Metro Station Operating Costs: An Econometric Analysis

Mohammed Quddus, Loughborough University, Leicestershire, UK
Nigel Harris and Daniel J Graham, Imperial College London, London, UK

Abstract

This article develops an econometric analysis of metro station operating cost to identify factors that create variation in cost efficiency. Stations operating costs can be classified amongst the semifixed costs that a metro faces in the sense that they do not vary proportionately with metro output. They may therefore be important in determining the degree of returns to density. This article seeks to provide an improved understanding of some of the major factors driving these costs. Empirical results show that strong system-specific influences impact costs but over and above these we detect positive associations from a range of station characteristics, including the length of passageways, number of platforms, peak-level service frequency, interchange demand, and the provision of toilet facilities. In addition, we find that the presence of air-conditioning has a substantial effect in increasing expected station operating cost by as much as 40 percent.

Introduction

The cost structure of the mainline railway industry has received a great deal of attention in the academic literature (e.g., Caves et al. 1980; Caves et al. 1981a, 1981b; Freeman et al. 1985; Caves et al. 1985; Dodgson 1985; McGeehan 1993; Bookbinder and Qu 1993; Oum and Yu 1994; Cowie and Riddington 1996; Wunsch 1996; Tretheway et al. 1997; Oum et al. 1999; Cantos et al. 1999; Cantos et al. 2002). Research has demonstrated the very large variance in cost efficiency, or

productivity, that is often present within a sample of rail firms and has developed cost and production function approaches to analyze the factors underpinning this variance.

A prominent theme in the rail efficiency literature is whether cost structures are subject to returns to scale (RTS) or returns to density (RTD). RTS describe the relationship between all inputs and the overall scale of operations, including both output and network size. RTD describe the relationship between inputs and outputs with the rail network held fixed. Evidence in the literature indicates that RTD are due to the prevalence of fixed costs in the rail industry and to a range of semifixed costs that do not vary proportionally with output. Less consistent evidence is available on the existence of scale economies, though the majority view is that railways operate under constant returns to scale. Few studies have been conducted on the costs structure of urban metros, though Graham et al. (2003) estimates increasing RTD and constant RTS.

Station operations may provide an important source of increasing RTD in metro operations. Stations must remain staffed and functioning, with all the energy and other resources required, throughout the duration of the metro operating hours. Moreover, costs may differ quite substantially from one station to another due to the nature of engineering, the depth of station, its size and dimensions, the technology employed, and so on. So we can conceive of station operating costs as semifixed costs that do not vary proportionately with system throughput and therefore may be instrumental in giving rise to increasing RTD.

In this study we develop an econometric model to analyze variance in station operating costs. An econometric model is essential to determine the effect of a particular characteristic of a metro station on its operating costs while controlling for all other factors affecting the metro station operating cost. The analysis of historical data fails to control for the effects of other factors while estimating the effect of a particular factor. We use data on 83 stations from 13 metro systems from around the world to estimate the main drivers of cost. Model specifications and the data used for estimation are discussed and results are presented.

Model Specification and Data

The data available for our analysis describe the total operating cost of each station and a range of station characteristics collected from a total of 13 metros (Buenos Aires, Dublin, Glasgow, Hong Kong KCR, Hong Kong MTR, Lisbon, London, Montreal, Naples, Sao Paulo, Singapore, Taipei, and Toronto). The analysis we develop

below regresses the total operating costs against these station characteristics to determine their role in influencing variance in costs.

It is important to stress that we do not adopt a conventional cost function approach. We do not have data on factor prices and therefore cannot estimate the cost function. However, another important consideration in this respect is that since the operating costs of any one particular station represents only a small fraction of total metro operating costs, individual stations cannot be regarded as the appropriate units over which cost decisions are made. For instance, metro operators do not demand factor inputs at the station level in accordance with prices but make rational decisions relating to costs and operations for the system as a whole. Furthermore, it would be wrong to ascribe any particular behavioral assumptions to individual stations (e.g., cost minimizing behavior). A metro may not seek to sustain a set level of station efficiency across the system but rather allow for disparities in efficiency to achieve some broader objectives relating to the appropriate level of system output given overall costs.

In this respect, it is mainly how the station *characteristics* serve to influence total cost that is of interest in the present analysis. One important issue, however, relates to the absence of factor price data, because this will certainly be important in determining station costs. To control for these omitted variables, which we cannot observe, we estimate the station operating costs model with a set of dummy variables for the 13 metro systems. We assume that these dummies will capture unobserved system-specific effects including factor prices.

