and roofing). The range of organisation forms available to the contractor is constrained by the clients' use of procurement systems. It follows that the contractor's choice of organisations has to be a response to the population of the procurement systems chosen by the client. For example, when management system is widely and increasingly selected, there is no reason for the main contractor to expand the organisation boundary to subsume subcontractors or designers under the rubric of 'hierarchy.

Whether the contractor has chosen a suitable organisation form has to be tested by the performance of the project and this experience can help the contractor formulate the
heuristic principles for organising the construction team in undertaking the next project.

Fig. 10-6 The contractor’s learning mechanism for organisation selection

Therefore, this study wants to convey an important message to people who are thinking about contractor’s strategies that, in modeling the interaction relations between the client and the contractor, the client takes an active role in making the first move and may take into account the accompanying repercussion from the contractor in the next move so as to avoid backlash from the wrong decision. This is a two-way adjustment process. The constraints imposed by the client are decisive factors in affecting the contractor’s decision.

10.3.3 Provide a new paradigm for analyzing the construction process

The goal of this study is an application of Williamson’s transaction cost reasoning to the analysis of construction process so as to rationalize the procurement system selection behaviour. Its potential influence will go beyond the particular scope of this thesis, and hopefully lay a foundation for prompting a paradigm shift for the way that the construction process should be examined. Conventionally in theoretical economics, the problems of construction procurement are treated as no more than those of selecting an appropriate contract form (incentive and risk allocation). The client is suggested to choose contract on the basis of risk preference of themselves and the contractor (e.g., McAfee and Mcmillan, 1987). The logic underlying this wisdom reflects the assumption of complete contracting, according to which contract is an effective tool to govern the transaction process so that contract itself becomes the decisive factor for the success of transactions. Then, it may be asked, why should contract be un-
derstood as incomplete; and what are the implications for analysis? First of all, the concept of contract incompleteness can be best understood as a problem of determining the optimal level of contract incompleteness. There are two forces at work: one force pushes forward the frontier that contract can contain as far as possible so as to alleviate the adverse impacts brought about by unforeseen events, while meanwhile the costs of formulating a contract will be pushed upward as well. The optimal level of incompleteness must be struck somewhere between two extremes: no contract at all and a perfect contract. Therefore choosing a contract with more or less incompleteness will best serve the interest of the decision maker, depending on the relative severity, in context of two problems: cost of information and cost of incompleteness. It can be clearly sensed that the fact that it is costly to formulate a contract is a key to the understanding of why the effectiveness of contract cannot be entirely relied upon. It may be of interest to see where these costs come from. This question can be approached from the opposite side. Intuitively, a contract can be said to be complete only if it is able to accommodate all the information relating to the transaction in question and transform it into enforceable courses of actions that the trading parties agree to abide by so that the contract can provide clear rules for them to follow as the transaction unfolds. In this description, the process of drawing up a contract actually constitutes three parts:

(1) Formulation of contract:
This requires collecting relevant information, then turning that information into written clauses. The information required in a contract actually is a set of information modules, consisting of the pair of contingency description and its corresponding actions. In classification, there are four categories of information modules: (1) information that is available to the transactor, (2) information that is available to others and can be accessed by the transactor at a cost, (3) information that is unknown yet but can be obtained by way of research, (4) information that cannot be known anyway. The costs of gaining and transforming each category of information module into contract clauses that are desirable to one party will be different, generally increasing on this order, so the composition of information units of each category becomes a critical factor for determining the costs of contract formulation.

(2) Negotiation:
The first step aims to discover the desirable contract terms to the party owning the private information, while whether an agreement can be reached usually depends on the bargaining power. Due to the presence of information asymmetry, mutually beneficial terms may fail to be achieved and efforts may be devoted to unnecessary haggling.

(3) Enforcement:
The third element highlights the enforceability of contract clauses. Only when the contract can be enforced by either private or public ordering systems if disputes arise will it be effectively binding for maintaining transaction order. Thus, the content of a contract should go beyond “if ..., then...” and consider the costs of verifying “ifs” and “thens”.

<table>
<thead>
<tr>
<th>Points of difference</th>
<th>Process perspective</th>
<th>Contract perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>The effectiveness of contract</td>
<td>Incomplete</td>
<td>Complete</td>
</tr>
<tr>
<td>Nature of procurement systems</td>
<td>A type of governance</td>
<td>A form of contract</td>
</tr>
<tr>
<td>Focus of analysis with respect to time</td>
<td>The process in its entirety</td>
<td>A point in time</td>
</tr>
</tbody>
</table>

Putting the above three elements together, there is little hope of expecting the contract can be perfectly drafted and frictionlessly implemented. It follows that the contract must be incomplete. Unfortunately, awareness of this fact is not always shown in studies of construction, while its full implications for theorizing haven’t been accorded apposite attention. Therefore, the perspective of modeling construction process as contract still prevails up to the present. This contract perspective is marked off from the process perspective espoused in this study, by its unrealistic assumption of perfect contract effectiveness. Once it is not taken for granted that the contract is an agreement that can solve any problems likely to happen in the course of transaction, the “black box” of the transaction process has to be unraveled. Now the place of time in the model will not be confined to the point of decision making, but the whole process of transaction (as illuminated in Sec5.2.1): (1) look forward from the outset to
predict what may happen; (2) then examine the dynamic process of transaction; and (3) look backward to see whether the original decision should be modified. Through the examination of transaction process, we can come to realise that the nature of procurement systems should be rightly understood as an organisational matrix governing information and material flows in the construction process by way of contractual arrangements between parties. The points of difference between the two perspectives are illustrated in Table 10-5.

10.4 Implications of this study

This study is an attempt to apply the transaction cost reasoning to give construction procurement behaviour a rational interpretation. The inference departs from the premise that, in pursuit of profit maximization, the client will choose a procurement route according to his/her available information. That is, the ex ante information efficiency is assumed. The profit of a project is made up from three components: revenue, less transaction cost, less production cost. Following the tradition of transaction cost economics, production costs are assumed to be invariant to institutional setting, so the focus of analysis is mainly on the role of revenue and transaction costs, especially the latter.

