Using a self-selection mechanism for tendering in the construction industry: A case study of railway renewal contracts

VTI Working Paper 2023:10

Jan-Eric Nilsson[1], Kristofer Odolinski[1], Johan Nyström[2]

[1] VTI, Swedish National Road and Transport Research Institute
[2] NYFOU, Nyström Forskning och Utredning

Abstract

One of the consequences of the institutional separation of railway infrastructure from train operations in Europe is a misalignment of incentives in which the actions of one party may create costs for the other. To internalise otherwise external costs of track-works experienced by train operators and customers, it is essential to reform the way in which project contracts are tendered. This study suggests a self-selection mechanism for tendering rail infrastructure activities. Bidders may therefore submit bids based on the industry's standard Unit Price Contract or a Fixed-Price Contract. The mechanism is designed to increase the possibility for a welfare maximising trade-off between construction and user costs. Using standard Benefit-Cost principles and parameter values, a case study where five switches are replaced provides substance to the discussion. The study provides a starting point for addressing risk in the construction industry and a blueprint for further development by professionals to fill in gaps and to test the approach under a controlled format before full-scale implementation.

Keywords

procurement; risk; rail infrastructure; vertical separation; delay fee; unit price contract; fixed-price contract

JEL Codes

H57; R42; R48

Using a self-selection mechanism for tendering in the construction industry: A case study of railway renewal contracts

Jan-Eric Nilsson[1], Kristofer Odolinski[1], Johan Nyström[2]

[1] VTI, Swedish National Road and Transport Research Institute

[2] NYFOU, Nyström Forskning och Utredning

Abstract

One of the consequences of the institutional separation of railway infrastructure from train operations in Europe is a misalignment of incentives in which the actions of one party may create costs for the other. To internalise otherwise external costs of track-works experienced by train operators and customers, it is essential to reform the way in which project contracts are tendered. This study suggests a self-selection mechanism for tendering rail infrastructure activities. Bidders may therefore submit bids based on the industry's standard Unit Price Contract or a Fixed-Price Contract. The mechanism is designed to increase the possibility for a welfare maximising trade-off between construction and user costs. Using standard Benefit-Cost principles and parameter values, a case study where five switches are replaced provides substance to the discussion. The study provides a starting point for addressing risk in the construction industry and a blueprint for further development by professionals to fill in gaps and to test the approach under a controlled format before full-scale implementation.

Keywords: procurement; risk; rail infrastructure; vertical separation; delay fee; unit price contract; fixed-price contract

1. Introduction

Sweden separated railway traffic operations from infrastructure management in 1988 and such vertical separation has subsequently become the norm for the European Union's railway industry (Nash et al. 2019). Later, the Swedish government instructed Trafikverket, its Infrastructure Manager (IM), to divest itself of all in-house construction and maintenance capacities. Today, the Swedish IM tenders contracts for these activities on a least-cost basis—see e.g. Odolinski (2015).

In contrast to road works, many maintenance and renewal measures in the railway industry cannot be implemented without blocking train traffic. In a vertically integrated industrial structure, the prioritisation of traffic versus renewal and line closures is managed within the same organisation. The new industrial structure means that this trade-off now involves multiple parties. The IM pays for using a builder's staff and equipment to perform a job while the train operators, passengers and freight customers carry the costs for traffic disruptions.

Currently, the renewal activity of switch replacement is tendered for implementation during weekends and requires comprehensive line closure while the work is performed. The IM specifies the timing for implementing the work. A commercial manufacturer of switches has suggested that it is feasible to replace old switches in-between scheduled trains with less disturbances to train traffic. This is an example of an innovative design that has not been tested in practice.

Like most projects in the railway industry, switch replacement is tendered and signed within the framework created by the Unit Price Contract (UPC)[1] institution, which is based on a detailed specification and quantification of all activities that the winning bidder must implement. Bidders submit a unit price for each quantity. Contracts are typically awarded to the lowest bidder after multiplying quantities by price.

[1] The International Federation of Consulting Engineers provides detailed information about UPCs which engineers refer to as Design–Bid–Build (DBB). The Federation also describes a fixed-price (FP) contract which is referred to as Design–Build (DB) as well as other types of contracts (cf. www.fidic.org). To focus on incentives, we use the economic concepts of UPC and FP contracts rather than DBB and DB, respectively.

This approach means that little scope is left for the agent awarded a contract to implement it in any other way than specified by the principal. Our hypothesis is that this hampers the dynamics of the industry, which is a potential reason for dismal cost performance. National accounts data from different European countries collected in the EU KLEMS[2] database (see Jäger 2017; The Economist, 2017; Salomonsson et. al., 2019) indicate that overall productivity growth in the European construction industry has been more broadly poor relative to other economic sectors;[3] however, no systematic explanation of this dismal performance has been suggested.

Based on the above conjecture, we subsequently propose that the tendering process be augmented to make it feasible to submit bids using a less restrictive contract format. The purpose of the study is to describe a self-selection mechanism in which bidders for a project may use the standard UPC format but where it is also possible to submit a fixed-price (FP) bid for the contract. The mechanism also includes a means for comparing bids submitted under the two contract formats which is designed to minimise social costs for the work done. Because of problems to establish an analytical model that would specify the efficiency properties of the proposed approach (cf. next section), and since data are not available for econometric analysis, a case study is used to illustrate the proposal.

The proposal makes use of contributions from three strands of the research literature. First, in the planning and prioritisation of infrastructure investment projects, Benefit-Cost Analysis methods have been transferred into a large set of principles and parameter values, including estimates of the value of passengers' travel times, which are also readily applicable for use in the planning and implementation of maintenance and renewal activities. Second, existing literature suggests ways to manage temporary congestion during road work; Ridderstedt et al. (2023) conducts a review of this literature. Nyström et al. (2019) suggest that the contractor pay a delay fee (DF) based on road users'

[2] The EU KLEMS is an industry-level growth and productivity research project and stands for EU level analysis of capital (K), labour (L), energy (E), materials (M) and service (S) inputs.