A log linear model is used to identify the factors that influence the operating cost of a metro station. The model can be written as:

$$1n\, y_i = \alpha + \beta 1n X_i + \theta D_i + \varepsilon_i \qquad (1)$$

where:

y_i	is the total operating cost of a metro station *i*
X_i	represents a kx1 vector of continuous explanatory variables describing the characteristics of station *i*
D_i	denotes a mx1 vector of dummy explanatory variables relating to metro systems
ε_i	is white noise
β	represents kx1 vector of parameters to be estimated
θ	is a mx1 vector of parameters to be estimated

The log linear model is used because it reduces the potential for multicollinearity and provides direct parameter estimates of the elasticities.

The dependent variable is the total cost of operating the station per year. This includes the costs associated with staff, utilities (e.g., electricity, gas, and water), the maintenance of lifts and escalators, and the maintenance of other systems such as CCTV, air-conditioning, ticketing equipment, and building. Two econometric models will be estimated using equation (1). The first model will be estimated without metro-specific dummy variables and the second model will be estimated with metro-specific dummies to control country-specific effects on metro stations' operating cost.

The explanatory variables, which describe the station characteristics, and the hypotheses we seek to test with each variable are described below.

Age of the station. Age of the station is taken as the number of years since the station opened. This figure is averaged if the station was opened in stages. Our hypothesis is that older stations will incur higher maintenance costs than new stations.

Lifts and escalators. The number of lifts and escalators within a station may influence the operating cost because this equipment needs to be in operation on a daily basis and frequently maintained.

Number of ticket machines/ticket offices/ticket sales windows/entry and exit gates. The number of ticket machines includes only those machines used by the public to purchase or validate tickets. The number of ticket offices is the number of areas in the station where ticket-selling takes place. The number of ticket sales windows relates to the number of potentially staffed positions used by the metro staff to sell tickets to passengers. We hypothesis that these factors will influence the staff costs of the station.

Number of opening hours per day. This variable is taken as the average number of metro station opening hours per day. The hypothesis is, of course, that longer operating hours induce higher costs.

Service frequency. Two service frequency variables—peak frequency and off-peak frequency— are considered. Frequency is calculated as the average number of trains per hour (each way) during peak periods (peak frequency) or off-peak periods (off-peak frequency). The inclusion of these variables will allow us to test whether costs are associated with frequency.

Length of trains. This is calculated as the total number of carriages of a train using the station. At stations with multiple lines averages are used.

Platform dimensions. Variables for width, length, and elevation of the platform are used to determine if these factors are important for maintenance and cleaning costs.

Roof length of platforms. For underground stations, this is clearly the same as platform length, but for at-grade and elevated stations only part of platforms may be covered by a canopy, shelter, or overall roof. This variable is included to understand if variation in the maintenance associated with roof lengths affects total station costs.

Length of passageways. This is measured as the total length of passageways, including escalator shafts, estimated by metros as an indicative proxy for the amount of cleaning and building repair that may need to be done. No account is taken of possible variations in passageway width. A better measure might have been the total floor area, but this would not have directly reflected the amount of walls and ceilings that need maintenance and cleaning; this is also discarded as a measure because it is more difficult for metros to estimate easily.

Station demand variables. The two main demand variables considered are entry demand and interchange demand. Station entry demand is the total number of passengers entering the station per year. This includes passengers changing modes at the station, and entering from mainline rail or bus stations, as well as those starting their journeys locally and entering the station on foot. Interchange demand relates only to those passengers changing metro lines at the station concerned. Two secondary variables—peak entry demand and peak interchange demand—are also considered. Peak entry (interchange) demand is calculated as the total number of entry (interchange) passengers for the busiest hour during a standard week, and is designed to test whether peak demand (entry/interchange) drives station capacity and hence costs, or total demand drives staffing levels and hence cost.

Types of metro stations. Dummy variables are used to reflect the overall type of metros in terms of being at-grade, elevated, subsurface (typically constructed by cut and cover, and 5–6m below ground), or deep tube. At-grade and subsurface stations can be managed without lifts or escalators for passengers to travel vertically, whereas elevated and deep tube stations normally need this equipment, which adds significantly to costs (e.g., electricity).

