Panel Data Analysis of Operating Costs in the Norwegian Car Ferry Industry

Terje Andreas Mathisen[a]* and Finn Jørgensen[a]

[a] Bodø Graduate School of Business, University of Nordland, NO-8049 Bodø, Norway

* Corresponding author: Email: tam@uin.no, Tel.: +47 75 51 76 37, Fax.: +47 75 51 72 35

Abstract

This paper demonstrates that panel data estimation methods can be applied to derive operating costs for the highly regulated car ferry industry in Norway. The balanced data set includes 360 observations of ferry crossings from 1995 to 2005. Compared with cross-sectional results from earlier studies, the random effects (GLS) model applied in this paper provides more efficient and unbiased cost estimates. With particular relevance for pricing according to welfare economic principles, our estimates indicate that previous studies have underestimated the marginal costs of transporting vehicles of different sizes. Moreover, the difference between fares and marginal costs increases with distance. Hence, vehicles transported on longer trips are, in contrast to the welfare principles forming the basis for the current fare scheme, charged relatively more and, thus, subsidize the fares of vehicles transported on shorter trips.

Keywords:

Costs, ferry industry, long-run marginal costs, fares, panel data analysis

Introduction

The Norwegian Public Roads Administration (NPRA) each year subsidizes car ferry transport services in the coastal areas of Norway to the extent of approximately NOK[1] 1300 million[2] each year. Even though all ferry crossings, with the exception of one, are not financially profitable, they do provide transport services considered indispensable for the welfare of society with an estimated social surplus amounting to about NOK 4.3 billion (Jørgensen *et al.*, 2011). The transport authorities are, therefore, willing to compensate the ferry industry for the deficits sustained through providing these services. Subsidies are distributed amongst the operators by the transport authorities using a policy of strict regulation of supplied services and fares (e.g. Jørgensen *et al.*, 2004). The objective of the regulation policy is to provide a coherent transport network with a fare system that charges equally throughout the country for a given distance and also generating sufficient revenues for transport companies to operate profitably at the given level of subsidies (NOU 30A, 1977).

Earlier studies carried out on the Norwegian ferry industry have used cross-sectional data to derive the relationship between the operating costs and the produced transport services (e.g. Jørgensen *et al.*, 2004; Mathisen, 2008). In these studies the econometric models explaining total costs of operation are designed so the coefficients of the explanatory variables can be interpreted directly as marginal costs for transporting vehicles onboard the ferries. The advancing panel data techniques enable the analysis of a combination of variation in costs and production measures between crossings and within a crossing over time, thus providing more efficient and unbiased estimation results compared to both cross-sectional and time-series studies (e.g. Baltagi, 2005). Panel data methods can also be applied in the ferry industry to derive more precise cost estimates. All companies operating ferry crossings subsidized by the Norwegian transport authorities are obliged to report cost- and traffic figures on a yearly basis. This is objective and computer-processed information which generates a data set enabling comparability both between crossings and between years. The data set enables the use of a simple multiproduct cost function based on well-recognized output measures of the ferry industry and it is well suited for basic panel data analysis.

The aim of this paper is to apply panel data analysis to derive better estimates of the operating costs in the ferry industry so that the transport authorities can obtain a better basis for predicting the subsidy requirement for every specific crossing. Also, the new panel data enable better estimates of long-run marginal costs which, according to the social-welfare maximization principles (e.g. Hensher and Brewer, 2001), should be used as the basis for the national fare scheme. The study, in contrast to most other analyses of costs in the transport industry, is not an efficiency or productivity study, but rather focuses on providing new and better estimates of the actual costs of producing the current transport services.

The paper first discusses cost functions with special emphasis on the car ferry industry followed by a presentation of the panel data specification of the cost model. Then, the data set is presented. The next section analyzes the empirical data, assesses the model fit and presents the estimation results. Finally, the concluding section highlights relevant policy implications.

Cost Functions in the Ferry Industry

It is important to define an appropriate functional form for the specific industry. Well recognized functional forms for analysing cost structures in transport industries are for example the linear, quadratic, Cobb-Douglas and translog specifications. Different cost functions, from the single product linear relationships to the more advanced multiproduct functions, are presented in a historical perspective by Braeutigam (1999), and in a more applied perspective by Pels and Rietveld (2008). Baumol et al. (1988) discuss several different functional forms for multiproduct industries in relation to certain desired mathematical properties. Common function specifications such as the Cobb-Douglas and Translog produce cost functions with weaknesses of non-negativity, and zero-value if an output measure is zero.