[3] Other than the ossification emanating from the UPC framework, costs in multi-year contracts are typically indexed against a sector-specific input price index. This shelters more risk-averse agents against external changes of input prices during longer construction projects. However, it also eliminates incentives for an agent to substitute less costly inputs for costlier inputs, adding to the poor cost reduction incentives of this mechanism.

loss of travel time due to construction. This would decentralise the trade-off between construction costs and user time to the contractor and increase social welfare. The contractor would not only account for user time during the preparation of the bid for each single project but is also incentivised to develop innovative construction techniques. The authors note that a DF cannot be implemented in combination with a UPC since this contract format precludes the delegation of authority to the builder. This causal inference is further developed in this study. We also describe the negative consequences of the vertical separation of the railway industry for facilitating the optimal design of infrastructure intervention tendering, adding to the previous theoretical speculation by adapting the DF to the idiosyncrasies of the railway sector. Third, a huge body literature–some of which is reviewed in the next section–addresses the design of the procurement process and the contracts which are signed by the parties. An important contribution by this literature is that a menu of contracts could be designed in a way which enhances efficiency. This is also our suggested approach to manage risk in the construction industry . Nevertheless, we have not seen the possibility of implementing a menu addressed in the literature based on the UPC framework.

The remainder of this paper continues with a short introduction to procurement in the construction industry, including a review of relevant research literature. Section 3 presents how the UPC and FP contract can be combined to establish cost-efficient implementation of rail infrastructure renewal. Section 4 applies this solution to switch renewal. Using standard Benefit-Cost Analysis parameter values, a case study is presented in section 5 which illustrates how the inability to account for the impact of line closures on train traffic generates a substantial risk of awarded contractors completing work in an inefficient way. A discussion is presented in section 6 and section 7 concludes.

2. Procurement in the construction industry

Laffont and Tirole (1993) analyse a range of challenges to a principal's ability to implement a preferred policy due to informational constraints. Since the outcome of an assignment is contingent upon information possessed by the agent, this jeopardises the principal's possibility of holding down the

costs for the job. Actions that require effort by the agent may affect costs, quality, and other unobservable considerations or at least those which are not verifiable by an outside party (moral hazard). An agent may also have better information than the principal on exogenous factors such as the technical possibilities or difficulties of implementing activities. This makes it possible for the agent to extract rent from the principal (adverse selection) even when the agent's bargaining position is poor.

Payments for tendered services are said to have different power, meaning that the link between government payment and firms' actual prices, costs and profits differs. A cost plus (C+) contract means that the government pays a firm for its realised costs plus a fixed fee which is independent of actual performance. This is a low-powered contract since it inherently has poor incentive for cost savings; however, it eliminates the possibility of excessive rents from the agent. In contrast to this is the FP contract, which pays the firm a fixed sum to implement a project. This is a high-powered contract since cost savings accrue to the agent who will therefore exert effort to reduce costs, but any savings are also retained by the agent. The principal would prefer to have both low construction costs and low rents for the agent, but information asymmetries mean that these priorities must be traded off against one another.

Both the industry and the engineering management literature highlight challenges for construction contracts beyond information asymmetry. Bajari and Tadelis (2001) suggest that time to completion, project complexity, construction change orders and risk management are fundamental problems in all contracting environments. However, the primary interest of this research is contracting in a private sector construction industry in which C+ and FP contracts dominate. In contrast, tendering in Sweden's construction industry is dominated by the UPC (cf. section 5 in Mandell & Nilsson 2010). We are unaware of any comprehensive international review of standard contracting procedures in the construction business; however, several studies address the use of UPC in other countries. For instance, Bajari et al. (2014) consider the competitive tendering of road renewal projects. The authors emphasise that auctions to procure standard goods such as stationery and off-the-shelf software are

straightforward; however procuring custom-made goods and services to fit a buyer's unique needs poses more extreme challenges.

Information asymmetry also affects incentives under a UPC. In the Quote for Bids for a UPC project, the principal specifies and quantifies (metre, square metre, cubic metre, etc.) the tasks that are tendered. Bidders submit single unit prices for each activity. Using $\boldsymbol{w}$ and $\boldsymbol{z}$ to represent the vectors of unit prices and inputs, respectively, the winning bid is denoted as $B_{UPC} = \boldsymbol{wz}$.

Under a UPC, the agent has incentive for cost savings of the same nature as an FP contract; however, this only concerns the unit prices ($\boldsymbol{w}$) submitted. To minimise costs, the agent must consider alternative ways to deploy staff and equipment. This logic applies before submitting a bid to increase the chances of winning the contest and after that the contract is signed, during implementation of a project. However, low unit prices in a bid create a risk of loss when the resulting revenue is insufficient to recover costs.

Furthermore, the principal has full responsibility for ensuring that ex-ante quantity estimates ($\boldsymbol{z}$) are accurate. Since ex-ante estimated quantities and ex-post actual quantities rarely perfectly match, the resulting final compensation is seldom equal to the winning bid. In addition, changes in scope may arise when some activities that were not specified in the Quote for Bids must be added for successful project implementation. Cost increases after contract signing due to the misspecification of $\boldsymbol{z}$ must be paid by the principal. We argue that this is an approach for reducing an agent's risk premium under a UPC.

An additional consequence of the separation of responsibility for $\boldsymbol{w}$ and $\boldsymbol{z}$ is that the agent cannot implement the assignment in any way other than what has been specified in the Quote for Bids. Even if the agent has innovative ideas to facilitate cost savings in the process, remuneration remains linked to the quantity of each activity, which diminishes the agent's interest in invoicing quantities below $\boldsymbol{z}$. This makes the UPC rigid and prevents innovation regarding the implementation process.

Bajari et al. (2014) address UPC in an empirical setting, with a starting point asserting that procurement contracts are often renegotiated because of required changes after a contract's award.

