ECONOMIC ANALYSIS OF HIGHWAY FRANCHISING

by

TAN Zhijia, B. SC, M. SC

A thesis presented to
The Hong Kong University of Science and Technology
in partial fulfillment of the requirement for
the degree of Doctor of Philosophy
in Civil Engineering

July, 2010, Hong Kong

AUTHORIZATION

I here declare that I am the sole author of this thesis.

I authorize the Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the Hong Kong University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutes or individuals for the purpose of scholarly research.

Tan Zhijia

SIGNATURE TAN Zhijia

ECONOMIC ANALYSIS OF HIGHWAY FRANCHISING

by

TAN Zhijia, B.SC, M. SC

This is to certify that I have examined the above PhD thesis and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the thesis examination committee have been made.

Approved by:

____

Prof. Hai YANG, Supervisor

____

Prof. H K LO, Acting Department Head

Department of Civil and Environmental Engineering

July 2010

TO MY FAMILY

ACKNOWLEDGEMENTS

First and foremost, I am heartily thankful to my supervisor, Prof. Hai Yang, who was abundantly helpful and offered invaluable assistance, guidance and support throughout my PhD study. His intelligent perspicacity and constructive comments not only helped me complete this thesis, but also shaped my understanding of scientific research. I want express my appreciation of his continual encouragement, consideration and patience, which benefited me greatly in my academic research and everyday life.

I am also deeply grateful to Prof. Hong Kam Lo, Prof. Donggen Wang, Dr. Xueqing Zhang and Dr. Hoyin Mak for their time and efforts to be committee members of my PhD examination. Their insightful comments and invaluable suggestions are very helpful to improve the quality of this thesis.

I specially thank Prof. William H. K Lam, Dr. Agachai Sumalee, Dr. Qiong Tian and Dr. Xiaolei Guo for their clever suggestions and discussions on many other topics within and beyond this thesis.

I also owe my thanks to all my friends and colleagues in our transportation lab. I am grateful for the wonderful time we spent together.

The most heartfelt thanks go to my wife and daughter for their unconditional love, encouragement and understanding. I am deeply indebted to my parents for their long time support.

Contents

List of Figures

List of Tables

ABSTRACT

Private provision of public roads through build-operate-transfer (BOT) contracts is increasing around the world in both developed and developing countries. Under a BOT contract, a private firm builds and operates roads in a road network at its own expense, and in return receives the revenue from road tolls for a number of years. Then these roads are transferred to the government. Important issues are the length of the concession period, capacity, toll charges and the benefits to the private investor, road users and the whole society under various market conditions and government regulations.

Several issues related to a BOT toll road project are studied. This thesis first develops a benchmark analysis of the BOT road contract viewed as a combination of concession period, road capacity and toll charge under the assumption that the demand and cost functions are known to both public and private sectors. A bi-objective optimization problem is proposed for maximizing social welfare and private profit. And the properties of the concession period, service quality, social welfare and profit gains along the Pareto-optimal frontier are studied. A variety of government regulatory regimes are also investigated. We prove that: the private sector tends to offer a lower road capacity and a lower service quality under the price-cap regulation, while it chooses a higher service quality, a higher capacity and a higher toll charge under the rate-of-return regulation than those associated with the corresponding Pareto-efficient solution. In contrast, we prove that both the demand and markup charge regulations lead to Pareto-optimal outcomes.

User heterogeneity in value-of-time (VOT) has long been a fascinating issue in road pricing

studies. This thesis investigates the effects of the user VOT heterogeneity on the properties of the Pareto-efficient BOT contracts. Under some technical conditions, we prove that, the Pareto-optimal road-life concession period is free from the effects of user heterogeneity, while, even with constant returns to scale, the service quality is not constant and dependent on the curvature of the mean residual VOT function. The effects of user VOT heterogeneity on the outcomes of various regulatory regimes are also investigated. We find that, both the demand and markup regulations fail to achieve the Pareto-optimal outcomes and result in a lower level of capacity and service quality.

In the presence of traffic uncertainty, this thesis proposes the BOT contract with full and partial flexibility according to the instruments adopted by the public and private sectors. Full flexibility refers to the case in which the public sector promises an exogenous rate of return on the private investment and in turn can freely ex post adjust the contract in a socially optimal manner according to the observed demand curve. Partial flexibility refers to the case where the public and private sectors agree on an ex ante demand risk allocation by contract and the ex post contract adjustment can be made contingent on a Pareto-improvement to both parties. A preferred Pareto-improvement can be selected from the Pareto-optimal solution set of a bi-objective programming problem equipped with a rational preference. In comparison with the traditional rigid contract, in which, private sector burdens all project risk, the proposed flexibility of the BOT contract is valuable for the contract adjustment mechanism to solve the discrepancy between the public and private sectors. Two other issues are also investigated in this thesis. One is how to select the optimal combination of concession period, capacity and toll levels for a BOT toll road project if the yearly increasing operation costs are incorporated in the project. Another is how to choose the capacity and toll levels of an "add-on" toll road parallel to an existing one with various ownership regimes, namely the existing road can be a free, public or private toll road.

1

INTRODUCTION

1.1 Background

How should society go about expanding its road systems? Who would decide where to provide more road capacity, and how much more? Where would funds for expansion come from? The recent world-wide tendency toward the introduction of commercially and privately provided public roads proves to be an efficient answer to these questions. Private participation in road construction and operations has the advantages of efficiency gains, private financing, and better identification of attractive investment projects. Such participation is generally implemented through a build-operate-transfer (BOT) contract, under which a private firm builds and operates roads in a road network at its own expense, and in return receives the revenue from road tolls for a number of years, and then these roads are transferred to the government. Such commercial and private provision of public roads has attracted growing interest in recent years and such plans are being used to finance modern road systems worldwide (Roth, 1996). Private participation in the form of BOT franchises has worked well in a number of projects such as the tunnels in Hong Kong. In mainland China, many local, mainly municipally affiliated companies have undertaken the development of toll roads in recent years, often in joint ventures with Hong Kong investors (Tam, 1998). Once road provision is market driven, many issues must be carefully addressed, because the interests of the private sector are different from those of the public sector. From the viewpoint of private sector, the profitability of a project is of great concern because private firms are put at risk. From the viewpoint of the public sector, it is meaningful to assess whether the construction of a road will lead to a positive welfare gain and also be profitable so that private provision is worthwhile. In a BOT project, private sector receives a concession to finance, build and operate a road over a set period of time, in exchange for the right to charge the users of the road at a rate which makes the investment commercially viable. At the end of the concession period the road is turned over to the state (Handley, 1997).

Private provision of roads is driven by a number of factors. A primary motivation is a widespread belief that the private sector is inherently more efficient than the public sector, and therefore builds and operates facilities at less cost than the public sector. Also, the public sector, facing taxpayer resistance, may simply be unable to finance facilities that the private sector would be willing and able to undertake for a profit (Gomez-Ibanez et al., 1991). In addition, if new road space is provided as an "add-on" to an existing network system, and if road users find it worthwhile to patronize this new road and pay charges, and if the charges cover all costs (including congestion and environmental costs), all may gain benefit, and there would be no obvious losers. Even those who do not use these new roads would benefit from reduced congestion on the old ones (Mills, 1995).

Once road provision moves to a market economy, there are many intriguing issues to be addressed. From the viewpoint of private investors, profitability of a project is of great concern because their private firms are put at risk. Specially, the paucity of experience with toll-charged roads and the long pay-back period after the initial investment make it more difficult to predict future revenues and thus make the uncertainty and risks of investment in roads very high compared to alternative investment options (Nijkamp and Rienstra, 1995). Therefore, careful project selection and a clear identification and assessment of the risks are important in order to guarantee that willing buyers are prepared to pay sufficiently to induce willing suppliers to provide it. From the viewpoint of society, it is meaningful to assess whether the construction of a road will give a positive welfare increment if it is profitable, compared with the do-nothing alternative, and vice versa whether any road which adds to welfare will be profitable, and hence can be provided privately. More importantly, under the commercial and private provision of roads, what roles can a government play? What government regulations concerning private toll roads should be imposed? For example, should the road charge level be controlled by the government to prevent private firms from abusing their monopoly over the road and for equity reasons (access to services for everyone)? It is obvious that these issues must be addressed carefully, because transport infrastructure has major strategic, economic, social, financial and environmental effects and because the interests of the private sector are different from those of the public sector (Nijkamp and Rienstra, 1995).

1.2 Objectives

Private provision of roads is typically through a build-operate-transfer (BOT) contract. The BOT projects use the market criterion of profitability for road development and rely on the voluntary participation of private investors who hope to benefit financially from their participation. A BOT contract generally involves three fundamental decision variables: the concession period, the road capacity and the toll charge. These three variables are crucial for both the private firm and the government to reach their respective objectives: the private firm wishes to undertake the road project for a maximum profit throughout the concession period; while the government aims to maximize social welfare throughout the whole life of the road when awarding the road concession contract. The concession period, the number of years operating the road for the private firm, directly governs the private firms total toll revenue as well as the total social welfare gain during the life of the road (the concession period and the post-concession period after the contract ends). The selected road capacity affects both the private firms profit and the total social welfare directly or indirectly. First, the road capacity determines the construction cost, the major investment cost of the private firm for the road project; second, the road capacity affects its congestion, and thereby travel times, which in turn affects the travel demand and, as a result, toll revenue and the social welfare. The toll charge to a large extent determines the total revenue that the private sector operator receives and it influences the social welfare gain during the concession period as well. In summary, each of the three fundamental variables plays an important role in forming a feasible BOT contract. Their values will determine the profitability and welfare gain of the project, and thus must be selected carefully, because transport infrastructure has major strategic, economic, social, financial and environmental effects, and because the interests of the private sector are different from those of the public sector.

This thesis intends to address these important issues in the development of private toll roads. The results can help potential private investors identify how, and under what circumstances, a highway BOT project is feasible and profitable, and assist the public sector in understanding how a proposed project will benefit the private investor, road users and the whole of society. Therefore, the proposed topic is of widespread interest and with many potential applications.

The main objectives of the research are: to use bi-objective optimization approach for setting Pareto-efficient BOT contracts that strike a balance between public and private interests in the procuring of road investment; to study how to reach the optimal BOT contracts through bilateral negotiations.

With a practical consideration, the study will be extended to cover the following topics: to examine the effects of the user heterogeneity on the properties of the Pareto-efficient BOT contracts and the outcomes of various regulatory regimes; to develop a theocratically tractable method to deal with the demand uncertainty for designing the contract of the BOT toll road project; to explore other possibility of extensions and inspire other intriguing theoretical and practical topics, such as, optimal BOT toll road contract design with road deterioration and user damage, transportation network capacity investment decision with the mixed ownership regimes.

1.3 Outline of the Thesis

The remainder of this thesis is organized as follows. Chapter 2 carries out a literature review on related fields. Four branches literatures most relevant to this thesis are reviewed: pricing, capacity choice and road financing; user heterogeneity and road financing; highway franchising; concessionaire selection and regulations.

In Chapter 3, we consider a bi-objective optimization problem for maximizing social welfare and private profit to capture the different interests between the public and private sectors. The properties of Pareto-efficient solutions for the bi-objective optimization problem are investigated, and the efficiency of various regulatory regimes are studied.

Chapter 4 extends the study of the previous chapter to consider the user heterogeneity by assuming that the value-of-time of the travelers follows a continuous distribution. Introducing two well-known concepts in reliability, we develop a theoretical method to examine the effects of user heterogeneity on the properties of the Pareto-optimal solutions and the efficiency of the regulatory regimes studied in the previous chapter.

In Chapter 5, we propose the full and partial flexibility of the BOT contract according to the instruments adopted by the public and private sectors to deal with the demand uncertainty. The public sector can freely ex post adjust the contract in a socially optimal manner according to the observed demand curve in fully flexible contract with a promised exogenous return rate on the private investment. While, in the partially flexible contract, an ex ante demand risk allocation is agreed by both parties and the ex post contract adjustment is required to be Pareto-improve. The optimal flexible BOT contracts with the two types of flexibility are examined in this chapter.

Two other issues are also investigated in this thesis. One is how to select the optimal combination of concession period, capacity and toll levels for a BOT toll road project if a yearly increasing operation cost resulted from road deterioration and maintenance effects is incorporated in the project. Chapter 6 models the BOT problem as an isoperimetric problem in calculus of variations to maximize the social welfare with a profit constraint explicitly incorporating the effect of road deterioration and maintenance over the years. The maintenance cost depends on the traffic loads, road capacity and road natural deterioration. We analyze the properties of the optimal pricing policy and propose an efficient regulatory to achieve the optimal contract. The effects of economic growth on the solution properties of the problem are investigated in this chapter. Another is how to choose the capacity and toll levels of the "add-on" links for the government once the emergence of the various ownership regimes in the transportation network because of the increasing private provision of public roads. Chapter 7 theoretically analyzes the problem of when and how to invest a new competitive toll road project to an existing alternative with various ownership regimes.

A summary of this thesis is given in Chapter 8. We also point out some topics that justify as the future research subjects.

2

LITERATURE REVIEW

It has long been regarded incorporating congestion costs into road prices as essential to an efficient use of road since road use has negative congestion externality (e.g., Knight, 1924; Pigou, 1920; Walters, 1961). To internalize the congestion externality and thereby maximize social welfare, the classic marginal cost pricing principle states that road users of congestion roads should pay a toll equal to the difference between the marginal social cost and the marginal private cost. Although the road pricing theory is directly related to the study, we intend not to repeat the review on road pricing studies. We should mentioned here some more comprehensive reviews or monographs on road pricing include (certainly not limited to) Lo and Hickman (1997), Yang and Verhoef (2004), Small and Verhoef (2007), Yang and Huang (2005), Lindsey (2006). In this chapter, we focus the literature review on the theory of road financing and highway franchising.

2.1 Pricing, Capacity Choice and Road Financing

Traditionally, large transportation infrastructure projects used to be funded primarily by governments. Therefore, naturally, there have been a number of studies looking at the linkage between the optimal congestion charge and road investment strategy in a first-best environment. In this environment, the public planner sets a user fee equal to the congestion externality. And, at the optimal capacity level, the ratio of the revenue from the congestion charge to the cost incurred for capacity equals the elasticity of capacity cost with respect to capacity. The ratio is the degree of self-financing (Mohring and Harwitz, 1962; Verhoef and Mohring, 2009). The special case with constant scale economics (the elasticity is equal to one) is the exact self-financing theorem. Namely, the revenues from optimal congestion pricing, equal to the difference between the marginal social cost and the marginal private cost at any level of traffic flow, will be just sufficient to finance the fixed costs associated with the optimal capacity supply. The conditions are satisfied when there are constant returns to scale in road construction and maintenance, and the capacity can be increased in continuous increments. There will be a deficit when there are economies of scale in capacity construction (the elasticity is less than one), and a surplus with dis-economies (the elasticity is less than one). These results are entirely consistent with basic micro economic insights that tell us that the private sector that is forced to apply marginal cost pricing will face a deficit under economies of scale, a zero-surplus under constant scale economies, and a surplus under dis-economies of scale.

This classical exact self-financing result was later expanded in various ways. Henderson (1985) showed that self-financing is achieved in a case where users have a choice between two modes, which could be either driving or taking transit, or choosing between two roads. Road damage cost was incorporated (Newbery, 1988, 1989) so that the road charge, equal to the sum of the optimal congestion charge and road damage charge, covers the fixed operating cost and the variable maintenance cost provided there are constant returns to scale in construction and use of road capacity. Oum and Zhang (1990) examined the relationship between congestion tolls and capacity expansion costs when capacity expansion is indivisible. Arnott and Kraus (1998b) addressed the issue that when anonymous congestion charges are consistent with marginal cost pricing in the presence of user heterogeneity. They find that the self-financing theorem remains valid in present value terms, provided the size of capacity additions is optimized conditional on the timing of investments. This is true whether or not capacity is added continuously or intermittently and whether or not the timing of investments is optimal. Small (1999) examined the conditions under which a congestible facility is self-financing in the presence of non-competitive factor markets. Yang and Meng (2002) showed that the self-financing theorem holds for each road in a full network individually and consequently for the network in aggregate, provided each link is optimally priced and all capacities are optimized.

More studies contributed to answer the question how the under-pricing alternative affects the optimal road width and the optimal fare in an imperfect or second-best environment, in which, there is a unalterably under-pricing congestion road (Levy-Lambert, 1968; Marchand, 1968). Verhoef et al. (1996) and Yang and Huang (2005) pointed out that in the presence of an un-tolled substitute to a new toll road, the second-best toll falls short of the marginal external congestion cost on the new route and the resulting revenue will fail to cover the capacity cost, and the second-best road width exceeds the first-best level. Arnott and Yan (2004) carried out a comprehensive review of the two-mode problem and pointed out that there is still a long way to understand the toll and/or capacity decision problem in the second-best environment because of many economic and political issues, such as, the demand of the two roads is imperfect substitution. They proposed two different functions to model the demand levels of the two modes both dependent on the full cost of the two roads and synthetically discussed this kind of problem.

To guarantee the efficient welfare outcome and financial equilibrium, another second-best solution concept is to maximize the social welfare with a zero- or positive profit constraint when considering the toll and capacity choice. Yang and Meng (2000) investigated the profitability or self-financing and social welfare gain of a single new toll road in a general network through numerical experiments. Verhoef (2008) and Verhoef and Rouwendal (2004) addressed some implications of second-best congestion pricing for the applicability of the self-financing theorem, using a numerical experiment approach.

2.2 User Heterogeneity and Road Financing

A key assumption of the many standard studies presented in the previous section is that users experience the same disutility from travel delay. The concept of value-of-time (VOT) as a pivotal toll in capturing the travelers' tradeoffs between money and time is now common in transportation network assignment literature. There are in general two lines of approaches to dealing with the tradeoffs between cost and time. A first line of approach consists of differentiating several discrete classes of users, each one with a VOT belonging to some interval (Dafermos, 1973; Daganzo, 1983; Guo and Yang, 2009; Yang and Huang, 2004). The second line of approach assumes a continuously distributed VOT across users that can be inferred from the income distribution of the population (Dial, 1999a,b; Leurent, 1996).

In fact, user heterogeneity in VOT has long been a fascinating issue of road pricing studies in the transportation literature. Edelson (1971) analyzed and compared the socially optimal and monopolistic toll levels in a two-mode network. He showed that, in his model formwork, the monopolistic toll level is identical with the socially optimal level when the commuters have a homogeneous VOT, but may be lower or higher than that at social optimum if the commuters vary in their VOTs. Mayet and Hansen (2000) investigated second-best congestion pricing with continuously distributed VOT for a highway with an unpriced substitute. They studied different optimal tolls depending on whether the social welfare function is measured in money or in time, and whether toll revenue is or is not included as part of the benefit. Small and Yan (2001) investigated the effects of the degree of user heterogeneity on the efficiency of the toll policy by enlarging the VOT difference of two user classes and keeping the average VOT as a constant. They suggested that ignoring user heterogeneity may lead to serious underestimation of the efficiency of a value pricing highway policy. Verhoef and Small (2004) examined the extent of the effects of user heterogeneity on pricing policy by varying the type of the VOT distribution used in their numerical examples. They claimed that VOT distribution as "imperfect information" should be considered when applying a road pricing policy. With explicit incorporation of user heterogeneity, Liu et al. (2009) and Nie and Liu (2010) examined the existence of Pareto-improving toll schemes in a two-mode network, and Guo and Yang (2010a) explored Pareto-improving road pricing and revenue refunding schemes in general networks.

In the context of simultaneous choice of toll level and road capacity, Yang et al. (2002)

examined the impact of user heterogeneity on the profitability and social welfare gain of a new toll road under various combination of toll level and capacity in a general network. Xiao and Yang (2008) examined the likely bias of a monopoly market away from the social optimum under the assumption that each trip-maker has a unique VOT when a new toll road is parallel to an existing free road connecting a given origin and destination. They investigated the efficiency loss and the road capacity and/or toll set by a monopolist under different kinds of government regulatory regimes. Light (2009) investigated the optimal toll and capacity decision with a continuous VOT distribution in the case of value pricing that involves dividing a highway into free and priced lanes so that in equilibrium the highway effectively operates at two levels of service. The planner sets the toll and chooses the amount of capacity to provide the two levels of route service (the express and high occupancy/toll lanes). He proved that it is beneficial to the planner to provide the differentiated service in the sense of the aggregate social cost.

2.3 Highway Franchising

Governments today are increasingly finding themselves either unwilling or unable to finance a growing number of new infrastructure activities. They are looking for new ways to provide public services at a lower cost to the taxpayers. Private provision of roads is typically undertaken through a build-operate-transfer (BOT) contract, which becomes a major means of the public-private-partnership (PPP) projects. Under a BOT contract, the private sector builds and operates the public facility at its own expense, and in turn receives the revenue from user fee for a period before the facility is transferred to the public sector. It is different from privatization and nationalization. In particular, BOT differs from privatization, in which, a government owns the entity first and then transfers (sells) it to the private sector (Qiu and Wang, 2009). The former refers to the bundling of the design, building, finance, and operation of the project, which are contracted out to a consortium of private firms for a long period of time, usually 25-30 years. Furthermore, the public sector specifies the service it wants, and some basic standards, but it leaves the private sector with control rights over how to deliver the service. In contrast, in privatization, the government only adopts the economic instruments to regulate the private firm, who buys the entity (Bennett and Iossa, 2006). Therefore, the first-of-all question is to answer whether the BOT scheme is better than the conventional provision (nationalization) and regulated privatization. Based on incentive analysis, Bennett and Iossa (2006) showed that the desirability of bundling the building and management operations depends on the externality of building investment. It is beneficial to the public sector to bundle the building and management operations if increasing the building investment leads to a fall in the cost of providing the service; and vice versa. Qiu and Wang (2009) proposed a new BOT framework to deal with the unverifiable project-quality, in which, the BOT project lasts for two operation periods. The private sector automatically operates the project in the first period after building it, and the public sector determines whether or not to extend the concession based on the observed quality and pre-determined rules. Engel et al. (2008) systematically studied the economic and political issues to answer when to adopt the PPP form and how to determine the optimal contract.

In this thesis, we suppose that a BOT toll road project has been adopted by the public sector, who contracts with the private sector to achieve several simultaneous objectives: construction and operation of the project at minimum cost; provision of quality services to drivers; efficient use of capacity by appropriate pricing; expansion of capacity according to social needs; and the financial equilibrium of the concessionaire (Nombela and de Rus, 2004). The above mentioned analyses in Sections 2.1 and 2.2 have focus on capacity choice and/or toll setting and the resulting profitability and social welfare gains. The concession period was usually assumed to be given and fixed (the converted unit cost of capacity per unit period was thus a given constant). Indeed, the fixed term concessions have been standard contracts between the public sector and private operators. However, demand uncertainty and fixed term contracts can make it impossible to fulfill the concession agreement in many cases, and contract renegotiation has been used to restore financial equilibrium. This has some undesirable economic consequences: selecting the most efficient concessionaire is no longer guaranteed, and prices lose their role as signals of allocative efficiency (Engel et al., 1997, 2001; Nombela and de Rus, 2004). The flexible-term contracts for road franchising have been proposed for achieving efficient pricing and cost recovery without contract renegotiation, and they can be implemented fairly easily. For example, in a least-present-value-of-revenue auction, the bidding variable is the present value of toll revenues, the lowest bid wins, and the franchise ends when the amount has been collected (Engel et al., 2001; Nombela and de Rus, 2004). In this case the linkage between traffic uncertainty and revenue uncertainty is effectively broken; the contract term is endogenously determined by the realized level of future demand, so it is shortened in case of high demand and extended if traffic levels are low. The simulation-based approach and Monte Carlo simulation technique are common tools in engineering literature to quantify construction and market risks for decision making (Zhang, 2009). A critical simplifying assumption of those studies is that traffic congestion is ignored or the travel time and thus traffic demand is independent of the road capacity or initial investment level for the new road (equivalent to assuming that the road capacity was predetermined and large enough). Recently, Guo and Yang (2010b) conducted a preliminary study on the selection of the concession period with deterministic demand and homogeneous users. They incorporated all three essential variables (concession period, road capacity and toll charge) and explicitly considered traffic congestion and demand elasticity for unconstrained and profit-constrained welfare-maximizing BOT contracts. They found that it is better to set the concession period as long as possible in the sense of the social welfare. Although their analysis is based on the perfect information (demand and cost are known to both parties), the result is still practically meaningful even the demand is uncertainty according to Engel et al. (2008). The latter showed that, under uncertain demand, given a minimum revenue guarantee and a revenue cap restriction, the flexible concession period is last definitely except the realized demand level is high enough to exceed the revenue cap. They assumed an exogenous upfront investment (no congestion).

2.4 Concessionaire Selection and Regulations

Concessionaire selection and regulation are two significant topics once the public sector decides to provide the road service using BOT schemes. In highway franchising, concessionaires are usually selected through auctions at which candidates submit bids for tolls or payments to the government. Engel et al. (2001) investigated the optimal full-information contract for each demand state by assuming that the travel demand is revealed, immediately, after the highway is built. Since the auction takes place before demand is realized, the bidding variable cannot be state-contingent. They showed that the toll and franchise term can not achieve the desirable optimal contract. They then proposed a Least-Present-Value-of-Revenue (LPVR) bidding mechanism: the public sector first announces the discount rate and the toll associated with each demand state; and the private sectors bid the least present value of toll revenue; the lowest wins the bid. They proved that the LPVR auction implements the optimal contract. In the spirit of Engel et al. (2001), also allowing a flexible franchise term, Nombela and de Rus (2004) proposed a bi-dimensional auction bidding for total net revenue and maintenance cost. Since the maintenance cost is fixed and invariant to the level of demand, the bi-dimensional bidding, essentially, is similar to the LPVR auction because that the public sector can promise to burden the maintenance cost and only let the potential concessionaires bid the least present toll revenue. Another problem is that the maintenance cost is hard ex post verified. But, the bundling the maintenance cost in bidding is important and practically meaningful. Recently, without considering the concession period and demand uncertainty, Ubbels and Verhoef (2008) analyzed capacity choice and toll setting by private investors for a toll road competing a free alternative in a competitive bidding framework organized by the government. They considered capacity and toll selection based on four auction mechanisms including minimizing the full cost of the users, maximizing capacity, minimizing subsidy (second-best solution results in a negative surplus in presence of the free alternative), minimizing travel costs and subsidy divided by total traffic demand. They compared the resulting welfare gains (or losses) from the first-best welfare outcome. In a numerical example, they claimed that the fourth auction mechanism is most efficient and nearly results in the socially optimal welfare level in the second-best environment. Furthermore, when the capacity level of the free road approaches to zero (no competitive alternative), both the second and fourth mechanisms are efficient and result in first-best welfare outcome. The degenerative situation corresponds to the single-road network. The result has been proved by Guo and Yang (2010b) and says that the highway patronage is an efficient bidding variable resulting in the optimal contract. Differentiating to Engel et al. (2001), Guo and Yang (2010b) assumed that the capacity or the upfront investment is a contract variable besides the franchise term and toll charge, and the demand is deterministic.

Government economic regulations, besides the safety and environment intervene, are necessary once she faces the profit-maximizing natural monopoly to provide the road service, which might result in abusive pricing and harm to movability. The economic regulations are increasingly adopted by the governments as more and more private participant in transportation industries (railway, airport and roadway). Traditionally, governments have relied on rate-of-return (ROR) regulation, under which they guarantee that operators will recover their costs and make enough money to remunerate investors, and thus, also called cost-plus regime. under this regime, the operators do not have a strong incentive to cut costs. As a result, the government introduced the price-cap regulation, which results in a high incentive to cut costs, but, simultaneously, a low investment on service quality or disregarding safety. Furthermore, the ROR regulation entails little risk since cost recovery is almost guaranteed whatever the demand, while price-cap regulation shifts all risk onto the private sectors (Train, 1991). For a single highway, Guo and Yang (2010b) proved that regulating the travel demand (ensuring the minimal user number) is efficient even when the profit-maximizing private sectors are allowed to select any combination of the franchise term, toll and investment level. Considering the two-road network, Tsai and Chu (2003) studied regulation alternatives for governments on private highway investment under a BOT scheme set up to protect both road users and private sector. The impact of various circumstances (such as a minimum flow constraint) on traffic demand, as well as users' cost, profit levels, and welfare were explored. They found that a BOT project with regulation performs between the cases of maximizing welfare and that of maximum profit.

The competition between several franchisers in road market sometimes may substitute for the regulations since there are various alternatives to traverse between one OD pair (Engel et al., 2004). An increasing interest is viewed recently in the topic of toll road competition. The issue was first addressed by De Vany and Saving (1980). In their model a fixed number of firms operate identical parallel roads between a single origin-destination pair with elastic demand. De Palma (1992) also examined the competition in terms of toll charges on two parallel private roads. The analyses showed that private competitive supply and pricing of highways can be welfare improving, but much should be done to understand what types of regulation have to be imposed on the private oligopoly or monopoly. Yang and Woo (2000) examined graphically the competitive Nash equilibrium by considering two toll roads provided and operated by two profit-maximizing private firms. Engel et al. (2004) and Acemoglu and Ozdaglar (2007) examined efficiency losses of toll road competition for some simplified case under certain restrictive assumptions. The capacity investment decision incorporated in the private toll roads competition is conducted in recent research. De Borger and Van Dender (2006) analyzed a model with two substitute congestible facilities under three administrative regimes: (a) social optimum, (b) monopoly, and (c) duopoly in a sequential capacity-then-toll game. Verhoef (2007) examined the efficiency impacts of private roads in initially un-priced networks. His study allows for capacity and toll choice by private operators, and endogenise entry and therewith the degree of competition, allowing for both parallel and serial competition. Xiao et al. (2007a,b) studied both toll and capacity competition among private asymmetric roads with congestion in a network with parallel links. They showed that the volume-capacity ratio or service quality of each road makes no difference for a network of parallel links between the oligopolistic market and the social optimum under the assumption of homogeneous of degree zero link travel time functions and constant return to scale in road construction. Mills (1995) discussed the possibility of divergence between profit and welfare for a tolled link in a simple road network. In particular, he showed that a link that is profitable can nevertheless entail a welfare decrement. De Palma et al. (2007a) studied maintenance and tolling decisions by two competing private operators of roads that experience depreciation and congestion. They found that the system performs better when maintenance and tolling decisions are made simultaneously rather than sequentially, because firms in the sequential game curtail maintenance in the first stage in order to soften toll competition in the second stage.

However, in reality, it is impossible for many parallel roads to enter the market simultaneously because of the high up-front investment. The entry-time id different and thus decision variables are asymmetric: an "add-on" road can be determined by capacity and toll while only the toll level can be adjusted for an existing toll road. And the ownership regimes in the road networks tend to be diversified, which results in a mixed oligopoly market, such as the private firms aim to maximize profit while the government cares about the social welfare of the whole network. The well-known BOT schemes are the three harbour tunnels in Hong Kong (Hong Kong Cross-Harbour Tunnel, Eastern Harbour Crossing and Western Harbour-Crossing), operating from beginning 1972, 1989 and 1997 and owning for a 30-year concession period, sequentially. Because of the early termination of the concession, two of the tunnels are operated by two different private firms, respectively, and one is controlled by the government. De Palma and Lindsey (2000) investigated the toll mechanism and efficiency gains under the environment of two parallel roads with three private ownership regimes: (1) a private road on one route and free access on the other, (2) a private roads duopoly, and (3) a mixed duopoly with a private road competing with a public toll road.

2.5 Summary

This chapter gives a literature review on the past studies and figures the background for this thesis. Most previous analyses of road investment have focused on capacity choices and setting tolls and the resulting profitability and social welfare gain, specially, in a first-best or second-best environment. A lot of issues challenge the researchers once the private sector participates the public highway projects: the public sector is not be the unique decision-maker in a highway project; the concession term plays a key role as the capacity and toll levels to affect the welfare and profit gains of the project; the strategic behavior between parties affects the welfare and profit outcomes of the project. Although those issues attracted the attentions of several economists, their analysis followed a general model framework by simply assuming a given capacity investment. This thesis conducts an investigation on the highway franchising problem by incorporating the concession term, capacity and toll as the decision variables and taking into account both profit and welfare outcomes.

3

PARETO-EFFICIENT BOT CONTRACTS AND REGULATIONS

Two objectives associated with a build-operate-transfer (BOT) toll road project, should be considered, namely, to maximize the social welfare and to maximize the profit. This chapter proposes a bi-objective programming problem to analyze the benefit of a BOT toll road project. We investigate the properties of the Pareto-efficient solution of the problem, and then discuss various regulatory regimes including the rate-of-return, price-cap, demand and markup charge regulations. The behavior of the profit-maximizing private sector under those regulations and the resulting efficiency are studied.

3.1 Introduction

Private provision of roads is typically undertaken through a build-operate-transfer (BOT) contract. Under a BOT contract, the private sector builds and operates the road at its own expense, and in turn receives the revenue from road tolls for a period before the road is transferred to the government. BOT projects use the market criterion of profitability for road development and rely on the voluntary participation of private investors who hope to benefit financially from their participation. A BOT contract generally involves three fundamental decision variables: the concession period, the road capacity and the toll charge. These three variables are crucial for both the private firm and the government to reach their respective objectives: the private firm wishes to undertake the road project for a maximum profit throughout the concession period; while the government aims to maximize social welfare throughout the whole life of the road when awarding the concession contract. The concession period, the number of years operating the road for the private firm, directly governs the private firms total toll revenue as well as the total social welfare gain during the life of the road (the concession period and the post-concession period after the contract ends). The selected road capacity affects both the private firms profit and the total social welfare directly or indirectly. First, the road capacity determines the construction cost, the major investment cost of the private firm for the road project; second, the road capacity affects its congestion, and thereby travel times, which in turn affects the travel demand and, as a result, toll revenue and social welfare. The toll charge to a large extent determines the total revenue that the private firm receives and it influences the social welfare gain during the concession period as well. In summary, for a congestible toll road with elastic demand, each of the three fundamental variables plays an important role in forming a feasible BOT contract. Their values will determine the profitability and welfare gain of the project, and thus must be selected carefully, because transport infrastructure has major strategic, economic, social, financial and environmental effects, and because the interests of the private sector are different from those of the public sector.

In this chapter, we propose to view a BOT contract as a combination of concession period, road capacity and toll charge. In view of the different interests between the public and private sectors, we consider a bi-objective optimization problem for maximizing social welfare and private profit, with respect to the three primary variables of concession period, road capacity and toll charge, for a given toll road. By assuming that the government and the private firm both have perfect information on the project cost and future traffic demand, we examine the properties of the Pareto-optimal solution set. Each Pareto-optimal solution dictates a Pareto-efficient BOT contract that leads to an efficient outcome in the sense that neither social welfare nor private profit can be further enhanced without reducing the other.

Another concern of the BOT toll road project is how to achieve the Pareto-efficient contracts. This chapter focus on the discussion of the government economic regulations by assuming that the private sector is the natural monopoly. The behavior of the profit-maximizing private sector and the welfare outcomes under those regulations should be first understood and investigated by the public sector. The efficient regulatory mechanisms then be selected to achieve a predetermined desirable Pareto-optimal outcome.

The chapter is organized as follows. Section 3.2 introduces our bi-objective programming formulation of a BOT toll road scheme and the definition of Pareto-efficient contracts. Some important properties of the set of Pareto-efficient contracts are explored in Section 3.3. A further analysis of the efficiency of a given Pareto-optimal contract in comparison with the social optimum is conducted in this section. Section 3.4 extends the analysis to the general cases of decreasing and increasing returns to scale in road construction. Section 3.5 investigates a variety of regulatory regimes for the government to achieve a predetermined Pareto-efficient contract. Numerical examples are used to elucidate our results in Section 3.6, and, finally, conclusions are presented in Section 3.7.

3.2 Basic Definitions and Assumptions

Assume that the government wants to get a private firm to build a new highway whose technical characteristics are exogenous. Let $y \geq 0$ be the capacity of the new road, $q \geq 0$ be the travel demand and $B(q)$ be the inverse demand function (or the marginal benefit function), and $t(q, y)$ be the link travel time function. Note that q and y are measured in the number of vehicles per unit period. The following demand-supply equilibrium condition always holds:

$$B\left(q\right) = p + \beta t\left(q, y\right), \tag{3.1}$$

where p is the toll charged to each user of the road and β is the value-of-time to convert time into an equivalent monetary cost (we consider homogeneous users only). Condition (3.1) simply means that travel demand for the new road is determined by the full price for a trip. Let $I(y)$ be the construction cost of the highway as a function of capacity. The following assumption is made about $B(q)$, $I(y)$ and $t(q,y)$ throughout the chapter.

Assumption 3.1. *(a) The inverse demand function, $B(q)$, is a strictly decreasing and differentiable function of q for $q \geq 0$; $qB(q)$ is a strictly concave function of q for $q \geq 0$.*
(b) The road construction cost function, $I(y)$, is a continuously increasing and differentiable function of y for $y \geq 0$.
(c) The travel time function, $t(q,y)$, is a continuously differentiable function of (q,y) for $q \geq 0$ and $y \geq 0$; for any $q > 0$, $t(q,y)$ decreases with y; for any $y > 0$, $t(q,y)$ is a convex and increasing function of q.

From equilibrium condition (3.1), the toll, p, can be viewed as the following function of the demand, q, and the capacity, y:

$$p(q,y) = B(q) - \beta t(q,y). \tag{3.2}$$

For a given y, the toll, p, is uniquely determined by the demand, q. Therefore, determining the variables p and y is essentially equivalent to selecting the variables q and y. Hereafter, the demand, q, is substituted for the toll, p, for convenience of exposition.

We first consider the private firm's problem. Let $\hat{T}$, $\hat{T} > 0$, be the life of the road under consideration. The private firm must choose a combination of the BOT variables, including the concession period, T, with $0 \leq T \leq \hat{T}$, the travel demand q (or equivalently, the toll charge, p) and the road capacity, y, to maximize its profit, $P(T,q,y)$, during the concession period T:

$$P(T,q,y) = Tqp - I(y), \tag{3.3}$$

where the travel demand, q, is determined by condition (3.1), the first term of Eq. (3.3) is the total toll revenue collected by the private firm during the concession period and the second term is the construction cost, which is fully borne by the private firm. With Eq.

(3.2), problem (3.3) can be rewritten as:

$$P(T,q,y) = TqB(q) - \beta Tqt(q,y) - I(y). \tag{3.4}$$

From Part (c) of Assumption 3.1 that $t(q,y)$ is convex, $qt(q,y)$ is convex in q for any given $y > 0$. From the strict concavity of $qB(q)$, the profit function, $P(T,q,y)$, is strictly concave in q for any given y and T.

Next, we consider the government's problem of choosing the best combination of the BOT variables (T,q,y) to maximize the social welfare during the whole road life, $\hat{T}$:

$$W(T,q,y) = TS(q,y) + \left(\hat{T} - T\right)\tilde{S}(y) - I(y), \tag{3.5}$$

where $I(y)$ is again the construction cost, $S(q,y)$ and $\tilde{S}(y)$ are, respectively, the unit-time social welfare during the concession and post-concession periods and determined below:

$$S(q,y) = \int_0^q B(w)\,\mathrm{d}w - \beta qt(q,y); \tag{3.6}$$

$$\tilde{S}(y) = \max_{q_1 \geq 0} S(q_1,y) = \max_{q_1 \geq 0} \int_0^{q_1} B(w)\,\mathrm{d}w - \beta q_1 t(q_1,y). \tag{3.7}$$

Equation (3.7) implies that, during the post-concession period, the road capacity is given and fixed and the government can select the optimal traffic volume to maximize the unit-time social welfare only.

Therefore, the BOT problem can be defined as selecting simultaneously the combination of the three variables (T,q,y) to maximize the total social welfare and the private firm's profit, which can be formulated as the following bi-objective programming problem:

$$\max_{(T,q,y)\in\Omega} \begin{pmatrix} W(T,q,y) \\ P(T,q,y) \end{pmatrix}, \tag{3.8}$$

where $\Omega = \left\{(T,q,y) : 0 \leq T \leq \hat{T}, q \geq 0, y \geq 0\right\}$ and social welfare, $W(T,q,y)$, and profit, $P(T,q,y)$, are defined by (3.5) and (3.4), respectively.

Our task here is to seek the set of the Pareto-efficient solutions of the bi-objective optimization problem (3.8). Since a BOT contract is essentially an outcome of negotiation between the government and the private firm and can be characterized by a combination of (T, q, y), we are now ready to define a Pareto-efficient contract for the BOT problem (3.8) as follows.

Definition 3.1. Pareto-efficient BOT Contract *A BOT triple $(T^*, q^*, y^*) \in \Omega$ is called a Pareto-efficient contract if there is no other feasible BOT triple $(T, q, y) \in \Omega$ such that $W(T, q, y) \geq W(T^*, q^*, y^*)$ and $P(T, q, y) \geq P(T^*, q^*, y^*)$ with at least one strict inequality.*

The Pareto-efficient BOT contract is an important and meaningful concept that represents the situation in which no party can be made better off without making the other one worse off.

For simplicity, we do not adopt an interest rate to discount future revenues to their equivalent present values. In fact, the use of a discounting rate does not alter our results since both social welfare and profit in this chapter are invariant with the calendar time. The result is shown in the following lemma for the same reason given by Guo and Yang (2010b).

Lemma 3.2.1. *Both the public and private sectors' decisions are free from the effect of discounting factor.*

Proof. Here we show that the use of a discounting factor on the stream of future revenue does not alter our results. Assume time τ is continuous and let r be an interest rate of reference used for discounting all monetary units to equivalent values at $\tau = 0$. The social welfare (3.5) and the profit (3.4) functions should be rewritten as

$$\begin{aligned} W(T, q, y) = & \int_0^T S(q, y)\, e^{-r\tau} \mathrm{d}\tau + \int_T^{\hat{T}} \tilde{S}(y)\, e^{-r\tau} \mathrm{d}\tau - I(y) \\ = & \frac{1-e^{-rT}}{r} S(q, y) + \left(\frac{1-e^{-r\hat{T}}}{r} - \frac{1-e^{-rT}}{r} \right) \tilde{S}(y) - I(y) \end{aligned} \tag{3.9}$$

and

$$P(T,q,y)=\int_0^T R(q,y)\,e^{-r\tau}\mathrm{d}\tau-I(y)=\frac{1-e^{-rT}}{r}R(q,y)-I(y)\,. \qquad (3.10)$$

Denote $L=\left(1-e^{-rT}\right)/r$ and $\hat{L}=\left(1-e^{-r\hat{T}}\right)/r$. Then, the bi-objective programming problem (3.8) would not change. □

It is clear, from the discussion, that the discount rate has not impact on our remaining analysis of this chapter.

3.3 Constant Returns to Scale in Road Construction

3.3.1 Properties of Pareto-efficient BOT contracts

In this section, we examine the properties of Pareto-efficient contracts in the BOT problem (3.8) under perfect information, namely, the demand and construction costs are common knowledge to both the public and the private sectors. We begin with the following proposition.

Proposition 3.3.1. *Under Assumption 3.1, if a triple* $(T^*,q^*,y^*)\in\Omega$ *is a Pareto-efficient BOT contract, then* $T^*=\hat{T}$.

Proof. Suppose that $(T^*,q^*,y^*)\in\Omega$ is a Pareto-efficient BOT contract and $T^*<\hat{T}$. Denote q_1 as the maximizer of $S(q,y^*)$ given by (3.6). From the first-order condition, we have

$$p=B(q_1)-\beta t(q_1,y^*)=\beta q_1\frac{\partial t(q_1,y^*)}{\partial q}>0,$$

which implies that, if $q^*=q_1$, then $T^*=\hat{T}$, because if $T^*<\hat{T}$ then (T^*,q^*,y^*) is strictly dominated by $\left(\hat{T},q^*,y^*\right)$, or the private firm can increase its profits by prolonging the concession period without changing the interest of the public sector.

We now prove $T^* = \hat{T}$ if $q^* \neq q_1$. To see this, we show that any (T^*, q^*, y^*) with $T^* < \hat{T}$ must not be Pareto-optimal, i.e., it must be dominated by another feasible BOT triple. First, from Assumption 3.1, we know that, for any given $y > 0$, both the unit-time social welfare $S(q, y)$ given by (3.6) and the following unit-time toll revenue

$$R(q, y) = q(B(q) - \beta t(q, y))$$

are strictly concave in q. Let T^{**} be $T^{**} = T^* + \Delta T$, $\Delta T > 0$, and denote any convex combination of q^* and q_1 as $q^{**} = \eta q^* + (1-\eta) q_1$, $\eta \in (0, 1)$. From the concavities of $S(q, y)$ and $R(q, y)$, we have

$$\begin{aligned} W(T^{**}, q^{**}, y^*) = &\ (T^* + \Delta T) S(\eta q^* + (1-\eta) q_1, y^*) \\ &+ \left(\hat{T} - (T^* + \Delta T)\right) S(q_1, y^*) - I(y^*) \\ > &\ W(T^*, q^*, y^*) \\ &+ ((1-\eta) T^* - \eta \Delta T)(S(q_1, y^*) - S(q^*, y^*)) \end{aligned} \tag{3.11}$$

and

$$\begin{aligned} P(T^{**}, q^{**}, y^*) = &\ (T^* + \Delta T) R(\eta q^* + (1-\eta) q_1, y^*) - I(y^*) \\ > &\ P(T^*, q^*, y^*) \\ &+ (\eta \Delta T - (1-\eta) T^*)(R(q^*, y^*) - R(q_1, y^*)) \\ &+ \Delta T R(q_1, y^*). \end{aligned} \tag{3.12}$$

Since $T^* < \hat{T}$, we can always choose η in $(0, 1)$ and a positive ΔT such that $\eta \Delta T - (1-\eta) T^* = 0$ or $\eta = T^*/(T^* + \Delta T)$. As a result, (T^{**}, q^{**}, y^*) dominates (T^*, q^*, y^*), which conflicts the Pareto-optimality of (T^*, q^*, y^*). Note that capacity is fixed at y^* throughout the proof even though optimal capacity generally changes when q and T change. This does not invalidate the proof because it is only necessary to show that a contract (T^{**}, q^{**}, y^*) exists that Pareto-dominates contract (T^*, q^*, y^*) when $T^* < \hat{T}$. The proof is completed. □

Proposition 3.3.1 states that any Pareto-efficient BOT contract requires a whole road life concession period. This "lifetime concession period" result seems to be realistic because several BOT contracts around the world have been awarded for 99 years, including Highway 407 in Toronto, the Chicago Skyway and the Pocahontas Parkway (Virginia Route 495) in Richmond, Virginia.

The economic logic of Proposition 3.3.1 becomes clear in the cases of monopoly and socially optimal solutions corresponding to the two polar points of the Pareto-efficient frontier of the bi-objective programming problem (3.8). First, maximizing the profit for a monopoly solution clearly requires as long a concession period as possible since profit in each operating period is positive (only the initial road construction costs are considered and maintenance and operating costs are ignored). Second, we note that the social welfare, given by Eq. (3.5), comes in the concession and the post-concession periods. If the concession period is less than the lifetime of the road, and the capacity and price of the road are fixed at their optimal (welfare-maximizing) values, it is a matter of indifference for the government how the transfer time is determined. However, the firm can reach a higher profit or Pareto improvement can be made when the transfer time is extended. In this case, the socially optimal and Pareto-efficient contracts must extend for the full lifetime of the road. Next, we consider a concession period less than the road life for any Pareto-efficient solution other than the monopoly and social optimum. Since the franchising firm realizes a positive contribution to its profit in each period, extending the concession period will therefore certainly increase profits at the prevailing price/capacity that differs from the socially optimal value. Such an extension therefore provides room for price and/or capacity changes that are more in the interest of social welfare without lowering private profits during the concession period. However, there may still be a loss associated with this change. Extending the concession period means that the government is no longer able to set a welfare-maximizing price during the extension of the concession period. In this case, it is not intuitively clear if extending the concession period will result in a net social welfare gain. From the proof of the proposition, we can see that the property of Pareto-efficient BOT contracts with full life-concession period, essentially, is guaranteed by the concavities of the unit-time social surplus and revenue functions in demand. This implies that a contract in which the firm operates at a moderate or compromise output throughout the road's life is Pareto superior to a two-stage contract with the firm earning a relatively high profit in the first stage and nothing in the second stage, while welfare suffers badly in the first stage because of the low output.

With Proposition 3.3.1, it is sufficient to set $T = \hat{T}$ in our subsequent analysis of the bi-objective problem (3.8). To obtain useful insights into Pareto-efficient BOT contracts, the following two common assumptions in the literature are introduced and used hereafter unless otherwise explicitly noted.

Assumption 3.2. *Link travel time function, t, is homogeneous of degree zero in the link flow, q, and the link capacity, y, i.e., $t(\alpha q, \alpha y) = t(q, y)$ for any $\alpha > 0$.*

Clearly, the widely used BPR (Bureau of Public Roads) link travel time functions satisfy this assumption. Note that, with this assumption, for any $y > 0$, $t(q, y) = t(q/y, 1)$. With a slight abuse of the notation, we denote $t(q, y)$ as $t(\gamma)$ for convenience, where γ, $\gamma = q/y$, is the volume-capacity (v/c) ratio. The v/c ratio, γ, is an important index to capture the service quality of the highway: the larger the value of γ, the longer the travel time and thus the worse the service quality and vice versa. Since the concession period extends to the whole road life in the static environment considered here, the v/c ratio under a given BOT contract is unchanged over time.

Assumption 3.3. *Constant return to scale in road construction, namely, $I(y) = ky$, where k denotes the constant cost per unit of capacity.*

Let $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$ and $\left(\hat{T}, \bar{q}, \bar{y}\right)$ be the socially optimal (SO) and monopoly optimal (MO) solutions, which maximize social welfare, $W\left(\hat{T}, q, y\right)$, and profit, $P\left(\hat{T}, q, y\right)$, respectively, or they meet the following first-order conditions, respectively:

$$\frac{\partial W}{\partial q} = B(\tilde{q}) - \beta t(\tilde{\gamma}) - \beta \tilde{\gamma} t'(\tilde{\gamma}) = 0, \tag{3.13}$$

$$\frac{\partial W}{\partial y} = \hat{T}\beta(\tilde{\gamma})^2 t'(\tilde{\gamma}) - k = 0, \tag{3.14}$$

and

$$\frac{\partial P}{\partial q} = B(\bar{q}) + \bar{q}B'(\bar{q}) - \beta t(\bar{\gamma}) - \beta\bar{\gamma}t'(\bar{\gamma}) = 0, \tag{3.15}$$

$$\frac{\partial P}{\partial y} = \hat{T}\beta(\bar{\gamma})^2 t'(\bar{\gamma}) - k = 0, \tag{3.16}$$

where $\tilde{\gamma}$ and $\bar{\gamma}$ denote the SO and MO v/c ratios, respectively, namely, $\tilde{\gamma} = \tilde{q}/\tilde{y}$ and $\bar{\gamma} = \bar{q}/\bar{y}$. By comparing conditions (3.14) and (3.16), it is readily seen that $\tilde{\gamma} = \bar{\gamma}$ from the fact that $\gamma^2 t'(\gamma)$ is strictly increasing in γ and thus both equations have the same unique solution of γ. More generally, we draw the following conclusion on the v/c ratio for any Pareto-efficient BOT contract.

Proposition 3.3.2. *Under Assumptions 3.1-3.3, the v/c ratio, γ^*, for any Pareto- efficient BOT contract $\left(\hat{T}, q^*, y^*\right)$ solves*

$$\hat{T}\beta(\gamma^*)^2 t'(\gamma^*) = k. \tag{3.17}$$

Thus, it is constant along the Pareto-optimal frontier and equals the socially optimal v/c ratio, $\tilde{\gamma}$.

Proof. Suppose that $\left(\hat{T}, q^*, y^*\right)$ is a Pareto-efficient contract. Then, (q^*, y^*) solves the following Lagrange problem

$$\begin{aligned} L(q, y, \eta) = &\ \hat{T}\left(\int_0^q B(w)\,\mathrm{d}w - \beta q t\left(\frac{q}{y}\right)\right) - I(y) \\ + &\ \eta\left(\left(\hat{T}q\left(B(q) - \beta t\left(\frac{q}{y}\right)\right) - I(y)\right) - P^*\right), \end{aligned} \tag{3.18}$$

where $P^* = P\left(\hat{T}, q^*, y^*\right)$ and $\eta \geq 0$ is the Lagrange multiplier. We have the following first-order conditions:

$$\begin{aligned} \frac{\partial L}{\partial q} = &\ (1+\eta)\hat{T}B(q^*) - (1+\eta)\hat{T}\left(\beta t\left(\frac{q^*}{y^*}\right) + \beta\frac{q^*}{y^*}\frac{\partial t(q^*/y^*)}{\partial q}\right) \\ &+ \eta\hat{T}q^*B'(q^*) = 0 \end{aligned} \tag{3.19}$$

and

$$\frac{\partial L}{\partial y} = (1+\eta)\left(\hat{T}\beta\left(\frac{q^*}{y^*}\right)^2 \frac{\partial t\left(q^*/y^*\right)}{\partial y} - k\right) = 0. \tag{3.20}$$

Denote $\gamma^* = q^*/y^*$ as the v/c ratio. Since $\eta \geq 0$, condition (3.20) can be reduced to

$$\hat{T}\beta(\gamma^*)^2 t'\left(\gamma^*\right) = k. \tag{3.21}$$

Since $t\left(\cdot\right)$ is strictly convex, Eq. (3.21) admits a unique solution, which implies that $\gamma^* = \tilde{\gamma}$. This completes the proof. □

The v/c ratio, γ, governs the travel time or delay of road users. Xiao et al. (2007b) compared the service quality levels of a congested highway offered by a profit-maximizing monopoly and by the public sector. If users value product quality equally (corresponding to homogeneous users with identical values of time), Xiao et al. (2007b) proved that the monopoly firm would offer the same service quality as the public sector, which is in line with the well-known economic findings (Spence, 1975). Proposition 3.3.2 further reveals that the service quality will coincide with that preferred by the public sector whenever the BOT contract is Pareto-efficient.

The structure of Pareto-efficient BOT contracts turns out to be very simple from Propositions 3.3.1 and 3.3.2: it includes a whole road-life concession period, $\hat{T}$, and a constant v/c ratio, γ. We define the contract curve in the demand-capacity space or the Pareto-optimal solution set of problem (3.8) (with concession period $T = \hat{T}$) as:

$$\Theta = \left\{ (q^*, y^*) \;\middle|\; \left(\hat{T}, q^*, y^*\right) \text{ is a Pareto-optimal solution} \right\}. \tag{3.22}$$

Any efficient bargaining between the public and the private sectors should result in an agreement on the contract curve. Any feasible BOT contract off the contract curve would be inefficient.

From the assumption that $B\left(q\right)$ is strictly decreasing and $\bar{\gamma} = \tilde{\gamma}$, we immediately ob-

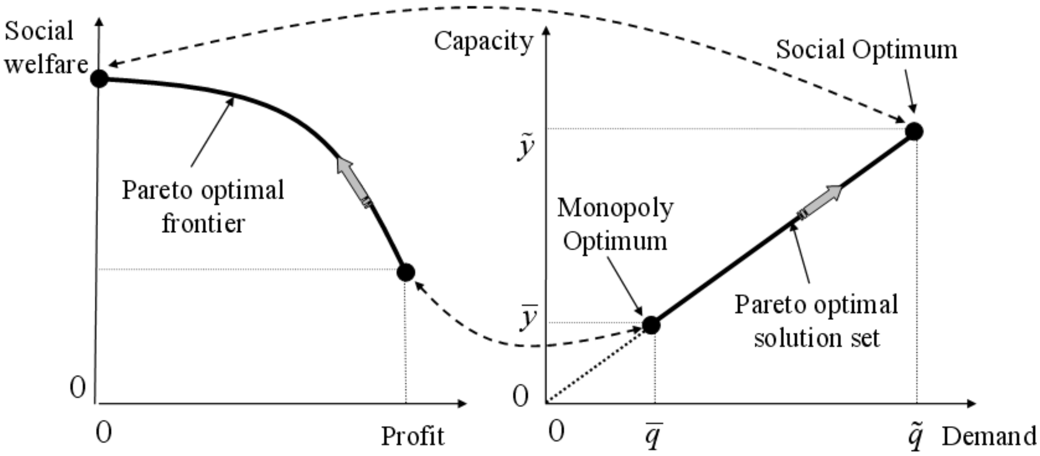


Figure 3.1: The contract curve and Pareto-optimal frontier

tain $\tilde{q} \geq \bar{q}$ by comparing Eqs. (3.13) and (3.15). Note that under Assumptions 3.1-3.1, both $W\left(\hat{T}, q, y\right)$ and $P\left(\hat{T}, q, y\right)$ given by the bi-objective programming problem (3.8) are jointly concave in (q, y), and, thus, the Pareto-optimal solution set is connected (Warburton, 1983). Therefore, the contract curve defined by (3.22) is the portion of the line connecting $(\bar{q}, \bar{y})$ and $(\tilde{q}, \tilde{y})$ with slope $\tilde{\gamma}$. Figure 3.1 shows the contract curve and the corresponding Pareto-optimal frontier in the decision space (q, y) and the objective space (P, W) as bold curves. The arrows indicate increasing social welfare. For any BOT contract, (T, q, y), the average social cost (ASC) (per user per unit time or per trip during the concession period) is defined as:

$$\text{ASC} = \frac{\beta T q t\left(q/y\right) + I\left(y\right)}{Tq} = \beta t\left(\frac{q}{y}\right) + \frac{I\left(y\right)}{Tq}. \tag{3.23}$$

The ASC does not change over calendar time in the static case considered; it equals the sum of two terms: the average travel time (in monetary units) and the construction cost allocation per trip. For any Pareto-efficient BOT contract, $\left(\hat{T}, q^{*}, y^{*}\right)$, by Assumption 3.1 and condition (3.17), the ASC given by (3.23) for $T = \hat{T}$ can be calculated as:

$$\text{ASC} = \beta t\left(\gamma^{*}\right) + \beta \gamma^{*} t'\left(\gamma^{*}\right). \tag{3.24}$$

From Proposition 3.3.2, we readily obtain the following result.

