Improving the payment mechanism in transport public–private partnerships

Antonio Sánchez Soliño and Vicente Alcaraz Carrillo de Albornoz

Universidad Politécnica de Madrid, Spain

ABSTRACT

Based on agency theory, this paper shows that the payment mechanism in a transport public–private partnership (PPP) should generally combine a fixed payment to the contractor, a payment based on service quality and a payment relating to the number of users. The transfer of demand risk can be totally excluded only if the public authority can define and verify a series of indicators that cover all the performance dimensions of the service.

IMPACT

The payment mechanism in transport PPPs is a key element of the contracts, since it defines the system of incentives and risks transferred to the contractor. Starting from the idea that incentives and risks are closely related to each other, this paper addresses the problem of transfering demand and performance risks. The authors assume that the public authority sets the parameters of the payment mechanism included in the contract in order to maximize social benefit. As a result, this paper establishes the general principles for an optimal payment mechanism in PPP contracts.

There has been a growing tendency in numerous countries over recent decades to externalize infrastructures and public services traditionally provided by local and central government. This has particularly been the case in the field of transport with the development of contract formulas based on public–private partnerships (PPPs) (Hodge & Greve, 2007; Yang, Hou, & Wang, 2013). The PPP concept is based on the idea that the public sector will no longer construct or purchase infrastructure that will later serve to support public services, but that a non public sector agent will be contracted to provide that service, under specific performance conditions, over an extended period of time (Grout, 1997).

Within this broad definition of PPPs, this paper focuses on private finance initiative (PFI) contracts, where a public sector organization (and not the users) pays for the services provided (Yescombe, 2007). According to the European PPP Expertise Center, PFI contracts are long-term agreements (typically 20 to 35 years) where the private sector builds the project's assets and raises the required funding, and where contractual payments from the public sector are the primary security for lenders (EPEC, 2012).

In PFI agreements, the payment mechanism from the public authority to the private partner is key in ensuring the economic efficiency of the contract, since it defines the distribution of risks between the parties and the contractor's incentives. Although some public bodies have paid attention to this issue (HM Treasury, 2007; EPEC, 2011), the design of the

payment mechanism is usually based on private sector practice. For example, in the first PFI road projects in the UK the payment mechanism was based on shadow tolls. However, more recent projects tend to be based only on performance measures (Standard & Poor's, 2003), when there is no reason for these extreme solutions. This paper presents an analytical tool for the public sector to set up an optimal payment mechanism in transport projects.

The academic literature has tended to look at PPPs from the perspective of economic efficiency. It focuses on the drivers for making the decision about the most suitable procurement procedure: PPP or traditional procurement. Payment mechanisms are generally analysed in this context—not as the object of study, as we do here.

This paper is based on a principal–agent approach, where the principal is the public authority and the agent is the contractor in a PPP agreement. The model provided is based on the measure of demand and performance as indicated by the (imperfect) signals of the effort made by the private partner. Consequently, there will be a certain degree of transfer of demand and performance risks, but the risk aversion of the agent will limit the extent of this transfer. Following the multi-task agency approach developed by Holmström and Milgrom (1991), this paper considers that performance is a complex variable. In a general case, there will be some service performance dimensions that may, to a certain extent, be verified and contracted by the public

authority, while there will be others that may be observed by the users, but not verified by that authority (Iossa & Martimort, 2009). Setting up and managing a PPP project will generally require important resources—such as a group of public servants who will, on a regular basis, physically inspect the underlying asset and control the performance of the service. Even when public resources are plentiful (rarely the case), it would not make economic sense to conduct a fully comprehensive overview of all the aspects related to the PPP contract, since the cost of such measures would be offset by the benefits obtained. In many cases, this overview can be replaced by a payment mechanism which is partly based on usage.

Additionally, our model considers two random sources in the contract:

- The presence of unpredictable factors that affect the service demand.
- A possible variation between the efforts made by the agent and the effective results obtained with respect to performance levels.

Literature review

In the recent academic literature on PPPs, there has been a lot of work reported on effective risk allocation between the public and the private sectors. Any contract to provide public services implies risk sharing between the contracting parties. These risks are particularly significant in PPPs (Ke, Wang, & Chan, 2012; Carbonara, Constantino, Gunnigan, & Pellegrino, 2015), which tend to last for long periods of time and frequently require considerable initial investment that is often irreversible. A broad description and classification of these risks may be found in Loosemore (2007) and in Palma, Leruth, and Prunier (2009).