Examining highway paving contracts, the authors demonstrate that renegotiation due to imperfect ex-ante quantity estimates imposes significant adaptation costs. The findings suggest that bidders respond strategically to contractual incompleteness and that adaptation costs account for 7.5%–14% of the winning bids. By comparison, markups based on private information and market power, which is the focus of much of the bidding auction literature, are much smaller.

The two papers briefly described above related to contracts in the construction industry assume that both principal and agent are risk neutral. Arve and Martimort (2016) contend that while the literature on risk aversion in moral hazard environments is huge, analyses of risk aversion in adverse selection settings, and particularly in regulation and procurement, remains surprisingly scarce. The authors argue that this is likely best explained by the technical complexity encountered when risk aversion and incentives interact. The general finding of this scant literature is that risk aversion shifts optimal contracts towards being low-powered, which reduces the agent's incentives for cost-efficient production and agent's information rents compared to settings under risk neutrality. Implicitly, this also suggests a hypothesis for understanding the prevalence of UPC in the construction industry in which the external aspects of risk are pervasive.

Notably, not all contracts are equally risky and there may be room for alternative contract designs that can produce welfare improvement. Providing the possibility for contractors to choose between contracts with different incentive intensities (low- to high-powered contracts) may unlock potential efficiency gains. Laffont and Tirole (1986) present an early application which is subsequently elaborated on in their textbook (Laffont and Tirole 1993). Self-selection from a menu of contracts is also analysed by Rogerson (2003), who considers a menu including FP and cost-reimbursement contracts, analytically demonstrating that it performs well compared with the first best menu proposed by Laffont and Tirole (1986); cf. also Chu and Sappington (2007).

3. Cost minimisation of switch replacements: An analytical solution

Under the aegis of an IM, the government seeks to maintain the national track system to maximise social welfare. One dimension of these responsibilities is the identification and implementation of activities for periodic replacement of existing assets. To this end, it is essential to establish the lowest possible social costs for replacing an asset that has started to wear out. Since this means that the infrastructure must be closed for traffic while work is performed, it involves a trade-off between (higher) renewal costs vs. (lower) costs for tracks' users.

To elaborate on this trade-off, consider the replacement of q switches on a railway line. A composite commodity ($\boldsymbol{z_1}$) represents the physical activities to be implemented, each with unit price $\boldsymbol{w_1}$. Specifically, $\boldsymbol{z_1^i} = \{z_1^1, \dots, z_1^I\}$ represents the $i = 1, \dots, I$ activities required to complete the required work. $\boldsymbol{w_1^i} = \{w_1^1, \dots, w_1^I\}$ is the unit cost for each activity.

In addition to the work on the tracks, line closure means that a composite commodity ($\boldsymbol{z_2}$) is necessary for completing the job. This represents the impact on users' (train operators, travellers and freight forwarders) time because of the replacement activity. The price of $\boldsymbol{z_2}$ is the users' cost for being delayed ($\boldsymbol{w_2}$).

The switch replacement project is characterised by a production function ($f(\boldsymbol{z_1}, \boldsymbol{z_2})$), and the principal seeks a combination of $\boldsymbol{z_1}$ and $\boldsymbol{z_2}$ that minimises the social costs (W) for substituting $\bar{q}$ units, where the bar over q emphasises that the purpose of the project is to replace a certain number of switches. Referencing Nyström et al. (2019), we consider the following optimisation problem and the consequent Langrangian where λ is the multiplier.

$$min\, W = \boldsymbol{w_1 z_1} + \boldsymbol{w_2 z_2}$$

$$s.t.\, f(\boldsymbol{z_1}, \boldsymbol{z_2}) = \bar{q}$$

$$L = \boldsymbol{w_1}\, \boldsymbol{z_1} + \boldsymbol{w_2}\, \boldsymbol{z_2} - \lambda(f\,(\boldsymbol{z_1}, \boldsymbol{z_2}) - \bar{q}\,)$$

The solution to the value function $W^*(\bar{q})$ defines the minimum value of social cost W. The optimal value of λ^* is the shadow price of the total resources, or the marginal production cost of $\bar{q}$. Given the optimal values of the IM's choice variables, the envelope theorem indicates that

$$\frac{\partial W^*(q)}{\partial q} = \lambda^*$$

Subsequently, the three first-order conditions are now as follows:

$$w_1 + \lambda \frac{\partial f}{\partial \mathbf{z_1}} = 0 \quad (1)$$

$$w_2 + \lambda \frac{\partial f}{\partial \mathbf{z_2}} = 0 \quad (2)$$

$$f(\mathbf{z_1}, \mathbf{z_2}) = \bar{q} \quad (3)$$

These give

$$\lambda = \frac{\mathbf{w_1}}{\delta f / \delta \mathbf{z_1}} = \frac{\mathbf{w_2}}{\delta f / \delta \mathbf{z_2}} \quad (4)$$

$$\mathrm{MRTS}_{\mathbf{z_2},\mathbf{z_1}} = \frac{\delta f / \delta \mathbf{z_2}}{\delta f / \delta \mathbf{z_1}} = \frac{\mathbf{w_2}}{\mathbf{w_1}} \quad (5)$$

Hence, at optimum, the marginal rate of technical substitution shall be proportional to the ratio of the prices. Intuitively, if the impact on users has a high cost, it may be reason to use more resources such as labour hours to complete the job. In section 5, a case study is used considering whether this condition can be met in that specific application.

4. Tendering project implementation

Much of the asymmetric information literature implicitly assumes that a principal interacting with a random agent is unaware of the agent's actual costs to implement an assignment. This does not fit with the industrial procedure in which contracts are competitively tendered. The procurement process is initiated by the principal publicising a Quote for Bids to implement an activity. The following sections describe the standard approach using the UPC (4.1) and the FP alternatives (4.2). Section 4.3 then describes how these two options can be simultaneously included in a Quote for Bids.