Proposition 3.3.3. *Under Assumptions 3.1-3.3, the average social cost defined by (3.23)* *in any Pareto-efficient BOT contract* $\left(\hat{T}, q^*, y^*\right)$ *is the same as the socially optimal ASC.*

Like the v/c ratio in Pareto-efficient BOT contracts, the ASC is constant along the Pareto-efficient frontier and equals the socially optimal ASC. Let C_0 denote the constant ASC under Pareto-efficient BOT contracts. From Eq. (3.24) and condition (3.13), we have $C_0 = \beta t(\tilde{\gamma}) + \beta\tilde{\gamma}t'(\tilde{\gamma}) = B(\tilde{q})$. Also, from Eq. (3.2), condition (3.13), Assumption 3.3 and the definition of ASC in (3.23), we readily know that in a socially optimal BOT contract $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$, the toll charge equals the allocation of the construction cost per trip:

$$\tilde{p} = B(\tilde{q}) - \beta t(\tilde{\gamma}) = \frac{I(\tilde{y})}{\hat{T}\tilde{q}}.$$

The toll revenue just covers the construction cost of the road, which is the classical self-financing result (Mohring and Harwitz, 1962). However, for any other Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$, other than the socially optimal contract with $q^* < \tilde{q}$,

$$B(q^*) > B(\tilde{q}) = C_0, \tag{3.25}$$

which means that the average generalized travel cost would exceed the average social cost. By subtracting the average travel time (in monetary units) on both sides of inequality (3.25), we have:

$$p^* > B(\tilde{q}) - \beta t(\gamma^*) = B(\tilde{q}) - \beta t(\tilde{\gamma}) = \frac{I(\tilde{y})}{\hat{T}\tilde{q}} = \frac{I(y^*)}{\hat{T}q^*}. \tag{3.26}$$

The last two equalities follow the result from Proposition 3.3.2. Equation (3.26) reveals that the Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$, with $q^* < \tilde{q}$ will be strictly profitable. Hence, we can view the corresponding toll charge, p^*, as the sum of two distinct parts: one for recovering the road construction cost, denoted as p_1, and the other for gaining profits on the investment, denoted as p_2 and called the markup charge. From condition (3.17), p_1 can be expressed as

$$p_1 = \frac{I(y^*)}{\hat{T}q^*} = \frac{\beta\hat{T}(\gamma^*)^2 t'(\gamma^*)\, y^*}{\hat{T}q^*} = \beta\gamma^* t'(\gamma^*) = \beta q^* \frac{\partial t(q^*, y^*)}{\partial q}. \tag{3.27}$$

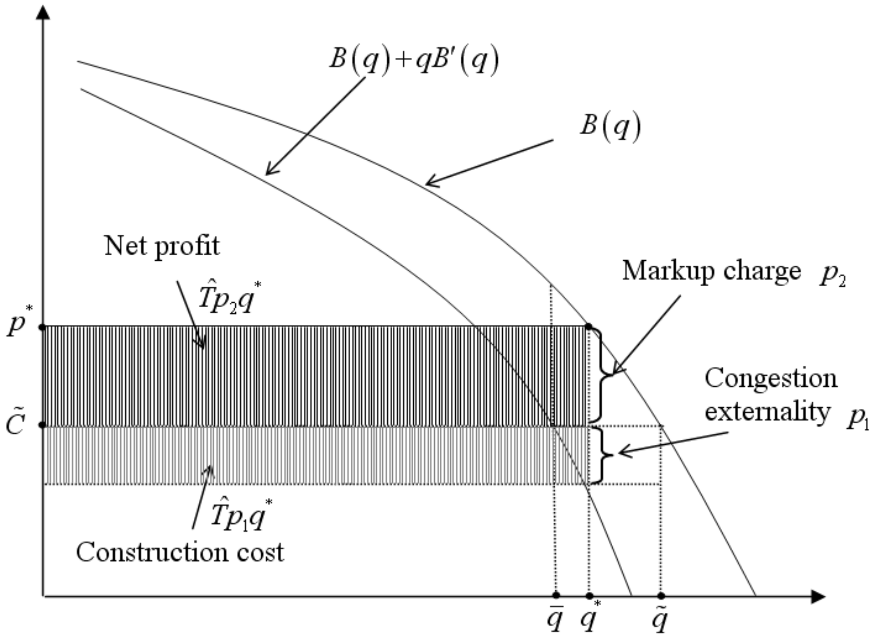


Figure 3.2: Geometric illustration of Pareto-efficient toll charges, demand and profits

Thus, for any Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$, p_1 is constantly equal to the socially optimal toll since $\gamma^* \equiv \tilde{\gamma}$ and it exactly equals the congestion externality. The markup charge imposed on each trip, p_2, can be calculated as

$$p_2 = p^* - p_1 = B(q^*) - \beta t(\gamma^*) - \beta\gamma^* t'(\gamma^*) = B(q^*) - C_0, \tag{3.28}$$

where the last equality follows from Proposition 3.3.3.

From the above analysis, we have clear knowledge of the profitability of a private road when a Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$, is awarded to a private firm: the private firm invests $\hat{T}q^*p_1$ and earns a profit equal to $\hat{T}q^*p_2$. In particular, with a socially optimal BOT contract, $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$, $p_2 = 0$ and the profit is nil. These observations are illustrated geometrically in Figure 3.2.

3.3.2 Efficiency of Pareto-efficient BOT contracts

In this section, we devote our analysis to the divergence between the socially optimal BOT contract and the other Pareto-efficient BOT contracts in terms of the realized social welfare. The degree of divergence is measured by the following ratio of social welfare:

$$\rho = \frac{W^*}{\tilde{W}} \leq 1.0, \tag{3.29}$$

where W^* is the total social welfare realized under a Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$, and $\tilde{W} = W\left(\hat{T}, \tilde{q}, \tilde{y}\right)$ is the maximized social welfare in a socially optimal solution.

From Proposition 3.3.3 and after simple manipulation, we have

$$W^* = W\left(q^*, y^*\right) = \hat{T}\left(\int_0^{q^*} \left(B\left(w\right) - C_0\right) \mathrm{d}w\right),$$

and

$$\tilde{W} = W\left(\tilde{q}, \tilde{y}\right) = \hat{T}\left(\int_0^{\tilde{q}} \left(B\left(w\right) - C_0\right) \mathrm{d}w\right).$$

Thus, the efficiency ratio, ρ, can be expressed as:

$$\rho = \frac{\int_0^{q^*} \left(B\left(w\right) - C_0\right) \mathrm{d}w}{\int_0^{\tilde{q}} \left(B\left(w\right) - C_0\right) \mathrm{d}w}. \tag{3.30}$$

It is noted that any Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$, of problem (3.8) must uniquely solve the following scalar programming problem (Geoffrion, 1967):

$$\max_{(T,q,y)\in\Omega, T=\hat{T}} \left(1 - \lambda\right) W\left(T, q, y\right) + \lambda P\left(T, q, y\right), \tag{3.31}$$

where $\lambda \in [0, 1]$ is a weighting parameter of social welfare and profit. From Propositions 3.3.2 and 3.3.3, we readily obtain

$$B\left(q^*\right) + \lambda q^* B'\left(q^*\right) = C_0, \tag{3.32}$$

which can be rearranged as

$$B\left(q^*\right) = \frac{C_0 E_{q^*}^{B^*}}{E_{q^*}^{B^*} + \lambda}, \tag{3.33}$$

where $E_{q^*}^{B^*}$ is the point price elasticity of demand at $\left(B\left(q^*\right), q^*\right)$, defined by

$$E_{q^*}^{B^*} = \frac{B\left(q^*\right)}{q^* B'\left(q^*\right)} \ (< 0)\,. \tag{3.34}$$

On the other hand, function $\int_0^q \left(B\left(w\right) - C_0\right) \mathrm{d}w$ is strictly concave in q (because $B\left(\cdot\right)$ is a decreasing function) and thus has a unique maximum at $q = \tilde{q}$. Note that $q^* \in [0, \tilde{q}]$, and thus q^*, can be expressed as a convex combination of 0 and $\tilde{q}$:

$$q^* = \left(1 - \frac{q^*}{\tilde{q}}\right) \cdot 0 + \frac{q^*}{\tilde{q}} \cdot \tilde{q}.$$

From the concavity of $\int_0^q \left(B\left(w\right) - C_0\right) \mathrm{d}w$, we readily have

$$\int_0^{q^*} \left(B\left(w\right) - C_0\right) \mathrm{d}w > \frac{q^*}{\tilde{q}} \int_0^{\tilde{q}} \left(B\left(w\right) - C_0\right) \mathrm{d}w,$$

which implies that

$$\rho = \frac{\int_0^{q^*} \left(B\left(w\right) - C_0\right) \mathrm{d}w}{\int_0^{\tilde{q}} \left(B\left(w\right) - C_0\right) \mathrm{d}w} > \frac{q^*}{\tilde{q}}. \tag{3.35}$$

We rewrite the last term of Eq. (3.35) as

$$\frac{q^*}{\tilde{q}} = \frac{1}{1 - \frac{q^* - \tilde{q}}{B(q^*) - B(\tilde{q})} \frac{B(q^*)}{q^*} \left(1 - \frac{B(\tilde{q})}{B(q^*)}\right)}. \tag{3.36}$$

Substituting Eq. (3.33) and $B\left(\tilde{q}\right) = C_0$ into Eq. (3.36) gives rise to

$$\rho > \frac{q^*}{\tilde{q}} = \frac{E_{q^*}^{B^*}}{E_{q^*}^{B^*} + \lambda Esh_{q^*}^{B^*}}, \tag{3.37}$$

where $Esh_{q^*}^{B^*}$ is the price elasticity of demand measured by a shrinkage ratio at $\left(B\left(q^*\right), q^*\right)$ and $\left(B\left(\tilde{q}\right), \tilde{q}\right)$ and defined by

$$Esh_{q^*}^{B^*} = \frac{q^* - \tilde{q}}{B\left(q^*\right) - B\left(\tilde{q}\right)} \frac{B\left(q^*\right)}{q^*} \ (< 0)\,. \tag{3.38}$$

The bound of the efficiency ratio, ρ, associated with a Pareto-efficient BOT contract can be stated in the following proposition.

Proposition 3.3.4. *Under Assumptions 3.1-3.3, for any Pareto-efficient BOT contract,* $\left(\hat{T}, q^*, y^*\right)$*, the efficiency ratio defined by (3.29) is bounded by*

$$\frac{1}{1+\lambda\xi} < \rho \leq 1.0, \tag{3.39}$$

where $\lambda \in [0,1]$ *satisfies condition (3.32) and* $\xi = Esh_{q^*}^{B^*} / E_{q^*}^{B^*} > 0$.

From this proposition, the bound of the efficiency ratio for a Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$, can be determined by the weighting parameter, λ, of social welfare and profit associated with $\left(\hat{T}, q^*, y^*\right)$ and parameter ξ, which depends on the convexity of the benefit function.

With a convex benefit function $B(q)$ or $B'' \geq 0$, we can further tighten the lower bound explicitly. In this case, we have the following two inequalities. A graphical illustration can be found in Xiao et al. (2007b):

$$\int_0^{q^*} \left(B(w) - C_0\right) \mathrm{d}w \geq \frac{1}{2} q^* \left(-q^* B'(q^*)\right) + \left(B(q^*) - C_0\right) q^* > 0$$

and

$$0 \leq \int_{q^*}^{\tilde{q}} \left(B(w) - C_0\right) \mathrm{d}w \leq \frac{1}{2} \left(\tilde{q} - q^*\right) \left(B(q^*) - C_0\right).$$

Based on the relationship between the above two inequalities,

$$a \geq \bar{a} > 0 \quad \& \quad 0 \leq b \leq \bar{b} \;\Rightarrow\; \frac{a}{a+b} \geq \frac{\bar{a}}{\bar{a}+\bar{b}}$$

and in view of definition (3.30), we immediately obtain

$$\rho \geq \frac{\frac{1}{2} q^* \left(-q^* B'(q^*)\right) + \left(B(q^*) - C_0\right) q^*}{\frac{1}{2} q^* \left(-q^* B'(q^*)\right) + \left(B(q^*) - C_0\right) q^* + \frac{1}{2} \left(\tilde{q} - q^*\right) \left(B(q^*) - C_0\right)}. \tag{3.40}$$

By rewriting (3.32) as

$$B(q^*) - C_0 = -\lambda q^* B'(q^*)$$

and substituting it into (3.40), we have

$$\frac{\frac{1}{2}+\lambda}{\frac{1}{2}+\lambda+\frac{\lambda^2}{2}\xi} \le \rho \le 1.0.$$

which is a bound that is tighter than that given by (3.39).

With a linear benefit function, $\xi = 1$ and (3.40) is strictly equal. Thus, the bound of the efficiency ratio can be simply reduced to

$$\frac{3}{4} \le \frac{\frac{1}{2}+\lambda}{\frac{1}{2}+\lambda+\frac{1}{2}\lambda^2} = \rho \le 1.0.$$

The lower bound 3/4 corresponds to the monopoly situation with $\lambda = 1$, which is consistent with that derived in Xiao et al. (2007b); the upper bound 1.0 corresponds to the socially optimal situation with $\lambda = 0$. When a BOT toll road is faced with a trade-off parameter of $\lambda = 0.5$, the efficiency ratio can be calculated as $\rho = 8/9$.

3.4 Effects of returns to scale in road construction

So far, we have examined the properties and the efficiency of Pareto-efficient contracts under constant returns to scale in road construction by assuming that $I(y) = ky$. To look into the effects of decreasing and increasing returns to scale in road construction, we now relax Assumption 3.3 to consider the following specific construction cost function

$$I(y) = ky^{\alpha}, \alpha > 0. \tag{3.41}$$

Road construction exhibits decreasing returns to scale when $\alpha > 1$ and increasing returns to scale when $0 < \alpha < 1$.

3.4.1 Pareto-efficient BOT contracts with non-constant returns to scale

We now examine how the returns to scale in road construction affect the properties of the Pareto-efficient BOT contract set. From Proposition 3.3.1, we know that $T = \hat{T}$ for any Pareto-efficient BOT contract. Let $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$ and $\left(\hat{T}, \bar{q}, \bar{y}\right)$ be the SO and MO solutions, which maximize social welfare, $W(T, q, y)$, and profit, $P(T, q, y)$, respectively. By subtracting $P\left(\hat{T}, q, y\right)$ from $W\left(\hat{T}, q, y\right)$, we readily obtain

$$W\left(\hat{T}, q, y\right) - P\left(\hat{T}, q, y\right) = \hat{T}\left(\int_0^q B(w)\,\mathrm{d}w - qB(q)\right). \tag{3.42}$$

Note that the term on the right-hand-side of Eq. (3.42) is strictly increasing in q, which implies that $\tilde{q} \geq \bar{q}$, because, if $\tilde{q} < \bar{q}$, then

$$\begin{aligned} W\left(\hat{T}, \bar{q}, \bar{y}\right) &= P\left(\hat{T}, \bar{q}, \bar{y}\right) + \hat{T}\left(\int_0^{\bar{q}} B(w)\,\mathrm{d}w - \bar{q}B(\bar{q})\right) \\ &> W\left(\hat{T}, \tilde{q}, \tilde{y}\right) \end{aligned}, \tag{3.43}$$

where the last inequality follows from the fact that $\left(\hat{T}, \bar{q}, \bar{y}\right)$ maximizes profit, $P(T, q, y)$. Equation (3.43) contradicts the assumption that $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$ maximizes the social welfare.

Suppose that $\left(\hat{T}, q^*, y^*\right)$ is a Pareto-efficient BOT contract. Like the constant returns to scale case, (q^*, y^*) solves the Lagrange problem (3.18) for a certain Lagrange multiplier, λ, and taking the derivative in y yields the following first-order condition:

$$\hat{T}\beta(\gamma^*)^2 t'(\gamma^*) = k\alpha(y^*)^{\alpha-1}, \tag{3.44}$$

where $\gamma^* = q^*/y^*$ is the v/c ratio under the Pareto-efficient BOT contract $\left(\hat{T}, q^*, y^*\right)$. Equation (3.44) under assumption (3.41) is the counterpart of the Pareto-efficiency condition (3.17) associated with $I(y) = ky$.

The following proposition reveals a few important properties of Pareto-efficient BOT contracts. Their departure from the case with constant returns to scale in road construction is self-evident.

Proposition 3.4.1. *Under Assumptions 3.1 and 3.2, for any two distinct Pareto-efficient BOT contracts,* $\left(\hat{T}, q^*, y^*\right)$ *and* $\left(\hat{T}, q^{**}, y^{**}\right)$*, with* $q^* < q^{**}$*, we must have*

(a) $y^* < y^{**}$, $p^* > p^{**}$*;*

(b) $\gamma^* < \gamma^{**}$ *and* $ASC^* < ASC^{**}$ *for* $\alpha > 1$*;* $\gamma^* > \gamma^{**}$ *and* $ASC^* > ASC^{**}$ *for* $0 < \alpha < 1$*, where ASC is the average social cost defined by Eq. (3.23);*

(c) $P^* > P^{**}$, $W^* < W^{**}$*;*

(d) $ROR^* > ROR^{**}$*, where ROR is the rate of return on investment defined as the ratio of the profit to the construction cost*

$$ROR = \frac{P}{I}\ (\times 100\%)\,. \tag{3.45}$$

Proof. By the definition of γ, the Pareto efficiency condition (3.44) can be rewritten as

$$\hat{T}\beta(\gamma^*)^{\alpha+1}t'(\gamma^*) = k\alpha(q^*)^{\alpha-1}.$$

Note that the function $\hat{T}\beta(\gamma)^{\alpha+1}t'(\gamma)$ is strictly increasing in γ since $\alpha > 0$ and $t(\gamma)$ is increasing and convex. Therefore, for any two distinct Pareto-efficient BOT contracts, $\left(\hat{T}, q^*, y^*\right)$ and $\left(\hat{T}, q^{**}, y^{**}\right)$, when $\alpha > 1$, $q^* < q^{**}$ implies that $\gamma^* < \gamma^{**}$, which, from condition (3.44), implies that $y^* < y^{**}$; while when $0 < \alpha < 1$, $q^* < q^{**}$ implies that $\gamma^* > \gamma^{**}$, which yields $y^* < y^{**}$. Thus, $y^* < y^{**}$ for any $\alpha > 0$ in (a) is obtained. In addition, the average social cost defined by (3.23) for any given $\left(\hat{T}, q^*, y^*\right)$ can be calculated as

$$\mathrm{ASC}^* = \beta t(\gamma^*) + \frac{\beta}{\alpha}\gamma^* t'(\gamma^*)\,.$$

Since ASC is a strictly increasing function of γ, we thus conclude that (b) is true.

From (3.42), if $P^* \leq P^{**}$, then $q^* < q^{**}$ must induce $W^* < W^{**}$, which contradicts the Pareto-optimality of $\left(\hat{T}, q^*, y^*\right)$. Thus $P^* > P^{**}$. Similarly, we also have $W^* < W^{**}$ whenever $q^* < q^{**}$. Thus, (c) is proved.

From Assumption 3.1, $B(q)$ is strictly decreasing in q and $t(\gamma)$ is strictly increasing in

γ. For $\alpha > 1$, from Eq. (3.2), $q^* < q^{**}$ and $\gamma^* < \gamma^{**}$ directly derive $p^* > p^{**}$. To prove the case with $0 < \alpha < 1$, using Eq. (3.44), we first rewrite the profit function for a given $\left(\hat{T}, q^*, y^*\right)$ as

$$P^* = \hat{T}q^* \left(p^* - \frac{\beta}{\alpha}\gamma^* t'(\gamma^*)\right). \tag{3.46}$$

From (c), we know that $q^* < q^{**}$ implies that $P^* > P^{**}$. Thus,

$$p^* - \frac{\beta}{\alpha}\gamma^* t'(\gamma^*) > p^{**} - \frac{\beta}{\alpha}\gamma^{**} t'(\gamma^{**}), \tag{3.47}$$

Function $\gamma t'(\gamma)$ is strictly increasing and $\gamma^* > \gamma^{**}$. We obtain $p^* > p^{**}$. Therefore, $p^* > p^{**}$ in (a) for any $\alpha > 0$ is proved.

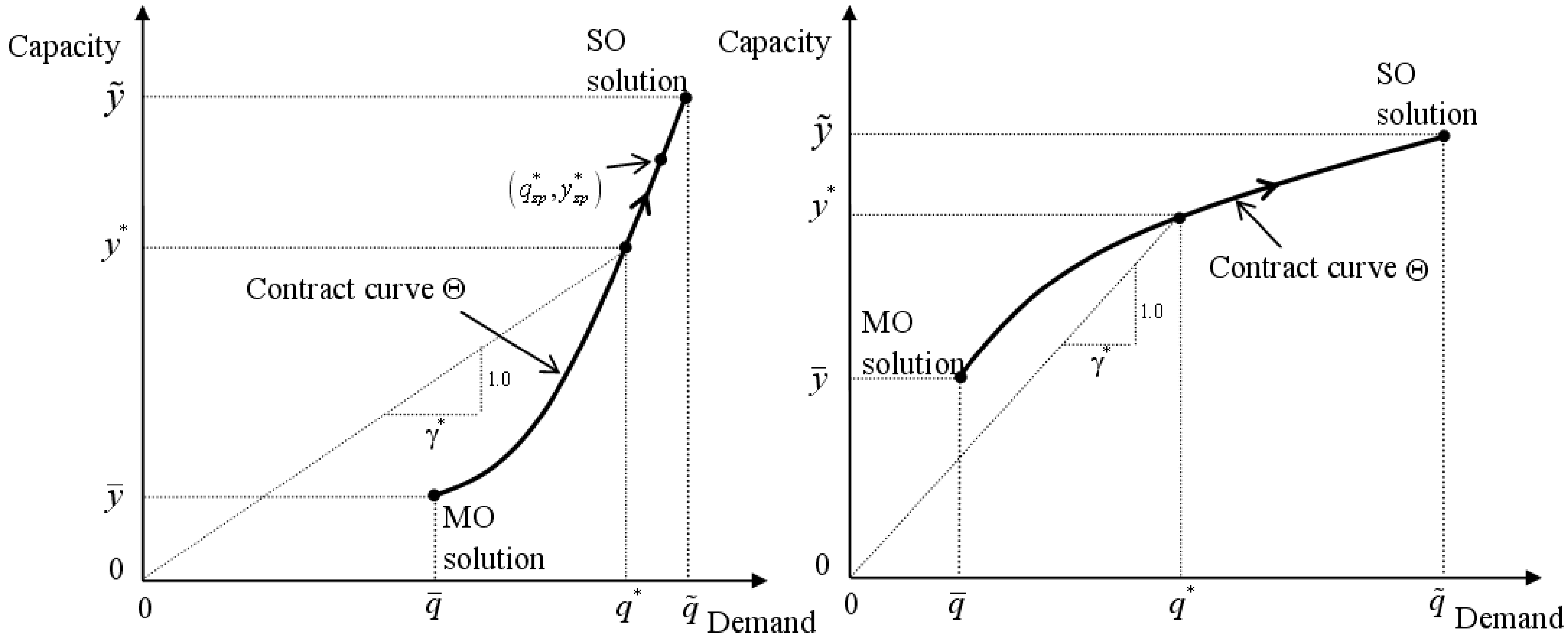
Finally, (a) and (c) together imply (d) since the construction cost, $I(y)$, strictly increases with y. The whole proof of Proposition 3.4.1 is completed. □

The properties of the contract curve, Θ, defined by (3.22) are explained geometrically in Figure 3.3. Along the direction shown in the figure, the travel demand, road capacity and social welfare increase, while the profit decreases. However, the v/c ratio increases when $\alpha > 1$ (decreasing returns to scale) and decreases when $0 < \alpha < 1$ (increasing returns to scale). As a result, the service quality decreases or increases along the Pareto-efficient frontier from monopoly to the social optimum, namely, the private firm tends to offer higher or lower service quality than the socially optimal level.

3.4.2 The profit properties at the social optimum

We go further to investigate the profit properties of the socially optimal BOT contract, $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$. In this case, the toll charge is exactly equal to the congestion externality,

$$\tilde{p} = B(\tilde{q}) - \beta t(\tilde{\gamma}) = \beta\tilde{\gamma} t'(\tilde{\gamma}).$$



(a) Increasing returns to scale

(b) Decreasing returns to scale

Figure 3.3: Contract curves with non-constant returns to scale in road construction

From Eq. (3.44), the profit of the private sector can be expressed as

$$P = \left(1 - \frac{1}{\alpha}\right)\hat{T}\tilde{q}\tilde{p}, \tag{3.48}$$

which reveals that, with as socially optimal BOT contract, $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$, the private firm would earn a positive profit with decreasing returns to scale in road construction ($\alpha > 1$); zero profit with constant returns to scale ($\alpha = 1$); and negative profit with increasing returns to scale ($0 < \alpha < 1$). The profit (positive or negative) is proportional to the total toll revenue with a proportional constant, $(1 - 1/\alpha)$. Equation (3.48) also implies that the cost recovery ratio (total toll revenue relative to the initial capacity cost) is exactly equal to α, which is consistent with the general rules derived by Mohring and Harwitz (1962).

3.4.3 The zero-profit constrained Pareto-efficient BOT contract

From the above analysis, the private sector would encounter a negative profit with a socially optimal BOT contract $\left(\hat{T}, \tilde{q}, \tilde{y}\right)$, when there are increasing returns to scale in road construction ($0 < \alpha < 1$). In this case it is of interest to look into the zero-profit-constrained Pareto-efficient BOT contract, $\left(\hat{T}, q^*_{zp}, y^*_{zp}\right)$, where "zp" stands for "zero-profit".

Before we seek such a zero-profit contract, $\left(\hat{T}, q^*_{zp}, y^*_{zp}\right)$, we provide the conditions for its existence. Clearly, this requires that the profit of the private sector with a monopoly optimal BOT contract, $\left(\hat{T}, \bar{q}, \bar{y}\right)$, be positive. In this case, by combining Eq. (3.2) and Eq. (3.44), the monopoly profit can be calculated as:

$$P\left(\hat{T}, \bar{q}, \bar{y}\right) = \hat{T}\bar{q}\left(B\left(\bar{q}\right) - \beta t\left(\bar{\gamma}\right) - \frac{\beta}{\alpha}\bar{\gamma}t'\left(\bar{\gamma}\right)\right). \tag{3.49}$$

Note that, for any road capacity, $y > 0$, when $q \to 0$, the average travel time approaches the free-flow travel time, $t\left(0\right)$, and the congestion externality approaches zero. Therefore,

$P\left(\hat{T}, \bar{q}, \bar{y}\right) \geq 0$ is guaranteed under the following condition:

$$B(0) - \beta t(0) > 0. \tag{3.50}$$

Intuitively, condition (3.50) is not practically restrictive, because we can reasonably expect that there is a positive potential traffic demand with a free-flow travel time. Otherwise, it is meaningless to build a new highway.

When condition (3.50) is met, there exists a Pareto-efficient contractwith zero profit for $0 < \alpha < 1$, $\left(\hat{T}, q_{zp}^*, y_{zp}^*\right)$. In this case, the solution, $\left(q_{zp}^*, y_{zp}^*\right)$, or equivalently $\left(q_{zp}^*, \gamma_{zp}^*\right)$ can be determined by Eq. (3.44) and the following zero-profit condition:

$$P\left(q_{zp}^*, y_{zp}^*\right) = \hat{T} q_{zp}^* \left(B\left(q_{zp}^*\right) - \beta t\left(\gamma_{zp}^*\right)\right) - I\left(y_{zp}^*\right) = 0.$$

The corresponding toll charge is given by

$$p_{zp}^* = B\left(q_{zp}^*\right) - \beta t\left(\gamma_{zp}^*\right) = \frac{1}{\alpha} \beta \gamma_{zp}^* t'\left(\gamma_{zp}^*\right) > \beta \gamma_{zp}^* t'\left(\gamma_{zp}^*\right). \tag{3.51}$$

The charge is thus higher than the corresponding congestion externality for cost recovery. With increasing returns to scale in road construction, the zero-profit Pareto-efficient contract, $\left(\hat{T}, q_{zp}^*, y_{zp}^*\right)$, with projection $\left(q_{zp}^*, y_{zp}^*\right)$ in the (q, y) space shown in Figure 3.3b, is thus a critical point; and any Pareto-efficient contract $\left(\hat{T}, q^*, y^*\right)$ with $q^* < q_{zp}^*$ would result in a positive profit for the private firm. Otherwise, the profit becomes negative.

3.5 Governmental Regulations

So far, we have examined the basic properties of Pareto-efficient BOT contracts for a road project. In this section, we focus on the regulatory mechanism that induces the private firm to choose a predetermined Pareto-optimal solution voluntarily, and we identify the regulatory regime that establishes a situation in which the regulatory outcomes are efficient. In all regulations considered below, we assume that the government already predetermines a targeted Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$ or $\left(\hat{T}, p^*, y^*\right)$, where

$$p^* = B(q^*) - \beta t(q^*/y^*) = B(q^*) - \beta t(\gamma^*). \tag{3.52}$$

3.5.1 Rate-of-return regulation

We first investigate a rate-of-return (ROR) regulatory mechanism, under which the private firm is allowed to earn no more than a "fair" rate of return on its investment. The private firm is free to choose a combination of the BOT variables as long as its profits do not exceed this fair rate.

Let s denote the ROR on the firm's investment and s^*, $s^* \geq 0$, be the ROR determined by (3.45) for a given Pareto-efficient BOT contract, $\left(\hat{T}, q^*, y^*\right)$. Under the ROR regulation, the government restricts the ROR on the investment of the private firm as follows:

$$s = \frac{Tqp - I(y)}{I(y)} \leq s^*. \tag{3.53}$$

Under the above ROR regulation, the problem of the profit-maximizing private firm can be expressed as:

$$\max_{0 \leq T \leq \hat{T}, p \geq 0, y \geq 0} \quad Tpq - I(y) \tag{3.54}$$

subject to condition (3.53). Now we have the following proposition disclosing the behavior of the private sector under the ROR regulation (3.53).

Proposition 3.5.1. *Let $\left(\hat{T}, q^*, y^*\right)$ be a non-monopoly Pareto-efficient solution with a corresponding ROR s^*. Then, under the ROR regulation (3.53), the private sector would choose a non-Pareto-efficient BOT contract, $\left(\hat{T}, q, y\right)$, with $q < q^*$ and $y > y^*$.*

Proof. We first note that condition (3.53) is binding under the best response of the private firm since s^* is a realizable rate-of-return associated with the predetermined Pareto-efficient contract, $\left(\hat{T}, q^*, y^*\right)$. We thus have

$$Tqp - I\left(y\right) = s^* I\left(y\right). \tag{3.55}$$

Therefore, under the binding condition (3.53), the profit-maximizing problem (3.54) is equivalent to maximizing the investment, $I\left(y\right)$, or equivalently maximizing the capacity, y, since $I\left(y\right)$ is strictly increasing in y. Now we view y as a function of T and q as determined by (3.55) and p is a function of q given by (3.2). We take the derivatives of y with respect to T and q, yielding

$$\left(\left(s^*+1\right) I'\left(y\right) - \beta T \frac{q^2}{y^2} t'\left(\frac{q}{y}\right)\right) \frac{\partial y}{\partial T} = q\left(B\left(q\right) - \beta t\left(\frac{q}{y}\right)\right) \tag{3.56}$$

$$\begin{aligned} &\left(\left(s^*+1\right) I'\left(y\right) - \beta T \tfrac{q^2}{y^2} t'\left(\tfrac{q}{y}\right)\right) \tfrac{\partial y}{\partial q} = \\ &\qquad T\left(B\left(q\right) + qB'\left(q\right) - \beta t\left(\tfrac{q}{y}\right) - \beta \tfrac{q}{y} t'\left(\tfrac{q}{y}\right)\right). \end{aligned} \tag{3.57}$$

If the concession period, T, for maximizing the profit or equivalently for maximizing the capacity is an interior point or $T \in \left(0, \hat{T}\right)$, then $\partial y / \partial T = 0$. As a result, the right-hand side term of Eq. (3.56) must be zero or we must have $p\left(q, y\right) = 0$ from Eq. (3.2), which conflicts with the binding condition (3.53). This means that the private firm must choose the concession period to be $\hat{T}$ or zero under the ROR regulation (3.53). Clearly, the choice of no concession period is out of question. We thus conclude that the private sector will choose a whole road-life concession period, $T = \hat{T}$, under the ROR regulation (3.53).

If $\left(\hat{T}, q^*, y^*\right)$ happens to be a solution of problem (3.53), we have $\partial y / \partial q = 0$ and hence the term in the bracket on the right-hand side of Eq. (3.57) equals zero, which corresponds to the monopoly optimality conditions (3.15). In addition, $\left(\hat{T}, q^*, y^*\right)$ is a given Pareto-efficient solution and thus satisfies the Pareto-efficiency condition (3.44). These two observations allow us to conclude that $\left(\hat{T}, q^*, y^*\right)$ is the MO solution. This conclusion in turn implies that, if $\left(\hat{T}, q^*, y^*\right)$ is a non-MO Pareto-efficient solution, then it must not be the optimal solution of problem (3.54) or $\left(\hat{T}, q, y\right) \neq \left(\hat{T}, q^*, y^*\right)$ and $\partial y / \partial q \neq 0$.

If $\left(\hat{T}, q^*, y^*\right)$ is a non-MO Pareto-efficient solution, then we have $\partial y / \partial q < 0$ at $\left(\hat{T}, q^*, y^*\right)$

for the following reason. First, $\partial y/\partial q = 0$ is excluded from the above analysis. In addition, $\partial y/\partial q > 0$ at $\left(\hat{T}, q^*, y^*\right)$ implies that increasing the demand can increase the capacity under condition (3.55). As a result, the profit, $P\left(\hat{T}, q, y\right) = s^* I(y)$, will increase. Also, from Eq. (3.42), the social welfare must strictly increase from $W\left(\hat{T}, q^*, y^*\right)$ when both the demand and profit increase. These results contradict the fact that $\left(\hat{T}, q^*, y^*\right)$ is Pareto-optimal. Therefore, under the ROR $s \leq s^*$ associated with the given non-MO Pareto-efficient solution, $\left(\hat{T}, q^*, y^*\right)$, the private sector will choose a BOT contract, $\left(\hat{T}, q, y\right)$, with $q < q^*$ and $y > y^*$. Thus, Proposition (3.5.1) is proved. □

Proposition 3.5.1 states that the private sector's choice deviates from the Pareto-efficient outcome. Meanwhile, the resulting v/c ratio, $\gamma = q/y < \gamma^*$, or the service quality of the highway under the ROR regulation (3.53) will be higher than the Pareto-efficient level. Since the private sector makes more profit than at the Pareto-efficient level, we have

$$\hat{T}pq - I(y) > \hat{T}p^*q^* - I(y^*) . \tag{3.58}$$

Inequality (3.58) implies that $p > p^*$ from the strictly increasing assumption of $I(\cdot)$.

In summary, the ROR regulation is inefficient since the private firm would select a higher road capacity and a higher toll level than the targeted Pareto-efficient BOT solution, $\left(\hat{T}, q^*, y^*\right)$. It is worth noting that the ROR regulation creates an incentive for the private sector to choose an inefficiently high capacity. This over-investment is, in fact, an instance of the Averch-Johnson effect (Averch and Johnson, 1962).

3.5.2 Price-cap regulation

The price-cap regulation allows the private sector to set a price below or equal to a price ceiling set by the government. For a targeted Pareto-efficient BOT solution, $\left(\hat{T}, q^*, y^*\right)$, the toll ceiling, p^*, is given by (3.52). In this case one can easily see that the private firm will choose the concession period to be $\hat{T}$ whenever the solution, $\left(\hat{T}, q^*, y^*\right)$, is profitable and selects a toll equal to p^*. Therefore, under the price-cap regulation constraint, the private firm's problem is to maximize its profit as given below:

$$\max_{q\geq 0, y\geq 0} \; P = \hat{T}p^*q - I\left(y\right)$$

subject to $B\left(q\right) - \beta t\left(q/y\right) = p^*$. By viewing y as a function of q and taking the derivative of P in q at $q = q^*$, we have

$$\left.\frac{\mathrm{d}P}{\mathrm{d}q}\right|_{q=q^*} = B\left(q^*\right) + q^*B'\left(q^*\right) - \beta t\left(\frac{q^*}{y^*}\right) - \beta\frac{q^*}{y^*}t'\left(\frac{q^*}{y^*}\right).$$

If the targeted Pareto-efficient solution, $\left(\hat{T}, q^*, y^*\right)$, is the MO solution, then the private firm will choose $\left(\hat{T}, q^*, y^*\right)$ straightforwardly to maximize its profits.

If $\left(\hat{T}, q^*, y^*\right)$ is a non-MO Pareto-efficient solution, we must have $\mathrm{d}P/\mathrm{d}q < 0$ at $q = q^*$ for the following reason. First, note that $\mathrm{d}P/\mathrm{d}q = 0$ at $q = q^*$ is excluded since it is a non-MO solution. If, however, $\mathrm{d}P/\mathrm{d}q > 0$ at $q = q^*$, then increasing q from q^* will increase profits and social welfare from (3.42). Therefore, the private sector would select $\left(\hat{T}, q, y\right)$ with $q < q^*$ to earn more profits under the price cap regulation of $p \leq p^*$. In addition, from the binding price-cap constraint, it is clear to see that $\gamma > \gamma^*$, and, as a result, that $y < y^*$.

In summary, the price-cap regulation is also inefficient and cannot induce the private firm to choose a Pareto-efficient BOT solution unless the targeted Pareto-efficient contract is the MO solution. Generally, under the price-cap regulation, the private sector would offer lower road capacity and lower service quality than would the Pareto-efficient solution, $\left(\hat{T}, q^*, y^*\right)$.

3.5.3 Capacity regulation

For any Pareto-efficient BOT solution, $\left(\hat{T}, q^*, y^*\right)$, from Proposition 3.4.1, $\bar{y} \leq y^*$, which means that it is pointless to regulate $y \leq y^*$ because the private firm would surely choose the monopoly optimal $\left(\hat{T}, \bar{q}, \bar{y}\right)$. Thus, in what follows, we examine the behavior of the private firm under the capacity regulation of $y \geq y^* \ (> \bar{y})$ only.

To be practically sensible, we suppose that the government predetermines a profitable target solution, $\left(\hat{T}, q^*, y^*\right)$, in setting up the capacity regulation. In this case, the private sector will certainly choose the concession period to be $\hat{T}$. Clearly, at any capacity level, the best response of the private firm is to set a monopoly price; that is, for any $y > 0$, the private firm will choose demand q such that the first-order condition (3.15) is satisfied. Viewing q as a function of y given by (3.15) and taking the derivative of profit, P, $P = \hat{T}q\left(B\left(q\right) - \beta t\left(q/y\right)\right) - I\left(y\right)$, in y gives rise to

$$\frac{\mathrm{d}P}{\mathrm{d}y} = \frac{\partial P}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}y} + \frac{\partial P}{\partial y} = \frac{\partial P}{\partial y} = \hat{T}\beta\left(\frac{q}{y}\right)^2 t'\left(\frac{q}{y}\right) - I'\left(y\right), \tag{3.59}$$

where $\partial P/\partial q = 0$ from the first-order condition (3.15). If the private sector chooses $\left(\hat{T}, q, y\right)$ with $y > y^*$ to maximize its profits, then $\mathrm{d}P/\mathrm{d}y = 0$ in (3.59), or the Pareto-efficiency condition (3.44) is satisfied. With Eq. (3.15), we conclude that $\left(\hat{T}, q, y\right)$ is the MO Pareto-efficient contract with $y > y^*$, which conflicts with Part (a) of Proposition 3.4.1. Therefore, under the capacity regulation of $y \geq y^*$, the private sector must choose $y = y^*$.

Conditional on $y = y^*$, if the private firm can increase profits by lowering the toll, the demand will increase and social welfare will increase as well, contradicting the Pareto efficiency of $\left(\hat{T}, q^*, y^*\right)$. The private firm would not choose $p = p^*$ either, because, otherwise, $\left(\hat{T}, q^*, y^*\right)$ must be the MO Pareto-efficient contract for the same reason above. Therefore, the private sector will choose a profit-maximizing toll that is certainly higher than p^* conditional on $y = y^*$. Meanwhile, the service quality is made higher than the Pareto-efficient level.

3.5.4 Demand regulation

Under the demand regulation by the government for a targeted Pareto-efficient solution, $\left(\hat{T}, q^{*}, y^{*}\right)$, the private firm is allowed to make choices subject to the resulting realized traffic volume level, $q \geq q^{*}$.

Like before, we consider the practically meaningful case in which the solution $\left(\hat{T}, q^{*}, y^{*}\right)$ is profitable. The private firm will choose the concession period to be $\hat{T}$. From Eq. (3.42), we can see that there is no strictly profitable deviation from the Pareto-efficient solution, $\left(\hat{T}, q^{*}, y^{*}\right)$, for the private firm to choose $\left(\hat{T}, q, y\right)$ with $q \geq q^{*}$, because the social welfare must be improved under constraint $q \geq q^{*}$ whenever the profit is increased, which contradicts the Pareto-optimality of $\left(\hat{T}, q^{*}, y^{*}\right)$.

We further confirm that the Pareto-efficient BOT solution, $\left(\hat{T}, q^{*}, y^{*}\right)$, is uni- que and maximizes private profits under regulation $q \geq q^{*}$. Since there is no profitable deviation, any choice of $\left(\hat{T}, q, y\right)$ by the private sector must realize the same profit and social welfare as $\left(\hat{T}, q^{*}, y^{*}\right)$, simultaneously, namely,

$$\hat{T}\left(\int_{0}^{q} B(w)\,\mathrm{d}w - \beta q t\left(\frac{q}{y}\right)\right) - I(y) = W\left(q^{*}, y^{*}\right) \tag{3.60}$$

and

$$\hat{T}\left(qB(q) - \beta q t\left(\frac{q}{y}\right)\right) - I(y) = P\left(q^{*}, y^{*}\right). \tag{3.61}$$

From Definition 3.1, we know that $\left(\hat{T}, q, y\right)$ is Pareto-efficient since no other BOT triple dominating $\left(\hat{T}, q, y\right)$. Subtracting Eq. (3.61) from Eq. (3.60) on each side gives rise to:

$$\int_{0}^{q} \left(B(w) - wB(q)\right)\mathrm{d}w = \int_{0}^{q^{*}} \left(B(w) - wB\left(q^{*}\right)\right)\mathrm{d}w. \tag{3.62}$$

Note that function $\int_{0}^{q} B(w)\,\mathrm{d}w - qB(q)$ is strictly increasing in q. Thus, Eq. (3.62) implies that $q = q^{*}$. By substituting $y = q^{*}/\gamma$ into condition (3.44), we readily obtain

$\gamma = \gamma^*$, and thus $y = y^*$. Therefore, under demand regulation $q \geq q^*$, the Pareto-efficient solution, $\left(\hat{T}, q^*, y^*\right)$, is the unique profit-maximizing BOT solution to be selected by the private firm.

Proposition 3.5.2. *Under Assumptions 3.1 and 3.2, the resulting choice by the private firm under the demand regulation, $q \geq q^*$, is $\left(\hat{T}, q^*, y^*\right)$ or Pareto-efficient.*

Proposition 3.5.2 shows that the demand regulation is an appealing regulatory choice. The government needs only to set a minimum level of demand and to let the private firm freely choose a preferable combination of road capacity, toll charge and concession period for profit maximization.

3.5.5 Markup charge regulation

Under a markup charge regulation, the firm is allowed to earn a certain amount of profit on each unit of output it sells in an economic setting. This is equivalent to the return-on-output (ROO) regulation (Train, 1991). In the earlier special case with constant returns to scale in road construction, the markup charge, p_2, is given by (3.28). In the general case, the markup charge can be defined as

$$p_2 = \frac{P(T, q, y)}{Tq},$$

which means that the markup charge is the amount of profit earned from each unit of realized demand (each trip) during the concession period. For a given Pareto-efficient BOT contract $\left(\hat{T}, q^*, y^*\right)$, the markup charge regulation refers to setting a ceiling of the markup charge during the concession period as follows:

$$p_2 \leq p_2^* = \frac{P\left(\hat{T}, q^*, y^*\right)}{\hat{T} q^*}. \tag{3.63}$$

From Proposition 3.5.2, we know that $q \leq q^*$ for the private sector to earn more profit. In view of $T \leq \hat{T}$, it is impossible for the private firm to choose (T, q, y) to earn more

Table 3.1: Summary of regulatory outcomes

Regulatory Regime	Choices of Private Firm	Pareto Efficiency
Price-Cap ($p \leq p^*$)	$y < y^*, p > p^*, T = \hat{T}$	No
Rate of return ($s \leq s^*$)	$y > y^*, p > p^*, T = \hat{T}$	No
Capacity ($y \geq y^*$)	$y = y^*, p > p^*, T = \hat{T}$	No
Demand ($q \geq q^*$)	$y = y^*, p > p^*, T = \hat{T}$	Yes
Markup charge ($p_2 \leq p_2{}^*$)	$y = y^*, p > p^*, T = \hat{T}$	Yes

profit under (3.63), namely,

$$P(T, q, y) = Tqp_2 \leq \hat{T}q^*p_2^* = P\left(\hat{T}, q^*, y^*\right),$$

for any (T, q, y) under the markup charge regulation condition (3.63). In fact, if the markup charge is fixed, the private sector must increase the total demand during the concession period to maximize the profit, which, equally, is to minimize the total user cost. We thus have the following proposition.

Proposition 3.5.3. *Under Assumptions 3.1 and 3.2, the resulting choice by the private firm under the markup charge regulation, $p_2 \leq p_2^*$, is $\left(\hat{T}, q^*, y^*\right)$ or Pareto-efficient.*

Proposition 3.5.3 reveals an alternative regulatory regime for the government to induce Pareto-optimal outcomes just by restricting the markup charge and letting the private firm choose any combination of the BOT variables.

3.5.6 Further discussion of regulatory issues

For a targeted Pareto-efficient solution, $\left(\hat{T}, q^*, y^*\right)$, the outcomes of the five alternative regulatory regimes examined so far are summarized in Table 3.1, where $p^* = B(q^*) - \beta t(q^*/y^*)$, $\text{ROR}^* = P\left(\hat{T}, q^*, y^*\right)/I(y^*)$ and $p_2^* = P\left(\hat{T}, q^*, y^*\right)/\hat{T}q^*$. The results of ROR and price-cap regulations are actually instances of the general results in economic settings (Train, 1991). Ubbels and Verhoef (2008) discussed the efficiencies of various auction mechanisms for a private road with an un-tolled alternative through numerical examples. Their simulation results revealed that the auction to minimize generalized travel costs and subsidies divided by total traffic demand would approach the social optimum (they called it the second-best outcome), and the auction to minimize generalized travel costs could result in the social optimum with the zero-profit constraint (the second-best outcome with the zero-profit constraint). For a single road with constant returns to scale in road construction, the social optimums with and without the zero-profit constraint are identical. Therefore, the auction to minimize generalized travel costs does approach the social optimum, which is equivalent to maximizing travel demands. Proposition 3.5.2 shows that, for a more general construction cost function, any Pareto optimum including the social optimum can be achieved via the demand regulation, and Proposition 3.5.3 is an alternative to the demand regulation.

3.6 Numerical Examples

In this section, three simple examples with different returns to scale in road construction are presented to demonstrate the results obtained so far. The following BPR link travel time function is used:

$$t\left(q,y\right)=t_0\left(1.0+0.15\left(\frac{q}{y}\right)^4\right),$$

where the free-flow travel time for the new highway is $t_0 = 0.5$ (h). Time is converted into money with a value-of- time of $\beta = 100$ (HK\$/h). The benefit function takes the following negative exponential form:

$$B\left(q\right)=-\frac{1}{b}\ln\left(\frac{q}{Q}\right),\ b>0,$$

where Q is the potential demand, $Q = 1.0\times10^4$ (veh/h), b is a scaling parameter reflecting the sensitivity of demand to the full trip price, $b = 0.04$. It is easy to check that the benefit function satisfies Assumption 3.1(a). The construction cost function for the highway is assumed to take the following power form of capacity:

$$I\left(y\right)=k_\alpha t_0 y^\alpha, \alpha>0,$$

Table 3.2: Returns to scale and corresponding parameter values

Returns to scale in road construction	α	k_α $\left(10^6\text{HK\$/h} \cdot (\text{veh/h})\right)$
Increasing returns to scale (IRS)	0.80	1.50
Constant returns to scale (CRS)	1.00	1.20
Decreasing returns to scale (DRS)	1.20	0.25

where parameter α captures the returns to scale in road construction: increasing returns to scale (IRS) with $0 < \alpha < 1$, constant returns to scale (CRS) with $\alpha = 1$ and decreasing returns to scale (DRS) with $\alpha > 1$, as mentioned before. The free-flow travel time, t_0, is proportional to the length of the road and k_α is the construction cost parameter corresponding to the returns to scale, α. The values of parameter α and k_α in Table 3.2 are used without necessarily representing a realistic setting. Finally, the road life, $\hat{T}$, is assumed to be 30 (years), or $\hat{T} = 1.314 \times 10^5$ (h) by assuming the number of operating hours per year to be $4380 = 12 \times 365$ (h). For the three cases with IRS, CRS and DRS listed in Table 3.2, we first view the maximized social welfare with a certain profit constraint, $P(T, q, y) \geq \tilde{P}$, as a parametric function of the concession period, T. Namely, we look at the result of the following maximization problem for any predetermined concession period:

$$W_{\max}(T) = \max_{q \geq 0, y \geq 0} \left\{ W(T, q, y) : P(T, q, y) \geq \tilde{P} \right\}.$$

Figure 3.4 shows how the maximized social welfare changes with the concession period, T, when $\tilde{P} = 0$ and $\tilde{P} = 0.8 \times 10^9$(HK\$), respectively. In the figure, it is clear that $W_{\max}$ is increasing in T and it reaches the maximum at $T = \hat{T}$ with a binding profit constraint of $P = \tilde{P}$. The Pareto-efficient BOT contract does require the concession period to be exactly the whole road life. Figures 3.5-3.7 report the Pareto-optimal solution sets or contract curves in the two-dimensional (γ, q) space, with bold curves connecting the MO and SO points for the three Cases of IRS, CRS and DRS, respectively, where the thick and thin contours represent social welfare and profit, respectively. The corresponding

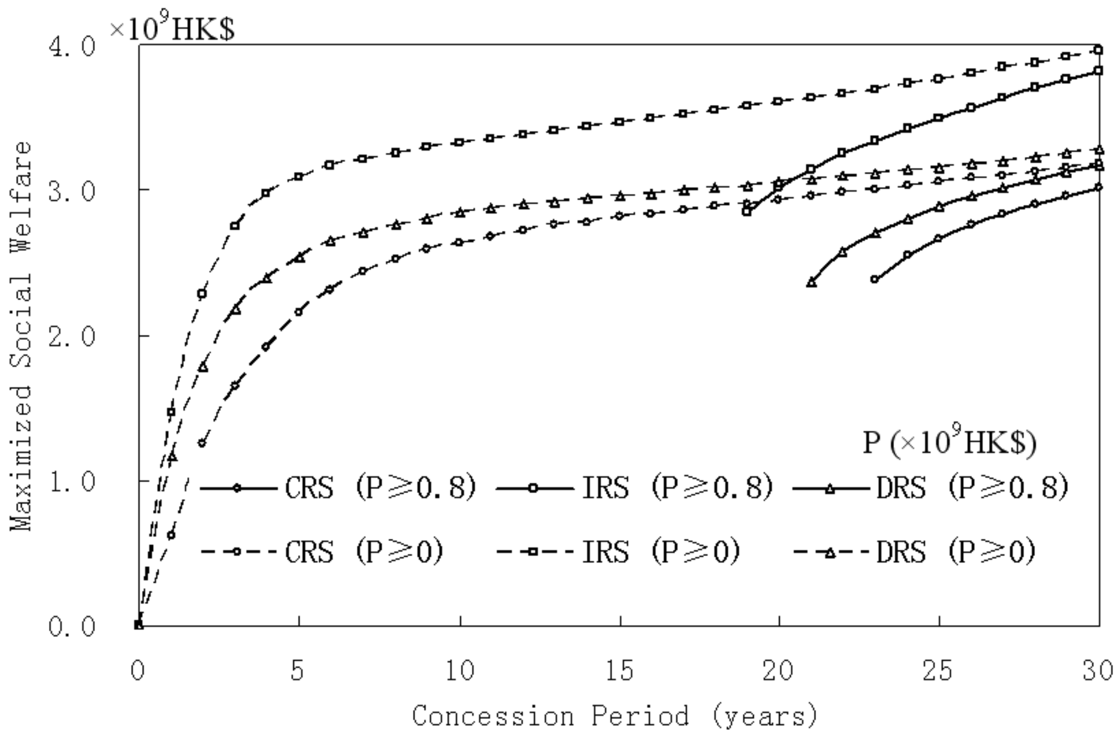


Figure 3.4: Maximized social welfare versus concession period with profit constraints

Table 3.3: Numerical results at SO and MO

	IRS		CRS		DRS	
	SO	MO	SO	MO	SO	MO
Demand (veh/h)	1230	445	970	357	955	368
Capacity (veh/h)	2448	851	1414	520	1375	549
Toll (HK$)	1.90	27.25	6.65	31.65	6.97	31.05
Social welfare (10^8HK$)	3.96	2.89	3.19	2.34	3.28	2.47
Profit (10^8HK$)	-0.07	1.43	0	1.18	0.15	1.26
Volume to capacity ratio	0.50	0.52	0.68	0.68	0.69	0.67
ASC (HK$)	52.9	53.4	58.5	58.5	60.5	59.1
ROR (%)	-0.8	57.6	0	108.3	20.3	519.2
Markup charge (HK$)	-0.45	24.41	0	25.00	1.19	26.03

representative numerical results for MO and SO are shown in Table 3.3. It is clear that, when moving from SO to MO, the service quality decreases (the v/c ratio increases) with increasing returns to scale in road construction; it remains unchanged with constant returns to scale; and it increases with decreasing returns to scale. For the IRS case, the zero-profit Pareto-efficient BOT contract, $\left(\hat{T}, q_{zp}^*, y_{zp}^*\right)$, can be readily obtained with $\hat{T} = 30$ (year), $q_{zp}^* = 1206$ (veh/h), $y_{zp}^* = 2399$ (veh/h), and the corresponding toll charge of $p_{zp}^* = 2.40$ (HK$). Figure 3.8 compares the outcomes of the private firm's choices under various regulatory regimes in the case with constant returns to scale in road construction (similar results can be obtained for the other two cases). For a predetermined Pareto-efficient solution, $\left(\hat{T}, q^*, y^*\right)$ = (30 years, 849 veh/h, 1230 veh/h), denoted as point A on the contract curve, we obtain the corresponding toll charge of $p^* = 9.96$ (HK\$), the rate of return of $s^* = 12.16\%$ and the markup charge of $\tau_2^* = 3.34$(HK\$). The five bold curves, L_1 L_5, indicate the binding constraints in the (γ, y) space, associated with the five regulatory regimes: (1) $p \leq p^* = 9.96$ (HK\$); (2) $s \leq s^* = 12.16$ (%); (3) $y \geq y^* = 1230$ (veh/h); (4) $q \geq q^* = 849$ (veh/h) and (5) $p_2 \leq p_2^* = 3.34$ (HK\$). The corresponding feasible domains are located above these curves in the (γ, q) space. As seen from the figure, each regulatory binding curve is tangent to a profit contour; the corresponding profit represents the maximum profit earned by the private firm under the given regulatory control. The choices made by the private firms under each regulatory regime are identified by the corresponding tangent point, denoted as $A_1 \sim A_5$ respectively. Specifically, under the price-cap regulation, the private firm chooses point A_1 (30 years, 723 veh/h, 769 veh/h), with a lower capacity and a lower service quality, producing a maximum profit of about4.90×10^8(HK\$). Under the capacity regulation, the private firm chooses point A_3 (30 years, 487 veh/h, 1230 veh/h), with a higher toll charge and a higher service quality, producing a maximum profit of about 8.90×10^8 (HK\$). A_4 and A_5, chosen by the private firm, coincide with Point A, or the private firm chooses the predetermined Pareto-efficient BOT solution, which is the unique choice to maximize its profits.