By way of the transaction cost reasoning, this study is intended to be a manifestation of the implication of incomplete contracting in construction procurement behaviour, which means that procurement system selection should be regarded as strategy of first-order importance in achieving the goal of economising. The role of contract and tendering system is relegated to second-order importance. This assertion will have profound repercussions on the basic thinking about several crucial issues that will be discussed in this chapter. Some implications proposed are in some sense conjectural and their actual influence awaits further investigations, while it seems quite safe to say that from a theoretical perspective, there is real need for paradigm shift in thinking about the construction process.

10.4.1 Procurement route selection as strategy of first-order economising

In the transaction cost economics, getting the selection of governance structures right is regarded as strategy of first-order economising, implying that the way that a trans-
action is governed has the greatest impact upon the efficiency of organisation. As applied to construction procurement behaviour, we would like to highlight that the form of contract cannot replace the function of procurement systems in three aspects: speed of completion, certainty on completion time and certainty on cost, as shown in Table 10-6.

Table 10-6 Comparison of strategy of governance and strategy of contract in the context of construction project

<table>
<thead>
<tr>
<th>Points of considerations</th>
<th>Strategy of governance</th>
<th>Strategy of contract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earlier delivery of the project</td>
<td>Choose a fast-track route.</td>
<td>Include reward clauses into the contract</td>
</tr>
<tr>
<td>Reduce uncertainties in total payment to contractors and designer</td>
<td>Choose a procurement route according to the characteristics of the project and the client itself.</td>
<td>Use a fixed-cost contract.</td>
</tr>
<tr>
<td>Reduce uncertainties in the date of completion</td>
<td>Choose a procurement route according to the characteristics of the project and the client itself.</td>
<td>Use the type of contracts by which the date of completion can be specified in advance.</td>
</tr>
</tbody>
</table>

First of all, to advance the delivery of the project, it is natural to adopt a fast-track route to make design and construction overlap. Perhaps, it can be achieved by including bonus clauses into contract to encourage the contractor to shorten the total duration. However, in some projects where the scope and quality of projects cannot clearly be specified before contracting, the enforcement of contract would be troublesome if the status of completion is hard to verify. Thus, bonus clauses are likely to bring about deterioration of construction quality on the one hand, and delay arising from disputes on the confirmation of completion and claims for allowance of extra time in consequence of client-generated changes.

The second function of contract in construction is to give the client a price commitment before the realization of actual costs. Restricting the total final payments to a controllable range can alleviate financial uncertainty of the client. Nonetheless, the fixed cost often means the fixed scope. Without clear delineation of scope, unexpected variations are likely to occur, raising a question about how to determine the price of added components of the project. The efficiency of ex post renegotiation
price of added components of the project. The efficiency of ex post renegotiation on the share of profit will be subject to many constraints. First, even if the ex ante profit margin is considered as acceptable, this private information is unknown to the client. The split of gains from trade depends on their relative bargaining power. For the project with high process specificity, the client is relatively vulnerable and renegotiation will lead to the possibility of the client being overcharged. Thus, whether the fixed-cost contract is efficacious is dependent on how likely is it for design to change.

Fig. 10-7 The logic of claiming procurement system selection as strategy of first-order economising

Contract also serves the function of fixing contractor's commitment to completion. For assuring earlier completion, bonus clauses are often contained in the contract. To what extent this measure works depends on the costs of verifying the status of completion. Accelerating the completion of the project may be traded off by the downgra-dation of quality. Wherever attributes of quality are hard to define in advance or costly to monitor, the quality rent will be present. Coupled with the manipulation of design rent, the implementation of bonus clauses may become the site of disputes. In contrast, choosing the procurement systems with the potential of making design and
construction overlap will be the most effective way of achieving the goal of earlier completion.

From this simple comparison of contract strategy and governance strategy, it can be realised that managerial actions on the basis of complete contracting are perilous in the sense that placing too much demand on the efficacy of contract could lead to the worse repercussion than not doing so. A sensible sequence for determining construction organisation strategy is to select a right procurement system and under this constraint determine the right form of contract.

10.4.2 Implications for construction risk management

According to the mainstream perspectives on construction risk management, identification of risk sources should be followed by picking a suitable risk strategy. One category of instrument for controlling risks is to choose the right contract type so as to allocate or transfer risk to the parties who are willing and able to take it. It is generally believed that the lump-sum contract can completely transfer risks to the contractor. It follows that, if a project is identified as highly risky and the client is risk averse, the best risk strategy should be the lump-sum contract. However, there are two potential pitfalls inherent in this line of thinking: (1) unreal assumption of complete contracting and (2) negligence of hazards from behavioural uncertainty. A combined effect of these two factors leads to failure to recognize the relative importance of contract and procurement systems. The price that has to be paid for this misconception would prevent the realization of the original intention.

As expounded in Chapter 4, decision makers are constrained both by physiological and language limits as well as ignorance of the future, so it is costly to collect, process and verify information. The presence of information costs makes it desirable to keep contract incomplete since the benefits of increasing the completeness of contract will be offset by the increased costs of information. Given the prevalence of incomplete contract and the assumption of opportunism, it will be no longer reliable to use the contract as instrument for allocating risks. This study proposes instead a hierarchical structure of strategy. First, the selection of procurement systems is a decisive factor in attaining efficiency. The determination of contract forms should be founded on the given procurement system. The tendering system is a mechanism for choosing the
most suitable contractor. Thus, the decision making of tendering system will take care of itself, seeking a balance between competence of contractor and costs of tendering.

10.4.3 Implications for the procurement route selection techniques
Through the examination of the most advanced techniques for procurement systems selection, it is found that the MAUA is no better than bare-footed empiricism without any attempt to explore causality behind empirical facts. This is a sensible starting point, but definitely not the end. We shouldn’t underestimate the power of deductive methods, like economic approach, to advancement of our understanding of the nature of procurement route selection. In fact, without reference to the results of economic analysis, there is no way to check whether the pure empiricism would lead clients astray. We are not promoting the mono-methodology or economic imperialism, but instead emphasize both ways of induction and deduction should be respected.