4.1 Tendering a Unit Price Contract

The process towards tendering a UPC starts with the principal using an expert to undertake the projecting (i.e. preparing drawings and information related to a site and the activities which are subsequently to be tendered). This results in a Bill of Quantities detailing the activities that the winning

bidder is expected to complete (i.e. $\boldsymbol{z_1}$). Bajari and Tadelis (2001) provide an elaborate account of the significance of spending resources for reducing the risk that a Bill is inaccurate or incomplete.

The principal then publishes the Quote for Bids which includes the Bill of Quantities and the timing of implementation of the project(s), τ^{UPC}. Referencing this information, interested construction firms submit bids by entering unit prices ($\boldsymbol{w_1}$), and the bid with the lowest cost wins the contract under option A, $B_{UPC} = \widehat{\boldsymbol{w}}_{\boldsymbol{1}}\boldsymbol{z_1}$.[4] In railway sector construction, a line is closed to traffic while a job is performed. Therefore, the IM (implicitly) preordains the number of hours that trains are affected during project implementation, which is not a decision parameter for the bidders.

Once work starts, payment is made against actual quantities.[5] As previously noted, this shelters the bidder from any extra costs that may appear because of incorrect quantity estimates in the Bill of Quantities. Moreover, remuneration for activities that was not specified in the Bill of Quantities but proves to be necessary for successful project implementation, must be negotiated between the parties and may add substantial costs to the final Bill. The need to negotiate Change Orders is another way to protect the builder from risk under the UPC format. Except for due diligence in the implementation of all activities, the agent's risk is related to whether the unit prices in the winning bid ($\widehat{\boldsymbol{w}}_{\boldsymbol{1}}$) are sufficient to recover the total costs for completing the job. This separation of quantity risk produces a low-powered UPC.

4.2 Tendering a fixed-price contract

While a UPC requires the IM to sign contracts with two separate entities (a designer and a builder), the FP contract tendering combines these two activities. Therefore, each bidder must collaborate with project experts to produce the design to be implemented. Different bidders may consequently submit varied proposed design implementations for a job.

[4] Since focus is on the winning bid, bids from the different contestants are not indexed.

[5] Different countries apply different rules for implementing this principle. In Sweden, deviations from the ex-ante estimate in the $+/-25$ percent range are paid for with the winning unit price, whereas unit prices for larger deviations may be renegotiated.

The Quote for Bids for FP contracts specifies two issues. The first is a description of what the builder is supposed to accomplish (in the case study in section 5 this is to replace five old switches with new switches of a certain standard). Second, the Quote also includes a detailed specification of the circumstances of replacement at the site, including a description of the qualities of the project, the detailed location of signalling wires and other technical installations and other relevant considerations.

Each bidder prepares a projection based on these requirements which may or may not be based on a Bill of Quantities, but this document is not a formal part of the subsequent FP bid, $B_{FP} = f(\widehat{\boldsymbol{w}}_{\boldsymbol{1}}, \boldsymbol{z}_{\boldsymbol{1}})$. Again, the lowest bidder wins the contract. Notably, the IM directly pays the costs for projection under the UPC, while this cost is part of the bid under the FP framework.

One consequence of using the FP format is that the winning bidder assumes all risks related to the current circumstances. However, since the desired output from the job, rather than the prescribed list of activities to be implemented makes it feasible, this also provides incentives for bidders to adapt implementation to the actual situation at the site (i.e. to find ways to complete the job at lowest possible cost). The possible presence of adverse selection to navigate the construction site circumstances in a novel way compared with the approach published by the IM under the UPC option promotes cost-efficiency in the implementation, while the competition when bids are submitted for the contract provides the means for the principal to benefit from unique solutions.

4.3 An augmented tendering process

Other than the generic trade-off between profits, incentives and risk, the vertical separation of train operations from infrastructure management makes the costs for completing the job fall under different subjects. To manage these features, we suggest that the tendering process be broadened so that bidders may (but must not) place bids for both types of contracts.

Bids for a UPC are still based on a Bill of Quantities prepared by the IM and on a pre-set time for project implementation (τ^{UPC}). The difference compared with the standard procedure is that proposals are evaluated not only on the bid (B_{UPC}) but also on the cost for displacing track users

($\hat{B}_{UPC} = B_{UPC} + \boldsymbol{w_2 z_2}$). Therefore, the IM determines the costs for such disturbances under the design defined in the Bill of Quantities and publishes that number as part of the Quote for Bids.

A construction firm may alternatively place a FP bid. To this end, the Quote for Bids also includes information about $\boldsymbol{w_2}$ and bidders are informed that the winner will be charged for disturbances inflicted on traffic ($\boldsymbol{z_2}$), which corresponds to the DF proposed by Nyström et al. (2019). Therefore, the bid for option B will be $\hat{B}_{FP} = f(\boldsymbol{w_1, z_1} \parallel \boldsymbol{w_2})$, establishing a direct link between the construction firm and the consequences of project activities for track users. Consequently, bidders will consider the possibility of implementing the project at the same time as under option A (τ^{UPC}), or at a different time (τ^{FP}). Choosing the same time could be worthwhile if it is feasible to implement the project in a different way than is specified under the Bill of Quantities ($\boldsymbol{z_1}$) for the UPC. It is also feasible for bidders to consider implementation at some other time of the year (τ^{FP}) if this would limit disturbances. The IM will select the lowest bid $\{\hat{B}_{UPC}, \hat{B}_{FP}\}$.

To submit a bid for the FP contract, all bidders must incur costs for projecting, while this cost has been sunk when a UPC is tendered. However, the extra costs for preparing this bid may be minimal since it can piggy-back on the Bill of Quantities prepared by the IM. This means that the first motivation for bidders to consider submitting a FP bid is a clear assumption that it is technically feasible to implement the project in a way that differs from the plan under the UPC format, including the timing established by the IM. Based on the nature of the FP contract, this will only occur if a bidder determines that it will save on costs relative to the UPC alternative. A smarter implementation means lower cost and a higher chance of winning a contract; however, it also means higher risk for potentially unpleasant surprises during project implementation, resulting in a higher mark-up as compensation for accepting the risk.

