The Basic Public Finance of Public-Private Partnerships

Eduardo M.R.A. Engel, Ronald D. Fischer and Alexander Galetovic[1]

January 9, 2006

Abstract

Public-private partnerships (PPPs) have become increasingly popular in recent years. We show that for these arrangements to be desirable from a public finance point of view, private firms must be productively more efficient than the public sector. In particular, PPPs are not a means to save on distortionary taxation.

We also characterize the contract that trades off optimally demand risk, user-fee distortions and the opportunity cost of public funds, under the assumption that the private sector is more efficient. The private firm is fully insured against demand risk in the case of large and small projects, but bears risk for projects of intermediate size. For small projects, no subsidies are required and the optimal contract length is demand contingent. By contrast, demand contingent subsidies are handed out in every state of demand for large projects and the contract lasts indefinitely. For projects of intermediate size the optimal contract involves a "minimum income guarantee" and states where the contract lasts indefinitely coexist with those where it is finite—the private firm collects more revenue in the latter than in the former.

For large and small projects the optimal contract can be implemented with an auction where the bidding variable is the present value of use fee revenue. Having firms bid on the lowest subsidy in this case is not only suboptimal, but also leads to a subsidy that is higher, on average, than the one obtained under the optimal auction. Finally, a bidding variable different from the present value of user fee revenue is needed to implement the optimal contract for intermediate size projects.

Key words: cost of public funds, Demsetz auction, subsidies, user fee, distortionary taxation.
JEL classification:

[1]Engel: Department of Economics, Yale University, 28 Hillhouse Ave., New Haven, CT 06511. Fischer: Center for Applied Economics (CEA), Department of Industrial Engineering, University of Chile, Av. República 701, Santiago, Chile. Galetovic: Facultad de Ciencias Económicas y Empresariales, Universidad de Los Andes, Av. San Carlos de Apoquindo 2200, Santiago, Chile. E-mails: eduardo.engel@yale.edu, rfischer@dii.uchile.cl, agaletovic@uandes.cl.

1 Introduction and motivation

There is an increasing interest in public-private partnerships (PPPs) around the world.[2] In a typical project of this type, a private firm builds and finances the infrastructure and then collects user fees for many years. Eventually, the franchise ends and the infrastructure reverts to the state.[3]

PPPs have been used to finance toll-roads, provide sanitation services, sports stadiums, train lines, seaports and airports, and even to develop so-called orphan drugs, i.e., neglected disease drug development projects.[4] They have been hallowed as a third way between public provision and privatization, potentially combining the strengths of both. But the new trend raises many questions. When is private financing of infrastructure projects desirable? When is public financing optimal? When are public subsidies warranted? More generally, when is a public-private partnership best and how should the corresponding contract be designed? This paper provides a public-finance framework to answer these questions.

We study a model where a risk neutral government must contract a risk-averse firm to build and finance an infrastructure project with uncertain demand that requires a large up-front investment.[5] The firm (or *franchise holder* in what follows) can be compensated with a combination of subsidies that are paid out of the general budget, and user fees. At one extreme are arrangements where subsidies are the only source of income for the franchise holder. This is the "traditional" approach or public model, where firms build the infrastructure project and then hand it over to a public agency. At the other extreme is the case where all of the franchise holder's income comes from user fees. A variety of public-private financing arrangements are possible in between.

Our first result is that the usual justification for PPPs—relieving public budgets and substituting cheap private funding for distortionary tax finance—is suspect. To see why, note that it implies that the franchise holder should finance as much as possible of the project's construction cost and, consequently, the government should subsidize as little as possible. Yet this argument overlooks an essential point. At the margin, extending the concession term has an opportunity cost, since the government foregoes the revenues generated by the project during this period and this revenue could have been used to reduce distortionary taxation. Hence, the opportunity cost of \$1 in user fees is the shadow cost of public funds. For this reason, if the public and the private sector are equally efficient, user fees and subsidies

[2] For example, articles in the Financial Times mentioning this concept increased twenty-fold over the last decade, from 50 in 1995 to 1,153 in 2004.

[3] The term "Public-Private Partnership" does not have an unambiguous meaning and definitions abound. In this paper we will have in mind an infrastructure project such that (i) assets are possibly temporarily owned by the private firm; (ii) both the private firm and the government are residual claimants, often in ambiguous terms; and (iii) there is substantial public planning involved.

[4] The case of PPPs in the transportation sector is particularly compelling. Growing congestion, budgetary problems, and a major decrease in toll collection costs have led more than 20 U.S. states to pass legislation permitting the operation of public-private partnerships (PPPs) to build, finance and operate toll-roads, bridges and tunnels. See "Paying on the Highway to Get Out of First Gear." New York Times, April 28, 2005. Congestion costs in the top U.S. metro areas have grown steadily, reaching \$63.1 billion in 2003, 60% higher (in real terms) than a decade earlier (see Schrank and Lomax, 2005).

[5] As in principal-agent models, the less risk averse party —in our case the government— is assumed to be risk neutral. Assuming a risk averse firm is a shortcut for agency problems preventing risk diversification, see Appendix D in the working paper version of Engel, Fischer and Galetovic (2001) for a model along these lines.

are perfect substitutes at the margin and a continuum of revenues/subsidy combinations implement the optimum.

We then show that private participation is warranted only if franchise holders are productively more efficient and can deliver the infrastructure at a lower cost. This is not terribly surprising, because it is the standard argument in favor of privatization. Nevertheless, the optimal contract has quite specific features, most of which seem to be absent from contracts observed in practice.

To begin, the optimal contract remunerates the franchise holder as much as possible with revenues from user fees. Because the public sector spends inefficiently (otherwise a PPP is not warranted) one would like to subsidize as little as possible. It follows that no subsidies are granted if user revenues are sufficient to pay for the infrastructure in all states. On the contrary, if a subsidy is warranted in some state, then the concession should last indefinitely, thus minimizing the subsidy payment.

Given a perfectly inelastic stochastic demand structure, optimal contracts can be classified into three groups, depending on the size of the upfront investment. For small projects, defined as those where user fees are enough to pay for the infrastructure in all states of demand, the franchise holder receives full insurance and the franchise term is finite and flexible. Franchises last longer in states where demand is lower.

For large projects, defined as those where user fees cannot finance the project in any state of demand, subsidies are paid in all states and the contract lasts indefinitely. Again, the franchise holder is provided full insurance.

Yet for intermediate size projects there coexist states where the franchise term is finite with states where it lasts indefinitely. Subsidies are paid, if at all, only in states where the franchise lasts indefinitely. The franchise holder receives the same total income in all states where subsidies are optimal and this income is strictly less than total income in the states without a subsidy. It follows that in this case the franchise holder receives a so-called minimum income guarantee i.e. a state contingent subsidy that stabilizes income in low-demand states.

The third set of results relaxes the assumption of infinitely inelastic demand and considers optimal pricing of infrastructure services provided with a PPP. In general, prices should be set above the marginal cost of production even if no scale economies are present. The reason is that user fees substitute for distortionary taxation at the margin, both during and after the franchise. Thus it pays to distort pricing a little to reduce the need of distortionary taxation. Optimal prices are even higher in states where subsidies are paid, because they substitute for inefficient subsidies at the margin. In these states it pays to distort beyond what is warranted by the cost of public funds.

What is the economics behind these results? The optimal contract trades off three margins. First, the regulator can distort user fees to raise revenue to cover the up front investment. A second margin is the extent to which the franchise holder is forced to bear demand risk. And third, the government may use subsidies to insure the franchise holder and reduce user fee distortions, but it must collect distortionary taxes and bear the inefficiencies of public spending. In principle, these three margins suggest a complicated optimal combination of distortions; in practice the solution is quite simple and has a structure similar to that described above for the perfectly inelastic demand case.

If user fees that distort as much as the shadow cost of public funds can pay for the infrastructure in all states, the franchise holder receives full insurance, no subsidies are paid, and the infrastructure is priced optimally to substitute for public funds at the margin. This is the case of a project of small size, relative to its demand. At the opposite extreme, if subsidies are paid in all states, the franchise holder receives full insurance and the infrastructure is priced optimally, to substitute for public subsidies at the margin. This is the case of a large project. In between, when subsidies are paid in low demand states, but not in high-demand states, the inefficiency of public spending introduces a wedge between the marginal opportunity cost of public funds and the marginal opportunity cost of public spending. Thus, it pays to depart from optimal pricing in some states, and to introduce some risk.

In Engel, Fischer and Galetovic (2001), henceforth EFG, we imposed a "self-financing constraint" that ruled out subsidies by assumption, and studied the optimal *private* provision of infrastructure projects solving a Ramsey problem with variable franchise lengths. We go beyond that paper by allowing the government to grant subsidies, so that any combination along the public-private continuum is now possible. Thus we can answer the question of when the private provision of infrastructure is desirable, and derive several new insights for the optimal contract in this case.

This paper is also related to the literature on franchise bidding pioneered by Chadwick (1859) and Demsetz (1968), according to which competition for a monopoly infrastructure project will reproduce the competitive outcome (see Stigler [1968], Posner [1972], Riordan and Sappington [1987], Spulber [1989, ch. 9], Laffont and Tirole [1993, chs. 7 and 8], Harstad and Crew [1999] for important papers within this tradition, and Williamson [1976, 1985] for a critique). We contribute to this literature by including cases where projects cannot self-finance and government subsidies are necessary to make them feasible.

The remainder of the paper is organized as follows. We begin with a simple model with perfectly inelastic demand and only two states of nature (Section 2). This simplification allows us to study the basic public finance problem without the additional complications introduced by multiple states and a demand that responds to prices. The model is generalized in various directions in section 3, increasing the number of demand states, incorporating price-responsive demand and allowing for moral hazard. Section 4 concludes. A technical appendix follows with the main proofs.

2 Benchmark model

To better appreciate the basic public finance of infrastructure PPPs it is convenient to start with a perfectly inelastic demand and two possible states. In the next section we generalize our results to the case of price-responsive demand and an arbitrary number of states.

2.1 The model

A benevolent social planner hires a private firm to build an infrastructure project whose technical characteristics are exogenous. The firm can be compensated with sales revenues and subsidies. The planner's objective is to maximize the expected present value of users' welfare, subject to finding a firm that is willing to build the project, and considering the shadow cost of the public funds needed to pay for the subsidies.[6] When the franchise ends, the project reverts to the government and any future revenues are used to reduce distortions elsewhere in the economy.