The cost function applied in this study is a linear multiproduct specification using untransformed variables, and it was chosen after having assessed different specifications and transformations of the variables. A special characteristic of the Norwegian car ferry industry is the stable production of transport services over time due to rigid- and long-term governmental regulation. The regulation of fares, schedules and quality using negotiated contracts and tenders of long duration, could be an important reason for the limited variation in traffic and cost figures over time. The limited variation can in turn be the reason why all model specifications which further reduce variation, such as logarithmic or "deviation from mean" transformations, produce too little divergence of individual variation to achieve proper panel data estimation results.

Generally, a well-suited basis for framing a study of costs is the stochastic frontier analysis (SFA) (e.g. Kumbhakar and Lovell, 2000). Differences between companies usually exist within an industry with respect to both technical and cost efficiency. In contrast to ordinary least square (OLS) analysis, the SFA calculates a disturbance term with two components, of which one is restricted to non-negativity while the other component has the traditional symmetrical distribution. SFA is commonly used in efficiency studies because the non-negative component represents an inefficiency term. SFA has been applied in the Norwegian ferry industry and identified considerable efficiency improvements (Odeck, 2008).

Despite the seemingly general appropriateness of SFA, particularly with regard to the output distance function, on cost analysis of the current data set, this study applies a basic panel data analysis method using untransformed variables. Panel data analysis is applied because the aim of this study is to analyze the cost structure of the industry in general, with special attention paid to welfare-optimal design of fare schemes and comparison towards cross-sectional estimates, rather than to measure and identify efficiency differences amongst companies or crossings in the ferry industry. Additionally, it is not recommended to apply SFA on linear model specifications (Syrjänen *et al.*, 2006). Analysis of the panel data model using ln-transformed variables does not provide good results and the statistical requirements for the random effects (GLS) model are not satisfied using such specifications.

Car ferries carry both passengers and vehicles, the latter differing considerably both in size and weight. Ferries can thus be categorized as a multiproduct industry. In Norway, the different vehicle groups are based on the passenger car equivalent (PCE) converting system - a fundamental concept when dealing with output in the ferry industry (see e.g. Mathisen, 2008). This concept handles the multiproduct problem and indicates how much burden a vehicle brings upon the ferry. By considering length, breadth and weight, each vehicle category is converted to a corresponding number of regular private cars ranging from 1.025 to 10.682 (The Norwegian Public Roads Administration, 2006). Even though the PCE concept is purely based on rough approximations, it enables the uniform measurement of the outputs of the ferry industry. Such combined output measures are common also in other transport industries (Winston, 1985). The PCE concept is often related to the transported distance in order to create a more complete concept of the transport work performed at a crossing. This output measure is called PCE-kilometres and is the product of the number of PCE's transported multiplied by the distance of the crossing. Both PCE and PCE-km are well recognized output measures in the ferry industry both politically (NOU 30A, 1977) and in earlier studies (Jørgensen *et al.*, 2004).

Earlier studies on operating costs in the Norwegian ferry industry have been linked to the optimal setting of fares. Even though different model specifications have been applied, they have generally been based on the PCE and PCE-km output measures, with an additional dummy variable indicating whether weather conditions are rough in the sailing area. Such a linear multiproduct function was used by Solvoll and Jørgensen (2001) to design a new fare scheme. Further, Jørgensen et al. (2004) derived, using a linear approximation of a modified translog function, the optimal fares under budget restrictions by applying the Ramsey pricing principles. The results from these cost estimations were used in trial arrangements made for the tendering of ferry crossings (Bråthen *et al.*, 2004) and electronic fares (Bråthen and Lillebakk, 2005). An expanded cost function including capacity variables was presented by Mathisen (2008) with the aim of calculating marginal costs for short, medium, and long time horizons. However, also this study was based on a cross-sectional data set.

Based on data available from 1995 to 2005 for this panel data study, it would appear appropriate to apply the same explanatory variables as those previously analyzed in cross-sectional studies, e.g. Jørgensen et al. (2004). The untransformed variables PCE and PCE-km and a dummy variable indicating rough weather provide a model with a good fit and easy interpretation. The econometric properties of such an output-based model have also proved satisfactory in earlier studies of the ferry industry. Because the variables consider both the amount of traffic and the transport distance, misspecifications should not be a problem. The annual operating costs (dependent variable) are adjusted to the 2006 price level using the general consumer price index from Statistics Norway (2006) to consider changes in the general price level.