The inherent difference between UPC and FP in terms of incentives and risk is crucial to reiterate. The UPC framework provides bidders with some scope for adapting the process for completing the job at the construction site. For instance, different bidders may organise the logistics for transporting old filling from and new macadam to the construction site in different ways, and

respective unit prices for these activities may differ. However, payment for completing a project is related to the observed quantities of each activity. Therefore, bidders lack the possibility or interest in suggesting that could reduce unit quantity and consequently reduce the principal's costs; rather, incentives move in the opposite direction (i.e. arguing that ex-ante quantities have been underestimated).

5. The case study

A tender for any project that requires shorter or longer line closure must accommodate the idiosyncrasies of traffic planning in the railway sector. This framework is described in section 5.1, while the application is described in section 5.2.

5.1 Preparations for a tender in the railway sector

The starting point for the process that will result in the replacement of a switch during year *t* is the decision to allocate resources for this purpose. This choice is made two years before the project (i.e. year $t-2$; cf. Figure 1). With access to information about this timing, train operators that prepare demand for timetable slots during year *t* are made aware of the necessity to adjust the standard timetable on lines where major interventions are required. Operators submit their demand for timetable slots during year *t* to the IM in April (year $t-1$) based on this information. The new timetable is subsequently established in September and traffic for year *t* commences in mid-December $t-1$.[6] Once the timetable is established, the Quote for Bids for the renewal project is published, which occurs approximately at 'Quote' in Figure 1.

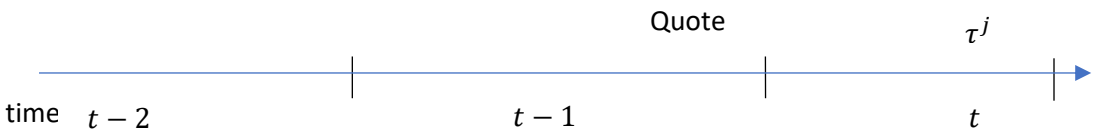

Figure 1*: Timeline for procurement of switch replacement.*

[6] Further details regarding the design of the timetabling process are available from Nilsson et al. (2023).

The Quote for Bids establishes that the project may be performed at date (τ^j) during year *t*. The timing of the line closure under the UPC option (τ^{UPC}) is an external input for bidders when the job is tendered. The FP option differs in that the timing of the project (τ^{FP}) is included in each bid submitted. τ^{FP} may (but must not) coincide with τ^{UPC}.

Figure 2 illustrates one aspect of this possible difference in timing. The far-left box represents off-the-tracks preparations for the upcoming work, assuming that these are the same under both options. The vital difference between the options is that the subsequent four activities depicted by the boxes in Figure 2 can be implemented in direct sequence under the UPC approach since the line is completely closed for traffic. The FP option requires that bidders schedule interruptions between the different work steps, as suggested by the figure. This means that the maintenance crew must terminate activities and possibly undertake interim measures to make it possible for trains to pass at pre-determined times. This provides the basis for predicting that tendering a FP contract requires more work hours and costs than a UPC.

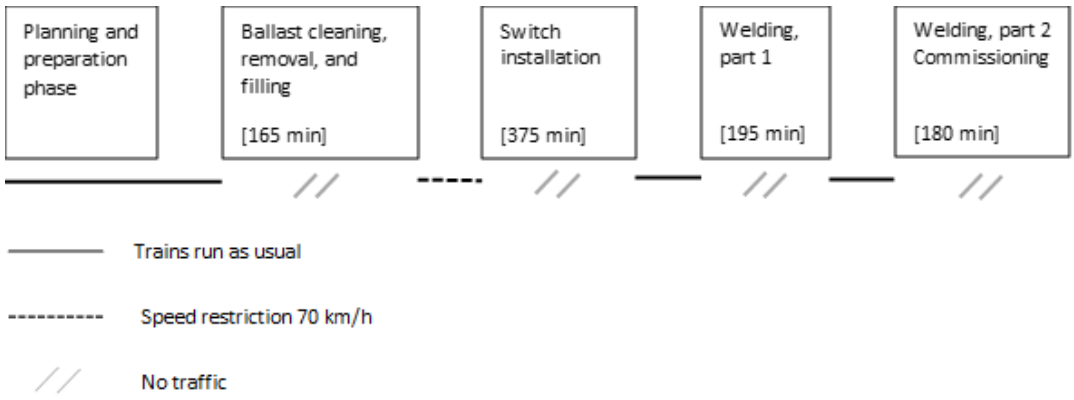

***Figure 2**: Replacing switches between trains.*
Source: Based on proposal made by the switch manufacturer, Vossloh.

5.2 The case of replacing five switches

At a meeting point on a double-track railway line, five switches (referred to in Figure 3 as 132a, 132b, 134a, 134b and 136) must be replaced, corresponding to $\overline{q}$ in the theoretical model.

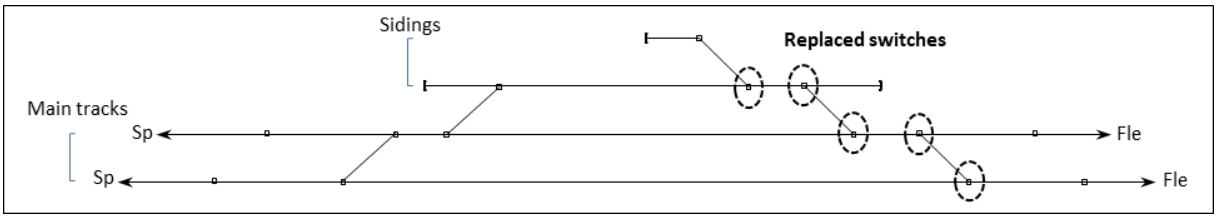

***Figure 3**: Replacing five switches on a double-track line at the Skebokvarn operating point.*

Table 1 indicates that the IM's cost expectation at the start of the process was SEK 22,7 million. Items (a) and (b) represent costs for ex-ante preparation for tendering the project while items (m) and (n) are costs for termination plus a component for uncertainty. Determining these components, the expected entrepreneurial cost is (SEK 22,6-1,7-1,3-1-1,1 =) approximately SEK 17,5 million, which is the numerical representation of $\boldsymbol{z_1 w_1}$ in the theoretical model.