3.7 Conclusions

The concession period, capacity and toll charge are three primary variables for a BOT toll road project. They determine the social welfare for the whole society during the whole life of the road and the profit of the private firm during the concession period. We analyzed the properties of Pareto-efficient BOT contracts via a bi-objective programming approach and established several key results. First, any Pareto-efficient BOT contract requires that the

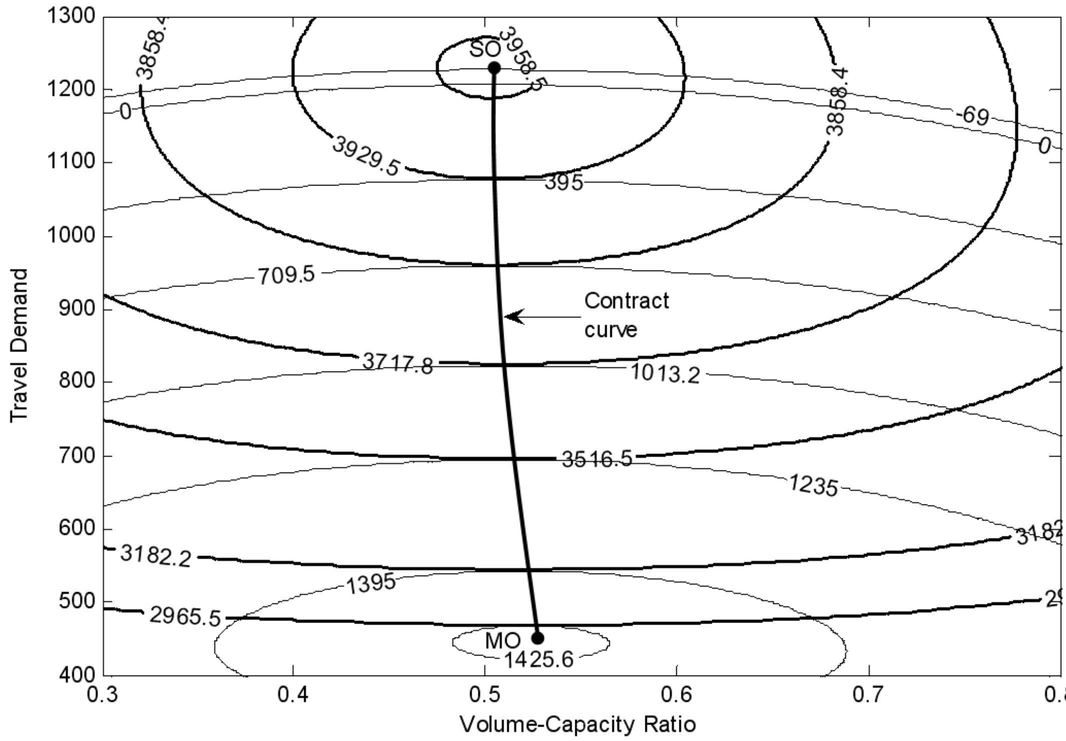


Figure 3.5: Contract curve with increasing returns to scale in road construction

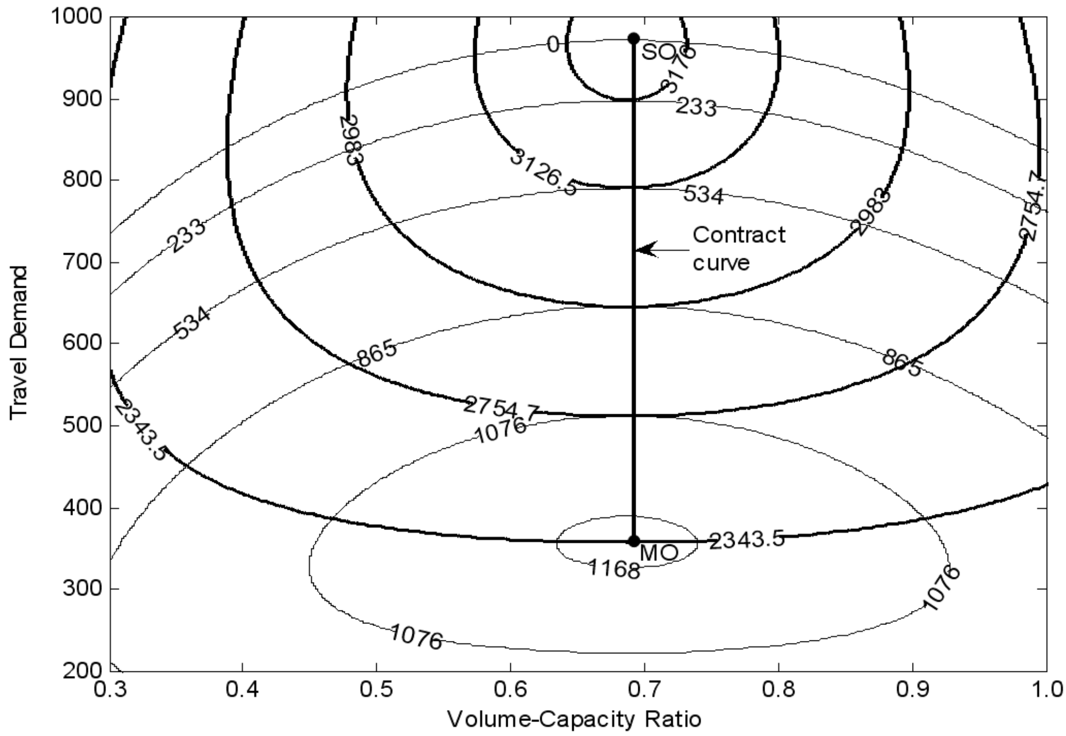


Figure 3.6: Contract curve with constant returns to scale in road construction

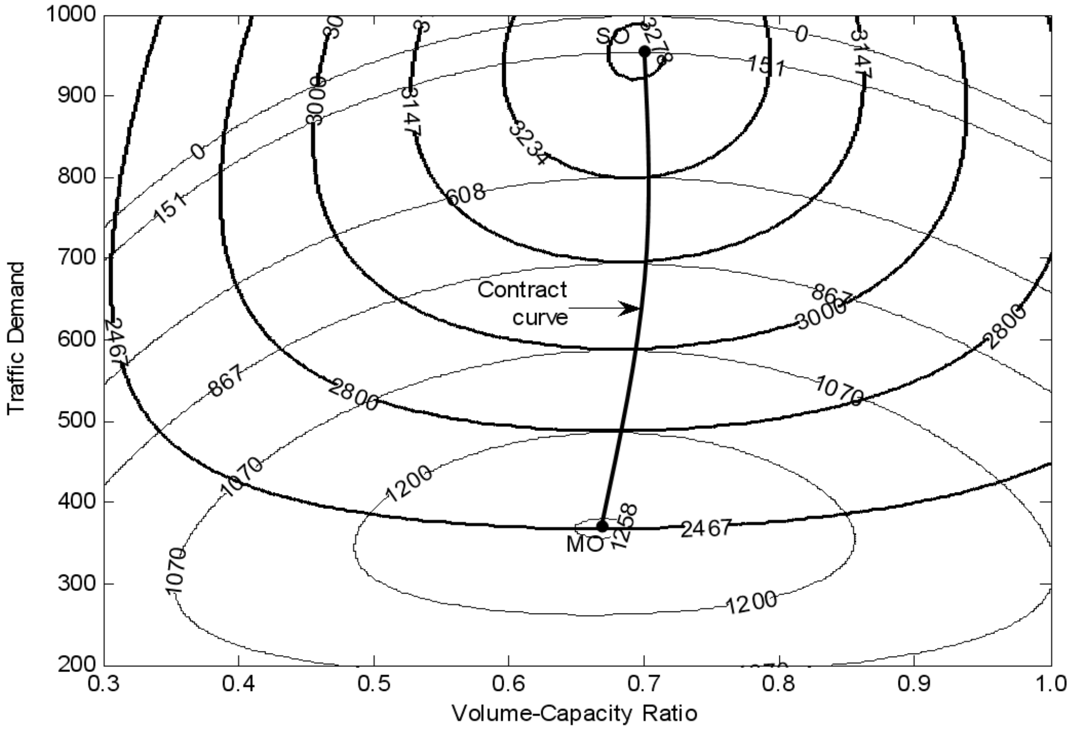


Figure 3.7: Contract curve with decreasing returns to scale in road construction

concession period should be the whole life of the road. Secondly, with constant returns to scale, the volume-capacity ratio and thus the service quality at any Pareto-efficient BOT contract coincide with the socially optimal levels; and the average social cost per trip is also constant along the Pareto-optimal frontier. We further established the efficiency bound of any Pareto-efficient BOT contract in terms of social welfare in comparison with the perfect social optimum. With a simple extension of the construction cost function, we proved that the private firm prefers to offer lower (higher) service quality than the socially optimal level when there are increasing (decreasing) returns to scale in road construction.

A variety of government regulatory regimes were investigated. Generally, both price-cap and rate-of-return regulations result in inefficient outcomes: the private firm tends to offer a lower road capacity and a lower service quality under the price-cap regulation, while it chooses a higher service quality, a higher capacity and a higher toll charge under the rate-of-return regulation than under the corresponding Pareto-efficient solution. The road capacity regulation is also inefficient. In contrast, we proved that both the demand and markup charge regulations lead to Pareto-optimal outcomes.

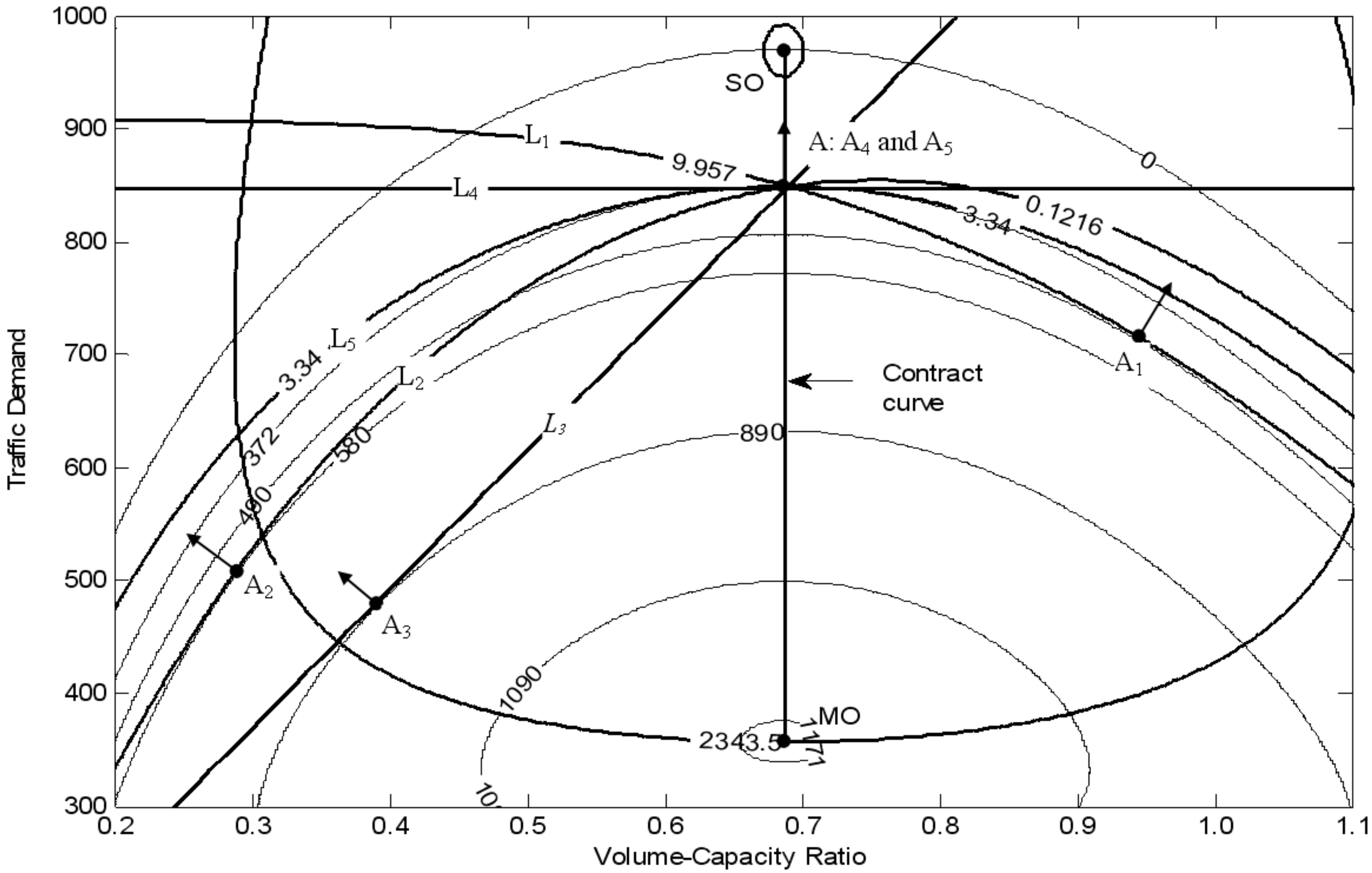


Figure 3.8: Regulations and the corresponding outcomes

4

PARETO-EFFICIENT BOT CONTRACTS WITH USERS HETEROGENEITY

This chapter re-examines the properties of BOT contracts with heterogeneous users whose value-of-time (VOT) has a continuous distribution. Two concepts of failure rate and mean residual VOT functions are introduced to develop a tractable analytical framework for the bi-objective programming formulation of the BOT problem. For the VOT distributions satisfying certain assumptions, we prove that the optimal concession period is free from the effects of user heterogeneity and should be set to be the road life as the case with homogeneous users. However, the service quality measured in terms of the volume-to-capacity ratio is dependent on the curvature of the mean residual VOT function. We prove that: if the VOT distribution has a convex (affine, concave) mean residual VOT function, then the service quality is increasing (identical, decreasing) along the Pareto-optimal frontier from the social optimal solution to the monopoly optimal solution. The effects of user VOT heterogeneity on the outcomes of various regulatory regimes are also investigated. Like the case with homogeneous users, the price-cap and rate-of-return regulations with heterogeneous users do not give rise to a Pareto-optimal outcome either. Nonetheless, we find that, the regulatory regimes of demand and markup regulations, which are efficient to achieve any pre-determined Pareto-efficient BOT contract in the case with homogeneous users, fail to result in any efficient outcome. Both two regimes result in a lower investment and lower service quality. It is also found that the markup regulation is equivalent to a demand regulation with a higher demand level associated with the pre-determined Pareto-efficient BOT contract.

4.1 Introduction

The previous chapter investigated the Pareto efficiency of the build-operate-transfer (BOT) contracts for a toll road project considering the social welfare and the profit as two objectives. The road users are assumed to experience the same disutility from travel delay, or, they have an identical value-of-time (VOT) to convert the time cost to monetary cost. The concept of VOT plays a pivotal role in analyzing travelers' route choices, particularly in simulating travelers' responses to toll roads, as it describes how users make tradeoffs between money and time in response to toll charges. Generally, travelers value their travel time differently, depending on their income levels and travel purposes. Many practical highway projects, such as the State Route 91 in California and Interstate 394 in Minnesota, are adopting the tolling scheme taking account of the users' VOT heterogeneity. In these two projects, parallel free lanes are simultaneously provided by the highway agency to the road users, and the users can enjoy the choice according to their values of travel time savings.

This chapter is intended to develop a tractable analytical framework to study the bi-objective BOT problem proposed in Chapter 3 in the presence of user VOT heterogeneity. By introducing two important concepts, the failure rate and mean residual VOT functions, this chapter contributes to answering the following questions to a certain extent: how does the user heterogeneity affect the optimal concession term, the optimal toll and capacity levels? How does the highway service quality vary along the Pareto-optimal frontier in the presence of user VOT heterogeneity, comparing to the case with homogeneous users? More important, whether are the regulatory regimes, proved to be efficient to achieve any predetermined Pareto-efficient BOT contract in the case with homogeneous users, still efficient in the presence of the user heterogeneity?

The next section, we revise the bi-objective programming proposed in the previous chapter to model a BOT toll road scheme with heterogenous users whose VOT continuously distributed. Section 4.3 investigates properties of the Pareto-efficient contracts. A significant different will be discussed in this section to compare with the results derived in the previous chapter. Section 4.4 examines the effects of user heterogeneity on the government regulations. Finally, Conclusions are given in Section 4.5.

4.2 The BOT Problem with Heterogenous Users

Suppose that a new BOT toll road is to be built in parallel to an existing free alternative road (or the set of roads). The alternative road (or the set of roads) is assumed to have a sufficiently large capacity without traffic congestion and its travel time is constantly equal to a fixed travel time t_0. The BOT toll road is assumed to have lower free flow travel time than t_0 (otherwise it does not attract any user). Let q be the travel demand of the BOT toll road and $t(q, y)$ be its corresponding travel time of the BOT toll road. The travel time function $t(q, y)$ is assumed to be a continuously differentiable function of (q, y) for $q \geq 0$ and $y \geq 0$; for any $q > 0$, $t(q, y)$ decreases with y; for any $y > 0$, $t(q, y)$ is a strictly convex and increasing function of q. Suppose that the value-of-time (VOT) of the road users follows a continuous and differentiable cumulative distribution $F(\beta)$ among the total fixed population Q with density function $f(\beta)$and a support $\Theta = (\beta_0, \beta^0)$ where $0 \leq \beta_0 \leq \beta^0 \leq +\infty$. As discussed by Verhoef and Small (2004), the party determining the road capacity y and toll charge level p induces two types of equilibrium: pooling equilibrium and separating equilibrium. The former means that all users choose either the free alternative or the toll road. In this case, either the new toll road needs not to be built or the existing road should be closed for all users. In the latter case of separating equilibrium, the users choose the two parallel routes according to their VOTs. From now on we shall focus our discussion on this case only. It is clear that the user will opt to pay the toll to use the toll road if the value of the travel time savings enjoyed on the toll road instead of the free alternative exceeds the toll cost. Suppose the value of VOT across population is sorted in decreasing order, let β give the VOT of the q-th user. Clearly, the demand of the toll road, q, is decreasing in β and has a relation with the cumulative function (Mayet and Hansen, 2000; Xiao and Yang, 2008)

$$q = Q\left(1 - F\left(\beta\right)\right). \tag{4.1}$$

Given the demand of the toll road, q, we know that the toll road is more attractive for the users with VOTs lager than β and not attractive for the users with VOTs less than β. At the separating equilibrium, the toll can be determined as

$$p = \beta\left(t_0 - t\left(q, y\right)\right). \tag{4.2}$$

Hereinafter, we will call β as the equilibrium VOT associated with the toll charge p and the equilibrium demand q that satisfy Eqs. (4.1) and (4.2).

Notably, for given toll and capacity levels, the travelers with a VOT higher than β will use the highway, and thus, a proportion of the total population, $F\left(\beta\right) = 1 - q/Q$, will suffer a loss to use the time-saving toll road. In this case $F\left(\beta\right)$ can be viewed as the failure probability of the decisions of toll and capacity, while, analogously, $\bar{F}\left(\beta\right) = 1 - F\left(\beta\right)$ can be viewed as the survival probability. To develop a tractable analytical framework to analyze the BOT problem, we introduce two concepts commonly used in reliability theory: the failure rate and mean residual VOT functions. Given q users using the toll road, the failure rate is the probability of failing to join the road for the next $q+1$-th user, mathematically, $h\left(\beta\right) \equiv f\left(\beta\right)/\bar{F}\left(\beta\right)$. Given the users with a VOT higher than β using the toll road, the mean residual VOT is the average excessive VOT of those users in comparison with the equilibrium VOT, β, mathematically,

$$m\left(\beta\right) = \mathrm{E}\left[x \middle| x \geq \beta\right] - \beta, \tag{4.3}$$

where E[ ] is the expectation operator. It is well known that, like a distribution function, the failure rate and mean residual VOT functions are two important concepts in reliability theory. Given the existence of the three functions, the distribution function $F\left(\beta\right)$, the failure rate function $h\left(\beta\right)$, and the mean residual VOT function $m\left(\beta\right)$, any one of the three functions can completely determine the other two (Lai and Xie, 2006). The failure rate and mean residual VOT functions have the following relation (Bryson and Siddiqui, 1969)

$$h(\beta)\, m(\beta) = m'(\beta) + 1. \tag{4.4}$$

Using the survival and mean residual VOT functions, the unit-time toll revenue and social surplus can be rewritten as

$$R(\beta, y) = Q\bar{F}(\beta)\, p = Q\beta\bar{F}(\beta)\,(t_0 - t(q, y)) \tag{4.5}$$

and

$$\begin{aligned} S(\beta, y) &= \int_{\beta}^{+\infty} Q\omega\,(t_0 - t(q, y))\, dF(\omega) \\ &= Q\bar{F}(\beta)\,(t_0 - t(q, y))\, \frac{\int_{\beta}^{+\infty} \omega dF(\omega)}{\bar{F}(\beta)} \\ &= Q\,(m(\beta) + \beta)\,\bar{F}(\beta)\,(t_0 - t(q, y))\,, \end{aligned} \tag{4.6}$$

where demand q is viewed as a function of β, given by Eq. (4.1), and toll p is viewed as a function of β and y, given by Eq. (4.2).

With the above preliminary notation and definitions, we are now ready to introduce the BOT problem studied in Chapter 3. The net profit of the private sector is the total toll revenue during the concession period minus the total up-front investment, namely,

$$P(T, \beta, y) = TR(\beta, y) - I(y)\,, \tag{4.7}$$

where the road construction cost function, $I(y)$, is continuously differentiable and increasing in capacity y for $y \geq 0$. The public sector aims to maximize the total social welfare during the whole road life, $\hat{T}$. Note that, after the ownership of the road is transferred from the private sector to the public sector, the toll can be adjusted in the interest of the public interest. Therefore, during the post-concession period, the public sector selects the toll charge (or equivalently, traffic volume or separating equilibrium VOT) to maximize the unit-time consumer surplus, $S(\beta, y)$, for the given road capacity, y. Namely,

$\tilde{S}(y) = \max_{\beta_1 \in \Theta} S(\beta_1, y)$. The total social welfare is the total social surplus during the concession period and the post-concession period minus the road investment, or,

$$W(T, \beta, y) = TS(\beta, y) + \left(\hat{T} - T\right)\tilde{S}(y) - I(y). \tag{4.8}$$

The BOT problem with heterogeneous users is defined as selecting the combination of the BOT variables: the concession period T, the equilibrium VOT β and the road capacity y, to maximize the total social welfare and the profit of the private sector s imultaneously. Thus we consider the following bi-objective programming problem:

$$\max_{(T,\beta,y)\in\Omega} \begin{pmatrix} W(T,\beta,y) \\ P(T,\beta,y) \end{pmatrix}, \tag{4.9}$$

where $\Omega = \left\{(T, \beta, y) : 0 \le T \le \hat{T}, \beta \in \Theta, y \ge 0\right\}$; the social welfare $W(T, \beta, y)$ and the profit $P(T, \beta, y)$ are defined by (4.8) and (4.7), respectively. With a simple modification, the Pareto-efficient BOT contract is defined as Definition 3.1. Furthermore, we assume that the travel time function $t(q, y)$ is a continuously differentiable function of (q, y) for $q \ge 0$ and $y \ge 0$; for any $q > 0$, $t(q, y)$ decreases with y; for any $y > 0$, $t(q, y)$ is a strictly convex and increasing function of q; the road construction cost function, $I(y)$, is a continuously increasing and differentiable function of y for $y \ge 0$. Namely, Assumption 3.1 (b) and (c) is true. Note that, BOT problem (4.9) is only a variant of the BOT problem studied in details in Chapter 3. At the center of the issue here is to examine the effects of the user VOT heterogeneity on the properties of the Pareto-efficient BOT contracts and various economic regulatory mechanisms.

4.3 Pareto-Efficient BOT Contracts with Heterogenous Users

In this section, we examine the effects of user heterogeneity on the properties of Pareto-efficient contracts in the BOT problem (4.9) using the continuously distributed VOT functions. The results can be used to contrasted with those described in the case of homogeneous users (Chapter 3).

One of the most used assumptions in revenue management, auction and mechanism design literature is to assume that the revenue function is quasi-concave in demand to ensure theoretically treatable (see Gallego and Ryzin, 1994; McAfee and McMillan, 1987; Ziya et al., 2004, for a comprehensive literature review). Obviously, function, $q\beta(q)$, determines the curvature of the "revenue function" in q. From Eq. (4.1), we know

$$\frac{\mathrm{d}\beta}{\mathrm{d}q} = -\frac{1}{Qf(\beta)},$$

and

$$\frac{\mathrm{d}^2\beta}{\mathrm{d}q^2} = -\frac{f'(\beta)}{Q^2(f(\beta))^3}.$$

By direct calibration, we obtain

$$\frac{\mathrm{d}^2(q\beta(q))}{\mathrm{d}q^2} = 2\frac{\mathrm{d}\beta}{\mathrm{d}q} + q\frac{\mathrm{d}^2\beta}{\mathrm{d}q^2} = -\frac{\bar{F}(\beta)^2\left(h'(\beta) + (h(\beta))^2\right)}{Q(f(\beta))^3}.$$

Therefore, the concavity of $q\beta(q)$ is equivalent to the following assumption.

Assumption 4.1. $h'(\beta) + h^2(\beta) \geq 0$.

Different from the homogeneous user case, the unit-time revenue function here, $R(\beta, y)$, under Assumption 4.1, may be not concave in q for any given y. This calls for a revision in applying and extending the proof of Proposition 3.3.1 with homogeneous users to the case here with heterogeneous users.

Lemma 4.3.1. *Under Assumption 4.1, if a triple $(T^*, q^*, y^*) \in \Omega$ is a Pareto-efficient BOT contract, then the concession period must equal the whole road life, i.e., $T^* = \hat{T}$.*

Proof. For convenience, we use q as the decision variable since q and β have one-to-one correspondence, given by Eq. (4.1). With a slight abuse of notation, let $S(q,\gamma)$, $W(T,q,y)$, $R(q,\gamma)$ and $P(T,q,y)$ denote $S(\beta,y)$, $W(T,\beta,y)$, $R(\beta,y)$ and $P(T,\beta,y)$, respectively. From Assumption 3.1 (b) and (c), we know that, for any given y, function $t_0 - t(q,y)$ is strictly decreasing and concave in q, while function $\int_0^q \beta(w)\,\mathrm{d}w$ is strictly increasing and concave in q. Thus, $S(q,y)$ is strictly concave in q for any given y. Suppose (T^*, q^*, y^*) is a Pareto-efficient BOT contract. Let $q_1 = \arg\max_{q\geq 0} S(q, y^*)$ and $\bar{q} = \arg\max_{q\geq 0} R(q, y^*)$. If $q^* = q_1$, then prolonging the concession period can improve the profit without changing the social welfare but increasing the profit of the private sector. Thus, $T^* = \hat{T}$. We look into the case $q^* \neq q_1$. Suppose $T^* < \hat{T}$ and $q^* \neq q_1$. We go to choose a feasible solution (T, q, y^*) such that both social welfare and profit can be improved, namely, (T, q, y^*) strictly dominates (T^*, q^*, y^*). For any given T and y, function $t_0 - t(q, y)$ is strictly decreasing and concave with respect to q, while function $\int_0^q \beta(w)\, dw$ is strictly increasing and concave in q. And thus, $W(T, q, y)$ is strictly concave with respect to q, for any given T and y. Denote $T = T^* + \Delta T \leq \hat{T}$ $(\Delta T > 0)$ and set $\alpha = T^*/(T^* + \Delta T)$. Notably, $\alpha \in (0, 1)$. Let q denote the convex combination of q^* and q_1 as $q = \alpha q^* + (1-\alpha) q_1$. From the concavity of $W(T, q, y)$ with respect to q, we obtain

$$W(T, q, y^*) > \alpha W(T, q^*, y^*) + (1-\alpha) W(T, q_1, y^*) = W(T^*, q^*, y^*).$$

In addition, since $t_0 - t(q, y^*)$ and $q\beta(q)$ are concave in q at domain (q^*, q_0), we have

$$\begin{aligned} P(T, q, y^*) > \;& P(T^*, q^*, y^*) \\ & + \Delta T \left\{ R(q^*, y^*) + \Delta_1 (q^* \beta(q^*)) + \Delta_2 (t_0 - t(q^*, y^*)) + \Delta_1 \Delta_2 \right\} \end{aligned}$$

where $\Delta_1 = (t_0 - t(q_1, y^*)) - (t_0 - t(q^*, y^*))$ and $\Delta_2 = q_1 \beta(q_1) - q^* \beta(q^*)$. Note that both Δ_1 and Δ_2 approach to zero as $\Delta T \to 0$. Therefore, there is a positive ΔT such that

$$|\Delta_i| \leq \xi \min\left\{ q^* \beta(q^*), t_0 - t(q^*, y^*) \right\}, \; i = 1, 2, \; \xi \in (0, 1), \tag{4.10}$$

and

$$R(q^*, y^*) + \Delta_1 (q^* \beta(q^*)) + \Delta_2 (t_0 - t(q^*, y^*)) + \Delta_1 \Delta_2 \geq R(q^*, y^*) \left(1 - 2\xi - \xi^2\right) \geq 0.$$

In fact, equation $\Psi(\xi) = 1 - 2\xi - \xi^2 = 0$ has a solution $\xi_0 \in (0, 1)$ since $\Psi(0) = 1 > 0$ and $\Psi(1) = -2 < 0$. In addition, since both Δ_1 and Δ_2 approach to zero as $\Delta T \to 0$, we can find, $\Delta T_1, \; \Delta T_2 \in \left(0, \hat{T} - T^*\right)$ such that inequalities (4.10) for $i = 1, 2$ are satisfied simultaneously for ξ with $0 < \xi < \xi_0$. Set $\Delta T = \min\{\Delta T_1, \Delta T_2\}$, we obtain inequalities

$W(T,q,y^*) > W(T^*,q^*,y^*)$ and $P(T,q,y^*) > P(T^*,q^*,y^*)$. Therefore, (T,q,y^*) strictly dominates (T^*,q^*,y^*), which conflicts the Pareto-optimality of (T^*,q^*,y^*). This completes the proof. □

It must be pointed out that Assumption 4.1 is one of the sufficient conditions for Lemma 4.3.1 to hold. It is still unclear whether the concession term is free from the effect of the user heterogeneity without Assumption 4.1. Fortunately, there are many types of commonly used distributions that meet this assumption, such as, the family of VOT distributions with increasing failure rate (IFR) property (the uniform, exponential, normal, truncated normal and Gumbel distributions, and Weibull and gamma distributions with shape parameters not less than one). A thorough study on the properties of the probability distributions can be found in the monograph by Lai and Xie (2006).

To obtain further insights, Assumptions 3.1 (b), (c), 3.2 and 3.3 are similarly introduced throughout this chapter unless otherwise explicitly noted. According to Lemma 4.3.1, let the concession period be the road lifetime $\hat{T}$, it is convenient to consider the following bi-objective maximization problem

$$\max_{(\beta,\gamma)\in\tilde{\Omega}} \begin{pmatrix} W(\beta,\gamma) \\ P(\beta,\gamma) \end{pmatrix} = \begin{pmatrix} Q\bar{F}(\beta)\left(\hat{T}(m(\beta)+\beta)(t_0 - t(\gamma)) - k/\gamma\right) \\ Q\bar{F}(\beta)\left(\hat{T}\beta(t_0 - t(\gamma)) - k/\gamma\right) \end{pmatrix}, \tag{4.11}$$

where $\tilde{\Omega} = \{(\beta,\gamma) : \beta \in \Theta, \gamma > 0\}$, and $\eta = k/\hat{T}$ is the unit-time unit-capacity construction cost of the toll road. With a slight abuse of notation, let $W(\beta,\gamma)$ and $P(\beta,\gamma)$ denote $W(T,\beta,y)$ and $P(T,\beta,y)$, respectively, in view of the definition $\gamma = q/y$. It is clear that any feasible solution of (4.11), (β,γ), corresponds to the VOT and capacity combination, $(\beta, q(\beta)/\gamma)$ and vice versa. For the bi-objective BOT problem (4.11), A BOT pair $(\beta^*,\gamma^*) \in \tilde{\Omega}$ is called a Pareto-optimal solution if and only if there is no other pair $(\beta,\gamma) \in \tilde{\Omega}$ such that $W(\beta,\gamma) \geq W(\beta^*,\gamma^*)$ and $P(\beta,\gamma) \geq P(\beta^*,\gamma^*)$ with at least one strict inequality.

Let $\left(\tilde{\beta},\tilde{\gamma}\right)$ and $(\bar{\beta},\bar{\gamma})$ be the social optimum (SO) and monopoly optimum (MO) solutions that maximize $W(\beta,\gamma)$ and $P(\beta,\gamma)$, respectively. The corresponding demand levels are denoted as $\tilde{q}$ and $\bar{q}$ at $\beta=\tilde{\beta}$ and $\beta=\bar{\beta}$, respectively. The first-order conditions for $\left(\tilde{\beta},\tilde{\gamma}\right)$ and $\left(\bar{\beta},\bar{\gamma}\right)$ can be expressed as:

$$\tilde{\beta}=\frac{k}{\hat{T}\tilde{\gamma}\left(t_0-t\left(\tilde{\gamma}\right)\right)}, \tag{4.12}$$

$$m\left(\tilde{\beta}\right)+\tilde{\beta}=\frac{k}{\hat{T}\tilde{\gamma}^2t'\left(\tilde{\gamma}\right)}, \tag{4.13}$$

and

$$\bar{\beta}\left(1-\frac{1}{\bar{\beta}h\left(\bar{\beta}\right)}\right)=\frac{k}{\hat{T}\bar{\gamma}\left(t_0-t\left(\bar{\gamma}\right)\right)}, \tag{4.14}$$

$$\bar{\beta}=\frac{k}{\hat{T}\bar{\gamma}^2t'\left(\bar{\gamma}\right)}. \tag{4.15}$$

Denote $\hat{\gamma}=\arg\max_{\gamma\geq 0}\left[\gamma\left(t_0-t\left(\gamma\right)\right)\right]$ and $\underline{\gamma}$ be the minimum v/c ratio solving the equation $\hat{T}\beta^0\gamma^2t'\left(\gamma\right)=k$, which is a specified positive number if $\beta^0<\infty$; otherwise $\underline{\gamma}=0$ if $\beta^0=\infty$. From first-order conditions (4.12)-(4.15) and the definitions of $\hat{\gamma}$ and $\underline{\gamma}$, it is intuitive to see that $\underline{\gamma}<\tilde{\gamma},\bar{\gamma}<\hat{\gamma}$.

Viewing β and γ as the functions of q and y, from Assumptions 3.1 (b), (c), 3.2, 3.3 and 4.1, we know that the welfare function W in problem (4.11) is jointly concave in q and y since the Hessian matrix of W in q and y is strictly negatively definite. Therefore, there exists a unique pair, (q,y) maximizing the welfare function, and thus, the SO solution, $\left(\tilde{\beta},\tilde{\gamma}\right)$, is a unique locally extreme point of W. In addition, under the same assumptions, $\beta+q\mathrm{d}\beta/\mathrm{d}q\geq 0$ at any extreme point of the profit function in problem (4.11) to guarantee the necessity of the extreme condition. Therefore, at any locally extreme point, the corresponding demand, q, must satisfy $\beta+q\mathrm{d}\beta/\mathrm{d}q\geq 0$. It is easy to check that the profit function in problem (4.11) is also jointly concave in q and y in the domain $\{(q,y):\beta+q\mathrm{d}\beta/\mathrm{d}q\geq 0, y\geq 0\}$. And thus, the monopoly optimal solution, $\left(\bar{\beta},\bar{\gamma}\right)$, is also a unique locally extreme point of P.

Using condition (4.12), the SO toll, $\tilde{p}$, can be rewritten as

$$\tilde{p} = \tilde{\beta}\left(t_0 - t\left(\tilde{\gamma}\right)\right) = \frac{k}{\hat{T}\tilde{\gamma}} = \frac{k\tilde{y}}{\hat{T}\tilde{q}}, \tag{4.16}$$

which means that the toll exactly equals the average allocation of construction cost to each user per unit time. In this case, the toll charge determined by Eq. (4.16) results in a zero profit for the private sector, or the toll revenue will exactly covers the construction cost, and thus the self-financing result holds. With condition (4.13), the SO toll can also be expressed as:

$$\tilde{p} = \frac{k}{\hat{T}\tilde{\gamma}} = \left(m\left(\tilde{\beta}\right) + \tilde{\beta}\right)\tilde{\gamma}t'\left(\tilde{\gamma}\right) = \mathrm{E}\left(x|\, x \geq \tilde{\beta}\right)\tilde{\gamma}t'\left(\tilde{\gamma}\right), \tag{4.17}$$

where $\mathrm{E}\left(x|\, x \geq \tilde{\beta}\right)$ is the conditional expectation of the VOT with a value larger than $\tilde{\beta}$ (average VOT of the users of the BOT toll road users). Equation (4.17) is an extension of the marginal cost pricing to the case of heterogeneous users, in which the average VOT value is used in ascertaining the monetary value of congestion externality. Equations (4.16) and (4.17) imply the well known first-best pricing and capacity rule: the socially optimal toll equals the congestion externality in the sense of the average VOT for all travelers using the toll road, and at optimal capacity level, the total toll revenue exactly covers the total construction cost.

Combining conditions (4.14) and (4.15), the MO toll can be expressed as:

$$\bar{p} = \frac{k}{\hat{T}\bar{\gamma}} + \frac{t_0 - t\left(\bar{\gamma}\right)}{h\left(\bar{\beta}\right)} = \bar{\beta}\bar{\gamma}t'\left(\bar{\gamma}\right) + \frac{t_0 - t\left(\bar{\gamma}\right)}{h\left(\bar{\beta}\right)}, \tag{4.18}$$

The MO toll consists of two components. The first component is the congestion charge equal to the marginal external cost of congestion evaluated at the average VOT. The second component is a strictly positive monopolistic markup. The portion of the revenue associated with the first term of congestion charge is exactly equal to the highway investment.

We now proceed to examine the property of the highway service quality in term of volume/capacity ratio for Pareto-efficient BOT contracts under Assumptions 3.1 (b), (c), 3.2, 3.3 and 4.1. Let $\left(\hat{T}, q^*, y^*\right)$ be any Pareto-optimal solution of problem (4.9), and denote the corresponding Pareto-optimal solution of problem (4.11) as (β^*, γ^*). Consider the following Lagrange problem:

$$L = W(\beta, \gamma) + \lambda \left(P(\beta, \gamma) - P(\beta^*, \gamma^*)\right), \tag{4.19}$$

where $\lambda \geq 0$ is a Lagrange multiplier. The following first-order conditions are satisfied:

$$\left.\frac{\partial L}{\partial q}\right|_{(q^*, \gamma^*)} = \hat{T}\beta^* \left(1 + \lambda \left(1 - \frac{1}{\beta^* h(\beta^*)}\right)\right)(t_0 - t(\gamma^*)) - \frac{(1+\lambda)k}{\gamma^*} = 0 \tag{4.20}$$

and

$$\left.\frac{\partial L}{\partial \gamma}\right|_{(q^*, \gamma^*)} = -\hat{T}\left(m(\beta^*) + \beta^* + \lambda\beta^*\right) t'(\gamma^*) + \frac{(1+\lambda)k}{(\gamma^*)^2} = 0. \tag{4.21}$$

Eliminating λ from Eqs. (4.20)-(4.21), we obtain the following relationship between the VOT and v/c ratio at any Pareto-optimal solution (β^*, γ^*):

$$\frac{k}{\hat{T}\gamma^* t'(\gamma^*)} \frac{1}{h(\beta^*) m(\beta^*)} - \frac{k}{\hat{T}(t_0 - t(\gamma^*))} = \gamma^* \beta^* \left(\frac{1}{h(\beta^*) m(\beta^*)} - 1 + \frac{1}{\beta^* h(\beta^*)}\right). \tag{4.22}$$

Viewing the Pareto-optimal v/c ratio γ^* as a function of the corresponding equilibrium VOT β^*, determined by implicit equation (4.22), $\gamma^* = \gamma^*(\beta^*)$. The continuity and differentiability of $\gamma^*(\beta^*)$ in β^* can be directly derived from the implicit function theorem. Taking derivatives on both sides of Eq. (4.22) in β^*, we have

$$\begin{aligned}&\left(\left(\left(\frac{\mathrm{d}}{\mathrm{d}\gamma}\left(\frac{k}{\hat{T}\gamma^* t'(\gamma^*)} - \frac{k}{\hat{T}(t_0 - t(\gamma^*))}\right) + \beta^*\right)\frac{1}{h(\beta^*)m(\beta^*)}\right) - \left(\frac{\beta^*}{h(\beta^*)m(\beta^*)} + \frac{1}{h(\beta^*)}\right)\right)\frac{\mathrm{d}\gamma}{\mathrm{d}\beta} \\ &= -\frac{\lambda\gamma^*}{1+\lambda}\frac{m''(\beta^*)}{h^2(\beta^*)m(\beta^*)}.\end{aligned} \tag{4.23}$$

From condition (4.21), we know that $\mathrm{d}\left(\eta/\gamma^* t'(\gamma^*)\right)/\mathrm{d}\gamma + \beta^* \leq 0$, and thus, the bracket term in the left-hand side of Eq. (4.23) is non-positive. The following lemma shows the relationship of the service quality and VOT along the Pareto-optimal frontier of the bi-objective maximization problem (4.11).

Lemma 4.3.2. *Along the Pareto-optimal frontier of problem (4.11), the relationship be-*

tween the v/c ratio and VOT is contingent on the the curvature of the mean residual VOT function, $m(\beta)$, in the following three distinct manners:

(i) If $m(\beta)$ is strictly convex in β, then, along the Pareto-optimal frontier, the v/c ratio, γ, is positively related to the VOT, β.

(ii) If $m(\beta)$ is affine in β, then, along the Pareto-optimal frontier, the v/c ratio, γ, is independent of the VOT, β.

(iii) If $m(\beta)$ is strictly concave in β, then, along the Pareto-optimal frontier, the v/c ratio, γ, is negatively related to the VOT, β.

The v/c ratio γ is a direct measure of the degree of congestion or the service quality provided by the toll road: the larger the v/c ratio, the lower the service quality and vise versa. In comparison with Proposition 3.3.2, we see that, the user heterogeneity has a significant effect on the highway service quality. For the case with homogeneous users or when the VOT distribution degenerates to a constant, $\beta(q) \equiv \beta_0$ in problem (4.11), it is clear that $\tilde{\gamma} = \bar{\gamma}$ and $\bar{q} = \tilde{q}$, which means that both the public sector and the private sector would offer the same service quality of the toll road. This is a known result in several previous studies (Wu et al., 2010; Xiao et al., 2007b). As indeed pointed out by Edelson (1971), even when capacity level is given, the profit-maximizing private firm may set a higher or lower toll level than that at social optimum. Thus, Lemma 4.3.2 is, in fact, nothing but an observation already made by Edelson (1971). To offer further insights into the effects of user heterogeneity on the Pareto-efficient BOT contracts, we make the following assumption on the mean residual VOT and failure rate functions.

Assumption 4.2. *$m(\beta) \geq m'(\beta)\beta$ and $\beta h(\beta)$ is increasing in β.*

The condition $m(\beta) \geq m'(\beta)\beta$ implies that the mean residual VOT from β, $\mathrm{E}(x| x \geq \beta) - \beta$, is not less than the linear estimation, $m'(\beta)\beta$, where $m'(\beta)$ is the increasing rate of the mean residual VOT at β. On the other hand, $\beta h(\beta)$ is the generalized failure rate function proposed by Lariviere (1999) that is required to be increasing with β in Assumption 4.2. Note that Assumption 4.2 is not restrictive. First, the IFR-distribution implies a decreasing mean residual VOT function (Lai and Xie, 2006) and an increasing generalized failure rate function, and thus, Assumption 4.2 includes the IFR distributions (uniform, exponential, normal, truncated normal, Erlang). Second, many widely used distributions, such as, Weibull, gamma, lognormal distributions and Pareto distributions, listed in Table 4.1, all meet Assumption 4.2.

For the gamma and Weibull distributions, $h'(\beta) \geq 0$ with $\alpha \geq 1$, Evidently, $m'(\beta) \leq 0$. It is clear condition $m(\beta) - m'(\beta)\beta \geq 0$ is naturally true. Furthermore, when $0 < \alpha < 1$, from the discussion below in Examples 2 and 3, we know that $m(\beta)$ is concave with $m(0) = \mathrm{E}(\beta) > 0$, and thus, it is true that $m(\beta) - m'(\beta)\beta \geq 0$ is true. For the lognormal distribution, from the direct calibration, we obtain

$$\begin{aligned}
& m(\beta) - m'(\beta)\beta \geq 0 \\
& \Leftrightarrow \sigma \geq \tfrac{\exp(-x^2/2)}{\int_x^{+\infty}\exp(-u^2/2)\mathrm{d}u} - \tfrac{\exp(-(x-\sigma)^2/2)}{\int_{x-\sigma}^{+\infty}\exp(-u^2/2)\mathrm{d}u} \\
& \Leftrightarrow \left(\tfrac{\exp(-x^2/2)}{\int_x^{+\infty}\exp(-u^2/2)\mathrm{d}u} - x\right) \leq \left(\tfrac{\exp(-(x-\sigma)^2/2)}{\int_{x-\sigma}^{+\infty}\exp(-u^2/2)\mathrm{d}u} - (x-\sigma)\right) \\
& \Leftarrow \int_x^{+\infty}\exp(-u^2/2)\mathrm{d}u \geq \tfrac{\exp(-x^2/2)}{\sqrt{1+x^2/4+x/2}} \\
& \Leftrightarrow \min_{x\geq 0}\left(\int_x^{+\infty}\exp(-u^2/2)\mathrm{d}u - \tfrac{\exp(-x^2/2)}{\sqrt{1+x^2/4+x/2}}\right) = 0.0147 > 0.
\end{aligned}$$

We now analyze the relationship between the MO and SO solutions based on the Assumption 4.2. From Eq. (4.11), we know that the social welfare can be expressed as

$$W(\beta,\gamma) = P(\beta,\gamma) + \hat{T}Q\bar{F}(\beta)\, m(\beta)\,(t_0 - t(\gamma)). \tag{4.24}$$

In view of relationship (4.4), we know that $\bar{F}(\beta)\, m(\beta)$ is strictly decreasing in β. Equation (4.24) implies that $\tilde{\gamma} \geq \bar{\gamma} \Rightarrow \tilde{\beta} < \bar{\beta}$ and $\tilde{\beta} \geq \bar{\beta} \Rightarrow \tilde{\gamma} < \bar{\gamma}$ since $W\left(\tilde{\beta},\tilde{\gamma}\right) > W\left(\bar{\beta},\bar{\gamma}\right)$ and $P\left(\bar{\beta},\bar{\gamma}\right) > P\left(\tilde{\beta},\tilde{\gamma}\right) = 0$. We now show that the latter case never occurs under Assumption 4.2. Dividing the both sides of Eqs. (4.12) and (4.13), respectively, we have

$$\frac{\tilde{\beta}}{m\left(\tilde{\beta}\right) + \tilde{\beta}} = \frac{\tilde{\gamma}t'(\tilde{\gamma})}{t_0 - t(\tilde{\gamma})}.$$

Similarly, dividing the both sides of Eqs. (4.13) and (4.14), respectively, and using the hazard rate function, we have

$$1-\frac{1}{\bar{\beta}h\left(\bar{\beta}\right)}=\frac{\bar{\gamma}t'\left(\bar{\gamma}\right)}{t_0-t\left(\bar{\gamma}\right)}.$$

Combining $\tilde{\gamma}\leq\hat{\gamma}$, $\bar{\gamma}\leq\hat{\gamma}$ and $\gamma\left(t_0-t\left(\gamma\right)\right)\geq\gamma^2t'\left(\gamma\right)$, we know that, if $\tilde{\gamma}<\bar{\gamma}$, then

$$\frac{\tilde{\beta}}{m\left(\tilde{\beta}\right)+\tilde{\beta}}<1-\frac{1}{\bar{\beta}h\left(\bar{\beta}\right)}$$

since $\gamma t'\left(\gamma\right)$ is strictly increasing and $t_0-t\left(\gamma\right)$ is strictly decreasing in γ at $\gamma\leq\hat{\gamma}$. From Assumption 4.2, $\beta h\left(\beta\right)$ is increasing in β, and thus, for $\tilde{\beta}>\bar{\beta}$, we obtain

$$0<\frac{\tilde{\beta}}{m\left(\tilde{\beta}\right)+\tilde{\beta}}<1-\frac{1}{\bar{\beta}h\left(\bar{\beta}\right)}\leq1-\frac{1}{\tilde{\beta}h\left(\tilde{\beta}\right)}. \tag{4.25}$$

Inequality (4.25) implies that $\tilde{\beta}m'\left(\tilde{\beta}\right)>m\left(\tilde{\beta}\right)$, which conflicts Assumption 4.2. Therefore, under Assumption 4.2, $\tilde{\beta}\leq\bar{\beta}$ whenever $\tilde{\gamma}\leq\bar{\gamma}$. Since demand q is a decreasing function of VOT β, and thus, $\tilde{q}\geq\bar{q}$.

Furthermore, if $\tilde{\gamma}\leq\bar{\gamma}$, then $\tilde{y}=\tilde{q}/\tilde{\gamma}\geq\bar{q}/\bar{\gamma}=\bar{y}$. We now prove that $\tilde{y}\geq\bar{y}$ even if $\tilde{\gamma}>\bar{\gamma}$. Using first order conditions (4.11)-(4.14), the social welfare at SO and MO can be respectively expressed as

$$W\left(\tilde{\beta},\tilde{\gamma}\right)=k\tilde{y}\left(\frac{t_0-t\left(\tilde{\gamma}\right)}{\tilde{\gamma}t'\left(\tilde{\gamma}\right)}-1\right)$$

and

$$W\left(\bar{\beta},\bar{\gamma}\right)=k\bar{y}\left(\frac{t_0-t\left(\bar{\gamma}\right)}{\bar{\gamma}t'\left(\bar{\gamma}\right)}\left(\frac{m\left(\bar{\beta}\right)}{\bar{\beta}}+1\right)-1\right).$$

Note that $\left(t_0-t\left(\gamma\right)\right)/\gamma t'\left(\gamma\right)$ is strictly decreasing in γ at $\gamma\leq\hat{\gamma}$ and $m\left(\bar{\beta}\right)/\bar{\beta}\geq0$. Therefore,

$$0<\frac{t_0-t\left(\tilde{\gamma}\right)}{\tilde{\gamma}t'\left(\tilde{\gamma}\right)}-1\leq\frac{t_0-t\left(\bar{\gamma}\right)}{\bar{\gamma}t'\left(\bar{\gamma}\right)}\left(\frac{m\left(\bar{\beta}\right)}{\bar{\beta}}+1\right)-1$$

which implies $\tilde{y} \geq \bar{y}$ since $\left(\tilde{\beta}, \tilde{\gamma}\right)$ maximizes social welfare.

We move to compare the toll levels at SO and MO . If $\tilde{\gamma} \geq \bar{\gamma}$, from the definition of toll, $p = \beta\left(t_0 - t\left(\gamma\right)\right)$, and $\tilde{\beta} \leq \bar{\beta}$, we know $\tilde{p} \leq \bar{p}$. We now prove that $\tilde{p} \leq \bar{p}$ for $\tilde{\gamma} \leq \bar{\gamma}$. From Assumption 4.2, $m\left(\beta\right) \geq m'\left(\beta\right)\beta$, we know

$$m\left(\beta\right) \geq m'\left(\beta\right)\beta \Leftrightarrow \frac{m\left(\beta\right)}{m\left(\beta\right)+\beta} \leq \frac{1}{\beta h\left(\beta\right)} \Leftrightarrow \frac{\beta}{m\left(\beta\right)+\beta} \geq 1 - \frac{1}{\beta h\left(\beta\right)}.$$

Combining the first-order conditions (4.11) and (4.13), we get

$$\frac{\tilde{p}}{\bar{p}} = \frac{\bar{\gamma}}{\tilde{\gamma}}\left(1 - \frac{1}{\bar{\beta}h\left(\bar{\beta}\right)}\right) \leq \frac{\bar{\gamma}}{\tilde{\gamma}}\frac{\bar{\beta}}{m\left(\bar{\beta}\right)+\bar{\beta}} \leq \frac{\bar{\gamma}}{\tilde{\gamma}}\frac{\bar{\beta}}{m\left(\tilde{\beta}\right)+\tilde{\beta}}.$$

where the last inequality follows the increasing monotonicity of $m\left(\beta\right)+\beta$ and $\tilde{\beta} \leq \bar{\beta}$. Again, using the first-order conditions (4.12) and (4.14), we obtain

$$\frac{\tilde{p}}{\bar{p}} \leq \frac{\bar{\gamma}}{\tilde{\gamma}}\frac{\left(\tilde{\gamma}\right)^2 t'\left(\tilde{\gamma}\right)}{\left(\bar{\gamma}\right)^2 t'\left(\bar{\gamma}\right)} = \frac{\tilde{\gamma} t'\left(\tilde{\gamma}\right)}{\bar{\gamma} t'\left(\bar{\gamma}\right)} \leq 1,$$

since $\tilde{\gamma} \leq \bar{\gamma}$.

The above discussion is summarized into the following proposition.

Proposition 4.3.1. *Under Assumptions 3.1(b), (c), 3.2, 3.3, 4.1 and 4.2, $\tilde{\beta} \leq \bar{\beta}$, and thus $\tilde{q} \geq \bar{q}$, $\tilde{y} \geq \bar{y}$, $\tilde{p} \leq \bar{p}$.*

Proposition 4.3.1 is a useful extension of the theoretical analysis by Xiao and Yang (2008), who proposed only some examples to illuminate the possible results of the SO and MO solutions. Proposition 4.3.1 points out that an analytical results exists when the continuous VOT distributions satisfy certain properties. Note that Assumptions 4.1 and 4.2 are commonly used in revenue management literature; neither of the two assumptions is more restrictive than the other (see Ziya et al., 2004, for more discussion and some counterexamples). If the concession length is pre-determined (typically, 25 to 30 years), then Lemma 4.3.1 and Assumption 4.1 become irrelevant, in this case Assumption 4.2 is sufficient to guarantee Lemma 4.3.2 and Proposition 4.3.1.

We know that the Pareto-efficient v/c ratio, γ^*, viewing as a function of Pareto-efficient VOT, β^*, given by Eq. (4.22), is a continuously differentiable function. The sign of $\mathrm{d}\gamma^*/\mathrm{d}\beta^*$ is completely determined by the curvature of the mean residual function $m(\beta)$, such as, $\mathrm{d}\gamma^*/\mathrm{d}\beta^* > 0$ when $m''(\beta) > 0$; $\mathrm{d}\gamma^*/\mathrm{d}\beta^* = 0$ when $m''(\beta) = 0$; and $\mathrm{d}\gamma^*/\mathrm{d}\beta^* < 0$ when $m''(\beta) < 0$. On the other hand, as we have discussed above that, from Assumptions 3.1(b), (c), 3.2, 3.3 and 4.1, both the welfare and profit function, W and P, in problem (4.11) are jointly concave in q and y since the Hessian matrices of W and P in q and y are strictly negatively definite. Therefore, any Pareto-efficient profit level is associated with a unique Pareto-efficient solution in (γ, β)-space because (β, γ) is uniquely determined by (q, y). The following proposition is the direct result of Lemma 4.3.2 and Proposition 4.3.1 whenever the Pareto-efficient solution set is connected.

Proposition 4.3.2. *Under Assumptions 3.1(b), (c), 3.2, 3.3, 4.1 and 4.2, along the Pareto-efficient frontier from monopoly optimum to social optimum, the volume-capacity ratio is decreasing (identical or increasing), and thus, the service quality measured in volume-capacity ratio is increasing (identical or decreasing) if the mean residual VOT function, $m(\beta)$, is convex (affine, concave).*

From Eq. (4.23), we know that, if the distribution $F()$ is nondegenerate (users are strictly heterogeneous), the v/c ratio γ along the Pareto-optimal frontier from MO to SO is decreasing, constant or increasing (the service quality is increasing, identical or decreasing) whenever the mean residual VOT function, $m()$, is convex, affine or concave. In particular, by comparing the two polar cases of MO and SO, the private sector offers a lower, identical or higher service quality than the public sector does when $m()$ is convex, affine or concave. The results are similar to those of Xu (2009) on the supply chain, in which the manufacturer sells his products to the customers of heterogeneous tastes directly or via a retailer. At the optimal decision of the manufacturer, the product quality turns out to be lower, identical or higher, if the curvature of the reciprocal of the hazard rate function is convex, affine or concave, due to the effect of the distribution channel (retailer).

In fact, in our case the government can be regarded as the manufacturer, providing the highway service to the road users directly (social optimum solution) or by contracting with a private firm (Pareto-optimal solution or contract). The major difference from the supply chain distributing channel lies in that the road service in our case is a public good and thus the objective function is an integral form that involves the mean residual VOT function. We now provide several examples to depict the three cases in Proposition 4.3.2.

Example 1. If the mean residual VOT function $m(\beta)$ is affine, such as, exponential and uniform distributions, namely, $m(\beta) = a\beta + b$, $(a > -1, b > 0)$, then the hazard rate function $h(\beta) = (m'(\beta) + 1)/m(\beta) = (a+1)/(a\beta + b)$ (Lai and Xie, 2006). Equation (4.22) is reduced to

$$b\gamma = \frac{(a+1)\eta}{\gamma t'(\gamma)} - \frac{\eta}{t_0 - t(\gamma)}. \tag{4.26}$$

Since the left-hand-side term is strictly increasing and the right-hand-side term is strictly decreasing in γ, Eq. (4.26) admits a unique solution, implying that the v/c ratio is a constant along the Pareto-optimal frontier (Case (ii) in Lemma 4.3.2). In particular, we have $\tilde{\gamma} = \bar{\gamma}$. This leads to $\tilde{\beta} < \bar{\beta}$, $\tilde{q} > \bar{q}$, $\tilde{y} > \bar{y}$ and $\tilde{p} < \bar{p}$. Therefore, if $m(\beta)$ is affine, the private sector tends to choose a lower capacity but a higher toll than does the public sector, resulting in the same service quality level as that at the social optimum.

Example 2. For the Gamma distribution, the hazard rate and mean residual VOT functions are given in Table 4.1. The curvature of $m(\beta)$ is completely determined by the generalized hazard rate function, $\beta h(\beta)$, which is free from the effect of the scaling parameter μ. Setting $\mu = 1$ and taking the derivative of $\beta h(\beta)$, we have

$$\begin{aligned}(\beta h(\beta))' &= \frac{-\int_1^{+\infty}(1-x)x^{\alpha-1}\exp(\beta(1-x))\mathrm{d}x}{\left(\int_1^{+\infty}x^{\alpha-1}\exp(\beta(1-x))\mathrm{d}x\right)^2} \\ &= \left(\frac{\alpha}{\beta} - 1\right)\frac{1}{\int_1^{+\infty}x^{\alpha-1}\exp(\beta(1-x))\mathrm{d}x} + \frac{1}{\beta}\frac{1}{\left(\int_1^{+\infty}x^{\alpha-1}\exp(\beta(1-x))\mathrm{d}x\right)^2}\end{aligned}$$

and

$$(\beta h(\beta))'' = \frac{1}{\beta^2}\frac{(\alpha-1)}{\left(\int_1^{+\infty} x^{\alpha-1}\exp(\beta(1-x))\mathrm{d}x\right)^3}\left\{\alpha\left(\int_1^{+\infty}(x-1)\,x^{\alpha-2}\exp\left(\beta\left(1-x\right)\right)\mathrm{d}x\right)^2 + \int_1^{+\infty}(x-1)\,x^{\alpha-2}\exp\left(\beta\left(1-x\right)\right)\mathrm{d}x + \left(\int_1^{+\infty} x^{\alpha-2}\exp\left(\beta\left(1-x\right)\right)\mathrm{d}x\right)\left(\int_1^{+\infty}(x-1)\,x^{\alpha-2}\exp\left(\beta\left(1-x\right)\right)\mathrm{d}x\right)\right\}$$

which implies that $m(\beta)$ is convex when $\alpha > 1$; affine when $\alpha = 1$; concave when $0 < \alpha < 1$. Therefore, gamma distribution with $\alpha > 1$, $\alpha = 1$ and $0 < \alpha < 1$ falls into Case (i), (ii) and (iii) in Lemma 4.3.2, respectively.

Example 3. If β follows a Weibull distribution, the hazard rate and mean residual VOT functions are given in Table 4.1 (Nassa and Eissa, 2003). From Eq. (4.4), we know

$$m'(\beta) = h(\beta)\,m(\beta) - 1 = \alpha\rho^{1-1/\alpha}e^{\rho}\Gamma\left(1+\frac{1}{\alpha}\right)\left(1-\Gamma_1\left(\rho,\frac{1}{\alpha}\right)\right) - 1.$$

To check the convexity of $m(\beta)$, we just need to check the increasing monotonicity of function $m'(\beta)$. Taking derivative of $m'(\beta)$ with respect to β, we get

$$m''(\beta) = \frac{\alpha(\alpha-1)\,\Gamma(1+1/\alpha)\,\rho^{2(1-1/\alpha)}e^{\rho}}{\mu\Gamma(1/\alpha)}\int_{\rho}^{+\infty}\left(\frac{x}{\rho}-1\right)x^{1/\alpha-2}e^{-x}\mathrm{d}x,$$

which Weibull distribution with $\alpha > 1$, $\alpha = 1$ and $0 < \alpha < 1$ falls into Case (i), (ii) and (iii) in Lemma 4.3.2, respectively.

To examine the results of Proposition 4.3.2, we consider the Weibull distributions with a scaling parameter $\mu = 1$. Without necessarily representing a realistic setting, we assume that $t_0 = 120$ (min), $t(\gamma) = 60(1+\gamma)$, $Q = 1$ $k = 1$ and $\hat{T} = 1$. With shape parameter $\alpha = 2.0$ and $\alpha = 0.8$, respectively, Figures 4.1 and 4.2 depict the contours of social welfare and profit in the two-dimensional (γ, q) space. The Pareto-optimal solution sets or contract curves, are denoted in bold curves connecting the MO and SO points. It is clear that, when moving from SO to MO in the two cases shown in the two figures, the v/c ratio increases and decreases, respectively, which is consistent with Proposition 4.3.2.