Other variables. Presence of air-conditioning, toilets for public use, platform screen doors, and shops are all included in modeling through a dummy variables. Each is thought to generate costs (electricity, cleaning, maintenance, and management time).

Results

Prior to model fitting, a number of statistical tests were performed to determine the nature of the data. For example, it is possible that the explanatory variables may be correlated with each other (the effect of multicollinearity) or that the data exhibits heteroskedasticity (the effect of nonconstant variance).

Although imperfect multicollinearity does not violate the assumptions of the classical model, if its presence is sufficiently acute, it can lead to biased, inefficient, and even wrongly signed estimates. If the overall goodness of fit, R^2, is relatively high (say more than 0.8) but only few explanatory variables are significantly different from 0 or there are high pair-wise correlations among the regressors, then it is possible that multicollinearity may be present. Here, we use the variance inflation factor (VIF) proposed by Chatterjee et al. (2000) to determine the presence of multicollinearity. The number of ticket gates at a station, for example, is found to be highly correlated with the entry demand at the station, and the length of the platform at a station is correlated with the length of the longest train passing the station. Based on the VIF test, the highly correlated variables are excluded from the explanatory variables used in the final model. The problem of Omitted Variable Bias (OVB) is addressed in the conventional way by the use of proxy variables and fixed effects to control for unobserved metro-specific variables. We have no evidence that multicollinearity affects the parameter estimates.

Data from London metro stations are not included in the model as operating costs are not obtainable at the station level for the categories which are consistent with the other metros. This reduces the total number of observations to 83. However, we still have to estimate more than 30 parameters which are found to be uncorrelated with each other. Some of the explanatory variables such as entry and interchange demand, lifts, and escalators are then combined to minimize the number of parameters to be estimated. A dummy variable is used to represent the presence of lifts or escalators within a station in the model. This variable takes on a value of 1 if there are any lifts or escalators in a metro station and a value of 0 otherwise. Summary statistics (observations, mean, standard deviation, minimum, and maximum) of the final explanatory variables used in the model are shown in Table 1.

Table 1. Summary Statistics of Explanatory Variables Used in the Model

Variables	***Obs***	***Mean***	***Std. Dev.***	***Min***	***Max***
Operating cost (US$)	83	933731.1	831858.3	148085.7	4347232
Age of station (years)	83	37.265	39.326	1	170
Ticket office (1=Yes, 0=otherwise)	83	0.892	0.313	0	1
Total number of entrances	83	3.434	2.232	1	11
Opening hours per day	83	18.800	0.961	17	21
Presence of escalators or lifts (1=Yes, 0=otherwise)	83	0.928	0.261	0	1
Total passages (m)	83	349	448	0	2808
Total number of platforms	83	2.337	0.994	1	6
The width of the platform (m)	83	8.923	7.684	3	30
The length of the longest train (carriages)	83	6.394	1.848	3	12
Peak frequency (each way train/h)	83	17.845	8.245	4	36
Demand (entry+interchange) (passenger p.a.)	83	14200000	19700000	270000	98900000
Roof length (m)	83	122.024	62.853	0	300
Air-conditioning (1=Yes, 0=otherwise)	83	0.229	0.423	0	1
Presence of toilets (1=Yes, 0=otherwise)	83	0.434	0.499	0	1
Types of metro stations					
At grade	83	0.108	0.313	0	1
Elevated	83	0.265	0.444	0	1
Subsurface	83	0.072	0.261	0	1
Tube	83	0.554	0.500	0	1
Metros 1					
Metro-1	83	0.072	0.261	0	1
Metro-2	83	0.157	0.366	0	1
Metro-3	83	0.072	0.261	0	1
Metro-4	83	0.036	0.188	0	1
Metro-5	83	0.084	0.280	0	1
Metro-6	83	0.084	0.280	0	1
Metro-7	83	0.072	0.261	0	1
Metro-8	83	0.084	0.280	0	1
Metro-9	83	0.084	0.280	0	1
Metro-10	83	0.084	0.280	0	1
Metro-11	83	0.072	0.261	0	1
Metro-12	83	0.096	0.297	0	1

1. Metro names are omitted to preserve confidentiality.

Another important assumption of the classical linear regression model is that the disturbances appearing in the regression function are homoskedastic. The problem of heteroskedasticity is common in cross-sectional analysis because the data usually involves observations from heterogeneous units (i.e., stations from different metros), and therefore heteroskedasticity may be expected if data from small, medium, and large stations are sampled together. In conducting the Park Test (Park 1966), we find that our data are not characterized by heteroskedasticity. This may be due to the use of the log linear model, which reduces the variances among the variables.