Table 2 illustrates the way in which the cost for one of the activities in Table 1–namely item (e) Track-work is estimated. This illustrates the modus operandi for estimating all cost components in Table 1. Specifically, Table 2 enumerates all activities the builder must implement for replacing the switches as well as the expected quantity of each activity $(i = 1, \dots, I$ for $\boldsymbol{z_1^i} = \{z_1^1, \dots, z_I^1\})$.[7] As an example, consider entry 9 in Table 2, delivering the ballast to the site. It is estimated that this activity requires 1250 m³ filling and that the unit cost ($\boldsymbol{w_1^i}$) is SEK 70000 per m³ adding up to SEK 14.4 million.

While Figure 2 provides an example of one single switch being replaced, Figure 3 illustrates the five switches that were replaced under the case study. The combination of two main tracks and a siding may make it possible to allow trains use one of the main tracks while work is performed on the other track and/or on the siding. If this is possible, it may be feasible to implement work with less consequences for track users than under a full line closure. We lack information regarding how the ideas behind the proposal outlined by Figure 2 could be adapted to the circumstances of a double-track line. The example still illustrates the potential benefits of tendering and compare technically feasible while not necessarily welfare improving ways to do the job.

[7] Information in the table regarding the values of $\boldsymbol{z_1}$ and $\boldsymbol{w_1}$ emanate from the IM's estimate before the tender is published, and the same format is subsequently used when the Quote for Bids for the project is published.

Eq. (5) establishes that the first statistic required for estimating the benefit loss because of the track-works is z_2, representing the total time of traffic disturbances. Work was performed during a weekend with all traffic cancelled. Assuming the same traffic as the preceding weekend, 208 trains were affected. In many cases, this means that passengers must get off the train, board buses and then board another train at the next station. In this specific case, trains were rerouted, adding 45 km to trip length and 40 minutes travel time for each train.

Table 1: Ex-ante estimate of cost for switch replacement using option A, in 1000 SEK.

Item	Activity	Cost
(a)	Project administration	1700
(b)	Projecting	1300
(c)	Environmental measures	140
(d)	Earthwork	902
(e)	Trackwork	14442
(f)	Electricity work	1084
(g)	Signalling work	920
(h)	Telecom work	125
(m)	Submission & termination	1000
(n)	General uncertainty	1051
	Total	22664

Source: Trafikverket (2017)

The operator of the regional trains which were rerouted shared information about its service that must pass the site.[8] The line closure affected five return trains on Saturday and six on Sunday, with an average of 84 passengers per train. Except for the regional service, five different commercial train service providers operate the remaining 186 trains over longer distances. Since no information is available regarding the occupancy rate, benchmark values established by the IM are used, indicating an average 165 passengers per train. The two first rows of Table 3 refer to the benefit loss based on 40 minutes of extra travel time for passengers. In addition, operators incur costs related to time and

[8] The regional operator used buses to replace trains for parts of the line. Information regarding the amount of this non-marginal extra cost is unavailable. Few freight trains are operated during the weekend.

distance. The table suggests that user costs increased by some SEK 1,8 m because of the temporary line closure.

Table 2: Estimated cost for trackwork, in SEK 1000.

Item #	Description	Unit	Quantity	Unit price	Cost
1	Demolition & disassembly of existing track switches	No.	5	75	375
2	Demolition & disassembly of existing tracks	Track m	130	0,6	78
3	New track switches, type 1	No.	4	1700	6800
4	New track switches, type 2	No.	1	800	800
5	Kirow crane	Work shifts	3	310	930
6	Costs for transport of switches and crane	No.	1	100	100
7	Switches and level crossings, composite	No.	5	400	2000
8	New tracks	Track m	130	5,5	715
9	Filling incl. Macadam	m^3	1250	0,7	875
10	Complementary filling of ballast layer macadam	Tonne	10	0,85	9
11	Track/switch direction after new construction and after 3–6 months	Work shifts	3	300	900
12	Grinding	Work shifts	2	250	510
13	Measurement of the track infrastructure. Track position control	Work shifts	1	300	300
14	Basis for relationship document	No.	1	50	50
	Total				14442

The case study illustrates how to estimate a comprehensive cost for project implementation, including the users' additional distance and travel time because of associated line closure. Our best estimate is that the disturbance for track users implies a cost of at least SEK 1,8 m. Any effort to reduce this cost by using more resources for construction such as additional time for staff that perform the job would therefore be worthwhile up to this level.

The unit cost for each activity enumerated in Table 1 comprises the sum of costs for basic inputs macadam, equipment (truck, excavator, etc.) and manpower ($\boldsymbol{w_1^i} = f^i(materials, equipment, labor)\ \forall i$). The unit price estimates include costs for all these inputs, and

the proportion of each component cannot be deduced from the information that is available. For the sake of illustration and based on back-of-the-envelope calculations, we estimate labour costs for all items (c)–(h) in Table 2 to be approximately SEK 4.5 million of the expected entrepreneurial cost of SEK 17,5 million. This indicates that if it is feasible to implement work in a way that reduces train disruptions, it might be feasible to reduce the aggregate social costs for the project.

Table 3: Estimated extra costs for travellers and train operators because of detours.

	No. of trains	Passengers per train	Unit value; SEK/minute or km*	Cost, SEK
Passengers: Extra time, 40 min				
Regional trips	22	84	62	76384
National trips	186	165	85	1739100
Train operators: Extra time 40 minutes	208		73.5	10192
Train operators: Extra distance 45 km	208		23.52	3261
Total				1828937

Source: Parameter values from Trafikverket (2023).