Generally speaking, it is accepted that the majority of the technical and economic risks in PPPs (as is the situation with construction risks at the initial stage of the contract, or the risks of variation in operating costs) are going to be assumed by the contractor, while the public sector takes the political risks (such as opposition to the project by the general public) or those arising from situations of *force majeure*, such as natural disasters (Yescombe, 2007).

However, the greatest difficulty in the allocation of risks is the demand risk, i.e. the risk posed by a significantly different service demand to that initially forecasted. This explains the numerous studies that have been conducted on the characterization and management of demand risk. By way of example, Roumboutsos and Pantelias (2015), have shown that theory and practice are divergent in the allocation of revenue risk (a broader concept that includes demand

risk and counter-party risk), exposing projects to potential failure. Flyvbjerg, Skamris, and Buhl (2005) have shown that, in procurement processes for transportation infrastructure projects, tenderers tend to be overly-optimistic and present overestimated demand forecasts in order to win a contract, in the hope that this can then be renegotiated at a later stage (see also Eliasson & Fosgerau, 2013). However, the presence of this biased and overly-optimistic conduct should not hide the presence of a real demand risk which should be allocated by the contract to one of the parties, whether in full or shared.

Together with demand risk, another key aspect in PPPs is the transfer of the performance risk to the private sector. This risk essentially refers to the consequences for the agent in the event of failing to supply the services specified in the contract or doing so at a lower standard to the one required. In many recent contracts, transfer of the performance risk has taken the form of a payment to the private agent in accordance with performance levels. These performance levels are measured in relation to indicators that are specified in the contract.

A considerable amount of specialized literature has been published on the use of performance indicators in the field of transportation (Mandri-Perrot, 2010; Harding, Bodarwé, & Cadež, 2010; Federal Highway Administration, 2011). In parallel, a suitable theoretical framework in which to conduct economic analyses of PPPs is agency theory, which has undergone considerable development since its origins in the 1970s. Agency theory attempts to explain the agency relationship whereby one party (the principal) delegates certain tasks to another party (the agent), and where this relation is regulated by a contract or by similar means (Jensen & Meckling, 1976). The agency problem arises under conditions of asymmetric information between the parties to a contract, when there is a conflict of interest and where it is difficult, or expensive, for the principal to verify what the agent is actually doing (Fama & Jensen, 1983; Eisenhardt, 1989).

The problem of the agency relationship becomes more complex when one or more of the variables involved (such as demand) has a random component. In these cases the efficient sharing of risks between the principal and the agent requires an incentive mechanism to be put in place. This is because, in agency theory, it is generally assumed that the agent is risk averse and the principal is risk neutral or, at least, that the agent is more averse to risk (Eisenhardt, 1989; Dewatripont & Legros, 2005; Sadka, 2007).

Agency theory has been applied to PPPs by Dewatripont and Legros (2005), Martimort and Pouyet (2008) and Iossa and Martimort (2015) in order to compare PPPs with more traditional public procurement.

Research to date has tended to focus on two phases in the provision of public sector infrastructure and public services: a construction phase and a service operation phase. PPP contracts are characterized by the bundling of both activities, while under traditional contracting, construction and operation would be contracted by the government separately. The optimal choice of one or other procedure will depend on the existence of externalities (positive or negative) between both phases, taking into account the different incentives for contractors in a PPP arrangement versus traditional contracting (Hart, 2003).

This paper, in contrast, looks at an existing PPP contract focusing on the payment mechanism to the private partner with the aim of maximizing a social benefits.

Model formulation

Initial assumptions

Assume that a certain public authority (the principal) is responsible for the provision of a specific transport service, which may entail the need to build a particular infrastructure. The availability of the infrastructure and the provision of the service are made possible by a PPP contract (a PFI). The public authority wants to maximize the social benefit by optimizing the parameters that define the payment mechanism of the contract.

The service in this case is free to users, but our results can be extended to PPPs where users pay a fixed fee and the service is subsidized by the public sector. We assumed that service demand would depend to some extent on the performance levels offered, and that these performance levels, in turn, would depend on the efforts made by the agent to improve the service. Generally speaking, an agent's activity will include many dimensions and the public partner might be able to observe and verify some of these through the definition and monitoring of performance indicators (Iossa & Martimort, 2009). However, the authority may have difficulty in establishing suitable indicators or controlling the compliance which might affect the performance perceived by the users.

A good example is an underground (subway) transport service. Some performance parameters (for example service frequency, punctuality) can be readily observed and checked by the authorities. However, there are other performance dimensions that may be observed by the users but which are far more difficult to define in terms of performance standards and which are more difficult to control (for example surveillance on stations and in trains, user information, queues at ticket dispensers or booths).