The major pitfalls of applying the MAUA to procurement route selection lie in its reliance upon the spurious association of procurement route with the attributes of the client’s preference, i.e. priority variables. Those documented variables, say in Love et al. (1998), can meet the basic premise of the MAUA, that decision variables have to be attributes of the consequence of the decision, only when the construction contract is assumed to be complete. However, with the rise of transaction cost economics, this line of thinking has been demonstrated to be too unrealistic to be of practical use. It follows that the advice on the procurement route selection according to the client’s assignment of weights to priority variables deserves scepticism. A proposed alternative is to employ three-step transaction cost reasoning framework – operationalisation of construction transaction, determination of linking coefficients and choice of the most efficient procurement route. There is no denying that this approach has its inherent difficulty in finding the reliable linking coefficients, but by no means is this insurmountable. It is a way worthy of venturing.

10.4.4 Implications for the study of construction disputes
Construction disputes, no doubt, are a topic arousing general concerns. The current research on this issue leaves out a very crucial explanatory variable: the divergence of ex ante expectations on the future and ex post realization. According to an assertion of this study, the mismatch of procurement systems with transaction attributes would
give rise to the aggravating effect on the likelihood and extent of disputes arising from rent-seeking behaviour. Unfortunately, this critical point is largely ignored by the existing literature on construction claims and disputes. For example, in Diekmann and Girard (1995), they found 21 out of 38 variables are relevant to the dispute performance, including owner's management and organization, contractor management and organization, project complexity, project size, financial planning and project scope definition. But no allowance is made for the effect of the procurement route of choice. If the possibility of misalignment of procurement routes is not taken into account, we may miss one of the ultimate reasons causing disputes. Similarly, Mitropoulos and Howell (2001) attempts to analyse construction disputes from the perspective of transaction cost economics, while the importance of aligning the chosen procurement system with project attributes is still ignored. In response to this paper, Chang and Ive (2001) set out a transaction cost based framework for understanding construction disputes, on the basis of which preventive strategies can be developed.

10.4.5 Implications for the philosophy of construction management

Efficiency is a welfare criterion of foremost importance in economics. For the whole economy, a desired outcome is Pareto optimality, where nobody can be better off without sacrificing other's interest. In mathematical terms, the equilibrium market price can bridge the equalization of the marginal rate of substitution (MRS) and the marginal rate of technical substitution (MRTS), channeling the scarce resources to their best uses. Whenever the outcome of markets falls short of the conditions of perfect competition, the practice is hypothesised to be inefficient. This kind of judgment against the benchmark of frictionless world is discredited as nirvana approach. With the deficiency of our understanding to the nature of hierarchy and market, there has arisen a comparative institution paradigm, asserting that the efficiency of governance structures should be assessed on the basis of the principle of remediableness, namely that the existing governance structure can be said to be the most efficient one only if there is no welfare improvable alternative. As a result, before jumping to the conclusion that there emerges inefficiency in a project and a need to take some managerial actions, a very important step is to take stock of the relative advantage of procurement systems and choose the suitable one. This logic seems to have not been fully used by construction practitioners. Take the example of gap analysis. In Winch et al. (1996), it is asserted that information gap is the site of problems where management should in-
tervene. The rationale of doing so stems from the implicit premise that information gap must cause inefficiency. This reasoning has something in common with a once dominant view on the role of government that market failure is a sufficient condition to introduce government intervention. The problem here is that we don’t know whether there is any alternative that would generate net gains. If not, according to the principle of remediableness, this is efficient. As applied to construction projects, information gaps are sources causing costs, while this cost is not necessary remediablable. We have to always remember that managerial actions are always accompanied by costs, so the appraisal of these interventions should be founded on the cost-and-benefit analysis, instead of claiming that the information gap should be as narrow as possible.

10.5 Remarks on the limitations of this study

10.5.1 The assumption of ex ante efficiency

A central assumption made in this study is that the client will choose the most desirable procurement system so as to maximize the expected profit from the project. This is an assumption of ex ante efficiency, i.e., optimising the objective function on the basis of ex ante expectations on decision variables. It is to be noted that the client is assumed to make a choice by referring to his/her expectations on transaction attribute variables rather than the components in the profit function, like expected revenue or transaction costs. In other words, whether the client’s choice is commensurate with theoretical prediction heavily depends upon the efficacy of the client’s learning mechanism as enunciated in Sec.5.4. There is no guarantee that the result of equilibrium must reach the state where the client always makes the theoretical right choice, whereas this assumption is pervasively made in the most mainstream economics studies. It seems to me that the most convincing defence has to resort to “sophisticated falsificationism”, meaning that no theory should be rejected if no better alternative exists [Vroey, 1998]. Under this doctrine, it can be considered as approximately true that the agent is assumed to behave as if an objective function is maximized. Of course, taking this view run the risks of ignoring the effects of incompleteness of the client’s learning mechanism. Though the empirical results of this study are reasonably well in line with predictions, how serious the errors would be still deserves further attention. In what follows, two plans for testing the validity of the assumption of ex ante efficiency are sketched out.
Plan I: Compare the performance of two contrasting groups of projects
If we can collect two large enough groups of samples, one group following TCE principles and another group inconsistent with TCE principles, the average performance of the projects in two populations will give clues as to whether the assumption of ex ante efficiency is sensible. For the null hypothesis, the proposition that the procurement system is not chosen on the basis of efficiency principle, then except for random disturbances, we would not find systematic differences in the aspects such as cost overrun, disputes and time delay. This method can be used to double-check the assumption.