Demand for the project is constant and completely inelastic. Demand may be high (Q_H), with probability π_H or low (Q_L), with probability π_L, where $\pi_L + \pi_H = 1$ and $Q_H > Q_L$. The upfront investment does not depreciate, and since we are not interested in construction cost uncertainty, we assume that there are many identical firms that can build the project at cost $I > 0$. We assume that there is a fixed price per unit of service equal to P, constant across demand states, this assumption is relaxed in the next section.

There are neither maintenance nor operation costs. There are two reasons why ignoring maintenance and operations costs is not a serious limitation. First, for most of the infrastructure projects of interest, the main costs are upfront, and maintenance and operation costs are relatively smaller (consider highways, dams, sport stadiums and rail lines). Second, and more important, if maintenance and operations costs are proportional to demand for the project, which is often a good approximation, then our framework extends trivially to the case with maintenance and operations costs, by substituting the price net of maintenance costs for the price in what follows.[7]

After the franchise ends, sales revenues revert to the government. All firms are identical, risk-averse expected utility maximizers, with preferences represented by the strictly concave utility function $u(\cdot)$.[8]

2.2 The optimal contract

It is often claimed that infrastructure franchising is desirable because private firms have access to funds at lower cost—they do not raise funds through distortionary taxation. By contrast, governments must resort to distortionary taxation to finance infrastructure project. Is this argument enough to make the case for franchising these projects?

It is useful to consider first the problem solved by a planner who knows I. Denote the present value of sales revenue received by the franchise-holder when demand is high by PVR_H and by PVR_L the amount received when demand is low. Then

$$\text{PVR}_i(T_i) \equiv \int_0^{T_i} PQ_i e^{-rt} dt = \frac{PQ_i(1-e^{-rT_i})}{r}, \qquad i = H, L; \tag{1}$$

where r is the discount rate, common across firms and the planner, and T_H and T_L denote the length of

[6]This objective function assumes that the income of users is uncorrelated with the benefit of using the project, so that if users spend a small fraction of their incomes on the services of the project they will value the benefits produced by the project as if they were risk neutral. See Arrow and Lind (1970).

[7]This assumption is true for highways and probably for rail lines.

[8]This should be interpreted as a reduced form for an agency problem that prevents the franchise-holder from diversifying risk. See Appendix D in Engel et al. (2001) for a model along these lines.

the franchise when demand is, respectively, high or low.

The planner can subsidize the project in the amounts S_H, $S_L \geq 0$. By a "subsidy" we mean any cash transfer from the government to the private firm. For example, it may involve the payment made upfront to procure the project in the traditional fashion or a cash transfer under a Build-Operate-and-Transfer (BOT) contract to supplement sales revenue from the project ('minimum income guarantees'). In any case, distortionary taxes that cost $\lambda > 1$ must be raised to pay one dollar of subsidy. We also assume that the planner collects revenues after the franchise ends. Each dollar raised in user fees is then used to reduce distortionary taxes collected elsewhere in the economy. For both reasons mentioned above, the planner wants to transfer the smallest possible subsidy to the project.

Since private participation is voluntary, the planner solves the following problem:

$$\min_{\{T_H,T_L,S_H,S_L\}} \sum_{i=H,L} \pi_i \left[\mathrm{PVR}_i + \lambda S_i - (\lambda - 1)\left(\frac{PQ_i}{r} - \mathrm{PVR}_i\right)\right] \tag{2}$$

$$\text{s.t.} \qquad \sum_{i=H,L} \pi_i u_i(\mathrm{PVR}_i + S_i - I) = u(0),$$

where $u(0)$ is the firm's outside option—the level of utility attained when not undertaking the project—and we have omitted the functional dependence of PVR_i on T_i for brevity.

The terms in the objective function and in the participation constraint are justified as follows: $\mathrm{PVR}_i + \lambda S_i$ is the public cost of the total amount transfered to the franchise holder in state i. Of course, the franchise holder receives only $\mathrm{PVR}_i + S_i$, precisely the amount that appears in the firm's participation constraint. Next, $\left(\frac{PQ_i}{r} - \mathrm{PVR}_i\right)$ is the total revenues collected by the government, in present value. This allows the government to reduce distortionary taxation, thereby saving resources to society in the amount $(\lambda - 1)\cdot\left(\frac{PQ_i}{r} - \mathrm{PVR}_i\right)$.

Minimizing the objective function (2) is equivalent to minimizing:

$$\sum_{i=H,L} \pi_i(\mathrm{PVR}_i + S_i), \tag{3}$$

where the term $(\lambda - 1)(PQ_i/r)$ is ignored because it does not depend on the problem's choice variables,[9] and a positive multiplicative constant λ is dropped as well. It can be seen that the per-dollar cost of paying for the project with sales revenues or subsidies is the same. Thus, social welfare depends on *total* transfers to the franchise-holder, no matter whether these come in the form of a subsidy or sales revenue. This is the fundamental insight behind the following proposition:

Proposition 1 (Irrelevance of the Public Cost of Funds Argument) *Any combination T_H, T_L, S_H, S_L such that $PVR_i + S_i = I$ for all i solves the planner's problem (2).*

Proof: Any of these combinations satisfies the firm's participation constraint, so they are feasible. Next note that these combinations of subsidies and sales revenue eliminate risk for the franchise holder. Because the franchise holder is risk averse, and this minimizes expected total transfers to the franchise

[9] This term will be important in our analysis of effort.

holder, it is optimal. ∎

What is the economics of this result? The standard reasoning in favor of Build-Operate-and-Transfer contracts points out that subsidies are an expensive means of financing projects, because they are paid with distortionary taxes. Yet the multiplicity of possible subsidy-sales revenue combinations indicates that distortionary taxation ($\lambda > 1$) is not sufficient to make BOT contracts preferable. For one possible solution is that $T_L = T_H = 0$ and $S_L = S_H = I$—the traditional approach to project financing where the government pays for the project upfront. At the other extreme is a BOT contract, where the franchise holder pays I, collects sales revenues, and no subsidies are paid. In addition, there is a continuum of intermediate solutions. What does the standard reasoning overlook?

An essential aspect of the infrastructure projects we consider is that the government foregoes sales revenues under BOT. Extending the concession term by Δt has an opportunity cost at the margin, for the government foregoes the sales revenues that the project generates during this extension. This income could have been used to reduce distortionary taxation. Hence, the opportunity cost of paying the franchise holder with \$1 out of subsidies, λ, is exactly the same as paying him with sales revenue. The point that is made clear by rewriting the objective function as in (3) is that $\lambda > 1$ justifies minimizing the total transfer to the firm that builds the project, but the revenue/subsidy mix is irrelevant.

Budget risk Even though absent in the government's objective function, budgetary uncertainty is unaffected by the option chosen along the public-private continuum of optimal contracts. For all such contracts, the total revenue that the government gives up in each state of demand is $\text{PVR}_i + S_i$, which is equal to I and does not depend in demand realization. It follows that the government budget bears no additional risk because of the project.

Shadow fees In some countries PPPs take the form of fixed term contracts where users pay no fees for the infrastructure service. The franchise holder is compensated via so-called "shadow fees," that is, by user fees paid directly by the government.

As discussed later in this section, shadow fees can be inefficient if there are productive inefficiencies or congestion effects. Proposition 1 provides an additional reasons to discourage fixed term contracts with shadow fees:

Corollary 1 *When no user fees can be charged (say because of political constraints), $S_i = I$ for all states i is optimal. Hence shadow prices which make payments contingent on the use of the infrastructure for a fixed and finite term T are never optimal.*

When P is forced to be zero, only one of the continuum of optimal contracts described in Proposition 1 is feasible. Furthermore, shadow prices combined with a finite term contract not only force the franchisee holder to bear risk, but also impose risk on the public budget.

2.3 Why PPPs? Productive efficiency

Our analysis shows that the justification of private participation in infrastructure cannot rest on the often claimed "fact" that private participation relieves strained budgets and reduces distortionary taxation. One of the main arguments in favor of franchises is that governments are unable to spend efficiently, perhaps because of political economy considerations or outright corruption. On the other hand, many argue that the experience with infrastructure PPPs has been unsatisfactory and that the traditional model may be more cost efficient after all. This controversy is, of course, about productive efficiency. In this section we explore the implications of differences in productive efficiency for the optimal contract.

To model productive efficiency we let ζ denote the number of dollars needed by the government to achieve with a subsidy what a private firm achieves by spending one dollar. If $\zeta > 1$ then private firms are more efficient, if $\zeta < 1$, then the traditional model is better. This leads to the following planner's problem:

$$\min_{\{T_H,T_L,S_H,S_L\}} \sum_{i=H,L} \pi_i \left[(\mathrm{PVR}_i + \zeta\lambda S_i) - (\lambda - 1)\left(\frac{PQ_i}{r} - \mathrm{PVR}_i\right)\right]$$

$$\text{s.t.} \sum_{i=H,L} \pi_i u_i(\mathrm{PVR}_i + S_i - I) = u(0).$$

Note that λ is multiplied by ζ in the planner's objective function, since the planner needs ζ dollars to increase the receipts of the franchise holder by one dollar and ζ dollars cost $\zeta\lambda$. By contrast, the second term in the objective function, that captures the reduction in tax distortions due to user fees collected by the government, continues being multiplied by $(\lambda - 1)$.

As before, the objective function can be replaced by

$$\sum_{i=H,L} \pi_i (\mathrm{PVR}_i + \zeta S_i). \tag{4}$$

It can be seen that when $\zeta \neq 1$ either subsidies are cheaper ($\zeta < 1$) or more expensive ($\zeta > 1$) than sales revenues as a means of financing the infrastructure project. The following result shows that the traditional approach to infrastructure financing is better if $\zeta < 1$.

Proposition 2 *If $\zeta < 1$, the optimal contract is such that all income received by the franchise-holder comes from subsidies. Thus, the traditional approach to infrastructure financing is strictly preferred to a BOT contract.*

Proof: It follows directly from simple inspection of (4). ∎

Now consider the case $\zeta > 1$, that is, private firms are more efficient. Clearly, $\zeta > 1$ is *not* a sufficient argument against subsidizing a road, for it may be the case that its social value exceeds I and user fee income is insufficient to pay for it in one or both states. In those cases (possibly state contingent) subsidies, large enough to make the project privately attractive, are warranted. The following proposition characterizes the solution.