Table 4.1: The hazard rate and mean residual VOT functions

Distribution	CDF or density	Hazard rate $h(\beta)$	Mean residual VOT $m(\beta)$
Exponential	$f(\beta) = \mu \exp(-\mu\beta)$ $\mu > 0, \beta \geq 0$	$h(\beta) = \mu$ (constant)	$m(\beta) = 1/\mu$ (constant)
Uniform	$U\left[\beta_0, \beta^0\right]$ $\beta^0 \geq \beta_0 \geq 0$	$h(\beta) = 1/\left(\beta^0 - \beta\right)$ (IFR)	$m(\beta) = \left(\beta^0 - \beta\right)/2$ (Affine)
gamma	$f(\beta) = \frac{\mu^\alpha \beta^{\alpha-1} \exp(-\mu\beta)}{\gamma(\alpha)}$ $\alpha, \mu > 0, \beta > 0$	$h^{-1}(\beta) = \int_1^{+\infty} \beta x^{\alpha-1}$ $\exp(\mu\beta(1-x))\,\mathrm{d}x$ ($\alpha > 1$ IFR; $\alpha = 1$ constant; $0 < \alpha < 1$ DFR)	$m(\beta) = \frac{\beta h(\beta)}{\mu} + \frac{\alpha}{\mu} - \beta$ ($\alpha > 1$ convex; $\alpha = 1$ affine; $0 < \alpha < 1$ concave)
Weibull	$f(\beta; \mu, \alpha) = \frac{\alpha}{\mu}\left(\frac{\beta}{\mu}\right)^{\alpha-1} \exp\left(-\left(\frac{\beta}{\mu}\right)^{\alpha}\right)$ $\beta \geq 0, \alpha > 0, \mu > 0$	$h(\beta) = \frac{\alpha}{\mu}\left(\frac{\beta}{\mu}\right)^{\alpha-1}$ ($\alpha > 1$ IFR; $\alpha = 1$ constant; $0 < \alpha < 1$ DFR)	$m(\beta) = \mu e^{\rho} \Gamma\left(1 + \frac{1}{\alpha}\right)$ $\left(1 - \Gamma_1\left(\rho, \frac{1}{\alpha}\right)\right)$ $\rho = (\beta/\mu)^{\alpha}$ $\Gamma(1 + 1/\alpha) = \int_0^{\infty} x^{1/\alpha} e^{-x} \mathrm{d}x$ $\Gamma_1(\rho, 1/\alpha) = \frac{\int_0^{\rho} x^{1/\alpha - 1} e^{-x} \mathrm{d}x}{\Gamma(1/\alpha)}$ ($\alpha > 1$ convex; $\alpha = 1$ affine; $0 < \alpha < 1$ concave)
Lognormal	$f(\beta; \mu, \sigma) = \frac{\exp\left(-x^2/2\right)}{\sqrt{2\pi}\sigma\beta}$, $x = (\ln\beta - \mu)/\sigma$	$h(\beta) = \frac{\exp\left(-x^2/2\right)}{\sqrt{\pi/2}\sigma\beta \mathrm{erfc}\left(x/\sqrt{2}\right)}$, (increasing then decreasing or upside-down bathtub shape failure rate)	$m(\beta) = e^{\mu + \sigma^2/2} \frac{\mathrm{erfc}\left((x-\sigma)/\sqrt{2}\right)}{\mathrm{erfc}\left(x/\sqrt{2}\right)} - \beta$, $\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-u^2} \mathrm{d}u$, (decreasing then increasing or bathtub shape mean residual VOT function)

IFR: increasing failure rate; DFR: decreasing failure rate

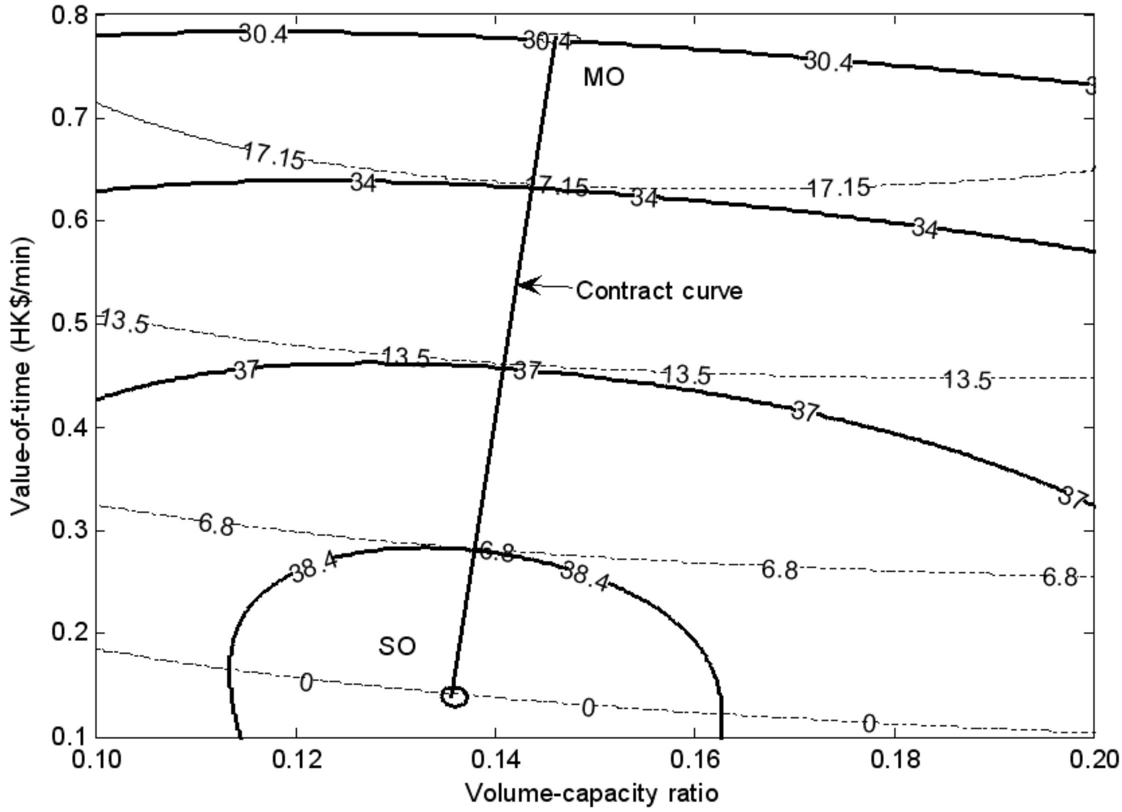


Figure 4.1: Contract curve with Weibull Distribution ($\mu = 1.0$, $\alpha = 2$)

The toll at any Pareto-optimal solution (β^*, γ^*) can be expressed as:

$$p^* = \frac{\beta^* h(\beta^*) m(\beta^*) \gamma^* t'(\gamma^*)}{1 + \left(\beta^* + \frac{1}{1+\lambda} m(\beta^*)\right) \left(m(\beta^*) + \beta^* - \beta^* h(\beta^*) m(\beta^*)\right)}. \tag{4.27}$$

Unlike the case with homogeneous users, here it is generally hard to compare the toll levels among the different Pareto-optimal solutions. Nonetheless, from $\lambda \geq 0$ and $m(\beta) + \beta \geq \beta h(\beta) m(\beta)$, we readily see $\tilde{p} \leq p^* \leq \bar{p}$. We know that, if the VOT distribution satisfies Assumptions 4.1 and 4.2, the private sector has incentive to set a higher toll level than the public sector and the intermediate Pareto-optimal toll must be higher than that at SO and lower than that at MO. This observation is useful for investigating the behavior of the private sectors under the regulations of the public sector. This result is restated in the following corollary for our late reference.

Corollary 4.3.1. *Under Assumptions 3.1(b), (c), 3.2, 3.3, 4.1 and 4.2, the toll level at any intermediate Pareto-efficient BOT contract must be larger than that at SO and less than that at MO, namely,* $\tilde{p} < p^* < \bar{p}$.

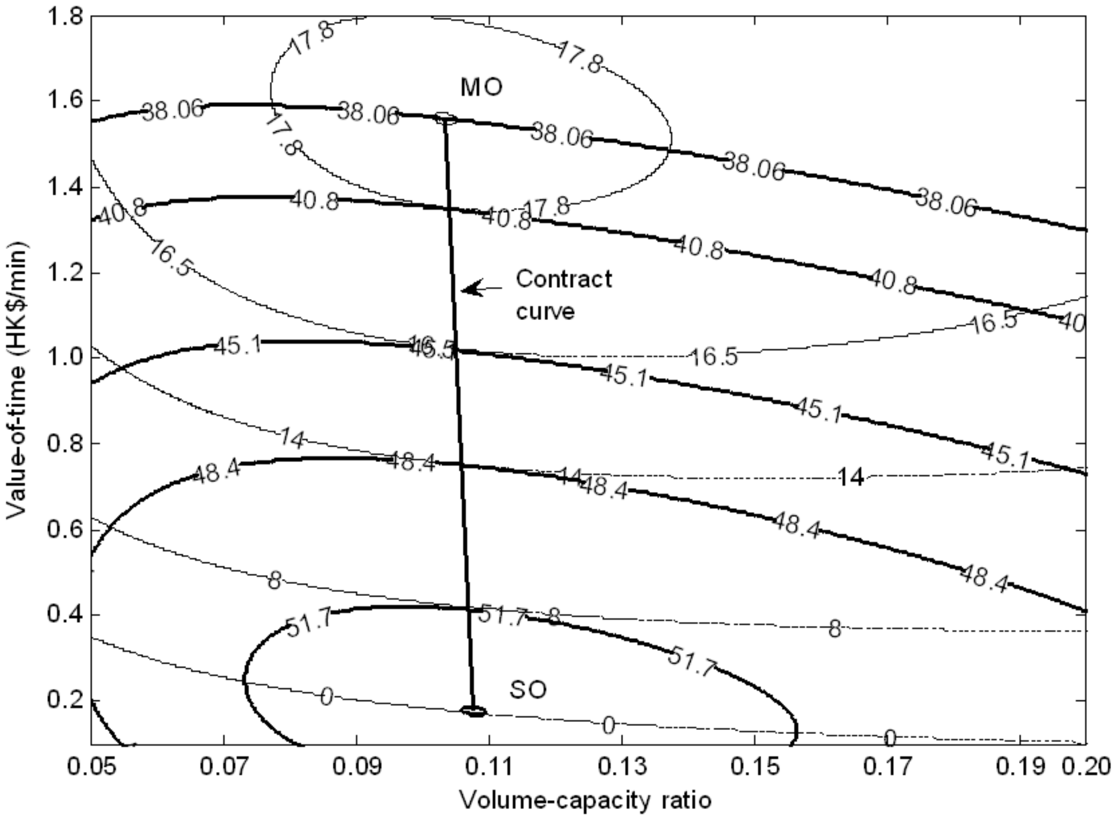


Figure 4.2: Contract curve with Weibull Distribution ($\mu = 1.0$, $\alpha = 0.8$)

4.4 Effects of User Heterogeneity on Regulations

The regulation of the public sector aims to induce the private sector to choose a predetermined Pareto-efficient BOT contract voluntarily. In more details, under the regulatory mechanism, the profit-maximizing private sector obtains greater profit when she chooses the Pareto-efficient BOT contract preset by the government than by any other alternative choice. In this section, we go to investigate the outcomes of the various regulatory mechanisms under the assumptions of Proposition 4.3.2. Namely, In addition to the standard Assumptions 3.1(b), (c), 3.2 and 3.3 about link cost and construction cost functions, we further assume that the VOT distribution function satisfies Assumptions 4.1 and 4.2.

We first discuss two classical regulatory mechanisms: price-cap and rate of return (ROR) regulations. Under the price-cap regulation, the private sector is allowed to set a price below or equal to the upper bound price imposed by the public sector; under ROR regulation, the private sector is allowed to earn no more than a "fair" rate of return on her investment, by freely choosing a combination of the BOT variables.

Let $\left(\hat{T}, \beta^*, y^*\right)$ be the predetermined Pareto-efficient BOT contract and the corresponding Pareto-optimal solution of problem (4.9) is (β^*, γ^*). For the price-cap regulation, the public sector sets a price constraint, $p \leq p^*$, where, $p^* = \beta^* (t_0 - t(\gamma^*))$. The problem of the private sector under the price-cap regulation is to select a combination of the concession period, capacity and toll levels to maximize the total profit under the toll restriction. Obviously, the private sector will choose a concession period equal to $\hat{T}$ whenever $p^* \geq 0$. She also chooses $p = p^*$, because choosing a lower toll level means that the price cap losses validity and she can obtain more profit return at a certain locally extreme solution. As discussed in Section 4.3, the local extreme solution must be MO solution since there only exist a unique locally extreme point under our assumptions. As a result, the MO toll $\bar{p}$ must be less than p^*, namely $\bar{p} \leq p^*$, which contradicts to Corollary 4.3.1. First we note that, under the binding price cap constraint, $\beta (t_0 - t(\gamma)) = p^*$, β can be regarded as an increasing function of γ. From the relationship (4.24) between the social welfare and the profit, we readily see that reducing γ results in a decrease of β, thereby yielding a larger difference between the total social welfare and profit, since $\bar{F}(\beta) m(\beta)$ is decreasing in β. Thus, reducing v/c ratio will surely improve the social welfare whenever the profit is enhanced, which conflicts the Pareto-optimality of (β^*, γ^*). As a result, the private sector may increase his profit gain only by increasing v/c ratio. Taking derivative of $P(\beta, \gamma)$ in γ yields

$$\frac{\mathrm{d}P}{\mathrm{d}\gamma} = \hat{T} Q \bar{F}(\beta) t'(\gamma) \left(\beta^2 h(\beta) - \frac{k \beta h(\beta)}{\hat{T} \gamma (t_0 - t(\gamma))} + \frac{k}{\hat{T} \gamma^2 t'(\gamma)} \right). \tag{4.28}$$

Substituting (4.20) and (4.21) into (4.28) leads to

$$\left. \frac{\mathrm{d}P}{\mathrm{d}\gamma} \right|_{(q^*, \gamma^*)} = \frac{\hat{T} Q \left(\bar{F}(\beta^*) \right) t'(\gamma^*) \left(m(\beta^*) + \beta^* \right)}{1 + \lambda} > 0 \tag{4.29}$$

which means that, under the price-cap regulation, the private sector tends to offer a higher v/c ratio or lower service quality. Since the private sector makes a profitable deviation to choose (β, γ) with $\gamma > \gamma^*$ and hence $\beta > \beta^*$, the resulting choice of road capacity would be lower or $y < y^*$. It should be noted that if the price ceiling is set to correspond to the MO solution $(\bar{\beta}, \bar{\gamma})$, the corresponding Lagrange multiplier $\lambda \to \infty$, and $\mathrm{d}P/\mathrm{d}\gamma|_{(\bar{\beta},\bar{\gamma})} = 0$, the private sector would not deviate since she can not earn any more profit. The result comfirms to the homogeneous case examined in Chapter 3 and the traditional economic literature (Train, 1991).

For a given Pareto-efficient BOT contract $\left(\hat{T}, \beta^*, y^*\right)$*, the private sector would choose a road-life concession period and offer a lower road capacity and lower service quality under the price-cap regulation* $p \leq p^*$*.*

Next we consider ROR regulation. For a BOT contract (T, β, y), we define the rate of return (ROR) on investment is given as:

$$\mathrm{ROR} = \frac{P(T, q, y)}{I(y)} = \frac{T\beta\gamma(t_0 - t(\gamma))}{k} - 1. \tag{4.30}$$

Let $s^* \geq 1$ be the ROR on investment at a predetermined Pareto-efficient contract $\left(\hat{T}, q^*, y^*\right)$. Under the ROR regulation, the public sector imposes an upper bound on the ROR:

$$\mathrm{ROR} \leq s^*. \tag{4.31}$$

It is clear that the profit-maximizing private sector will select a combination of the concession period, travel demand and road capacity such that constraint (4.31) becomes binding. Therefore, in view of $y = q/\gamma$, the profit-maximizing problem of the private sector under ROR regulation is equivalent to the following constrained optimization problem

$$\max_{0 \leq T \leq \hat{T}, \beta \geq 0, \gamma \geq 0} \quad s^* k Q \frac{\bar{F}(\beta)}{\gamma} \tag{4.32}$$

subject to the binding condition (4.31).

Viewing β as a function of concession period T and v/c ratio γ determined by the binding condition (4.31) and taking partial derivatives of β in T and γ, respectively, we obtain

$$\beta + T\frac{\partial\beta}{\partial T} = 0 \tag{4.33}$$

and

$$\beta\left(t_0 - t\left(\gamma\right) - \gamma t'\left(\gamma\right)\right) + \gamma\left(t_0 - t\left(\gamma\right)\right)\frac{\partial\beta}{\partial\gamma} = 0. \tag{4.34}$$

For any given γ, Eq. (4.33) implies that β is an increasing function of T. Thus, under ROR regulation, the private sector surely chooses the road life as the concession period or $T = \hat{T}$.

Furthermore, taking derivative of the objective function of problem (4.30) in γ, letting $T = \hat{T}$ and using $\partial\beta/\partial\gamma$ given by (4.34), we have

$$\frac{\mathrm{d}P}{\mathrm{d}\gamma} = -\frac{s^* kQ\bar{F}\left(\beta\right)\beta h\left(\beta\right)}{\gamma^2}\left(1 - \frac{1}{\beta h\left(\beta\right)} - \frac{\gamma t'\left(\gamma\right)}{t_0 - t\left(\gamma\right)}\right). \tag{4.35}$$

From Eqs. (4.22) and (4.21), we obtain

$$\left.\frac{\mathrm{d}P}{\mathrm{d}\gamma}\right|_{\left(\beta^*,\gamma^*\right)} = -\frac{s^* kq^*}{\left(\gamma^*\right)^2}\frac{\beta^* h\left(\beta^*\right)m\left(\beta^*\right)}{m\left(\beta^*\right) + \left(1+\lambda\right)\beta^*}\left(\frac{m\left(\beta^*\right) + \beta^*}{\beta^* h\left(\beta^*\right)m\left(\beta^*\right)} - 1\right) \leq 0. \tag{4.36}$$

The inequality follows from $m\left(\beta\right) + \beta \geq \beta h\left(\beta\right)m\left(\beta\right)$ and the strictly inequality is true if $m\left(\beta\right) > 0$ (in fact, $m\left(\beta\right) > 0$ is one of the necessary conditions for $m\left(\right)$ to be the mean residual VOT function of a certain continuously VOT distribution). Therefore, decreasing γ can improve the profit, giving rise to an increasing in β and y (note that increasing γ will decrease β under the binding condition (4.31), and decrease the resulting profit). Furthermore, viewing p as a function of γ and taking derivative of (4.2) in γ, we obtain $\mathrm{d}p/\mathrm{d}\gamma = \left(t_0 - t\left(\gamma\right)\right)\mathrm{d}\beta/\mathrm{d}\gamma - \beta t'\left(\gamma\right) < 0$. Thus, the toll charge will be increased when decreasing γ. In particular, if the ROR is set to correspond to the MO solution $\left(\bar{\beta},\bar{\gamma}\right)$, we have $\mathrm{d}P/\mathrm{d}\gamma|_{\left(\bar{\beta},\bar{\gamma}\right)} = 0$, and the private sector would not deviate since she can not earn more profit. Consequently, the outcome of ROR regulation is same as that with homogeneous users studied in Section 3.4.3

For a given Pareto-efficient BOT contract $\left(\hat{T}, \beta^*, y^*\right)$*, the private sector would choose a road-life concession period and offer a higher road capacity, higher toll level and higher service quality under the ROR regulation* $ROR \le ROR^*$*.*

For the case of homogeneous users, Chapter 3 proved that the demand and markup charge regulations can attain Pareto efficiency. The demand regulation is to set a minimal level of demand, while the markup charge regulation is to restrict the maximal amount of profit earned from each unit of realized demand (each trip) during the concession period. However, in the presence of user heterogeneity here, both mechanisms prove to be inefficient as elucidated below.

Under the demand regulation by the government for a targeted Pareto-efficient solution, $\left(\hat{T}, \beta^*, y^*\right)$, the private firm is allowed to make choices subject to the resulting realized traffic volume level, $q \ge q^*$. Denote (β^*, γ^*) as the corresponding Pareto-optimal solution of problem (4.11). The demand regulation is equivalent to restricting $\beta \le \beta^*$. Without loss of generality, suppose that the target Pareto-efficient BOT contract is profitable, otherwise, it would be unacceptable to the private sector. The private sector's problem is to consider the following maximization problem:

$$\max_{0 \le T \le \hat{T}, \beta \le \beta^*, \gamma \ge 0} Q\bar{F}(\beta)\left(T\beta\left(t_0 - t(\gamma)\right) - \frac{k}{\gamma}\right) \tag{4.37}$$

subject to $\beta \le \beta^*$. It is clear that the private firm will choose a life-time concession period $T = \hat{T}$ and select capacity and toll levels such that the VOT β and v/c ratio γ satisfy

$$\beta = \frac{k}{\hat{T}\gamma^2 t'(\gamma)}. \tag{4.38}$$

Comparing Eqs. (4.38) and (4.21), we know that $\gamma > \gamma^*$ whenever, $\beta^* < \bar{\beta}$. If $\beta^* = \bar{\beta}$, the private firm surely chooses the target contract since it is a monopoly optimum. In addition, note that the Assumptions 4.1 or 4.2 guarantees the unimodality of the profit function (Lariviere, 1999), we know that $\bar{\beta} > \beta^*$ for any non-monopoly Pareto-optimal BOT contract. Therefore, the private firm will not choose a lower VOT, or, equivalently, not offer service for more users. The reason is intuitive because if $\beta < \beta^*$, the problem (4.31) is equivalent to a unconstrained maximization problem, which gives rise to a monopoly optimal solution. As a result, the private firm will choose a lower capacity level and a lower toll charge since $\gamma > \gamma^*$ and $\beta = \beta^*$ (equivalently, $q = q^*$).

The markup charge regulation, as proposed in Chapter 3, means that the amount of profit earned from each unit of realized demand (each trip) during the concession period is restricted by a cap charge, p_2. It is given by:

$$p_2 \leq p_2^* = \frac{P\left(\hat{T}, \beta^*, y^*\right)}{\hat{T} q^*}, \tag{4.39}$$

where p_2 is the markup charge and is defined by

$$p_2 = \frac{P\left(T, \beta, y\right)}{T q}.$$

Under the markup regulation, the problem of the private firm is to maximize the net profit by selecting concession period, toll and capacity levels under condition (4.39). Note that the monopoly solution is the uniqueness of the locally extreme point. It is clear that the profit-maximizing firm surely does not allow a space between the final determined markup charge p_2 and the government regulated value p_2^*, and thus, condition (4.39) must be binding. After a simple rearrangement, the problem of the private firm under the markup regulation is equivalent to maximizing the demand or minimizing the corresponding VOT by selecting a combination of concession period and v/c ratio:

$$(T, \gamma) = \arg\min \left\{ \beta : \beta \left(t_0 - t\left(\gamma\right)\right) - \frac{k}{T\gamma} = p_2^*, T \in \left[0, \hat{T}\right], \gamma \geq 0 \right\}. \tag{4.40}$$

Following a similar discussion as the ROR regulation, we know that the private firm will choose a lifetime concession period, $T = \hat{T}$. Viewing β as a function of γ, the first-order condition implies that the private firm will select capacity and toll levels such that the frontier VOT β and v/c ratio γ satisfy condition (4.38). Substituting β determined by

Eq. (4.38) into the binding regulatory condition (4.39) yields

$$\frac{k\left(t_0-t\left(\gamma\right)\right)}{\hat{T}\gamma^2 t'\left(\gamma\right)}-\frac{k}{\hat{T}\gamma}=\beta^*\left(t_0-t\left(\gamma^*\right)\right)-\frac{k}{\hat{T}\gamma^*}\leq\frac{k\left(t_0-t\left(\gamma^*\right)\right)}{\hat{T}(\gamma^*)^2 t'\left(\gamma^*\right)}-\frac{k}{\hat{T}\gamma^*}. \tag{4.41}$$

The last inequality of Eq. (4.41) from the Pareto-optimal condition (4.21) and the inequality is strict whenever the target contract is not a monopoly optimum solution. Note that the function given by the left-hand-side term of Eq. (4.41) is a multiplication of two strictly decreasing and positive functions of γ, and thus $f\left(\gamma\right)$ is strictly decreasing with respect to γ, which, combining inequality in (4.41) implies that $\gamma > \gamma^*$ for any non-monopoly Pareto-optimal BOT contract. From $\gamma > \gamma^*$ and binding condition (4.39), we further obtain

$$p=p_2^*+\frac{k}{\hat{T}\gamma}<p_2^*+\frac{k}{\hat{T}\gamma^*}=p^*. \tag{4.42}$$

Since the problem of private sector is equivalent to maximizing the demand (minimizing the VOT), he must not choose a higher frontier VOT, and thus the demand will not decrease under the markup regulation, which results in a higher capacity level chosen by the private sector.

We now examine that whether the private firm will choose a strict lower VOT under the markup regulation. Consider function

$$\phi\left(\gamma\right)=\beta^*\left(t_0-t\left(\gamma\right)\right)-\frac{k}{\hat{T}\gamma}, \tag{4.43}$$

with $\phi\left(\gamma^*\right)=p_2^*$. Taking derivative of $\phi\left(\gamma\right)$ in γ gives rise to

$$\frac{\mathrm{d}\phi\left(\gamma\right)}{\mathrm{d}\gamma}=-t'\left(\gamma\right)\left(\beta^*-\frac{k}{\hat{T}\gamma^2 t'\left(\gamma\right)}\right). \tag{4.44}$$

It is clear that, from condition (4.21), $\mathrm{d}\phi\left(\gamma^*\right)/\mathrm{d}\gamma > 0$ whenever $\left(\hat{T},\beta^*,y^*\right)$ is a non-monopoly Pareto-optimal BOT contract. Thus, increasing γ from γ^* can strictly increase the markup charge if β^* is kept unchanged, which implies that the private sector can strictly decrease β from β^* to trade off the increase of markup charge and the resulting profit is strictly increased from the level associated with $\left(\hat{T}, \beta^*, y^*\right)$. Viewing β as a function of γ given by the binding condition (4.39), taking derivative of $y = Q\bar{F}(\beta)/\gamma$ directly, we get that $\mathrm{d}y/\mathrm{d}\gamma < 0$. To compare of the case with homogeneous users, we summarize the effects of demand and markup regulations into the following proposition.

Proposition 4.4.1. *Under Assumptions 3.1(b), (c), 3.2, 3.3, 4.1 and 4.2, for a given non-monopoly Pareto-efficient BOT contract* $\left(\hat{T}, \beta^*, y^*\right)$*, the private sector would choose a road-life concession period and offer a lower road capacity and toll levels, which result in a lower service quality under the demand and markup regulations; Both two have a positive profitable derivation.Furthermore, the resulting demand exactly equals the regulated level,* q^**, under demand regulation, while strictly higher than* q^* *under markup regulation; both regulations yield a positive profitable deviation.*

Given a predetermined Pareto-efficient BOT contract $\left(\hat{T}, \beta^*, y^*\right)$, the resulting frontier VOT and v/c ratio satisfy condition (4.38) under demand and markup regulations. Note that the latter induces a lower VOT or higher demand level (denoted by q^M). This means the private sector achieves more profit gain under the demand regulation. In addition, the consequent welfare gain of the markup regulation can be completely achieved by defining a new demand regulation using the demand level, q^M, resulted by the markup regulation. Therefore, the markup regulation dominates the demand regulation to achieve more efficiency in term of social welfare gain whenever $q^M \leq \tilde{q}$.

Example 4. Figure 4.3 compares the outcomes of the private firm's choices under various regulatory regimes for the example adopted in Example 3. For a predetermined Pareto-optimal solution, (β^*, γ^*) =(0.4\$/min, 0.1842), the rate of return of $s^* = 1.8584$ and the markup charge of $p_2^* = 13.447$ (\$). The four bold curves, L_1˜L_4, indicate the binding constraints in the (γ, β) space, associated with the five regulatory regimes: (1) $p \leq p^*$; (2) $s \leq s^*$; (3) $\beta \leq \beta^*$ or $q \geq q^*$ and (4) $p_2 \leq p_2^*$. The corresponding feasible domains are located above these curves in the (γ, β) space. As seen from the figure, each regulatory binding curve is tangent to a profit contour; the corresponding profit represents the maximum profit earned by the private firm under the given regulatory control. The

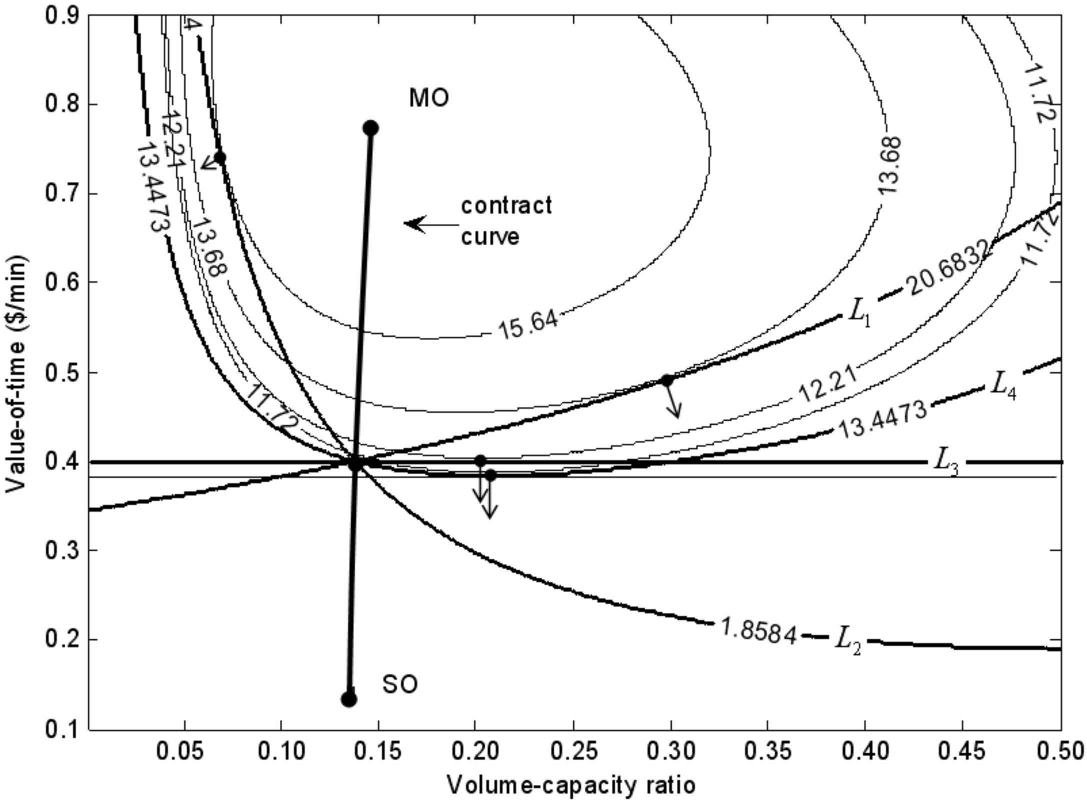


Figure 4.3: Regulations and outcomes with Weibull Distribution ($\mu = 1.0$, $\alpha = 2.0$)

choices made by the private firms under each regulatory regime are identified by the corresponding tangent point. In this numerical example, the markup regulation is more efficient than the other three alternatives. It is worth mentioning that more efficient regulation requires more information, and thus more costly to implement.

4.5 Conclusions

Built upon our work on the Pareto-efficient BOT contract and regulation for road franchising in the previous chapter, this part investigated the properties of a BOT toll road in the presence of user heterogeneity by employing a continuous VOT distribution. For the VOT distributions satisfying certain assumptions, we proved that the concession period is free from the effects of user heterogeneity, while the service quality (or, the volume-capacity ratio) is dependent on the curvature of the mean residual VOT function. We proved that: if the VOT distribution has a convex, affine, or concave mean residual VOT function, then the service quality is increasing, identical, or decreasing, respectively, along the Pareto-efficient frontier from social optimal solution to monopoly optimal solution. This chapter also investigated the behavior of the profit-maximizing private firm under various regulatory constraints. Distinguished from the earlier results derived in chapter 3, we found that, the regulatory regimes (demand and markup regulations), which are efficient to achieve any pre-determined Pareto-efficient BOT contract in the homogeneous case, fail to lead to an efficient outcome. Both regulatory regimes result in a lower investment and lower service quality. Furthermore, in contrast to the demand regulation, the markup regulation results in a higher demand level and a higher welfare gain.

5

FLEXIBLE BOT CONTRACTS WITH DEMAND UNCERTAINTY

This chapter investigates the optimal contract design for the highway franchising scheme incorporating demand uncertainty, which can not be ex ante forecasted by either of the public and private sectors. It is impossible to achieve the Pareto-efficient BOT contract in the incomplete environment. This chapter investigates the full and partial flexibility of the BOT contract according to the instruments adopted by the public and private sectors. Full flexibility refers to the case in which the public sector promises an exogenous rate of return on the private sector and in turn can freely ex post adjust the contract in a socially optimal manner according to the observed demand curve. Partial flexibility refers to the case where the public and private sector agree on an ex ante demand risk allocation by contract and the ex post contract adjustment can be made contingent on a Pareto-improvement to both sectors. A preferred Pareto-improvement can be selected from the Pareto-optimal solution set of a bi-objective programming problem equipped with a rational preference. The optimal BOT contracts with the two types of flexibility are examined by assuming that the public sector chooses the original contract variables to maximize the expected total social welfare while taking account of the ex post optimal adjustments.

5.1 Introduction

The studies in the previous chapter on the optimal BOT contracts (Homogeneous or heterogeneous users) critically rely on the availability of perfect information on project investment cost and future traffic demand. It is not unusual to encounter demand forecasts which differ from reality by 20%-60% (Skamris and Flyvbjerg, 1997). High estimation errors lead to either inefficient high levels of congestion or politically untenable levels of under-utilization, which require for a flexible toll schedule, except for a flexible concession period for cost recovery. Athias and Saussier (2007) provided various reasons for multiple toll adjustment provisions in order to adapt the contractual framework to unanticipated contingencies and to create incentives for cooperative behavior. Demand uncertainty and fixed term contracts can, indeed, make it impossible to fulfill the concession agreement in many cases, and contract renegotiation has been used to restore financial equilibrium. This has some undesirable economic consequences: selecting the most efficient concessionaire is no longer guaranteed, and prices lose their role as signals of allocative efficiency (Engel et al., 1997, 2001; Nombela and de Rus, 2004). For this reason, flexible-term contracts for road franchising have been proposed for achieving efficient pricing and cost recovery, without contract renegotiation, and they can be implemented fairly easily. For example, in a least-present-value-of-revenue (LPVR) auction, the bidding variable is the present value of the toll revenues; the lowest bid wins, and the franchise ends when the required amount has been collected (Engel et al., 2001; Nombela and de Rus, 2004). In this case, the linkage between traffic uncertainty and revenue uncertainty is effectively broken. The contract term is endogenously determined by the realized level of future demand, so it is shortened in case of high demand and extended if traffic levels are low. Following their earlier research (Engel et al., 2001), Engel et al. (2008) studied the construct of an optimal risk-sharing contract between the public and private sectors when there is substantial exogenous demand uncertainty, and then looked into the flexible term contract following the realization of demand. Their analyses of flexible-term contracts mainly focused on the concession period and demand uncertainty, a critical simplifying assumption is that traffic congestion and thus service quality is ignored or the travel time and thus traffic demand is independent of the road capacity or initial investment level for the new road (equivalent to assuming that the road capacity was predetermined and large enough).

In this chapter, we examine flexible BOT contracts for a new highway project, for which the risk-neutral public and private sectors need to determine concession period, road capacity and toll level in the face of demand uncertainty. It is assumed that, after the highway is built, the uncertain demand curve will be revealed and observed and one or more contract variables can be ex post adjusted according to the ex ante predetermined rules. It is also assumed that the concessionaire can be compensated with a combination of subsidies and future toll revenues, since the required up-front investment is generally huge for a highway project. The cost of one dollar of public finance or subsidy is assumed to be larger than one dollar of private spending to capture the inefficiency of the subsidy (de Palma et al., 2007b). The resulting BOT problem with flexible contracts is to select a combination of the BOT variables (capacity, toll and concession period) for maximizing the expected social welfare, under the constraint of the exogenous minimum attractive rate of return or minimum accepted rate of return (MARR) on private investment, while taking into account the ex post adjustment strategies. The strategies with postponements have been investigated in the marketing literature, in which, capacity and/or price decisions are made ex post till acquiring enough knowledge of demand (Chod and Rudi, 2005; Mieghem and Dada, 1999). We address two types of flexibility: full and partial flexibility. Full flexibility refers to the case in which the public sector guarantees an exogenous rate of return on the private investment and in return can freely ex post adjust the contract (concession period, toll, subsidy size). Partial flexibility refers to the case where demand risk is allocated between the public and private sectors and the ex post contract adjustment can be made contingent on a Pareto-improvement to both parties.

The chapter is organized as follows. Section 5.2 introduces the basic definitions and assumptions. Section 5.3 describes the BOT problems with and without flexibility and Section 5.4 provides some important properties of the optimal contracts for the benchmark with perfect information on demand. Section 5.5 investigates the public sector's optimal strategy for the BOT problem with full flexibility and Section 5.6 analyzes the BOT contract with partial flexibility and the corresponding ex post Pareto-improving adjustment strategies. A numerical example is provided to elucidate our results in Section 5.7. Finally, general conclusions are given in Section 5.8.

5.2 The Basic Definitions and Model

Assume that the public sector plans to get a private sector to build a new highway with an exogenously given life time, $\hat{T}$. It is assumed that there is a constant marginal construction cost, k $(k > 0)$, with respect to the highway capacity, y, namely, the initial construction cost is $I(y) = ky$. There are no operation/maintenance (O/M) costs. For many highway projects, the upfront construction costs are much higher than the O/M costs. If the O/M costs are constant or proportional to demand for the project, then the case with O/M costs is a trivial extension as suggested by Engel et al. (2008). Let p denote the toll charge for using the road and $t(q, y)$ denote the travel time through the road as a function of traffic demand, q, and road capacity, y. The price-dependent demand is modeled as

$$q = D(\mu, \theta), \tag{5.1}$$

where μ is the full price or the generalized travel cost of a trip through the highway and θ is a random variable that captures the demand uncertainty. It is assumed that θ is realized after the highway is built. We are concerned with the demand uncertainty or deviation that occurs before and after the highway is built. The effects of the demand fluctuation of the project after the highway is built, which is a multiple period dynamic problem, is not considered in this chapter. And, to avoid technicalities, it follows a cumulative distribution function, $F(\theta)$, with a continuous density function and a support $\Theta \subset (0, +\infty)$. The generalized travel cost, μ, is given by a linear combination of travel time, t, and toll charge, p, namely, $\mu = p + \beta t(q, y)$, where β is the value-of-time (VOT), converting time into equivalent monetary cost (we consider homogeneous users only). For any given θ, let $B(q, \theta)$ be the corresponding inverse demand function of $D(\mu, \theta)$.

Therefore, for any given toll charge, p, capacity, y and realization of θ, the demand, q, is determined by the following demand-supply equilibrium:

$$B(q, \theta) = p + \beta t(q, y). \tag{5.2}$$

Throughout this chapter, $B(q, \theta)$ is assumed to be strictly decreasing in q, strictly increasing in θ, and twice differentiable in q and θ.

Given the large up-front investment of the highway project. the private sector can be compensated with a combination of subsidy and toll revenue. Like Engel et al. (2008), to model the link of the project and the public funds, the marginal cost of public funds (MCPF), λ, is assumed to be larger than or equal to one (raising one HK$ will cost λ HK$ to society), which captures the inefficiency of the subsidy adopted in the project.

When the highway capacity, y, and the toll charge, p, are selected, the demand, q, is determined by Eq. (5.2), the unit-time (year or month) revenue for a realization of θ can be calculated as:

$$R(p, y, \theta) = qp. \tag{5.3}$$

Let P^{S} be the subsidy received from the public sector. The profit of the private sector during the concession period, T, for any realization of θ, can be calculated as

$$P\left(T, p, P^{\mathrm{S}}, y, \theta\right) = TR(p, y, \theta) + P^{\mathrm{S}} - I(y). \tag{5.4}$$

The unit-time social surplus during the concession period is the sum of the consumer surplus and the toll revenue

$$S(p, y, \theta) = \mathrm{CS}(p, y, \theta) + R(p, y, \theta), \tag{5.5}$$

where the unit-time consumers surplus, CS, can be expressed as

$$\mathrm{CS}(p, y, \theta) = \int_0^q B(w, \theta)\, \mathrm{d}w - q(p + \beta t(q, y)). \tag{5.6}$$

Substituting (5.3) and (5.6) into (5.5), we have

$$S\left(p,y,\theta\right)=\int_{0}^{q}B\left(w,\theta\right)\mathrm{d}w-\beta qt\left(q,y\right). \tag{5.7}$$

During the post-concession period, $\hat{T}-T$, the road capacity is fixed and the demand curve is realized, leading to a deterministic problem. In this case, the public sector can select a toll charge to maximize the unit-time social surplus

$$\tilde{S}\left(y,\theta\right)=\max_{p_1\geq 0}\ S\left(p_1,y,\theta\right). \tag{5.8}$$

The total social welfare can be expressed as

$$\begin{aligned} W\left(T,p,P^{\mathrm{S}},y,\theta\right)= \quad & TCS\left(p,y,\theta\right)+P\left(T,p,P^{\mathrm{S}},y,\theta\right) \\ & -\lambda P^{\mathrm{S}}+\left(\hat{T}-T\right)\tilde{S}\left(y,\theta\right), \end{aligned} \tag{5.9}$$

where the first two terms of Eq. (5.9) are the total consumers surplus and total producer surplus during the concession period; the third term is the cost of subsidy; the last term is the total social surplus during the post-concession period.

Note that Assumptions 3.1 (b), (c), 3.2 and 3.3 are still made throughout in this chapter. In addition, to consider the demand uncertainty, we revise the Assumption 3.1 (a) on the functions, $B\left(q,\theta\right)$.

Assumption 5.1. *The function $qB\left(q,\theta\right)$ is strictly concave in q for any θ.*

With Assumption 3.1 (b) and (c), $t\left(q,y\right)$ is convex and thus $qt\left(q,y\right)$ is convex in q for any given $y>0$. Combining the strict concavity of $qB\left(q,\theta\right)$, the revenue function $R\left(p,y,\theta\right)$ is strictly concave in q for any given y and θ. Assumption 5.1 is common in revenue management literature to assume the concavity of the "revenue function" (Ziya et al., 2004).

5.3 Models and Contract Flexibility

We assume that both the public sector and private firm are risk-neutral, and consider the following various strategies to deal with demand uncertainty.

5.3.1 The BOT contract without flexibility

Without allowing for contract flexibility, all contract variables of capacity, concession period, toll charge (or equivalently, demand), and subsidy size should be ex ante determined by both parties according to their prior knowledge on the demand. Let $\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)$ be the ex ante BOT contract, the BOT problem can be modeled as

$$\max_{0 \leq T_{\mathrm{I}} \leq \hat{T}, p_{\mathrm{I}} \geq 0, P_{\mathrm{I}}^{\mathrm{S}} \geq 0, y_{\mathrm{I}} \geq 0} \mathrm{E}\left(W\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right)\right) \tag{5.10}$$

with participation constraint of the private sector

$$\mathrm{E}\left(P\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right)\right) \geq rI\left(y_{\mathrm{I}}\right), \tag{5.11}$$

where $\mathrm{E}(\cdot)$ denotes the expectation and r is the minimum attractive rate of return (MARR) on private investment. Note that, different from the participation constraint of the private sector proposed by Guo and Yang (2010b), the present study adopts the financial indicator of MARR rather than the exogenously required profit (net prevent value of future revenue minus initial investment). In real world, the return rate of investment is an important index for the sectors to assess their investment, specially in the presence of investment risks. The MARR varies from high rates for those seeking high-risk investments to low rates for those looking for security. In this chapter, the MARR is assumed to be exogenously given. A number of models and processes are available to determine the value of MARR (Sullivan et al., 2006).

5.3.2 The BOT contract with full flexibility

In contrast with the case without flexibility, we suppose the public sector bears all the demand risk, and in return, after observed demand curve and contingent on the ex ante capacity choice, can ex post freely adjust the contract variables, T, p or q, and determine the subsidy size. Therefore, the BOT contract with full flexibility can be viewed as the following problem: the public sector ex ante selects capacity and ex post determines other contract variables after the demand curve is observed, while the return rate of the investment of the private sector is guaranteed at any realized demand state.

Let y_{I} be the ex ante capacity and $\left(T_{\mathrm{II}}, p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}\right)$ be the ex post choice of the remaining contract variables after the demand curve is observed. The BOT problem with full flexibility can be modeled as:

$$\max_{y_{\mathrm{I}} \geq 0} \mathrm{E}\left(W\left(y_{\mathrm{I}}, \theta\right)\right) \tag{5.12}$$

subject to

$$P\left(T_{\mathrm{II}}, p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right) \geq rI\left(y_{\mathrm{I}}\right), \forall \theta \in \Theta. \tag{5.13}$$

Constraint (5.13) means that the private sector can receive a predetermined return rate on its investment regardless of the future demand state, and thus it is risk-free. It is clear that $\left(T_{\mathrm{II}}, p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}\right)$ are functions of the demand state θ.

5.3.3 The BOT contract with partial flexibility

For the BOT contract with partial flexibility, it is assumed that both the public and private sectors share the project risk, a compromised case between the two polar cases of rigid and fully flexible BOT contracts. In this case the ex post contract adjustment should be Pareto-improving and thus would be acceptable to both parties. However, there is a discrepancy between the two parties in choosing their own preferred Pareto-improvement from a set of available Pareto-improvements. To resolve this discrepancy, we first specify the following rules of the ex post contract adjustment:

Rule a. *the private sector should adopt a non-dominated Pareto-improvement and alter the original toll as little as possible;*

Rule b. *the public sector can modify the contract for the interests of the public by compensating the concessionaire through changes in toll, concession period and subsidies.*

The first rule specifies the freedom of the private sector to adjust the original contract, which must guarantee the current realized public interests and adopt a least (positive or negative) amendment on toll within the Pareto-improvement domain with elastic demand. The second rule simply states that the public sector modifies the contract for the public interests. In this case, the public sector should compensate the loss of the private sector caused by the modification by a combinative strategy. It is common to predetermine some rules for the ex post contract adjustment in reality. In Chile, similar rules are specified in detail by the concession laws for the toll road projects, in which, the toll can be adjusted by the private sector within some predetermined limits (Lorenzen et al., 2003).

To simplify our discussion, we assume that toll is the only flexible instrument, which can be adjusted according to the observed demand curve. This can be the case in reality if the private sector does not accept an extension of the duration of concession period without increasing the minimum attractive rate of return of its up-front investment. Let $\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)$ denote the original contract and, for any realization of θ, let p_{II} and $P_{\mathrm{II}}^{\mathrm{S}}$ denote the ex post toll and additional subsidy (if necessary), we then have the following constrained bi-objective programming problem:

$$\max_{p_{\mathrm{II}}\geq 0, P_{\mathrm{II}}^{\mathrm{S}}\geq 0} \begin{pmatrix} W\left(p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}\middle| T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right) \\ P\left(p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}\middle| T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right) \end{pmatrix} \tag{5.14}$$

subject to

$$\begin{pmatrix} W\left(p_{\text{II}}, P_{\text{II}}^{\text{S}}\middle| T_{\text{I}}, p_{\text{I}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) \\ P\left(p_{\text{II}}, P_{\text{II}}^{\text{S}}\middle| T_{\text{I}}, p_{\text{I}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) \end{pmatrix} \geq \begin{pmatrix} W\left(T_{\text{I}}, p_{\text{I}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) \\ P\left(T_{\text{I}}, p_{\text{I}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) \end{pmatrix}, \tag{5.15}$$

where ex post social welfare and profit are

$$W\left(p_{\text{II}}, P_{\text{II}}^{\text{S}}\middle| T_{\text{I}}, p_{\text{I}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) = W\left(T_{\text{I}}, p_{\text{II}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) - \lambda P_{\text{II}}^{\text{S}}, \tag{5.16}$$

and

$$P\left(p_{\text{II}}, P_{\text{II}}^{\text{S}}\middle| T_{\text{I}}, p_{\text{I}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) = P\left(T_{\text{I}}, p_{\text{II}}, P_{\text{I}}^{\text{S}}, y_{\text{I}}, \theta\right) + P_{\text{II}}^{\text{S}}. \tag{5.17}$$

Note that problem (5.14)-(5.15) has a non-dominated or Pareto-optimal solution set. To incorporate adjustment Rules (a) and (b), we equip the Pareto-optimal solution set with the following preference:

$$\left(\breve{p}_{\text{II}}, \breve{P}_{\text{II}}^{\text{S}}\right) \prec \left(p_{\text{II}}, P_{\text{II}}^{\text{S}}\right) \Leftrightarrow \begin{pmatrix} \breve{P}_{\text{II}}^{\text{S}} > P_{\text{II}}^{\text{S}} \geq 0 \text{ or} \\ \breve{P}_{\text{II}}^{\text{S}} = P_{\text{II}}^{\text{S}} \geq 0 \text{ and } \left|\breve{p}_{\text{II}} - p_{\text{I}}\right| > \left|p_{\text{II}} - p_{\text{I}}\right| \end{pmatrix} \tag{5.18}$$

It is clear to see that preference relation (5.18) satisfies transitive and complete assumptions, and thus, it is rational (Varian, 1992). If $\breve{P}_{\text{II}}^{\text{S}} = P_{\text{II}}^{\text{S}} = 0$, the toll p_{II} is the non-dominated Pareto-solution of problem (5.14)-(5.15) with the minimal increment, therefore, according to Rule (a), p_{II} is allowed to be set by the private sector. As to be proved later, at the optimum of problem (5.14)-(5.15) equipped with preference (5.18), Rule (b) is fulfilled as well, namely, the public sector shall only subsidize the private sector to modify the contract if the resulting social welfare after the adjustment under Rule (a) can be further enhanced. These together imply that the optimal solution of problem (5.14)-(5.15) with preference (5.18) is the unique Pareto-improvement under Rules (a) and (b). It is difficult to design franchising contract with toll-flexibility. To allow user fees to be adjusted by the concessionaire would result in monopoly pricing as the case of the Orange County 91 Express Lanes (Engel et al., 2008). The restriction of Pareto-improvement mitigates to a certain degree the negative consequence of contract flexibility.

the public sector, considering the ex post adjustment performance, determines an original contract $\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}\right)$ to maximize the expected total social welfare with the expected profit constraint, namely,

$$\max_{\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)} \mathrm{E}\left(W\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right)\right) = \int_{\Theta} W\left(p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}} \middle| T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right) \mathrm{d}F\left(\theta\right) \tag{5.19}$$

with the constraint

$$\mathrm{E}\left(P\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right)\right) = \int_{\Theta} P\left(p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}} \middle| T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right) \mathrm{d}F\left(\theta\right) \geq rI\left(y_{\mathrm{I}}\right). \tag{5.20}$$

5.3.4 Further discussions of the BOT contract with variable degree of flexibility

For the fully flexible contract defined by problem (5.12)-(5.13), the private sectors are generally risk-free and the contract can be awarded through a competitive auction organized by the public sector, such as, the least-present-value-revenue (LPVR) auction proposed by Engel et al. (2001) after determining the capacity or investment level. Thus, a reasonable minimum attractive rate of return on the investment can be ascertained by the auction.

For the partially flexible contract defined by problem (5.14)-(5.15) with preference (5.18), any possible loss of the investors is not compensated by the public subsidy. The ex post subsidy in the second stage is offered by the public sector for compensation to the investors only if a welfare-improving adjustment causes a decrease from the investors' potential profit level based on Rule (a). The compensation mechanisms are adopted in practice including the provision of public subsidy, variation of tolls and change of contract length. The contract is renewed through the renegotiation between the parties concerned. Comparing to the renegotiation procedure, the partially flexible contract considered here is a self-renewal contract since the adjustment is Pareto-improving. Vassallo (2006) investigated three mechanisms applied in Chile to mitigate the traffic risk: the minimum income guarantee, the least present value of the revenue and the revenue distribution mechanism; Soong (1999, ADB report) listed numbers of examples including the three types of toll roads contracts. This type of self-renewing contract can potentially enhance those compensation mechanisms prevailing in practice. For example, in the presence of demand elasticity, without requirement for the Pareto-improvement, the minimum income guarantee mechanism could result in a huge welfare loss. A careful examination of the existence of Pareto-adjustment is surely favourable before a compensation mechanism is implemented. In the contract with minimum income guarantee, if the realized revenue is lower than 70 percent of the total up-front investment plus the estimated maintenance and operation cost, the public sector will provide a subsidy to the investor; whereas a more than 15% rate of return on investment will trigger a revenue-sharing mechanism (Vassallo, 2006).

The ex post adjustment for the BOT contract with full flexibility is also Pareto-improving, because the public sector will not lower the private benefit due to guaranteed MARR when determining the contract variables. It is worth mentioning here that the MARRs for the models with and without contract flexibility are different. Among the three kinds of BOT contracts, including the rigid one, the public sector should set a lowest MARR for the contract with full flexibility since the private sector bears no demand risk, and a highest MARR for the rigid contract since the private sector is not allowed to adjust the contract but bears all demand risk. The rigid contract would result in an underinvestment, while the fully flexible contract reduces the cost-cutting incentive of private sector and causes overrun of the project cost. The partially flexible contract is an ideal compromise between the two polar cases.

Generally speaking, the BOT problem without flexibility is of less interest unless in the absence of uncertainty or for whatever political reasons, since, for any given MARR, the outcomes with a partially flexible BOT contract will always dominate a rigid one. Hereafter, we will consider the latter two types of BOT contracts with full or partial flexibility.

5.4 Benchmark Analysis with Perfect Information

Consider the benchmark case with perfect information or parameter θ is known to both parties. In this case, the BOT problem becomes the following maximization problem:

$$\max_{0\leq T\leq \hat{T}, p\geq 0, P^{\mathrm{S}}\geq 0, y\geq 0} W\left(T, p, P^{\mathrm{S}}, y, \theta\right) \tag{5.21}$$

with the following participation constraint of the private sector

$$P\left(T, p, P^{\mathrm{S}}, y, \theta\right) \geq rI\left(y\right). \tag{5.22}$$

A pertaining issue here is to determine when a public subsidy is desired and what size is optimal in the BOT problem. The following proposition from model (5.21) and (5.22) offers the public sector a guideline on provision of subsidy.

Proposition 5.4.1. *Under Assumptions 5.1, there exists a unique $r_0 \geq 0$ such that the private sector must be compensated by a positive subsidy to implement the social welfare-maximizing BOT contract if and only if the minimum attractive rate of return (MARR) is lager than r_0.*

Proof. Firstly, it is clear that $T^{*} = \hat{T}$ at the optimal solution of problem (5.21). A rigorous proof of the optimal concession period can be found in Proposition 2 in Guo and Yang (2010b) in the deterministic optimum case, and Proposition 3.3.1 in Chapter 3 in the deterministic Pareto-optimum case. Rewriting Eq. (5.2) as

$$p\left(q, y\right) = B\left(q, \theta\right) - \beta t\left(q, y\right), \tag{5.23}$$

which means that determining the variables p and y is essentially equivalent to selecting the variables q and y for the benchmark case (θ is known to both parties). Therefore, we substitute BOT contract $\left(T, q, P^{\mathrm{S}}, y\right)$ for $\left(T, p, P^{\mathrm{S}}, y\right)$ and consider the following two polar cases of the original problem.

Relaxing the non-negative constraint of subsidy. In this case, the subsidy (positive or negative) always considered by the public sector, P^{S} is expressed as:

$$P^{\mathrm{S}} = (1+r)\,ky - \hat{T}qp. \tag{5.24}$$

After substituting P^{S} into the objective function (5.9) and in view of $T = \hat{T}$, problem (5.21) becomes the following unconstrained programming problem of q and y.

$$(\mathbf{P1}):\ \max_{p\geq 0, y\geq 0}\ \hat{T}S\,(p,y,\theta) + (r-\lambda-r\lambda)\,I\,(y) - (1-\lambda)\,\hat{T}qp. \tag{5.25}$$

A negative subsidy in ($\mathbf{P1}$) means that the public sector extracts revenue from the highway system, which is, however, not the optimal solution of the original problem (5.21)-(5.22), because the public sector would not value the revenue of one HK$ as one HK$.

No subsidy. For this problem, subsidy is not considered by the public sector and thus $P^{\mathrm{S}} = 0$. In this case, we consider the following problem:

$$(\mathbf{P2}):\ \max_{p\geq 0, y\geq 0}\ \hat{T}S\,(p,y,\theta) - ky \tag{5.26}$$

subject to

$$\hat{T}R\,(p,y,\theta) - ky \geq rky. \tag{5.27}$$

where p is the function of q and y, given by (5.23).

Let $\left(\hat{T}, q^*, P^{\mathrm{S}^*}, y^*\right)$ be the optimal solution of problem (5.21). We know that, if $P^{\mathrm{S}^*} > 0$, $\left(\hat{T}, q^*, y^*\right)$ must be the solution of problem ($\mathbf{P1}$), and vice versa; if $P^{\mathrm{S}^*} = 0$, $\left(\hat{T}, q^*, P^{\mathrm{S}^*}, y^*\right)$ must be the solution of problem ($\mathbf{P2}$), and vice versa. The proof is organized in the following several steps.

(1) The optimal demand, $\hat{q}^*$, and optimal capacity, $\hat{y}^*$ of ($\mathbf{P1}$) are continuously differentiable and strictly decreasing with r.

Evidently, the optimal demand $\hat{q}^*$ and capacity $\hat{y}^*$ of problem (**P1**) must satisfy the following first-order conditions

$$\left.\frac{\mathrm{CS}_q}{R_q}\right|_{q=\hat{q}^*,y=\hat{y}^*} = -\lambda \tag{5.28}$$

and

$$(R_y)_{q=\hat{q}^*,y=\hat{y}^*} = \frac{k}{\hat{T}} + \left(1 - \frac{1}{\lambda}\right)\frac{rk}{\hat{T}}, \tag{5.29}$$

where CS_q denote the partial derivative of consumer surplus, CS, with respect to q. Similarly, S_q, S_{qq}, S_{qy},R_q, R_{qq} and R_{qy} denote the first- and second-order partial derivatives of the corresponding functions in the subscript variables. Equation (5.28) states that the ratio of the marginal consumer surplus to the marginal revenue in demand at optimum is exactly equal to the marginal cost of public funds, λ. From Eqs. (5.28) and (5.29), using the implicit function theorem, we have

$$A\begin{pmatrix} \frac{\mathrm{d}\hat{q}^*}{\mathrm{d}r} \\ \frac{\mathrm{d}\hat{y}^*}{\mathrm{d}r} \end{pmatrix} = \begin{pmatrix} 0 \\ \left(1 - \frac{1}{\lambda}\right)\frac{k}{\hat{T}} \end{pmatrix}$$

where A is 2×2 matrix

$$A = \begin{pmatrix} \frac{S_{qq}R_q - R_{qq}S_q}{R_q^2} & \frac{S_{qy}R_q - R_{qy}S_q}{R_q^2} \\ S_{qy} & S_{yy} \end{pmatrix}.$$

By direct calculation, we know $|\mathrm{A}| < 0$ and

$$A\begin{pmatrix} \frac{\mathrm{d}\hat{q}^*}{\mathrm{d}r} \\ \frac{\mathrm{d}\hat{y}^*}{\mathrm{d}r} \end{pmatrix} < \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Therefore, $(\hat{q}^*, \hat{y}^*)$ is continuously differentiable in r and both $\hat{q}^*$ and $\hat{y}^*$ strictly decrease with r.