The production of ferry services is largely unchanged over the time period studied. According to Statistics Norway (www.ssb.no), the traffic measured in passenger kilometres has increased by 19% during the period 1995-2005. During the time period 1999-2005, the trend shows that important input factors like salaries and fuel have increased roughly by 33% and 25%, respectively, which is considerably higher than the corresponding 12.5% increase in the consumer price index. One of the major technological advances was the change in fuel from diesel to gas. Gas propulsion pollutes less than diesel, but increases the acquisition costs of ferries considerably (The Norwegian Ministry of Petroleum and Energy, 2002). However, this technological advance has little effect on the analysis because investment in new ferries has been limited. According to the ferry arrangement plans from the NPRA, the average age of the fleet was 18 years in 1995 and 22 years in 2006. Gas-fuelled ferries were first introduced on a single trial-crossing in 2000 with further implementations to other regular services from 2006.

Panel Data Model Specification

There is a potential benefit of getting more efficient and unbiased estimates using panel data because observations of the same individuals over time provide information that is undetectable using cross-sectional and time series methods. The cost function which is the basis for the panel data study is presented in equation (1).

(1) $$C_{it} = \alpha + \beta_1 X_{1it} + \beta_2 (X_{1it} X_{2i}) + \delta Z_{1i} + u_i + \varepsilon_{it} \quad i = 1,\ldots,N; t = 1,\ldots,T$$

The panel data cost function in equation (1) estimates the total operating costs (in 2006 prices) for crossing i at time t, C_{it}, where the subscript i denotes the cross-section dimension and t denotes the time-series dimension. The intercept term α is defined as independent of both time and crossing. X_{1it} is an observable time-varying factor, while Z_{1i} and X_{2i} are observable time-invariant factors. X_{1it} is the number of passenger car equivalents (PCE) transported over crossing i each year and X_{2i} is the length of the crossing for which the vehicles are transported. Multiplication of X_{1it} and X_{2i} derives an output measure of the total transported PCE kilometres, ($X_{1it}X_{2i}$). The dummy variable Z_{1i} indicates whether crossing i is located in sheltered waters, $Z_{1i} = 0$, or in open sea, $Z_{1i} = 1$. Considerably higher safety requirements for services on open sea make cost increases reasonable if $Z_{1i} = 1$. Furthermore, the disturbance term is divided between the unobservable company-specific effect, u_i, and the remainder of the disturbance, ε_{it}. u_i is time-invariant and accounts for the crossings' unobserved characteristics, whereas the remainder of the disturbance, ε_{it}, varies with companies and time and can be thought of as the usual disturbance term in the regression (see e.g. Wooldridge, 2006). The parameters α, β_1, β_2 and δ are assumed to be positive, implying that all factors are positively correlated with costs.

Attention will be directed at the within-group- (fixed effects or least-square dummy variable), between-group-, and random effects panel data models and how they make better use of the data compared to pooled regression. Whether the panel data methods analyzing purely within individual variation (fixed effects), or the method analyzing both within- and between variation (random effects) can be applied, depends mainly on the covariance assumptions of the unobserved effects, u_i, and the compound residual, $v_{it} = u_i + \varepsilon_{it}$.

The within-group regression in equation (2) implies that deviations from average values of explanatory variable j, X_j, explain deviations from the average in C. Subtraction of the mean evens out the variables with no variation over time. This relates, in the specification of equation (1), not only to the unobservable individual effects, u_i, but also to the intercept term, α, and the dummy variable, Z_i that are all time invariant. A requirement for achieving feasible results using within-group analysis is that there are variations within individuals over time.

(2)
$$C_{it} - \overline{C}_t = \beta_1(X_{1it} - \overline{X}_{1t}) + \beta_2((X_{1it}X_{2i}) - (\overline{X_{1t}X_2})) + \varepsilon_{it} - \overline{\varepsilon}_t$$
$$\Rightarrow \widetilde{C}_{it} = \beta_1 \cdot \widetilde{X}_{1it} + \beta_2 \cdot (\widetilde{X}_{1it}\widetilde{X}_{2i}) + \widetilde{\varepsilon}_{it}$$

The company specific abilities can be estimated by averaging the within-individual variations, so that only the between-individual variation is analyzed. The between effects estimator presented in equation (3) runs the regression of averages across time and is mainly used to produce the between variance for the random effects estimator.

