PERGAMON

Transport Policy 11 (2004) 43–53

TRANSPORT POLICY

www.elsevier.com/locate/tranpol

Cost overruns in road construction—what are their sizes and determinants?

James Odeck*

Department of Civil and Transport Engineering, The Norwegian University of Science and Technology, N-7491 Trondheim, Norway

Received 1 June 2002; revised 1 September 2002; accepted 1 January 2003

Abstract

This paper investigates the statistical relationship between actual and estimated costs of road construction using data from Norwegian road construction over the years 1992–1995. Based on this data a regression model is developed. The findings reveal a discrepancy between estimated and actual costs, with a mean cost overrun of 7.9% ranging from −59% to +183%. In absolute terms, cost overruns amounted to a formidable 519 million Norwegian kroners. One particular new finding that has not been shown before in previous studies is that cost overruns appear to be more predominant among smaller projects as compared to larger ones. This observation, for the Norwegian road sector in particular, leads us to assert that the greatest potential for cost savings lies in exerting pressure on smaller projects to control their costs. Other factors found to influence the size of cost overruns include completion time of the projects and the regions where projects are situated. Surprisingly, neither project type nor work force type seems to influence the level of cost overrun. Finally, the paper proposes some policy implications.

Keywords: Road construction; Cost estimates; Actual costs; Cost overrun

1. Introduction

Studies on the relationship between estimated costs and actual costs of road projects has left many policy makers stunned by suggesting that cost overrun is prevalent in the sector and that the magnitudes may be large (see for instance Skamris and Flyvbjerg, 1997). A general view, also held particularly by the media, is that cost overrun is predominant in the sector as opposed to under runs. In the wake of shrinking government revenues, common sense tells us that spending needs to be curbed. Overruns will imply wastage of resources which otherwise would have been used for productive purposes elsewhere.

A highly problematic situation occurs when the cost of a project is underestimated and presented to the decision-makers. Given that the decision maker, in evaluating the viability of projects may consider the net present value of projects (i.e. the difference between the net benefits and net cost of carrying out the project), underestimated costs may be deceptive. In this case, the net present value may be so large that had the actual viability of a project been known, the decision makers may have resolved to choose one of the following three alternatives: (a) not to implement the project at all, (b) to implement the project in another form, or (c) to implement other project or projects. The end result may be that nonviable projects are being implemented due to inaccurate estimates. This would lead to an inefficient allocation of resources. Thus, cost estimates presented to decision makers should be scrutinized. Ideally, ex post studies should be conducted seeking to explain the divergence between estimated costs and actual costs. The results should contribute to the improvement of cost estimation.

The subject matter of this paper is to contribute to the debate on the magnitude of cost overrun and eventually reveal their causes in the Norwegian road sector. To do this, we investigate the statistical relationship between estimated and actual costs using data from Norwegian counties for the years 1992–1995. We build an econometric model, which we estimate to reveal factors that may help explain the observed overruns. This yields several benefits. We can control a host of factors such as regions, time factors (estimated construction time and delay in completion time), the type of project (tunnels and bridges), and estimated

* Current address: The Norwegian Public Roads Administration, P.O. Box 8142, 0033 Oslo, Norway. Fax: +47-22-07-36-79.

E-mail address: james.odeck@vegvesen.no (J. Odeck).

0967-070X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0967-070X(03)00017-9

constructions costs. Thus, the following questions are addressed:

(i) The magnitude and direction if bias occurs in cost estimations.
(ii) The relationship between cost overrun and other factors such as completion time, size of estimated cost and regions where projects are situated.

Some caution is, however, necessary with respect to the results that may be derived in this study. First, the projects that we examine are those that were carried out by the Norwegian Public Roads Administration (NPRA), and not through a bidding process, and involve about 40% of the total budget for road projects. This may seem unrepresentative. However, 40% is still large enough to pinpoint the magnitude of overruns. Besides, it represents 100% of those projects carried out by the NPRA in the period considered, and therefore at least will reveal the magnitude of overruns among the NPRAs projects. How these perform in relation to other projects not included is hard to tell; we have no reason to suspect any particular bias. In addition, one particular reason for selecting only the NPRA projects is that they are the ones with accessible data. Projects carried out through bidding processes by the private sector are not readily accessible in details due to company confidentialities.

The remainder of this paper is organized as follows. Section 2 presents some previous research on the relationship between estimated and actual costs of road infrastructure projects. Section 3 presents the data and Section 4 the methodology. In Section 5, the results are presented while Section 6 offers some conclusions and their possible policy implications.

2. Previous studies

Efforts to assess the divergence between estimated costs and actual costs of roads are rare. Only a few studies exist that rigorously compare forecast costs and actual costs for large groups of road infrastructure projects.

One of the first attempts to gather and analyze information about uncertainties in the measurement of characteristics of transportation facilities, which planners supply to the decision-makers, was by Knudsen (1976). The study considered uncertainties in airport cost analysis and their effect on site selection. The study found out that the distribution of the ratio between actual and estimated cost was lognormal and not normal as is usually assumed. Second, the empirical results showed a considerable bias towards underestimation.

Odeck and Skjeseth (1995) assessed Norwegian toll roads to reveal whether planning procedure shortcomings experienced by Norwegian road agencies hsd resulted in poorer than projected financial performances for some of the toll roads. They found overestimation of traffic forecasts and underestimation of construction costs. In their small sample of 12 toll projects, they found cost overrun on average at about 5% but the interval was large at −10 to 170%. They claim that the uncertainties inherent in cost estimates are never brought to the attention of the decision makers even as sensitivity analysis.