*(2) The optimal demand, $\breve{q}^*$, and optimal capacity, $\breve{y}^*$, of (**P2**) are continuously differentiable and strictly decreasing in r; furthermore, at optimum, function, $(CS_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*}$, is strictly decreasing from -1 to $-\infty$ as $r \to +\infty$.*

Note that the first-best solution results in a zero profit under Assumptions 5.1 and 3.2. At the optimum of problem (**P2**), constraint (5.27) becomes binding:

$$\hat{T}R(p, y, \theta) = (1 + r)\,ky. \tag{5.30}$$

The optimal demand, $\breve{q}^*$, and capacity, $\breve{y}^*$ satisfy the following optimality condition:

$$\left.\frac{\mathrm{CS}_q}{R_q}\right|_{q=\breve{q}^*, y=\breve{y}^*} = \frac{rk\breve{y}^*}{\hat{T}(\breve{q}^*)^2 B_q(\breve{q}^*, \theta)} - 1. \tag{5.31}$$

Viewing $(\breve{q}^*, \breve{y}^*)$ as a function of r, the standard proof from the implicit-function theorem readily obtains the continuous differentiability of $\breve{q}^*$ and $\breve{y}^*$. Similar to the proof in step (1), we further have that both $\breve{q}^*$ and $\breve{y}^*$ strictly decrease in r. Taking derivative of both sides of Eq. (5.31) in r, we obtain that $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*, y=\breve{y}^*}$ also strictly decreases in r and $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*, y=\breve{y}^*} = -1$ with $r = 0$. Furthermore, we have that $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*, y=\breve{y}^*}$ approaches $-\infty$ as $r \to +\infty$. To prove this, we only need to prove that $g(\breve{q}^*, \breve{y}^*, r)$ approaches $-\infty$ as $r \to +\infty$, where

$$g(\breve{q}^*, \breve{y}^*, r) = \frac{rk\breve{y}^*}{\hat{T}(\breve{q}^*)^2 B_q(\breve{q}^*, \theta)}. \tag{5.32}$$

Since $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*, y=\breve{y}^*}$ is strictly decreasing in r, if $g(\breve{q}^*, \breve{y}^*, r)$ does not approach $-\infty$ as $r \to +\infty$, then $g(\breve{q}^*, \breve{y}^*, r)$ is lower bounded and thus convergent as $r \to +\infty$. Without loss of generality, we assume that $g(\breve{q}^*, \breve{y}^*, r)$ approaches $-g_0$, where g_0 is a certain positive number, mathematically, for any infinitely small positive number $\varepsilon > 0$ ($\varepsilon < 1/2$), there exists r_1 such that for any $r > r_1$, we have $g(\breve{q}^*, \breve{y}^*, r) > -g_0 + \varepsilon$. Taking derivative of the social welfare, W, in q under the binding condition (5.27), we have

$$\begin{aligned}\left.\tfrac{\mathrm{d}W}{\mathrm{d}y}\right|_{q=\breve{q}^*, y=\breve{y}^*} &= rk + \left(\tfrac{\mathrm{CS}_q}{R_q}\tfrac{\mathrm{d}q}{\mathrm{d}y}\right)_{q=\breve{q}^*, y=\breve{y}^*} \\ &> rk\left(1 + (-g_0 - \varepsilon)\tfrac{1}{1-(-g_0+\varepsilon)}\right) \\ &> 0,\end{aligned} \tag{5.33}$$

which contradicts the optimality of $\breve{q}^*$ and $\breve{y}^*$.

(3) The private sector must be compensated by a positive subsidy to implement the social

welfare-maximizing BOT contract if and only if the MARR is lager than r_0.

Proving the above statement is equivalent to proving that, there exists a unique $r_0 \geq 0$, problem (5.21)-(5.22) is equivalent to (**P1**) if $r > r_0$ and equivalent to (**P2**) if $r \leq r_0$. From step (2), setting $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*} = -\lambda$ induces a unique solution r_0 since $\lambda \geq 1$ and $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*}$ strictly decreases from -1 to $-\infty$ with r. From Eqs. (5.30) and (5.31), we obtain

$$(R_y)_{q=\breve{q}^*,y=\breve{y}^*} = \frac{k}{\hat{T}} + \left(1 - \frac{1}{-(\mathrm{CS}_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*}}\right)\frac{rk}{\hat{T}}. \tag{5.34}$$

Substituting $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*} = -\lambda$ into (5.34), we know that $(\breve{q}^*, \breve{y}^*)$ satisfies conditions (5.28) and (5.29), or $(\breve{q}^*, \breve{y}^*) = (\hat{q}^*, \hat{y}^*)$ at $r = r_0$.

For any r with $0 \leq r < r_0$, $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*} > -\lambda$, and thus $\frac{\breve{q}^*}{\breve{y}^*} < \frac{\hat{q}^*}{\hat{y}^*}$. From Eq. (5.34). Rewriting conditions (5.28) and (5.31), respectively, as

$$B(\hat{q}^*, \theta) + \left(1 - \frac{1}{\lambda}\right)\hat{q}^*\frac{\partial B(\hat{q}^*, \theta)}{\partial q} = \beta t\left(\frac{\hat{q}^*}{\hat{y}^*}\right) + \beta\frac{\hat{q}^*}{\hat{y}^*}t'\left(\frac{\hat{q}^*}{\hat{y}^*}\right) \tag{5.35}$$

and

$$\begin{aligned} &B(\breve{q}^*, \theta) + \left(1 - \tfrac{1}{-(\mathrm{CS}_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*}}\right)\breve{q}^*\tfrac{\partial B(\breve{q}^*,\theta)}{\partial q} \\ &= \beta t\left(\tfrac{\breve{q}^*}{\breve{y}^*}\right) + \beta\tfrac{\breve{q}^*}{\breve{y}^*}t'\left(\tfrac{\breve{q}^*}{\breve{y}^*}\right) \end{aligned} \tag{5.36}$$

From Assumption 3.2, combining $(\mathrm{CS}_q/R_q)_{q=\breve{q}^*,y=\breve{y}^*} > -\lambda$ and $\frac{\breve{q}^*}{\breve{y}^*} < \frac{\hat{q}^*}{\hat{y}^*}$, we know that $\breve{q}^* > \hat{q}^*$ and thus $\breve{y}^* > \hat{y}^*$. It is clear to see both the consumer surplus, CS, and the net profit, $rk\breve{y}^*$, at $(\breve{q}^*, \breve{y}^*)$ are larger than those at $(\hat{q}^*, \hat{y}^*)$. Therefore, $(\breve{q}^*, \breve{y}^*)$ is the optimal solution of problem (5.21)-(5.22). By the same token, we readily obtain that, for any r with $r > r_0$, $(\hat{q}^*, \hat{y}^*)$ is the optimal solution of problem (5.21)-(5.22). In summary, when the MARR is lager than r_0, a positive size of subsidy is adopted by the public sector at optimum; while when the MARR is less than or equal to r_0, no positive subsidy can enhance the social welfare. The proof is completed. □

Proposition 5.4.1 provides a criterion for the public sector to consider subsidy in the public-private partnership (PPP) projects. Designing the subsidy schedule is of importance to the public sector in the PPP highway projects: the public sector guarantees investment return, on one hand, to reduce the cost-cutting incentive of the private sector, on the other hand, to prevent exhausting the user fees to overcompensate the investment. When the exogenous MARR is less than or equal to r_0, subsidy can not enhance the social welfare; it is better for the public sector to compensate the concessionaire solely using the future toll revenue. However, when the MARR is larger than r_0, subsidy can improve the social welfare and, at optimum, the gain in social welfare will exceed the total investment cost (construction cost plus the weighted subsidy cost). The subsidy effect on the optimal social welfare can be explained geometrically in Figure 5.1 (the parameters take the values set in Section 5.7.

The logic of this proposition is intuitive. From the proof in Appendix 1, if the substitution ratio of the marginal consumer surplus to the marginal revenue in demand exceeds the marginal cost of public funds, λ, or if,

$$\left.\frac{\mathrm{CS}_q}{R_q}\right|_{q=q^*,y=y^*} < -\lambda, \tag{5.37}$$

then it is unwise for the public sector to refuse compensation to the private sector using public funds. The reason is, in this case, that one dollar private investment causes more than λ equivalent dollars of consumer' benefit, and thus the social benefit is enhanced by substituting one-dollar private investment with equivalent public fund. However, the MCPF, which is used to capture the inefficiency of public subsidy, is hard to estimate for a practical application. If the public sector values the MCPF too high, then the subsidy schedule would be Ruled out from the BOT problem. As a result, a second-best solution of the social welfare maximization problem with zero- or positive return rate constraint has to be achieved. If the public sector is efficient and the user fees and public funds are perfectly substitutive to each other (there is no additional cost using public funds), then the public sector always implements a welfare-maximizing BOT contract by virtue of subsidy if need regardless of the required rate of return.

From the classical road self-financing theorem under Assumptions 3.1-3.3 (Mohring and Harwitz, 1962), the profit at social optimum is zero, the reader may refer to Verhoef and Mohring (2009) for a latest exposition of the theorem. The theorem is also true even under optimal selection of the concession length for the welfare maximizing problem (5.21) without participation constraint (5.22) (Guo and Yang, 2010b, and Chapter 3 in this thesis). Thus, the strictly positive return constraint (5.22) must be binding at the optimum solution of problem (5.21) and (5.22). We note that, regardless of the necessity of subsidy, condition (5.22) is always binding at the optimum solution of problem (5.21) and (5.22). If no subsidy is needed at the optimal solution of problem (5.21)-(5.22), then the binding condition 5.22 reduces to:

$$\hat{T}R\left(p\left(q^*\right), y^*, \theta\right) = (1+r)\, k y^*. \tag{5.38}$$

which implies that the MARR is just satisfied. If a subsidy is required to compensate the private sector at the optimum solution, we know that the corresponding optimal demand must satisfy the first-order condition (5.28), In its details,

$$B\left(\hat{q}, \theta\right) + \left(1 - \frac{1}{\lambda}\right)\hat{q}\frac{\partial B\left(\hat{q}, \theta\right)}{\partial q} - \beta t\left(\frac{\hat{q}}{y}\right) - \beta\frac{\hat{q}}{y}t'\left(\frac{\hat{q}}{y}\right) = 0 \tag{5.39}$$

and the corresponding toll $\hat{p}$ is given by Eq. (5.2). For any given capacity y. let $\check{q}$ be the demand level determined by Eq. (5.39), at which the minimum attractive return rate is just satisfied, and the corresponding toll charge is denoted by $\check{p}$. Correspondingly, let $\tilde{q}$ and $\bar{q}$ be the first-best and monopoly demand levels, which maximize the unit-time social welfare, S, and unit-time revenue, R, respectively. The corresponding toll charges are denoted by $\tilde{p}$ and $\bar{p}$, for given capacity y. It is clear that $\tilde{q}$ and $\bar{q}$ solve the following equations, respectively,

$$B\left(\tilde{q}, \theta\right) - \beta t\left(\frac{\tilde{q}}{y}\right) - \beta\frac{\tilde{q}}{y}t'\left(\frac{\tilde{q}}{y}\right) = 0 \tag{5.40}$$

and

$$B\left(\bar{q}, \theta\right) + \bar{q}\frac{\partial B\left(\bar{q}, \theta\right)}{\partial q} - \beta t\left(\frac{\bar{q}}{y}\right) - \beta\frac{\bar{q}}{y}t'\left(\frac{\bar{q}}{y}\right) = 0. \tag{5.41}$$

Viewing $\tilde{q}$ and $\bar{q}$ as functions of y, we know that both $\tilde{q}$ and $\bar{q}$ are strictly increasing functions of y. Evidently, for any $y > 0$, $\tilde{q}$ is located between $\tilde{q}$ and $\bar{q}$, because, from the proof of Proposition 5.4.1, CS_q/R_q is strictly decreasing in q, $\mathrm{CS}_q/R_q = -1$ at $q = \tilde{q}$, and $\mathrm{CS}_q/R_q = -\infty$ at $q = \bar{q}$, respectively.

5.5 BOT Problem with Full Flexibility

In this section, we assume that the public sector bears all the demand risk, and in return has the power to ex post adjust all contract variables contingent on the ex ante selected capacity according to the realized demand curve. The BOT problem with full flexibility is modeled by programming problem (5.12) and (5.13). We first examine the public sector's optimal adjustment strategy after the highway is built and the capacity is given y_{I}. Recall that $\tilde{q}$ and $\hat{q}$ are the socially optimal demand and the demand with subsidy at capacity y_{I}, given by conditions (5.40) and (5.39), respectively, and $\tilde{p}$ and $\hat{p}$ are the corresponding toll charges. For given $y_{\mathrm{I}} > 0$, we define:

$$\bar{\theta}\left(y_{\mathrm{I}}\right) = \inf\left\{\theta \in \Theta : \hat{T}R\left(\tilde{p}, y_{\mathrm{I}}, \theta\right) \geq \left(1+r\right)ky_{\mathrm{I}}\right\} \tag{5.42}$$

and

$$\underline{\theta}\left(y_{\mathrm{I}}\right) = \sup\left\{\theta \in \Theta : \hat{T}R\left(\hat{p}, y_{\mathrm{I}}, \theta\right) \leq \left(1+r\right)ky_{\mathrm{I}}\right\}. \tag{5.43}$$

It is easy to determine $\bar{\theta}\left(y_{\mathrm{I}}\right)$ and $\underline{\theta}\left(y_{\mathrm{I}}\right)$ when support set Θ is given. Let $\Theta = (0, +\infty)$. If, for any $\theta \in \Theta$, $\hat{T} \cdot R\left(\tilde{p}, y_{\mathrm{I}}, \theta\right) > \left(1+r\right)ky_{\mathrm{I}}$, then $\bar{\theta}\left(y_{\mathrm{I}}\right) = +\infty$; if, for any $\theta \in \Theta$, $\hat{T}R\left(\tilde{p}, y_{\mathrm{I}}, \theta\right) < \left(1+r\right)ky_{\mathrm{I}}$, then $\bar{\theta}\left(y_{\mathrm{I}}\right) = 0$; otherwise, solving for $\tilde{q}$ from the following equation

$$\beta\hat{T}(q)^2\frac{\partial t\left(q, y_{\mathrm{I}}\right)}{\partial q} = \left(1+r\right)ky_{\mathrm{I}} \tag{5.44}$$

and substituting the solution into Eq. (5.40) give rise to $\bar{\theta}\left(y_{\mathrm{I}}\right)$. Similarly, $\underline{\theta}\left(y_{\mathrm{I}}\right)$ can be determined by combining Eqs. (5.38) and (5.44). The following lemma reveals the properties of the infimum and supremum of the high and low demand states, $\bar{\theta}(y_{\rm I})$ and $\underline{\theta}(y_{\rm I})$.

Lemma 5.5.1. *Both $\bar{\theta}(y_I)$ and $\underline{\theta}(y_I)$ are non-decreasing in y_I; for any capacity $y_I > 0$, $0 \le \underline{\theta}(y_I) \le \bar{\theta}(y_I) \le +\infty$ and*

(i) $\hat{T}R(\tilde{p}, y_I, \theta) > (1+r)\,ky_I$ *at* $\theta \in \left(\bar{\theta}(y_I), +\infty\right) \bigcap \Theta$;

(ii) $\hat{T}R(\breve{p}, y_I, \theta) = (1+r)\,ky_I$ *for a certain* $\breve{p} \in [\tilde{p}, \hat{p}]$ *at* $\theta \in \left[\underline{\theta}(y_I), \bar{\theta}(y_I)\right] \bigcap \Theta$;

(iii) $\hat{T}R(\hat{p}, y_I, \theta) < (1+r)\,ky_I$ *at* $\theta \in (0, \underline{\theta}(y_I)) \bigcap \Theta$.

Proof. Without loss of generality, suppose $\bar{\theta}(y_{\rm I}) < +\infty$ and $\underline{\theta}(y_{\rm I}) > 0$. If there is a θ with $\theta > \bar{\theta}(y_{\rm I})$, such that $\hat{T}R(\tilde{p}, y_{\rm I}, \theta) \le (1+r)\,ky_{\rm I}$, then for any $\theta^{\#} < \theta$, we have

$$\hat{T}R\left(\tilde{p}^{\#}, y_{\rm I}, \theta^{\#}\right) = \hat{T}\beta\left(\frac{\tilde{q}^{\#}}{y_{\rm I}}\right)^2 t'\left(\frac{\tilde{q}^{\#}}{y_{\rm I}}\right) < \hat{T}\beta\left(\frac{\tilde{q}}{y_{\rm I}}\right)^2 t'\left(\frac{\tilde{q}}{y_{\rm I}}\right) < (1+r)\,ky_{\rm I}, \tag{5.45}$$

where both $\left(\tilde{q}^{\#}, \theta^{\#}\right)$ and $(\tilde{q}, \theta)$ satisfy condition (5.40). Inequality (5.45) implies that θ is a lower bound of the set $\left\{\theta \in \Theta : \hat{T}R(\tilde{p}, y_{\rm I}, \theta) \ge (1+r)\,ky_{\rm I}\right\}$, which contradicts definition (5.42). Conclusion (i) is proved.

Viewing $\hat{q}$, given by (5.28), as a function of θ for given $y_{\rm I}$, we have

$$\frac{{\rm d}\hat{q}}{{\rm d}\theta} = \frac{R_q\left(\lambda B_\theta + (\lambda - 1)\left(B_\theta + qB_{q\theta}\right)\right)}{S_{qq}R_q - S_qR_{qq}}. \tag{5.46}$$

Taking the derivative of $R(\hat{p}, y_{\rm I}, \theta)$ in Θ, we obtain

$$\frac{{\rm d}R(\hat{p}, y_{\rm I}, \theta)}{{\rm d}\theta} = R_q\frac{{\rm d}\hat{q}}{{\rm d}\theta} + R_\theta. \tag{5.47}$$

If ${\rm d}\hat{q}/{\rm d}\theta \le 0$, we have ${\rm d}R(\hat{p}, y_{\rm I}, \theta)/{\rm d}\theta > 0$ since $R_q(\hat{p}, y_{\rm I}, \theta) < 0$ and $R_\theta > 0$; If ${\rm d}\hat{q}/{\rm d}\theta > 0$, or $\lambda B_\theta + (\lambda-1)\left(B_\theta + qB_{q\theta}\right) < 0$, we have

$$\begin{aligned}\frac{{\rm d}R(\hat{p}, y_{\rm I}, \theta)}{{\rm d}\theta} &= \frac{R_q\left(R_\theta(S_{qq} + (\lambda-1)R_{qq}) - R_q\left(B_\theta + (\lambda-1)\left(B_\theta + qB_{q\theta}\right)\right)\right)}{S_{qq}R_q - S_qR_{qq}} \\ &> 0.\end{aligned} \tag{5.48}$$

Since revenue $R(\hat{p}, y_{\rm I}, \theta)$ strictly increases in θ, if there is a θ with $\theta < \underline{\theta}(y_{\rm I})$, such that $\hat{T}R(\hat{p}, y_{\rm I}, \theta) \ge (1+r)\,ky_{\rm I}$, then for any $\theta^{\#} > \theta$, $\hat{T}R\left(\hat{p}^{\#}, y_{\rm I}, \theta^{\#}\right) > (1+r)\,ky_{\rm I}$. Therefore, θ is a upper bound of the set $\left\{\theta \in \Theta : \hat{T}R\left(\hat{p}, y_{\mathrm{I}}, \theta\right) \leq (1+r)\, k y_{\mathrm{I}}\right\}$, which contradicts definition (5.43). We obtain conclusion (iii). In addition, (i) and (iii) simply imply (ii).

Suppose $y_1 > y_2$, if $\bar{\theta}\left(y_1\right) < \bar{\theta}\left(y_2\right)$, then there is θ such that $\bar{\theta}\left(y_1\right) < \theta < \bar{\theta}\left(y_2\right)$. From definition (5.42) and conclusion (i), we know that

$$\hat{T}R\left(\tilde{p}_1, y_1, \theta\right) \geq (1+r)\, k y_1 \tag{5.49}$$

and

$$\hat{T}R\left(\tilde{p}_2, y_2, \theta\right) < (1+r)\, k y_2. \tag{5.50}$$

Together with (5.40), we obtain

$$\hat{T}\beta\left(\frac{\tilde{q}_1}{y_1}\right)^2 t'\left(\frac{\tilde{q}_1}{y_1}\right) \geq (1+r)\, k > \hat{T}\beta\left(\frac{\tilde{q}_2}{y_2}\right)^2 t'\left(\frac{\tilde{q}_2}{y_2}\right). \tag{5.51}$$

Since $\hat{T}\beta\gamma^2 t'(\gamma)$ is monotonically increasing in γ, Eq. (5.51) implies $\tilde{q}_1/y_1 = \tilde{\gamma}_1 > \tilde{\gamma}_2\tilde{q}_2/y_2$. Note that, at optimum, $B(q,\theta) = \beta t(\gamma) + \beta\gamma t'(\gamma)$. We further have $B\left(\tilde{q}_1,\theta\right) > B\left(\tilde{q}_2,\theta\right)$, and thus, $\tilde{q}_1 < \tilde{q}_2$ since $B(q,\theta)$ strictly decreases in q. These two relations, $\frac{\tilde{q}_1}{y_1} > \frac{\tilde{q}_2}{y_2}$ and $\tilde{q}_1 > \tilde{q}_2$, give rise to $y_1 < y_2$, which contradicts our assumption that $y_1 > y_2$. Therefore, $y_1 > y_2$ must imply $\bar{\theta}\left(y_1\right) \geq \bar{\theta}\left(y_2\right)$, or $\bar{\theta}(y)$ is non-decreasing in y. In a similar manner, we can readily prove that $\underline{\theta}(y)$ is non-decreasing in y. The proof is completed. □

Lemma 5.5.1 depicts that $\Theta \subset \left(0, \underline{\theta}\left(y_{\mathrm{I}}\right)\right) \bigcup \left[\underline{\theta}\left(y_{\mathrm{I}}\right), \bar{\theta}\left(y_{\mathrm{I}}\right)\right] \bigcup \left(\bar{\theta}\left(y_{\mathrm{I}}\right), +\infty\right)$ is a partition of the future demand state space. The domain $\left(0, \underline{\theta}\left(y_{\mathrm{I}}\right)\right) \bigcap \Theta$ represents a low demand state. When the realized θ locates in $\left(0, \underline{\theta}\left(y_{\mathrm{I}}\right)\right) \bigcap \Theta$, even setting toll level $\hat{p}$, corresponding to the demand level $\hat{q}$, the toll revenue is not enough to compensate the private sector's investment and the required return. In this case, setting a higher toll level to increase total revenue is inefficient than directly subsidizing the private sector. The domain $\left(\bar{\theta}\left(y_{\mathrm{I}}\right), +\infty\right) \bigcap \Theta$ represents a high demand state. When the realized θ locates in $\left(\bar{\theta}\left(y_{\mathrm{I}}\right), +\infty\right) \bigcap \Theta$, the toll revenue at the first-best or marginal cost pricing $\tilde{p}$ is enough to compensate the private sector's investment and the required return. In this case, it is evident for the public sector to set the first-best toll charge $\tilde{p}$. Between the two domains, or if the realized θ locates in $\left[\underline{\theta}\left(y_{\rm I}\right), \bar{\theta}\left(y_{\rm I}\right)\right] \bigcap \Theta$, the public sector can always select a toll level $\breve{p}$ higher than $\tilde{p}$ and lower than $\hat{p}$ such that the total toll revenue is exactly equal to the sum of the private sector's investment and the required return. In this case, from the discussions in Section 3, the best response for the public sector is to set toll level $\breve{p}$, and subsidy can not enhance the total social welfare. We term this domain $\left[\underline{\theta}\left(y_{\rm I}\right), \bar{\theta}\left(y_{\rm I}\right)\right] \bigcap \Theta$ as representing an intermediate demand state. In addition, $\underline{\theta}\left(y_{\rm I}\right)$ is the upper bound of the low demand state, which is useful to calculate the subsidy probability. No public funds should be transferred to the private sector if the realized θ is larger than $\underline{\theta}\left(y_{\rm I}\right)$. The subsidizing probability is $\Pr\left\{\theta \in \left(0, \underline{\theta}\left(y_{\rm I}\right)\right) \bigcap \Theta\right\}$.

According to the realized demand states, the public sector's optimal ex post adjustment procedure is stated in the following proposition.

Proposition 5.5.1. *For any given capacity y_I, suppose $\left(T_{II}^*, p_{II}^*, P_{II}^{S*}\right)$ is the optimal solution of problem (5.13), we have*

(i) if $\theta \in \left(\bar{\theta}\left(y_I\right), +\infty\right) \bigcap \Theta$, then $p_{II}^* = \tilde{p}$, $T_{II}^* \geq (r+1)\, k y_I / R\left(\tilde{p}, y_I, \theta\right)$ with $P_{II}^{S*} = 0$;

(ii) if $\theta \in \left[\underline{\theta}\left(y_I\right), \bar{\theta}\left(y_I\right)\right] \bigcap \Theta$, then $p_{II}^* = \breve{p}$, $T_{II}^* = \hat{T}$ with $P_{II}^{S*} = 0$;

(iii) if $\theta \in \left(0, \underline{\theta}\left(y_I\right)\right) \bigcap \Theta$, then $p_{II}^* = \hat{p}$, $T_{II}^* = \hat{T}$ with $P_{II}^{S*} = (1+r)\, k y_I - \hat{T}\hat{q}p^* > 0$.

Engel et al. (2008) investigated optimal flexible PPP contracts with inelastic demand. They also extended their method to consider price-response demand but without congestion externality. The partition of low, intermediate and high demand states in their method, in fact, corresponds to the three domains in Proposition 5.5.1 with a given and fixed capacity level $y_{\rm I} > 0$. Nevertheless, their method depends on the joint distribution of the critical profit levels corresponding to the critical demand states, $\bar{\theta}\left(y_{\rm I}\right)$ and $\underline{\theta}\left(y_{\rm I}\right)$, which is hard to determine. Under our assumption that the benefit function is strictly increasing in the uncertainty term θ, the partition of the three demand states is directly determined by the method proposed in Lemma 5.5.1 and Proposition 5.5.1.

After establishing the optimal strategy in the second stage when capacity $y_{\rm I}$ is given, the expected social welfare according to the prior knowledge on demand can be calculated as

$$\begin{aligned} \mathrm{E}\left(W\left(y_{\mathrm{I}},\theta\right)\right) &= \int_{\Theta_1}\left(\hat{T}S\left(\hat{p},y_{\mathrm{I}},\theta\right)+\left(1-\lambda\right)\hat{T}R\left(\hat{p},y_{\mathrm{I}},\theta\right)\right)\mathrm{d}F\left(\theta\right) \\ &+\int_{\Theta_2}\left(\hat{T}S\left(\breve{p},y_{\mathrm{I}},\theta\right)\right)\mathrm{d}F\left(\theta\right) \\ &+\int_{\Theta_3}\left(\hat{T}S\left(\tilde{p},y_{\mathrm{I}},\theta\right)\right)\mathrm{d}F\left(\theta\right) \\ &-\left(\lambda-1\right)\left(r+1\right)ky_{\mathrm{I}}\Pr\{\Theta_1\}-ky_{\mathrm{I}}, \end{aligned} \tag{5.52}$$

where $\Theta_1=\left(0,\underline{\theta}\left(y_{\mathrm{I}}\right)\right)\bigcap\Theta$, $\Theta_2=\left[\underline{\theta}\left(y_{\mathrm{I}}\right),\bar{\theta}\left(y_{\mathrm{I}}\right)\right]\bigcap\Theta$, and $\Theta_3=\left(\bar{\theta}\left(y_{\mathrm{I}}\right),+\infty\right)\bigcap\Theta$. The optimal capacity is selected to maximize the expected social welfare, namely, ${y^*}_{\max}=\arg\max_{y_{\mathrm{I}}\geq 0}\left\{\mathrm{E}\left(W\left(y_{\mathrm{I}},\theta\right)\right)\right\}$. Note that, because $\mathrm{E}\left(W\left(y_{\mathrm{I}},\theta\right)\right)$ is not concave in y_{I} for many distributions, there is no simple condition available to determine the optimal capacity analytically. However, for any given cumulative distribution function of the uncertain term, θ, various numerical approaches can be used to determine the optimal capacity.

5.6 BOT Problem with Partial Flexibility

In this section, we go further to consider the BOT problem with partial flexibility, or the problem (5.19)-(5.20). We first investigate the condition for the existence of Pareto-improvement for any given original contract $\left(T_{\mathrm{I}},p_{\mathrm{I}},P_{\mathrm{I}}^{\mathrm{S}},y_{\mathrm{I}}\right)$ with a realization θ.

If the original toll p_{I} is less than $\tilde{p}$, namely, $p_{\mathrm{I}}<\tilde{p}=B\left(\tilde{q},\theta\right)-\beta t\left(\tilde{q}/y_{\mathrm{I}}\right)$, then both unit-time social surplus, S, and unit-time revenue R, can be strictly improved by increasing p_{I}. In this case, the original toll is set too low at capacity y_{I}, a proper increase in toll will lead to a Pareto-improvement in social welfare and profit.

If the original toll p_{I} is set too high such that $p_{\mathrm{I}}>\hat{p}=B\left(\hat{q},\theta\right)-\beta t\left(\hat{q}/y_{\mathrm{I}}\right)$, then, the realized demand is lower than $\hat{q}$, similar to the proof of Proposition 5.4.1, we know that, $\left(\partial\mathrm{CS}/\partial q\right)/\left(\partial R/\partial q\right)$ strictly decreases in q for given y_{I} and θ, and thus $\left(\partial\mathrm{CS}/\partial q\right)/\left(\partial R/\partial q\right)<-\lambda$. This implies that the positive gain of the consumer surplus is larger than $\left(\lambda-1\right)\Delta R$ the additional social cost associated with the substitution of revenue ΔR with a subsidy.

Therefore, it is beneficial for the public sector to substitute the revenue with an equivalent amount of subsidy. In this case, a Pareto-improving toll adjustment thus exists.

Finally, it is clear to see that when the original toll p_{I} is in the range of $\hat{p} \leq p_{\mathrm{I}} \leq \tilde{p}$, then there is no Pareto-improving toll adjustment.

To facilitate our discussion, we now partition the state-space as $\Theta = \Theta_1 \bigcup \Theta_2 \bigcup \Theta_3 \bigcup \Theta_4$, where

$$\begin{aligned} &\Theta_1 (p_{\mathrm{I}}, y_{\mathrm{I}}) = \{\theta \in \Theta : q \leq \bar{q}\}, \qquad \Theta_2 (p_{\mathrm{I}}, y_{\mathrm{I}}) = \{\theta \in \Theta : \bar{q} < q \leq \hat{q}\}, \\ &\Theta_3 (p_{\mathrm{I}}, y_{\mathrm{I}}) = \{\theta \in \Theta : \hat{q} < q \leq \tilde{q}\}, \qquad \Theta_4 (p_{\mathrm{I}}, y_{\mathrm{I}}) = \{\theta \in \Theta : q > \tilde{q}\}. \end{aligned} \tag{5.53}$$

The four domains represent four possible outcomes. Domain Θ_1 and Θ_2 represent the situation that the original toll is set too high such that the realized demand is lower than the demand level with subsidy, $\hat{q}$. In domain Θ_4, the toll is too low and increasing the current toll charge to the first-best toll level is a Pareto-improving adjustment, because the profit of the private sector is also enhanced. However, in domain Θ_3, it is impossible to enhance both parties' interests by adjusting toll. According to the assumption that $B(q, \theta)$ is strictly increasing in θ, it is clear to see that, for any y_{I} and p_{I}, all four demand quantities, $\tilde{q}$, $\hat{q}$, $\bar{q}$ and q, increase with θ and they satisfy $\tilde{q} > \hat{q} > \bar{q}$ for any state θ. Therefore, domains Θ_1, Θ_2, Θ_3 and Θ_4 partition the state space sequentially from low to high demand level. Note that, depending on the original contract, one or more domains may be empty in the sense of zero probability of occurrence or some outcomes may not happen in practice. Furthermore, from the analysis in Section 4.3, $\hat{q}$ approaches $\tilde{q}$ as MCPF, λ decreases. In particular, domain Θ_3 degenerates to an empty set for $\lambda = 1.0$, which implies that the public sector tends to implement the first-best contract if there is no additional cost of using public funds for subsidy. Finally for given original capacity, y_{I}, and toll, p_{I}, all four demand quantities, $\tilde{q}$, $\hat{q}$, $\bar{q}$ and q, are determined uniquely by parameter θ from (5.40), (5.38), (5.41) and (5.2). As a result, it is easy to sort out the four domains, Θ_1, Θ_2, Θ_3 and Θ_4.

We now discuss how to ex post adjust the toll level based on the occurrence of demand state. When $\theta \in \Theta_1$ or p_{I} is larger than , we prove that the unique Pareto-improvement $\left(p_{\mathrm{II}}^*, P_{\mathrm{II}}^{\mathrm{S}*}\right)$ with $p_{\mathrm{II}}^* = \hat{p}$ and $P_{\mathrm{II}}^{\mathrm{S}*} = T_{\mathrm{I}}\left(R\left(\bar{p}, y_{\mathrm{I}}, \theta\right) - R\left(\hat{p}, y_{\mathrm{I}}, \theta\right)\right)$ is the optimal solution of the problem (5.14)-(5.15) with preference (5.18). Namely, $\left(p_{\mathrm{II}}^*, P_{\mathrm{II}}^{\mathrm{S}*}\right)$ is the adjusted contract adopted by both parties according to Rules (a) and (b) for the following reasons. Firstly, any feasible solution $\left(p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}\right)$ with $p_{\mathrm{II}} \in \left(\bar{p}, p_{\mathrm{I}}\right]$ and $P_{\mathrm{II}}^{\mathrm{S}} \geq 0$ must be dominated by $\left(\bar{p}, P_{\mathrm{II}}^{\mathrm{S}}\right)$ since decreasing the toll improves the profit and the social welfare, simultaneously because the ex ante toll level is larger than the monopoly toll level. Furthermore, according to the proof of Proposition 5.4.1, at $\bar{q}$, inequality (5.37) is satisfied since $\mathrm{CS}_q / R_q|_{q=\bar{q}, y=y_{\mathrm{I}}} = -\infty$, and thus, $\left(\bar{p}, P_{\mathrm{II}}^{\mathrm{S}}\right)$ is dominated by $\left(\hat{p}, P_{\mathrm{II}}^{\mathrm{S}} + P_{\mathrm{II}}^{\mathrm{S}*}\right)$. From preference relation (5.18), we know that $\left(\hat{p}, P_{\mathrm{II}}^{\mathrm{S}} + P_{\mathrm{II}}^{\mathrm{S}*}\right) \prec \left(\hat{p}, P_{\mathrm{II}}^{\mathrm{S}*}\right)$. Secondly, for any pair $\left(p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}\right)$ with $p_{\mathrm{II}} \leq \bar{p}$ and $P_{\mathrm{II}}^{\mathrm{S}} \geq 0$, again, from preference relation (5.18), $\left(p_{\mathrm{II}}, P_{\mathrm{II}}^{\mathrm{S}}\right) \prec \left(\bar{p}, P_{\mathrm{II}}^{\mathrm{S}}\right) \prec \left(\hat{p}, P_{\mathrm{II}}^{\mathrm{S}*}\right)$. Intuitively, based on Rule (a) proposed in Section 6.4.2, the private sector is allowed to adopt a Pareto-improvement adjustment and, as a result, sets the monopoly toll $\bar{p}$; then, the public sector enhances the social welfare by reducing the monopoly toll $\bar{p}$ to $\hat{p}$ and compensating the private sector with corresponding subsidy $P_{\mathrm{II}}^{\mathrm{S}*}$. Therefore, $\left(\hat{p}, P_{\mathrm{II}}^{\mathrm{S}*}\right)$ is the adjusted contract adopted by both parties according to Rules (a) and (b).

Similarly, $\left(\hat{p}, P_{\mathrm{II}}^{\mathrm{S}}\right)$, with $P_{\mathrm{II}}^{\mathrm{S}} = T_{\mathrm{I}}\left(R\left(p_{\mathrm{I}}, y_{\mathrm{I}}, \theta\right) - R\left(\hat{p}, y_{\mathrm{I}}, \theta\right)\right)$, and $(\tilde{p}, 0)$ dominate all other feasible solutions of problem (5.14)-(5.15) with preference (5.18) when $\theta \in \Theta_2$ or $\theta \in \Theta_4$, respectively. However, when $\theta \in \Theta_3$ or $\tilde{p} \leq p_{\mathrm{I}} \leq \hat{p}$, the private sector can not increase his/her profit without reducing the social welfare, and the public sector can not enhance the social welfare either by a subsidy schedule. In this case, neither wants to change the original BOT contract. We summarize the above analysis into the following proposition.

Proposition 5.6.1. *Given the original BOT contract* $\left(T_I, p_I, P_I^S, y_I\right)$ *and a realization* θ, $\left(p_{II}^*, P_{II}^{S*}\right)$ *is the optimal solution of problem (5.14)-(5.15) with preference (5.18), where*

$$\left(p_{II}^*, P_{II}^{S*}\right) = \begin{cases} \left(\hat{p}, T_I\left(R\left(\bar{p}, y_I, \theta\right) - R\left(\hat{p}, y_I, \theta\right)\right)\right) & \theta \in \Theta_1 \\ \left(\hat{p}, T_I\left(R\left(p_I, y_I, \theta\right) - R\left(\hat{p}, y_I, \theta\right)\right)\right) & \theta \in \Theta_2 \\ \left(p_I, 0\right) & \theta \in \Theta_3 \\ \left(\tilde{p}, 0\right) & \theta \in \Theta_4 \end{cases} \tag{5.54}$$

and Θ_i*,* $i \in \{1,2,3,4\}$ *are defined by Eq. (5.53).*

Proposition 5.6.1 states the preferred Pareto-improvement within each of the four domains adopted by both parties according to the predefined Rules (a) and (b). In domain Θ_1, the realized demand is too low, and the original toll is set even higher than the monopoly price. According to Rule (a), the private sector can adjust the initial toll to the monopoly price, which is Pareto-improving since the adjustment enhances the social welfare as well. Obviously, according to Rule (b), the public sector can still improve the social welfare by reducing the toll charge but compensating the private sector up to the level of the potential monopoly profit. In domain Θ_2, there is no Pareto-improving adjustment for the private sector to increase its profit, while the public sector can improve the social welfare by reducing the toll charge and compensating the private sector the loss from the current profit. In domain Θ_4, the realized demand is very high, according to Rule (a), the private sector can increase the toll to the first-best level, which improves both profit and social welfare. However, the private sector is not allowed to further increase toll from the first-best level according to the Rule (a) by which the original toll should be altered as little as possible. In domain Θ_3, it is impossible to enhance both parties' interests by adjusting tolls.

We now move on to investigate the optimality of the original BOT contract, $\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)$. First, the duration of the concession period is a key variable. Note that, after the concession period, the public sector can set a first-best toll and achieve the maximal social welfare at any demand state. This means that the shorter the concession period, the more room for social welfare maximization during the post-concession period. However, given the MARR constraint, a longer duration of the concession period would allow the public sector to set a low original toll and to achieve more benefit by ex post adjusting the toll. The following result shows the optimal strategy of the public sector to select the duration of the concession period.

Proposition 5.6.2. *At the optimum of the BOT problem with partial flexibility, the concession period must be set to be the lifetime of the road if the toll is the only flexible* *instrument.*

Proof. Suppose $\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)$ is the optimal original contract in the first stage. If $T_{\mathrm{I}} < \hat{T}$, extending the duration of concession period by ΔT, notably, there exists, p with $p < p_{\mathrm{I}}$, such that $(T_{\mathrm{I}} + \Delta T)\,\mathrm{E}\left(R\left(p, y_{\mathrm{I}}, \theta\right)\right) = (1+r)ky_{\mathrm{I}}$. Since $p < p_{\mathrm{I}}$, we know that for any state θ, $q\left(\theta\right) > q_{\mathrm{I}}\left(\theta\right)$, where $q\left(\theta\right)$ and $q_{\mathrm{I}}\left(\theta\right)$ are the demands corresponding to p and p_{I}, respectively. Therefore, $q\left(\theta\right)$ can be denoted as a convex combination of $q_{\mathrm{I}}\left(\theta\right)$ and $\tilde{q}\left(\theta\right)$, $\tilde{q}\left(\theta\right)$ is the first-best demand at state θ, or $q\left(\theta\right) = \alpha\left(\theta\right)\tilde{q}\left(\theta\right) + \left(1-\alpha\left(\theta\right)\right)q_{\mathrm{I}}\left(\theta\right)$, $\alpha\left(\theta\right) \in (0,1)$. Note that, the lower the original toll, the more the benefit obtained by the public sector in the second stage. Since both unit-time social surplus, S, and revenue, R, are strictly concave in demand q, for each demand state, θ, we have

$$\begin{aligned}
&(T_{\mathrm{I}}+\Delta T)\,\mathrm{E}\left(S\left(p, y_{\mathrm{I}}, \theta\right)\right) + \left(\hat{T} - T_{\mathrm{I}} - \Delta T\right)\mathrm{E}\left(S\left(\tilde{p}, y_{\mathrm{I}}, \theta\right)\right) \\
&= (T_{\mathrm{I}}+\Delta T)\,\mathrm{E}\left(S\left(q\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) + \left(\hat{T} - T_{\mathrm{I}} - \Delta T\right)\mathrm{E}\left(S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) \\
&> (T_{\mathrm{I}}+\Delta T)\,\mathrm{E}\left(\left(1-\alpha\left(\theta\right)\right)S\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right) + \alpha\left(\theta\right)S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) \\
&\quad + \left(\hat{T} - T_{\mathrm{I}} - \Delta T\right)\mathrm{E}\left(S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) \\
&= T_{\mathrm{I}}\mathrm{E}\left(S\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) + \left(\hat{T} - T_{\mathrm{I}}\right)\mathrm{E}\left(S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) \\
&\quad + \mathrm{E}\left(\left(T_{\mathrm{I}}\alpha\left(\theta\right) - \Delta T\left(1-\alpha\left(\theta\right)\right)\right)\left(S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right) - S\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right)\right).
\end{aligned} \tag{5.55}$$

and

$$\begin{aligned}
&(T_{\mathrm{I}}+\Delta T)\,\mathrm{E}\left(R\left(p, y_{\mathrm{I}}, \theta\right)\right) = (T_{\mathrm{I}}+\Delta T)\,\mathrm{E}\left(R\left(p, y_{\mathrm{I}}, \theta\right)\right) \\
&> (T_{\mathrm{I}}+\Delta T)\,\mathrm{E}\left(\left(1-\alpha\left(\theta\right)\right)S\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right) + \alpha\left(\theta\right)S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) \\
&= T_{\mathrm{I}}\mathrm{E}\left(R\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) + \Delta T\mathrm{E}\left(R\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right) \\
&\quad - \mathrm{E}\left(\left(T_{\mathrm{I}}+\Delta T\right)\alpha\left(\theta\right)\left(R\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right) - R\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right)\right).
\end{aligned} \tag{5.56}$$

If the following condition is true,

$$\begin{aligned}
&\mathrm{E}\left(\alpha\left(\theta\right)\frac{R\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right) - R\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right)}{\mathrm{E}\left(R\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right)}\right) < \frac{\Delta T}{T_{\mathrm{I}}+\Delta T} \\
&\quad < \mathrm{E}\left(\alpha\left(\theta\right)\frac{S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right) - S\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right)}{\mathrm{E}\left(S\left(\tilde{q}\left(\theta\right), y_{\mathrm{I}}, \theta\right) - S\left(q_{\mathrm{I}}\left(\theta\right), y_{\mathrm{I}}, \theta\right)\right)}\right).
\end{aligned} \tag{5.57}$$

or, equivalently,

$$\mathrm{E}\left(\alpha\left(\theta\right)\frac{\left(\tilde{S}-S\right)\mathrm{E}\left(R-\tilde{R}\right)-\left(R-\tilde{R}\right)\mathrm{E}\left(\tilde{S}-S\right)+\left(\tilde{S}-S\right)\mathrm{E}\left(\tilde{R}\right)}{\mathrm{E}\left(\tilde{S}-S\right)\mathrm{E}\left(R\right)}\right)>0, \quad (5.58)$$

then, there exist p and ΔT such that inequalities (5.55) and (5.56) are satisfied simultaneously, which contradicts the optimality of $\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)$. It should be pointed out that inequalities (5.55) and (5.56) exclude term $P_{\mathrm{I}}^{\mathrm{S}}$, because $P_{\mathrm{I}}^{\mathrm{S}}$ is constant after selection of the contract $\left(T_{\mathrm{I}}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)$ and our proof here aims to construct a contradiction whenever $T_{\mathrm{I}} < \hat{T}$. In fact, it is clear that, when $p = p_{\mathrm{I}}$, $\alpha\left(\theta\right) \equiv 0$ for any θ; when p approaches zero, $\alpha\left(\theta\right) \to 1$ for any θ and inequality (5.58) is satisfied. Since we can continuously decrease p, there must exist p satisfying inequalities (5.58). Furthermore, since $\Delta T/(T_{\mathrm{I}} + \Delta T) \to 0$ with $\Delta T \to 0$, we can choose ΔT such that the increase in the subsidy size, caused by the extension of the concession length in the second stage, is small enough in the sense of expectation. We can do this because the selection of p depends on the concavities of the unit-time social surplus and revenue functions, while the extension of the concession length can be freely chosen whenever a space caused by the adjustment of toll occurs. The proof is completed. □

Proposition 5.6.2 shows an ideal result that the public sector can freely select the duration of the concession period without considering the discount rate or the minimum attractive rate of return. Generally, if the concession period is set too long, the potential profits beyond that period must be so heavily discounted to the present value that they are become negligible. Therefore, there would be a tradeoff when selecting the ex ante selection of the concession length and would not be set to be truly equal to the whole lifetime.

Based on Proposition 5.6.2 and the adjustment method, for any original contract $\left(\hat{T}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}\right)$, the expected maximum social welfare can be expressed as:

$$\begin{aligned}&\mathrm{E}\left(W\left(\hat{T}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right)\right)\\ &=\int_{\Theta_1\bigcup\Theta_2\bigcup\Theta_3\bigcup\Theta_4} W\left(\hat{T}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right)\mathrm{d}F\left(\theta\right)\end{aligned} \quad (5.59)$$

Note that $\tilde{p}$, $\hat{p}$ and $\bar{p}$ are functions of y_{I} and θ, with Θ_i, i =1,2,3,4, are functions of y_{I} and p_{I}. Suppose $\mathrm{E}\left(W\left(\hat{T}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta\right)\right)$ is differentiable in p. A revised version of condition of the deterministic case, Eq. (5.38), for adopting a positive subsidy at optimum capacity is

$$\begin{aligned}\int_{\Theta_3} \left(S_p\left(p_{\mathrm{I}}, y_{\mathrm{I}}, \theta\right) + (\lambda - 1) R_p\left(p_{\mathrm{I}}, y_{\mathrm{I}}, \theta\right)\right) \mathrm{d}F(\theta) + \lambda \int_{\Theta_1} R_p\left(p_{\mathrm{I}}, y_{\mathrm{I}}, \theta\right) \mathrm{d}F(\theta) \\ + \int_{\Theta_2} R_p\left(p_{\mathrm{I}}, y_{\mathrm{I}}, \theta\right) \mathrm{d}F(\theta) + (\lambda - 1) \int_{\Theta_4} R_p\left(p_{\mathrm{I}}, y_{\mathrm{I}}, \theta\right) \mathrm{d}F(\theta) = 0.\end{aligned} \tag{5.60}$$

For the case with perfect information, at the optimum, the ex post adjustment is not needed. As a result, $\Pr\left(\Theta_3\right) = \Pr\left\{\hat{p}\left(y_{\mathrm{I}}\right)\right\} = 1$, and condition (5.60) is equivalent to (5.38) with perfect information.

In addition, according to condition (5.60), we can further determine the optimal toll, p_{I}, and the subsidy size, $P_{\mathrm{I}}^{\mathrm{S}}$, for given capacity y_{I}. Denote the original toll given by Eq. (5.60) as p_0 for given capacity y_{I}, and the corresponding expected unit-time toll revenue as R_0, R_0=$\mathrm{E}\left(R\left(p_0, y_{\mathrm{I}}, \theta\right)\right)$. If $\hat{T}R_0 < (1+r)\, k y_{\mathrm{I}}$, then the size of the ex ante subsidy is $P_{\mathrm{I}}^{\mathrm{S}} = (1+r)\, k y_{\mathrm{I}} - \hat{T}R_0$; Otherwise, the subsidy is not really needed for the original contract and condition (5.60) becomes invalid. In this case, the lower the original toll, the more the benefit to the society. Thereby, the optimal toll for given capacity is $p_{\mathrm{I}} = \arg\min\left\{p : \hat{T}\mathrm{E}\left(R\left(p, y_{\mathrm{I}}, \theta\right)\right) = (1+r)\, k y_{\mathrm{I}}\right\}$. In summary, the optimal original toll and subsidy size can be viewed as a function of capacity y_{I} and expressed as:

$$p_{\mathrm{I}} = \begin{cases} p_0 & \hat{T}R_0 < (1+\lambda)\, k y_{\mathrm{I}} \\ \arg\min\left\{p : \hat{T}\mathrm{E}\left(R\left(p, y_{\mathrm{I}}, \theta\right)\right) = (1+r)\, k y_{\mathrm{I}}\right\} & \hat{T}R_0 \geq (1+\lambda)\, k y_{\mathrm{I}} \end{cases} \tag{5.61}$$

and

$$P_{\mathrm{I}}^{\mathrm{S}} = \begin{cases} (1+r)\, k y_{\mathrm{I}} - \hat{T}R_0 & \hat{T}R_0 < (1+\lambda)\, k y_{\mathrm{I}} \\ 0 & \hat{T}R_0 \geq (1+\lambda)\, k y_{\mathrm{I}} \end{cases} \tag{5.62}$$

where p_0 is given by condition (5.60). The expected total social welfare depends solely on the choice of capacity y_{I}. Various numerical methods can be used to determine the optimal capacity level,

$$y^*{}_{\max} = \arg\max_{y_{\mathrm{I}} \geq 0} \left\{ \mathrm{E}\left(W\left(\hat{T}, p_{\mathrm{I}}, P_{\mathrm{I}}^{\mathrm{S}}, y_{\mathrm{I}}, \theta \right)\right)\right\}, \tag{5.63}$$

where p_{I} and $P_{\mathrm{I}}^{\mathrm{S}}$ are given by Eqs. (5.61)-(5.62).

5.7 Numerical Examples

Simple examples are presented in this section to elaborate the results obtained so far. The negative exponential demand function, $q = \theta \exp(-b\mu)$, is adopted, where q is the uncertain demand; μ is the full trip price (including toll and equivalent money of travel time); b is a scaling parameter with a value of $b = 0.04$, reflecting the sensitivity of demand to the full trip price; θ is the random potential demand following a known continuous distribution with a support $\Theta = (0, +\infty)$ and can be observed after the highway is built. It is assumed that $\theta = 10^4 x$ (veh/hr), where x is a random variable following a Weibull distribution. The following Bureau of Public Roads (BPR)-type link travel time function, $t(q, y) = t_0 \left(1.0 + 0.15\left(\frac{q}{y}\right)^4\right)$, is adopted, where the free-flow travel time for the new highway is $t_0 = 0.5$(h). The construction cost functions is $I(y) = ky$ with $k = 1.2 \times 10^6 \times t_0$ (HK\$/h). The marginal cost of public funds is assumed to be $\lambda = 1.05$. These values are selected for illustrative purpose without necessarily representing realistically reasonable values.

For the benchmark case in the absence of demand uncertainty, $\Pr\left\{\theta = 10^4\right\} = 1$, we have $r_0 = 18.8\%$ by direct calibration. If the MARR is less than 18.8%, then the subsidy is not needed at an optimal BOT contract; otherwise, adopting a combination of the toll revenues and subsidies will be better than using the toll revenue alone in the sense of maximizing social welfare. As depicted in Figure 5.1, Let the MARR $r = 15\%$, then the optimal capacity and toll are $y^* = 1357$ (veh/h) and $p^* = 7.65$ (HK\$), respectively. For the selected MARR, $r = 15\%$, no subsidy is needed, namely, $P^{\mathrm{S}*} = 0$. We now go to consider the BOT contract with full flexibility. For any ex ante determined capacity y_{I},

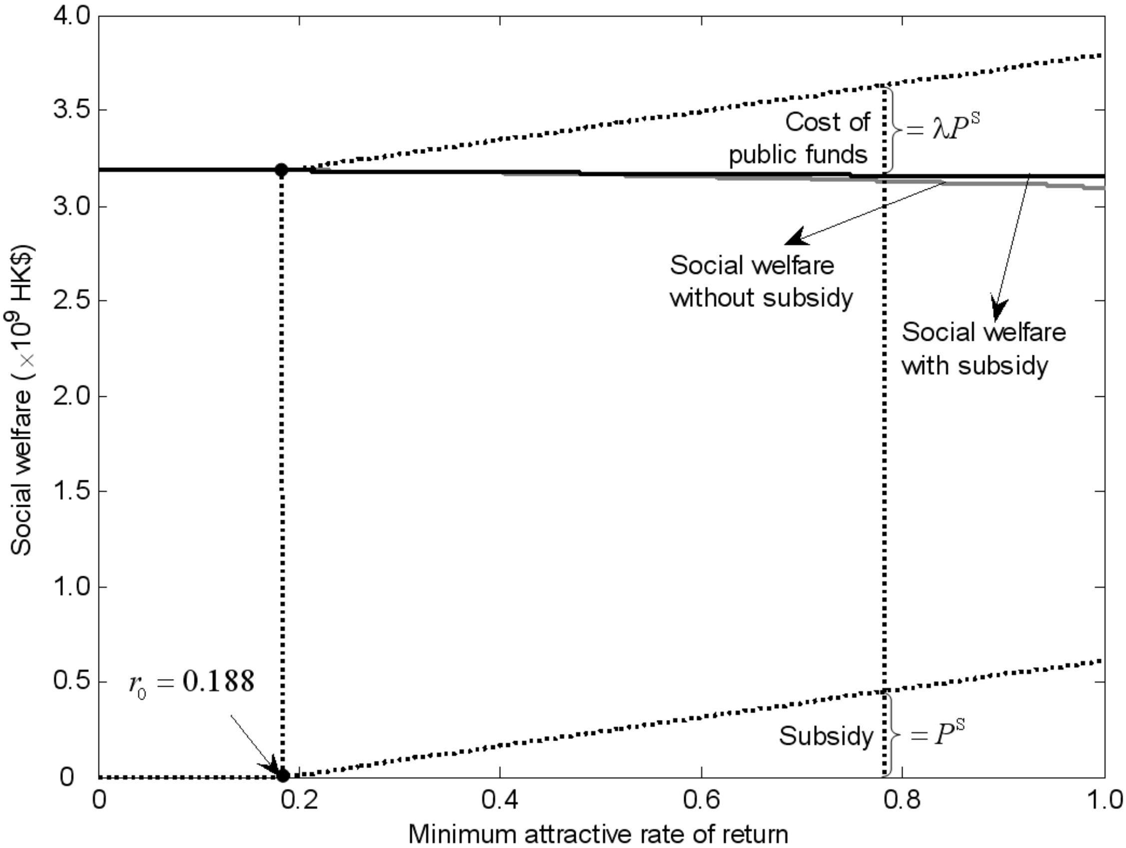


Figure 5.1: Welfare effects of Subsidy with various MARR (MCPF=1.05)

from definitions (5.42)-(5.43), by direct calculations, we have

$$\bar{\theta}\left(y_{\mathrm{I}}\right)=\frac{y_{\mathrm{I}}\tilde{\gamma}}{\exp\left(-b\beta\left(t\left(\tilde{\gamma}\right)+\tilde{\gamma}t'\left(\tilde{\gamma}\right)\right)\right)}$$

and

$$\underline{\theta}\left(y_{\mathrm{I}}\right)=\frac{y_{\mathrm{I}}\hat{\gamma}}{\exp\left(-b\beta\left(t\left(\hat{\gamma}\right)+\hat{\gamma}t'\left(\hat{\gamma}\right)\right)-\left(1-\frac{1}{\lambda}\right)\right)},$$

where $\tilde{\gamma}$ and $\hat{\gamma}$ are constants, respectively, given by

$$\beta(\tilde{\gamma})^2t'\left(\tilde{\gamma}\right)=\frac{\left(1+r\right)k}{\hat{T}}$$

and

$$\beta(\hat{\gamma})^2t'\left(\hat{\gamma}\right)+\left(1-\frac{1}{\lambda}\right)\frac{\hat{\gamma}}{b}=\frac{\left(1+r\right)k}{\hat{T}}.$$

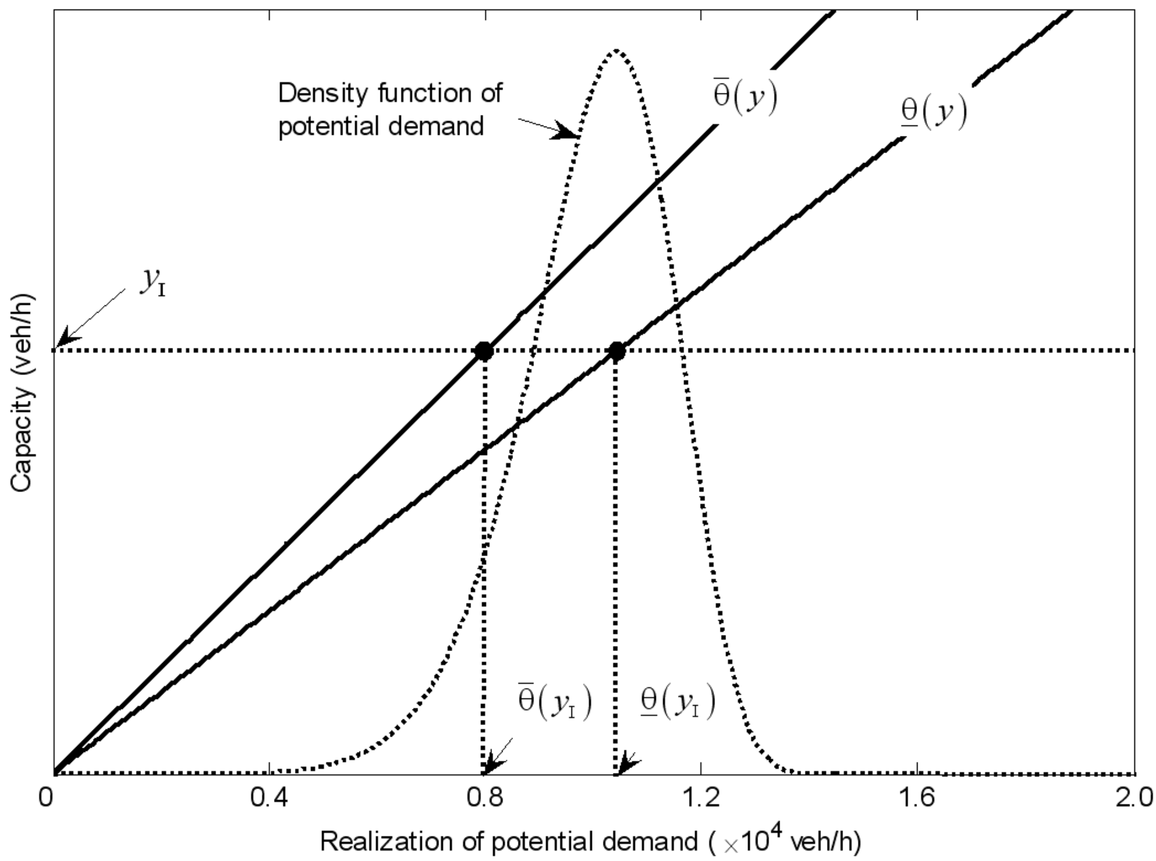


Figure 5.2: Determination of low, intermediate and high demand states with full flexibility

Figure 5.2 shows determination of the low, intermediate and high demand states for any given capacity $y_{\rm I}$. For the negative exponential demand function, the infimum and supremum of the high and low demand states, $\bar{\theta}(y_{\rm I})$ and $\underline{\theta}(y_{\rm I})$, are linear functions of capacity. The partition presented in Lemma 5.5.1 is clearly depicted in Figure 5.2 when the prior density function of θ is known to both parties. For MCPF, $\lambda = 1.05$, the gap between $\bar{\theta}(y_{\rm I})$ and $\underline{\theta}(y_{\rm I})$ is very small. To highlight the gap, a value of $\lambda = 1.5$ is taken in Figure 5.2.