Table 2 presents our results. Two models are considered: one without metro dummies and one with metro dummies. The second model includes the metro-specific effects to control for heterogeneous environments. Ramsey's RESET test (an F-test) is used to select the better model (Ramsey 1969) and this shows that the addition of metro station dummies significantly increases the goodness of fit of the model. Therefore, the model with the metro station dummies is used for the interpretation of the results.

The model goodness of fit, the adjusted R^2, is 0.88, which shows a good degree of explanatory power for a cross-sectional model. The comparison between the observed cost and the predicted cost is shown in Figure 1. The mean prediction error is found to be only 2.3 percent. Note that the names of the metros are omitted to preserve confidentiality.

Figure 1. Observed and Predicted Costs

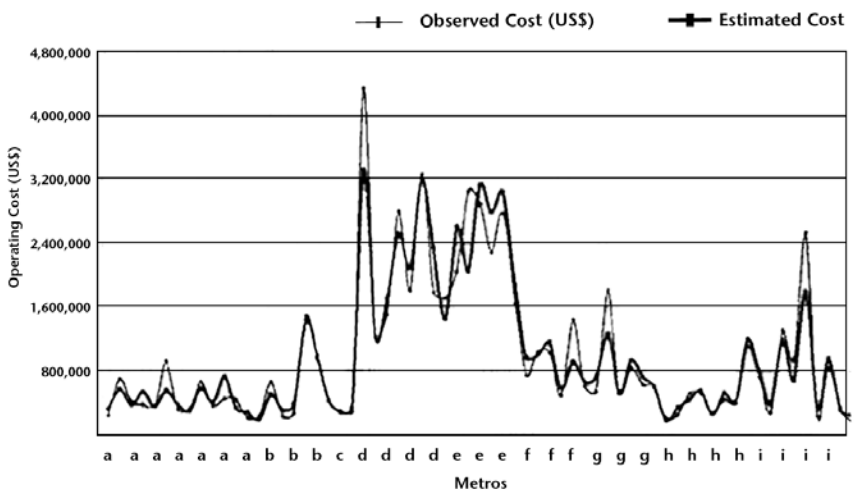

Table 2. Model Estimation Results for the Operating Cost of a Metro Station

Dependent Variable=ln [*total operation cost (US$)*]	Log Linear Model					
	Without Metro Dummies			***With Metro Dummies***		
Explanatory variables	Coeff	*t*-stat	*p*-value	Coeff	*t*-stat	*p*-value
ln [age of station (years)]	-0.070	-0.90	0.37	**-0.100**	**-1.71**	**0.09**
Ticket office (1=Yes, 0=otherwise)	0.073	0.29	0.77	0.070	0.28	0.78
ln (total number of entrances)	0.064	0.60	0.55	-0.033	-0.40	0.69
ln (opening hours per day)	0.055	0.03	0.97	0.447	0.12	0.90
Presence of Escalators or Lifts (1=Yes, 0=otherwise)	-0.301	-1.13	0.26	-0.229	-1.27	0.21
ln [total passages (m)]	**0.093**	**1.96**	**0.06**	**0.142**	**2.56**	**0.01**
ln (total number of platforms)	0.178	1.04	0.30	**0.270**	**2.25**	**0.03**
ln (width of the platform)	**-0.303**	**-2.39**	**0.02**	-0.113	-0.91	0.37
ln [length of the longest train (vehicles)]	0.258	0.86	0.39	0.423	1.35	0.18
ln [peak frequency (each way train/h)]	**0.510**	**3.01**	**0.00**	**0.479**	**3.63**	**0.00**
ln [demand (entry+interchange) (passenger p.a.)]	**0.150**	**2.11**	**0.04**	**0.096**	**2.08**	**0.04**
ln [roof length (m)]	0.100	1.43	0.16	**0.083**	**1.65**	**0.10**
Types of stations						
At grade (Reference)	-	-	-	-	-	-
Elevated	-0.180	-0.78	0.44	0.093	0.56	0.58
Subsurface	-0.364	-0.96	0.34	0.324	1.12	0.27
Tube	-0.312	-1.25	0.21	0.332	1.54	0.13
Air-conditioning (1=Yes, 0=otherwise)	**0.740**	**3.81**	**0.00**	**0.345**	**2.02**	**0.05**
Presence of toilets (1=Yes, 0=otherwise)	-0.106	-0.86	0.39	**0.324**	**2.85**	**0.01**
Metro-specific dummy 1						
Metro-1 (Reference)	-	-	-	-	-	-
Metro-2				-0.080	-0.21	0.84
Metro-3				**1.213**	**2.72**	**0.01**
Metro-4				0.537	0.93	0.36
Metro-5				**0.659**	**1.89**	**0.06**
Metro-6				**0.623**	**1.69**	**0.10**
Metro-7				**0.946**	**2.45**	**0.02**
Metro-8				-0.360	-0.77	0.44
Metro-9				-0.240	-0.57	0.57
Metro-10				-0.300	-0.71	0.48
Metro-11				0.134	0.47	0.64
Metro-12				0.165	0.61	0.55
Constant	9.051	1.91	0.06	7.27	0.68	0.50
Descriptive statistics						
Number of observations (valid)	83.00			83.00		
R-squared	0.70			0.92		
Adjusted R-squared	0.63			0.88		