A study by Skamris and Flyvbjerg (1996, 1997) compared the accuracy of traffic forecasts and cost estimates on large transportation projects in Denmark. The study considered cost estimates of seven bridges and tunnels dating from just before the decision was made to build and actual costs after completion. For the non-completed projects, developments in forecast were compared to the original forecasts. The main conclusion from this study is that cost overrun of 50–100% is common for larger transportation infrastructures and that overruns above 100% are not unusual. Further, they claim that the differences between forecast and actual costs (and traffic) they find cannot be explained primarily by the inherent difficulty in predicting the future.

Pickrell (1990) carried out a study for the US Department of Transportation covering US rail transit projects with a total value of £15.5 billion. The total capital cost overrun for eight of the projects was calculated to be 61% ranging from −10 to +106%. Another study by the Auditor general of Sweden (1994) covering 15 road and rail projects revealed that the average cost overrun of eight road projects was 86%. The range for road projects was from −2 to +182%, while the average overrun for the seven rail projects was 17%, ranging from −14 to +74%. Two-thirds of the projects were still under construction and it was concluded that final costs could turn out to be even higher.

Another study carried out by Fouracre et al. (1990) for the UK Transport and Road Research Laboratory (TRRL) covered 21 metro systems in developing countries. The results showed that six metros had cost overruns above 50%, two of these in the range from 100 to 500%. Three had overruns in the range 20–50%, and the remaining four ranged from −10 to +20%.

As concluded by Skamris et al. (1996), in most of the previous studies, changes in design, delays and technological innovation can explain some of the overruns. However, there remains a substantial portion of the divergence that cannot be explained by technological causes alone. In fact, one author (Wachs, 1990) implicates manipulated forecasts as a probable cause.

A most recent study on cost overrun of road projects is that by Flybjerg et al. (2002). Based on a sample of 258 transportation infrastructure projects worth US$90 billion and representing different projects worldwide and across historical periods, they conduct a statistical study of cost escalation. They find with overwhelming statistical significance that cost estimates used to decide whether such projects should be built are highly and systematically misleading. They conclude that the underestimations observed cannot be explained by error and is best explained by a strategic interpretation, which is tantamount to lying. They thus warn legislators, administrators and those who value honest numbers not to trust cost estimates and benefit-cost analysis produced by project promoters.

Another recent and interesting study is by Hecht and Niemeier (2002) which compares transportation project development efficiencies. In particularly they examines time and cost project development efficiencies between voter or legislatively approved projects, and projects with standard scope. Their data is based on interviews conducted with projects managers who were recently involved in project development of large transportation projects. They do not find significant evidence that state highway projects, with highly defined voter or legislatively approved project scopes, time or costs, are any more likely to have lower development costs or times than projects that are non-voter approved.

3. Data

The data for this study was obtained from the NPRAs database on completed trunk road projects. The database consists of annual reports from the 19 district headquarters of the NPRA, which again are grouped into five regional headquarters. All these projects were carried out and managed by NPRA. They do not include tendered projects. However, the NPRA can hire labor from private firms. The database contains information such as estimated cost, actual cost, completion time, project type etc., for 620 projects completed in the years 1992 through 1995.

The 620 projects included very small, small, medium and large projects. The very small category comprised projects costing less than 15 million NOK, category small comprised projects costing greater than 15 and less than 100 million NOK, category medium comprised projects costing between 100 and 350 million NOK, and category large comprised projects costing more than 350 million NOK. There were 420, 156, 33 and 11 projects in categories very small, small, medium and large, respectively. All the data were deflated using the construction cost index for 1995 obtained from the Central Bureau of Statistics. The complete set of variables available in the data set is shown in Table 1.

Table 1
List of variables

No	Variable	Measure	Explanation
1	Country	Categorical	County in which project is situated
2	Year	Categorical	Year of project completion
3	Road_nr	Categorical	Road identification number
4	Road class	Categorical	Class of road
5	Road Length	Scale	Road length in meters
6	Road width	Scale	Road width
7	Contractor	Categorical	Type of contractor (all NPRA projects)
8	Work force	Categorical	Type of work force on project
9	Planned start	Date	Planned start of construction work
10	Planned completion	Date	Planned opening of project
11	Actual completion	Date	Actual completion time
12	Delay in completion	Scale	Delay in completion time
13	Length of bridge	Scale	Total length of bridge on project
14	Length of tunnel	Scale	Total length of tunnel on project
15	Type of project	Categorical	Either road, tunnel or bridge of combinations
16	Estimated cost	Scale	Esitmated cost of project
17	Actual cost	Scale	Actual costs of project
18	Cost overrun (%)	Scale	Ratio of actual to estimated cost in %
19	Cost overrun (abs)	Scale	Difference between actual and estimated cost
20	Projects size	Categorical	Size of project as measured by class of estimated cost

A few words are in order concerning the estimation procedures used in the Norwegian road sector. To start with, there are up to several cost estimates associated with each road project. The first of such estimates is normally made at the time of feasibility studies. In Norway, the standard of accuracy at this stage is ±40%. The second estimate takes place at the corridor plan level, and these estimates are refined and the standard of accuracy is ±25%. The final refinements are made at a detailed planning level. This is the stage where design, specifications and final cost estimates are made. These are the cost estimates that are presented in the national budget, to be sanctioned by parliament normally a year before construction starts. The requirements set by the Ministry of Transport are that the uncertainties of cost estimates should be in the range ±5% for projects at detailed planning level (national budget level). In this paper, the cost estimates used are those from the final stage, i.e. the detailed plan level when decisions to build are made. Arguments may be raised against using the estimates at the final stage as proposed here, see for instance Flyvbjerg et al. (2002). As these authors have proposed, if we cannot use these estimates as the point of departure for assessing overruns, it would almost be impossible to make meaningful comparison of costs because no common standard of comparison would be available.

An interesting question is what happens if the cost overrun exceeds 5%. In principle, the regional NPRA responsible for construction is normally asked to give an explanation for the overrun after which the government may cover the losses. Thus, from the outset there is no clear-cut incentives for the regions to keep their overruns within the 5% level. From an economic point of view, cost overrun may here be viewed simply as a 'tolerated' inefficiency.