The following demand function can be formulated for this case:

$$D = D_0 + k_1 q_1 + k_2 q_2 + \varepsilon_d \qquad (1)$$

D represents the service demand (normally measured by the number of users) and D_0 is a minimum demand level. This minimum demand level takes into account the fact that there are a certain number of 'captive' users in a transport service, basically on account of the lack of alternatives. The variable q_1 represents the performance level of the service with respect to those aspects that may be verified by the authorities and k_1 is a constant (≥ 0) which considers the impact of these performance aspects on the level of service demand. The next variable (q_2) represents the performance level obtained by the agent for those aspects that may be observed by the users but which cannot be measured and verified by the authorities, and where k_2 (≥ 0) is a constant that reflects the impact of these performance aspects on demand. By definition, q_1 and q_2 are independent variables, since the dimensions of performance that are captured in q_2 cannot be measured by the authorities.

It is assumed that the service demand will have an additional, random, component, independent of the others, and represented by the variable ε_d. This random variable incorporates the effect of external factors on the service demand, such as macro-economic conditions, population shift, changes in the attitudes or customs of users and so on. This random component is considered to follow a probabilistic distribution with a mean of zero and a standard deviation equal to σ_d.

The performance levels q_1 and q_2 are related to the respective efforts made by the agent by means of the following expressions:

$$q_1 = e_1 + \varepsilon_{q1} \qquad (2)$$

$$q_2 = e_2 + \varepsilon_{q2} \qquad (3)$$

The variables e_1 and e_2 represent the efforts made by the agent, while the variables ε_{q1} and ε_{q2} represent a random component of the performance levels obtained in each case. This implies that the agent does not totally control the service performance, as this does not purely depend on the agent's efforts, but also on other unpredictable values that are contained within these random variables. From a different perspective, this could also be interpreted as implying in the case of q_1 that the authorities cannot directly observe the efforts made by the agent and can only verify the level reached in terms of a set of performance indicators. This can lead to a disparity between the performance level and the effort applied. A good example of this is PPP contracts for roads. One group of performance

indicators that are sometimes used are those related to the number of accidents and fatalities on part of a road (Rangel, Vassallo, & Arenas, 2012). However, accidents and fatalities do not depend solely on the state and management of the road, but on many other factors that are beyond the agent's control.

In the model, it is assumed that ε_{q1} and ε_{q2} follow a normal distribution pattern, with an average equal to zero and a standard deviation equal to σ_{q1} and σ_{q2} respectively. As such, in this model, the agent operates under risk conditions that are due to the randomness of the variables involved. The public sector authority, however, is likely to be sufficiently diversified, in terms of the number of projects and services for which they are responsible, to be risk neutral (Eisenhardt, 1989).

The model then considers the following cost function for the production of the service:

$$C = C_0 + cD + \Psi_1(e_1) + \Psi_2(e_2) \quad (4)$$

Where C represents the cost of producing the service and C_0 is a fixed cost term incorporating the cost derived from any necessary initial investment for delivering the service. This fixed cost term may be interpreted as the minimum cost necessary to put the service into operation at a minimum performance level. The variable costs include a cost term which is linearly dependent on demand. In this equation, the constant c represents the marginal cost of the service and D represents the service demand. Variable costs also include two cost functions that depend on the levels of effort e_1 and e_2 respectively, where $\Psi_i'(e_i) > 0$ and $\Psi_i''(e_i) > 0$. In this regard, we assumed that the variable costs associated with performance increase more than proportionally to efforts e_1 and e_2. This is a key assumption of the model, which is in line with the growing marginal cost assumption used in micro-economic theory.

In order to simplify the exposition, while retaining the general application of the main results of the model, it is possible to consider a specific cost function that satisfies the condition of being strictly positive in its first and second derivatives. In economic literature, it is common to take the following type of quadratic cost function (Arrow & Radner, 1979; Gibbons, 1998; Rob & Zemsky, 2002; Socorro, 2007; Martimort & Pouyet, 2008):

$$\Psi_i(e_i) = e_i^2/2 \quad (5)$$

So the cost function becomes:

$$C = C_0 + cD + e_1^2/2 + e_2^2/2 \quad (6)$$

The use of this cost function simplifies the model without loss of generality in the main results of the model.

Regarding remuneration, it is assumed that the agent will receive an amount from the authority, which in this model is given by a generic linear equation, taking into account the optimality of such linear incentives as demonstrated by Holmström and Milgrom (1987, 1991):

$$t(D, q_1) = \alpha + \beta D + \omega q_1 \quad (7)$$

In this expression, t is the amount paid to the agent, and depends on both the service demand (D) and the performance level (q_1) in respect to those aspects that may be observed and verified by the authority; α, β and ω are fixed parameters established in the contract. The agent will receive (or pay—since the optimal value of α can be negative) a fixed sum given by the parameter α, together with a variable amount that includes one component in accordance with demand and another component that depends on the performance that may be verified by the authorities.