Proposition 3 *If $\zeta > 1$, the optimal contract varies with I as follows:*

1. ***Small projects:*** *If $PVR_L(\infty) \geq I$, then the optimal contract is the unique pair (T_L, T_H) such that $PVR_L(T_L) = PVR_H(T_H) = I$. No subsidies are provided in this case.*

2. ***Large projects:*** *If $PVR_H(\infty) < I$, the optimal contract is such that the government provides a subsidy $S_i = I - PVR_i(\infty)$ and the franchise lasts indefinitely in all states.*

3. ***Intermediate size projects:*** *If $PVRI_L(\infty) < I \leq PVR_H(\infty)$, the optimal contract is such that the franchise holder receives less than I in the low demand state and more than I in the high demand state. As I increases within this range of values, the following scenarios attain:*

   (a) *First there is a range of values of I for which total revenue in the low state remain constant at $VPI_L(\infty)$, while the length of the franchise in the high state increases to the point where the participation constraint is satisfied.*

   (b) *A range of values for I where revenue in both states increases with I follows.*[10] *Subsidies are paid out in low state, while the franchise length continues increasing in the high state.*

   (c) *Finally comes a range of value of I where income in the high state remains constant and equal to $PVR_H(\infty)$. Subsidies in the low demand state continue increasing, to the point where the participation constraint is satisfied.*

Proof: For a formal proof, see the case with n demand states in the Appendix. The following arguments provide an informal reasoning, using Figure 1. The figure shows how total revenue—present value of user fees plus subsidy—varies with I in each state of demand. $\text{PVR}_i(\infty)$ is equal to 50 in the low demand state and equal to 100 in the high demand state. The solid lines correspond to a more risk averse utility function than the dashed lines.[11]

1. Consider the case $PVRI_L(\infty) \geq I$. Since the projects pays for itself in all states of the world, and subsidizing is costly because $\zeta > 1$, there should be no subsidy, and in this sense one can say that "privatization" is optimal. Rent extraction and risk aversion further imply that the present value of user fee income received by the franchise holder should be equal to I in all states. It follows that $T_L > T_H$ because revenues accrue at a slower rate when demand is low. Thus the franchise term will in general be finite but demand contingent.

2. If the project is such that $PVR_H(\infty) < I$, i.e., it never pays for itself, it is best to equate distortions across states of the world by subsidizing in all states. Since subsidies are an expensive way of remunerating the firm, it pays to minimize them. To do so, the franchise must last as long as possible. Moreover, the cost is minimized if we do not impose risk on the franchise holder, i.e., $S_i = I - PVR_i(\infty)$. Thus, when subsidies are paid in all states the firm receives full insurance and the franchise lasts indefinitely.

[10] As discussed below, with low degrees of risk aversion this set of values of I may be empty.

[11] Even though the figure only depicts total revenue for each state of demand, this amount can be decomposed into user fees and subsidies noting that, since $\zeta > 1$, subsidies are used only when user fee revenue cannot provide the required amount. That is, if Rev_i denotes optimal revenue in state i, then $S_i = \max(0, \text{Rev}_i - \text{PVR}_i(\infty))$.

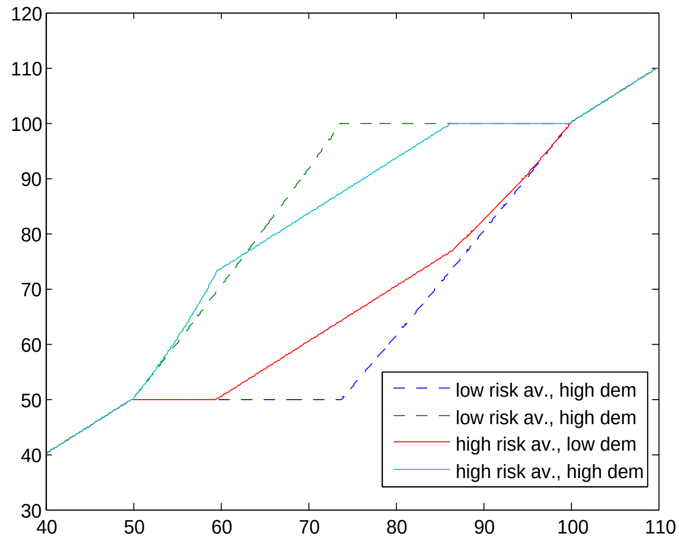

Figure 1: Revenue in both states, as a function of I. High (—) and low (- - -) risk aversion.

3. Consider first part (a). For I slightly larger than $VPI_L(\infty)$, it is more attractive to keep revenue in the low demand state fixed at $\text{PVR}_L(\infty)$ and increase revenue in the high demand state until the participation constraint is met. The reason for this is that the utility cost of having the franchise-holder bear demand risk is second order in this range of values for I. By contrast, providing a subsidy in the low demand state involves a jump in the distortion associated with the last dollar used for financing, from λ to $\lambda\zeta$, and therefore cannot be optimal.

   Put somewhat differently, suppose that full insurance was granted and $\text{PVR}_L(\infty)+S_L = I = \text{PVR}_H(T_H)$. By assumption, this requires a subsidy in the low demand state. But now consider the following trade: (i) reduce the subsidy in \$1 in the low demand state; this saves $\lambda\zeta$ to society; (ii) increase the concession term by Δt in the high-demand state to pay \$1 more to the firm; this costs society only λ. Because we start at full insurance, the risk effect is second-order while the welfare gain, equal to $\lambda\zeta - \lambda > 0$, is first order. Thus the fact that government is inefficient in spending funds is the reason why it is optimal to have the franchise holder bear risk.

   As I increases, the contract described in (a) runs into one of two problems. Either risk borne by the franchise holder reaches a utility cost equal to that associated with paying subsidies in the low demand state. Or user fee revenue in the high state runs out.

   In the first case—depicted by the solid line in Figure 1—a range of values of I where total income in both states of demand grows follows, this corresponds to part (b). User fee income in the high state is increasing and the government subsidizes the bad state to reduce risk. Since subsidies are an expensive way of remunerating the firm, it pays to minimize them. To do so, the franchise must last as long as possible. However, risk is not eliminated, due to the additional cost ζ of the subsidy that would be necessary in the L state. Eventually I reaches a value such that user fee income in the high demand state is exhausted. A range of values of I follows where income remains constant in this state while it continues increasing in the low state, this corresponds to part (c).

By contrast, if user fee revenue in the high state runs out before the cost of risk borne by the franchise holder equates the distortions associated with subsidies—as is the case for the dashed lines in Figure 1—then we go from the pattern described in (a) to that described in (c), skipping the intermediate range described in (b). This happens when risk aversion is low. ∎

What is the economics of propositions 2 and 3? When the government reduces the subsidy to the franchise holder by one dollar, it relaxes the government's intertemporal budget constraint by ζ dollars, which saves $\zeta\lambda$. On the other hand, the franchise holder must appropriate one additional dollar of user fee revenue in present value to meet her budget constraint. This forces the government to increase the tax burden by one dollar, which costs λ. Hence, the traditional approach is better if $\zeta < 1$, for then having the government build the project is cheaper.[12]

On the contrary, it will pay to avoid subsidies as much as possible if $\zeta > 1$. The surprising feature in this case is that the firm no longer receives full insurance when subsidies are paid in one state but not in the other. The economics is as follows. Low-demand states require a subsidy, which has a marginal cost of $\lambda\zeta$, so that *(i)* it is convenient to subsidize as little as possible, so the franchise in these states must be infinite in order to maximize sales revenue, and *(ii)* in order to satisfy the participation constraint it is preferable to increase revenues in the high demand state (which are "cheap", at a marginal cost of λ) instead of equating income in the two states. Thus the optimal contact considers a *minimum income guarantee* that reduces sales revenue risk, but does not eliminate it altogether.

Remark 1 *There exists values I_1 and I_2 such that for $I_1 < I < I_2$ the optimal contract resembles a variable-term concession with a minimum income guarantee. There also exists a minimum income guarantee for values of $I \geq I_2$, yet the concession lasts indefinitely in this case. By contrast, no guarantees are involved when $I \leq I_1$.*

2.4 The optimal auction

The informational requirements needed to implement the optimal contract might seem formidable, yet in most cases it can implement with a straightforward extension of the Present-Value-of-Revenue (PVR) auction proposed in EFG.

Consider first a small project, that is, a project that can pay its way with sales revenues in all states. Then an auction where the bidding variable is the total present value of sales revenues collected by the

[12] This is the case, for example, if the ministry has a technological advantage and is able to build cheaper than the private concessionaire.

franchise holder over the life of the concession, β, implements the optimal contract. This follows from noting that rents will be dissipated in a competitive auction, so that β will satisfy:

$$\pi_L u(\beta - I) + \pi_H u(\beta - I) = u(0). \tag{5}$$

Hence the winning bid will be $\beta = I$, which corresponds to the optimal contract derived in the preceding section. Since $Q_H > Q_L$, the franchise term is shorter when demand is high. Users pay the same amount in both states of nature and thus face no risk.[13] Furthermore, the planner can implement the optimal contract using a PVR auction even if she does not know the values of I, the π_i's or the Q_i's, $i = L, H$. All the planner needs to know is that the project is small, that is, that $\text{PVR}_L(\infty) \geq I$.

Consider next a large project, that is, a project where the optimal contract involves subsidies in all demand states. A PVR auction will implement the optimal contract in this case as well, as long as the government subsidizes the difference between the winning bid and the present value of user fees collected. Informational requirements are small again, since the planner only needs to know that the project is large, that is, that $\text{PVR}_H(\infty) < I$.

We note that not only does a PVR auction implement the optimal contract, both for large and for small projects, it also reveals to the government the value of I. We summarize these results as follows:

Proposition 4 *The optimal contract can be implemented with a PVR auction if either $S_i = 0$ or $S_i > 0$ in all states i. Bidders reveal their (common) value of I in the auction and informational requirements are weak.* ∎

A PVR auction does not implement the optimal contract for an intermediate size project, since total income received by the franchise holder is larger in the high demand state, so that the value of β determined from (5) cannot be optimal.[14]

Assuming the government has full knowledge of demand parameters, the optimal contract can be implemented in this case as follows. First note that, denoting total revenue in state i as a function of the upfront investment, I, by $\text{Rev}_i(I)$, we have that the function that assigns $(\text{Rev}_H(I), \text{Rev}_L(I))$ to each value of I is one-to-one.[15]

If firms compete on who bids the smallest β, and the government announces that, given a winning bid β_0, the franchise holder's revenue in the high and low states will be $\text{Rev}_H(\beta_0)$ and $\text{Rev}_L(\beta_0)$, respec-

[13] It should be noted that uncertainty in I, which may be important in some projects, cannot be eliminated with a variable term contract.