The implication of demanding a ±5% accuracy for road projects is that the distribution of uncertainties in estimated cost is assumed to be normally distributed, to meet the total budget constraint. The estimation technique used for project cost is the so-called *judgmental forecasting*. Here the forecast has the purpose of predicting the future factors that will influence the costs of the project. The goal of the forecasting is to get a reliable estimate of cost and/or duration of a project. The procedure consists of putting together a team of professionals with experience in cost estimates and construction work. The teams' main task is to identify factors that may contribute to accurate cost estimates. At their availability is the databank with past experience on road projects administered by the NPRA. The uncertainties in these factors must be such that the total uncertainties are in the range set by the Ministry of Transport at ±5%. A statistical model that ensures that the total cost estimate is within this range and that the uncertainty is normally distributed is used. The team draws both on past experience, knowledge, historical data and any other source of information available to them. The major problem however, is that this process is not free of subjective evaluations by the team. Many factors may be beyond past experience alone. In studying the divergence between actual and estimated costs we are in effect studying the extent to which estimates may be biased and thus call for corrective measures for subjective evaluation.

4. Methodology

Several approaches may be used in studying discrepancies between estimated costs and actual costs. The approach used in this study is that of using historical data for both estimated and actual cost to derive the explanation and magnitude of the ratio λ of divergence defined as:

$$\lambda_i = \left(\frac{K}{E}\right)_i, \qquad i = 1, \ldots, 620 \tag{1}$$

where K is the actual cost and E is the estimated cost.

By studying the variation of λ for the different road projects, a measure for the magnitude of cost overrun is obtained. Thus, to include the error term, Eq. (1) may be re-written as:

$$\lambda_i = \left(\frac{K}{E}\right)_i + \mu_i \tag{2}$$

In the sequel we postulate that the cost overrun may be explained by a host of factors, and the model to be estimated can be formulated as:

$$\begin{aligned}\lambda_i = {} & \alpha + \beta_1(\text{Estimated cost})_i + \beta_2(\text{Estimated cost})_i^2 \\ & + \beta_3(\text{Delay in completion})_i + \beta_4(\text{Delay in completion})_i^2 \\ & + \beta_5(\text{completion time})_i + \beta_6(\text{completion time})_i^2 \\ & + \delta_{1ij}(\text{year})_{ij} + \delta_{2ij}(\text{region})_{ij} + \delta_{3ij}(\text{work force type})_{ij} \\ & + \delta_{4ij}(\text{project size})_{ij} + \mu_i \end{aligned} \tag{3}$$

where λ_i is the cost overrun in percent for project i and β_i's and δ_i's are parameters to be estimated. Note that δ_i's are coefficients for the dummy variables, thus year of project completion, region, work force employed and project size are treated as dummy variables where for example, δ_{1ij} is equal to 1 if project i was completed in year j and equal 0 otherwise. Particularly, for the dummy variable year, which is a dummy for the year when the project was completed, we have four periods (1992–95), so that to avoid multicollinearity we use $(s-1)$ dummy variables to denote the s different periods. This rule applies for all dummy variables. For work force we have three categories and hence use two dummies, for regions we have five and use four dummies and for project size we have four and use three dummies. The coefficient of these dummies reflects the average difference in the independent variables between the category variables and the omitted variable. So for the dummy year, if we omit 1992, the coefficient for 1993 is a measure of the effect of 1993 on cost overrun compared to 1992. Thus the coefficients for the dummies will measure effects relative to these omitted categories.

Note that a few variables in the data set presented in Table 1 are not used in the model. These are road class, length and width of the road project. Since we are dealing with trunk roads, we found very minor differences with respect to class in the data. Concerning length and width, these are a direct function of estimated cost already included.

Eq. 3 is a second-degree polynomial (also called quadratic) equation. It is frequently used in analyses where slopes of a relationship are expected to depend on the level of variables itself. In our case, we expect the impact of certain variables in cost overrun to depend on the level of that variable. Such variables would include estimated cost, delay in completion and total length of construction period (completion time). Thus using this functional form allows us to test the hypothesis that the impact of these variables on cost overrun is dependent on their levels.

Some precautionary measures are, however, still necessary when studying the distribution of λ as expressed in Eq. (2) above. The possibility exists that when observing λ, the variation observed is not random but systematic. In that case, one of the assumptions necessary for proper estimation of the regression model in Eq. (3) will be violated. It concerns the assumption that the residuals have a constant variance, i.e. $E(\mu_i) = \sigma^2$.

Problems of this kind are termed heteroscedasticity in econometrics. Violation of this assumption will not render the estimation of Eq. (2) efficient if one uses the ordinary least squares method.

There are several reasons to suspect that the assumption of constant variance might not hold when studying cost overrun of road projects through model (2). These may be (1) delay in completion may incur additional costs resulting into cost overruns, due to for example increasing cost, (2) completion time may reflect uncertainty and hence overruns may be larger for projects with longer completion time and, (3) which is the most obvious one; in model 1 the estimated cost enters both the right left and right hand side of the equation and hence the residuals are expected to vary with this variable.

One way to uncover the problem of *heteroscedasticity* is to plot the absolute value of the residuals (μ) against the predicted values of cost overrun. This plot is given in Fig. 1.

Several problems appear in the plot depicted in Fig. 1. First, looking at the residuals versus the predicted value of cost overrun, the scatter fans out to the right until the predicted value of cost overrun gets closer to and above zero, and then fans in (note that the negative predicted values mean under run). What this indicates is heteroscedasticity where the residuals vary with the size of predicted overrun. Second, the residual distribution has a heavy median and a positive skew, meaning that cost overrun is slightly more frequent among projects as compared to under runs. Residual versus estimated cost overrun, not shown here due to space, showed a similar trend where the scatter fanning in was more clear-cut. The same trends were also observed with respect to delay and construction time.