The agent's revenue function implicitly includes the random variables defined above and the greater the incentives established, the greater the significance of these variables. The problem then is to establish the value of the α, β and ω parameters of the contract that will maximize social benefits, when taking into account the trade-off between the agent's efforts and their risk aversion.

Agent's utility and social benefit functions

When taking into account the random components of the agent's remuneration, the expected value would be as follows:

$$E[t] = \alpha + \beta(D_0 + k_1e_1 + k_2e_2) + \omega e_1 \quad (8)$$

In addition to the social benefit function, defined later on in this paper, it is also necessary to consider the restrictions established on the basis of the agent's utility function. We assumed that the preferences of the contractor are of the mean variance type (Markowitz, 2014), and therefore:

$$Ur = E[t] - E[C] - r\sigma_z^2 \quad (9)$$

Where Ur is the objective function of the agent providing the service, E[t] the expected value of the payment made by the principal, and E[C] the expected cost of producing the service; σ_z^2 is the variance of the variable $z = t - C$, and r is the parameter representing the attitude of the agent toward risk. As such, and when expressed in other terms, Ur would then be the certainty equivalent of the returns expected by the agent.

In the case of the risk aversion of the latter, this would imply an r value strictly greater than zero, while in the case of risk neutrality this would suppose an r equal to zero. In this model it is taken that the

agent will normally be risk averse, excluding the possibility that the agent is a risk lover:

$$r\,\sigma_z^2 \geq 0 \qquad (10)$$

The variance σ_z^2 would be as follows:

$$\sigma_z^2 = [(\beta-c)^2 k_1^2 + \omega^2 + 2(\beta-c)k_1\omega]\sigma_{q1}^2$$
$$+ (\beta-c)^2 k_2^2 \sigma_{q2}^2 + (\beta-c)^2 \sigma_d^2 \qquad (11)$$

On developing equation (9), when considering (6) and (8), we get:

$$Ur = \alpha + (\beta-c)(D_0 + k_1 e_1 + k_2 e_2)$$
$$+ \omega e_1 - C_0 - e_1^2/2 - e_2^2/2 - r\sigma_z^2 \qquad (12)$$

In order for the agent to carry out their activity, they will have to obtain a utility greater or equal to the alternative of doing nothing. If it is assumed that not carrying out any activity represents a utility equal to zero, the agent's participation is then dependent on:

$$Ur \geq 0 \qquad (13)$$

In addition, the agent will try to maximize their objective function, given by (12). The agent's efforts e_1 and e_2 should then comply with the following first-order conditions:

$$e_1 = (\beta - c)k_1 + \omega \qquad (14)$$

$$e_2 = (\beta - c)k_2 \qquad (15)$$

The social benefit function will, in turn, be formed by the value of the benefits produced by the provision of the service for society as a whole, minus the costs incurred in making the service available, including the cost derived from the risk aversion of the agent. In this way, the expected social benefit function is given by:

$$Ws = E[S] - E[C] - r\sigma_z^2 - \lambda(E[t]) \qquad (16)$$

In equation (16), Ws is the expected social benefit that is to be maximized. The variable S represents the value of all the benefits produced by the transport service in question, including the surplus of service users and all externalities (positive or negative) produced, and E[S] represents the expected value of S. These benefits will depend both on the performance of the service and on the number of users:

$$S = f(q_1, q_2, D) \qquad (17)$$

For simplicity, a linear form of the function f can be assumed:

$$E[S] = S_0 + a_1 e_1 + a_2 e_2 \qquad (18)$$

Where a_1 and a_2 are constants that may be positive or zero. S_0 is a minimum benefit produced by the service when the agent makes zero effort.

Equation (16) includes the excess burden of public funds which is given by the term $\lambda(E[t])$, where λ is a parameter (attributed with a value greater than zero) that characterizes the tax system of the country in question. The value $(1+\lambda)$ is normally referred to as the marginal cost of public funds, this being a concept that incorporates various aspects and including the distortion introduced by the tax system in the decisions of the economic agents and the tax administration cost. In other words, it is assumed that the disutility to taxpayers inflicted by levying an additional monetary unit would be equivalent to $(1+\lambda)$ monetary units, where the value of λ depends on the institutional framework of each country and the tax provisions used to obtain additional public funds. This weighting of public funds is not taken into account in most cost-benefit analysis and it is, instead, implicitly established that $\lambda = 0$. However, Laffont and Tirole (1993) considered it reasonable to establish a value $\lambda = 0.3$ for the American economy and Kleven and Kreiner (2003) estimate λ values of between 0.09 and 0.80 for different OECD countries, considering proportional tax increases for the tax system as a whole. A detailed study of the concept of the marginal cost of public funds may be found in Dahlby (2008).