[14] In EFG we showed that an extension of the PVR auction, where the government provides no compensation when total revenue collected in an indefinite franchise is less than the winning bid, implements the optimal contract when subsidies are ruled out by assumption. Such an extension is of no help here, since it cannot incorporate the subsidies that may be needed in the low demand state.

[15] The proof is by contradiction. If (x_L, x_H) corresponds to two values of I, $I_1 > I_2$, then the participation constraint can't be satisfied with equality for both values. Thus either it is not satisfied for I_1, or it is satisfied with slack for I_2. Both alternatives contradict the definition of $(\text{Rev}_H(I), \text{Rev}_L(I))$. Also note that the relation between I and each of the components of $(\text{Rev}_H(I), \text{Rev}_L(I))$ is not one-to-one, since there are flat portions for each component.

tively, then a competitive auction will lead to a winning bid such that

$$\sum_i \pi_i u(\mathrm{Rev}_i(\beta_0) - I) = u(0).$$

It follows that $\beta_0 = I$, since it can be easily shown that the left hand side of the identity above, as a function of β_0, is strictly increasing. The optimal contract then is implemented by adding the proviso that government subsidies are handed out only when $\mathrm{Rev}_i(\beta_0)$ cannot be collected when the franchise lasts indefinitely. Also note that this auction implements the optimal contract, and reveals the value of I, in the more general case where the project's size can be small, intermediate or large.[16]

Bidding on the smallest subsidy Consider next the auction where the government sets a fixed franchise term T and a user fee P, and firms bid on the subsidy they require for building, operating and maintaining the road. Under competition the winning bid S satisfies:[17]

$$\sum_{i=H,L} \pi_i u\left(\frac{PQ_i(1-e^{-rT})}{r} + S - I\right) = u(0),$$

which means that $PQ_H(1-e^{-rT}) + S > I > PQ_L(1-e^{-rT}) + S$.

For this auction, the subsidy is the same in all states of demand, which does not correspond to any of the optimal contracts described in Proposition 3. The winning bidder is required to face risk and the winning bid does not reproduce the planner's solution. Furthermore, risk implies that for projects that are either large or small the expected transfer to the franchise holder is not minimized.

3 Extensions

This section extends the results in three directions. First, we examine the case in which there are more than two states of the world. The results remain unchanged, and the optimal contract with n states is very similar to the two-state case. Second, we examine the issue of price-responsive demand, and show that the qualitative results obtained with perfectly inelastic demand follow through. Finally, we examine the issue of moral hazard: we assume that the demand that attains depends on the effort of the agent, but effort is costly and must be rewarded. The optimal contract in this case is somewhere in between the standard moral hazard contract and the case without moral hazard. As the effect of effort decreases, the optimal contract tends to the contracts studied in the previous section.

[16] The functions $\mathrm{Rev}_i(I)$ is equal to I for I in the range that corresponds to small and large projects.

[17] The winning bid S is negative in the case of a small project, firms then are bidding on an upfront transfer to the government. With some straightforward modifications the result that follows also holds in this case.

3.1 Optimal contract with more than two states of the world

Let $\mathrm{PVR}_i^{\infty} \equiv \mathrm{PVR}_i(\infty)$ be the present value of revenue generated by the road over its entire lifetime, and define $L_i = e^{-rT_i}$. The general problem is to choose $(S_i, L_i)_{i=1}^n$ to solve

$$\min_{L_i, S_i} \sum_i \pi_i [\mathrm{PVR}_i^{\infty}(1 - L_i) + \zeta S_i],$$

subject to

$$\sum_i \pi_i u(\mathrm{PVR}_i^{\infty}(1 - L_i) + S_i - I) = u(0),$$
$$0 \leq L_i \leq 1,$$
$$S_i \geq 0.$$

The first order conditions for this problem w.r.t. L_i imply that:

$$\mu u_i' \geq 1 \quad \text{if} \quad L_i = 0, \tag{6}$$
$$\mu u_i' = 1 \quad \text{if} \quad 0 < L_i < 1, \tag{7}$$
$$\mu u_i' \leq 1 \quad \text{if} \quad L_i = 1, \tag{8}$$

where μ denotes the Lagrange multiplier associated with the participation constraint and

$$u_i' \equiv u'(\mathrm{PVR}_i^{\infty}(1 - L_i) + S_i - I).$$

The first order conditions w.r.t. S_i imply:

$$\mu u_i' \leq \zeta \quad \text{if} \quad S_i = 0, \tag{9}$$
$$\mu u_i' = \zeta \quad \text{if} \quad S_i > 0. \tag{10}$$

These set of conditions are used next to formally derive the optimal contract.

3.1.1 The case with $\zeta \leq 1$

We first study the characteristics of the solution when the government is at least as efficient as private firms.

Proposition 5 *Let $\zeta < 1$. Then, for all states i, $T_i = 0$ and $S_i = I$.*

Proof: Conditions (6) and (7) are incompatible with the first order conditions for S_i. It follows that $L_i = 1$ in all states of demand, and thus $T_i = 0$. Thus the only source of income for the franchise holder are subsidies, and the problem reduces to finding state-contingent subsidies that minimize the objective function subject to the firm's participation constraint. That $S_i = I$ in all states of demand now follows immediately. ∎

The intuition is simple: because $\zeta < 1$, subsidy finance is cheaper. Thus, one wants the government's share to be as big as possible.

Proposition 6 *Let $\zeta = 1$. Then, for all states i, any combination such that* $\mathrm{PVR}_i(T_i) + S_i = I$ *maximizes the planner's objective function.*

Proof: This is (a straightforward extension of) Proposition 1. ∎

3.1.2 The case with $\zeta > 1$.

We now study the optimal contract when the private sector is productively more efficient. We first note that $L_i < 1$ for all i, that is, $T_i > 0$. The proof is by contradiction, showing that if $L_i = 1$ in one state, say state 1, then the firm's participation constraint cannot be met.

If $L_1 = 1$, then from (8) and the first order conditions for S_1 it follows that $S_1 = 0$. Hence the franchise holder receives no revenue in state 1. Furthermore, from (6) and (7) it follows that in states j with $L_j < 1$ we must have $\mu u'_j$ larger than $\mu u'_1$. Since u' is decreasing (and μ is positive), this implies that total income in all states is zero and the participation constraint cannot be satisfied.

Next we note that $S_i > 0$ if and only if $L_i = 0$. This follows from the first order conditions for S_i and L_i and reflects the fact that, since $\zeta > 1$, to collect a given amount of revenue it is always better to exhaust user fee income before resorting to subsidies.

We are now ready to characterize the optimal contract:

Proposition 7 *Given a (positive) value for the participation constraint multiplier μ, the corresponding optimal contract is obtained as follows.*[18]

1. *For states i such that $\mu u'(PVR_i^\infty) \leq 1$, the optimal contract involves no subsidies and the contract length, L_i, is chosen so that*
$$\mu u'(PVR_i^\infty(1 - L_i)) = 1.$$

2. *For states i such that* $1 < \mu u'(PVR_i^\infty) \leq \zeta$, *the optimal contract extends indefinitely and no subsidies are handed out.*

3. *Finally, for sates i such that $\mu u'(PVR_i^\infty) > \zeta$, the optimal contract lasts indefinitely and S_i is determined by*
$$\mu u'(PVR_i^\infty + S_i) = \zeta.$$

Since u' is decreasing, it follows that total revenue is the same in all states in group 1, and all states in group 3. Furthermore, the common value for total revenue in the latter states is smaller than in the former states. Finally, total revenue varies across states in group 2, but always takes values between the common values for states in groups 1 and 3.

[18] That this approach covers all possible contracts follows from the fact that μ is strictly increasing in I (see the Appendix).

Of course, it may happen that some of the classes of states mentioned above is empty. If all states belong to group 1, we have a small project and the franchise holder is fully ensured in all states of demand. No subsidies are needed and the franchise length varies inversely with demand. By contrast, subsidies are used in all states to fully ensure the franchise holder if all states belong to group 3. This is the case of a large project and the contract lasts indefinitely.

Proof: It is straightforward to show that all first order conditions are satisfied by the solution above. Since the problem is a standard convex optimization problem (minimize a linear objective over a convex set) the first order conditions characterize the optimal contract.

Denoting by $\text{Rev}_i(\mu)$ total revenue in state i as a function of the participation constraint multiplier, it can be shown that $\sum_i \pi_i u(\text{Rev}_i - I)$ is strictly increasing in μ. It follows that there exists a unique value of μ for which the participation constraint is met with equality. ∎

The intuition underlying this result is similar to the case with two demand states. Figure 2 provides additional insights. The horizontal axis shows I, and the vertical axis plots the total revenue obtained by the franchise holder in each state of demand. The values of PVR_i^∞ are 40, 60, 80 and 100. In what follows it will be convenient to order the states so that $\text{PVR}_1^\infty < \text{PVR}_2^\infty < \ldots < \text{PVR}_n\infty$.

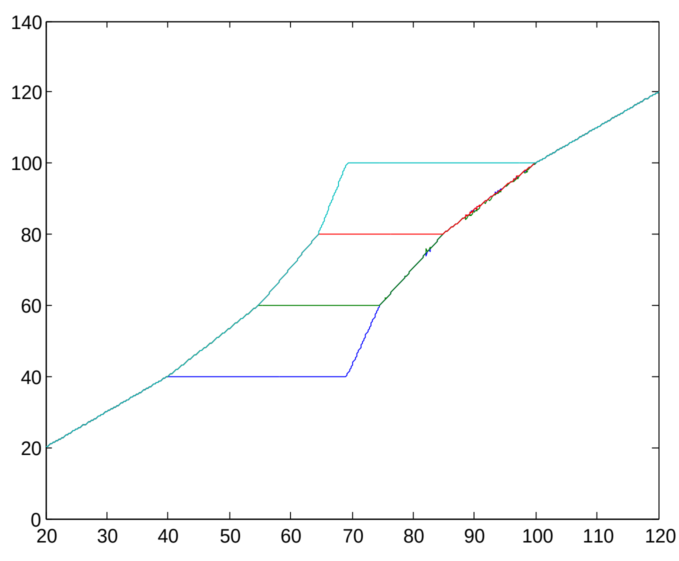

Figure 2: Revenue in each state of demand, as a function of I. Case with $n = 4$ and low risk aversion.

Consider first the optimal contract when $I \leq \text{PVR}_1(\infty) = 40$. It can be seen from Figure 2 that the franchise holder is fully insured and receives exactly I in each state. Hence, $T_i < \infty$ in the four states (except for the limit case when $I = 40$) and subsidies are never granted. Similarly, if $I \geq \text{PVR}_4(\infty)$ total revenues in each state are also equal to I. But now in all four states $T_i = \infty$ and $S_i > 0$ (except in the limit case when $I = 100$).