(3)
$$\overline{C}_i = \alpha + \beta_1\overline{X}_{1i} + \beta_2(\overline{X}_{1i}\overline{X}_{2i}) + \delta Z_{1i} + u_i + \overline{\varepsilon}_i$$

When considering both within-, and between-group variation, the best unbiased estimation of equation (1) can be derived using random effects estimation based on the Generalized Least Squares (GLS) model. The random effects (GLS) model assumes the properties of data being sampled randomly to construct a more efficient estimator, i.e. individual effects are uncorrelated with the regressors over all time periods, Cov(X_{1it}, u_i) = 0. In the random effects model of equation (4), the weight given to between-group variation is defined by theta, $\theta = 1 - \sqrt{\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_u^2}}$, where $\sigma_\varepsilon^2 = \text{var}(\varepsilon_{it})$, $\sigma_u^2 = \text{var}(u_i)$ and T is time periods. Theta is defined as a value between 0 and 1. If theta is close to 0, the unobserved effect is relatively unimportant, and the within and between variations should be given equal weight. This makes random effects estimates equal to pooled OLS estimates. A theta value close to 1 indicates that the random effects estimates are equal to the within-group estimates (Wooldridge, 2006).

(4)
$$C_{it} - \theta\overline{C}_t = \alpha(1-\theta) + \beta_1(X_{1it} - \theta\overline{X}_{1t}) + \beta_2((X_{1it}X_{2i}) - \theta(\overline{X}_{1t}\overline{X}_2)) + \delta(1-\theta)Z_i + (v_{it} - \theta\overline{v}_t)$$

The relatively strict assumptions imply that the model should be thoroughly tested to assess whether the random effects (GLS) can be applied or if pooled regression or fixed effects are more appropriate (e.g. Baltagi, 2005).

Data Examination

All ferry companies in Norway are obliged to report detailed accounting and traffic information to the Norwegian Directorate of Roads with regard to trips subsidized by the state. Consequently, high quality comparable data is collected to assess consistently the activities related to the different ferry crossings. The traffic information is publicly available in an annual statistical yearbook of ferry transport published by the NPRA (2006). Company sensitive data relating to crossing-specific accounting figures and ferry information was also gathered from the NPRA with regard to state subsidized crossings [3].

The two panel data dimensions relate to 1) a crossing identifier and 2) time periods indicated by waves. The crossing identifier (cid) is a unique number combining information about the county, company and crossing number. The waves consider time periods of one year and start with the value 1 in 1995 and end with 11 in 2005. Due to technical problems, the NPRA was unable to provide accounting information for 2001 and 2002, implying that waves 7 and 8 were omitted from the analysis. The short panel used in the following analysis thus considers only the years 1995-2000 and 2003-2005. A varying number of individual companies do not report their figures within the deadlines, thereby bringing unsystematic missing observations in the data set. Hence, the observations in the data set can be considered as a random draw from the total population of crossings.

The total number of crossings decreased from 145 in 1995 to 132 in 2005 because of improvements in infrastructure relieving ferries from service. Crossings with more than two ports of call were omitted because traffic and accounting figures cannot be allocated to the different parts of the crossing. The remaining observations have been reviewed for missing values and typing errors.

This study applies a balanced data set. However, unbalanced data sets can be used if the reason for missing data is not correlated with the random error term, ε_{it} (Wooldridge, 2006). The use of unbalanced data sets also avoids survivorship bias and the loss of sample size (Baum, 2006). Survivorship bias is, however, not a problem in the current analysis because the survival of the highly regulated ferry crossings in Norway is purely politically conditioned and, therefore, relatively uninfluenced by crossing or company specific abilities. That is, crossings are discontinued only if the politicians decide to and this takes place only in the rare occasions when ferries are substituted by bridges or tunnels. The balanced data set retains annual observations of 40 crossings (n), which is assumed to be a representative and sufficiently large sample providing 360 observations (N) over the studied waves (the missing waves 7 and 8 were excluded). Consequently, the applied data set is a balanced, but not compact, representative sample of the two-port crossings receiving subsidies from the Norwegian state.

Insert Table 1 about here.

The descriptive statistics for the variables are presented in Table 1 and indicate relatively large variation within the data set. For all variables, the within variation is considerably lower than the between variation and indicates large differences between crossings and relatively stable characteristics over the years. A comparison of the minimum and maximum values with the mean values indicates that the data set is “right skewed” with a few large crossings measured by PCE-km. The average crossing has total operating costs of NOK 22.8 million and transports 274,500 PCE and 2 million PCE-km. One of five crossings is located in unsheltered waters.

Applying the Panel Data Ferry Cost Model

The panel data specifications presented in equations (2), (3), and (4) are tested using Stata 9.2. A comparison of the estimates is performed to determine whether the random effects (GLS) estimates can be appropriately applied in the case of the Norwegian ferries.