Several approaches have been proposed to establish heteroscedasticity and to correct for it (see for instance Koutsoyiannis, 1977). We used the Glesjer test where a regression of the dependent variable on the independent one is performed and the residuals computed. Next, the absolute values of the residuals are regressed on the independent variables with which the variances of the residuals are suspected to be associated. Since the form of the regression is not known before hand, several formulations containing various powers of the independent variable are tested. The regression with best fit is chosen. This gives information on the form of heteroscedasticity. The Glesjer test performed for this paper gave the following results, (1) the variance of the residuals was significantly associated to only one variable; estimated cost and, (2) heteroscedasticity was of the form $E(\mu_i) = \sigma^2 = k^2(E\cos t)^{-1}$ implying that the variance of the residuals decreases with increasing estimated cost. Thus the appropriate transformation of our

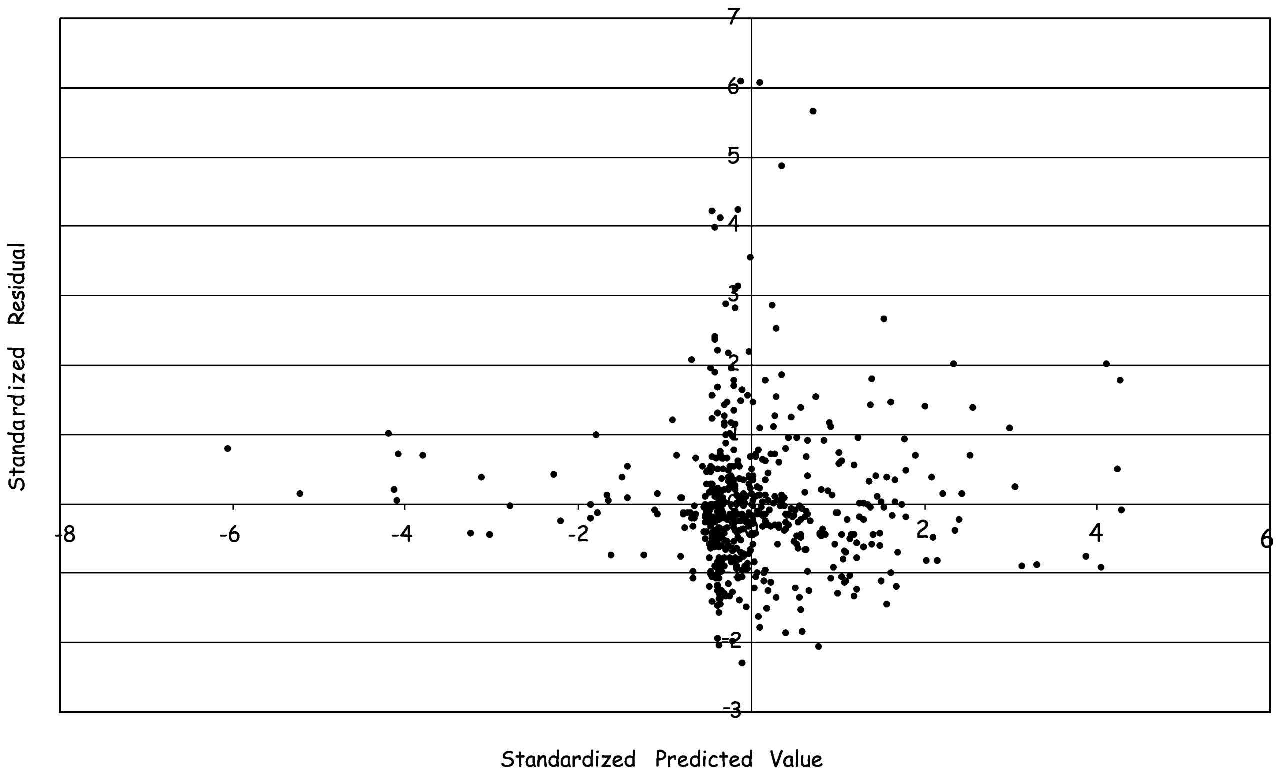

Fig. 1. Plots of residuals versus predicted value of the dependent variable.

original model consisted of dividing it by $\sqrt{E\cos t^{-1}}$ and the transformed model to be estimated is:

$$\frac{\lambda_i}{\sqrt{E\cos t_i^{-1}}} = \frac{\alpha}{\sqrt{E\cos t_i^{-1}}} + \beta_1 \frac{\text{Estimated cos } t_i}{\sqrt{E\cos t_i^{-1}}} + \beta_2 \frac{(\text{Estimated cos } t)_i^2}{\sqrt{E\cos t_i^{-1}}} + \beta_3 \frac{\text{Delay in completion time}_i}{\sqrt{E\cos t_i^{-1}}} + \beta_4 \frac{(\text{Delay in completion time})_i^2}{\sqrt{E\,\text{cos}t_i^{-1}}} + \beta_5 \frac{\text{completion time}_i}{\sqrt{E\cos t_i^{-1}}} + \beta_6 \frac{(\text{completion time})_i^2}{\sqrt{E\cos t_i^{-1}}} + \delta_{1ij}(\text{year})_{ij} + \delta_{2ij}(\text{region})_{ij} + \delta_{3ij}(\text{work force type})_{ij} + \delta_{4ij}(\text{project size})_{ij} + \frac{\mu_i}{\sqrt{E\cos t_i^{-1}}} \quad (4)$$

The transformed random term μ_i is now homoscedastic with constant variance equal to k^2. Our succeeding analysis is based on this transformation. In estimating this model, the original data must be divided all through by $\sqrt{E\cos t^{-1}}$. Note that the dummies need not be transformed.

5. Empirical results

The main objective of this study is to gain insight in the divergence between estimated costs and actual costs for different road projects in Norway by applying the regression model outlined in Section 3 above. In the following sections, we present the results of our findings.