By contrast, when $I \in (\text{PVR}_1(\infty), \text{PVR}_4(\infty))$ the franchise holder initially receives less revenue in the low-demand states 1 than in the high-demand states 2, 3 and 4. As I increases but remains close enough to $\text{PVR}_1(\infty)$, the revenue of the franchise holder remains constant in the low-demand state 1 (because $T_1 = \infty$ and it is optimal to impose some risk on the agent so as not to induce the additional marginal distortion $\lambda\zeta$ due to a subsidy) while it gradually increases in the high-demand states 2, 3 and 4, because T_2, T_3 and T_4 increase to compensate for the losses in state 1. But when I reaches $\bar{I}_1$, it becomes convenient to begin to subsidize in the low-demand state 1 in order to limit the cost due to risk on the franchise holder. From then on, as I continues to increase, revenue received by the franchise holder increases in all states: by means of extending the term of the franchise T_2, T_3, T_4 in the high-demand states, and by increasing the subsidy S_1 in the low-demand state. At some point, when I reaches $\bar{I}_2$, $T_2 = \infty$. Now revenue remains fixed in state 2 and increases, by means of the subsidy in the low-demand state 1, and by lengthening the franchise in states 3 and 4. Eventually, as I increases, it becomes convenient to subsidize in state 2. Thus income in states 1 and 2 grows at the same rate via subsidies while it grows in states 3 and 4 via increases in the length of the franchise. Eventually, the franchise in state 3 becomes infinitely long, so income in state 3 remains constant while income in the 2 states with less demand grows via subsidies. Eventually we reach a range of values of I where income across states 1, 2 and 3 is the same and grows via subsidies, while income in state 4 remains constant at PVR_4^{∞}.

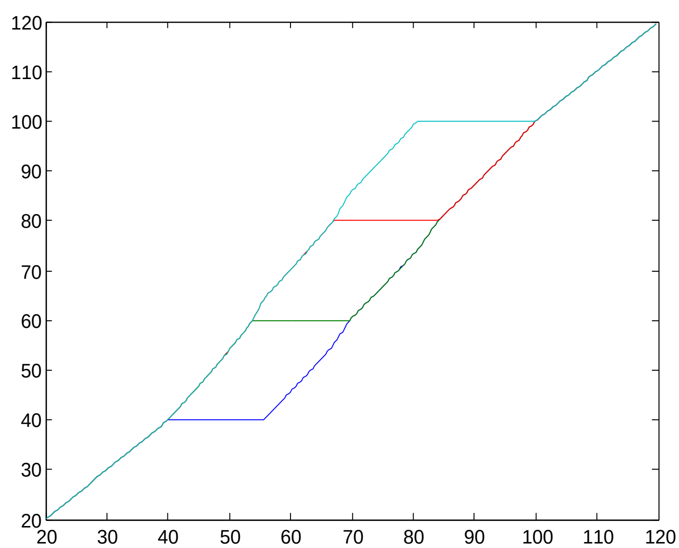

Figure 3: Revenue in each state of demand, as a function of I. Case with $n = 4$ and high risk aversion.

There is another possibility, illustrated in Figure 3, where income in state 2 (and 3) stops growing (because the franchise length becomes infinite) before the previous state receives the subsidy. This, however, does not change the intuition for the results. Whether this happens depends on the risk aversion of the agent. As risk aversion increases, the horizontal steps, where risk increases, become smaller.

In both cases considered above, for values of I where the franchise holder bears risk, the optimal contract provides a minimum income guarantee, depicted by the lower envelope in Figures 2 and 3. This guarantee comes together with an upside risk in high demand states.

The economics behind Proposition 7 is as follows. In states with revenues on the upper frontier, the term of the franchise is finite. It would not pay to differentiate revenues in those states because the franchise holder would bear risk, whereas the cost of giving her revenue in either state is the same—the cost of public funds. Similarly, on states on the lower frontier, the franchise holder receives the same revenue. Here the cost of giving her revenue in all states is also the same, but higher than the cost of public funds, because subsidies bear the inefficiency of public spending. The difference between the cost of subsidies and the cost of public funds introduces a wedge and justifies that the franchise holder bears some risk. Last, in between there are states where the concession lasts indefinitely, but still it does not pay to subsidize. These states emerge because in them the cost imposed by risk not so high to justify a subsidy.

3.2 Price-responsive demand

In this section we generalize the analysis by including price responsive demand and thus, allocative efficiency; and many possible states of demand. Essentially, none of the results obtained in the previous section changes. Once prices are set optimally to manage congestion, the marginal opportunity cost of funds follows essentially the same rules as when demand is perfectly inelastic

3.2.1 The model

There are n possible states of demand, which occur with probability $\pi_i, i = 1, ..., n$, with $\sum_i \pi_i = 1$. As before, the state becomes known immediately after the road is built, so that demand remains constant throughout time.

Demand for the infrastructure service in state i is given by $Q_i(P)$, the instantaneous surplus function corresponding to this state is

$$G_i(P) \equiv \mathrm{CS}_i(P) + PQ_i(P), \tag{11}$$

where $\mathrm{CS}_i(P)$ is consumer surplus in state i when the price is P.[19]

The function G_i is strictly concave and allows for congestion effects in the provision of the infrastructure's services (see EFG for a proof). It follows that $G_i(P)$ is decreasing for all P when there is no congestion in the infrastructure provision, and therefore attains its maximum at $P_i^* = 0$. On the other hand, when congestion matters, $G_i(P)$ has a unique interior maximum at $P_i^* > 0$. A user fee of $P_i = P_i^*$ makes users fully internalize the congestion externality they create. Thus, we call P_i^* the *congestion fee*

[19]For a derivation from first principles of G_i and Q_i considering congestion costs and the inelasticity of demand, see EFG.

in state i. It can also be shown that there exists a fee P_i^M, the *monopoly price*, that maximizes user fee revenues.

Finally, define sales revenues in state i to be $R_i(P) \equiv PQ_i(P)$. We assume that $R_i(P)$ is concave and strictly increasing in the range $[P_i^*, P_i^M]$. No (physical) cost of providing the infrastructure service is assumed in (11), beyond the upfront investment I and eventual congestion costs. This assumption is made for ease of notation. All the results that follow extend easily to the case with a (physical) cost function $C_i(Q_i)$ of providing services, as long as $\Pi_i(P) = R_i(P) - C_i(Q_i(P))$ is strictly concave in the range $[P_i^*, P_i^M]$.

If there were no need to finance the infrastructure project, of if $I = 0$, the user fee in state i should be chosen to maximize

$$G_i(P) + (\lambda - 1)R_i(P).$$

We denote this price by $P_i^*(\lambda)$.

Yet the project needs to be financed, and this involves distortions associated with the inefficiencies that come with subsidies and the cost of having the franchise holder bear income risk. For this reason, the user fee charged in some states may end up being above $P_i^*(\lambda)$. The following definition will be useful when describing these prices.

Definition 1 *For $\eta \geq 1$, define $H_i(P,\eta) \equiv G_i(P) + (\eta - 1)R_i(P)$ and $P_i^*(\eta) \equiv argmax\, H_i(P_i, \eta)$.*

The H function is a generalized welfare function for the government that can incorporate either the marginal cost of tax revenues (λ) and the marginal cost of subsidies ($\lambda\zeta$).[20] We define the value of the function $H_i(P,\eta)$ at $P_i^*(\eta)$ to be $H_i^*(\eta) = H_i(P_i^*(\eta), \eta)$.

It is easy to see that as η varies from 1 to infinity, $P_i^*(\eta)$ varies from P_i^* to P_i^M (see the Appendix for a proof of this and the following statements). Thus $P_i^*(\eta)$ is strictly increasing in η and

$$1 - \frac{G_i'(P_i^*(\eta))}{R'(P_i^*(\eta))} = \eta. \tag{12}$$

3.2.2 The planner's problem

We can now analyze the planner's problem. For each possible state of demand i the planner must choose two prices, the one that users pay during the life of the franchise and a second price that is charged by the government after the end of the franchise. The user fees in state i are denoted by P_i^F and P_i^A, where the superscripts "F" and "A" stand for *franchise* and *after*, respectively. Moreover, the planner must set the optimal contract lengths T_i (or L_i) and subsidies S_i. The subsidy is paid to the franchise holder immediately after demand is realized.[21]

[20] As well as congestion, which appears through the G_i function. The appendix characterizes the properties of the H function.

[21] The subsidy can be paid at any time during the franchise, as long as its discounted value at time zero is equal to S_i.

Considering the definition of H, the planner chooses P_i^F, P_i^A, L_i and S_i, $i = 1, \ldots, n$, to solve:

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{n} \pi_i \left[H_i(P_i^F, 0)(1 - L_i) \;+\; H_i(P_i^A, \lambda) L_i \;-\; \lambda\zeta r S_i \right] \\
\text{s.t.} \quad & \sum_{i=1}^{n} \pi_i u \left(\frac{R_i(P_i^F)}{r}(1 - L_i) + S_i - I \right) = u(0), \\
& 0 \le L_i \le 1, \\
& S_i \ge 0.
\end{aligned}
$$

Proposition 8 *In the optimal contract we have $P_i^A = P_i^*(\lambda)$.*

Proof: Since P_i^A does not appear in the problem's constraints, the optimal value of P_i^A is obtained by maximizing the subexpression in the objective function that includes this price. The result then follows from the definition of H and of P_i^*.[22] ∎

This result shows that after the franchise ends, the only consideration for the planner is that additional revenues from the franchise allow it to decrease the tax distortion at a rate of λ\$ per \$. Thus, the planner distorts the price for the services of infrastructure by an amount equal to the distortions induced by the user fee. Equivalently, if the project had an investment cost of zero, this would be the optimal toll.

In order to solve the planner's maximization problem, we consider the associated Lagrangian, taking advantage of the fact that we have obtained the optimal P_i^A:

$$
\mathscr{L} \;=\; \sum_{i=1}^{n} \pi_i \left[H_i(P_i^F, 0)(1 - L_i) \;+\; H^*(\lambda) L_i \;-\; \lambda\zeta r S_i \right] \;+\; r\mu \left[\sum_{i=1}^{n} \pi_i u \left(\frac{R_i(P_i^F)}{r}(1 - L_i) + S_i - I \right) \;-\; u(0) \right],
$$

where $r\mu$ denotes the multiplier associated with the firm's participation constraint. The optimality conditions derived from the Lagrangian provide some useful results.