1. Metro names are omitted to preserve confidentiality.

Table 2 shows a number of statistically significant effects on metro station operating cost that arise having controlled for unobservable system-specific effects.

The age of the station is found to be negatively associated with the operating cost of a metro station at the 90 percent confidence level. This is surprising as we would expect an older station to require more maintenance and hence be associated with higher costs. The explanation of this counterintuitive finding may be due to the fact that more recent stations (e.g., KCR, Hong Kong) tend to be larger and to have higher quality facilities, which also require a relatively high maintenance treatment.

Length of passageways, total number of platforms, peak-hour service frequency, and entry and interchange demand are found to be statistically significant at the 95 percent confidence level and positively associated with the operating cost. These results confirm our hypotheses. The elasticity associated with the peak-period service frequency is higher compared to others. The result suggests that a 10 percent increase in peak-period service frequency (each way, per hour) is associated with a 4.8 percent increase in the operating cost, and a 10 percent increase in the number of platforms leads to a 2.7 percent increase in the operating cost. The length of the roof is also found to be positively associated with the cost but only at the 90 percent confidence interval.

The effect of air-conditioning is captured by a dummy variable. This variable is found to be positively associated with the operating cost and is statistically significantly different from 0 at the 95 percent confidence level. This is an indication that average operating cost is high in a station with air-conditioning if all other factors remain constant. The coefficient (θ) of the effect of the air-conditioning is 0.35, indicating that the relative effect on the average operating cost due to the presence of air-conditioning is $100*\{\exp(\theta)-1\}$, or 41 percent. In other words, air-conditioning has an extremely large impact on costs, increasing the expected operating cost by 41 percent, holding all other factors included in the model constant.

The presence of toilets within a station is also found to be positively associated with the operating cost. This is expected as some costs are associated with the maintenance and staffing of toilets. However, the coefficient of this variable is unexpectedly high, perhaps because this variable represents the effects of some other factors that are not included in the model.

Interestingly, the type of metro station has little effect on operating cost. As explained previously, a categorical variable (grade, subsurface, elevated, and tube) is used to reflect the overall type of metro station. None of the coefficient estimates are statistically significant at the 95 percent confidence level. The tube-type metro station shows a positive coefficient relative to the at-grade-type station but only at the 87 percent confidence level.

The system-specific dummy variables are expressed relative to and intercept for Metro-1. The result suggests that Metro-3, Metro-5, Metro-6, and Metro-7 are costlier compared to Metro-1. The operating cost associated with Metro-5, for instance, is about 93 percent higher relative to Metro-1 if all other factors included in the model remain constant.

The number of ticket offices in a station, total number of entrances, operating hours per day, presence of lifts or escalators, width of platforms, and length of the longest train are found to be statistically insignificant. This is perhaps because the metro-specific dummies included in the model pick up the effects hypothesized from these factors.

The models are reestimated without the statistically insignificant variables (below 90% confidence level) of the models presented in Table 2 (with metro dummies). The results are shown in Table 3. Interestingly, the model goodness of fit remains the same after excluding five insignificant explanatory variables. The age of the station now becomes insignificant. As expected, the metro dummies now pick up most of these effects. The operating cost of Metro-2, Metro-8, Metro-9, and Metro-10 are now lower relative to Metro-1. The effects of all other factors remain invariable.