Results of the model

The solution to the maximization of Ws is given by:

$$Ur = 0 \qquad (19)$$

$$\beta^* = \frac{a_2 k_2 + c(1 + \lambda)(2rk_2^2\sigma_{q2}^2 + 2r\sigma_d^2)}{(1 + \lambda)(k_2^2 + 2rk_2^2\sigma_{q2}^2 + 2r\sigma_d^2)} \qquad (20)$$

$$\omega^* = \frac{a_1 + ck_1(1 + \lambda)2r\sigma_{q1}^2}{(1 + \lambda)(1 + 2r\sigma_{q1}^2)} - k_1\beta^* \qquad (21)$$

In this way, we obtain the values of the parameters of equation (7), which optimise the contract in terms of social benefits. The value of the fixed payment α^* is obtained in a residual manner as from the condition (19) and would largely depend on the fixed cost term C_0 and the minimum demand level D_0.

Therefore, in the event of the risk aversion of the agent $(r > 0)$, the transfer of the demand risk implies an additional cost term in the social benefit function, and the optimal situation would then be obtained by mitigating this transfer. However, the transfer of a certain degree of demand risk would be necessary in an optimal contract, providing that there were certain effects of the performance dimensions of the service that cannot be verified by the authorities, i.e. on the condition that $k_2 \neq 0$.

Table 1 summarizes the analysis of the derivatives of the incentive parameters with respect to the variables of the model.

The analysis of the derivatives obtained for the incentive parameters shows that the value of β^*

tends to decrease as the variances σ_{q2}^2 and σ_d^2 increase, as we expected. However, the value of β^* does not depend on the variance σ_{q1}^2: the optimal degree of transfer of the demand risk depends solely on the performance dimensions that cannot be verified by the authorities. Additionally, the value of β^* decreases as the agent's risk aversion (r) and the marginal cost of public funds $(1+\lambda)$ increase. On the other hand, the parameter β^* tends to increase when c (the marginal cost of the service) increases. However, no general conclusion can be reached on the sign of the derivative of β^* with respect to k_2. Nevertheless, for a rank of low values of k_2, it can be stated that β^* increases when k_2 increases, which is an intuitive result.

The value of the parameter ω^*, to the contrary, decreases as the variance σ_{q1}^2 increases. This means that, in the event of a high value of the variance σ_{q1}^2, it is inefficient to maintain a PPP contract fundamentally based on performance indicators with a high ω parameter. In addition, the value of ω^* increases with any increase in the value of the variance σ_d^2 (and, thus, a decrease in the value of β^*). As such, the contract should largely be based on performance payments when there is a high degree of randomness of the demand due to external factors. In other words, the optimal performance-based mechanism depends on both the system of performance indicators and on the parameters that define the optimal transfer of demand risk.

However, no conclusion can be reached with regards to the variation of ω^* when increasing the agent's risk aversion (i.e. increasing r). This means that an increase in the agent's risk aversion tends to decrease the transfer of the demand risk in an optimal contract, though the effect on the performance payment cannot be determined at the outset. In the same way, no conclusion may be reached with regards to the variation of ω^* when increasing the marginal cost of public funds or the marginal cost of the service.

Policy implications of the results

Some of the variables in our model are dependent on the general economic and institutional environment of the country in which the project is located. In

Table 1. Sign of the derivatives of optimal incentives.

Variables	β	ω
c	+	n.d.
k_1	0	−
k_2	n.d.	n.d
σ_{q1}^2	0	−
σ_{q2}^2	−	+
σ_d^2	−	+
r	−	n.d.
λ	−	n.d

n.d. = not determined

particular, the marginal cost of public funds $(1+\lambda)$ depends on the specific economic and fiscal system. For example, in countries with highly inefficient taxes, λ will be high and the optimal transfer of demand risk will therefore be lower. The evidence also suggests that the marginal cost of public funds tends to increase with the level of taxation burden in the economy (Barrios, Pycroft, & Savein, 2013).