Lemma 1 *We have:*

$$
\frac{\partial \mathscr{L}}{\partial L_i} = 0 \Longleftrightarrow P_i^F = P_i^*(\lambda), \tag{13}
$$

$$
\frac{\partial \mathscr{L}}{\partial S_i} = 0 \Longleftrightarrow P_i^F = P_i^*(\lambda\zeta). \tag{14}
$$

[22] The proof above applies if $L_i > 0$. If $L_i = 0$, then the objective function does not depend on P_i^A and this variable can take any value, including $P_i^*(\lambda)$.

Combining these results with that fact that $P_i^*(\eta)$ is increasing in η we have:

$$\frac{\partial \mathscr{L}}{\partial L_i} \leq 0 \Longleftrightarrow P_i^F \geq P_i^*(\lambda), \tag{15}$$

$$\frac{\partial \mathscr{L}}{\partial L_i} \geq 0 \Longleftrightarrow P_i^F \leq P_i^*(\lambda), \tag{16}$$

$$\frac{\partial \mathscr{L}}{\partial S_i} \leq 0 \Longleftrightarrow P_i^F \leq P_i^*(\lambda\zeta). \tag{17}$$

Proof: A straightforward calculation (see the Appendix for details) of the derivatives of the Lagrangian shows that

$$\frac{\partial \mathscr{L}}{\partial P_i^F} = \pi_i(1-L_i)\left[G_i' - R_i' + \mu u'\left(\frac{R_i}{r}(1-L_i) + S_i - I\right)R_i'\right], \tag{18}$$

$$\frac{\partial \mathscr{L}}{\partial L_i} = \pi_i\left[H_i^*(\lambda) - H(P_i^F, 0) + \left(\frac{G_i'}{R_i'} - 1\right)R_i\right],$$

$$\frac{\partial \mathscr{L}}{\partial S_i} = -\pi r\left[\lambda\zeta + \frac{G_i'}{R_i'} - 1\right],$$

where R_i, R_i' and G_i' are evaluated at P_i^F. Using (12), $\partial\mathscr{L}/\partial L_i$ evaluated at $P_i^*(\lambda)$ and $\partial\mathscr{L}/\partial S_i$ evaluated at $P_i^*(\lambda\zeta)$, are equal to zero. ∎

By Complementary Slackness (CS), $\partial\mathscr{L}/\partial P_i^F \leq 0$ implies $L_i = 0$ (similarly $\partial\mathscr{L}/\partial P_i^F \geq 0$ implies $L_i = 1$). It is only in the case when the partial derivative with respect to L_i is equal to 0 that the franchise length is a finite, positive value. Applying CS to the derivatives of the Lagrangian with respect to S_i, we finally obtain the following proposition that allows us to characterize the different demand states:

Proposition 9 *In the optimal contract:*

(a) If $L_i > 0$ then $P_i^F = P_i^*(\lambda)$.

(b) If $S_i > 0$ then $P_i^F = P_i^*(\lambda\zeta)$.

(c) If $L_i = 0$ and $S_i = 0$ then $P_i^*(\lambda) \leq P_i^F \leq P_i^*(\lambda\zeta)$.

(d) If $S_i > 0$ then $L_i = 0$.

(e) If $L_i > 0$ then $S_i = 0$.

Proof: Follows directly from Lemma 1. See the Appendix for details. ∎

Observe that there are three possible types of demand states. First, those where the contract length is positive and finite, and there are no subsidies ($L_i > 0, S_i = 0$); second, those where the term is infinite but there are no subsidies ($S_i = 0, L_i = 0$); and finally, those where the franchise term is infinite and there are positive subsidies ($L_i = 0, S_i > 0$). The first type is the *high demand state*, the second is the *medium demand state* and finally we have the *low demand state*.

Proposition 10 *Total income (present value of fees plus subsidies) and profits (total income minus investment) in high demand states is larger than in intermediate demand states, which is larger than in low demand states. Also, total income (and profits) is the same across high demand states, and across low demand states. Income may vary across intermediate demand states.*

Proof: From (18) we have that in the optimal contract:

$$\mu u'(\mathrm{Rev}_i - I) = 1 - \frac{G_i'(P_i^F)}{R_i'(P_i^F)}, \tag{19}$$

where Rev_i denotes total (present discounted) income in state i. It then follows from (12) and Proposition 9 that $\mu u'(\mathrm{Rev}_i - I) = \lambda$ when i is a high demand state, $\mu u'(\mathrm{Rev}_i - I) = \eta_i$, $\lambda \le \eta_i \le \lambda\zeta$ when i is an intermediate demand state, and $\mu u'(\mathrm{Rev}_i - I) = \lambda\zeta$ when i is a low demand state. ∎

This proposition is the reason for the names given to the different types of states. Note that in the high demand state, the optimal price is $P_i^*(\lambda)$, i.e. it only accounts optimally for the tax revenue distortion. Similarly, low demand state incorporates fully the additional cost of the subsidy that is required to reduce risk optimally for the franchise holder, with a price $P_i^*(\lambda\zeta)$.

The next result shows that the only way in which all states can be high demand states is that the state with the lowest demand generates sufficient revenues to finance the investment with an infinite contract.

Proposition 11 *All states are high demand states if and only if*

$$\min_i R_i(P_i^*(\lambda)) \ge rI. \tag{20}$$

Proof: From Proposition 10 and the firm's participation constraint we have that all states are high demand states if and only if $\mathrm{Rev}_i = I$ for all i.[23] From Proposition 9 it follows that for all i we must have $P_i^F = P_i^*(\lambda)$ and $S_i = 0$. Thus all states are high demand states if and only if the present value of user fees collected charging $P_i^*(\lambda)$ indefinitely is at least I. Condition (20) is necessary and sufficient for this to be the case. The optimal contract in this case sets $P_i^F = P_i^A = P_i^*(\lambda)$, $S_i = 0$ and $L_i = 1 - rI/R_i(P_i^*(\lambda))$. Condition (20) ensures that $L_i \ge 0$. ∎

We are now ready to characterize the optimal contract, which can be compared to the optimal contract for the inelastic demand case described in proposition 7. The strategy consists on defining, for each feasible value of μ, the revenues that would be obtained in an infinite length contract to determine the type of the state (high, medium or low demand) and then using the characteristics of the optimal solution in each type of state to define all the relevant characteristics of the contract.

Hence, given a value for the participation constraint multiplier, μ, the results above imply that the optimal contract can be fully specified (as a function of μ) as follows:

[23] Proposition **??** implies that total income is constant across states while the participation constraint implies that this constant is equal to I.

1. For $i = 1, \ldots, n$ find η_i that solves:

$$\mu u'\left(\frac{R_i(P_i^*(\eta_i))}{r} - I\right) = \eta_i. \tag{21}$$

2. Those states where $\eta_i \leq \lambda$ are high demand states. For these states $P_i^F = P_i^A = P_i^*(\lambda)$, $S_i = 0$, and $L_i \geq 0$ is chosen so that

$$\mu u'\left(\frac{R_i(P_i^*(\eta_i))}{r}(1 - L_i) - I\right) = \lambda.$$

3. Those states where $\lambda < \eta_i \leq \lambda\zeta$ are intermediate demand states. In these states $P_i^F = P_i^*(\eta_i)$, $L_i = 0$ and $S_i = 0$.

4. Those states where $\eta_i > \lambda\zeta$ are low demand states. For these states $L_i = 0$, $P_i^F = P_i^*(\lambda\zeta)$ and S_i is chosen so that

$$\mu u'\left(\frac{R_i(P_i^*(\eta_i))}{r} + S_i - I\right) = \lambda\zeta.$$

The following proposition uses the algorithm above and the participation constraint to characterize the optimal contract, by showing that there exists a fixed point (the optimal μ) to this algorithm.

Theorem 1 (Characterization of the Optimal Contract) *Denote by $Rev_i(\mu)$ total income in state i assigned by the algorithm above for a given value of μ and define*

$$C(\mu) \equiv \sum_{i=1}^{n} \pi_i u(Rev_i(\mu) - I).$$

Then there exists a unique value of μ that solves $C(\mu) = u(0)$. This is the participation constraint multiplier and the algorithm above, for this particular value of μ, fully characterizes the optimal contract.

Proof: The proof is by an application of the implicit function theorem (see the appendix). ∎

3.3 Moral hazard

[To be written up. For the time being you can read the formal results in the appendix and enjoy the figure below.]

4 Conclusions

As the worldwide enthusiasm about privatizations waned, PPPs began to boom. One of the main advantages from the point of view of governments is that PPPs allow them to transfer assets to private firms without transferring ownership, thus avoiding criticism from those who oppose privatizations. In addition, most PPPs involve a risk sharing agreement (many times not well specified in the contract) whereby

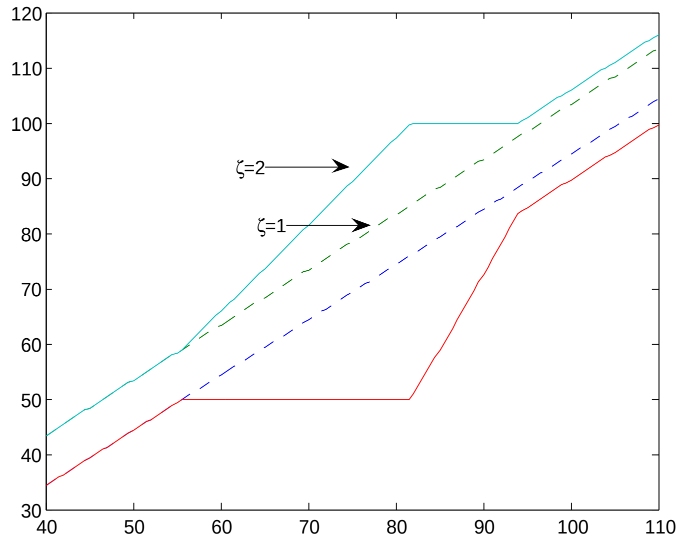

Figure 4: Revenue in both states, as a function of I, when effort matters.

both parties remain residual claimants. Thus, contrary to a privatization, a PPP maintains a direct link between the public budget and the infrastructure project.

This paper developed a simple analytical framework to study this link and its normative implications. Our model highlights that PPPs allow governments to make intertemporal transfers. A PPP may liberate public funds today, but it does so at the cost of fewer public revenues or higher subsidies in the future. At the margin, the marginal sacrifice of resources invested in the project equals the opportunity cost of public funds—exactly as in the traditional model—. Clearly, thus, the advantages from PPPs are not financial.