5.1. Summary statistics

A summary of cost overrun is presented in Tables 2–4. Consider first Table 2. The mean cost overrun is at 7.88% while the range is large from −58.5 to 182.7%. The standard deviation, which gives an indication of how closely or widely the individual cost overruns are spread around their mean value, is 29.2% indicating a rather large variation. Thus, from a cost estimation perspective the picture that emerges is that discrepancies between cost estimates and actual costs are skewed to the right, i.e. cost overrun is more predominant among projects as compared to cost savings. Consider now Table 3. The figures there illustrate the distribution of projects by cost overrun divided into three categories; (1) under run (negative overrun), (2) no overrun, and (3) overrun. For those projects with under runs, the mean is at −15% while for those with overrun the mean is at 25%. Again this is an indication of skewness in cost overrun. In terms of percentages in each group, it is observed that the majority of projects (52.4%) have in fact cost overrun. There are surprisingly a significant number of projects with under-runs (35.5%), implying that there are fairly a good number of projects being completed with costs less than estimated. Only 12.1% of the total projects were completed according to estimated costs i.e. neither over- or under run.

Next consider Table 4, which shows the distribution of overrun in both percentages and absolute terms in million Norwegian kroners (NOK), and categorized according to the size of projects divided in to four different categories. These categories are identical to those normally used by the NPRA in classification of road projects according to estimated costs. The upper part of the table shows the distribution by percentages. Here it is seen that the majority of projects in the portfolio can be classified as small projects with total estimated cost less than 100 million NOK. This represents about 93% of the whole portfolio. Medium and large projects thus make only about 7% of the total by number. In terms of mean overruns, again the small projects are the highest at 7.55 and 10.62% for projects with estimated cost at below 15 and those at greater than 15 to less than 100 million NOK, respectively. The mean overrun for medium sized projects are low at 2.46% while that of large sized projects are negative (i.e. under runs). In terms of percentage of net total overrun, these small projects together make about 99%. Thus medium sized projects represent

Table 2
A summary statistics of cost overrun

	Statistics
No. of projects	620
Mean	7.88
Std. error of mean	1.17
Median	1.96
Std. deviation	29.20
Minimum	−58.50
Maximum	182.67

Table 3
The distribution of projects by percentage cost overrun

	Mean % overrun	Number of projects	Percent of projects
Underrun	−15	220	35.48
None	0	75	12.10
Overrun	25	325	52.42
Total		620	100.00

Table 4
Distribution of overrun in percent and in million NOK by project size

	Percentage cost overrun				
	Very small (≤15 ml NOK)	Small (>15 ≤ 100 ml NOK)	Medium (>100 ≤ 350)	Large (>350 ml)	Total
Number of projects	420	156	33	11	620
% of projects	67.74	25.16	5.32	1.77	100
Mean overrun (%)	7.55	10.62	2.46	−2.50	7.88
Median	0.00	6.20	0.41	−4.02	1.05
% of total overrun	71.95	26.91	1.61	−0.48	100
	Absolute overrun in ml NOK				
Number of projects	420	156	33	11	620
% of projects	67.74	25.16	5.32	1.77	100
Mena overrun (%)	0.35	3.31	1.74	−18.32	0.84
Median	0.01	2.38	0.50	−17.40	0.20
% of total overrun	28.16	99.59	11.09	−38.84	100
Sum overrun (in ml NOK)	146.09	516.73	57.53	−201.51	518.84

about 1.6% while large projects represent a total under run of about −0.48% and therefore offsets total overruns.

Thus, judging from these observations, percentage cost overrun seem to be greater among smaller projects as compared to larger ones. What normally is of interest, especially in the media, is the budgetary implication of overruns in absolute terms. The lower part of Table 4 shows overruns in absolute terms i.e. in million NOK. Now we observe that the mean overrun is highest for small projects costing above 15 but less than 100 million NOK. With a mean of 3.3 million NOK, this group of project makes 99.6% of net overrun amounting to about 516.7 million NOK. Both in the upper and lower part of the table results of the hypothesis that the different types of projects produce the same distribution are reported. We see that the hypothesis is rejected at any reasonable significance level both by the Kruskal–Wallis and Median test. Another test, not reported in the table, was the hypothesis that cost overrun is as common as cost under run. The p-value for the Mann–Whitney test was at 0.0001 implying that under-estimation is more common.

In general, the absolute values of overruns shows that the costs incurred to Norwegian society due to overrun by road projects amounted to 518 million NOK kroners which otherwise, through careful estimations could have been avoided. These results indicate also that overruns are prevalent among smaller projects while the larger ones, although smaller in number, experience under runs on average. One remark however is order. The size of the so-called large projects is small as compared to the smaller ones. An argument can be raised to the effect that the large projects are outliers and should have been excluded in the analysis. Our objection to this argument is that we are considering a whole portfolio of projects whether large or small, and where the main aim is to explore the distribution of overruns. Statistically, there is no problem in comparing a small number against a large number of projects.

One other important observation is that the observed distribution of cost overrun is different from that which the policy makers demand at ±5%. This implies that policy makers are not given the correct information on the magnitude of uncertainties at the time decisions are made.