6. Discussion

Many aspects of construction industry activities are uncertain, which makes it difficult to prepare appropriate ex-ante cost estimates (i.e. $\boldsymbol{w_1 z_1}$). A common feature of investment in and maintenance of infrastructure is that it requires digging. More precise ex-ante information regarding what may be hidden under the ground's surface that will appear once digging starts lowers the risk for unpleasant surprises during project implementation (Ribeiro et al., 2020). In other words, the more comprehensive the ex-ante projecting and estimate of $\boldsymbol{z_1}$, the less risk there will be that final costs will deviate from the bid submitted by the winning contractor. Moreover, railway sector projects' risks differ widely because of different traffic frequencies (the gap between trains in the timetable), whether a line is single- or double-tracked, the exact physical circumstances at the site and other relevant factors.

To appreciate the nature of these uncertainties, our case study regarding the replacement of five switches can be contrasted with a related class of activities; namely, the day-to-day maintenance of the railway network. Odolinski (2019) analyses the design of 34 separate maintenance contracts tendered by Sweden's IM for the country's 14200 km state-owned railways, with each contract comprising varying track lengths. The assignment of a maintenance contract means that the winning bidder is to ensure that all installations of tracks, signalling, optical wiring and electricity are operational around the clock and throughout the year so that trains can operate according to the established timetable. Despite extensive documentation regarding the stock of equipment, its location, age and other characteristics, information about actual quality along the line remains incomplete. Since this is a generic aspect of a line's standard, it should be thought of as a common value uncertainty, wherein neither the principal nor any of the firms that prepare a bid for the contract knows more than the others regarding the actual standard.

The distinction between a spot measure such as the replacement of switches and a comprehensive maintenance contract illustrates that the risk is smaller for some types of tenders than in others. That is, given the uncertainties of railway maintenance and the contracts used for this activity, the uncertainties associated with switch renewal could be manageable.

Example: Consider a Quote for Bids for a project which includes $\boldsymbol{z_1^i} = \{z_1^1, \dots, z_1^I\}$ the $i = 1, \dots, I$ quantified activities required for completing work as specified by the principal. A bidder estimates the bid $\hat{B}_{UPC}$ would suffice to implement the project under the UPC option and $\hat{B}_{FP}$ under a solution of the firm's own design. The latter is obviously riskier than the UPC problem but would improve the chances of winning the contract. Therefore, the bidder would choose this option if $\hat{B}_{UPC} - \hat{B}_{FP} > k$, where k is the premium necessary for accepting the increased risk. The design of the UPC option with w^I activities means that the principal has established the activities to be implemented and the technical standard such as the type of switch to be installed, technical equipment and other specified parameters. This means that the bidder may use other types of activities or a higher or lower

quantity of each but cannot deviate from the technical specification. It is the how dimension of the task which could be designed differently if the bidder accepts that the k threshold is met.

It may be difficult for the IM to determine whether there is more than one way to undertake an assignment. Our case study of executing switch renewal in a non-standard way is an obvious example. Another possibility is that a bidder for an investment or renewal contract may have completed previous jobs at the same location where the new job is to be implemented, and construction firms with activities in different countries may have encountered alternative methods for managing a certain type of activity that could also be considered in Sweden. The self-selection design in our proposal compels bidders with more or better (private) information regarding the job that is announced to express the cost implications of this information when submitting a bid for the contract.

Other than explaining the dominance of UPC in the industry as a means for allocating risk, its use may also be rationalised as a way to handle non-contractible quality. In an empirical application for labour contracts, Prendergast (2002) suggests that the existing literature fails to account for the important effect of uncertainty on incentives through the allocation of responsibility to employees. When workers operate in certain settings, firms are content to assign tasks to workers and monitor their inputs. In contrast, when a situation is more uncertain, they delegate responsibility to workers, but to constrain their discretion, base related compensation on observed output. Baker (2002) presents a similar argument but emphasises the possibility of the multi-task nature of employment. For instance, a public sector organisation with a welfare maximisation objective may include dimensions related to quality that are not possible to measure, which may lead to an employment contract with weak incentives reducing employees' tendency to neglect quality under a regime with strong cost incentives.

For the present application, the IM's overall objective is to minimise social costs over the lifetime of long-lived assets. It may be difficult to allocate control over quality to a firm that does not have control over the asset's life cycle costs. Rather than delegating the method for replacing a switch and risking the solution becoming costly after a few years, the implementation is guided by a UPC

where the principal establishes most aspects of the work design. Therefore, the prevalence of the UPC could be understood as a way to constrain the agent when the switch is being replaced to reduce the risk of future cost increases.

This may provide a complementary motive for low-powered contracts of the same nature such as the risk hypothesis we suggest. Under the quality hypothesis, quality is safeguarded through the detailed specification in the Bill of Quantities for a UPC. Instructing the contractor to use the same equipment and processes dictated under a UPC will still leave the door open to search for non-standard methods for implementation (i.e. the *how* question).

7. Conclusions

Other than external restrictions that establish a level playing field in construction tendering (under the procurement Directive 2014/24/EU in Europe), the process to build, renew and maintain infrastructure is rigid and strongly controlled by the tendering principal in other ways. The challenges of incomplete information and about the quality of the substructure where work is to be completed provide one possible reason for this stringency. We consider the rigidity of the standard UPC to be an instrument for shielding contractors from excessive risk, and to avoid bids becoming excessively high because of a risk premium that would be necessary under a FP design. Under this institution, it is the principal who tenders a contract that carries this risk. Moreover, this contract format ensures that the principal has full control over how the work is performed, which increases the possibility to control life cycle cost properties.