We have also shown that a PPP is justified only if the private firm is productively more efficient. This should not be surprising after all, but it has many implications. One is that whenever government provision is more efficient there is no case for a PPP. Moreover, if PPP is warranted and able to pay its way, it should look like a privatization, in that no risks should be shared. Last, if a PPP is warranted but unable to pay its way completely out of user fees, concessions should last as long as possible so as to minimize the subsidy transfer to the franchise holder.

PPPs have boomed around the world but many have been far from successful. One common occurrence is that contracts have been renegotiated, often involving term extensions or even direct bailouts paid out of the budget. Moreover, there is quite a lot of anecdotal evidence that governments have used PPPs to sidestep normal budgetary provisions. A PPP allows the current government to spend in infrastructure without adding the expenditure into the current budget. It also allows the current government to grant future subsidies—hence the popularity of minimum income guarantees—. The absence of rigorous accounting rules that force governments to account for these intertemporal transfers may stem from the belief that PPPs are a sort of imperfect privatization. Our model suggests that there is no conceptual reason whatsoever to excuse PPPs from normal budgetary practices.

References

[1] Arrow-Lind (1970).

[2] E. Chadwick, Results of different principles of legislation in Europe, Journal of the Royal Statistical Society A 22 (1859), 381-420.

[3] H. Demsetz, Why Regulate Utilities, Journal of Law and Economics 11 (1968), 55-66.

[4] E. Engel, R. Fischer, A. Galetovic, Infrastructure franchising and government guarantees, in T. Irwin, M. Klein, G. E. Perry, M. Thobani (Eds.), Dealing with Public Risk in Private Infrastructure, The World Bank, Washington, D.C., 1997.

[5] E. Engel, R. Fischer, A. Galetovic, Least-present-value-of-revenue auctions and highway franchising, Journal of Political Economy 109 (2001), 993-1020.

[6] J. L. Guasch, Concessions: boom or bust? Determinants of failure and success: an empirical study of a decade of experience, paper presented at the Fifth Annual Meeting of the Latin American and the Caribbean Economic Association (LACEA), Rio de Janeiro, Brazil, 2000.

[7] R. Harstad, M. Crew, Franchise bidding without holdups: utility regulation with efficient pricing and choice of provider, Journal of Regulatory Economics 15 (1999), 141-163.

[8] J.J Laffont, J. Tirole, A Theory of Incentives in Procurement and Regulation. MIT Press, Cambridge 1993.

[9] R. Posner, The appropriate scope of regulation in cable television, The Bell Journal of Economics, 3 (1972), 335-358.

[10] M.H. Riordan, D.E.M. Sappington, Awarding monopoly franchises, American Economic Review 77 (1987), 375-387.

[11] Spengler, J., Vertical integration and anti-trust policy, Journal of Political Economy 58 (1950), 347-352.

[12] D. Spulber, Regulation and Markets, MIT Press, Cambridge 1989.

[13] G. Stigler, The Organization of Industry, Richard D. Irwin, Homewood 1968.

[14] O.E. Williamson, Franchise bidding for natural monopolies-in general and with respect to CATV, Bell Journal of Economics 7 (1976), 73-104.

[15] O.E. Williamson, The Economic Institutions of Capitalism, The Free Press, New York 1985.

Appendix

A The general case with toll-responsive demand

The assumptions and notation were introduced in the main text.

Lemma 2 *For $\eta \geq 1$ define $H_i(P,\eta) \equiv G_i(P) + (\eta - 1)R_i(P)$. We assume $P_i^*(\eta) \equiv$ argmax $H_i(P,\eta)$ is well defined (that is, there exists a unique P where the maximum is attained) and that the maximum is interior. [These assumptions should be derived from first principles]. Then:*

(a) $P^*(\eta_i)$ satisfies:

$$G_i'(P_i^*(\eta)) + (\eta - 1)R_i'(P_i^*(\eta)) = 0 \tag{22}$$

and hence

$$1 - \frac{G_i'(P_i^*(\eta))}{R'(P_i^*(\eta))} = \eta, \tag{23}$$

(b) $P_i^*(\eta)$ is strictly increasing in η,

(c) $H_i^*(\eta) \equiv H_i(P_i^*(\eta),\eta)$ is strictly increasing in η.

(d) The function $\eta \longrightarrow P_i^*(\eta)$ maps the interval $[\lambda, \lambda\zeta]$ onto the interval $[P_i^*(\lambda), P_i^*(\lambda\zeta)]$. That is, the inverse function, mapping $P \in [P_i^*(\lambda), P_i^*(\lambda\zeta)]$ to $\eta \in [\lambda, \lambda\zeta]$ is well defined.

Proof: We drop the subindex i in what follows.

(a) Trivial.

(b) Applying the Implicit Function Theorem to (22), we obtain:

$$[P^*]'(\eta) = -\frac{R'}{G'' + (\eta - 1)[2Q' + PQ'']},$$

where all functions on the right hand side are evaluated at $P^*(\eta)$. That $P^*(\eta)$ is strictly increasing in η now follows from Assumptions A1 and A2.

(c) Assume $\eta_1 > \eta_2$. We then have:

$$H^*(\eta_1) \; = \; H^*(P^*(\eta_1),\eta_1) \; \geq \; H^*(P^*(\eta_2),\eta_1) \; =$$

$$= \; G(P^*(\eta_2)) + (\eta_1 - 1)R(P^*(\eta_2)) \; > \; G(P^*(\eta_2)) + (\eta_2 - 1)R(P^*(\eta_2)) \; = \; H^*(\eta_2).$$

The fist inequality follows from $H(\cdot,\eta_1)$ attaining its maximum at $P^*(\eta_1)$. The second inequality because $\eta_2 > \eta_1$ and $R > 0$.

(d) Follows from (a), (b) and smoothness properties of R_i and G_i. ∎

Planner's Problem

The planner chooses P_i^F, P_i^A, L_i and S_i, $i = 1, \ldots, n$, to solve:

$$\begin{aligned}
\max \quad & \sum_{i=1}^{n} \pi_i \left[H_i(P_i^F, 0)(1 - L_i) + H_i(P_i^A, \lambda) L_i - \lambda \zeta r S_i \right] \\
\text{s.t.} \quad & \sum_{i=1}^{n} \pi_i u \left(\frac{R_i(P_i^F)}{r} (1 - L_i) + S_i - I \right) = u(0), \\
& 0 \le L_i \le 1, \\
& S_i \ge 0.
\end{aligned}$$

Proposition 12 *In the optimal contract we have $P_i^A = P_i^*(\lambda)$.*

Proof: Since P_i^A does not appear in the problem's constraints, the optimal value of P_i^A is obtained by maximizing the subexpression in the objective function that includes this price. The result then follows from Lemma 2.[24] ∎

Lagrangian and Necessary Optimality Conditions

Using the result in Proposition 12 we have that the problem's Lagrangian is:

$$\mathscr{L} = \sum_{i=1}^{n} \pi_i \left[H_i(P_i^F, 0)(1 - L_i) + H^*(\lambda) L_i - \lambda \zeta r S_i \right] + r\mu \left[\sum_{i=1}^{n} \pi_i u \left(\frac{R_i(P_i^F)}{r} (1 - L_i) + S_i - I \right) - u(0) \right],$$

where $r\mu$ denotes the multiplier associated with the firm's participation constraint.

Proposition 12 allows us to ignore first order conditions w.r.t. P_i^A. The remaining first order conditions are the participation constraint and

$$\frac{\partial \mathscr{L}}{\partial P_i^F} = 0, \tag{24}$$

$$\left[L_i = 0 \text{ and } \frac{\partial \mathscr{L}}{\partial L_i} \le 0 \right] \text{ or } \left[0 < L_i < 1 \text{ and } \frac{\partial \mathscr{L}}{\partial L_i} = 0 \right] \text{ or } \left[L_i = 1 \text{ and } \frac{\partial \mathscr{L}}{\partial L_i} \ge 0 \right], \tag{25}$$

$$\left[S_i = 0 \text{ and } \frac{\partial \mathscr{L}}{\partial S_i} \le 0 \right] \text{ or } \left[S_i > 0 \text{ and } \frac{\partial \mathscr{L}}{\partial s_i} = 0 \right], \tag{26}$$

$$\frac{\partial \mathscr{L}}{\partial \mu} = 0 \Longleftrightarrow \sum_{i=1}^{n} \pi_i u \left(\frac{R_i(P_i^F)}{r} (1 - L_i) + S_i - I \right) = u(0). \tag{27}$$

The last condition is the firm's participation constraint, where we incorporate immediately that it holds with equality.

[24] The proof above applies if $L_i > 0$. If $L_i = 0$, then the objective function does not depend on P_i^A and this variable can take any value, including $P_i^*(\lambda)$.

Lemma 3 *We have:*

$$\frac{\partial \mathscr{L}}{\partial P_i^F} = \pi_i(1-L_i)\left[G_i' - R_i' + \mu u'\left(\frac{R_i}{r}(1-L_i)+S_i-I\right)R_i'\right], \tag{28}$$

$$\frac{\partial \mathscr{L}}{\partial L_i} = \pi_i\left[H_i^*(\lambda) - H(P_i^F,0) + \left(\frac{G_i'}{R_i'}-1\right)R_i\right], \tag{29}$$

$$\frac{\partial \mathscr{L}}{\partial S_i} = -\pi r\left[\lambda\zeta + \frac{G_i'}{R_i'} - 1\right], \tag{30}$$

where R_i, R_i' and G_i' are evaluated at P_i^F.

Proof: The derivation of (28) is straightforward. When deriving (29) and (30), equation (28) is used to get rid of the participation constraint multiplier. ∎

Corollary 2 *We have:*

$$\frac{\partial \mathscr{L}}{\partial L_i} = 0 \iff P_i^F = P_i^*(\lambda), \tag{31}$$

$$\frac{\partial \mathscr{L}}{\partial S_i} = 0 \iff P_i^F = P_i^*(\lambda\zeta). \tag{32}$$

Combining these results with Lemma 2b we have:

$$\frac{\partial \mathscr{L}}{\partial L_i} \leq 0 \iff P_i^F \geq P_i^*(\lambda), \tag{33}$$

$$\frac{\partial \mathscr{L}}{\partial L_i} \geq 0 \iff P_i^F \leq P_i^*(\lambda), \tag{34}$$

$$\frac{\partial \mathscr{L}}{\partial S_i} \leq 0 \iff P_i^F \leq P_i^*(\lambda\zeta). \tag{35}$$

Proof: A straightforward calculation, based on Lemma 2, shows that $\partial\mathscr{L}/\partial L_i$ evaluated at $P_i^*(\lambda)$, and $\partial\mathscr{L}/\partial S_i$ evaluated at $P_i^*(\lambda\zeta)$, are equal to zero. ∎
exist no other prices s.t. partial derivatives are equal to zero]

Proposition 13 *In the optimal contract we have $L_i < 1$.*

Proof: Analogous to the proof provided in the main text in the case of perfectly inelastic demand. ∎

Proposition 14 *In the optimal contract:*

(a) If $L_i > 0$ then $P_i^F = P_i^*(\lambda)$.