5.2. The regression model results

We now turn to the results of the regression model expressed in Eq. (4) and thus explore factors that may help explain the observed cost overruns. The results of the estimations are presented in Table 5. In regressing the model, we used the stepwise method as an independent variable selection procedure (see for instance, SPSS Base System Users Guide, Release 6.0, 1993). In this way, the unnecessary variables can be excluded from the model. From the original model the following variables were excluded (meaning that they were not relevant explanatory variables in the model); Dummies for all regions except

Table 5
Estimation results from the regression model

Variables	Coefficient	Std. error	*t*	Sig.
(Constant)	7.1371	1.5049	4.7429	0.0000
Estimated cost	−0.0096	0.0032	−2.9844	0.0030
Completion time	0.0110	0.0042	2.6048	0.0094
Estimated cost squared	0.0000	0.0000	−2.7262	0.0066
Completion time squared	−0.0009	0.0004	−1.9535	0.0512
Year 1993 (1 if vintage year is 1993,0 otherwise)	4.7784	2.5294	1.8891	0.0593
Region middle (1 if middle, 0 otherwise)	−7.3908	3.0323	−2.4374	0.0151
Adjusted R-squared	0.21	88.45		

the region of middle Norway, all dummies for the completion year of projects, all dummies for work force type, dummies for project size and the second order term for delay in completion time. The reason for exclusion is that these variables are least significant in explaining overrun and some of them are correlated with variables already included in the model. For example, the dummies for projects size were all excluded because project size is already included in the variable estimated cost.

Table 5 suggests several interesting findings. Notice first that the adjusted R-squared is about 0.21 implying the estimated equation explains about 20% of the variations observed in cost overrun. The interpretation here is that the observed variation in cost overrun is also explained by other factors beyond those that are captured in the equation estimated. It is not the intention of this paper to delve into which other factors that explain overrun. Instead we leave it to readers to speculate. In this respect, the recent work by Flyvbjerg et al. (2002) is a good starting point for exploring such factors which these authors classify into four groups: (1) technical explanations, (2) economic explanations, (3) psychological explanations and (4) political explanations. Unlike these authors, we have no reason to speculate on which one(s) applies for our data set.

The size of estimated cost has a significant negative impact on percentage cost overrun indicating that percentage overrun tend to be lower the higher estimated costs are. This verifies our earlier observation that overruns tend to be higher for smaller projects. Looking at the second-order term, which has a zero and significant coefficient, this implies that the slope of the relationship between percentage overrun and estimated cost does not depend the level of estimated cost, i.e. the sign of the slope does not change in any range of estimated cost. A plot of the relationship between the estimated cost and percentage overrun is plotted in Fig. 2(a). In that figure, cost overrun is seen to fall with estimated cost. Note that there are many very small projects and these may, due to the scale used, appear as having estimated cost very close to zero. Fig. 2(b) shows the same relationship but in absolute terms (i.e. overrun defined as the difference between actual and estimated cost in mill NOK).

Since overrun is here measured in absolute terms, we would expect a marked positive relationship. However, what we observer is in fact a slight negative relationship as shown by the dotted line. As expected, there are some outliers. In general, the figure conforms to our earlier observations in Table 4 that larger projects on average in the period studied had a under runs while the smaller projects had overruns. Thus from this observation one is led to conclude that, in the Norwegian road sector percentage cost overrun is more predominant among smaller projects. Even when overrun is considered in absolute terms, smaller projects still do not outperform larger ones as far as under run is concerned. One possible explanation as to why overrun is less predominant among larger projects is the amount of attention larger projects are given. Thus, larger projects are most probably under much better management as compared to smaller ones.

Another question that can be raised is to what extent the time lapse from when a project is started until the project is completed influences cost overrun. There are several reasons why a correlation between completion time and cost overrun may exist. As an example, uncertainties in prices and wages normally increase with length of time until a project is completed. However, the Pearson correlation coefficient between completion time and estimated cost gave a relatively low, but significant $r = 0.356$. In addition our stepwise procedure left these variables in the equation meaning that these two variables do not necessarily explain the same thing. The first order coefficient for completion time is positive and significant implying that overrun increases with completion time. The second order coefficient is negative and significant implying that the relationship between completion time and overrun is quadratically formed. What this observation implies is that cost overrun increases within a certain range of completion time after which it declines. This relationship is depicted in Fig. 3 where cost overrun increases until a completion time of about 200 weeks (about 2.7 years) is reached and then falls.

A conclusion that may be drawn is that cost overrun tends to be higher the shorter the completion time is expected to take. A possible explanation is that accuracy in cost prediction is difficult the shorter the construction is expected to take, meaning that uncertainties decreases with time in the Norwegian road construction. Another explanation could be that project management get better oversight over causes of overrun and are able to control for them the longer the project lasts.

The included dummy variables also offer some interesting results. Considering first the regional dummy, cost overrun may be expected to vary according regions due to different reasons. One possible reason is that different managers some of which may be more efficient in management as compared to others run different regional road administrations. Another difference may be that some regions may be more susceptible to uncertainties as compared to others due to terrain, e.g. mountains and fjords, which may make cost estimations more difficult. Among the dummies for regions, only the region of middle Norway was found to be different from the base region of eastern Norway. Thus there are indications that there are some regional differences with respect to cost overrun.

Considering the dummies for vintage year of project, only 1993 were found to differ from the base year of 1992. We note in passing that the type of work force employed and project type, i.e. whether bridge, tunnel or ordinary road does not seem to influence on overruns as these were dropped in the stepwise procedure.