Assessing the Model

The results of the within-group, between-group and random effects (GLS) estimates including parameters and main characteristics are presented in Table 2. A critical question is whether the relatively simple model presented in (1) is able to capture a relevant proportion of the variation in production and costs. Even though factors that are not included in the model, such as frequency and capacity, would change on occasion, the proportion of variance explained, indicated by R^2, is generally high and the F-test indicates good model fit. The coefficients were significant at a 1% level for all models, except the area dummy coefficient which was significant at a 10% level for between-effects. The within-group estimation provides a correlation between the unobserved individual effects with the explanatory variables on -0.44. A little more than two thirds of the variation in operating costs is explained by the unobserved individual effects (rho), for instance by properties such as the quality of ferries and the organizational structure.

Insert Table 2 about here.

All three methods of analysis returned estimates with quite similar coefficient values for the parameters α, β_1, β_2 and δ. The intercept parameter, α, in the random effects estimates, is the remaining constant after the unobserved effects have been accounted for. Furthermore, parameter β_1, indicating the costs of taking on board one more PCE, varies from 37 NOK, using the between-group analysis, to 40 NOK using the within-group analysis. Parameter β_2, indicating the costs of carrying a PCE unit one more kilometre, varies from 2.94 NOK using, the between-group analysis, to 3.10 NOK using the within-group analysis. The extra costs of operating a crossing in unsheltered waters, indicated by parameter δ, vary from about 3.9 million NOK, using between-group, to 4.3 million NOK using random effects.

The random effects combination of between-group and within-group estimates gives a median theta value of 0.754, thus giving high emphasis to the within estimator and indicating that the random effects method using both within and between variations will be optimal. However, some tests of the underlying assumptions must be carried out to determine whether the random effects model is appropriate.

Some degree of positive serial correlation is indicated.[4] The correlation between the error term, ε_{it}, and both the explanatory variables and the unobserved effects, u_i, is close to zero and not significant at a 5% level. Furthermore, the Breusch-Pagan Lagrange multiplier test, assessing whether pooled data should be used instead of random effects, strongly rejects the null hypothesis and favours the random effects model. The Hausman test, checking whether regressors are uncorrelated with individual effects, states that the probability is 0.10 (p-value) for returning a chi-squared value of 4.56, given the null hypothesis of non-systematic differences between the within-group and random effects methods. This p-value satisfies the 5% significance level and indicates that the random effects model is consistent. Also, the Mundlak test provides significant results similar to the Hausman test. Consequently, the random effects model is appropriate and gives the best unbiased estimates.

Empirical Results

Using the random effects coefficient estimates from Table 2, equation (5) presents the cost function from equation (1) with corresponding marginal costs (MC). Attention is given to the interpretation of marginal costs because they, as the principle basis for determining welfare optimal prices, can be related to fares and can offer relevant policy implications. According to the derived marginal costs, the distance-independent costs related to carrying one extra PCE unit on a ferry is 39.58 NOK and the distance-dependent costs for each kilometre a PCE unit is transported is 3.10 NOK.

$$C = 4942403 + 39.58 \cdot X_1 + 3.10 \cdot (X_1 X_2) + 4311391 \cdot Z_1$$

(5)

$$\frac{\partial C}{\partial X_1} = MC = 39.58 + 3.10 \cdot X_2$$

The results presented in equation (5), even though they are not directly comparable, could be related to earlier cross-sectional studies of costs in the ferry industry. Based on data from 1995 and 2003, respectively, Jørgensen et al. (2004) applied a modified translog function and Mathisen (2008) a power function. Adjusted to comparable 2006 prices, a linear approximation of the estimates provided by Jørgensen et al. (2004) provide the long-run marginal costs for transporting a PCE of $20.58 + 4.50 \cdot X_2$. The price adjusted estimate from Mathisen (2008) is $21.70 + 1.41 \cdot X_2$ with an additional increasing element related to capacity, making us unable to compare the parameter values directly. The new estimates from (5) using random effects (GLS) indicate a higher constant than the earlier studies and a less steep increase with respect to distance. Evidence indicates, thus, that the previous long-run marginal cost estimates have underestimated the distance-independent costs, while there is no clear conclusion with respect to the increasing slope.

Performing a pooled data analysis on the data set derives marginal costs of $37.43 + 2.95 \cdot X_2$. The coefficients indicate both a lower constant and a lower increase with respect to distance, compared to the random effects (GLS) estimates in equation (5). Hence, for all crossings, ranging from 1 km to 28 kms, the simple pooled data estimates indicate too low marginal cost estimates. Figure 1 illustrates the development of long-run marginal costs with respect to distance for the random effects (GLS) estimates, pooled data estimates and the current fare scheme - all in 2006 prices.

Insert Figure 1 about here.