A higher value of λ will lead to a lower value of β^*, i.e. the optimal transfer of demand risk will be lower. This does not contradict the result obtained by Auriol and Picard (2013), who compared build-operate-transfer (BOT) concessions with public management in a given facility, and found that the higher the cost of public funds the more governments were likely to choose BOT concessions. Indeed, since in a BOT concession there are no public funds involved, a higher λ implies a higher social cost of the public management compared to the alternative of a BOT concession. However, in PPP projects where there is a subsidy from public funds, a higher value of λ implies a lower transfer of demand risk.

Other variables depend on each particular project. Table 2 provides an assessment of the level of some key variables in urban rail and in road projects. These variables are the marginal cost of the service (variable c), the variable related to the elasticity of demand with respect to performance dimensions that cannot be verified by the authorities (k_2), and the variance of demand (σ_d^2).

The projects in Table 2 are ones where PPP agreements have undergone considerable expansion over recent decades. As can be seen, the demand risk (given by the variance of demand) is rated as high or very high in these kind of projects, according to the empirical evidence available (see, for example, Flyvbjerg et al., 2005). Generally speaking, this would mean that the transfer of demand risk should be limited in transport PPPs, taking into account equations (20) and (21). This result is in line with the research carried out by Engel, Fischer, and Galetovic (2001), in which the authors propose a contract based on a variable concession term (adjusted in turn according to the present value of revenues) as a mechanism to mitigate the demand risk.

However, there are some important differences in other aspects of each type of project, which are analysed below.

Urban rail projects are particularly complex, and the system of incentives for the private partner should be carefully designed (Sánchez Soliño & Vassallo, 2009).

Table 2. Assessment of the level of key variables in transport PPP projects.

Type of project	c	k_2	σ_d^2
Urban rail systems	High	Low	High/very high
Road PPPs	Low/medium	Very low	High/very high

In this type of project, the marginal cost of the service is rated as high and the payment mechanism should take into account the number of users, even when the variance of demand is high. It is also to be expected in this kind of project that the service performance may have a certain influence on demand. In this regard, HM Treasury (2007) states that 'payments linked to usage can bring advantages when the contractor's performance can influence the level of usage, since customers can "vote with their feet" on the availability and quality of the service'. This is in line with the model described in this paper, with the difference that, in the latter, only dimensions of performance that cannot be verified by the authorities must be considered to determine the payment linked to usage. When it is possible for the authorities to establish an appropriate indicator to verify the level reached by the private partner on a specific performance dimension, the latter is not taken into account in the calculation of the parameter β*.

In summary, the optimal incentives in urban rail projects would normally be mixed, based on demand and on performance. In fact, this is the situation in reality, where the actual arrangement in many urban rail projects takes into account both the number of users and a key performance indicator (KPI) system. A mixed payment arrangement (and/or a demand risk mitigation mechanism) can be found in a number of tram/light rail systems in, for example, the UK (Manchester Metrolink, Nottingham Express Transit), France (Rouen Metrobus) and Spain (Madrid Light Rail, Barcelona Tram) (National Audit Office, 2004; Mandri-Perrot, 2010; Carpintero & Petersen, 2014).

In *road projects*, the variable k_2 can be rated as very low (as it is possible, for roads, to establish a set of indicators, verifiable by the authorities, that covers the majority of the dimensions of performance: Federal Highway Administration, 2011) and the marginal cost is rated as low to medium, because most maintenance costs are partially fixed and do not depend on the traffic volume (Doll & van Essen, 2008). As a result, the parameter β*can be expected to be very low in road PPPs.

In the case of PFI road projects in the UK, there has been a clear evolution in the payment mechanism established in contracts (Standard & Poor's, 2003). In a first phase (1994–1999), the payment mechanism was primarily based on shadow tolls, where the public authority paid a fee for each vehicle or vehicle-km. This mechanism was applied in projects like the widening, operation and maintenance of the A1(M) motorway (128 million GBP of capital value) and the construction of the M1–A1 motorway link (214 million GBP), both awarded in 1996.

In 2000, the contract for the upgrading, operation and maintenance of the A13 Thames Gateway was awarded, with a capital value of 411 million GBP. The payment mechanism established in the contract considers shadow tolls exclusively for heavy vehicles, while the main component is based on several criteria that measured the performance of the private partner. In particular, approximately 70% of the private partner's income is linked to road availability which, in turn, incentivizes the project company to maximize the time that the road is available to road users, particularly during peak hours.

More recent contracts, since 2002, have eliminated the shadow toll component, and the payment mechanism relies exclusively on performance indicators. The latter have been extended to new dimensions of performance, like the reduction of congestion. One of the most important projects carried out under these criteria was the widening, operation and maintenance of the M25 awarded in 2009 (3,400 million GBP of present value of the payments to the private partner; National Audit Office, 2010).