Plan II: Develop a complete model for the client’s learning mechanism
The assumption of efficiency as the only objective downplays the influence of dynamic equilibrium process. In general, the client will try to improve the decision quality in procurement system selection by way of learning from the past experiences, other people’s experiences or consulting with guide books and experts, while how effective is the learning mechanism and how important is the effect of path dependence on the equilibrium outcomes of procurement system selection are still unknown. Providing we have a well-developed model for answering these questions, we will be more confident about the possible outcomes of equilibrium.

10.5.2 Lack of better measures for variables
Another possible limitation may come from lack of objective measure for some key variables in this study. It has been a popular practice for the soft studies in the social science that measuring variables by way of eliciting the respondents’ subjective evaluation can be accepted as a substitute method in dealing with the case where there is no alternative objective measure that can be employed. However, relative to objective variables, such as price, unemployment, inflation rate, subjective variables must be subject to the influence of some uncontrollable factors, such as the respondent’s mood in answering the questions, understanding of the questions and selection of benchmark in making comparisons. Technically, the errors resulting from these factors can be minimized if the rigorous standard procedure of survey is followed, while we still have to put at the top of our research agenda the development of objective measures for variables that are very complicated but very critical to the analysis of
construction, such as uncertainty, complexity or performance of a project. The potential difficulty of this undertaking is well evidenced by its slow progress, while it is not hopeless. For example, a feasible way for representing the project risk is to estimate its cost variations relative to the average curve of the population (see Fig. 10-8). It can be estimated by

$$\text{risk} = \left[ \sum_i (C_i^c - C_i^b)^2 \right]^{1/2}$$

where $C_i^c$ and $C_i^b$ are the points on the cost accumulation curve of the compared project and benchmark curve.

![Fig. 10-8 A measure of project risk](image)

This proposed index bears resemblance to the standard measure of financial risk, Sharpe’s beta, whereas there is one point of difference worth noting. Shape’s beta is derived by comparing the return of a portfolio with that from riskless bonds while the project risk is estimated by comparing with the average performance of the similar projects. That is, the benchmark is varying with the selection of the sample projects. To get over this, on the one hand, a methodology with solid theoretical foundation of statistical inference is called for; on the other, the need for a database designed to collect the detailed records of construction projects should be appreciated.

10.5.3 The effect of $\text{TC}_1$ on $\text{TC}_{II}$

In Sec.3.3.2, a novel distinction between $\text{TC}_1$ and $\text{TC}_{II}$ is proposed to explore the nature of transaction costs that a transaction incurs. In construction, $\text{TC}_1$ may not only take the form of tendering costs and measurement costs, but also costs for contract
design. Contract is an instrument by which the range of actions under different circumstances that each party is allowed to take can be specified. A perfect contract must contain a complete list of mapping between contingencies and actions and thus everything expected ex ante can be enforced ex post. However, the costs of contract design will increases as the completeness of a contract increases. In principle, the more contingencies a contract can govern, the fewer loopholes can be exploited and the less severe is the rent-seeking behaviour. Accordingly, there is likely to be a tradeoff relation between TC_i and TC_{ii}. This tradeoff is relevant to the selection of procurement systems if a change in the degree of completeness of a contract has a differential effect on alternative procurement systems. Putting it more explicitly, the comparison of procurement systems should be made under the condition of the same level of completeness. When the increase in TC_i (i.e., the degree of completeness) can lead to more savings in TC_{ii} and the savings significantly vary as procurement systems, the selection of procurement systems and contract design has to be determined simultaneously. The problem as to whether the effect of the tradeoff between TC_i and TC_{ii} is significant enough to alter the relative efficacy of procurement systems requires further investigation. It is too premature to take any position. However, we may find some interesting clues from the observation that standard forms of contract are more and more popularly used in practice. Perhaps, it is a good starting point to explore why few clients or contractors are willing to spend resources in improving the completeness of contracts.

10.6 Suggestions for policies

Rent-based arguments are characterized by their stress on the motivation of exploiting other parties’ vulnerability and its accompanying consequences. To many, excluding good will out of human nature is too simplified and gloomy a perspective. As Arrow (1974) claims, trust is the fabrication oil in the working of an economy. Without it, many transactions will not take place. This also provides the rationale why partnering agreements can alleviate construction disputes and increase efficiency. Is there any inconsistency between these two views? The answer lies in the definition of trust. In fact, if we examine the rationale of TCE carefully, the magnitude of transaction costs depends on a key variable: asset specificity k. Functionally,

\[ TC = f(k) \quad f'(k) > 0 \]
This equation implies that, along with the interdependency of trading parties deepening, the costs arising from rent-seeking behaviour will rise because of appropriable rents becoming larger. That is, the larger is the bait, the more lucrative is the attempt to snap it up, and the more serious is the consequence. In this setting, observing the degree of asset specificity is equivalent to predicting the similar choice of governance structures that will be observed. However, only when the trading relations are assumed to be constant across companies is this view true. A more general model is to introduce a variable $s$ to denote the degree of trading parties trust in each other, that is,

$$TC = f(s,k) \quad f_s(s,k) < 0 \quad f_k(s,k) < 0$$

This model implies that transaction hazards will be lower when trust prevails in the course of transaction. Accordingly, the measures you should take to protect your stakes in this transaction are chiefly dependent upon the expectation on the reliability of your trading partners. For a given degree of asset specificity, the anticipation of low possibility of breach of promise will shift the selection of governance structures to the more vulnerable one. Graphically, with reference to Fig.10-9, higher expectation on trust can move the curve of $\Delta G$ from $\Delta G(s_1)$ to $\Delta G(s_2)$, shifting the joint curve of production costs and transaction costs up from $\Delta C + \Delta G(s_1)$ to $\Delta C + \Delta G(s_2)$. We can examine the effect of trust from two ways: (1) the range suitable for the use of
market expands from $[0,k_1]$ to $[0,k_2]$; implying market is more vulnerable as hostility dominates. (2) At trust level of $s_1$, $k=k_1$ is a condition that makes market and hierarchy efficiency-indifferent, while with higher trust level, say, $s_2$, market is more efficient than hierarchy. In a word, the function of trust can oppress the hazards arising from vulnerable governance structures being employed.