In addition to generating revenues to meet budget restrictions, fares should be based on marginal costs in order to comply with welfare optimal principles. It is thus interesting, from a policy perspective, to compare the current fare system in the Norwegian ferry industry with the estimated values. Even though the transport authority states that fares should be optimally set according to long-run marginal costs (LRMC), they also have to consider regional policy goals such as keeping settlements in rural areas.

The current fare scheme for transporting one PCE is directly related to distance and derives a constant of 35.73 NOK and a distance-dependent element of 4.64 NOK per kilometre (2006 prices). Compared to the estimates in equation (5), the current fare scheme has a higher distance-dependent element and a slightly lower distance-independent element. As illustrated by intersection A in Figure 1, the current fare scheme charges above marginal costs for all distances above 2.5 km and the difference increases with respect to distance, both in absolute and relative values. This suggests that vehicles on longer crossings subsidize the vehicles on the shorter crossings.

Using data from 2005, the annual welfare loss through not charging fares according the random effects (GLS) estimates can be approximated to 90,000 NOK for an average crossing in the data set and a total of 8 million NOK for the entire Norwegian ferry industry. This value is calculated using mean distance (7 km) for an average sized crossing (315,000 PCE) in 2005. Changes in demand are calculated based on the experiences of Bråthen et al. (2006). The loss for the total industry is corrected to consider that the average size of all crossings is smaller than the average of the selection.

Conclusions and Implications

The applied panel data method analyzes observations along two dimensions and provides extended information about the costs of the Norwegian car ferry industry compared to previous studies using cross-sectional analysis. The linear multiproduct cost function for estimating operating costs of the Norwegian ferry industry consists of two explanatory variables and a dummy variable. Consisting of variations between crossings and within crossings, the random effects (GLS) model meets the econometric requirements and provides efficient and unbiased estimates.

According to the new estimates, the yearly operating costs have a constant of about NOK 4.9 million with an add-on of NOK 4.3 million for crossings located in open sea with rougher weather conditions. Further, the costs increase by about 61 NOK for each PCE at a mean distance crossing of 7 km.

The marginal cost of transporting one passenger car equivalent (PCE) is especially interesting for the policy makers of this highly regulated industry, because it should, principally, be providing the basis for setting fares to maximize social welfare. Comparison of the panel data marginal cost estimates is conducted in relation to a pooled data analysis, earlier cross-sectional studies and the current fare scheme. The pooled data estimates provide too low estimates of marginal costs for all distances and would give too low fares in the case of marginal cost pricing. Generally speaking, this is also the case with estimates from the earlier cross-sectional studies. The current fare scheme has a little lower constant and a steeper increase with respect to distance than the panel data estimates, making fares higher than marginal costs for all distances exceeding 2.5 km. The difference between the fare scheme and the marginal costs increases with distance and charges vehicles on longer crossings relatively higher prices than vehicles on shorter trips. In isolation, this is a contradiction of the welfare principles on which the fare scheme is based. However, this conclusion considers neither regional policy goals, such as keeping settlements in rural areas, nor possible pricing strategies, such as Ramsey pricing, to meet budget restrictions.

The proposed cost function explains more than 90% of the variance in costs using only three explanatory variables and it has good econometric properties. This method is simple and can be applied for deriving the expected operating costs on a crossing. It provides the transport authorities with more precise estimates of the operating costs in the car ferry industry. Such knowledge can prove important for example in assessing the economic effects and efficiency of the planned increased exposure to competition of the Norwegian ferry industry, using competitive tendering.

End notes

[1] Currency rates in 2010: 1 € ≈ 8 NOK, 1 $ ≈ 6 NOK.

[2] This was the subsidy level in 2004.

[3] Data related to one of the ferry companies is omitted because of systematic manipulation of accounting figures in an attempt of subsidy fraud (see e.g. Jørgensen and Mathisen, 2010).

[4] The Durbin Watson analysis for panel data was performed with the add-on file xtdw.ado in STATA version 9.2 (Nunziata, 2002).

References

Baltagi, B. H. (2005) *Econometric analysis of panel data* (3rd ed.). Chichester: Wiley.

Baum, C. F. (2006) *An introduction to modern econometrics using Stata*. College Station, Texas: Stata Press.

Baumol, W. J., Panzar, J. C. and Willig, R. D. (1988) *Contestable markets and the theory of industry structure*. San Diego: Harcourt Brace Jovanovich.

Braeutigam, R. R. (1999) Learning about transport costs. In J. A. Gomez-Ibanez (ed.), *Essays in transportation economics and policy: A handbook in honor of John R. Meyer*, pp. 57–97. Washington DC: Brookings Institution Press.