However, projects in the railway industry may not all carry equal risk. One distinction is that spot measures come with site-specific risks, whereas larger investments and much of the contracting for day-to-day maintenance may involve risks which are disproportionately costly to manage by spending more on ex-ante projecting to reduce uncertainty. Moreover, it may be challenging to determine which projects are less risky than others beforehand. In addition, it is unclear whether

alternative methods for implementing a project are available since by nature innovations cannot be ordered beforehand.

Our case study of an alternative method for replacing old switches with new switches demonstrates that a more general format for tendering may enhance efficiency. The starting point for our inquiry is that this implementation of the work has not been tested and the rigid UPC format for tendering blocks the possibility, and indeed the incentives, for testing any alternative method. Moreover, the principal may have less control over production methods and the potential to adapt implementation to the precise circumstances at the worksite than a construction firm.

The study proposes a means for tendering railway projects in a way which makes the choice of contract format–allocating risks and providing incentives for alternative ways to implement a job–an endogenous result of the tendering process. Incentives for examining alternative designs are built into the tendering structure by making it feasible for construction companies to bid for the standard type of UPC and a riskier FP contract.

The case study also illustrates that the benefits of solutions which reduce train disturbances during implementation of track-works may be substantial for track users, motivating the search for alternative methods. The interaction between costs for the construction per se and costs for users affected by a line closure is one example of a quality dimension that must be controlled to enhance efficiency. Our application benefits from the development of Benefit-Cost Analysis that makes it possible to quantify the consequences (time losses) and apply a set of parameter values to estimate losses and make them an aspect of the decision regarding the appropriate construction method to use. If it is feasible to quantify and value other dimensions of quality in tendered contracts, our suggested approach would have even larger applicability.

A possible extension of the choice between two types of contracts would be to let bidders combine a fixed-price and unit price components according to their own preferences, the bid ($\mathcal{B}$) being $\mathcal{B} = f + \boldsymbol{w_1 z_1}$ where f is the fixed-price component. This is the way in which most tenders are designed by the IM, where the relative size of f differs across types of projects. Before doing so, a

more profound analysis of the incentives when the firms could choose between more or less risky dimensions of projects would be necessary.

Our blueprint provides a starting point for a process towards testing the practical consequences of making it feasible to let bidders chose between types of contracts. While the blueprint per se should be straightforward, and indeed adapted to the timetabling realities of railway sector applications, it must be confronted by the nitty-gritty reality of tendering in the sector.

Acknowledgements

This work was financed by Sweden's Innovation Agency (Vinnova). We are grateful for the valuable comments by Ivan Ridderstedt on a previous version. We also wish to thank Arne Nissen and Anders Boëthius, both at Trafikverket, for providing data and for their helpful comments and answers to our questions. We are also grateful for the traffic information provided by Mälardalstrafik. Any remaining errors are the responsibility of the authors.

References

Arve, M., Martimort, D., 2016. Dynamic Procurement under Uncertainty: Optimal Design and Implications for Incomplete Contracts. *American Economic Review*, 106(11), 3238–3274. DOI: http://dx.doi.org/10.1257/aer.20150275

Bajari, P., Houghton, S., Tadelis, S., 2014. Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs. *American Economic Review*, 104(4), 1288–1319. DOI: http://dx.doi.org/10.1257/aer.104.4.1288

Baker, G., 2002. Distortion and Risk in Optimal Incentive Contracts. *The Journal of Human Resources*, 37(4), 728–751. DOI: https://doi.org/10.2307/3069615

Chu, L.Y., Sappington, D.E.M., 2007. Simple Cost-Sharing Contracts. American Economic Review, 97(1), 419–428. DOI: https://doi.org/10.1257/aer.97.1.419

Laffont, J-J., Tirole, J., 1986. Using Cost Observation to Regulate Firms. *Journal of Political Economy*, 94(3), 614–641. DOI: https://doi.org/10.1086/261392

Laffont, J-J., Tirole, J., 1993. A theory of incentives in procurement and regulation. The MIT Press, Cambridge, Massachusetts, London, England.

Mandell, S., Nilsson, J-E., 2010. A Comparison of Unit Price and Fixed Price Contracts for Infrastructure Construction Projects. Working Paper, Swedish National Road & Transport Research Institute (VTI), No. 2010:13.

Nash, C., Smith, A., Crozet, Y., Nilsson, J-E., 2019. How to liberalise rail passenger services? Lessons from European experience. *Transport Policy*, 79, 11–20. DOI: https://doi.org/10.1016/j.tranpol.2019.03.011

Nyström, J., Nilsson, J-E., Börjesson, M., 2019. Delay fees in procurement of constructions projects. *Journal of Transport Economics and Policy*, 53(4), 348–363.

Odolinski, K., 2015. Reforming a publicly owned monopoly: costs and incentives in railway maintenance. Doctoral dissertation, Örebro Studies in Economics 30.

Odolinski, K., 2019. Contract design and performance of railway maintenance: Effects of incentive intensity and performance incentive schemes. *Economics of Transportation*, 18, 50–59. DOI: https://doi.org/10.1016/j.ecotra.2019.05.001

Prendergast, C., 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy*, 110(5), 947–1174. DOI: https://doi.org/10.1086/341874

Ribeiro, J.A., Pereira, P.J., Brandão, E.M., 2020. A real options approach to optimal bidding in construction projects considering volume uncertainty. *Managerial and Decision Economics*, 41(4), 631–640. DOI: https://doi.org/10.1002/mde.3126

Rogerson, W.P., 2003. Simple Menus of Contracts in Cost-Based Procurement and Regulation. *American Economic Review*, 93(3), 919–926. DOI: https://doi.org/10.1257/000282803322157160

Trafikverket, 2017. Underlagskalkyl Skebokvarn - spårväxelbyten 2017-05-30 (Trafikverket's internal document, in Swedish).

Trafikverket, 2023. Samhällsekonomiska kalkylvärden. Bilaga till Trafikverkets publikation "Beräkningsmetodik för transportsektorns samhällsekonomiska analyser". ASEK 7.1. (In Swedish)