In recent years, to soften the confrontational atmosphere between the client and his/her agents, in particular contractors, partnering agreement is upheld as nearly panacea for malaise in construction. Borrowing Klein’s terminology, the use of partnering agreements can help enlarge the range of contract self-fulfilling. Use the case in Fig.10-9 as an example. The effect of making market suitable for the transactions with asset specificity in the interval of $k_1$ and $k_2$ is a manifestation of trust at work, namely that trust makes vulnerable positions less dangerous than otherwise it would be. But we also want to emphasise that maintaining trust is more difficult than creating it in the business world. Maybe it is too bold and too premature to make a final comment on the effects of partnering. At least, we can be sure about one thing: that the limitation of partnering will be sharpened gradually as it is widely employed without prudent assessment. And finally, we still need to go back to explore its economic principles and conduct a research with the purpose similar to that of this study.

10.7 Suggestions for further studies
10.7.1 Organisational studies in construction
Following the standard practice of TCE, this study also treats the construction transaction process as its unit of analysis, restricting the focus of analysis to the cases where the client selects the procurement system project by project, while the basic framework is in principle extendible to handle all construction organisation issues. On the one hand, we can attempt to enlarge the horizon and scope of consideration to consider the project as a part of company operation strategy, paying more attention to the learning mechanism of the client’s organisation and construction firms and the effect of organisational competence on the choice of procurement systems. On the other, we can apply the same line of reasoning to build up a transaction cost-based theory of subcontracting.
10.7.2 Improve procurement system selection technique

This study in essence is positively oriented, i.e., towards providing a rational model to explain and predict the empirical observations. However, once the theoretical hypotheses are proven to be consistent with facts, the underlying reasoning can be transformed to normative principles of procurement system selection.

As discussed in Sec. 2.4, the multiple-attribute utility approach is a useful tool for aiding the selection of procurement systems. However, what factors should be considered is a point of contention. The current use of this approach based on the premise of complete contracting is misleading. A possible remedy is to choose the decision factors that are identified and demonstrated to be significant in this study. Now the problem left is how to determine the linking coefficient matrix.

Based on the theory set out in this study, we can develop a transaction cost approach (TCA) in which construction procurement systems are considered as instruments for achieving the client’s goal - maximising the profits from the project. To pick a suitable procurement system, a three-step scheme is suggested to be followed (see Fig. 10-11).

![Fig. 10-10 The theoretical foundation of TCA in procurement system selection](image)

The first step is to operationalise the construction transaction, i.e., to find out the transaction attributes affect the differential efficacy of procurement systems in carrying out projects of different attributes. Next, we have to identify the inherent charac-
teristics of procurement systems that can demonstrate direct bearing on the performance of procurement systems in terms of the extent that the client's objective can be satisfied. The last step is to single out the conditions to which each procurement system is most suited by aligning the transaction attributes with procurement systems.

Thus, to anticipate the next steps of the argument: for projects with attributes such that the quality enforcement problem is likely to be more serious than the hold-up problem, a preference for an integrated procurement system is rational, on TCE grounds; for a project with attributes such that the hold-up problem is likely to be more serious than the quality enforcement problem a fragmented system is an appropriate choice, whilst for projects with attributes such that the benefits of earlier completion outweigh either or both problem, a management system is appropriate.

Now, the project attribute that will make hold-up a severe problem is, above all others, the degree of likelihood that the specification will have to be varied post-contract. Once the client steps outside the protection afforded by sticking to the specification explicitly embedded in the (incomplete) contract, their vulnerability to hold-up increases by an order of magnitude.

Likewise, the project attribute that will make quality enforcement a severe problem, above all others, is the degree to which the specification is unusual, innovative or complex.

Finally, the project attribute that will make achievement of earlier completion a benefit sufficient to outweigh these other problems, is essentially the relative insensitivities of total project rates of return to increases in construction cost and management cost compared to their sensitivity to a delay in the inflow of project revenues.

The above discussion can be summarised as three points: (1) high opportunity costs of time is an incentive for the client to use fast-track strategy, like management system; (2) high costs of switching to the replacement contractor in the course of construction will form an incentive to decentralize the responsibility of design and construction to different parties; (3) high costs of identifying the liability of subcontractors can give the single-responsibility system an advantage. Thus, theoretically, opportunity costs
of time, degree of difficulty in replacing the contractor and degree of difficulty in identifying the liability of subcontractors (or trade contractors) are the tripod for determining the relative desirability of procurement systems. Empirically, apart from these three variables, related variables that can aggravate or alleviate the severity of the effect caused by these three variables should be taken into account as well. The additional factors include uncertainty, the use of partnering agreement, the scale of the client. Putting them together, the utility $U_j$ that procurement system $j$ can deliver to the client in dealing with the project with attributes $A_k$ can be expressed as

$$U_j = \sum_k A_k \times R_{jk} \tag{Eq. 10-1}$$

where the linking coefficient $R_{jk}$ is determined by the features of procurement systems, indicating the extent that procurement system $j$ can satisfy the client's objective in terms of projects with the attributes $A_k$. Though apparently both of Eq.10-1 and Eq.2-3 are of a similar form, the multiplication of utility coefficients and the value of attributes, the logic behind them is drastically distinct.

MAUA uses those factors that the client is likely to consider in selecting procurement system as priority variables ($a_i$), the value of which are given by the client's subjective evaluation or preference. However, the weighting factors ($w_y$) relating priority variables to procurement systems are generic and determined independently of the specific attributes of the project. In contrast, the TCA filters out the transaction attributes $A_k$ by operationalising construction transactions so as to identify the conditions that are most critical in differentiating the relative efficacy of procurement systems. Next, it determines the linking coefficients $R_{jk}$ with reference to the performance of procurement systems in terms of project with the specific dimensions $A_k$. Finally, choosing the procurement system with highest utility index is in conformance with the spirit of aligning the procurement system with transaction attributes. As a result, our proposed transaction-cost-based procurement system selection method is theoretically well founded, closely linked to Williamson's transaction cost reasoning framework (Williamson, 1985,1996).