Bråthen, S., Fuglseth, A., Hervik, A., Larsen, O. I., Lyche, L., Odeck, J. and Sandvik, E. T. (2004) Anbud i ferjesektoren - erfaringer og utsikter (Tendering in the ferry industry - experiences and expectations). Molde, Norway: Møre Research. 0408 (In Norwegian).

Bråthen, S., Husdal, J. and Rekdal, J. (2006) Gratis bruk av ferjer. Noen mulige varianter (Free fares on ferries . Some options). Molde, Norway: Møre Research (In Norwegian) Free fares on ferries. Some options. Report 0614 (In Norwegian).

Bråthen, S. and Lillebakk, L. M. (2005) Nytt riksregulativ for ferjetakster - Samordning med AutoPASS (Revised national ferry fare scheme - Coordination with AutoPASS). Molde, Norway: Møre Research. 0510 (In Norwegian).

Hensher, D. A. and Brewer, A. (2001) *Transport: an economics and management perspective*. Oxford: Oxford University Press.

Jørgensen, F. and Mathisen, T. A. (2010) Using Standardized Revenue and Cost Norm Analyses to Reveal Subsidy Fraud in Contracted Public Transport Services. *Transport Reviews,* 30(3): 299-313.

Jørgensen, F., Mathisen, T. A. and Larsen, B. (2011) Evaluating transport user benefits and social surplus in a transport market - the case of the Norwegian ferries. *Transport Policy* 18(1): 76-84.

Jørgensen, F., Pedersen, H. and Solvoll, G. (2004) Ramsey pricing in practice: the case of the Norwegian ferries. *Transport Policy,* 11(3): 205–214.

Kumbhakar, S. and Lovell, C. A. K. (2000) *Stochastic frontier analysis*. Cambridge: Cambridge University Press.

Mathisen, T. A. (2008) Marginal Costs and Capacity Utilization: Calculating Short-Run, Medium-Term, and Long-Run Marginal Costs in the Ferry Industry. *International Journal of Transport Economics,* XXXV(3): 373-389.

NOU 30A (1977). *Norsk samferdselsplan* (Public Report of Norway no. 30A - Plan for Transport and Communcation in Norway). (In Norwegian).

Nunziata, L. (2002) xtdw.ado: a Stata command for Durbin Watson statistic in fixed effects panel data models. Version 2.0. http://www.nuff.ox.ac.uk/users/nunziata/publications.htm, accessed 15th November 2002.

Odeck, J. (2008) Efficiency measurement of ferries serving road networks in Norway: A stochastic frontier approach. *Maritime Economics and Logistics,* 10(4): 409-428.

Pels, E. and Rietveld, P. (2008) Cost functions in transport. In K. Button & D. A. Hensher (eds.), *Handbook of transport modelling*, (2nd ed.), Vol. 1, pp. 381–394. Amsterdam: Pergamon.

Solvoll, G. and Jørgensen, F. (2001) Ferjetakster: takstmodell for prøvesamband (Ferry fares: fare scheme for trial ferry services). Bodø: Nordland Research Institute. Report 13/2001 (In Norwegian).

Statistics Norway (2006). Statistical yearbook of Norway 2006. Statistics Norway, Oslo.

Syrjänen, M., Bogetoft, P. and Agrell, P. (2006). *Efficiency benchmarking project B: analogous efficiency measurement model based on Stochastic Frontier Analysis. Final Report, 11.12.2006.*

The Norwegian Ministry of Petroleum and Energy (2002). *Storting Report No. 9 (2002-2003) Om innenlands bruk av naturgass mv.*

The Norwegian Public Roads Administration. (2006) Konsekvensanalyse. Håndbok 140 - Veiledning (Handbook in consequence analysis). (In Norwegian).

Winston, C. (1985) Conceptual developments in the economics of transportation: an interpretive survey. *Journal of Economic Literature,* 23(1): 57–94.

Wooldridge, J. M. (2006) *Introductory econometrics: a modern approach* (3rd ed.). Mason, OH: Thomson South-Western.