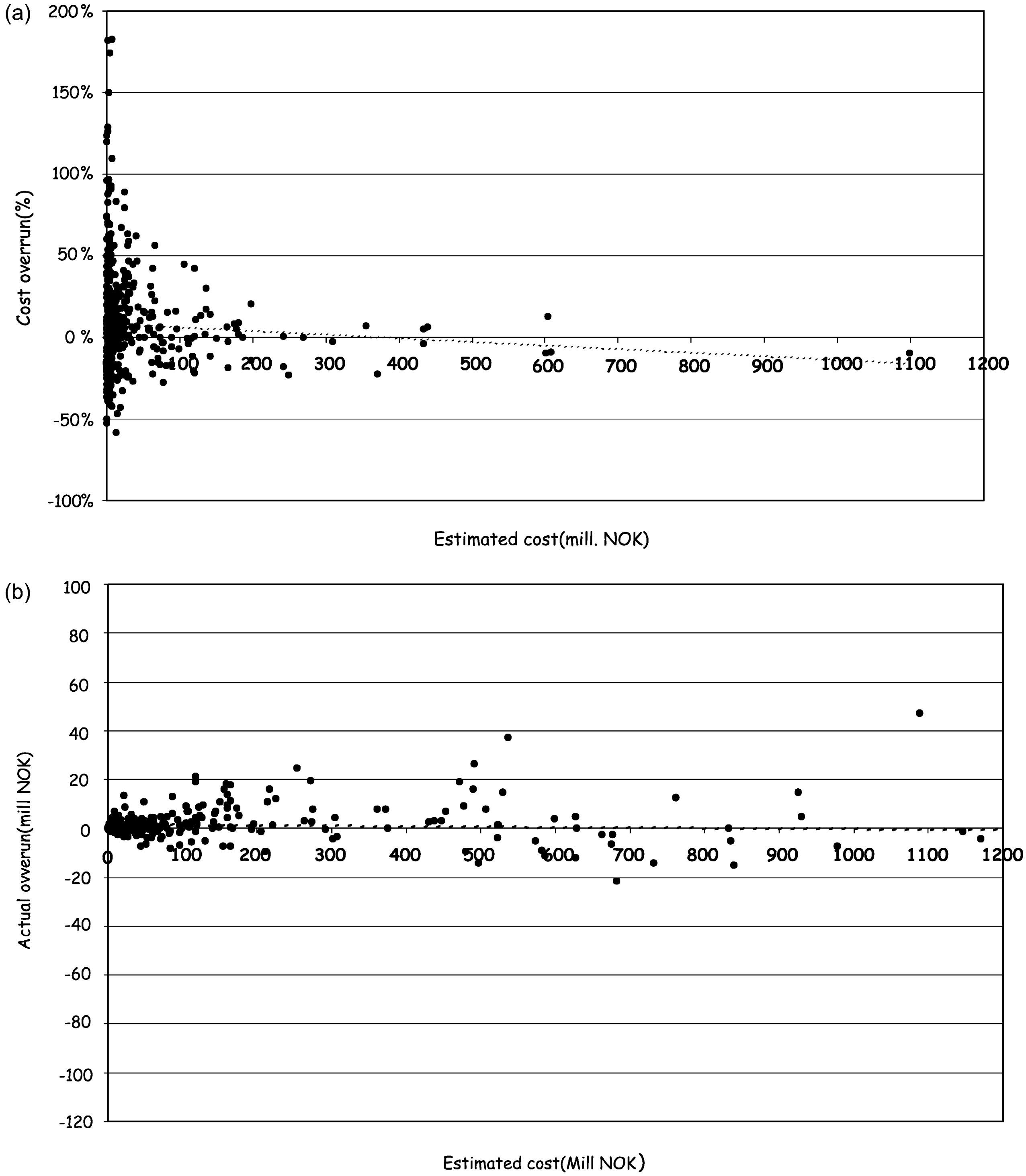

Fig. 2. (a) Plot of cost overrun (%) versus estimated cost. (b) Plot of cost overrun in mill NOK versus estimated cost.

6. Concluding remarks and policy implications

The objective of this paper has been to investigate the statistical relationship between actual and estimated cost. A regression model has been estimated using data from Norwegian road construction over the years 1992–1995. This has made it possible to identify both the magnitudes of cost overrun and the impacts that different variables have on the observed overruns themselves.

The data has revealed that there is a discrepancy between estimated and actual costs of projects constructed by NPRA. With overwhelming statistical