To investigate the effects of the prior information on optimal capacity choices, we select the random variable, x, following Weibull distributions with the same mean value of 1.0, but various variances increasing from 0.02 to 0.32 with a step 0.03. Figure 5.3 shows the change of the expected social welfare with the ex ante capacity choice for different variances. The optimal capacity level under certainty increases from 1376 to 1448 (veh/h), higher than the deterministic optimal capacity of 1357 (veh/h). The corresponding expected total

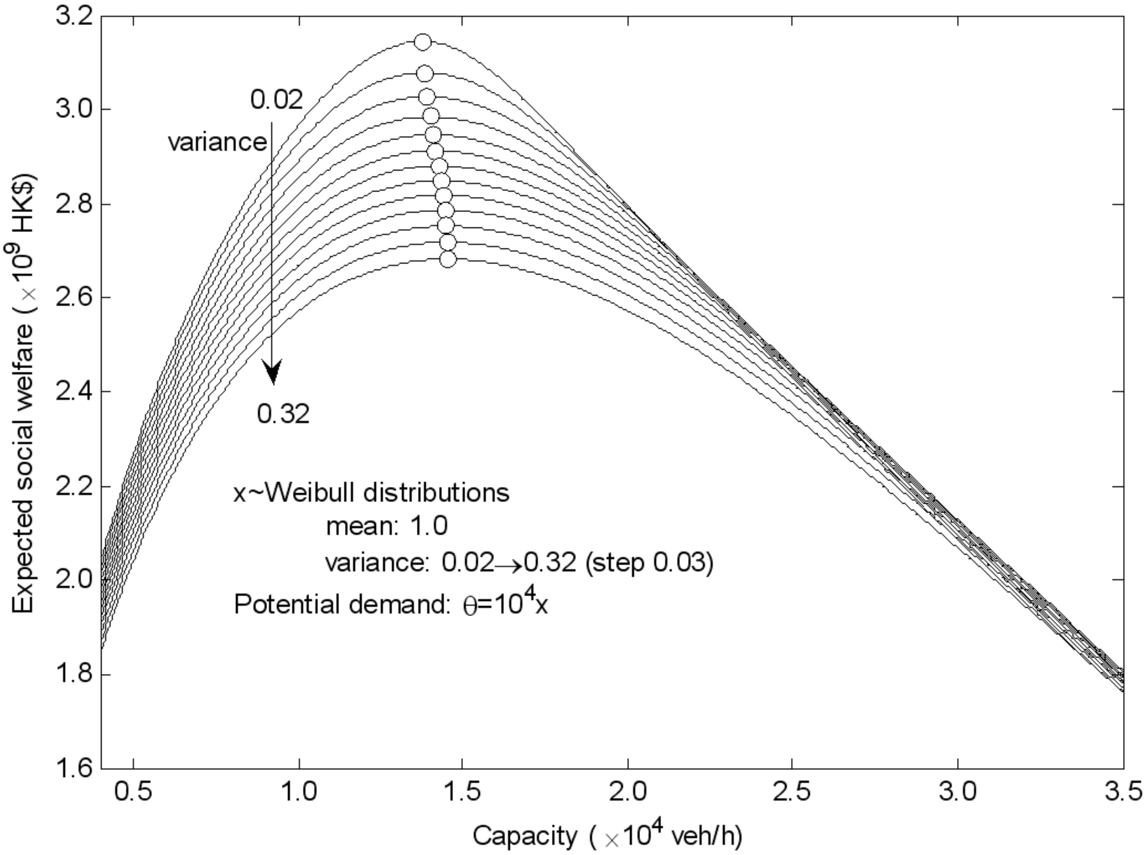


Figure 5.3: Effects of information on optimal capacity choice and expected social welfare with full flexibility

social welfare decreases from 3.145×10^9 to 2.753×10^9 (HK\$). Clearly, the larger the variability of the potential demand θ, the higher capacity the public sector to choose. Also, the expected social welfare decreases with the variability of the potential demand.

For the BOT contract with partial flexibility, we first explain how to partition the demand state space proposed in Proposition 5.6.1. Note that the original capacity y_{I} and toll p_{I} together determine the partition (5.53). We only need calculate θ_1, θ_2 and θ_3 from $y - \mathrm{I}$ and p_{I}, where $\theta_1 = \Theta_1 \bigcap \Theta_2$, $\theta_2 = \Theta_2 \bigcap \Theta_3$ and $\theta_3 = \Theta_3 \bigcap \Theta_4$, since parameter θ corresponds to the demand level (the larger the value θ, the higher the demand level for any given toll charge). By direct calculation, we have $\theta_1 (p_{\mathrm{I}}, y_{\mathrm{I}}) = 0$ if $p_{\mathrm{I}} \leq 1/b$ and $\theta_1 (p_{\mathrm{I}}, y_{\mathrm{I}}) = y_{\mathrm{I}} \bar{\gamma} \exp (b (p_{\mathrm{I}} + \beta t (\bar{\gamma})))$ if $p_{\mathrm{I}} > 1/b$; if $p_{\mathrm{I}} \leq (1 - 1/\lambda)/b$ and $\theta_2 (p_{\mathrm{I}}, y_{\mathrm{I}}) = y_{\mathrm{I}} \hat{\gamma} \exp (b (p_{\mathrm{I}} + \beta t (\hat{\gamma})))$ if $p_{\mathrm{I}} > (1 - 1/\lambda)/b$; $\theta_3 (p_{\mathrm{I}}, y_{\mathrm{I}}) = y_{\mathrm{I}} \tilde{\gamma} \exp (b (p_{\mathrm{I}} + \beta t (\tilde{\gamma})))$, where the corresponding v/c ratio can be determined by $\beta \bar{\gamma} t' (\bar{\gamma}) = p_{\mathrm{I}} - 1/b$, $\beta \hat{\gamma} t' (\hat{\gamma}) = p_{\mathrm{I}} -$

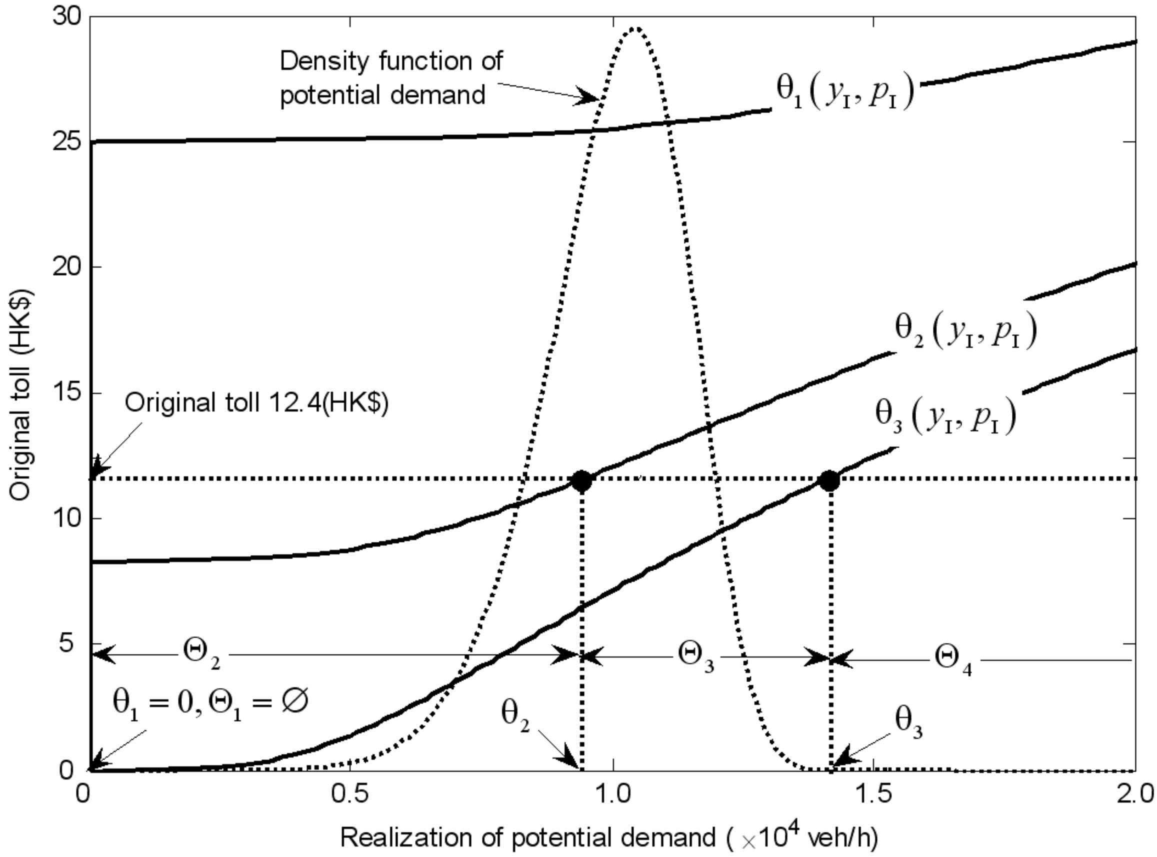


Figure 5.4: Partition of demand state space with partial flexibility ($y_I = 1357$ veh/h)

$(1 - 1/\lambda)/b$ and $\beta\tilde{\gamma}t'(\tilde{\gamma}) = p_I$. Figure 5.4 depicts the curves of $\theta_1(p_I, y_I)$, $\theta_2(p_I, y_I)$ and $\theta_3(p_I, y_I)$, which depend only on the original toll p_I, since the original capacity level is given and fixed at $y_I = 1357$ (veh/h). Note that, to highlight the differences among $\theta_1(p_I, y_I)$, $\theta_2(p_I, y_I)$ and $\theta_3(p_I, y_I)$, we take a value of $\lambda = 1.5$. For any given original toll, the unique partition of the demand state space can thus be determined. Figure 5.4 also shows a partition of the state space with respect to the original toll $p_I = 12.4$ (HK$). For any given y_I, we can solve for the corresponding toll, p , from the integral equation (5.60). By combining conditions (5.61) and (5.62), the optimal original toll and the corresponding subsidy size can be determined. Figure 5.5 shows a comparison of the outcomes (expected social welfare and profit) with partial flexibility and without flexibility. In the calculation, the random variable, x, is assumed to follow the Weibull distribution with a mean value 1.0 and a variance 0.02. The Pareto-improving adjustment of the toll does enhance both parties' interests than those without contract flexibility at any capacity level.

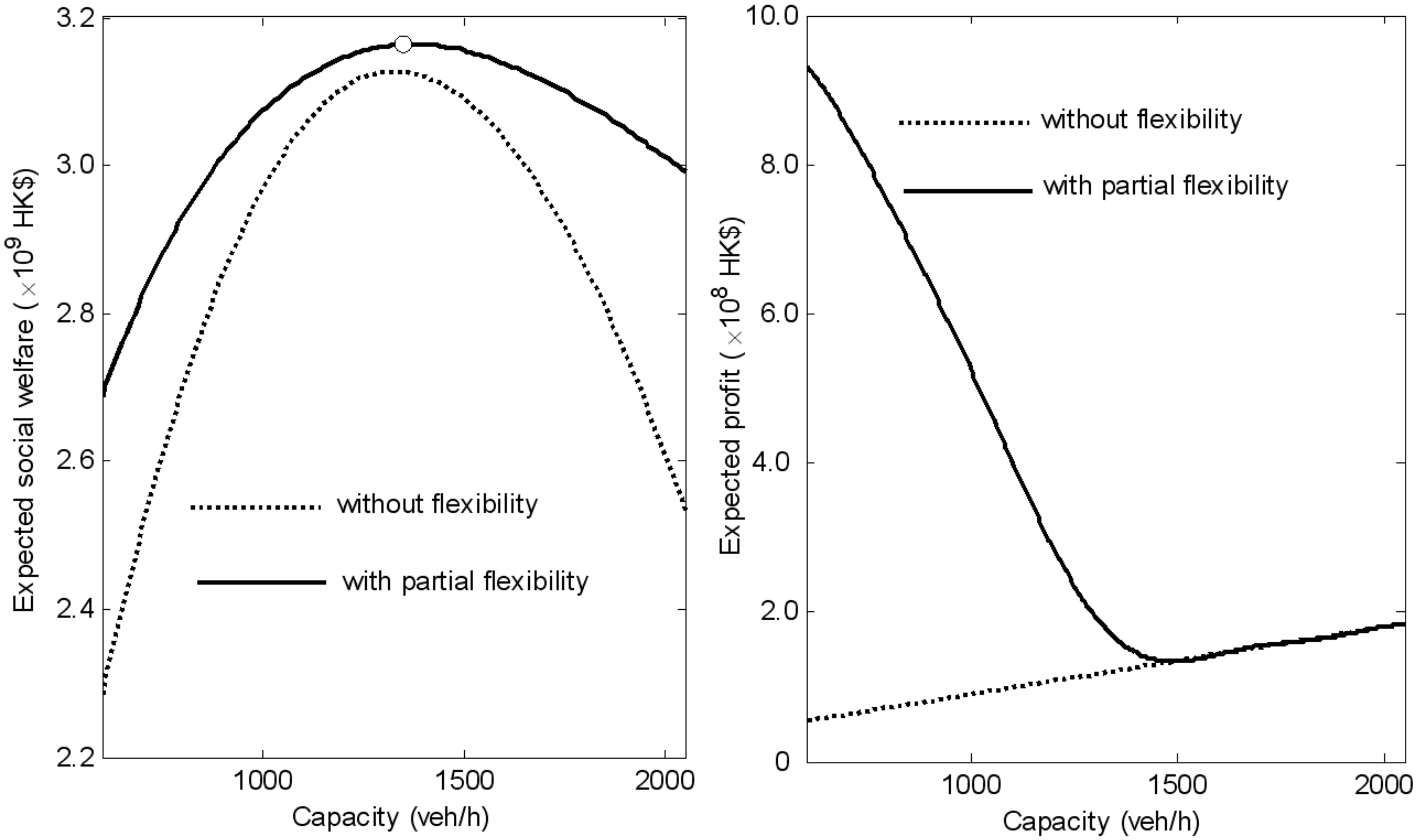


Figure 5.5: Outcomes without and with partial flexibility varying with capacity level

5.8 Conclusions

To deal with traffic risk pertaining to highway franchising contracts involving concession period, road capacity and toll level, this chapter investigated two types of flexible contracts between the public and private sectors under demand uncertainty, by assuming that the uncertain demand curve is ex post observed, and thus the contract variables are allowed to be ex post adjusted according to certain ex ante agreed Rules

For the benchmark case with perfect information, we proved that the subsidy is not always bad for the optimal capacity and toll (or demand) choices in the sense of maximizing social welfare. The optimal subsidy size depends on the MARR on the investment of the private sector. Under certain assumptions, a unique critical MARR is derived. At optimum, the subsidy is not needed if the MARR does not exceed the critical value; otherwise, a positive subsidy is desirable to increase the total weighted social welfare.

In the presence of demand uncertainty, the public sector guarantees the return rate of the private sector and bears all the demand risk, and in return has full flexibility to determine the optimal toll, the duration of the concession period, and the corresponding optimal subsidy size after the demand curve is observed. For any predetermined capacity level, the demand state space can be partitioned as high, intermediate and low domains. In the high demand domain, the public sector should set a first-best toll, a proper duration of the concession period and does not need to introduce subsidy; in the low demand domain, the public sector should set a concession period of the road life and, the subsidy size is the gap between the required profit and the total toll revenue, furthermore, at optimum, the ratio of the marginal consumer surplus to the marginal revenue exactly equals the marginal cost of public funds; in the intermediate demand domain, the public sector should set a concession period equal to the road life and a toll such that the total toll revenue just covers the investment cost plus the required profit of the private sector.

We also investigated the BOT problem with partial flexibility by assuming that the toll is the only flexible instrument. In this case, both parties share the demand risk. We proposed a bi-objective programming model to select a Pareto-improvement for the ex post contract adjustment from the Pareto-optimal solution set equipped with a rational preference relation. Under the preference structure of the Pareto-improvement set, a unique Pareto-improvement can be selected by the bi-objective programming model for any given original contract. We also proved that, for the original contract, the concession period should be set to equal the road life if the toll is the only flexible adjustment instrument.

6

OPTIMAL BOT CONTRACTS WITH ROAD DAMAGE AND MAINTENANCE COST

In the analysis of highway franchising project, the operation and maintenance costs are either neglected or assumed fixed over time in the previous chapters. This chapter models the BOT problem as the isoperimetric problem in calculus of variations to maximize the social welfare with a profit constraint. The model explicitly incorporates the effect of road deterioration and maintenance over the years, which is assumed to depend on the traffic loads, road capacity and road natural deterioration. We find that an optimal pricing policy requires toll increase over calendar time to reduce traffic load due to time-increasing and load-increasing maintenance cost. If, however, the marginal user damage on road is independent of time, then the optimal toll charge is free from the effect of road natural deterioration and thus time-invariant. We also discuss how to reach an optimal contract through government regulations and investigate the effects of economic growth on the solution properties of the problem.

6.1 Introduction

In the analysis of highway franchising project, the operation and maintenance costs are either neglected or assumed fixed over time in the previous chapters and many other literature (Engel et al., 2001; Guo and Yang, 2010b; Nombela and de Rus, 2004). The maintenance cost should not be ignored in a BOT contract. Figure. 6.1 shows a statistic analysis

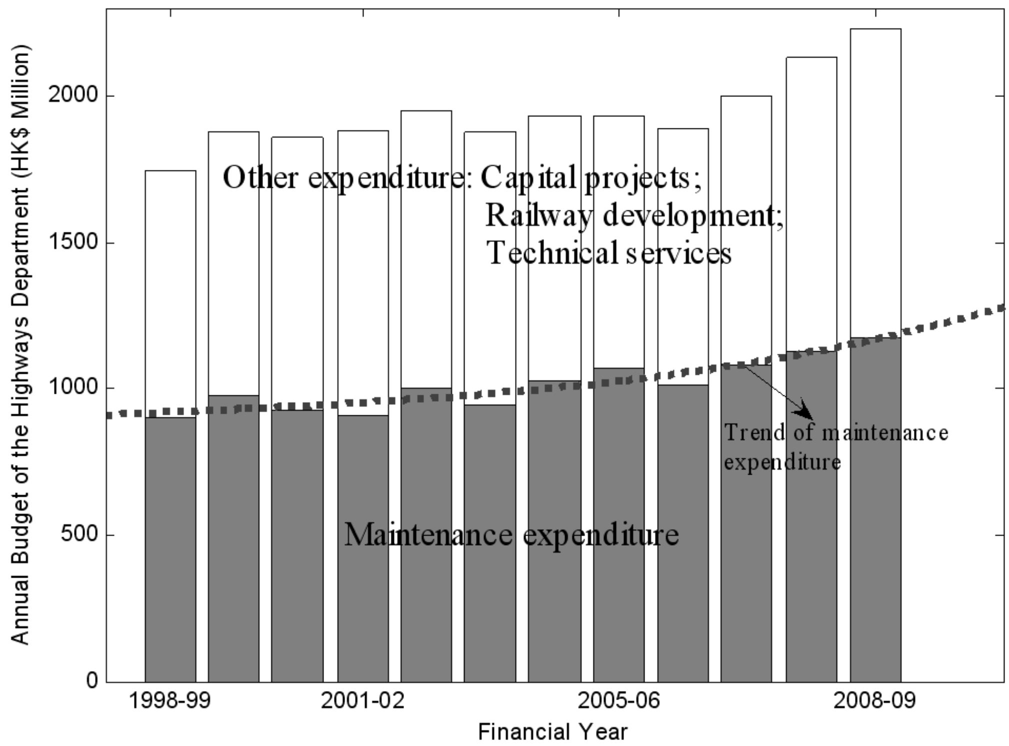


Figure 6.1: Annual maintenance expenditure of public roads in Hong Kong

of the road network maintenance costs in Hong Kong (http://www.budget.gov.hk). A significant increasing ratio of the maintenance costs to the total highway budget is observed. The technical report of the life-cycle cost analysis for the highway investment, published by FHWA (2002), pointed out that the efficiency of rehabilitation will decrease, while the social cost (user cost plus the agency cost) will increase because of the road deterioration. Small et al. (1989) claimed that the road wear caused by heavy loads is exacerbated by time and weather. Newbery (1988, 1989) studied road maintenance and damage using the roughness model of Paterson (1987). De Palma et al. (2007a) investigated maintenance and tolling decisions under the environment with two competing private roads. In their model, the road quality was assumed to be sustained by maintenance intensity, thus the cost of maintenance was determined by the road capacity and the maintenance intensity. However, even if the traffic load is fixed; the road would always deteriorate with the effects of time and weather (Paterson, 1987; Small et al., 1989). It is more practical to assume that the yearly maintenance cost increases over the years due to road deterioration.

In line with the study of Guo and Yang (2010b), we consider all the key variables of concession period, road capacity and toll charge in forming a BOT contract for a congestible highway with the road deterioration and maintenance effects. We assume that the maintenance cost to sustain the road quality is a function of road capacity, traffic load and time, and we can reasonably assume that the life time of the road can be determined endogenously as the time when the net economic benefit becomes nil as maintenance cost increases. Firstly, the unconstrained social-welfare maximizing and unconstrained profit maximizing problems can be viewed as the functional extreme problems in calculus of variations. The BOT contract is modeled as an isoperimetric problem to maximize the social welfare within the concession period with a profit constraint. The Euler-Lagrange equation is introduced to obtain the optimal toll policy, which can be expressed by a differentiable equation. We go further to analyze the properties of the optimal dynamic pricing policy under a steady economic environment in the sense that the travel demand depends only on generalized travel cost, and then in a growing economic environment in the sense that the travel demand can be viewed as a function of the generalized travel cost and the calendar time. Two strategies are proposed for the government to obtain the optimal BOT contract: full regulation and partial regulation. For the latter, with increasing maintenance cost over time, the government may tend to set a concession period longer than desired by the private firm, and set a time-varying traffic load such that the traffic load is restricted to a level larger than desired by the private firm in any time.

The chapter is organized as follows: we begin with some basic assumptions and introduce definition of road life and its property in the next section. In Sections 6.3 and 6.4 we formulate the unconstrained social-welfare maximizing and profit maximizing problems, respectively. The BOT problem is modeled and analyzed by applying the functional extreme theory in Section 6.5, in which the regulation of the government is investigated. Section 6.6 looks into the effects of the economic growth on the solution properties of the problems proposed in earlier sections. Finally, Conclusions are given in Section 6.7.

6.2 Basic Assumptions and Road Life

6.2.1 Basic assumptions

Assume that a single highway connecting two cities is to be built. The capacity of the new road $y \geq 0$ would not be changed along the calendar time t and its length is normalized to 1. We consider a steady economic environment where the travel demand depends only on the generalized travel cost (price elastic). For any time interval $(t, t + \mathrm{d}t)$ the following demand-supply equilibrium condition is assumed to always hold:

$$B\left(q\left(t\right)\right) = p\left(t\right) + \beta C\left(q\left(t\right), y\right), \tag{6.1}$$

where $q\left(t\right)$ is the equilibrium travel demand at time t; $B\left(\cdot\right)$ is the inverse demand function (or the willing-to-pay function), and its shape and location will not change along t (this assumption will be relaxed by considering a growing economic environment or the inverse demand function shift with time t); and $C\left(q\left(t\right), y\right)$ is the travel time function whose arguments include travel demand, $q\left(t\right)$ and road capacity, y; $p\left(t\right)$ is the toll charged on each user of the road at time t and β is the value-of-time (VOT) to convert time into equivalent monetary cost (we consider homogeneous users only). Condition (6.1) means that the equilibrium travel demand at each time interval (year or month) for the new road is completely determined by the generalized travel cost. Let $I\left(y\right)$ be the road construction cost function. In this chapter, Assumption 3.1 is also made on functions $B\left(q\right)$, $I\left(y\right)$ and $C\left(q, y\right)$.

The consumer surplus, S, measured in monetary unit during the interval, $(t, t + \mathrm{d}t)$, can be expressed as

$$S\left(q\left(t\right), y\right) \mathrm{d}t = \left(\int_0^{q(t)} B\left(\omega\right) \mathrm{d}\omega - \beta q\left(t\right) C\left(q\left(t\right), y\right)\right) \mathrm{d}t. \tag{6.2}$$

Under Assumption 3.1, we can readily observe that $S(q, y)$ is strictly concave in q, namely,

$$\frac{\partial^2 S(q, y)}{\partial q^2} = B'(q) - 2\beta \frac{\partial C(q, y)}{\partial q} - \beta q \frac{\partial^2 C(q, y)}{\partial q^2} < 0. \tag{6.3}$$

Correspondingly, the toll revenue, R, during the interval $(t, t + \mathrm{d}t)$ can be calculated using the following formulation by Eq. (6.1):

$$R(q(t), y)\,\mathrm{d}t = p(t)\, q(t)\,\mathrm{d}t = (q(t)\, B(q(t)) - \beta q(t)\, C(q(t), y))\,\mathrm{d}t. \tag{6.4}$$

Assumption 6.1. *For any given $y > 0$, the revenue function $R(q, y)$ is a strictly concave function of q for $q \geq 0$, i.e., $\partial^2 R/\partial q^2 < 0$*

With part (c) of Assumption 3.1, $C(q, y)$ is strictly increasing and convex and thus $qC(q, y)$ is strictly convex in q for any given $y > 0$, which means that the term, $-\beta qC(q, y)$ of $R(q, y)$ in (6.4) is strictly concave in q for any given $y > 0$. Thus $R(q, y)$ is strictly concave if the first term $qB(q)$ is concave in q. In other words, with Assumption 3.1, Assumption 6.1 holds if $qB(q)$ is concave. Indeed, in the literature it is common to assume that $qB(q)$ is concave. Thus Assumption 6.1 is not restrictive.

Like de Palma et al. (2007a), we assume that the road quality including various road characteristics can be sustained by the maintenance activity. We further assume that the maintenance cost, M, to sustain the initial road quality during the interval $(t, t + \mathrm{d}t)$ is a function of traffic load, $q(t)$, capacity, y, and time, t. This function is expressed as:

$$M\mathrm{d}t = M(q(t), y, t)\,\mathrm{d}t. \tag{6.5}$$

The following assumption is needed on the maintenance function $M(q, y, t)$.

Assumption 6.2. *$M(q, y, t)$ is strictly increasing in q, y and t, separately, and twice differentiable with the properties: $\partial^2 M/\partial q^2 \geq 0$; $\partial^2 M/\partial q \partial t \geq 0$; for any $q > 0$ and $y > 0$, $\lim_{t\to\infty} M(q, y, t) = +\infty$.*

Where $\partial^2 M/\partial q^2 \geq 0$ and $\partial^2 M/\partial q \partial t \geq 0$ mean that an aging road is more susceptible to the damage of traffic load and more expensive to maintain the same road quality.

6.2.2 Road life and its properties

Since the road wear caused by heavy loads is exacerbated by time and weather, or equivalently, the maintenance cost increases with time and quickened dramatically by the damage of users, it is reasonable to assume that the road life is limited. We define the road life as the longest time during which operating the road can bring a positive social welfare gain. In other words, beyond the road life, operating the road would always bring a negative social welfare gain no matter what number of users is using the road. A rigorous mathematical definition is given below:

Definition 6.1. *The road life, $\tilde{T}$, for a given road capacity, y, is given as:*

$$\tilde{T} = \max\left\{T \geq 0 : S(q,y) \geq M(q,y,T), \exists\, q > 0\right\}, \tag{6.6}$$

where $S(q,y)$ is the consumer surplus defined by Eq. (6.2).

Definition 6.1 implies that the road should be closed or rebuilt after $\tilde{T}$ since social welfare gain becomes negative for whatever positive traffic load. By Assumption 6.1, it is easy to show that the road life $\tilde{T}$ is uniquely determined in terms of road capacity only. By the definition, it is always beneficial for the government to operate a road till the road life. For given $\tilde{T}$, there is $\tilde{q} > 0$ such that

$$S(\tilde{q},y) \geq M\left(\tilde{q},y,\tilde{T}\right). \tag{6.7}$$

From the continuity of $M(q,y,\cdot)$ and the definition of $\tilde{T}$, we also have

$$S(q,y) \leq M\left(q,y,\tilde{T}\right), \forall q > 0. \tag{6.8}$$

The reason is that, if condition (6.8) is violated, then prolonging operating time still produces a positive social welfare gain, and thus $\tilde{T}$ is not the road life for a given road capacity y. Equations (6.7) and (6.8), in fact, reveal that $\tilde{q}_{\tilde{T}}$ maximizes function $f(\cdot) = S(\cdot, y) - M\left(\cdot, y, \tilde{T}\right)$ and $f(\tilde{q}_{\tilde{T}}) = 0$. We thus have the following property of the road life.

Lemma 6.2.1. *For any given road capacity y and with Assumption 3.1, $\tilde{T}$ is the road life if and only if there is a unique positive traffic load $\tilde{q}_{\tilde{T}}$ at time $\tilde{T}$ such that*

$$\frac{\partial S(\tilde{q}_{\tilde{T}}, y)}{\partial q} - \frac{\partial M\left(\tilde{q}_{\tilde{T}}, y, \tilde{T}\right)}{\partial q} = 0 \tag{6.9}$$

and

$$S(\tilde{q}_{\tilde{T}}, y) - M\left(\tilde{q}_{\tilde{T}}, y, \tilde{T}\right) = 0. \tag{6.10}$$

The uniqueness of $\tilde{q}_{\tilde{T}}$ in Lemma 6.2.1 steams from the fact that the function $f(\cdot) = S(\cdot, y) - M\left(\cdot, y, \tilde{T}\right)$ is strictly concave. Hereinafter, $\tilde{T}$ and $\tilde{q}_{\tilde{T}}$ always denote the road life and the corresponding traffic load, both are functions of road capacity y only.

6.3 The Unconstrained Social-Welfare Maximizing Problem

6.3.1 Problem formulation and optimality conditions

The unconstrained social welfare maximizing problem is to select the road capacity and toll charge (and thereby the resulting traffic load) to maximize the total social welfare throughout the road life. The total social welfare during the whole road life is the total economic benefit minus the total maintenance cost and the initial construction cost. The problem can be formulated as:

$$\max_{y \geq 0, q(\cdot) \in C^2\left[0, \tilde{T}\right]} W(q(\cdot), y) = \int_0^{\tilde{T}} \left(S(q(t), y) - M(q(t), y, t)\right) e^{-\lambda t} \mathrm{d}t - I(y), \tag{6.11}$$

where the first term in the integral is defined in Eq. (6.2); the traffic load, $q(\cdot)$, is assumed to be a twice differentiable function of time t; and λ is an interest rate of reference time used for discounting all future revenues and costs to their equivalent values at initial time $t = 0$.

Problem (6.11) is a functional extreme problem in terms of $q(\cdot)$ with a parameter y. Supposing $(q^*(\cdot), y^*)$ is the optimal solution of problem (6.11). From the Euler-Lagrange equation, we have:

$$\frac{\partial S\left(q^*\left((t)\right), y^*\right)}{\partial q} = \frac{\partial M\left(q^*(t), y^*, t\right)}{\partial q}, \ \forall t \in \left[0, \tilde{T}\right]. \tag{6.12}$$

Note that, different to Eq. (6.9) in Lemma 6.2.1, condition (6.12) reveals that each point on the extreme curve $q^*(\cdot)$ must maximize the function $f(\cdot) = S(\cdot, y) - M(\cdot, y, t)$. Thus, the extreme curve $q^*(\cdot)$ must satisfy the boundary condition $q\left(\tilde{T}\right) = \tilde{q}_{\tilde{T}}$, where the road life $\tilde{T}$ and traffic load $\tilde{q}_{\tilde{T}}$ are predetermined by Eqs. (6.9) and (6.10) at the corresponding road capacity y^*. The implicit function $q^*(\cdot)$ defined by condition (6.12), together with $\tilde{T}$ and $\tilde{q}_{\tilde{T}}$, completely determine the extreme curve $q^*(\cdot)$ for given y^*. Thus, we can view the extreme curve q^* as a function of time t and road capacity y^*. In addition, from the definition of social welfare in (6.1) and the equilibrium condition (6.1), condition (6.12) can be rewritten as:

$$p^*(t) = \beta q^*(t) \frac{\partial C\left(q^*(t), y^*\right)}{\partial q} + \frac{\partial M\left(q^*(t), y^*, t\right)}{\partial q}, \tag{6.13}$$

which implies that the socially optimal toll must equal the sum of the congestion externality and the marginal maintenance cost in traffic load at any time t. This result is a straightforward extension of the standard optimal pricing result with maintenance cost in a static framework (e.g., Chu and Tsai, 2004; Newbery, 1988; Yang and Huang, 2005) to the dynamic situation considered here.

It must be pointed out that the functional problem (6.11) with free horizon and free endpoint values require $f(q^*(t)) = S(q^*(t), y) - M(q^*(t), y, t) = 0$ at time $t = \tilde{T}$. Thus, Eq. (6.10) in Lemma 6.2.1, in fact, is the boundary condition of problem (6.11). To find

y^*, taking derivative of W with respect to y and setting the derivative equal to zero give rise to:

$$
\begin{aligned}
0 = \tfrac{\mathrm{d}W(q^*(\cdot),y^*)}{\mathrm{d}y} = & \left(S\left(q^*\left(\tilde{T}\right),y^*\right) - M\left(q^*\left(\tilde{T}\right),y^*,\tilde{T}\right)\right)e^{-\lambda\tilde{T}}\tfrac{\mathrm{d}\tilde{T}}{\mathrm{d}y} \\
& + \int_0^{\tilde{T}}\left(\tfrac{\partial S(q^*(t),y^*)}{\partial q} - \tfrac{\partial M(q^*(t),y^*,t)}{\partial q}\right)\tfrac{\partial q^*}{\partial y}e^{-\lambda t}\mathrm{d}t \\
& + \int_0^{\tilde{T}}\left(\tfrac{\partial S(q^*(t),y^*)}{\partial y} - \tfrac{\partial M(q^*(t),y^*,t)}{\partial y}\right)e^{-\lambda t}\mathrm{d}t - I'\left(y^*\right).
\end{aligned}
\tag{6.14}
$$

From (6.10) and (6.12), condition (6.14) can be reduced to:

$$
\int_0^{\tilde{T}}\left(\frac{\partial S\left(q^*\left(t\right),y^*\right)}{\partial y} - \frac{\partial M\left(q^*\left(t\right),y^*,t\right)}{\partial y}\right)e^{-\lambda t}\mathrm{d}t - I'\left(y^*\right) = 0. \tag{6.15}
$$

Equations (6.12) and (6.15) are the optimality conditions for problem (6.11) with respect to function $q\left(\cdot\right)$ and variable y. From Lemma 6.2.1 and condition (6.11), we have the following proposition.

Proposition 6.3.1. *The socially optimal traffic load, $q^*\left(t\right)$, is non-increasing and the socially optimal toll charge, $p^*\left(t\right)$, is non-decreasing over calendar time t.*

Proof. Consider Euler-Lagrange Eq. (6.12). From (6.3) and Assumption 6.2, we have

$$
\partial^2 S\left(q\left(t\right),y\right)/\partial q^2 - \partial^2 M\left(q\left(t\right),y,t\right)/\partial q^2 < 0. \tag{6.16}
$$

From implicit function theorem, we conclude that $q^*\left(t\right)$ is continuously differentiable. Taking derivatives of both sides of Eq. (6.12) in t yields:

$$
\left(\frac{\partial^2 S\left(q^*\left(t\right),y^*\right)}{\partial q^2} - \frac{\partial^2 M\left(q^*\left(t\right),y^*,t\right)}{\partial q^2}\right)\frac{\mathrm{d}q^*\left(t\right)}{\mathrm{d}t} - \frac{\partial^2 M\left(q^*\left(t\right),y^*,t\right)}{\partial q\partial t} = 0. \tag{6.17}
$$

From (6.16) and Assumption 6.2 again we thus have $\mathrm{d}q^*\left(t\right)/\mathrm{d}t \leq 0$ for $t \in \left[0,\tilde{T}\right]$. Furthermore, from equilibrium condition (6.1), it follows that

$$
\frac{\mathrm{d}p^*\left(t\right)}{\mathrm{d}t} = B'\left(q^*\left(t\right)\right)\frac{\mathrm{d}q^*\left(t\right)}{\mathrm{d}t} - \beta\frac{\partial C\left(q^*\left(t\right),y\right)}{\partial q}\frac{\mathrm{d}q^*\left(t\right)}{\mathrm{d}t} \geq 0. \tag{6.18}
$$

The proof is completed. □

Proposition 6.3.1 simply shows that, to maximize the total social welfare, the government should set the toll charge to be continuously increasing over time to achieve a time-decreasing traffic load, because of the time-increasing and load-increasing maintenance cost. By Definition 6.1 and Lemma 6.2.1, $q^* (\cdot)$ must satisfy the boundary condition (6.10) at the social optimum capacity $y = y^*$, this condition is explicitly denoted as:

$$q^* \left(\tilde{T}\right) = \tilde{q}_{\tilde{T}}. \tag{6.19}$$

If we further assume that, for any given y, the maintenance cost $M(q, y, t)$ can be represented as the sum of two separable terms, with one dependent on traffic load q and the other one on time t (natural degradation):

$$M(q, y, t) = M_1(q, y) + M_2(y, t), \tag{6.20}$$

then we have

$$\frac{\partial^2 M(q^*(t), y^*, t)}{\partial q \partial t} = 0, \tag{6.21}$$

which means that the marginal user damage on road is time-independent. Substituting (6.21) into (6.17), we obtain $\mathrm{d}q^*(t)/\mathrm{d}t = 0$. In view of the boundary condition (6.19), we readily have:

$$q^*(t) \equiv \tilde{q}_{\tilde{T}}, \forall t \in \left[0, \tilde{T}\right], \tag{6.22}$$

which, from (6.18), implies that the socially optimal toll charge is time-invariant. Therefore, we have the following corollary.

Corollary 6.3.1. *If the marginal user damage on road is independent of time, then the socially optimum pricing policy will not change over calendar time.*

From Eqs. (6.12), (6.15) and boundary condition (6.19), we can numerically obtain the solution of the social welfare maximizing problem (6.11). In this case both road life $\tilde{T}$ and endpoint traffic load $\tilde{q}$ are function of road capacity y, which can be predetermined from Eqs. (6.9) and (6.10) for any given y.

6.3.2 The self-financing theorem

It is assumed that the link travel time function, $C(q,y)$ is homogeneous of degree zero in both traffic flow q and road capacity y

$$C(\alpha q, \alpha y) = C(q,y), \ \forall \alpha > 0, \tag{6.23}$$

and there are constant returns to scale in road construction

$$I(\alpha y) = \alpha I(y). \tag{6.24}$$

In addition, we assume that the maintenance cost function, $M(q,y,t)$ is homogeneous of degree one with respect to demand q and capacity y for any given time $t \geq 0$

$$M(\alpha q, \alpha y, t) = \alpha M(q,y,t), \ \forall \alpha > 0. \tag{6.25}$$

From Euler equation, we can obtain the following two conditions:

$$q\frac{\partial C(q,y)}{\partial q} + y\frac{\partial C(q,y)}{\partial y} = 0, \tag{6.26}$$

and

$$q\frac{\partial M(q,y,t)}{\partial q} + y\frac{\partial M(q,y,t)}{\partial y} = M(q,y,t). \tag{6.27}$$

In addition, from Eqs. (6.2) and (6.12), for any time $t \in \left[0, \tilde{T}\right]$ we also have:

$$\begin{aligned} \tfrac{\partial S(q^*(t),y^*)}{\partial q} &= B(q^*(t)) - \beta C(q^*(t), y^*) - \beta q^*(t)\tfrac{\partial C(q^*(t),y^*)}{\partial q} \\ &= p^*(t) - \beta q^*(t)\tfrac{\partial C(q^*(t),y^*)}{\partial q} \\ &= \tfrac{\partial M(q^*(t),y^*,t)}{\partial q}. \end{aligned} \tag{6.28}$$

Thus, at time t, the toll revenue can be expressed as:

$$\begin{aligned} R\left(q^{*}\left(t\right),y^{*}\right) &= p^{*}\left(t\right)q^{*}\left(t\right) \\ &= q^{*}\left(t\right)\left(\beta q^{*}\left(t\right)\frac{\partial C\left(q^{*}\left(t\right),y^{*}\right)}{\partial q}+\frac{\partial M\left(q^{*}\left(t\right),y^{*},t\right)}{\partial q}\right). \end{aligned} \tag{6.29}$$

Substituting (6.26) and (6.27) into (6.29), we obtain,

$$\begin{aligned} R\left(q^{*}\left(t\right),y^{*}\right) = & -\beta y^{*}\left(q^{*}\left(t\right)\frac{\partial C\left(q^{*}\left(t\right),y^{*}\right)}{\partial y}-\frac{\partial M\left(q^{*}\left(t\right),y^{*},t\right)}{\partial y}\right) \\ & +M\left(q^{*}\left(t\right),y^{*},t\right) \\ = & \; y^{*}\left(\frac{\partial S\left(q^{*}\left(t\right),y^{*}\right)}{\partial y}-\frac{\partial M\left(q^{*}\left(t\right),y^{*},t\right)}{\partial y}\right)+M\left(q^{*}\left(t\right),y^{*},t\right), \end{aligned} \tag{6.30}$$

where we have used the relation, $-\beta q^{*}\left(t\right)\partial C\left(q^{*}\left(t\right),y^{*}\right)/\partial y=\partial S\left(q^{*}\left(t\right),y^{*}\right)/\partial y$, which is obtained by taking the derivative of $S\left(q\left(t\right),y\right)$ in (6.2) with respect to y. Furthermore, substituting (6.30) into (6.29) and in view of Assumption (6.24), we readily obtain:

$$\int_{0}^{\tilde{T}}\left(R\left(q^{*}\left(t\right),y^{*}\right)-M\left(q^{*}\left(t\right),y^{*},t\right)\right)e^{-\lambda t}\mathrm{d}t=y^{*}I'\left(y^{*}\right)=I\left(y^{*}\right). \tag{6.31}$$

Equation (6.31) means that the revenue generated from optimum road user charge just covers the road construction and maintenance cost. Namely, the classical self-financing rule still holds in a dynamic environment under the common assumptions about investment cost and link travel time functions. We have the following self-financing rule.

Proposition 6.3.2. *Supposing that assumptions (6.23)-(6.25) are satisfied in the sense that the road construction exhibits constant returns to scale and the link travel time and maintenance cost functions are homogeneous of degree zero and one in traffic load and road capacity, respectively, then the revenue from socially optimal pricing on a road during its whole life just covers the capital cost for constructing and maintaining the road.*

Note that Arnott and Kraus (1998a) already showed that self-financing results for congestible facilities do extend to inter-temporal economic environments; in their extension, maintenance cost was assumed to be the function of time. In contrast, the above self-financing result is established under the assumption that the maintenance cost is a function of both time and traffic load. The self-financing results hold under assumptions

(6.23)-(6.25). In the case with increasing (resp., decreasing) returns to scale in road construction and road maintenance, namely,

$$I(\alpha y) < (\text{resp.} >)\, \alpha I(y), \forall \alpha > 1, y > 0 \tag{6.32}$$

and

$$M(\alpha q, \alpha y, t) < (\text{resp.} >)\, \alpha M(q, y, t), \forall \alpha > 1, q, y, t > 0. \tag{6.33}$$

It can be easily observed from (6.28)-(6.31) that the revenue will fail to cover all the capital costs and thus the road will have to be subsidized for efficient operation (respectively, with decreasing returns to scale, the revenue will exceed the total capital cost and the road will earn a surplus). If, however, road construction and maintenance exhibit opposite returns to scales, the results become indeterminate. In reality, the full costs of highway transportation consist of various components (Levinson and Gillen, 1998). Conditions (6.24) and (6.25) are, generally, not satisfied. As the model proposed by Li and Madanat (2002), viewing the maintenance cost as the increasing function of roughness, which is the function of traffic load, capacity and calendar time (formula 4 in Newbery, 1989) and roughness is assumed to be uniformly distributed along the width of the road), then it is easy to see there is an decreasing return to scale in road maintenance. The revenue will exceed the total capital cost whenever there is a non-increasing return to scale in road construction.

6.4 The Unconstrained Profit Maximizing Problem

The unconstrained profit maximizing problem considered here is (for the private firm) to choose road capacity y, toll charge, $p(t)$ and a concession period T to maximize the total profit throughout the concession period, which is given by the total toll revenue minus total maintenance and construction costs.

$$\max_{y\geq 0, T\geq 0, q(\cdot)\in C^2[0,T]} P\left(q\left(\cdot\right), y, T\right) = \int_0^T \left(R\left(q\left(t\right), y\right) - M\left(q\left(t\right), y, t\right)\right) e^{-\lambda t} \mathrm{d}t - I\left(y\right), \tag{6.34}$$

where the first term in the integral is defined in (6.4). Note that traffic load $q\left(\cdot\right)$ is taken to be decision function in (6.34) in replace of toll function $p\left(t\right)$, and assumed to be twice differentiable in t. Along the same line of examination of the social-welfare maximizing problem, we have the following proposition for the unconstrained profit maximization problem.

Proposition 6.4.1. *Suppose $\left(\hat{q}\left(\cdot\right), \hat{y}, \hat{T}\right)$ is the optimal solution of the unconstrained profit maximizing problem (6.34), and let $\hat{q}_{\hat{T}} = \hat{q}\left(\hat{T}\right)$ represents the traffic load solution at time $t = \hat{T}$, then $\hat{q}\left(t\right)$ is a non-increasing function over calendar time with boundary condition:*

$$R\left(\hat{q}_{\hat{T}}, \hat{y}\right) - M\left(\hat{q}_{\hat{T}}, \hat{y}, \hat{T}\right) = 0, \tag{6.35}$$

the corresponding toll charge $\hat{p}\left(t\right)$ is non-decreasing over calendar time.

Proof. The proof is similar to that for Proposition 6.3.1 and outlined here. First, the optimality conditions for (6.34) are:

$$\frac{\partial R\left(\hat{q}\left(t\right), \hat{y}\right)}{\partial q} - \frac{\partial M\left(\hat{q}\left(t\right), \hat{y}, t\right)}{\partial q} = 0, \forall t \in \left[0, \hat{T}\right] \tag{6.36}$$

and

$$\int_0^{\hat{T}} \left(\frac{\partial R\left(\hat{q}\left(t\right), \hat{y}\right)}{\partial y} - \frac{\partial M\left(\hat{q}\left(t\right), \hat{y}, t\right)}{\partial y}\right) e^{-\lambda t} \mathrm{d}t - I'\left(\hat{y}\right) = 0. \tag{6.37}$$

Taking derivatives of both sides of Eq. (6.36) with respect to time t yields

$$\frac{\mathrm{d}\hat{q}\left(t\right)}{\mathrm{d}t} = \frac{\frac{\partial^2 M\left(\hat{q}(t), \hat{y}, t\right)}{\partial q \partial t}}{\frac{\partial^2 R\left(\hat{q}(t), \hat{y}\right)}{\partial q^2} - \frac{\partial^2 M\left(\hat{q}(t), \hat{y}, t\right)}{\partial q^2}} \leq 0. \tag{6.38}$$

In the same vein as Lemma 6.2.1 for social welfare maximization, we have the boundary condition (6.35), stating that the profit must be nil or revenue exactly equals maintenance cost at the end of concession period $t = \hat{T}$. Finally, like (6.18) we have $\mathrm{d}\hat{p}(t)/\mathrm{d}t \leq 0$ for $t \in \left[0, \hat{T}\right]$. □

Similarly, if the maintenance cost is separable as given in (6.20), then $\hat{q}(\cdot)$ is a constant and $\hat{q}(t) \equiv \hat{q}_{\hat{T}}$ and $\hat{p}(t) \equiv$ constant, $\forall t \in [0, \hat{T}]$.

6.5 The Profit-Constraint BOT Problem

The BOT problem considered here is a little different from that examined by Guo and Yang (2010b), in which the private firm builds and operates the road for a number of years and then the toll road is transferred to and operated by the government during the post-concession period. Here we define our BOT problem as choosing a combination of concession period, road capacity and toll charge to maximize the total social welfare subject to a required level of profit to the private firm during the concession period. Its rationality and connection with the model in Guo and Yang (2010b) will be made clear later by exploring its solution properties.

6.5.1 Problem formulation

The BOT problem is to select the concession period, road capacity and toll charge jointly to maximize the total social welfare subject to a required level of profit to the private firm during the whole concession period, and can be formulated below:

$$\max \ W(q(\cdot), y, T) \tag{6.39}$$

subject to

$$P(q(\cdot), y, T) \geq \tilde{P} \tag{6.40}$$

$$q(t) \geq 0,\ q(\cdot) \in C^2[0, T] \tag{6.41}$$

$$y \geq 0,\ T \geq 0 \tag{6.42}$$

where the total social welfare W and the total profit P are defined by Eqs. (6.11) and (6.34), respectively; $\tilde{P} \geq 0$ is the minimum profit margin acceptable to the private firm; the traffic load function $q(\cdot)$ is limited in $C^2[0, T]$, or twice differentiable functions in $[0, T]$. Let $P\left(q^*(\cdot), y^*, \tilde{T}\right)$ be the realized profit for the unconstrained social-welfare maximizing problem in Section 6.3 and define $\underline{P} = \max\left\{0, P\left(q^*(\cdot), y^*, \tilde{T}\right)\right\}$, also let $\bar{P}$ be the realized profit for the unconstrained profit maximizing problem in Section 6.4, it is assumed that $\bar{P} \geq 0$ (otherwise, private provision of the road becomes infeasible). Clearly, for any level of profit $\tilde{P}$ with $\underline{P} \leq \tilde{P} \leq \bar{P}$, existence of a feasible solution to the BOT problem (6.39)-(6.42) is ensured. If $P\left(q^*(\cdot), y^*, \tilde{T}\right) \geq 0$ and $\tilde{P}$ is set to be $0 \leq \tilde{P} \leq \underline{P}$, then the BOT problem (6.39)-(6.42) and the unconstrained social-welfare maximizing problem (6.11) have identical solution, and as a result, a first-best contract (Guo and Yang, 2010b) can be reached. If $\underline{P} < \tilde{P} \leq \bar{P}$, then constraint (6.40) becomes binding and the solution of the BOT problem (6.39)-(6.42) differs from the unconstrained social-welfare maximizing (first-best) solution, as a result, the optimal contract from (6.39)-(6.42) becomes the second-best (Guo and Yang, 2010b). Evidently, the BOT problem (6.39)-(6.42) is infeasible for $\tilde{P} > \bar{P}$.

6.5.2 Solution Properties under Binding Profit Constraint

Here we look into the properties of the second-best BOT solution with $\underline{P} < \tilde{P} \leq \bar{P}$, as is the most case in practice. Let $(q^{**}(\cdot), y^{**}, T^{**})$ be a second-best solution, then we have $P(q^{**}(\cdot), y^{**}, T^{**}) = \tilde{P}$ by definition. With given optimal concession period T^{**} and road capacity y^{**}, traffic load $q^{**}(\cdot)$ must be a solution of the following functional extreme with iso-perimetric constraint (van Brunt, 2004):

$$\max J(q(\cdot)) = \int_0^{T^{**}} \left(S(q(t), y^{**}) - M(q(t), y^{**}, t)\right) e^{-\lambda t} \mathrm{d}t - I(y^{**}) \tag{6.43}$$

subject to

$$\int_0^{T^{**}} \left(R(q(t), y^{**}) - M(q(t), y^{**}, t)\right) e^{-\lambda t} \mathrm{d}t = I(y^{**}) + \tilde{P}, \tag{6.44}$$

where $q(\cdot)$ is twice differentiable with variable endpoints. Let

$$F = S(q(t), y^{**}) - M(q(t), y^{**}, t) - \mu \left(R(q(t), y^{**}) - M(q(t), y^{**}, t)\right), \tag{6.45}$$

where μ is the Lagrange multiplier associated with (6.44). By Euler-Lagrange equation, $q^{**}(\cdot)$ must satisfy the following condition:

$$\begin{aligned} 0 = \frac{\partial F}{\partial q} = & \frac{\partial S(q^{**}(t), y^{**})}{\partial q} - \frac{\partial M(q^{**}(t), y^{**}, t)}{\partial q} \\ & -\mu \left(\frac{\partial R(q^{**}(t), y^{**})}{\partial q} - \frac{\partial M(q^{**}(t), y^{**}, t)}{\partial q} \right). \end{aligned} \tag{6.46}$$

Since the extremum of $J(q)$ at q^{**} depends on T^{**}, y^{**} and $\tilde{P}$, the Lagrange multiplier depends on those parameters as well. With given T^{**} and y^{**}, we can regard J as a function of parameter $\tilde{P}$ only. From conditions (6.44) and (6.46), we can obtain

$$\frac{\mathrm{d}J}{\mathrm{d}\tilde{P}} = \int_0^{T^{**}} \frac{\mathrm{d}}{\mathrm{d}\tilde{P}} \left(F e^{-\lambda t} + \mu \frac{I(y^{**}) + \tilde{P}}{T^{**}} \right) \mathrm{d}t = \mu. \tag{6.47}$$

Thus the Lagrange multiplier corresponds to the rate of change of total social welfare with respect to the profit for given T^{**} and y^{**}. Since increasing the profit level $\tilde{P}$ would decrease the optimal social welfare gain $J(q^{**}(\cdot))$, we must have $\mu \leq 0$. With this observation, we can readily establish the following property.

Property 6.5.1. *Suppose$(q^{**}(\cdot), y^{**}, T^{**})$ is a second-best contract of the BOT problem (6.39)-(6.42), then $q^{**}(t)$ is non-increasing and $p^{**}(t)$ is non-decreasing over time t.*

Proof. Taking derivations of both sides of Eq. (6.46) in time t and in view of Eq. (6.3) and Assumptions 6.1 and 6.2, we can also get

$$\frac{\mathrm{d}q^{**}(t)}{\mathrm{d}t} = \left[\frac{(1-\mu)\frac{\partial^2 M}{\partial q \partial t}}{\frac{\partial^2 S}{\partial q^2} - \frac{\partial^2 M}{\partial q^2} - \mu\left(\frac{\partial^2 R}{\partial q^2} - \frac{\partial^2 M}{\partial q^2}\right)} \right]_{(q^{**}(t), y^{**}, t)} \leq 0. \tag{6.48}$$

From equilibrium condition (6.1), the corresponding derivative of toll charge $p^{**}(t)$ is obtained:

$$\frac{\mathrm{d}p^{**}}{\mathrm{d}t} = \left(B'(q^{**}(t)) - \beta \frac{\partial C(q^{**}(t), y^{**})}{\partial q} \right) \frac{\mathrm{d}q^{**}}{\mathrm{d}t} \geq 0. \tag{6.49}$$

The proof is completed. □

It is notable again that, if the marginal user damage on road is independent of time t or Eq. (6.21) holds, then, q^{**} and p^{**} become constant.

Property 6.5.2. *For a second-best BOT contract* $(q^{**}(\cdot), y^{**}, T^{**})$*,*

$$\hat{q}_{y^{**}}(t) \leq q^{**}(t) \leq q^{*}_{y^{**}}(t), \ \forall t \in [0, T^{**}], \tag{6.50}$$

where $q^{*}_{y^{**}}(\cdot)$ *and* $\hat{q}_{y^{**}}(\cdot)$ *are the social-welfare and profit maximizing traffic loads, respectively at given capacity* $y = y^{**}$*.*

Proof. It is easy to see that $q^{*}_{y^{**}}(t)$ and $\hat{q}_{y^{**}}(t)$ maximize $S(q, y^{**}) - M(q, y^{**}, t)$ and $R(q, y^{**}) - M(q, y^{**}, t)$, respectively, or they satisfy conditions (6.12) and (6.36), respectively, for $y = y^{**}$. From definitions (6.2) and (6.4),

$$S(q(t), y^{**}) - R(q(t), y^{**}) = \int_0^{q(t)} B(w)\,\mathrm{d}w - q(t)B(q(t)). \tag{6.51}$$

For any $0 \leq q \leq \hat{q}_{y^{**}}(t)$, we thus have

$$\begin{aligned} \tfrac{\partial S(q,y^{**})}{\partial q} - \tfrac{\partial M(q,y^{**},t)}{\partial q} = {} & \tfrac{\partial R(q,y^{**})}{\partial q} \\ & -\tfrac{\partial M(q,y^{**},t)}{\partial q} + (-qB'(q)) > 0, \end{aligned} \tag{6.52}$$

where we have used $\partial(R - M)/\partial q > 0$ for $0 \leq q \leq \hat{q}_{y^{**}}(t)$, because $R - M$ is strictly concave in q and has a maximum at $q = \hat{q}_{y^{**}}(t)$. Since $S - M$ is strictly concave in q as well and has a maximum at $q = q^{*}_{y^{**}}(t)$, we must have

$$q^{*}_{y^{**}}(t) > \hat{q}_{y^{**}}(t), \ \ \forall t \in [0, T^{**}]. \tag{6.53}$$

From the above discussions, we see that $R-M$ and $S-M$ are increasing for $0 \leq q < \hat{q}_{y^{**}}(t)$ and decreasing for $q > q^{*}_{y^{**}}(t)$ and thus conclude that Property 6.5.2 is true. Otherwise, we can always modify $q^{**}(\cdot)$ to improve both objectives, which conflicts with the optimality of $(q^{**}(\cdot), y^{**}, T^{**})$. □

Property 6.5.2 tells us that the traffic load under a second-best con-tract is between the socially desirable demand and the demand preferred by the private firm.