To be of practical value, the TCA logic expressed in Eq.10-1 has to be transformed into a practical tool for aiding the client's choice of procurement systems. There are
several possible ways of carrying out this transformation. One is to replace the original variables in the statistical approach with transaction attributes derived from transaction cost reasoning as a means for predicting the expected performance of each procurement system in dealing with projects with certain set of attributes. Another possibility is to take frequency of each system being chosen in the past successful project as a guide for assessing the prospect that the new project in question turn out to be successful if that system is adopted.

10.7.2.1 Transaction cost-based performance prediction method

As proposed in Molenaar and Songer (1998), the efficacy of design and build can be assessed in five dimensions: (1) budget variance; (2) schedule variance; (3) conformance to expectations; (4) administrative burden; (5) overall user satisfaction. They try to regress these five explained variables with a few explanatory variables such as scope definition, schedule definition, budget definition, project complexity, owner experience and so on. Those variables with statistical significance become the explanatory variables of the optimum regression model. Put formally,

\[ S = V\beta + \epsilon \]

where \( S \) is a 5×1 column matrix, standing for five performance criteria; \( V \) is a 1×n row matrix, denoting n explanatory variables; \( \beta \) is a n×5 coefficient matrix; \( \epsilon \) is a 5×1 column matrix for random errors. In applying to the practical case, their proposal contains the following steps:

(1) Input the value of \( V \) of a new project in question into the regression model obtained in their research.

(2) Convert five performance indicators into an overall score and compare the score of the new project with an average score of the projects surveyed in their study.

This comparison will show whether this new project if carried out by design and build will be more likely to succeed than other projects.

Under this framework, the transaction cost approach can be easily accommodated by just changing the original explanatory variables with a set of transaction attributes derived from transaction cost reasoning. However, there is a side of contention in this method. To give the construction client a clear guide, the performance indicators have to be as simple as possible, i.e., as Molenaar and Songer (1998) suggests, a single in-
dex. But the problem is the justification of this conversion. No matter what method is used, simple summation, weighted average or something else, all of them are disputable. Moreover, this point is floating along with the addition and deletion of samples. It doesn’t mean that this benchmark has to be a constant, but it must display stability so that practical users can get a consistent guidance.

**10.7.2.2 Transaction cost-based frequency-guided method**

To avoid the dispute over the legitimacy of conversion system, we may take frequency-guided method. By this, we mean the client can take as the principle of selection the frequency of each system being chosen in dealing with the profit for given set of attributes. The basic idea is as follows. Suppose we can separate the past projects into two groups, success or failure, in some way and we possess data on project attributes and procurement system for a large sample of past successful projects. We can proxy the utility index as the probability of procurement system $R_i$ being chosen, $P(R_i)$, by assuming that the relative frequency of each procurement system being chosen among the successful projects reflects the probability that the project for given transaction attributes can be successfully delivered by this system. This means that, among all the systems able to successfully deliver the projects of given attributes in the past, the most frequently chosen one will be the one making your project most likely to be successfully completed in the future. Thus, this method suggests choosing an appropriate system on the basis of frequency of each procurement appearing in the past successful projects and this is where its name comes from.

To link $P(R_i)$ with transaction attributes, we have to assign a probability distribution to $P(R_i)$, say, a logistic function. That is, $P(R_i)$ can be expressed as

$$P(R_i) = \frac{e^{a'X + \beta_i}}{1 + e^{a'X + \beta_i}}$$  \hspace{1cm} \text{Eq.10-2}$$

, where $a$, $X$, $\beta_i$ are vectors for a constant, transaction attributes and coefficients. $R_i$, $i=0, 1, 2, \ldots, n$ represents $n$ types of alternative procurement systems. By taking the transaction attributes of the past $m$ successful projects as input data, we can estimate $\beta_i$ by way of statistical inference techniques, such as maximum likelihood method. Theoretically, $\beta_i$ play the role of bridging transaction attributes, $X$, and the efficacy of
procurement systems in carrying out the project with these attributes. This relation can be shown more explicitly by deriving the odd-ratio from Eq.10-2:

\[
\ln \frac{P(R_i)}{1 - P(R_i)} = a + X\beta_i
\]

Eq. 10-3

A positive coefficient of the transaction attribute, \(X_k\), for procurement system \(i\) means that the higher is \(X_k\), the more likely for procurement system \(i\) to be chosen and so its degree of desirability. It implies that procurement system \(i\) is relatively more capable of dealing with the projects with this attribute. Thus, the coefficients indicate the conditions under which a procurement system is more advantageous.

Taking the transaction attributes of a new project as input to the set of formulae with estimated coefficients \(\beta\), for each procurement system, we can find the procurement system most likely to have been chosen for similar past successful projects. The procurement system with highest probability is then the most appropriate procurement system for the new project.

Ideally, we need to select the successful projects (in the ex post sense) as our sample. The input data of this model would include: (1) the client’s actual (discrete) choice of procurement systems \(R_i\) and (2) the evaluation of transaction attributes of the project in question. To obtain reliable estimates of \(\beta\), a large good-quality sample is required. In practice, in collecting these data, we have to get over two obstacles. One is the definition of a successful project. The other is about the objective evaluation of the transaction attributes in a comparative basis across projects. First of all, though it is generally agreed that cost, time and quality are the most important dimensions for measuring project performance, there is still much work to be done, to devise a benchmark system, with clear cutoff points to tell failed projects from successful ones. Second, most clients need a supporting tool in order to assess the transaction attributes of a project. If there is no tool of this sort on hand, we cannot be certain whether any collected data is based on the consistent criteria. However, a theoretically sound tool serving this purpose is still awaited.

For these two reasons, the TCA-based procurement system selection method is not ready to be applied in practice yet. More academic efforts are called for to investigate
the fundamental issues of construction management, such as complexity, uncertainty and performance of the project. This knowledge will be an essential foundation for many applied studies in the field of construction management, including but not confined to construction procurement system selection.