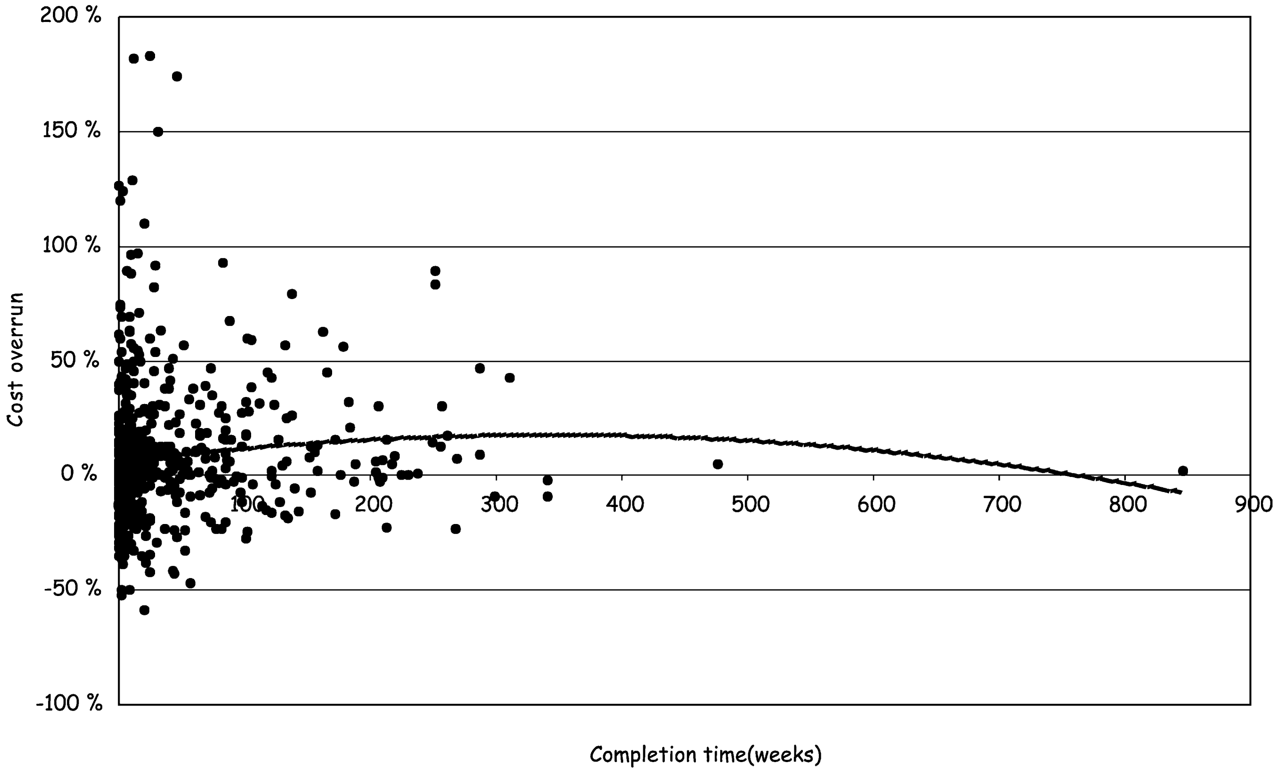

Fig. 3. The relationship between cost overrun and completion time.

significance, cost overrun is found to be more predominant as compared to cost savings. The mean cost overrun for projects completed in the period 1992 through 1995 is at 7.88% ranging from −59 to 183%. Although not very large in percent, the overruns in real terms for the portfolio of projects considered here sums up to 519 million NOK. This amount is large enough to raise concern. On the positive side however, there are significant number of projects being completed with actual costs less than estimated. These results confirm previous findings by other authors elsewhere, e.g. Flyvbjerg et al. (2002) and Skamris and Flyvbjerg (1997) that cost overrun are predominant in the road construction. One particular finding that differs or that has not been unveiled before is that overruns appear to be more predominant among smaller projects as compared to larger ones. This observation, for the Norwegian road sector in particular, leads us to assert that the greatest potential for cost savings lies in exerting pressure on smaller projects to control their costs.

Second, since estimated cost overrun decreases with estimated cost of projects, this may indicate that larger projects generally are under better management as compared to smaller ones. Third, overrun increases with completion time up to medium sized projects and then decreases. This may indicate that longer completion time offers opportunity for adjustments that may help control costs. Finally, there are indications that there are regional differences with respect to the magnitude of cost overrun.

The findings in this paper not only outline the discrepancy between estimated and actual costs of road projects and some factors explaining them. In our view, there are definite policy implications of this study. These are as follows:

1. The potential for resource saving lies in controlling for overruns among small projects in the Norwegian road sector. It is the relatively small overruns those sums up to huge sums. Thus, cost estimates for smaller projects should be set under the same scrutiny as those of larger projects.
2. A study of a set of projects throughout the construction period to investigate in greater details how factors, including those revealed in this study, influences construction costs. The aim should be to gain deeper understanding as to why cost overrun occurs predominantly and to improve cost estimations.
3. It should be demanded that all cost estimates be accompanied with some thorough sensitivity analysis. The idea is to enlighten the policy makers on the possible consequences for risks involved.
4. Since costs overruns are of a great magnitude in many projects, there is a need for a control mechanism that ensures sober estimates. An establishment of a standing advisory committee on cost estimates within the sector is one way to go about it.

Although we stress the points above, the current regime in the Norwegian road sector has been much concerned about the magnitude of cost overrun in the sector. In the 1990 s a lot of effort was exerted in proper management of large projects with a focus on keeping actual costs within the estimated costs. This effort has probably resulted to what we observe; lower overruns as compared to smaller projects. Our major recommendation therefore is that this effort be exerted also on smaller projects.

Other factors not observable in the data set that can help explain some of the variation in cost overrun has not been discussed due to scope of the current paper. This, however, is the focus of an on going research in NPRA where projects are being studied in all stages from the planning to completion combined with in-depth interview of those involved.

Acknowledgements

Thanks are due to two anonymous referees and the editor of this journal for improving the quality of this paper. Thanks are also due to Tore Knudsen at SINTEF (Norway), Rune Elvik at the Norwegian Institute of Transport Economics (Norway), Svein Bråthen at Molde Research (Norway) for useful comments on an earlier version of this paper. We also thank the NPRA for supplying the data and allowing the completion of this paper.

References

Auditor General Sweden, 1994. Riksrevisionsverket. Infrastrukturinvesteringar: En kostnadsjämörelse mellan plan og utyfall i 15 större prosjekt innöm Vägverket og Banverket. RRV 1994: 23.

Fouracre, P.R., Allport, R.J., Thomson, J.M., 1990. The performance and impact of rail mass transit in developing countries. TRRL Research Report 278. Transport Research Laboratory, Crowthorn, Berkshire, England.

Flyvbjerg, B., Holm, M.S., Buhl, S., 2002. Underestimating costs in public works: errors or lie? Journal of the American Planning Association 68 (3), 200.

Knudsen, T., 1976. Uncertainties in airport cost analysis and their effect on site selection. PhD dissertation, University of California, Berkeley.

Koutsoyiannis, A., 1977. Theory of Econometrics. An Introductory Exposition of Econometric Methods, Macmillan.

Hecht, H., Nemeier, D.A., 2002. Comparing Transportation project development efficiencies: California department of transportation and the California county sales tax agenicies. Transport Policy 9, 347–356.

Odeck, J., Skjeseth, T., 1995. Assessing Norwegian toll roads. Transportation Quarterly 49 (2), 89–98.

Pickrell, D.H., 1990. Urban rail transit projects: forecasts versus actual ridership and cost. US Department of Transportation, Washington DC.

Skamris, M.K., Flyvbjerg, B., 1996. Accuracy of traffic forecasts and cost estimates on large transportation projects. Transportation Research Record, 1518.

Skamris, M.K., Flybjerg, B., 1997. Inaccuracy of traffic forecasts and cost estimates on large transport projects. Transport policy 4 (3), 11–146.

SPSS Base System Users Guide (1993). Release 6.0.

Wachs, M., 1990. Ethics and advocacy in forecasting for public policy. Business and Professional Ethics Journal 9 (1–2).