This evolution in the payment mechanism used in the UK can be considered as part of a learning process in PFI contracts. According to Standard & Poor's (2003), in early-generation PFI road projects, the revenue stream was traffic-dependent. This proved problematic, given the uncertainty about revenue forecasts. In the most recent projects, however, performance-based mechanisms have reduced the volatility of expected revenues.

Overall, this is in line with the results obtained in this paper. A payment mechanism based primarily on usage represents a considerable amount of risk to the project company when the variance of demand is high. Therefore, the component of cost due to the risk aversion of the agent is large in the social benefit function. For this reason, the weight of the usage component should be low in road PPPs, according to the model presented in this paper. Nevertheless, at least a part of the payment should be based on the traffic volume, taking into account that the marginal cost of maintenance and operation of the road is by no means negligible. In other words, the total exclusion of a payment component based on usage would not be justified from the point of view of social benefit.

Conclusions

Our results show that in a PPP contract between a public sector authority and a private sector company for the provision of public transport, there is a base theory that justifies, under conditions of asymmetric information, a certain degree of transfer to the contractor of both the performance risk and the demand risk, due to the incentives this provides to the contractor to improve their performance. Furthermore, an optimal contract should also include a fixed

payment to the contractor, although the optimal fixed payment can be negative, i.e. a payment from the private partner to the public authority.

This means that the tendency in some countries to eliminate any payment component based on the number of users cannot be justified. To the extent that there are certain service performance dimensions that may be observed by the users, but not verified by the authorities, and that these performance dimensions have an influence on service demand, an optimal contract should incorporate some transfer of the demand risk. In many public services, including transport, there are some performance aspects that are very difficult to verify and to control by the authorities, but which are clear to the users.

However, it would also be incorrect in the majority of cases to base an agent's remuneration system only on the basis of the number of users, and the transfer of the demand risk should decrease as the variance of traffic and the agent's risk aversion increases.

Finally, the effects of the incentive system on the behaviour of the public administration itself have not been addressed in this paper. In particular, a PPP based solely on the transfer of performance risk (excluding the transfer of demand risk) could potentially constitute a perverse incentive for public authorities. They could find themselves in a situation in which they could avoid the budgetary discipline mechanism in the short term and, at the same time, avoid market control as the agent would not be subject to the demand risk; this could lead to the establishment of unnecessary and possibly unjustified projects. This issue is pointed out as a possible subject for further research.

Acknowledgements

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Arrow, K., & Radner, R. (1979). Allocation of resources in large teams. *Econometrica*, *47*(2), 361–385.

Auriol, E., & Picard, P. M. (2013). A theory of BOT concession contracts. *Journal of Economic Behavior & Organization*, *89*(C), 187–209.

Barrios, S., Pycroft, J., & Savein, B. (2013). *The marginal cost of public funds in the EU: The case of labour versus green taxes*, Taxation Papers, Working Paper n. 35. Luxembourg: European Commission.

Carbonara, N., Constantino, N., Gunnigan, L., & Pellegrino, R. (2015). Risk management in motorway PPP projects: Empirical-based guidelines. *Transport Reviews*, *35*(2), 162–182.

Carpintero, S., & Petersen, O. H. (2014). PPP projects in transport: Evidence from light rail projects in Spain. *Public Money & Management*, *34*(1), 43–50.

Dahlby, B. (2008). *The marginal cost of public funds: Theory and applications*. Cambridge, Massachusetts: The MIT Press.

Dewatripont, M., & Legros, P. (2005). Public–private partnerships: Contract design and risk transfer. *European Investment Bank Papers*, *10*(1), 120–145.

Doll, C., & van Essen, H. (2008). *Road infrastructure cost and revenue in Europe*. The Netherlands: CE Delft, Delft.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, *14*(1), 57–74.

Eliasson, J., & Fosgerau, M. (2013). Cost overruns and demand shortfalls – deception or selection? *Transportation Research Part B*, *57*, 105–113.

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

EPEC. (2011). *The guide to guidance: How to prepare, procure and deliver PPP projects*. Luxembourg: European PPP Expertise Centre.

EPEC. (2012). *United Kingdom-England. PPP units and related institutional framework*. Luxembourg: European PPP Expertise Centre.

Fama, E., & Jensen, M. (1983). Separation of ownership and control. *Journal of Law & Economics*, *26*, 301–324.

Federal Highway Administration. (2011). *Key performance indicators in public–private partnerships*. Washington DC: FHWA, U.S Department of Transportation.