Property 6.5.3. *For a second-best BOT contract $(q^{**}(\cdot), y^{**}, T^{**})$, profit is always negative for any toll policy and anytime after the concession period T^{**}:*

$$R(q, y^{**}) - M(q, y^{**}, t) < 0, \ \forall q > 0, \ \forall t > T^{**}. \tag{6.54}$$

Proof. Suppose there exist $q_0 > 0$ and $t_0 > T^{**}$ that deviate condition (6.54), namely, $R(q_0, y^{**}) - M(q_0, y^{**}, t_0) \geq 0$, because $M(q, y, t)$ is monotonically increasing in t, we have

$$R(q_0, y^{**}) - M(q_0, y^{**}, t) > 0, \ \forall t \in (T^{**}, t_0). \tag{6.55}$$

Without difficulty, we can extend the solution function $q^{**}(\cdot)$ smoothly (a polynomials of degree 3 is enough) from T^{**} to t_0 such that $q^{**}(t_0) = q_0$ and

$$\int_{T^{**}}^{t_0} \left(R(q^{**}(t), y^{**}) - M(q^{**}(t), y^{**}, t)\right) e^{-\lambda t} \mathrm{d}t > 0. \tag{6.56}$$

In view of (6.51), we further obtain

$$\int_{T^{**}}^{t_0} \left(S(q^{**}(t), y^{**}) - M(q^{**}(t), y^{**}, t)\right) e^{-\lambda t} \mathrm{d}t > 0. \tag{6.57}$$

Equations (6.56) and (6.57) contradict the optimality of $(q^{**}(\cdot), y^{**}, T^{**})$. Therefore, condition (6.54) is true. □

Property 6.5.3 implies that operating the road will always yield negative profit after concession period T^{**} under a second-best contract. This in turn implies that the profit earned by the private firm can never be positive at the end of concession period $t = T^{**}$ due to road deterioration. To avoid the risk of subsidizing a deteriorated road by the government, it might be preferable to rebuild the road for a new lifecycle.

Finally we note that, if the maintenance cost is neglected in the BOT problem (6.39)-(6.42) and the road life is assumed to be given (finite or infinite), then the optimal concession period must be the road life since we can always improve the social welfare and the profit via prolonging the concession period for any given q and y. In this environment, the BOT problem (6.39)-(6.42) and the static model proposed by Guo and Yang (2010b) have identical solution, or the dynamic model reduces to the static one in Guo and Yang (2010b).

6.5.3 Governmental Regulation for Optimal BOT Contract

We look into how to reach an optimal BOT contract through a bilateral negotiation between the government and the private firm. We treat the bilateral negotiation as a sequential game in which the government is the regulator or leader and the private firm is the follower. The government first sets (or imposes restrictions on) the values of the regulatory variables, while taking account of the responses of the private firm, then the private firm freely chooses the values of the rest of decision variables. We first consider full regulation in which all variables are regulated by the government. For given road capacity and concession period T^{**}, from relation (6.50) the second-best toll $p^{**}(t)$ must satisfy

$$\hat{p}_{y^{**}}(t) \geq p^{**}(t) \geq p^{*}_{y^{**}}(t), \; \forall t \in [0, T^{**}], \tag{6.58}$$

where $\hat{p}_{y^{**}}(t)$ and $p^*_{y^{**}}(t)$ are the socially optimal and profit-maximizing toll at time t for given $y = y^{**}$. In fact, $p^{**}(\cdot)$ can be adopted as the dynamic price cap once y^{**} and T^{**} are determined. Namely, the private firm must choose $p(t) = p^{**}(t)$ at time t to maximize its profit if just allowed to set toll charge under the following restriction:

$$p(t) \geq p^{**}(t), \ \forall t \in [0, T^{**}]. \tag{6.59}$$

We now move to consider the more realistic partial regulation. The following proposition is first given.

Proposition 6.5.1. *The social welfare, $W(q^{**}(\cdot), y^{**}, T^{**})$ at either first-best or second-best optimal contract $(q^{**}(\cdot), y^{**}, T^{**})$, can be realized whenever the government restricts the concession period $T \geq T^{**}$ and traffic load $q(t) \geq q^{**}(t)$ for $t \in [0, T]$.*

Proof. Suppose the private firm chooses a more profitable deviation $(q(\cdot), y, T)$ with $q(t) \geq q^{**}(t), \ \forall t \in [0, T]$ and $T \geq T^{**}$. In this case, the social welfare can be expressed as:

$$W(q(\cdot), y, T) = P(q(\cdot), y, T) + \int_0^T \int_0^{q(t)} (B(w) - B(q(t))\, w)\, e^{-\lambda t} \mathrm{d}w\mathrm{d}t. \tag{6.60}$$

Note that the function $\int_0^q (B(w) - B(q)\, w)\, \mathrm{d}w$ is strictly increasing with q. Thus, $q(t) \geq q^{**}(t), \ \forall t \in [0, T]$ and $T \geq T^{**}$ gives rise to

$$\begin{aligned} &\textstyle\int_0^T \int_0^{q(t)} (B(w) - B(q(t))\, w)\, e^{-\lambda t} \mathrm{d}w\mathrm{d}t \\ &\qquad \geq \textstyle\int_0^{T^{**}} \int_0^{q^{**}(t)} (B(w) - B(q^{**}(t))\, w)\, e^{-\lambda t} \mathrm{d}w\mathrm{d}t. \end{aligned} \tag{6.61}$$

By assumption, $(q(\cdot), y, T)$ is more profitable:

$$P(q(\cdot), y, T) \geq P(q^{**}(\cdot), y^{**}, T^{**}). \tag{6.62}$$

Adding both sides of (6.61) and (6.62) yields:

$$W(q(\cdot), y, T) \geq W(q^{**}(\cdot), y^{**}, T^{**}). \tag{6.63}$$

The optimality of $(q^{**}(\cdot), y^{**}, T^{**})$ implies that:

$$W(q(\cdot), y, T) = W(q^{**}(\cdot), y^{**}, T^{**}). \tag{6.64}$$

Therefore, the optimal social welfare is achieved under the restrictions of concession period, $T \geq T^{**}$, and traffic load $q(t) \geq q^{**}(t)$, $t \in [0, T]$. □

A major difference from the results of Guo and Yang (2010b) is that the private firm would prefer a limited concession period and leave a deteriorated road to the government once receiving a satisfactory profit return whenever maintenance cost comes into effect.

6.6 Effect of a Growing Economy

The demand for transport is a derived demand, which depends on various factors including income level, cost structure of all travel modes and other socioeconomic factors. The analysis so far pertains to a certain time frame within which the various factors determining the shape and positions of the demand curve can not change. However, a BOT scheme would normally run for few decades, over such a long time period travel demand will certainly shift. In this section, we ex-tend our model by considering the situation with a continually changing economy, in a similar spirit of the dynamic pricing policies applied successfully in the air-line industry (Talluri and van Ryzin, 2004). Let $B = B(q, t)$ denote the inverse demand or benefit function. We consider the case where the demand curve shifts upwards over time in a growing economy or $B(q, t)$ is increasing in t ($\partial B / \partial t \geq 0$) for any given $q \geq 0$. Also, we assume that there is a steady growth of economy with a declining growth rate. Namely, for $q \geq 0$, $B(q, t)$ is upper bounded with $\partial^2 B / \partial t^2 \leq 0$. Together with Assumption 3.1 and Assumptions 6.1-6.2, we can obtain the corresponding optimality conditions under the inter-temporal economies by following the line of earlier discussions. However, the monotonicity of the corresponding optimal traffic load functions would become undetermined. The important parallel conclusions are listed as below without tedious mathematical formulations:

(1) Along the unconstrained social-welfare maximizing traffic load path, the optimal traffic load would be non-increasing (non-decreasing) over time if the marginal user damage on road increases faster (slower) than the marginal benefit does over time.

(2) Along the unconstrained profit-maximizing traffic load path, the profit-maximizing traffic load would be non-increasing (non-decreasing) over time if the marginal user damage on road increases faster (or slower) than the marginal revenue over time.

(3) For the profit-constrained second-best BOT problem with economic growth, it is more complicated to discern the monotonicity of the optimal traffic load path. Nevertheless, in the special case when the marginal user damage on road is independent of calendar time, then the optimal traffic load functions in-creases over time.

(4) Finally, it is easy to check that the self-financing result still holds if the conditions of Proposition 2 are satisfied. Yet, both full and partial regulations are efficient in terms of social welfare attainment in the growing economic environment.

6.7 Conclusions

In this chapter, we examined the BOT problem a dynamic environment with road deterioration and maintenance effects. By assuming that the maintenance cost to sustain the road quality is a natural function of road capacity, traffic load and calendar time, the road life is endogenously determined as the time when the net economic benefit becomes nil as maintenance cost increases. We established several meaningful results. Firstly, under a first-best BOT contract, the concession period should be equal to the road life; the discounted toll revenue just covers the total capital cost including both maintenance and construction cost under a certain assumption (the classic self-financing theorem still holds). Secondly, the second-best optimal BOT contract can be determined by an isoperimetric problem to maximize the social welfare with a profit constraint. The optimal dynamic toll charge is increasing over time, as deemed necessary for reducing traffic load and thereby lowering down the road damage. The optimal BOT contract would not reach the road life. Thirdly, with increasing maintenance cost over time, the government would set a concession period longer than preferred by the private firm and set a minimum time-varying traffic load higher than desired by the private firm at any time. We also proved that if the marginal user damage on road is in-dependent of time, then the optimal toll charge is free from the effect of road natural deterioration and thus would be time-invariant. Furthermore, although economic growth has an ambiguous effect on pricing policies, the full and partial regulations remain effective.

7

PARETO EFFICIENCY OF TOLL ROAD PROJECTS UNDER VARIOUS OWNERSHIP REGIMES

As the increasing of the private provision of the public roads, specially, by highway franchising contract, the ownership regimes of the transportation network tend to be complicated, which has significant impacts on the decision of the new "add-on" toll road projects. This chapter investigates the effects of ownership regimes of the existing road on the decision of the capacity and toll choices of the new "add-on" toll road project, where the alternative road may be free, public or private toll road. When the alternative road is a private toll road, we assume that the operators of the two roads, simultaneously, set toll and/or capacity levels, respectively, following the one-shot game mechanism. The public sector is assumed to consider both the total social welfare of the whole network and net profit of the new toll road project when facing various ownership regimes. The properties of the Pareto-optimal solution sets under different ownership regimes are investigated. We also examine the existence of the Pareto improving schemes, which enhance the total social welfare with a positive profit gain for the public sector, of the new road project with various ownership regimes.

7.1 Introduction

As the increasing of the private provision of the public roads, specially, by highway franchising contract, the ownership regimes of the transportation network tend to be complicated, which has significant impacts on the decision of the new "add-on" toll road projects. As mentioned in Section 2.4, many previous studies on the private toll competition, either tolls or capacities-and-tolls are the decision variables in road competition among the firms. The asymmetry of the competition was captured by the performance function of the road and/or investment cost. However, in reality, the capacity/toll choice problem is seemed to be a little complicated because of road differentiation: the road ownership differentiation, and therefore objective differentiation (the private firms aim to maximize profit while the government cares about social welfare); the entry-time differentiation and thus decision-variable asymmetry (an "add-on" new road can be determined by capacity and toll, while only the toll level can be adjusted for an existing road). Inescapably, a BOT contract must face various market environments. The famous example is the autoroutes in France, which consists largely of private-owned toll roads, above 90% of the whole road network. Figure 7.1 shows the map of autoroute network of France. Another example is the three harbour tunnels in Hong Kong (Hong Kong Cross-Harbour Tunnel, Eastern Harbour Crossing and Western Harbour-Crossing), operating from beginning 1972, 1989 and 1997 and owning for a 30-year concession period, sequentially (Tam, 1998).

De Palma and Lindsey (2000) investigated the toll settings and efficiency gains of the competition between two parallel roads with three private ownership regimes: (1) a private road on one route and free access on the other, (2) a private roads duopoly, and (3) a mixed duopoly with a private road competing with a public toll road. Yang and Huang (2005) and Verhoef (2007) analyzed the socially optimal toll/capacity choice with an existing free alternative. They revealed that the optimal toll equal to the marginal external congestion costs on the new tolled road minus a (positive) term consisting of a fraction (between 1 and 0) of the marginal external congestion costs on the un-tolled route. The government planner should take into account the flow "spill-over" effect on the free road in order to maximize the social welfare. They also pointed out that the resulting revenue would fail to cover the capacity cost.

In this chapter, we focus on the choice of the capacity and toll for an "add-on" toll road parallel to an existing road with various ownership regimes, which can be a free, public or

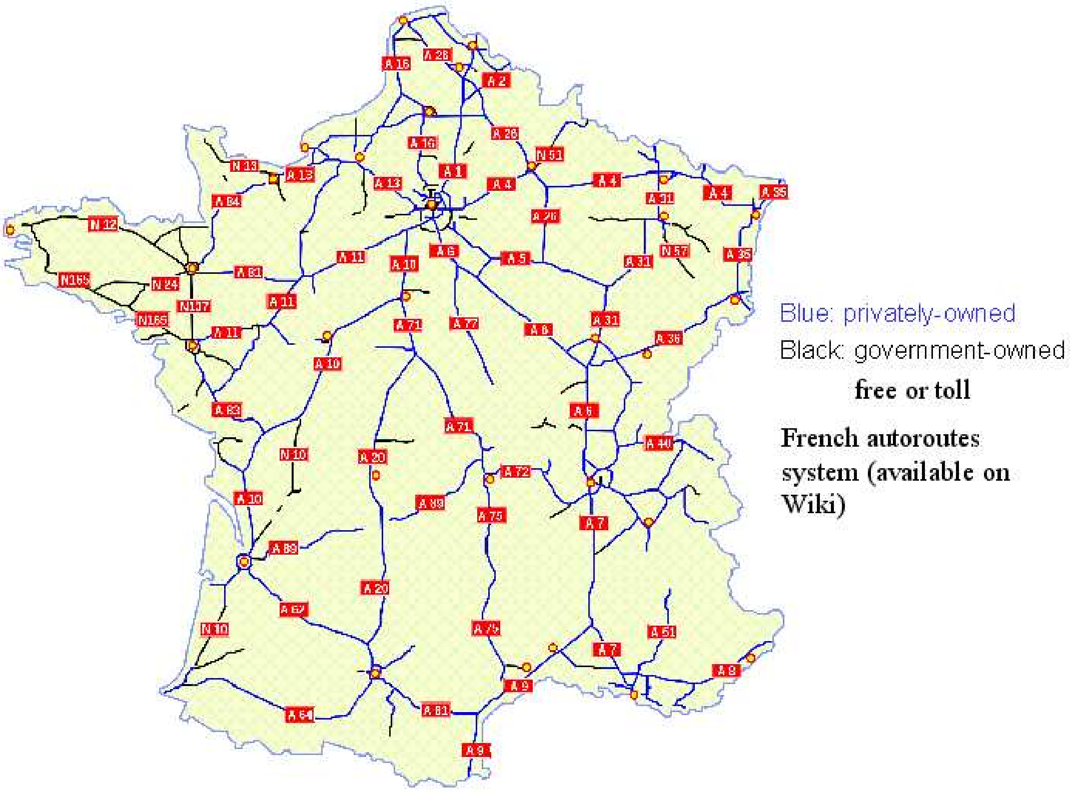


Figure 7.1: Map of Autoroutes in France

private toll road. The public sector is assumed to consider the total social welfare gain of the network and the net profit return of the new road project, simultaneously. When the alternative road is a public toll road, the desirable outcomes are, generally, easy to obtain since the public sector can freely select the toll of the existing road. In this case, the welfare-maximizing toll and capacity levels corresponds to the first-best solution. When the alternative road is a private toll road, we assume that the operators of the two roads following the one-shot game mechanism. The private sector chooses the toll level of the existing road to maximize the toll revenue of the existing road, while the public sector selects a combination of the toll and capacity levels of the "add-on" road to achieve a desirable outcomes of total social welfare and net profit. Our aims are: to examine the properties of the toll and capacity levels of the "add-on" toll road at any Pareto optimum under various ownership regimes and to investigate the effects of the ownership regimes on those properties. When a new toll road project is considered by the public sector, the Pareto-improving outcomes are significant important, which can enhance the total social welfare with a positive profit gain. To examine the existence of the Pareto-improving schemes under different ownership regimes is also our major objective in the current study.

The chapter is organized as follows. Section 7.2 introduces the problems and some basic settings. The properties of the Pareto-optimal solution sets under different ownership regimes are investigated in Section 7.3. Section 7.4 examines the existence of the Pareto-improving schemes for the public sector to introduce the new road project with various ownership regimes of the alternative road. The effects of the ownership regimes on the public sector's decisions, including the incentives to invest the new toll road project in the term of welfare maximization and Pareto improvement, optimal choice of capacity level, are studied in Section 7.5. A simple analytical example is used to elucidate and test those results in this section. Finally, conclusions are presented in Section 7.6.

7.2 The Problems and Models

Suppose that the public sector plans to build a new toll road (A) connecting two cities besides an existing alternative road (E), which can be free, public or private toll road. The public sector cares of the total social welfare of the whole network and the profit of the new toll road project. Let p_E, p_A, y_E and y_A be the toll and capacity levels of two roads, respectively. We assume that the capacity of road E is predetermined and can not be adjusted. Denote q_E, t_E and q_A, t_A as the flows and travel time functions of roads E A, respectively. It is assumed that the link travel functions are convex and increasing functions with respect to the corresponding link flows, or, ${t'}_{iq_i} > 0$ and ${t''}_{iq_i} > 0$, $i \in \{E, A\}$. For the given two-link network, the following user equilibrium (UE) condition always holds:

$$\begin{cases} p_i + \beta t_i\left(q_i, y_i\right) = \mu & q_i > 0 \\ p_i + \beta t_i\left(q_i, y_i\right) \geq \mu & q_i = 0 \end{cases}, i \in \{E, A\}, \tag{7.1}$$

where β is the value-of-time (VOT) to convert time into equivalent monetary cost (we consider homogeneous users only), and capacity levels are measured in the same unit as the link flow, and μ is the minimal generalized travel cost.

Denote the benefit function (or inverse demand function) as $B(Q)$, which is strictly decreasing in Q, or $B'(Q) < 0$. The unit-time social welfare of the whole network in monetary is:

$$W = \int_0^Q B(w)\mathrm{d}w - \beta q_E t_E(q_E, y_E) - \beta q_A t_A(q_A, y_A) - \eta I(y_A) \tag{7.2}$$

where $I(y_A)$ is the construction cost of road A, which is an increasing function of capacity, y_A; η is the discount factor converting the total construction cost into each unit time during a planning period. The toll revenue of road E and the net profit of road A (suppose the operate cost is zero), respectively, are:

$$P_E = p_E q_E \tag{7.3}$$

and

$$P_A = p_A q_A - \eta I(y_A). \tag{7.4}$$

We assume that the net profit of road A and the social welfare of the whole network are concerned by the public sector when introducing the new project. If road E is free, the public sector's problem is to select a combination of the toll and capacity levels of road A to maximize the total social welfare and net profit of road A, simultaneously:

$$\max_{p_A \geq 0, y_A \geq 0} \begin{pmatrix} W(p_A, y_A) \\ P_A(p_A, y_A) \end{pmatrix}. \tag{7.5}$$

When road E is a public toll road, and the public sector is a "monopoly" of the system. She can select the toll levels of road E and A, together with the capacity level of road

A,to maximize the total social welfare net profit of road A, simultaneously,

$$\max_{p_E \geq 0, p_A \geq 0, y_A \geq 0} \begin{pmatrix} W(p_E, p_A, y_A) \\ P_A(p_E, p_A, y_A) \end{pmatrix}. \tag{7.6}$$

If road E is a private toll road, it corresponds to a simultaneous game between the public sector and the private sector of road E. The public sector selects a combination of toll and capacity levels of road A to maximize the total social welfare and net profit of road A, while the private firm of road E chooses a toll level of road E to maximize her own toll revenue, namely,

$$\max_{p_A \geq 0, y_A \geq 0} \begin{pmatrix} W(p_A, y_A) \\ P_A(p_A, y_A) \end{pmatrix} \text{ vs. } \max_{p_E \geq 0} P_E(p_E). \tag{7.7}$$

Note that problems (7.5)-(7.7) are subject to the constraint of the UE condition (7.1). Denote "I", "II", "III" as the scenarios that road E is a free, public or private toll road, respectively and a superscript with "I", "II", "III" is used to depict the solution variables of those problems.

7.3 Pareto-optimal Solutions under Various Ownership Regimes

In this section, we investigate the properties of the Pareto-optimal solution sets of problems (7.5)-(7.7) under various ownership regimes of road E by assuming the existence of the inner solutions of the three problems. In this case, the UE condition (7.1) is expressed as:

$$p_E + \beta t_E(q_E, y_E) = p_A + \beta t_A(q_A, y_A) = B(Q), Q = q_E + q_A. \tag{7.8}$$

Suppose (p_A, y_A) be any Pareto-optimal toll and capacity pair of problems (7.5)-(7.7). Denote q_A and q_E as the corresponding demand levels of road A and E. From the standard optimality method, we obtain that, for any given toll charge of road E, capacity

y_A must satisfy the following optimal condition at Pareto optimum:

$$\beta q_A \frac{\partial t_A\left(q_A, y_A\right)}{\partial y_A} + \eta \frac{\mathrm{d}I\left(y_A\right)}{\mathrm{d}y_A} = 0. \tag{7.9}$$

Again, Assumptions 3.2 and 3.3 in Chapter 3 are introduced in this chapter to derive some useful insights unless otherwise explicitly noted. Under Assumptions 3.2 and 3.3, Eq. (7.9) is reduced as:

$$\left(\gamma_A\right)^2 \frac{\mathrm{d}t_A\left(\gamma_A\right)}{\mathrm{d}\gamma_A} = \frac{\eta k}{\beta}. \tag{7.10}$$

Since the function of the left-hand side of Eq. (7.10) is strictly increasing in γ_A, the Pareto-optimal volume-capacity ratio γ_A of bi-objective problems (7.5)-(7.7) is constant and solves Eq. (7.10) whenever there are positive flows on both roads. Xiao et al. (2007b) demonstrated that each profit-maximizing firm in a network with parallel toll roads chooses toll and capacity to keep the v/c ratio at a reasonable level being independent on the other firms' choices. The result of Xiao et al. (2007b) has been derived in general transportation network by Wu et al. (2010). Eq. (7.10) means that, under Assumptions 3.2 and 3.3, the v/c ratio along the Pareto-optimal frontiers of problems (7.5)-(7.7) is always constant and equals the socially optimal level. We now denote the v/c ratio, determined by Eq. (7.10), as $\tilde{\gamma}_A$. The average social cost per trip of road A at Pareto-optimum is calibrated as:

$$C_0 = \beta t_A\left(q_A, y_A\right) + \frac{\eta I\left(y_A\right)}{q_A} = \beta t_A\left(\tilde{\gamma}_A\right) + \beta \tilde{\gamma}_A t'_A\left(\tilde{\gamma}_A\right), \tag{7.11}$$

Equation (7.11) shows that the average social cost per trip of road A at Pareto-optimum is only dependent on the v/c ratio of road A, and thus is constant along the Pareto-optimal frontiers of problems (7.5)-(7.7).

For any given toll and capacity levels of road A, p_A and y_A, in view p_E as the function of traffic flow, q_E, determined by UE condition (7.8), the toll revenue of private sector, defined by Eq. (7.3), is now expressed as the function of q_E,

$$P_E = q_E \left(B\left(q_A + q_E\right) - \beta t_E\left(q_E, y_E\right)\right). \tag{7.12}$$

In addition, since the public sector always selects a combination of toll and capacity levels to keep the optimal v/c ratio, $\tilde{\gamma}_A$, using definition (7.11), the net profit of road A, determined by Eq. (7.4), is reduced as:

$$P_A = p_A q_A - \eta I\left(y_A\right) = q_A\left(B\left(q_A + q_E\right) - C_0\right). \tag{7.13}$$

For the same reason, the total social welfare, given by Eq. (7.2), is rearranged as:

$$W = \int_0^Q \left(B\left(w\right) - C_0\right) \mathrm{d}w + q_E\left(C_0 - \beta t_E\left(q_E, y_E\right)\right). \tag{7.14}$$

To guarantee the analytical treatability of problems (7.5)-(7.7), we assume that, under UE condition (7.8), the toll revenue of road E, P_E, is concave in q_E for any given p_A and y_A; both the profit of road A and social welfare of the network, P_A and W, are concave in q_A for any given p_E.

7.3.1 Pareto-optimal solutions when the alternative road is free

In this subsection, we examine the properties of the Pareto-optimal solutions when the alternative road is free. It is necessary to investigate the welfare-maximizing and profit-maximizing solutions. Denote $\left(\tilde{p}_A^{\mathrm{I}}, \tilde{y}_A^{\mathrm{I}}\right)$ and $\left(\bar{p}_A^{\mathrm{I}}, \bar{y}_A^{\mathrm{I}}\right)$ as the welfare-maximizing and profit-maximizing solutions of problem (7.5), respectively. Correspondingly, denoted the traffic flow of road A and E as $\tilde{q}_A^{\mathrm{I}}$ and $\bar{q}_A^{\mathrm{I}}$, respectively. From the optimality condition, we know that the best response of the public sector to maximize the total social welfare is to set the toll charge of road A such that

$$\tilde{p}_A^{\mathrm{I}} = \beta \tilde{q}_A^{\mathrm{I}} \frac{\partial t_A\left(\tilde{q}_A^{\mathrm{I}}, \tilde{y}_A^{\mathrm{I}}\right)}{\partial q_A} - \tilde{\alpha}_1^{\mathrm{I}} \beta \tilde{q}_E^{\mathrm{I}} \frac{\partial t_E\left(\tilde{q}_E^{\mathrm{I}}, y_E\right)}{\partial q_E} \tag{7.15}$$

where $\tilde{\alpha}_1^{\mathrm{I}} = B'\left(\tilde{Q}^{\mathrm{I}}\right) / \left(B'\left(\tilde{Q}^{\mathrm{I}}\right) - \beta \partial t_E\left(\tilde{q}_E^{\mathrm{I}}, y_E\right) / \partial q_E\right) \in [0,1]$ and $\tilde{Q}^{\mathrm{I}} = \tilde{q}_A^{\mathrm{I}} + \tilde{q}_E^{\mathrm{I}}$. Equation (7.15) means that the optimal toll equal to the marginal external congestion cost on the new toll road minus a positive term consisting of a fraction (between 1 and 0) of the marginal external congestion cost on the un-tolled road. The welfare-maximizing public sector should take into account the flow "spill-over" effect on the free road in order to maximize the social welfare (Verhoef, 2007; Yang and Huang, 2005). The best response of the public sector to maximize the net profit of road A is to set the toll charge of road A as

$$\bar{p}_A^{\mathrm{I}} = \beta \bar{q}_A^{\mathrm{I}} \frac{\partial t_A\left(\bar{q}_A^{\mathrm{I}}, \bar{y}_A^{\mathrm{I}}\right)}{\partial q_A} + \bar{\alpha}_1^{\mathrm{I}} \beta \bar{q}_A^{\mathrm{I}} \frac{\partial t_E\left(\bar{q}_E^{\mathrm{I}}, y_E\right)}{\partial q_E}, \tag{7.16}$$

where $\bar{\alpha}_1^{\mathrm{I}} = B'\left(\bar{Q}^{\mathrm{I}}\right) / \left(B'\left(\bar{Q}^{\mathrm{I}}\right) - \beta \partial t_E\left(\bar{q}_E^{\mathrm{I}}, y_E\right) / \partial q_E\right) \in [0,1]$ and $\bar{Q}^{\mathrm{I}} = \bar{q}_A^{\mathrm{I}} + \bar{q}_E^{\mathrm{I}}$.

Viewing the different aims of the public sector to maximizing the total social welfare or net profit, and using definition (7.11), UE condition (7.8) can be expressed as, respectively

$$\beta t_E\left(\tilde{q}_E^{\mathrm{I}}, y_E\right) = C_0 - \tilde{\alpha}_1^{\mathrm{I}} \beta \tilde{q}_E^{\mathrm{I}} \frac{\partial t_E\left(\tilde{q}_E^{\mathrm{I}}, y_E\right)}{\partial q_E} = B\left(\tilde{q}_A^{\mathrm{I}} + \tilde{q}_E^{\mathrm{I}}\right), \tag{7.17}$$

and

$$\beta t_E\left(\bar{q}_E^{\mathrm{I}}, y_E\right) = C_0 + \bar{\alpha}_1^{\mathrm{I}} \beta \bar{q}_A^{\mathrm{I}} \frac{\partial t_E\left(\bar{q}_E^{\mathrm{I}}, y_E\right)}{\partial q_E} = B\left(\bar{q}_A^{\mathrm{I}} + \bar{q}_E^{\mathrm{I}}\right). \tag{7.18}$$

From Eq. (7.17), it is easy to check that the derivative of the profit function , given by Eq. (7.13), $\mathrm{d}P_A/\mathrm{d}q_A \leq 0$ at $\tilde{q}_A^{\mathrm{I}}$. Therefore, the public sector can not increase his profit gain from the level at the welfare-maximizing solution by increasing capacity investment because P_A is concave in q_A. In addition, since the profit-maximizing public sector can earn strictly positive profit by selecting solution $\left(\bar{p}_A^{\mathrm{I}}, \bar{y}_A^{\mathrm{I}}\right)$, we know $\tilde{q}_A^{\mathrm{I}} \geq \bar{q}_A^{\mathrm{I}}$ and $\tilde{y}_A^{\mathrm{I}} \geq \bar{y}_A^{\mathrm{I}}$. Furthermore, the Pareto-optimal solution of problem (7.5) results in a Pareto-optimal demand-capacity pair set belongs to domain $\Omega = \left\{(q_A, y_A) : q_A = y_A \tilde{\gamma}_A, \bar{y}_A^{\mathrm{I}} \leq y_A \leq \tilde{y}_A^{\mathrm{I}}\right\}$. Evidently, directly from UE conditions (7.17) and (7.18), the toll level at the profit-maximizing solution is higher than that at the welfare-maximizing one, $\tilde{p}_A^{\mathrm{I}} \leq \bar{p}_A^{\mathrm{I}}$, which results a lower total demand, $\tilde{Q}^{\mathrm{I}} \leq \bar{Q}^{\mathrm{I}}$, and makes the alternative free road more congested,

$\tilde{q}_E^{\mathrm{I}} \leq \bar{q}_E^{\mathrm{I}}$.

7.3.2 Pareto-optimal solutions when the alternative is a public toll road

We now consider problem (7.6), in which, road E is a public toll road. In this case, the public sector is a "monopoly" in the network and can freely select the toll level of road E. Notably, the public sector can determine the first-best solution by setting the congestion externalities on both road to maximize the total social welfare, which corresponds to the well-known first-best solution in the highway financing problem (Verhoef and Mohring, 2009). At welfare-maximizing solution, we know

$$\tilde{p}_E^{\mathrm{II}} = \beta \tilde{q}_E^{\mathrm{II}} \frac{\partial t_E\left(\tilde{q}_E^{\mathrm{II}}, y_E\right)}{\partial q_E}, \tilde{p}_A^{\mathrm{II}} = \beta \tilde{q}_A^{\mathrm{II}} \frac{\partial t_A\left(\tilde{q}_A^{\mathrm{II}}, \tilde{y}_A^{\mathrm{II}}\right)}{\partial q_A}, \tag{7.19}$$

where $\tilde{q}_E^{\mathrm{II}}$ and $\tilde{q}_A^{\mathrm{II}}$ are the corresponding link flow of road E and A associated to the welfare-maximizing solution. UE condition (7.8) is now reduced as

$$\beta t_E\left(\tilde{q}_E^{\mathrm{II}}, y_E\right) + \beta \tilde{q}_E^{\mathrm{II}} \frac{\partial t_E\left(\tilde{q}_E^{\mathrm{II}}, y_E\right)}{\partial q_E} = C_0 = B\left(\tilde{q}_A^{\mathrm{II}} + \tilde{q}_E^{\mathrm{II}}\right). \tag{7.20}$$

Equation (7.20) implies that both the total demand and flow of road E are completely determined by the average social cost C_0. Specially, the total demand is independent on the travel delay function of road E, or the capacity level of road E.

It is intuitive to examine that the profit-maximizing solution $\left(\bar{p}_E^{\mathrm{II}}, \bar{p}_A^{\mathrm{II}}, \bar{y}_A^{\mathrm{II}}\right)$ is to set the toll level of road E high enough such that road A produces the monopoly profit, namely,

$$\bar{p}_E^{\mathrm{II}} + \beta t_E\left(0, y_E\right) \geq C_0 - \bar{q}_A^{\mathrm{II}} B'\left(\bar{q}_A^{\mathrm{II}}\right) = B\left(\bar{q}_A^{\mathrm{II}}\right). \tag{7.21}$$

Because if there is a positive flow on road E for a certain feasible solution $\left(p_E^{\mathrm{II}}, p_A^{\mathrm{II}}, y_A^{\mathrm{II}}\right)$, then increasing the toll charge of road E can increase the net profit production of road A. Therefore, at profit-maximizing solution, there is no positive flow on road E. It must be pointed out that the profit-maximizing toll charge of road E, $\bar{p}_E^{\mathrm{II}}$, may not be unique, while the profit-maximizing capacity $\bar{y}_A^{\mathrm{II}}$ is always constant and equals $\bar{q}_A^{\mathrm{II}}/\tilde{\gamma}_A$, where $\bar{q}_A^{\mathrm{II}}$

is determined by condition (7.21).

From conditions (7.20) and (7.21), we know that, when the alternative road is a public toll road, the profit-maximizing demand and capacity levels are the same as those at monopoly optimum in the case with a single highway project studied in Chapter 3, while the welfare-maximizing demand level is lower than that at social optimum in the case with a single highway project. Furthermore, since $\bar{q}_A^{\mathrm{II}}$, determined by condition (7.21), is independent of the capacity of road E, while $\tilde{q}_A^{\mathrm{II}}$ may decrease to zero as the increase of capacity y_E. And thus, it is possible that $\bar{q}_A^{\mathrm{II}} > \tilde{q}_A^{\mathrm{II}}$ for certain conditions. However, there is no sense for such profit-maximizing solution in practice. In particular, when the travel demand turns to be inelastic, the profit-maximizing public sector tends to set an infinitely high toll level and an infinitely low capacity level of road A.

We know that, the public sector always selects a combination of toll and capacity level to keep the optimal v/c ratio, $\tilde{\gamma}_A$ at any Pareto-optimal solution, $\left(p_E^{\mathrm{II}}, p_A^{\mathrm{II}}, y_A^{\mathrm{II}}\right)$. It is convenient to investigate the properties of the intermediate Pareto-optimal solution using the welfare and profit functions, given by (7.14) and (7.13), respectively. Denote the flow of road A as q_A^{II}, associated with the intermediate Pareto-optimal solution $\left(p_E^{\mathrm{II}}, p_A^{\mathrm{II}}, y_A^{\mathrm{II}}\right)$, which solves the following programming

$$\max_{q_A \geq 0} W\left(q_A\right), \text{ s.t}, P_A\left(q_A\right) \geq P_A^*, \tag{7.22}$$

where $0 \leq P_A^* \leq -\left(\bar{q}_A^{\mathrm{II}}\right)^2 B'\left(\bar{q}_A^{\mathrm{II}}\right)$. Note that the public sector can freely adjust the toll level of road E. Thus, UE condition (7.8) now have no effect on her decision to obtain profit P_A^*. q_E can be viewed as the function of q_A determined by the binding constraint $P_A\left(q_A\right) = P_A^*$. Using definition (7.11), we have the following Pareto-optimal condition at any intermediate Pareto-optimal solution, $\left(p_E^{\mathrm{II}}, p_A^{\mathrm{II}}, y_A^{\mathrm{II}}\right)$

$$\frac{B\left(Q^{\mathrm{II}}\right) - MSC_A\left(q_A^{\mathrm{II}}\right)}{B\left(Q^{\mathrm{II}}\right) - MSC_E\left(q_E^{\mathrm{II}}\right)} + \frac{B\left(Q^{\mathrm{II}}\right) - C_0}{-q_A^{\mathrm{II}} B'\left(Q^{\mathrm{II}}\right)} = 1 \tag{7.23}$$

where $MSC_i()$ is the marginal cost function

$$MSC_i(q_i) = \beta t_i(q_i, y_i) + \beta q_i \frac{\partial t_i(q_i, y_i)}{\partial q_i}, i \in \{A, E\}. \tag{7.24}$$

In Eq. (7.23), the first term is the ratio between the marginal social cost savings gained from roads E and A; the second term is the ratio of the Pareto-optimal markup level to the maximal markup charge (monopoly pricing) to realize the traffic volume levels q_A^{II} and q_E^{II}. The former captures the inefficient marginal social cost saving, and the latter depicts the degree of the inefficiency of using the user fee. Moving from the welfare-maximizing solution to profit-maximizing one along the Pareto-optimal frontier, the public sector extracts more and more consumers' surplus. And thus, the second term in Eq. (7.23) approaches to one. The second term turns to zero, which means the inefficient marginal social cost saving gained from road E increases more quickly than that from road A.

Furthermore, given the existence of the welfare-maximizing and profit-maximizing solutions, $\tilde{q}_A^{\mathrm{II}}$ and $\bar{q}_A^{\mathrm{II}}$, any Pareto-optimal traffic volume of road A, q_A^{II}, must locate between the two levels. Using condition (7.23), we can determine the corresponding demand level, Q^{II}, and the traffic volume of road E, q_E^{II}. And thus, the Pareto-optimal toll levels of roads A and E are easy to calibrate.

7.3.3 Pareto-optimal solutions when the alternative is a private toll road

In this subsection, we go to investigate the properties of Pareto-optimal solutions of problem (7.7) where road E is a private toll road. We assume a one-shot game between the public sector and private sector, namely, the operators of the two roads, simultaneously, set toll and/or capacity levels, respectively. The public sector aims to obtain a Pareto-efficient outcome of the total social welfare and the net profit gains by selecting a combination of toll and capacity levels of road A. And the private sector chooses the toll level of road E to maximize her own toll revenue. Either party makes her decision considering the action of another guy. Again, from the standard optimality method, for any given toll and capacity pair of road A, $\left(p_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)$, the best response of the revenue-maximizing private sector is to set a toll charge such that

$$p_E^{\mathrm{III}} = \beta q_E^{\mathrm{III}} \frac{\partial t_E\left(q_E^{\mathrm{III}}, y_E\right)}{\partial q_E} + \alpha_2^{\mathrm{III}} \beta q_E^{\mathrm{III}} \frac{\partial t_A\left(q_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)}{\partial q_A}, \tag{7.25}$$

where $\alpha_2^{\mathrm{III}} = B'\left(Q^{\mathrm{III}}\right) / \left(B'\left(Q^{\mathrm{III}}\right) - \beta \partial t_A\left(q_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right) / \partial q_A\right) \in [0,1]$.

We first consider two special cases by assuming that the public sector aims to maximize only the social welfare or net profit of the road project. Denote $\left(\tilde{p}_E^{\mathrm{III}}, \tilde{p}_A^{\mathrm{III}}, \tilde{y}_A^{\mathrm{III}}\right)$ as the competitive equilibrium between the welfare-maximizing public sector and the revenue-maximizing private sector. In this case, the public sector sets the toll level of road A as

$$\tilde{p}_A^{\mathrm{III}} = \beta \tilde{q}_A^{\mathrm{III}} \frac{\partial t_A\left(\tilde{q}_A^{\mathrm{III}}, \tilde{y}_A^{\mathrm{III}}\right)}{\partial q_A} + \tilde{\alpha}_1^{\mathrm{III}} \left(\tilde{p}_E^{\mathrm{III}} - \beta \tilde{q}_E^{\mathrm{III}} \frac{\partial t_E\left(\tilde{q}_E^{\mathrm{III}}, y_E\right)}{\partial q_E}\right), \tag{7.26}$$

where $\tilde{\alpha}_1^{\mathrm{III}} = B'\left(\tilde{Q}^{\mathrm{III}}\right) / \left(B'\left(\tilde{Q}^{\mathrm{III}}\right) - \beta \partial t_E\left(\tilde{q}_E^{\mathrm{III}}, y_E\right) / \partial q_E\right) \in [0,1]$ and $\tilde{Q}^{\mathrm{III}} = \tilde{q}_A^{\mathrm{III}} + \tilde{q}_E^{\mathrm{III}}$. Denote $\left(\bar{p}_E^{\mathrm{III}}, \bar{p}_A^{\mathrm{III}}, \bar{y}_A^{\mathrm{III}}\right)$ as the competitive equilibrium between the profit-maximizing public sector and the revenue-maximizing private sector. In this case, the public sector sets the toll level of road A as

$$\bar{p}_A^{\mathrm{III}} = \beta \bar{q}_A^{\mathrm{III}} \frac{\partial t_A\left(\bar{q}_A^{\mathrm{III}}, \bar{y}_A^{\mathrm{III}}\right)}{\partial q_A} + \bar{\alpha}_1^{\mathrm{III}} \beta \bar{q}_A^{\mathrm{III}} \frac{\partial t_E\left(\bar{q}_E^{\mathrm{III}}, y_E\right)}{\partial q_E}, \tag{7.27}$$

where $\bar{\alpha}_1^{\mathrm{III}} = B'\left(\bar{Q}^{\mathrm{III}}\right) / \left(B'\left(\bar{Q}^{\mathrm{III}}\right) - \beta \partial t_E\left(\bar{q}_E^{\mathrm{III}}, y_E\right) / \partial q_E\right) \in [0,1]$ and $\bar{Q}^{\mathrm{III}} = \bar{q}_A^{\mathrm{III}} + \bar{q}_E^{\mathrm{III}}$. Using competitive equilibrium tolls (7.25)-(7.26) and definition (7.11), UE condition (7.8) can be rewritten as

$$\begin{aligned} B\left(\tilde{q}_A^{\mathrm{III}} + \tilde{q}_E^{\mathrm{III}}\right) &= \beta t_E\left(\tilde{q}_E^{\mathrm{III}}, y_E\right) + \beta \tilde{q}_E^{\mathrm{III}} \frac{\partial t_E\left(\tilde{q}_E^{\mathrm{III}}, y_E\right)}{\partial q_E} + \tilde{\alpha}_2^{\mathrm{III}} \beta \tilde{q}_E^{\mathrm{III}} \frac{\partial t_A\left(\tilde{q}_A^{\mathrm{III}}, \tilde{y}_A^{\mathrm{III}}\right)}{\partial q_A} \\ &= C_0 + \tilde{\alpha}_1^{\mathrm{III}} \tilde{\alpha}_2^{\mathrm{III}} \beta \tilde{q}_E^{\mathrm{III}} \frac{\partial t_A\left(\tilde{q}_A^{\mathrm{III}}, \tilde{y}_A^{\mathrm{III}}\right)}{\partial q_A}, \end{aligned} \tag{7.28}$$

and

$$\begin{aligned} B\left(\bar{q}_A^{\mathrm{III}} + \bar{q}_E^{\mathrm{III}}\right) &= \beta t_E\left(\bar{q}_E^{\mathrm{III}}, y_E\right) + \beta \bar{q}_E^{\mathrm{III}} \frac{\partial t_E\left(\bar{q}_E^{\mathrm{III}}, y_E\right)}{\partial q_E} + \bar{\alpha}_2^{\mathrm{III}} \beta \bar{q}_E^{\mathrm{III}} \frac{\partial t_A\left(\bar{q}_A^{\mathrm{III}}, \bar{y}_A^{\mathrm{III}}\right)}{\partial q_A} \\ &= C_0 + \bar{\alpha}_1^{\mathrm{III}} \beta \bar{q}_A^{\mathrm{III}} \frac{\partial t_E\left(\bar{q}_E^{\mathrm{III}}, y_E\right)}{\partial q_E} \end{aligned} \tag{7.29}$$

The relationship between the welfare-maximizing and profit-maximizing solutions is indeterministic. Generally, the welfare-maximizing public sector is more efficient than the profit-maximizing one does in the sense of social welfare. However, it is not always true. When the capacity of road E is a little higher, the welfare-maximizing public sector has no a leading competitive power to operate the new road, and is less efficient than the public sector to maximize the net profit. For the fixed-capacity case, de Palma and Lindsey (2000), using numerical examples, compared the toll levels and efficiencies under those ownership regimes in which toll is the only instrument for both operators. Very strangely, they demonstrated in their numerical examples that the efficiency of mixed duopoly (public sector vs. private sector) would be lower than the pure duopoly (private sector vs. private sector) when the public road has a small share of capacity. Therefore, if the public road has a lower capacity or the public road performs very badly, it is better for the government to privatize the public toll road or to compete with his rival as a firm. In economic literature, the un-intuitive result was also found by de Fraja and Delbono (1989) in the mixed oligopoly market. If the market is competitive enough and the public sector has no a leading competitive power, it is socially optimal for the public sector to try to maximize its own profit not the total social welfare. They found there is a threshold number of private firms, M, such that privatization of the public sector improves the total social welfare if and only if the number of private firms exceeds M.

We now investigate the properties of the competitive equilibrium solution, $\left(p_E^{\mathrm{III}}, p_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)$, in which, the public sector tends to obtain the tradeoff efficient outcomes by selecting a combination of toll and capacity levels of the new toll road and considering the competitive effect of the private sector. Denote the flow of roads E and A associated to the competitive equilibrium solution as q_E^{III} and q_A^{III}, respectively. At competitive equilibrium, given toll level, p_E^{III}, $\left(p_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)$ is a certain selection of the public sector to achieve an efficient outcomes: either welfare or profit will be decreased for any other combination of p_A^{III} and y_A^{III}; given $\left(p_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)$, p_E^{III} is one of the best strategy of the private sector: any other toll level set by the private sector will not increase the toll revenue obtained from road E. It must be pointed out that, for given p_E^{III}, there may be various selections for the public sector to achieve different efficient outcomes, namely, p_E^{III} may correspond to a set of the efficient payoff pairs. While for given $\left(p_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)$, the private sector may have various optimal strategies, which result in the same payoff level.

On the other hand, for a certain p_E^{III}, if there exists $\left(p_E^{\mathrm{III}}, p_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)$ resulting in a zero flow of road E, $q_E^{\mathrm{III}} = 0$, then the corresponding flow of road A, $q_A^{\mathrm{III}} \in \left[\bar{q}_A^{\mathrm{II}}, \tilde{Q}^{\mathrm{II}}\right]$. The competitive outcome only happens when the free-flow travel time cost of road E is higher than the average travel cost of road A, or, $\beta t_E\left(0, y_E\right) \geq C_0$. Because, first, for $q_E^{\mathrm{III}} = 0$, at any Pareto-optimal solution, the generalized travel cost of road A is not less than the average social cost since $B\left(\tilde{Q}^{\mathrm{II}}\right) = C_0$ and $q_A^{\mathrm{III}} \in \left[\bar{q}_A^{\mathrm{II}}, \tilde{Q}^{\mathrm{II}}\right]$. Second, if $\beta t_E\left(0, y_E\right) < C_0$, the revenue-maximizing private sector surely sets a lower toll level on road E to attract a positive link flow and earn a strictly positive toll revenue. Therefore, the competitive equilibrium solution $\left(p_E^{\mathrm{III}}, p_A^{\mathrm{III}}, y_A^{\mathrm{III}}\right)$ can not result in $q_E^{\mathrm{III}} = 0$ whenever $\beta t_E\left(0, y_E\right) < C_0$. In fact, it is not difficult to make decision for the public sector if condition $\beta t_E\left(0, y_E\right) \geq C_0$ does exist because the private sector has no competitive power in this case. Without loss of generality, we assume that condition $\beta t_E\left(0, y_E\right) < C_0$ is true, which is not practically restrict because the new toll project may not so cheap in the sense of social cost even lower than the free-flow travel time cost of the existing road.

In Subsections 7.3.2, the toll level of the alternative road E can be freely set by the public sector, in which, the flow of road E is a function of q_A determined by the binding profit constraint, and the bi-objective problem can be changed as a single-objective programming problem. Differently, since q_E^{III} is positive under our assumption, and the toll of road E is set by the private sector and can not be changed by the public sector. Problem (7.7) gets more complicated than problems (7.5) and (7.6) discussed in the previous two subsections because of the competitive effect. We adopt the simple weighted sum method to determine the Pareto-optimal solution. The topic on the strategies of the public toll road in parallel roads competition is left for our future research. We assume that the Pareto-optimal solutions associated to p_E^{III} solve the following problem.

$$\max_{q_A \geq 0} \ \lambda W\left(q_A\right) + \left(1 - \lambda\right) P_A\left(q_A\right), \tag{7.30}$$

where the weighted parameter $\lambda \in [0,1]$. Note that, if $\left(p_E^{\text{III}}, p_A^{\text{III}}, y_A^{\text{III}}\right)$ is one of solutions of problem (7.7), then q_A^{III} solves problem (7.30) with UE constraint (7.8) by setting $p_E = p_E^{\text{III}}$. However, not all solutions of problem (7.30), together with p_E^{III}, are the competitive equilibrium ones of problem (7.7). When the capacity level of road E is a little higher, the weighted sum method (7.30) may result in unpredictable outcomes. That is to say, to set a higher weight on the welfare in problem (7.30) may induce less total social welfare gain. However, the weighted sum method can provide various alternative choices for the public sector when both social welfare and profit are their concerns. Consider problem (7.30), from the standard optimality method, we obtain

$$\begin{aligned} &\lambda \left\{ \left(B\left(Q^{\text{III}}\right) - MSC_A\left(q_A^{\text{III}}\right)\right) - \alpha_1^{\text{III}}\left(B\left(Q^{\text{III}}\right) - MSC_E\left(q_E^{\text{III}}\right)\right)\right\} \\ &\qquad + (1-\lambda)\left\{\left(B\left(Q^{\text{III}}\right) - C_0\right) + \left(1-\alpha_1^{\text{III}}\right) q_A^{\text{III}} B'\left(Q^{\text{III}}\right)\right\} = 0, \end{aligned} \tag{7.31}$$

where $\alpha_1^{\text{III}} = B'\left(Q^{\text{III}}\right)/\left(B'\left(Q^{\text{III}}\right) - \beta \partial t_E\left(q_E^{\text{III}}, y_E\right)/\partial q_E\right) \in [0,1]$. Considering the response of the private sector, described as Eq. (7.25), the public sector sets toll

$$\begin{aligned} p_A^{\text{III}} = &\ \beta q_A^{\text{III}} \frac{\partial t_A\left(q_A^{\text{III}}, y_A^{\text{III}}\right)}{\partial q_A} + \lambda \alpha_1^{\text{III}} \left(p_E^{\text{III}} - \beta q_E^{\text{III}} \frac{\partial t_E\left(q_E^{\text{III}}, y_E\right)}{\partial q_E}\right) \\ &+ (1-\lambda)\,\alpha_1^{\text{III}} \beta q_A^{\text{III}} \frac{\partial t_E\left(q_E^{\text{III}}, y_E\right)}{\partial q_E}. \end{aligned} \tag{7.32}$$

Observe Eqs. (7.26) and (7.27), it is a compromising response of the public sector to derive an efficient outcomes. Given weighted parameter λ, substituting toll charges, p_E^{III} and p_A^{III}, determined by Eqs. (7.25) and (7.32), respectively, into equilibrium condition (7.8), we can determine the corresponding Pareto-optimal solution. The toll setting rule (7.32) is also valid by setting $p_E = 0$ when the alternative road is free.

7.4 Pareto-improving Schemes under Various Ownership Regimes

We so far investigated the Pareto-optimal capacity and toll choices of a toll road parallel to an existing alternative road with various ownership regimes. The analysis is based on the assumption that the all Pareto-optimal solution of problems (7.5)-(7.7) are inner solutions. However, several questions are still not clear. Whether is it necessary to introduce road

A, or, whether can it improve the efficiency of the whole network in the sense of social welfare? Is there any Pareto-improving scheme for the public sector to obtain positive gains in both the total social welfare and profit by selecting a combination of capacity and toll levels of road A? In this section, we proceed to examine those questions. If there is a Pareto-improving scheme, we are interested in the non-dominated Pareto-improving scheme, which locates in the Pareto-optimal solution set. Hereinafter, we define the Pareto-improving scheme as the Pareto-optimal solution which results in positive welfare and profit gains, simultaneously.

7.4.1 Pareto-improving schemes when the alternative road is free

When road E is free, before introducing road A, the flow of road E, $\hat{q}_E^{\text{I}}$, satisfies $\beta t_E\left(\hat{q}_E^{\text{I}}, y_E\right) = B\left(\hat{q}_E^{\text{I}}\right)$. Notably, $\hat{q}_E^{\text{I}}$ is larger than the level when setting a toll equal to the congestion externality on road E. From condition (7.17), define the distorted marginal social cost (MSC) of road E as:

$$MSC_E^{\text{I}}\left(q_E\right) = \beta t_E\left(q_E, y_E\right) + \alpha_1 \beta q_E \frac{\partial t_E\left(q_E, y_E\right)}{\partial q_E}, \tag{7.33}$$

where $\alpha_1 = B'\left(q_E\right)/\left(B'\left(q_E\right) - \beta \partial t_E\left(q_E, y_E\right)/\partial q_E\right)$. We now prove that $MSC_E^{\text{I}}\left(\hat{q}_E^{\text{I}}\right) > C_0$ is a sufficient condition to invest road A for the welfare-maximizing public sector. That is to say, under condition $MSC_E^{\text{I}}\left(\hat{q}_E^{\text{I}}\right) > C_0$, the public sector can always improve the social welfare by properly selecting a combination of capacity and toll levels of road A. Suppose that road A is introduced and the public sector chooses combination of (p_A, y_A) such that $y_A = q_A/\tilde{\gamma}_A$, then the total social welfare is given by (7.14). Viewing q_E as a function defined by UE condition $\beta t_E\left(q_E, y_E\right) = B\left(q_A + q_E\right)$ and taking derivative of social welfare W in q_A, we obtain

$$\frac{\text{d}W}{\text{d}q_A} = \beta t_E\left(q_E, y_E\right) + \frac{B'\left(q_A + q_E\right)}{B'\left(q_A + q_E\right) - \beta \partial t_E\left(q_E, y_E\right)/\partial q_E} \beta q_E \frac{\partial t_E\left(q_E, y_E\right)}{\partial q_E} - C_0. \tag{7.34}$$

From Eq. (7.34), we know that $\left(\text{d}W/\text{d}q_A\right)_{q_A=0} = MSC_E^{\text{I}}\left(\hat{q}_E^{\text{I}}\right) - C_0$. We obtain the sufficiency. Furthermore, if welfare function W is concave in q_A, then $\left(\text{d}W/\text{d}q_A\right)_{q_A=0} \leq 0$

implies that increasing capacity investment of road A will not improve the total social welfare. Thus, condition $MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$ is also necessary for the welfare-maximizing public sector to implement a welfare-improving scheme.

In addition, if $ASC\left(\hat{q}_E^{\mathrm{I}}\right) = \beta t_E\left(\hat{q}_E^{\mathrm{I}}, y_E\right) > C_0$, the profit-maximizing public sector can achieve a strictly positive net profit by setting the toll level of road A, given by Eq. (7.16), and the capacity level satisfying condition (7.18). Furthermore, it is intuitive to see that $ASC\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$ is also necessary for the public sector to earn positive net profit by introducing road A. We know that, $ASC\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$ implies $MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$ since $MSC_E^{\mathrm{I}}\left(q_E\right) > ASC\left(q_E\right)$ for any $q_E > 0$. Therefore, there is a Pareto-improving scheme with position welfare and profit gains if and only if $ASC\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$.

In summary, if $ASC\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$, then there exists a Pareto-improving scheme for the public sector to achieve the positive social welfare and profit gains; if $ASC\left(\hat{q}_E^{\mathrm{I}}\right) \leq C_0 < MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right)$, the public sector can improve total social welfare but not obtain a positive profit; the public sector may not improve the total social welfare by introducing road A when $MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) \leq C_0$.

7.4.2 Pareto-improving schemes when the alternative is a public toll road

When road A is a public toll road, the sufficient and necessary condition for the public sector to invest road A is $MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0$, where MSC_E is the marginal social cost of road E, given by Eq. (7.24), and $\hat{q}_E^{\mathrm{II}}$ is the socially optimal flow of road E, defined by $MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) = B\left(\hat{q}_E^{\mathrm{II}}\right)$. Condition $MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0$ means that the existing road does not satisfy all demand at the socially optimal toll, the public sector can invest road A to improve the total social welfare. In addition, for any road capacity, $y_A > 0$, when $q_A \to 0$, the average travel time approaches the free-flow travel time, $t_A(0)$, and the congestion externality approaches zero. Therefore, the monopoly profit $P_A\left(\bar{p}_E^{\mathrm{II}}, \bar{p}_A^{\mathrm{II}}, \bar{y}_A^{\mathrm{II}}\right) \geq 0$ is guaranteed by condition (3.50). As discussed in Chapter 3, condition (3.50) is not practically restrictive, because we can reasonably expect that there is a positive potential traffic demand with a free-flow travel time. Otherwise, it is meaningless to build a new highway. Therefore, the public sector can always gain positive profit if the alternative is a public toll road.

It is clear that, when road E is a public toll road, there exists a Pareto-improving scheme for the public sector to achieve positive social welfare and net profit gains if $MSC_E\left(\hat{q}_E^{\text{II}}\right) > C_0$; if $MSC_E\left(\hat{q}_E^{\text{II}}\right) \leq C_0$, the total social welfare can not be improved even the public sector can obtain positive net profit by introducing road A.

7.4.3 Pareto-improving schemes when the alternative is a private toll road

We now examine the existence of the Pareto-improving schemes when the alternative road is a private toll road. Note that, from UE condition (7.28), at the critical case with zero traffic flow on road A, $\alpha_2^{\text{III}} = 0$. Thus, the following distorted MSC function of road E is considered:

$$MSC_E^{\text{III}}\left(q_E\right) = \beta t_E\left(q_E, y_E\right) + \beta q_E \frac{\partial t_E\left(q_E, y_E\right)}{\partial q_E} + \alpha_1^{\text{III}} \beta q_E \frac{\partial t_E\left(q_E, y_E\right)}{\partial q_E}, \tag{7.35}$$

where $\alpha_1^{\text{III}} = B'\left(q_E\right) / \left(B'\left(q_E\right) - \beta \partial t_E\left(q_E, y_E\right) / \partial q_E\right) \in [0,1]$. It is clear that, the unregulated private sector of road E surely sets a monopoly toll level whenever road A is not introduced into the network. The resulted link flow, denoted as, $\hat{q}_E^{\text{III}}$, satisfies

$$\beta t_E\left(\hat{q}_E^{\text{III}}, y_E\right) + \beta \hat{q}_E^{\text{III}} \frac{\partial t_E\left(\hat{q}_E^{\text{III}}, y_E\right)}{\partial q_E} = B\left(\hat{q}_E^{\text{III}}\right) + \hat{q}_E^{\text{III}} B'\left(\hat{q}_E^{\text{III}}\right). \tag{7.36}$$

We now prove that if $MSC_E^{\text{III}}\left(\hat{q}_E^{\text{III}}\right) > C_0$, then the existing road does not satisfy all demand at the monopoly toll and the public sector can invest road A to improve the total social welfare. First, suppose that the private sector keeps the monopoly toll of road E, $p_E = B\left(\hat{q}_E^{\text{III}}\right) - \beta t_E\left(\hat{q}_E^{\text{III}}, y_E\right)$. In this case, since the optimal capacity and toll choice of the public sector results in the socially optimal v/c ratio, $\tilde{\gamma}_A$, we consider welfare function, given by (7.14). Taking the derivative of W with respect to q_A, we obtain

$$\frac{\mathrm{d}W}{\mathrm{d}q_A} = B\left(q_A + q_E\right) - C_0 + \left(B\left(q_A + q_E\right) - \beta t_E\left(q_E, y_E\right) - \beta q_E \frac{\partial t_E\left(q_E, y_E\right)}{\partial q_E} \right) \frac{\mathrm{d}q_E}{\mathrm{d}q_A}. \quad (7.37)$$

Using UE condition (7.8) by setting $p_E = B\left(\hat{q}_E^{\mathrm{III}}\right) - \beta t_E\left(\hat{q}_E^{\mathrm{III}}, y_E\right)$, we have, at $q_A = 0$,

$$\left.\frac{\mathrm{d}W}{\mathrm{d}q_A}\right|_{q_A=0, q_E=\hat{q}_E^{\mathrm{III}}} = MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) - C_0. \quad (7.38)$$

Equation (7.38) implies that, if $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$ and the private sector keeps the monopoly toll, then the public sector can improve the total social welfare by introducing road A. When the private sector increases or decreases the toll level of road E following a strategy action, it is clear that the welfare-maximizing solution with strictly positive capacity level does exist and satisfying equilibrium condition (7.28). Therefore, condition $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$ is sufficient to guarantee that the total social welfare can be improved by properly choosing the combination of capacity and toll levels. Furthermore, if the social welfare W is concave in q_A, then increasing capacity investment of road A will not improve the total social welfare since $\left(\mathrm{d}W/\mathrm{d}q_A\right)_{q_A=0} \leq 0$. Therefore, condition $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$ is also necessary.