10.7.2.3 Comparison of three approaches
A comparison of four approaches is summarized in Table 10-7, which we interpret as showing that a TCA-based approach is not only theoretically well founded and but also potentially able to improve some weaknesses of the current approaches. However, between the two possible application methods, their relative advantages are not clear. Performance prediction method uses performance measures as predictors aiming to show the possibility of the project in question being successfully delivered by way of a system. But how to justify the conversion of several performance indicators into an overall score is a problem.

In contrast, frequency-guided method can shrug off the worry as to the legitimacy of conversion system, but suffer the doubt as to how to draw a line between success and failure projects. Which one is better? There is no answer provided here. Our intention is to reveal the possible ways ahead and provoke thought so as to develop a theoretically sound and at same time practically applicable procurement selection technique.

Though several variables, like time, flexibility and complexity, used by the MAUA are also taken as crucial by the TCA, the reason why they are subsumed is quite different. Following the logic of multi-attribute utility approach, only the attributes of consequence of a decision can be justified as decision variables. That is, those variables relating to the appraisal of project implementation performance are what should be counted. However, the proposed variable in the literature displays a mixture of factors relating to the client’s requirement and project attributes. This has invalidated the aforementioned basic premise of multi-attribute, weakening the credibility of the derived conclusion. In contrast, every step of the TCA is closely linked with the postulated goal of profit maximisation, so the conclusion is relatively trustworthy.
Table 10-7 A summary table for the comparison of three approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>Strong points</th>
<th>Weak points</th>
</tr>
</thead>
</table>
| Multiple-attribute utility approach | A framework that can develop the selection technique for a wide range of procurement system alternatives | 1. Eclectic selection of priority variables;  
2. Spurious association of priority variables and procurement systems;  
3. Utility coefficients \((w, y)\) rely on the experts' opinions  
4. The method able to assess the attributes of a new project on the comparable basis across project is yet to be developed. |
| Statistical approach              | Utility coefficients can be objectively estimated                             | 1. Lack of a rigorous theoretical foundation.  
2. The method able to assess the attributes of a new project on the comparable basis across projects is yet to be developed.                                                                                       |
| TCA-based approach                | 1. Founded on a well-developed theory;  
2. Able to consider a wide range of procurement system alternatives;  
3. Utility coefficients can be objectively estimated | 1. The system for converting multi-dimension performance indicators into a single index is disputable (e.g. client's weight on each criterion is not considered)  
2. Benchmark points are floating.  
3. The method able to assess the attributes of a new project on the comparable basis across projects is yet to be developed.                                                                                       |
| performance prediction method      | 1. Founded on a well-developed theory;  
2. Able to consider a wide range of procurement system alternatives;  
3. Utility coefficients can be objectively estimated | 1. The criteria for success or failure of a project are debatable;  
2. The method able to assess the attributes of a new project on the comparable basis across projects is yet to be developed.                                                                                       |
| frequency-guided method            | 1. Founded on a well-developed theory;  
2. Able to consider a wide range of procurement system alternatives;  
3. Utility coefficients can be objectively estimated | 1. The criteria for success or failure of a project are debatable;  
2. The method able to assess the attributes of a new project on the comparable basis across projects is yet to be developed.                                                                                       |
Bibliography


Chang, C.Y. and G.Ive, 2001, Model For Understanding, Preventing, and Resolving Project Disputes: Discussion, mimeo, Bartlett School of Graduate Studies, University College London (UCL).


Franks, J. 1993, The Design-and-Build Approach to Procurement, Chartered Institute of Building, UK.


Grueneberg, S. L. and G. J. Iwe, 2000, The Economics of the Modern Construction Firm, Macmillan, Basingstoke, UK.


Hubbard, T.N., 1999, How Wide Is The Scope of Hold-up-based Theories? Contractual Form and Market Thickness in Trucking, working paper 7347, NBER.


Naoum, S.G. 1991, Procurement and Project Performance, Chartered Institute of Building (CIOB).


Appendix

Introduction to this questionnaire

1. This is a survey for investigating construction client procurement practices. The data collected by the questionnaire will be used to test a newly developed theory. Your answer will be extremely valuable to this study.
2. The preferable person to fill out this questionnaire is the in-house project manager or the person responsible for the project on the behalf of the client (from here on, your organisation is used to represent the client’s company or organisation).
3. The Glenigan database of industrial and commercial construction projects shows that you have recently carried out the following project, so would you please answer the questionnaire on the basis of the project with the site address Swinging Bridge I.
   If you weren’t involved in the above project as Glenigan database shows, or the project was cancelled or postponed, would you instead choose a project that you are handling or have handled recently. If there are several projects proceeding in parallel, please choose the project with larger value scale.
4. Please would you give your answer by ticking the boxes, circling the numbers or by writing in the spaces provided.
5. For helping you complete this questionnaire, stages in RIBA’s plan of work are listed as below. Would you please refer to the number of each stage (say, S2) in answering the questions, if applicable.

<table>
<thead>
<tr>
<th>Usual Terminology</th>
<th>Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Briefing Stage</td>
<td>(S1) Inception</td>
</tr>
<tr>
<td></td>
<td>(S2) Feasibility</td>
</tr>
<tr>
<td>Sketch Plans</td>
<td>(S3) Outline Proposals</td>
</tr>
<tr>
<td></td>
<td>(S4) Scheme Design</td>
</tr>
<tr>
<td>Working Drawings</td>
<td>(S5) Detail Design</td>
</tr>
<tr>
<td></td>
<td>(S6) Production Information</td>
</tr>
<tr>
<td></td>
<td>(S7) Bill of Quantities</td>
</tr>
<tr>
<td></td>
<td>(S8) Tender Action</td>
</tr>
<tr>
<td>Site Operations</td>
<td>(S9) Project Planning</td>
</tr>
<tr>
<td></td>
<td>(S10) Operations on Site</td>
</tr>
<tr>
<td></td>
<td>(S11) Completion</td>
</tr>
<tr>
<td></td>
<td>(S12) Feedback</td>
</tr>
</tbody>
</table>