Figures

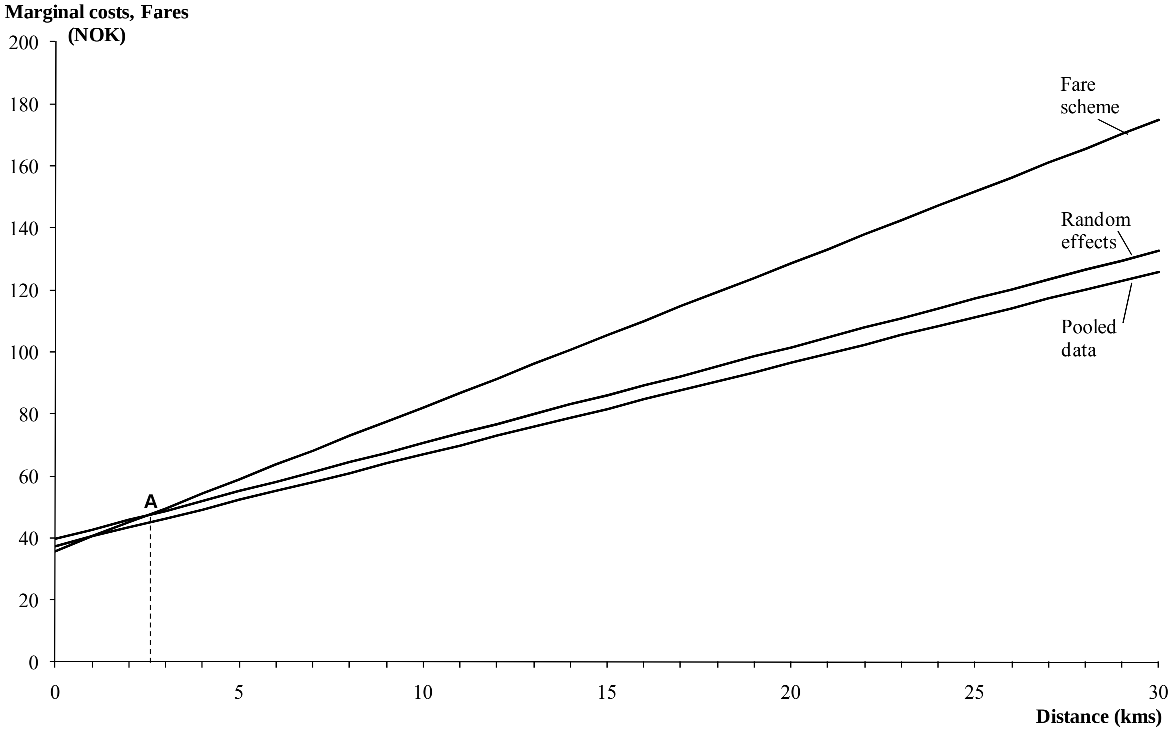

Figure 1 – Estimated marginal costs and current fare scheme of the Norwegian ferry industry

Tables

Table 1 – Panel data descriptive statistics for the variables

Variable		Mean	Std. Dev.	Min.	Max.	Observations
Costs (NOK 2006 prices)	Overall	$2.28 \cdot 10^7$	$2.09 \cdot 10^7$	5 487 968	$1.37 \cdot 10^8$	N = 360
	Between		$2.01 \cdot 10^7$	6 985 473	$9.51 \cdot 10^7$	n = 40
	Within		6 487 995	$-1.51 \cdot 10^7$	$6.46 \cdot 10^7$	T = 9
X_1 (PCE)	Overall	274 480	272 664	16 753	1 525 943	N = 360
	Between		266 664	20 204	1 059 214	n = 40
	Within		69 430	-50 143	741 209	T = 9
X_1X_2 (PCE-kms)	Overall	1 979 944	3 838 653	48 331	$2.77 \cdot 10^7$	N = 360
	Between		3 778 115	56 896	$2.06 \cdot 10^7$	n = 40
	Within		882 717	-3 265 721	9 052 223	T = 9
Z_1 (rough weather)	Overall	0.20	0.40	0.00	1	N = 360
	Between		0.41	0.00	1	n = 40
	Within		0.00	0.20	0.20	T = 9

Table 2 – Estimation results for within-group, between-group and random effects (GLS)

	Within-group			Between-group			Random effects (GLS)		
	Coeff.	**Std.err.**	**t-value**	**Coeff.**	**Std.err.**	**t-value**	**Coeff.**	**Std.err.**	**t-value**
α (Constant)	5 037 586	852 315	5.9	6 025 567	1 290 667	4.7	4 942 403	1 089 624	4.5
β_1 (X_1)	40.49	4.87	8.3	37.07	4.62	8.0	39.58	3.34	11.9
β_2 (X_1X_2)	3.36	0.38	8.8	2.94	0.32	9.1	3.10	0.25	12.5
δ (Z_1)	(dropped)	–	–	3 940 114	1 972 758	2.0	4 311 391	1 961 708	2.2
rho	0.70			–			0.63		
R^2 (overall)	0.92			0.92			0.92		
Model fit	0.00			0.00			0.00		