Flyvbjerg, B., Skamris, M. K., & Buhl, S. L. (2005). How (In)accurate are demand forecasts in public works projects? The case of transportation. *Journal of the American Planning Association*, *71*(2), 131–146.

Gibbons, R. (1998). Incentives in organizations. *Journal of Economic Perspectives*, *12*(4), 115–132.

Grout, P. A. (1997). The economics of the private finance initiative. *Oxford Review of Economic Policy*, *13*(4), 53–66.

Harding, J., Bodarwé, H., & Cadež, I. (2010). Evaluation of availability and service performance based payment mechanism for PPP road traffic infrastructure projects. In *Transportation Research Board 89th Annual Meeting*. Washington DC.

Hart, O. (2003). Incomplete Contracts and Public Ownership: Remarks, and an Application to Public–private Partnerships. *The Economic Journal*, *113*, 69–73.

HM Treasury. (2007). *Standardisation of PFI contracts, version 4*. London: The Stationery Office.

Hodge, G. A., & Greve, C. (2007). Public–private partnerships: An international performance review. *Public Administration Review*, *67*(3), 545–558.

Holmström, B., & Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, *55*(2), 303–328.

Holmström, B., & Milgrom, P. (1991). Multitask principal–agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics and Organization*, *7*, 24–52.

Iossa, E., & Martimort, D. (2009). *The theory of incentives applied to the transport sector*, Working Paper 09/210. CMPO, Bristol University.

Iossa, E., & Martimort, D. (2015). The Simple Micro-Economics of Public–private Partnerships. *Journal of Public Economic Theory*, *17*(1), 4–48.

Jensen, M., & Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*, 305–360.

Ke, Y., Wang, S., & Chan, A. P. C. (2012). Risk management practice in Chinás public–private partnership projects. *Journal of Civil Engineering and Management*, *18*(5), 675–684.

Kleven, H. J., & Kreiner, C. T. (2003). *The marginal cost of public funds in OECD countries: Hours of work versus labor force participation*. CESifo Working Paper no. 935.

Laffont, J.-J., & Tirole, J. (1993). *A theory of incentives in procurement and regulation*. Cambridge, Mass.: MIT Press.

Loosemore, A. (2007). Risk allocation in the private provision of public infrastructure. *International Journal of Project Management*, *25*, 66–76.

Mandri-Perrot, C. (2010). *Private sector participation in light rail/light metro transit initiatives*. DC, Washington: The World Bank.

Markowitz, H. (2014). Mean-variance approximations to expected utility. *European Journal of Operational Research*, *234*(2), 346–355.

Martimort, D., & Pouyet, J. (2008). Build It or Not: Normative and positive theories of public–private partnerships. *International Journal of Industrial Organization*, *26*, 393–411.

National Audit Office. (2004). *Improving public transport in England through light rail*. London: The Stationery Office.

National Audit Office. (2010). *Procurement of the M25 private finance contract*. London: The Stationery Office.

Palma, A. d., Leruth, L., & Prunier, G. (2009). *Towards a principal–agent based typology of risks in public–private partnerships, IMF Working Paper 09/177*. DC, Washington: International Monetary Fund.

Rangel, T., Vassallo, J. M., & Arenas, B. (2012). Evaluation of the effectiveness of safety-based incentives in public–private Partnerships: Evidence from the case of Spain. *Transportation Research Part A: Policy and Practice*, *46*(8), 1166–1176.

Rob, R., & Zemsky, P. (2002). Social capital, corporate culture and incentive intensity. *RAND Journal of Economics*, *33*(2), 243–257.

Roumboutsos, A., & Pantelias, A. (2015). Allocating revenue risk in transport infrastructure public private partnership projects: How it matters. *Transport Reviews*, *35*(2), 183–203.

Sadka, E. (2007). Public–private partnerships – a public economics perspective. *CESifo Economic Studies*, *53*(3), 466–490.

Sánchez Soliño, A., & Vassallo, J. M. (2009). Using public private partnerships to expand subways: Madrid-barajas international airport case study. *Journal of Management in Engineering*, *25*(1), 21–28.

Socorro, M. P. (2007). Optimal technology policy under asymmetric information in a research joint venture. *Journal of Economic Behavior and Organization*, *62*(1), 76–97.

Standard & Poośs. (2003). *The evolution of DBFO payment mechanisms: One more for the road?* London: Standard & Poor's.

Yang, Y., Hou, Y., & Wang, Y. (2013). On the development of public–private partnerships in transitional economies: An explanatory framework. *Public Administration Review*, *73*(2), 301–310.

Yescombe, E. R. (2007). *Public–private partnerships: Principles of policy and finance*. Oxford: Elsevier.