6. If you have any difficulty in completing the questionnaire, would you get in touch with Mr. Chen-Yu Chang
   Tel/Fax: 0208-4458145
   mobile: 07932-751648
   Email: ucfycyc@ucl.ac.uk
Appendix

Section 1: Basic Information

1-1. Information about surveyee
(A) Name of your organisation ________________________________
(B) Name of person completing this questionnaire ____________________________
An easy way to fill out item (A)-(E) is to affix your card
(C) Position within firm ________________________________
(D) Telephone number ________________________________
(E) Business address ________________________________________
____________________________________________________________________________________

(F) Years of experience in construction projects ____________ years
(G) Date of completing this questionnaire (month/yr) ____________

1-2 Name of this project ____________________________________________________________

1-3. Scale of the project: £ __________ million

1-4. Duration of construction (detailed design and construction): __________ months

1-5. What category does the procurement method you used fall into?
☐ Traditional method (design then tender)
☐ Accelerated Traditional method (design and construction overlapped)
☐ Design and Build (D&B) ☐ Influenced Design and Build
☐ Design and Manage ☐ Management contracting
☐ Construction Management
☐ Others, please specify ____________________________________________

1-6. Did you choose the procurement method of this project specifically to fit project circumstances?
☐ Yes ☐ No

1-7. Did you use a partnering agreement in this project?
☐ No ☐ Yes, informal agreement
☐ Yes, formal agreement.

1-8. Month and year in which detailed designer or D&B contractor was appointed (month/yr)
Appendix

Section 2: Questions

Note: Please answer the following sets of questions on the basis of the information that was available to you at the point of appointing detailed designer or D&B contractor (i.e. please try not to answer with benefit of hindsight).

Question 2-1–2-4 are concerned with the importance of earlier completion of the project to your organisation.

To what extent do you agree with the following statement:

<table>
<thead>
<tr>
<th>Question</th>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>You are willing to pay a higher price in exchange for earlier completion of this project.</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
<tr>
<td>2-2</td>
<td>Getting the project completed as early as possible, even with higher costs, would do much good for the success of this project.</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>Earlier delivery of the project will greatly increase the positive benefits of this project to your organisation.</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
<tr>
<td>2-4</td>
<td>The benefits of getting the project completed earlier would by far outweigh the extra costs of achieving it.</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

Question 2-5–2-13 are concerned with the possible sources of uncertainty concerning events that might arise during construction that would then cause you to change the requirements in the original brief.

To what extent do you agree with the following statement:

<table>
<thead>
<tr>
<th>Question</th>
<th>Statement</th>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-5</td>
<td>It is highly likely that the original brief of this project may need to be changed to cope with changing external market requirements after engagement of the detailed designer.?</td>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
<tr>
<td>2-6</td>
<td>It is highly likely that this project may encounter ground conditions that may</td>
<td>Strongly Disagree</td>
<td>Strongly Agree</td>
</tr>
</tbody>
</table>
Appendix

change the original brief after the engagement of the detailed designer.

2-7 The revenue stream from the project, compared with other commercial or industrial buildings in general, is highly likely to be revised during the construction period.

2-8 It is highly likely that the enforcement of statutory regulations (such as planning, building and fire regulations) in the region of your site may require the original brief to be modified after the engagement of the detailed designer.

2-9 It is highly likely that you may face a more stringent budget constraint than originally foreseen that may cause you to change the original brief after the engagement of the detailed designer.

2-10. It is highly likely that design variations would happen during construction in terms of the general characteristics of your project.

2-11. The design of this project was expected to be very complex.

2-12 The construction technology to be employed in your project was anticipated to be relatively advanced.

2-13. Roughly estimate what (value) percent of the project will contain the innovative elements that haven't been built ever?

<table>
<thead>
<tr>
<th>Percent of the value of innovative elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 6 11 16 21 25 31 36 &gt; 5 10 15 20 25 30 35 40 41</td>
</tr>
</tbody>
</table>
Appendix

Question 2-14–2-17 are concerned with the extent to which your project posed unique or special requirements which only a limited number of designers or contractors might be able to meet.

To what extent do you agree with the following statement:

2-14 It is difficult to find a designer competent to undertake the detailed design for the project.

- Strongly Disagree
- Strongly Agree

1  2  3  4  5  6  7  8  9  10

2-15 It might be difficult to find a contractor competent to undertake your project.

- Strongly Disagree
- Strongly Agree

1  2  3  4  5  6  7  8  9  10

2-16 The contractor will need to have specific construction technology to undertake your project.

- Strongly Disagree
- Strongly Agree

1  2  3  4  5  6  7  8  9  10

2-17 The construction technology that your project was perceived to need is not easily available from the UK construction industry.

- Strongly Disagree
- Strongly Agree

1  2  3  4  5  6  7  8  9  10

Question 2-18 are concerned with total expenditure of your organisation on construction projects in the last five years.

2-18 Your organisation’s total expenditure on construction projects in the last five years? (please fill in approx. number) £m

Question 2-19–2-21 are concerned with difficulties in clearly identifying the liabilities of trade or subcontractors. Would you please answer the following questions even if you didn’t use management contracting. Please provide your answer on the basis of expectations beforehand instead of observations after the procurement route has been chosen, namely on the basis of the available information at the point of selecting detailed designer or D&B contractor.

To what extent do you agree with the following statement:

2-19 It might be difficult to clearly identify the liability of subcontractors involved in this project due to complicated interfaces between their works.

- Strongly Disagree
- Strongly Agree

1  2  3  4  5  6  7  8  9  10
Appendix

2-20 It could be expected that the overlapped area of responsibility between subcontractors might lead to serious difficulties in ensuring the quality standard of this project.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

2-21 This project involves the instalment of complicated facility equipment, such as communication systems so that this project might be more likely to encounter difficulty in getting over the interface problems among subcontractors.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>