(b) If $S_i > 0$ then $P_i^F = P_i^*(\lambda\zeta)$.

(c) If $L_i = 0$ and $S_i = 0$ then $P_i^*(\lambda) \le P_i^F \le P_i^*(\lambda\zeta)$.

(d) If $S_i > 0$ then $L_i = 0$.

(e) If $L_i > 0$ then $S_i = 0$.

Proof: $L_i > 0$ implies that $\partial\mathcal{L}/\partial L_i = 0$ and part (a) follows from (31). $S_i > 0$ implies that $\partial\mathcal{L}/\partial S_i = 0$ and part (b) follows from (32). If $L_i = 0$ and $S_i = 0$, we have $\partial\mathcal{L}/\partial L_i \le 0$ and $\partial\mathcal{L}/\partial S_i \le 0$, and (c) follows from (33) and (35). Finally, parts (a) and (b) imply that $L_i > 0$ and $S_i > 0$ cannot occur simultaneously. Parts (d) and (e) then follow. ∎

The results above allow us to partition demand states into three categories, depending on whether the contract is finite or not, and on whether it involves subsidies:

1. States where $L_i > 0$ and $S_i = 0$.
2. States where $L_i = 0$ and $S_i = 0$.
3. States where $L_i = 0$ and $S_i > 0$.

For reasons that will become apparent in the next proposition, we call these states *high, intermediate* and *low* demand states.

Proposition 15 *Total income (present value of fees plus subsidies) and profits (total income minus investment) in high demand states is larger than in intermediate demand states, which is larger than in low demand states. Also, total income (and profits) is the same across high demand states, and across low demand states. Income may vary across intermediate demand states.*

Proof: From (28) we have that in the optimal contract:

$$\mu u'(\mathrm{Rev}_i - I) = 1 - \frac{G_i'(P_i^F)}{R_i'(P_i^F)}, \tag{36}$$

where Rev_i denotes total (present discounted) revenue in state i. It then follows from (23) and Proposition 14 that $\mu u'(\mathrm{Rev}_i - I) = \lambda$ when i is a high demand state, $\mu u'(\mathrm{Rev}_i - I) = \eta_i$, $\lambda \le \eta_i \le \lambda\zeta$ when i is an intermediate demand state, and $\mu u'(\mathrm{Rev}_i - I) = \lambda\zeta$ when i is a low demand state. All results now follow from the fact that u' is strictly decreasing. ∎

Proposition 16 *All states are high demand states if and only if*

$$\min_i R_i(P_i^*(\lambda)) \ge rI. \tag{37}$$

Proof: From Proposition 15 and the firm's participation constraint we have that all states are high demand states if and only if $\mathrm{Rev}_i = I$ for all i.[25] From Proposition 14 it follows that for all i we must have

[25] Proposition 15 implies that total income is constant across states while the participation constraint implies that this constant is equal to I.

$P_i^F = P_i^*(\lambda)$ and $S_i = 0$. Thus all states are high demand states if and only if the present value of user fees collected charging $P_i^*(\lambda)$ indefinitely is at least I. Condition (37) is necessary and sufficient for this to be the case.

The optimal contract in this case sets $P_i^F = P_i^A = P_i^*(\lambda)$, $S_i = 0$ and $L_i = 1 - rI/R_i(P_i^*(\lambda))$. Condition (37) ensures that $L_i \geq 0$. ∎

Characterization of the Optimal Contract

Given the participation constraint multiplier, μ, the results above imply that the optimal contract can be fully specified as follows:

1. For $i = 1, ..., n$ find η_i that solves:

$$\mu u'\left(\frac{R_i(P_i^*(\eta_i))}{r} - I\right) = \eta_i. \tag{38}$$

2. Those states where $\eta_i \leq \lambda$ are high demand states. For these states $P_i^F = P_i^A = P_i^*(\lambda)$ and $S_i = 0$, and $L_i \geq 0$ is chosen so that

$$\mu u'\left(\frac{R_i(P_i^*(\eta_i))}{r}(1 - L_i) - I\right) = \lambda.$$

3. Those states where $\lambda < \eta_i \leq \lambda\zeta$ are intermediate demand states. In these states $P_i^F = P_i^*(\eta_i)$, $L_i = 0$ and $S_i = 0$.

4. Those states where $\eta_i > \lambda\zeta$ are low demand states. For these states $L_i = 0$, $P_i^F = P_i^*(\lambda\zeta)$ and S_i is chosen so that

$$\mu u'\left(\frac{R_i(P_i^*(\eta_i))}{r} + S_i - I\right) = \lambda\zeta.$$

The following proposition uses the algorithm above and the participation constraint to characterize the optimal contract.

Theorem 2 (Characterization of the Optimal Contract) *Denote by $Rev_i(\mu)$ total income in state i assigned by the algorithm above for a given value of μ and define*

$$C(\mu) \equiv \sum_{i=1}^{n} \pi_i u(Rev_i(\mu) - I).$$

Then there exists a unique value of μ that solves $C(\mu) = u(0)$. This is the participation constraint multiplier and the algorithm above, for this particular value of μ, fully characterizes the optimal contract.

Proof: The Implicit Function Theorem and (38) imply that

$$\eta_i'(\mu) = \frac{u_i'}{1 - \mu u_i'' \frac{R_i'[P_i^*]'}{r}},$$

where u_i' and u_i'' denote u' and u'' evaluated at $R_i(P_i^*(\eta_i))/r - I$ and the remaining notation (hopefully) is obvious. It follows that η_i, and therefore C are strictly increasing in μ. ∎

Theorem 3 (Implementation) *Optimal contract can be implemented via a PVR auction. [Details to be provided].*

Proof: The crucial element is that total revenue in states of demand with a finite franchise length is larger than total revenue in states of demand with an infinite term contract. [Details to be provided]. ∎

Theorem 4 (Comparative Statics) *Rev_i is strictly increasing in I, P_i^F is increasing in I, S_i is increasing in I and L_i is decreasing in I.*

Proof: The Implicit Function Theorem and (38) lead to:

$$\eta_i'(I) = \frac{\mu'(I)u_i' - \mu(I)u_i''}{1 - \mu u_i'' \frac{R_i'[P_i^*]'}{r}}.$$

Since we have $\mu'(I) > 0$ it follows that η_i is increasing in I. The remaining results now follow from the implementation algorithm described above. ∎

B A model with effort

Define TR_i, $i = 1, \ldots, n$ to be total revenue in state i, and assume that effort e affects the probabilities of the different states $\pi_i(e)$, with $\sum_1^n \pi_i(e) = 1$ and $\pi_i > 0$, $\forall i$. Assume that utility of the franchise holder is separable in net income and effort thus: $U(y, e) = u(y) - ke$. Then the planner's problem can be written as:

$$\begin{aligned} \min \quad & \lambda \sum_i \pi_i(e)[TR_i + \zeta S_i] - (\lambda - 1)\sum \pi_i(e) TR_i^\infty \\ s.t. \quad & \sum_i \pi_i(e) u(TR_i + S_i - I) \geq u(0) + ke && (39) \\ \forall e' \neq e \quad & \sum_i \pi_i(e') u(TR_i + S_i - I) - ke' \leq \sum_i \pi_i(e) u(TR_i + S_i - I) - ke && (40) \end{aligned}$$

where (39) is the participation constraint and (40) is the incentive compatibility constraint. Under standard assumption we can use the first order approach and replace (40) by

$$\sum_i \pi_i'(e) u(TR_i + S_i - I) - k = 0 \qquad (41)$$

This is equivalent to maximizing

$$\min \lambda \sum_i \pi_i(e)\,[TR_i + \zeta S_i] - (\lambda - 1)\sum_i \pi(e)\,TR_i^{\infty} - \mu\left[\sum_i \pi_i(e) u(TR_i + S_i - I) - ke\right] - \tau\left[\sum_i \pi_i'(e) u(TR_i + S_i - I) - k\right] \quad (42)$$

The FOC with respect to e are:

$$\lambda \sum_i \pi_i'(e)\,[TR_i + \zeta S_i] - (\lambda - 1)\sum_i \pi_i'(e)\,TR_i^{\infty} - \tau \sum_i \pi_i''(e) u(TR_i + S_i - I) = 0 \quad (43)$$

with respect to S_i:

$$\lambda\zeta\pi_i(e) - \mu\pi_i(e) u'(TR_i + S_i - I) - \tau\pi_i'(e) u'(TR_i + S_i - I) = 0 \quad (44)$$

using (43):

$$\tau = \frac{\sum_i \pi_i'(e)[\lambda(TR_i + \zeta S_i) - (\lambda - 1)\,TR_i^{\infty}]}{\sum_i \pi_i''(e) u(TR_i + S_i - I)} \quad (45)$$

The FOC with respect to TR_i:

$$\lambda\pi_i(e) - \mu\pi_i(e) u'(TR_i + S_i - I) - \tau\pi_i'(e) u'(TR_i + S_i - I) \quad (46)$$

from (43):

$$u_i' = \frac{\lambda\zeta\pi_i(e)}{\mu\pi_i(e) + \tau\pi_i'(e)}$$

from which

$$u_i' = \begin{cases} \dfrac{\lambda\zeta}{\mu + \tau(\pi_i'(e)/\pi_i(e))} & \text{if } S_i \text{ is interior.} \\ \dfrac{\lambda}{\mu + \tau(\pi_i'(e)/\pi_i(e))} & \text{if } TR_i \text{ is interior.} \end{cases} \quad (47)$$

Since we have assumed that higher states involve higher demand, this means that higher states have smaller corresponding u_i'. By (47), this is equivalent to having an increase in i leading to an increase in $\pi_i'(e)/\pi_i(e)$. Hence, in both cases represented in (47) we have that increases in i lead to increases in $TR_i + S_i$.

Assuming that $\pi_i'(e)/\pi_i(e)$ is increasing in i (the Monotone likelihood ratio property?), and that higher effort leads to a higher probability of the high demand states, we have that, in order to command effort, the planner must introduce (additional) risk for the franchise holder.