It must be pointed out that, condition $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$ is sufficient and necessary for the welfare-maximizing public sector to build road A, but which does not mean that the public sector has no other alternative strategy to improve the social welfare when the condition is violated. Notably, if $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) \leq C_0 < B\left(\hat{q}_E^{\mathrm{III}}\right)$, road E can not satisfy the monopoly demand, and the profit-maximizing public sector can obtain positive profit by introducing road A. And if the competition, simultaneously, results in a positive welfare gain comparing to that in the monopoly case (the unregulated private sector operates the single road and sets a monopoly toll level), then a Pareto-improving scheme exists. It is in-deterministic that whether the total social welfare is surely enhanced under this condition. However, the consumer surplus is surely improved since the total demand is increased (Farrell and Shapiro, 1990). In this case, we still assume that the public sector prefers to introduce road A. Nevertheless, the private sector obtains the monopoly revenue if the generalized travel cost of road E is lower than C_0 at $\hat{q}_E^{\mathrm{III}}$, namely, $B\left(\hat{q}_E^{\mathrm{III}}\right) \leq C_0$. And thus, the public sector never achieves a positive net profit gain. There is no Pareto-improving scheme.

According to the above discussion, we know that, if $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$, then there exists a Pareto-improving scheme for the public sector to achieve positive social welfare and profit gains; if $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) \leq C_0 < B\left(\hat{q}_E^{\mathrm{III}}\right)$, the total social welfare may be improved and the public sector can obtain positive net profit gain by introducing road A; while $B\left(\hat{q}_E^{\mathrm{III}}\right) \leq C_0$, the public sector can never improve social welfare with the positive profit gain by building road A, and the private sector achieves monopoly revenue from road E.

7.5 Effects of Ownership Regimes

We so far investigated the Pareto-optimal solution sets of problems (7.5)-(7.7) and the existence of the Pareto-improving schemes under various ownership regimes, separately. One of the most important issues is how the ownership regimes affect the decisions of the welfare- and profit-maximizing public sector. In this section, we proceed to examine the following questions: how do the ownership regimes affect the capacity choices and incentive to investment a new road for the public sector? what about the existence of Pareto-improving schemes under various ownership regimes?

7.5.1 Effects of network ownership regimes on Pareto-optimal solution sets

The effects of the ownership regimes of the alternative road on the Pareto-optimal solution set are investigated in this subsection. We are particularly interested in the capacity choices of road A under different ownership regimes for the welfare-maximizing public sector.

From UE conditions (7.17) and (7.20), we know that $\tilde{q}_E^{\mathrm{I}} > \tilde{q}_E^{\mathrm{II}}$ and $\tilde{Q}^{\mathrm{I}} > \tilde{Q}^{\mathrm{II}}$. By direct calibration, we obtain

$$\tilde{Q}^{\mathrm{I}} - \tilde{Q}^{\mathrm{II}} = \left(1 - \tilde{\alpha}_1^{\mathrm{I}}\right) \tilde{q}_E^{\mathrm{I}} \frac{E_{sh}{}_{\tilde{Q}^{\mathrm{I}}}^{B}}{E_{\tilde{Q}^{\mathrm{I}}}^{B}}, \tag{7.39}$$

and

$$\tilde{q}_E^{\mathrm{I}} - \tilde{q}_E^{\mathrm{II}} = \left(1 - \tilde{\alpha}_1^{\mathrm{I}}\right) \tilde{q}_E^{\mathrm{I}} \frac{\beta \partial t_E \left(\tilde{q}_E^{\mathrm{I}}, y_E\right) / \partial q_E}{\partial \left(MSC_E \left(\xi_1\right)\right) / \partial q_E} \tag{7.40}$$

where $\xi_1 \in \left(\tilde{q}_E^{\mathrm{II}}, \tilde{q}_E^{\mathrm{III}}\right)$; MSC_E is the marginal social cost defined by Eq. (7.24); the point and shrinkage-ratio price elasticities of demand function, $B\left(Q\right)$ are defined as Eqs. (3.33) and (3.37), respectively. Subtracting both sides of (7.39) from (7.40) gives rise to

$$\tilde{q}_A^{\mathrm{I}} - \tilde{q}_A^{\mathrm{II}} = \left(1 - \tilde{\alpha}_1^{\mathrm{I}}\right) \tilde{q}_E^{\mathrm{I}} \left(\frac{E_{sh}{}_{\tilde{Q}^{\mathrm{I}}}^{B}}{E_{\tilde{Q}^{\mathrm{I}}}^{B}} - \frac{\beta \partial t_E \left(\tilde{q}_E^{\mathrm{I}}, y_E\right) / \partial q_E}{\partial \left(MSC_E \left(\xi_1\right)\right) / \partial q_E} \right). \tag{7.41}$$

Two sufficient conditions are derived to guarantee $\tilde{q}_A^{\mathrm{I}} \geq \tilde{q}_A^{\mathrm{II}}$ and $\tilde{q}_A^{\mathrm{I}} \leq \tilde{q}_A^{\mathrm{II}}$, respectively,

$$\frac{E_{sh}{}_{\tilde{Q}^{\mathrm{I}}}^{B}}{E_{\tilde{Q}^{\mathrm{I}}}^{B}} \geq \frac{\beta \partial t_E \left(\tilde{q}_E^{\mathrm{I}}, y_E\right) / \partial q_E}{\partial \left(MSC_E \left(\tilde{q}_E^{\mathrm{II}}\right)\right) / \partial q_E} \tag{7.42}$$

and

$$\frac{E_{sh}{}_{\tilde{Q}^{\mathrm{I}}}^{B}}{E_{\tilde{Q}^{\mathrm{I}}}^{B}} \leq \frac{\beta \partial t_E \left(\tilde{q}_E^{\mathrm{I}}, y_E\right) / \partial q_E}{\partial \left(MSC_E \left(\tilde{q}_E^{\mathrm{I}}\right)\right) / \partial q_E} \tag{7.43}$$

Condition (7.42) is satisfied for the line benefit functions and BPR-type link travel time functions. Therefore, for the BPR-type link travel time functions, condition (7.42) is surely true for the concave benefit functions. However, if condition (7.43) is satisfied, then we derive the opposite conclusions. Furthermore, it is easy to compare the capacity levels under scenarios I and II with conditions (7.42) and (7.43) since the v/c ratio is identical and equal to $\tilde{\gamma}_A$.

Following similar method, from UE conditions (7.20) and (7.28), we obtain the following two sufficient conditions to guarantee $\tilde{q}_A^{\mathrm{II}} \geq \tilde{q}_A^{\mathrm{III}}$ and $\tilde{q}_A^{\mathrm{II}} \leq \tilde{q}_A^{\mathrm{III}}$, respectively

$$\frac{E_{sh}{}^{B}_{\tilde{Q}^{\mathrm{III}}}}{E^{B}_{\tilde{Q}^{\mathrm{III}}}} \geq \frac{\beta \partial t_E\left(\tilde{q}_E^{\mathrm{III}}, y_E\right)/\partial q_E}{\partial\left(MSC_E\left(\tilde{q}_E^{\mathrm{III}}\right)\right)/\partial q_E} \tag{7.44}$$

and

$$\frac{E_{sh}{}^{B}_{\tilde{Q}^{\mathrm{III}}}}{E^{B}_{\tilde{Q}^{\mathrm{III}}}} \leq \frac{\beta \partial t_E\left(\tilde{q}_E^{\mathrm{III}}, y_E\right)/\partial q_E}{\partial\left(MSC_E\left(\tilde{q}_E^{\mathrm{II}}\right)\right)/\partial q_E}. \tag{7.45}$$

The conditions are sufficient because $\tilde{q}_E^{\mathrm{II}} \geq \tilde{q}_E^{\mathrm{III}}$. Again, for the linear or convex benefit functions and BPR-type travel time functions, condition (7.44) is true. Since $\tilde{Q}^{\mathrm{III}} \leq \tilde{Q}^{\mathrm{II}} \leq \tilde{Q}^{\mathrm{I}}$, we know that, a general concave or convex benefit functions may not guarantee conditions (7.44) and (7.42) are satisfied, simultaneously.

In summary, for the special case with the linear or constant-elasticity benefit functions and BPR-type link travel time functions, we know that $\tilde{q}_A^{\mathrm{I}} > \tilde{q}_A^{\mathrm{II}} > \tilde{q}_A^{\mathrm{III}}$ and $\tilde{y}_A^{\mathrm{I}} > \tilde{y}_A^{\mathrm{II}} > \tilde{y}_A^{\mathrm{III}}$. Therefore, in contrast to the scenario II (road E is a public toll road), the welfare-maximizing public sector tends to select a higher capacity level under scenario I (road E is free); while prefers to select a lower capacity level under scenario I (when road E is a private toll road). For this special case, Figure 7.2 depicts the Pareto-optimal demand-capacity pair set under scenario I, Ω. From the above discussion, we know that, the welfare-maximizing demand-capacity pairs when road E is a public or private toll road, $\left(\tilde{q}_A^{\mathrm{II}}, \tilde{y}_A^{\mathrm{II}}\right)$ and $\left(\tilde{q}_A^{\mathrm{III}}, \tilde{y}_A^{\mathrm{III}}\right)$, are two special points in domain Ω. However, it is hard to compare the capacity levels of road A selected by the profit-maximizing public sectors under the three scenarios, even with the special assumptions of the benefit and travel delay functions. In addition, for the case with fixed demand, the price elasticity of demand can be viewed as zero or $B' = \infty$, as a result, $\alpha_1 = \alpha_2 = 1$ for all three scenarios in conditions (7.17), (7.20) and (7.28). The welfare-maximizing public sector selects an identical capacity level, and thus, the capacity decision of the public sector is free from the effect of the ownership regimes of road E.

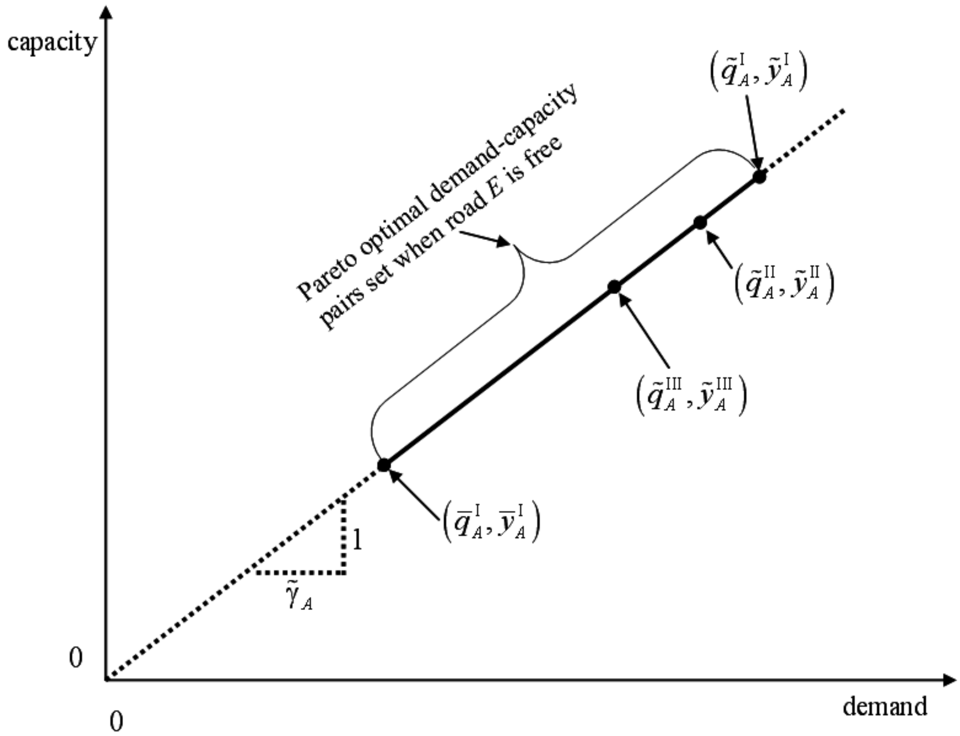


Figure 7.2: Pareto-optimal demand-capacity pairs under various ownership regimes

7.5.2 Effects of network ownership regimes on Pareto-improving schemes

The ownership regimes do have significant effect on the efficient outcomes of investing a new competitive toll road project parallel to an existing alternative, as discussed in Section 7.4. Under different ownership regimes, Table 7.1 lists the conditions for the public sector to achieve positive welfare and/or positive profit gains by investing road A. It must be pointed out that, under condition $B\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$, the profit-maximizing public sector tends to build road A to make money. In this case, the competition may lower the monopoly revenue of the private sector, and the total social welfare may be worsened. However, the total consumer' surplus is surely increased. We assume that the public sector still prefers to build road A to compete with the private sector, who is a natural monopoly.

For the special case with the linear benefit functions and BPR-type link travel time functions, proposed in Subsection 7.5.1, we now prove that

$$MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0 \Rightarrow MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0 \Rightarrow MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) > C_0. \tag{7.46}$$

Table 7.1: Conditions and outcomes under various ownership regimes

Ownership regime of road E	Profit-maximizing	Welfare-maximizing	Pareto-improving	Non-positive profit and welfare gains
Free (I)	$ASC\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$	$MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$	$ASC\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$	$MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) \leq C_0$
Public toll road (II)		$MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0$	$MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0$	
Private toll road (III)	$B\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$	$MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$	$B\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$	$B\left(\hat{q}_E^{\mathrm{III}}\right) \leq C_0$

Suppose that $MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) < MSC_E\left(\hat{q}_E^{\mathrm{II}}\right)$, following the discussion in Subsection 7.5.1, we obtain

$$\hat{q}_E^{\mathrm{I}} - \hat{q}_E^{\mathrm{II}} < \left(1 - \hat{\alpha}_1^{\mathrm{I}}\right) \hat{q}_E^{\mathrm{I}} \frac{\beta \partial t_E\left(\hat{q}_E^{\mathrm{I}}, y_E\right)/\partial q_E}{\partial\left(MSC_E\left(\hat{q}_E^{\mathrm{II}}\right)\right)/\partial q_E}. \tag{7.47}$$

However, from the definitions of $\hat{q}_E^{\mathrm{II}}$ and $\hat{q}_E^{\mathrm{I}}$, we also have

$$\hat{q}_E^{\mathrm{I}} - \hat{q}_E^{\mathrm{II}} > \left(1 - \hat{\alpha}_1^{\mathrm{I}}\right) \hat{q}_E^{\mathrm{I}} \frac{E_{sh}{}_{\hat{q}_E^{\mathrm{I}}}^{B}}{E_{\hat{q}_E^{\mathrm{I}}}^{B}}. \tag{7.48}$$

It is easy to check that, condition (7.42) is true at $\hat{q}_E^{\mathrm{I}}$. Therefore, $MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) \geq MSC_E\left(\hat{q}_E^{\mathrm{II}}\right)$, and thus, $MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0 \Rightarrow MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$. Similarly, we can also derive that $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$ implies $MSC_E\left(\hat{q}_E^{\mathrm{I}}\right) > C_0$ for the special case.

Relationship (7.46) implies that, in comparison with scenario II, the welfare-maximizing public sector has more incentive to invest the new road project when the existing road is free, and less incentive to invest the new project when the existing road is a private toll road. However, when the demand is inelastic, $B' = -\infty$, $MSC_E^{\mathrm{I}}() = MSC_E^{\mathrm{III}}() = MSC_E()$. Therefore, the conditions in Eq. (7.46) are equivalent to each other and the ownership regimes of road E have no impact on the incentive of the welfare-maximizing public sector to invest road A.

Furthermore, it is easy to check the following relationship

$$ASC\left(\hat{q}_E^{\mathrm{I}}\right) > C_0 \Rightarrow MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0 \Rightarrow B\left(\hat{q}_E^{\mathrm{III}}\right) > C_0. \tag{7.49}$$

Relationship (7.49) means that any Pareto-improving scheme under scenario I must be one of the Pareto-improving schemes under scenario II; and any Pareto-improving scheme under scenario II must be one of the Pareto-improving schemes under scenario III. It is more restrictive for the public sector to achieve a Pareto-improving outcomes when the existing road is free, and, generally, less restrictive to achieve a Pareto-improving outcomes when the existing road is a private toll road.

7.5.3 A simple analytical example

Consider the following numerical example: the Benefit function $B(Q) = 5 - 4Q$; the link travel time function is used in both roads with same free-flow travel time, $t_0 = 1.0$ (h), $t_i(q_i, y_i) = 1 + q_i/y_i$, $i \in \{E, A\}$; and construction cost function $I(y_A) = y_A$. It is assumed that the value-of-time, $\beta = 1$, and unit construction cost, $\eta = 1$. The socially optimal volume-capacity ratio and the average social cost of road A are, $\tilde{\gamma}_A = 1.0$ and $C_0 = 3$. The capacity of road E, y_E, as a variable, is adopted to identify the various investment conditions. According to the definitions of $\hat{q}_E^{\mathrm{I}}$, $\hat{q}_E^{\mathrm{II}}$ and $\hat{q}_E^{\mathrm{III}}$, we get that $\hat{q}_E^{\mathrm{I}} = 4/(4 + 1/y_E)$, $\hat{q}_E^{\mathrm{II}} = 4/(4 + 2/y_E)$, $\hat{q}_E^{\mathrm{III}} = 4/(8 + 2/y_E)$. From the conditions discussed in Sections 7.4.1-7.4.3, we know that, when the capacity level of the existing road is lower than $\left(1 + \sqrt{2}\right)/4$, $1/2$ or $\left(1 + \sqrt{5}\right)/8$, the public sector considers to invest the new road under scenarios I, II or III, respectively.

Figures 7.3-7.5 depict the change of the Pareto-optimal frontier with various capacity levels and different ownership regimes of road E. The ownership regimes of road E have significant effects on the Pareto-optimal outcomes and domains of Pareto-improving schemes. As shown in Figure 7.3, when road E is free, the public sector can obtain positive social welfare and profit gains at any Pareto-optimal solutions whenever $0 \leq y_E < 1/4$, or equivalently, $ASC\left(\hat{q}_{\mathrm{E}}^{\mathrm{I}}\right) > \mathrm{C}_0$; while the public sector can still enhance the total social welfare with a negative profit gain whenever $1/4 \leq y_E < \left(1 + \sqrt{2}\right)/4$, or equivalently, $ASC\left(\hat{q}_E^{\mathrm{I}}\right) \leq C_0 < MSC_E^{\mathrm{I}}\left(\hat{q}_E^{\mathrm{I}}\right)$; Any combination of toll and capacity levels of road A will result in a loss both in the welfare and profit when $y_E > \left(1 + \sqrt{2}\right)/4$. In contrast,

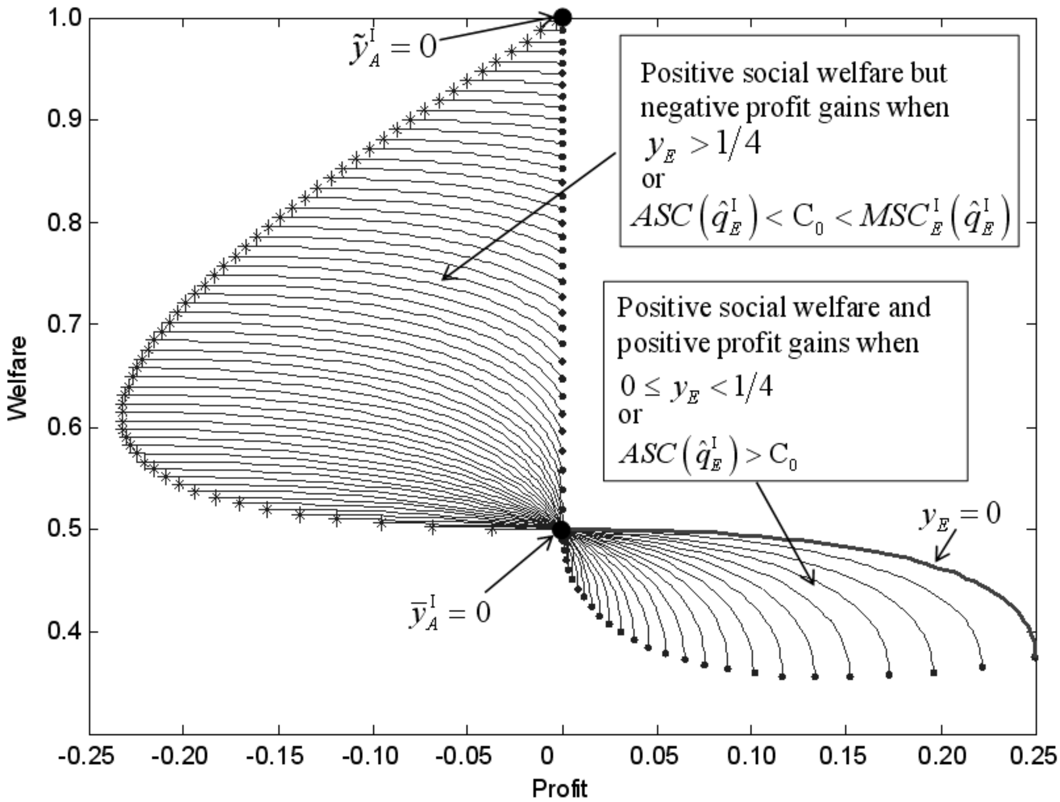


Figure 7.3: Pareto-improving scheme when road E is free

when road E is a public toll road, shown in Figure 7.4, the Pareto-improving scheme always exists whenever $0 \leq y_E < 1/2$, or equivalently, $MSC_E\left(\hat{q}_E^{\mathrm{II}}\right) > C_0$. Beyond this domain, the social welfare will never be improved. It is more complicated when road E is a private sector, as revealed in Figure 7.5. As we have discussed in Subsection 7.4.3, when $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) > C_0$, the public sector can enhance the total social welfare and obtain positive profit return. Although the welfare-maximizing solution does not exist when $y_E > \left(1+\sqrt{5}\right)/8$, or equivalently, $MSC_E^{\mathrm{III}}\left(\hat{q}_E^{\mathrm{III}}\right) < C_0$, there are the Pareto-improving schemes that enhance the total social welfare with a positive profit gain. We also see from Figure 7.5 that, when $y_E > 0.255$, depicted by point A, the profit-maximizing public sector performs better than the welfare-maximizing public sector in the sense of social welfare. In this case, the public sector should act as a private sector to maximize the profit when facing the competition of the private sector of road E.

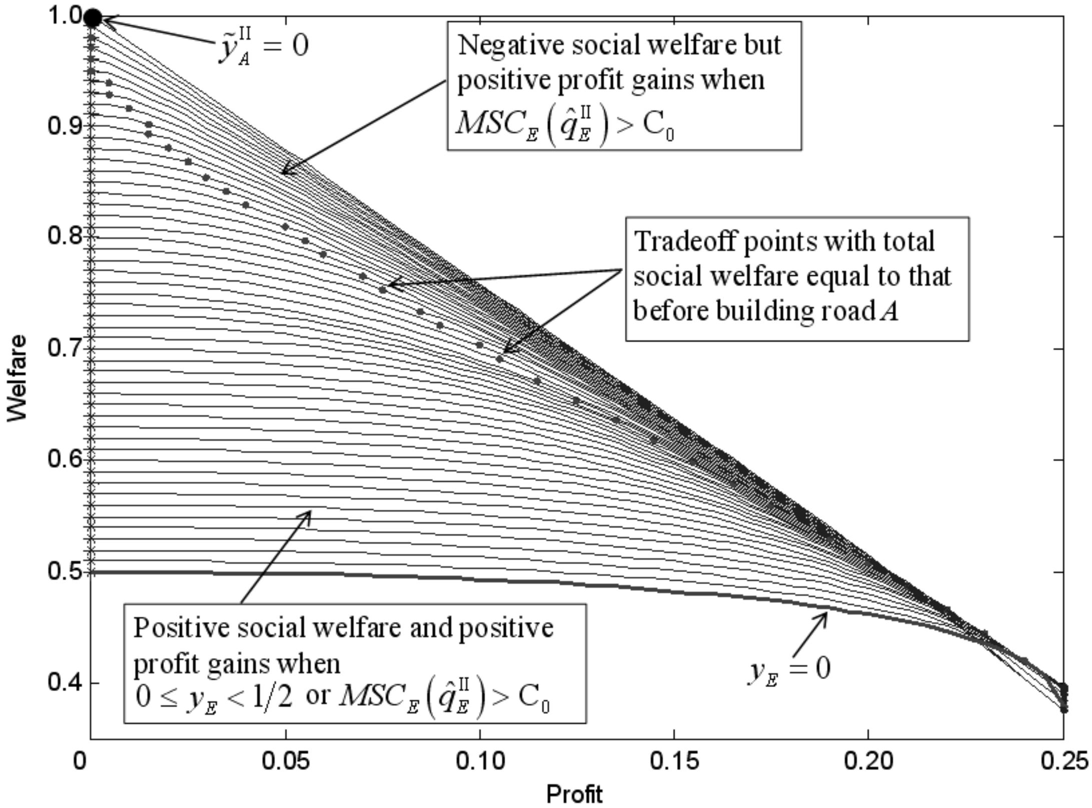


Figure 7.4: Pareto-improving scheme when road E is a public toll road

7.6 Conclusions

An important issue is how and when to introduce a new "add-on" road into an existing transportation network. This chapter investigated the choices of the toll and capacity for a new toll road added to an existing one-link transportation network with various ownership regimes: the existing road can be free, public or private toll road. When the alternative road is a private toll road, we assume that the operators of the two roads simultaneously set toll and/or capacity levels, respectively, following the one-shot game mechanism. The public sector is assumed to consider, simultaneously, the total social welfare and net profit of the new toll road project.

We investigated the properties of the Pareto-optimal solution sets under different ownership regimes and investigated the existence of the Pareto-improving schemes under different ownership regimes, which can enhance the total social welfare with a positive

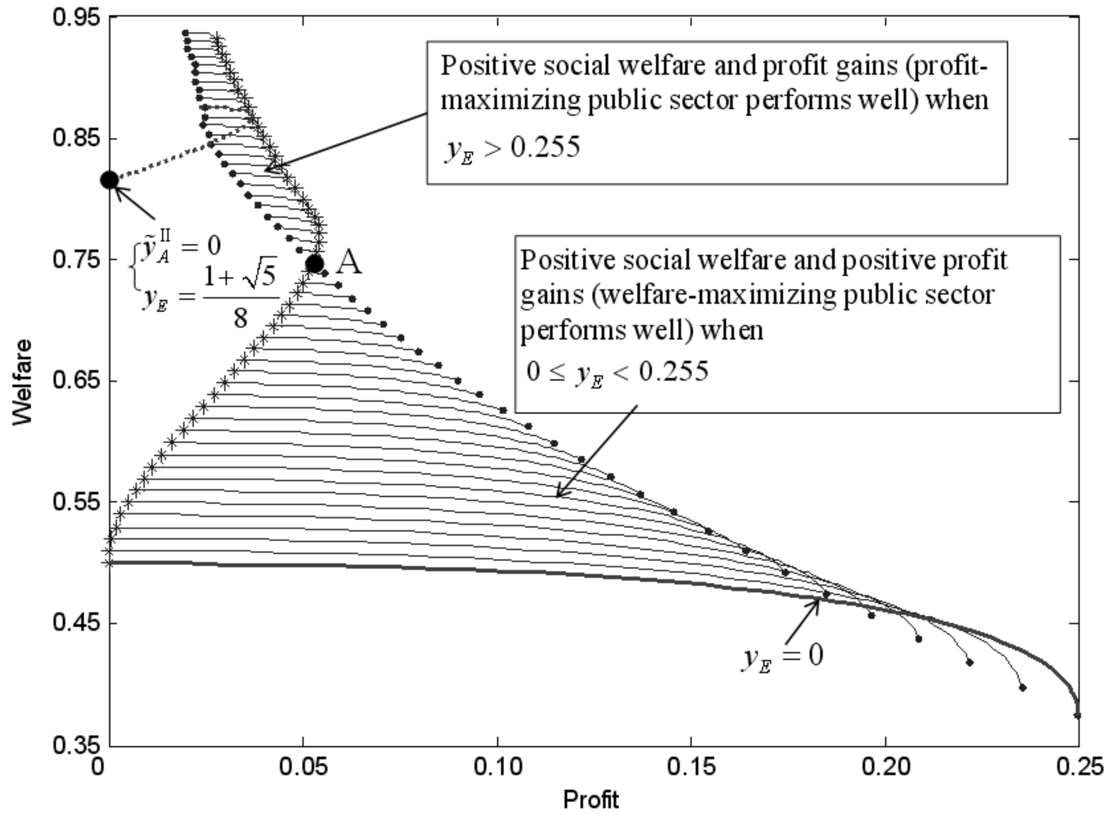


Figure 7.5: Pareto-improving schemes when road E is a private toll road

profit gain. The effects of the ownership regimes on the Pareto-optimal solution sets and existence of the Pareto-improving schemes were also investigated in this chapter. For the special case, such as, linear and constant-elasticity benefit function and BPR (Bureau of Public Roads) type link travel time function, we found that, in comparison with the first-best environment, where both roads are public toll roads, a welfare-maximizing public sector has more incentive to invest the new road project and selects a higher capacity level when the existing road is free, and less incentive to invest the new project and selects a lower capacity level when the existing road is a private toll road; it is more restrictive for the public sector to achieve a Pareto-improving outcomes when the existing road is free, and, generally, less restrictive to achieve a Pareto-improving outcomes when the existing road is a private toll road.

8

MAJOR FINDINGS AND FUTURE RESEARCH

This chapter summarizes the major contributions of this thesis, and discusses some unsolved problems and possible extensions.

8.1 Major Findings

This thesis contributes to road franchising studies in the following aspects.

For a highway franchising contract, the concession period, capacity and toll charge are three primary variables, which together determine the social welfare for the whole society during the whole life of the road and the profit of the private firm during the concession period. This thesis built a bi-objective programming problem to model the different objectives of the participants, the public and private sectors under deterministic demand and homogeneous trip-makers. The important concept associated with the bi-objective programming problem, the Pareto-efficient BOT contract, is proposed to capture the outcomes of the highway project. We analyzed the properties of Pareto-efficient BOT contracts and established several key results on the optimal concession period, service quality and the average social cost per trip for any Pareto-efficient BOT contract. First, any Pareto-efficient BOT contract requires that the concession period should be the whole life of the road. Secondly, with constant returns to scale, the volume-capacity ratio and thus the service quality at any Pareto-efficient BOT contract coincide with the socially optimal levels; and the average social cost per trip is also constant along the Pareto-optimal frontier. A variety of government regulatory regimes were also investigated. Following classic results in economic regulations, we discussed the outcomes of both price-cap and rate-of-return regulations, and proved that: the private firm tends to offer a lower road capacity and a lower service quality under the price-cap regulation, while it chooses a higher service quality, a higher capacity and a higher toll charge under the rate-of-return regulation than those under the corresponding Pareto-efficient solution. The road capacity regulation is also inefficient. In contrast, we proved that both the demand and markup charge regulations lead to Pareto-optimal outcomes.

To make a more practical application, this thesis examined the effects of user heterogeneity on properties of the Pareto-efficient BOT contracts and the outcomes of various regulatory regimes. The user heterogeneity captured by the value-of-time is a crucial factor in determining toll scheme and highway investment. We introduced two important concepts-failure rate and mean residual VOT functions-to develop a theoretically tractable method for the bi-objective programming problem with user VOT heterogeneity. Under mild conditions commonly adopted in economics and revenue management literature, we proved that the Pareto-optimal road-life concession period is free from the effects of user heterogeneity, while the service quality (or, the volume-capacity ratio) is dependent on the curvature of the mean residual VOT function. If the VOT distribution has a(n) convex (affine, concave) mean residual VOT function, then the service quality is increasing (identical, decreasing) along Pareto-optimal frontier from social optimal solution to monopoly optimal solution. We also investigated the effects of the user heterogeneity on the various regulatory regimes. The user heterogeneity does have significant effects on the efficiency of the regulations. The demand and markup regulations, which are efficient to achieve any pre-determined Pareto-efficient BOT contract in homogeneous case, are failure to achieve efficient outcome as the price-cap and rate-of-return regulations. Both two regimes result in a lower investment and lower service quality. Furthermore, in contrast to the demand regulation, the markup regulation results in a higher demand level and a higher welfare gain.

Another extensive study is to consider the demand uncertainty in the highway franchising contract. It is impossible to achieve the Pareto-efficient BOT contract in the incomplete environment. We adopted the principal-agent problem to model the BOT problem, in which, the public sector (the principal) maximizes the total social welfare with the participant constraint of the private sector (the agent). The complexity of the market and unexpected return requirement of the private sector would cause the exhaustion of user fee. Therefore, in our model, the public sector is allowed to participate the project using public funds associated with an additional marginal cost. That is to say, the public sector can compensate the private sector using subsidy to guarantee the interest of road users. Following a benchmark study under perfect information, we found that the Pareto-efficient contracts are not always desirable for the public sector because, along the Pareto-optimal frontier, the transfer ratio of the marginal consumer surplus to marginal revenue in demand (the monetary consumer surplus is required to equalize one HK$ revenue) tends to exceed the marginal cost of the public funds. That is to say, given the marginal cost of public funds, a unique critical minimum attractive rate of return (MARR) was derived. The subsidy is not needed if the MARR does not exceed the critical value; otherwise, a positive subsidy is desirable to increase the total weighted social welfare. In the presence of demand uncertainty, we proposed the full and partial flexibility of the BOT contract according to the instruments adopted by the public and private sectors, namely, a postponement strategy is adopted in a two-period decision problem. Full flexibility refers to the case in which the public sector promises an exogenous rate of return on the private investment and in turn can freely ex post adjust the contract in a socially optimal manner according to the observed demand curve. Partial flexibility refers to the case where the public and private sectors agree on an ex ante demand risk allocation by contract and the ex post contract adjustment can be made contingent on a Pareto-improvement to both parties. A preferred Pareto-improvement can be selected from the Pareto-optimal solution set of a bi-objective programming problem equipped with a rational preference. Based on our two-period flexible BOT model, the optimal concession period should be set as long as possible even when the demand is uncertain. Based on the degree of contract flexibility (full or partial), according to the pre-determined capacity and/or toll levels, we explicitly partition the demand state space in to several domains presenting the high, intermediate and low demand states. When the demand occurs in the different domains, the corresponding strategies adopted by the public and private sectors to improve their objective(s). In comparison with the traditional rigid contract, in which, private sector burdens all project risk, for the contract with full flexibility, the public sector bears all risk, but, in return has full flexibility to ex post adjust the contract variables to achieve the desirable outcomes; while for the contract with partial flexibility, both parties share the demand risk, and thus, the Pareto-improving ex post adjustment is required. The flexibility of the BOT contract is valuable for contract adjustment mechanism to solve the discrepancy between the public and private sectors and the further theoretical research.

To incorporate the road deterioration and maintenance effects into the BOT problem, this thesis considered the dynamic BOT problem by assuming that the maintenance cost to sustain the road quality is a natural function of road capacity, traffic load and calendar time. In this case, the road life is endogenously determined as the time when the net economic benefit becomes nil as maintenance cost increases. The BOT problem is modeled as an iso-perimetric problem to maximize the social welfare with a profit constraint. As the effect of the increasing maintenance cost over time, the private sector tends to set a shorter concession period than the public sector does. The proposed model can be used by the government to determine the optimal concession period, that is proved to be longer than preferred by the private sector. We also proved that the optimal dynamic toll charge is increasing over time, as deemed necessary for reducing traffic load and thereby lowering down the road damage. Specially, if the marginal user damage on road is independent of time, or each vehicle damages an old and new roads in a same way, then the optimal toll charge is free from the effect of road natural deterioration and thus would be time-invariant. With the effect of the increasing maintenance cost, the minimal concession term should be imposed, together with the time-varying the demand regulation regime, to achieve the optimal BOT contract.

As the increasing of the private provision of the public roads, specially, by highway franchising contract, some links are owned by the private sectors and the transportation network has a mixed ownership regime. The private and public toll roads co-exist, together with the free state-owned roads. It gets more complicated when introducing some new links or expanding the capacities of some existing links. The thesis finally addressed this topic and investigated the effects of the network ownership regimes on toll and capacity choices of a new link added to an existing one-link transportation network. The existing road can be free, public or private toll road. The public sector is assumed to consider, simultaneously, the total social welfare and net profit of the new toll road project. When the alternative road is a private toll road, we assume that the operators of the two roads simultaneously set toll and/or capacity levels, respectively, following the one-shot game mechanism. We investigated the properties of the Pareto-optimal solution sets under different ownership regimes, and analyzed the effects of the ownership regimes on those properties. This thesis also investigated the existence of the Pareto improving schemes under different ownership regimes, which can enhance the total social welfare with a positive profit gain.

8.2 Future Research

Though we have explored many important issues of highway franchising, those issues are indeed still need further studies, both in practical and theoretical aspects. It is valuable to highlight those unsolved problems and issues for possible future extensions research.

The concession term is one of the important variables in setting highway franchising contract between the public and private sectors. The fixed-term contracts are subject to frequent renegotiations because of unforeseen contingencies (Guasch, 2004). The economic solution proposed by Engel et al. (1997, 2008), is to set a flexible-term contract, in which, the public road service is offered without any public expense and the contract lasts till the private sector obtains his reserved return. As assumed in Chapter 5, the demand uncertainty is "static", namely, the unknown demand can be realized or observed after the highway is built. Under this assumption, the optimal concession period is easy to determine. If the demand uncertainty is, more practically, "infinite", namely, both parties never reveal the true demand or demand curve (Bain, 2009), the real option theory is a useful toll to determine the optimal concession period (Dixit and Pindyck, 1994; Galera and Solino, 2009). The problem of the determinant of the optimal concession period is still challenge in many aspects. If, for certain practical considerations, the concession period is subject to some constraints (such as, must be less than thirty years), then how those constraints affect the optimal decision of the parties? As pointed in Chapter 4, it is still not clear that whether the optimal concession period is free from the effect of user value-of-time (VOT) heterogeneity. In addition, if the payoff of private sector is decreasing over years, such as, the maintenance cost is increasing over year, discussed in Chapter 7, then the willingness of the private sector to operate the depreciated road is decreasing. In this case, it is interesting to examine the effects of the yearly-increasing operating costs on the government regulations and auction mechanism design.

The risk identification and analysis are very complicated and challenging tasks for a BOT toll road scheme. Miller and Lessard (2001) classified the project risks into three types: market-related (demand, supply and financial), completion (technical, construction and operational) and institutional risks. They pointed that transportation projects face highest market-related risks than the other two. The issue of demand uncertainty is addressed in Chapter 5, however, in which, the risk allocation is assumed to be pre-determined, and the public and private sectors are assumed to be risk neutral. We only focus the analysis on the flexibility of contract adjustment taking account of the demand realization. In fact, how to allocate the project risk among parties is one of the difficult questions, which has attracted an abundant of studies (Alonso-Conde et al., 2007; Dewatripont and Legros, 2005; Guthrie, 2006; Ng and Loosemore, 2007). When the contract is awarded, the public sector expects to transfer the investment, together with its risk, to the private sector of the economy. However, in the presence of unforeseen contingencies, the frequent renegotiations induce many political and economic side effects. The flexible-term contract method, together with the Minimum Income Guarantee and Revenue Distribution Mechanism, are current popular methods, which can transfer a part risk back to the public sector. A further consideration for the public sector is how to capture the risk perception of the private sector in his investment decision, which has significant impacts on those risk allocation methods.

There actually remains a lot to do on this subject. In Chapter 3, the Pareto-efficient BOT contracts are studied in an ideal world without asymmetric information and moral hazard problem, specially, the information on cost is perfectly known to both parties. And both parties do not take strategic action ins setting the contract. The hidden strategies are not seldom and have resulted in cost overrun (Alexander et al., 2000; Bajari et al., 2006; Odeck, 2004). The regulatory regimes and auction mechanisms should be designed to consider those asymmetries between the public and private sectors. In Chapter 6, the maintenance cost is assumed to keep the initial quality of the highway (roughness, safety) as de Palma et al. (2007a). The more practical method to view the maintenance performance as a decision activity of the highway agency (Durango-Cohen and Madanat, 2008; Li and Madanat, 2002), in which the highway agency chooses the frequency and intensity of the pavement rehabilitation activities to minimize the discounted social (agency plus users) cost over a long planning horizon for a given and known deterioration curve and rehabilitation effectiveness. The meaningful challenge is to extend the model to add other decision dimensions (frequency and intensity of maintenance performance) for the BOT problem. The limitations of the model in this chapter include that the homogeneous users and the assumption of deterministic world. To develop a model to consider those problems is another direction of our future research. Finally, in Chapter 7, the ownership regimes of the transportation network tends to be complicated, which has significant impacts on the transportation policy. Such as, the equity issue of the traditional network design problem should guarantee the interests of the existing private operators and road users in the network; the toll scheme should take in account the action of the private operators in the network; the planner should efficiently use the public funds and available private funds to tradeoff the budget and exhausted use of user fare in road investment.

Bibliography

Acemoglu, D., Ozdaglar, A., 2007. Competition and efficiency in congested markets. Mathematics of Operations Research 32 (1), 1–31.

Alexander, I., Estache, A., Oliveri, A., 2000. A few things transport regulators should know about risk and the cost of capital. Utilities Policy 9 (1), 1–13.

Alonso-Conde, A. B., Brown, C., Rojo-Suarez, J., 2007. Public private partnerships: Incentives, risk transfer and real options. Review of Financial Economics 16 (4), 335–349.

Arnott, R., Kraus, M., 1998a. Self-financing of congestible facilities in a growing economy. In: Pines, D., Sadka, E., Zilcha, I. (Eds.), Topics in Public Economics: Theoretical and Applied Analysis. Cambridge, Cambridge University Press, pp. 161–184.

Arnott, R., Kraus, M., 1998b. When are anonymous congestion charges consistent with marginal cost pricing? Journal of Public Economics 67 (1), 45–64.

Arnott, R., Yan, A., 2004. The two-mode problem: second-best pricing and capacity. Review of Urban & Regional Development Studies 12 (3), 170–199.

Athias, L., Saussier, S., 2007. Contractual flexibility or rigidity for public private partnership? theory and evidence from infrastructure concession contracts, mPRA working paper.

Averch, H., Johnson, L. L., 1962. Behavior of the firm under regulatory constraint. The American Economic Review 52 (5), 1052–1069.

Bain, R., 2009. Error and optimism bias in toll road traffic forecasts. Transportation 36 (5), 469–482.

Bajari, P., Houghton, S., Tadelis, S., 2006. Bidding for incomplete contracts: An empirical analysis. NBER Working Papers 12051, National Bureau of Economic Research, Inc.

Bennett, J., Iossa, E., 2006. Building and managing facilities for public services. Journal of Public Economics 90 (10-11), 2143–2160.

Bryson, M. C., Siddiqui, M. M., 1969. Some criteria for aging. Journal of the American Statistical Association 64 (328), 1472–1483.

Chod, J., Rudi, N., 2005. Resource flexibility with responsive pricing. Operations Research 53 (3), 532–548.

Chu, C. P., Tsai, J. F., 2004. Road pricing models with maintenance cost. Transportation 31 (4), 457–477.

Dafermos, S. C., 1973. Toll patterns for multiclass-user transportation networks. Transportation Science 7 (3), 211–223.

Daganzo, C. F., 1983. Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost jacobians. Transportation Science 17 (3), 282–300.

De Borger, B., Van Dender, K., 2006. Prices, capacities and service levels in a congestible bertrand duopoly. Journal of Urban Economics 60 (2), 264–283.

de Fraja, G., Delbono, F., 1989. Alternative strategies of a public enterprise in oligopoly. Oxford Economic Papers 41 (2), 302–311.

de Palma, A., 1992. A game-theoretic approach to the analysis of simple congested networks. The American Economic Review 82 (2), 494–500.

de Palma, A., Kilani, M., Lindsey, R., 2007a. Maintenance, service quality and congestion pricing with competing roads. Transportation Research Part B 41 (5), 573–591.

de Palma, A., Lindsey, R., 2000. Private toll roads: competition under various ownership regimes. The Annals of Regional Science 34 (1), 13–35.

de Palma, A., Lindsey, R., Proost, S., 2007b. Investment and the use of tax and toll revenues in the transport sector. Vol. 19 of Research in Transportation Economics. Elsevier.

de Vany, A., Saving, T. R., 1980. Competition and highway pricing for stochastic traffic. The Journal of Business 53 (1), 45–60.

Dewatripont, M., Legros, P., 2005. Public-private partnerships: contract design and risk transfer. EIB Papers 5/2005, European Investment Bank, Economic and Financial Studies.

Dial, R. B., 1999a. Network-optimized road pricing: part i: a parable and a model. Operations Research 47 (1), 54–64.

Dial, R. B., 1999b. Network-optimized road pricing: part ii: algorithms and examples. Operations Research 47 (2), 327–336.

Dixit, A. K., Pindyck, R. S., 1994. Investment under Uncertainty. Princeton University Press.

Durango-Cohen, P. L., Madanat, S. M., 2008. Optimization of inspection and maintenance decisions for infrastructure facilities under performance model uncertainty: A quasi-bayes approach. Transportation Research Part A 42 (8), 1074–1085.

Edelson, N. M., 1971. Congestion tolls under monopoly. The American Economic Review 61 (5), 873–882.

Engel, E., Fischer, R., Galetovic, A., 1997. Highway franchising: pitfalls and opportunities. The American Economic Review 87 (2), 68–72.

Engel, E., Fischer, R., Galetovic, A., 2008. Public-private partnerships: when and how? SCIP working paper 379.

Engel, E. M., Fischer, R., Galetovic, A., 2004. Toll competition among congested roads. Topics in Economic Analysis and Policy 4.

Engel, E. M. R. A., Fischer, R. D., Galetovic, A., 2001. Least-present-value-of-revenue auctions and highway franchising. The Journal of Political Economy 109 (5), 993–1020.

Farrell, J., Shapiro, C., 1990. Horizontal mergers: An equilibrium analysis. American Economic Review 80 (1), 107–26.

FHWA, 2002. Life cycle cost analysis primer, office of asset management. Tech. Rep. FHWA-IF-02-047, US. Federal Highway Administration,FHWA-IF-02-047.

Galera, A. L. L., Solino, A. S., 2009. A real options approach for the valuation of highway concessions. Transportation Science In Press.

Gallego, G., Ryzin, G. v., 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. Management Science 40 (8), 999–1020.

Geoffrion, A. M., 1967. Solving bicriterion mathematical programs. Operations Research 15 (1), 39–54.

Gomez-Ibanez, J., Meyer, J., Luberoff, D., 1991. The prospects of privatizing infrastructure. Journal of Transport Economics and Policy 25, 259–278.

Guasch, J. L., 2004. Granting and renegotiating infrastructure concessions: doing it right. World Bank Publications.

Guo, X., Yang, H., 2009. User heterogeneity and bi-criteria system optimum. Transportation Research Part B 43 (4), 379–390.

Guo, X., Yang, H., 2010a. Pareto-improving congestion pricing and revenue refunding with fixed demand. Transportation Research Part B In Press.

Guo, X. L., Yang, H., 2010b. Analysis of a build-operate-transfer scheme for road franchising. International Journal of Sustainable Transportation 3 (5-6), 312–338.

Guthrie, G., 2006. Regulating infrastructure: The impact on risk and investment. Journal of Economic Literature 44 (4), 925–972.

Handley, P., 1997. A critical view of the build-operate-transfer privatisation process in asia. Asia Journal of Public Administration 19 (2), 203–243.

Henderson, J., 1985. Economic Theory and the Cities, 2nd Edition. Academic Press, New York.

Knight, F. H., 1924. Some fallacies in the interpretation of social cost. The Quarterly Journal of Economics 38 (4), 582–606.

Lai, C., Xie, M., 2006. Stochastic Ageing and Dependence for Reliability. Springer, New York.

Lariviere, M., 1999. Supply chain contracting and coordination with stochastic demand in Quantitative models for supply chain man agement. Kluwer Academic Publishers, Boston.

Leurent, F., 1996. The theory and practice of a dual criteria assignment model with a continuously distributed value-of-time. In: Lesort, J. B. (Ed.), Transportation and Traffic Theory. Pergamon, Exeter, pp. 455–477.

Levinson, D. M., Gillen, D., 1998. The full cost of intercity highway transportation. Transportation Research Part D 3 (4), 207–223.

Levy-Lambert, H., 1968. Tarification des services a qualite variable-application aux peages de circulation. Econometrica 36 (3/4), 564–574.

Li, Y., Madanat, S., 2002. A steady-state solution for the optimal pavement resurfacing problem. Transportation Research Part A 36 (6), 525–535.

Light, T., 2009. Optimal highway design and user welfare under value pricing. Journal of Urban Economics 66 (2), 116–124.

Lindsey, R., 2006. Do economists reach a conclusion on road pricing? Econ Journal Watch 3 (2), 292–379.

Liu, Y., Guo, X., Yang, H., 2009. Pareto-improving and revenue-neutral congestion pricing schemes in two-mode traffic networks. Netnomics 10 (1), 123–140.

Lo, H., Hickman, M., 1997. Toward an evolution framework for road pricing. Journal of Transportation Engineering 123 (4), 316–323.

Lorenzen, C. C., Barrientos, M. E., Babbar, S., 2003. Toll road concessions: the chilean experience. PFG Discussion Paper Series, World Bank 124.

Marchand, M., 1968. A note on optimal tolls in an imperfect environment. Econometrica 36 (3/4), 575–581.

Mayet, J., Hansen, M., 2000. Congestion pricing with continuously distributed values of time. Journal of Transport Economics and Policy 34 (3), 359–370.

McAfee, R. P., McMillan, J., 1987. Auctions and bidding. Journal of Economic Literature 25 (2), 699–738.

Mieghem, J. A. V., Dada, M., 1999. Price versus production postponement: capacity and competition. Management Science 45 (12), 1631–1649.

Miller, R., Lessard, D., 2001. Understanding and managing risks in large engineering projects. International Journal of Project Management 19 (8), 437–443.

Mills, G., 1995. Welfare and profit divergence for a tolled link in a road network. Journal of Transport Economics and Policy 29, 137–146.

Mohring, H., Harwitz, M., 1962. Highway Benefits: an Analytical Framework. Northwestern University Press, Evanston, IL.

Nassa, M., Eissa, F., 2003. On the exponentiated weibull distribution. Communications in statistics Theory and Methods 32 (7), 1317–1336.

Newbery, D. M., 1988. Road damage externalities and road user charges. Econometrica 56 (2), 295–316.

Newbery, D. M., 1989. Cost recovery from optimally designed roads. Economica 56 (222), 165–85.

Ng, A., Loosemore, M., 2007. Risk allocation in the private provision of public infrastructure. International Journal of Project Management 25 (1), 66–76.

Nie, Y. M., Liu, Y., 2010. Existence of self-financing and pareto-improving congestion pricing: Impact of value of time distribution. Transportation Research Part A 44 (1), 39–51.

Nijkamp, P., Rienstra, S. A., 1995. Private sector involvement in financing and operating transport infrastructure. The Annals of Regional Science 29 (2), 221–35.

Nombela, G., de Rus, G., 2004. Flexible-term contracts for road franchising. Transportation Research Part A 38 (3), 163–179.

Odeck, J., 2004. Cost overruns in road construction–what are their sizes and determinants? Transport Policy 11 (1), 43–53.

Oum, T. H., Zhang, Y., 1990. Airport pricing : congestion tolls, lumpy investment, and cost recovery. Journal of Public Economics 43 (3), 353–374.

Paterson, W. D. O., 1987. Road Deterioration and Maintenance Effects, Models for Planning and Management. the Johns Hopkins University Press.

Pigou, A., 1920. The Economics of Welfare. London, Macmillan.

Qiu, L. D., Wang, S., 2009. Bot projects: incentives and efficiency. Journal of Development Economics In Press.

Roth, G., 1996. Roads in a Market Economy. Avebury Technical, Ashgate Publishing Limited.

Skamris, M. K., Flyvbjerg, B., 1997. Inaccuracy of traffic forecasts and cost estimates on large transport projects. Transport Policy 4 (3), 141–146.

Small, K., Yan, J., 2001. The value of "value pricing" of roads: second-best pricing and product differentiation. Journal of Urban Economics 49 (2), 310–336.

Small, K. A., 1999. Economies of scale and self-financing rules with non-competitive factor markets. Journal of Public Economics 74 (3), 431–450.

Small, K. A., Verhoef, E. T., 2007. The Economics of Urban Transportation. Handbook of transport systems and traffic control. Routledge.

Small, K. A., Winston, C., Evans, C. A., 1989. Road Works. The Brookings Institution Washington, D. C.

Soong, E., 1999. Key fundamentals and examples from asia. In: First Workshop on Economic Cooperation in Central Asia: Challenges and Opportunities in Transportation. Asian Development Bank.

Spence, A. M., 1975. Monopoly, quality, and regulation. The Bell Journal of Economics 6 (2), 417–429.

Sullivan, W. G., Wicks, E. M., Luxhoj, J. T., 2006. Engineering Economy, 13th Edition. Upper Saddle River, N.J. : Pearson/Prentice Hall.

Talluri, K. T., van Ryzin, G. J., 2004. The Theory and Practice of Revenue Management. Kluwer Academic. Publishers.

Tam, A., 1998. Developing toll roads in china. Asia Engineer 26, 9–10.

Train, K. E., 1991. Optimal Regulation: The Economic Theory of Natural Monopoly. Vol. 1 of MIT Press Books. The MIT Press.

Tsai, J.-F., Chu, C.-P., 2003. The analysis of regulation on private highway investment under a build-operate-transfer scheme. Transportation 30 (2), 221–243.

Ubbels, B., Verhoef, E. T., 2008. Auctioning concessions for private roads. Transportation Research Part A 42 (1), 155–172.

van Brunt, B., 2004. The Calculus of Variations. Springer.

Varian, H. R., 1992. Microeconomic Analysis, 3rd Edition. W.W. Norton & Company, Inc.

Vassallo, J. M., 2006. Traffic risk mitigation in highway concession projects: the experience of chile. Journal of Transport Economics and Policy 40 (3), 359–381.

Verhoef, E., Nijkamp, P., Rietveld, P., 1996. Second-best congestion pricing: the case of an untolled alternative. Journal of Urban Economics 40 (3), 279–302.

Verhoef, E. T., 2007. Second-best road pricing through highway franchising. Journal of Urban Economics 62 (2), 337–361, essays in Honor of Kenneth A. Small.

Verhoef, E. T., 2008. Private roads: auctions and competition in networks. Journal of Transport Economics and Policy 42 (3), 463–493.

Verhoef, E. T., Mohring, H., 2009. Self-financing roads. International Journal of Sustainable Transportation 3 (4-5), 293–311.

Verhoef, E. T., Rouwendal, J., 2004. Pricing, capacity choice, and financing in transportation networks. Journal of Regional Science 44 (3), 405–435.

Verhoef, E. T., Small, K. A., 2004. Product differentiation on roads. Journal of Transport Economics and Policy 38 (1), 127–156.

Walters, A. A., 1961. The theory and measurement of private and social cost of highway congestion. Econometrica 29 (4), 676–699.

Warburton, A. R., 1983. Quasiconcave vector maximization: connectedness of the sets of pareto-optimal and weak pareto-optimal alternatives. Journal of Optimization Theory and Applications 40 (4), 537–557.

Wu, D., Yin, Y. F., Yang, H., 2010. The independence of volume?capacity ratio of private toll roads in general networks. Transportation Research Part B In press.

Xiao, F., Yang, H., 2008. Efficiency loss of private road with continuously distributed value-of-time. Transportmetrica 4 (1), 19–32.

Xiao, F., Yang, H., Guo, X., 2007a. Bounding the inefficiency of toll competition among congested roads. In: Bell, M., Heydecker, B. (Eds.), Proceedings of 17th International Symposium on Transportation and Traffic Theory. London, pp. 27–54.

Xiao, F., Yang, H., Han, D., 2007b. Competition and efficiency of private toll roads. Transportation Research Part B 41 (3), 292–308.

Xu, X., 2009. Optimal price and product quality decisions in a distribution channel. Management Science 55 (8), 1347–1352.

Yang, H., Huang, H.-J., 2004. The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. Transportation Research Part B 38 (1), 1–15.

Yang, H., Huang, H. J., 2005. Mathematical and Economic Theory of Road Pricing. Elsevier.

Yang, H., Meng, Q., 2000. Highway pricing and capacity choice in a road network under a build-operate-transfer scheme. Transportation Research Part A 34 (3), 207–222.

Yang, H., Meng, Q., 2002. A note on "highway pricing and capacity choice in a road network under a build-operate-transfer scheme". Transportation Research Part A 36 (7), 659–663.

Yang, H., Tang, W. H., Man Cheung, W., Meng, Q., 2002. Profitability and welfare gain of private toll roads in a network with heterogeneous users. Transportation Research Part A 36 (6), 537–554.

Yang, H., Verhoef, E., 2004. Guest editorial: road pricing problems: recent methodological advances. Networks and Spatial Economics 4 (2), 131–133.

Yang, H., Woo, K. K., 2000. Competition and equilibria of private toll roads in a traffic network. Journal of the Transportation Research Board 1733, 15–22.

Zhang, X., 2009. Win-win concession period determination methodology. Journal of Construction Engineering and Management 135 (6), 550–558.

Ziya, S., Ayhan, H., Foley, R. D., 2004. Relationships among three assumptions in revenue management. Operations Research 52 (5), 804–809.