A limitation of the analysis is that not all of the characteristics associated with cost are readily alterable. For instance, the length of a metro station's roof plays a major role in increasing the station's operating cost.

The standard errors associated with the parameter estimates give us a guide as to how confident we can be in the magnitudes indicated by our results. Of course, econometric models can be revised given better data or new hypotheses to test, but an econometric model should be assessed based on both the "significance" of a variable and "the estimated magnitude of the effects" of the variable as this is one of the fundamental objectives of estimating an econometric model.

Table 3. Reestimated Models with Significant Variables of the Models Shown in Table 2

Dependent Variable=ln [*total operation cost (US$)*]	***Log Linear Model***					
	Without Metro Dummies			***With Metro Dummies***		
Explanatory variables	Coeff	*t* -stat	*p*-value	Coeff	*t*-stat	*p* -value
Age of station (years)	0.010	0.16	0.87	-0.059	-1.21	0.23
Total passages (m)	**0.123**	**3.12**	**0.00**	**0.090**	**2.08**	**0.04**
Total number of platforms	0.193	1.15	0.25	**0.317**	**2.89**	**0.01**
Peak frequency (each way train/h)	**0.342**	**2.50**	**0.02**	**0.477**	**3.88**	**0.00**
Demand (entry+interchange) (passenger p.a.)	**0.129**	**2.08**	**0.04**	**0.088**	**2.27**	**0.03**
Roof length (m)	**0.122**	**1.99**	**0.05**	**0.115**	**2.72**	**0.01**
Types of stations						
At grade (reference)	-	-	-	-	-	-
Elevated	**-0.446**	**-2.15**	**0.04**	0.002	0.01	0.99
Subsurface	**-0.523**	**-1.80**	**0.08**	0.203	0.80	0.43
Tube	**-0.499**	**-2.43**	**0.02**	0.209	1.11	0.27
Air-conditioning (1=Yes, 0=otherwise)	**0.631**	**4.25**	**0.00**	**0.304**	**1.95**	**0.06**
Presence of toilets (1=Yes, 0=otherwise)	-0.164	-1.44	0.15	**0.280**	**2.69**	**0.01**
Station-specific dummy 1						
Metro-1 (reference)	-	-	-	-	-	-
Metro-2				**-0.374**	**-2.10**	**0.04**
Metro-3				**0.981**	**3.74**	**0.00**
Metro-4				0.018	0.06	0.95
Metro-5				**0.464**	**1.76**	**0.08**
Metro-6				**0.584**	**1.90**	**0.06**
Metro-7				**0.467**	**2.06**	**0.04**
Metro-8				**-0.453**	**-2.15**	**0.04**
Metro-9				**-0.739**	**-2.54**	**0.01**
Metro-10				**-0.475**	**-1.94**	**0.06**
Metro-11				-0.045	-0.22	0.82
Metro-12				-0.018	-0.09	0.93
Constant	9.481	12.49	0.00	9.388	16.90	0.00
Descriptive statistics						
Number of observations (valid)	83			83		
R-squared	0.69			0.92		
Adjusted R-squared	0.64			0.88		

1. Metro names are omitted to preserve confidentiality.

Conclusions

We have developed an econometric model to investigate variance in metro station operating costs. The model regresses total metro station operating costs on a series of station characteristics and a set of metro systems' specific dummy variables. The results show strong unobserved system-specific effects, confirming the need to differentiate the data in this way. Over and above the system-specific

effects, we have identified some factors that appear to have an important influence on the levels of station costs. These include length of passageways, number of platforms, peak-level service frequency, interchange demand, and the provision of toilet facilities. In addition, we find a very strong effect from the existence of air-conditioning, which raises the expected station operating cost by as much as 40 percent.

Stations operating costs can be classified amongst those semifixed costs that do not vary proportionately with metro output. For this reason, they may be very important in determining the magnitude of RTD on the costs structure and productive efficiency of the firms. This article has provided an improved understanding of some of the major factors driving these costs.

References

Bookbinder, J., and W. Qu. 1993. Comparing the performance of major American railroads. *Transportation Research Forum* 33: 70–85.

Cantos, P., J. Pastor, and L. Serrano. 1999. Productivity, efficiency and technical change in the European railways: A non-parametric approach. *Transportation* 26: 337–357.

Cantos, P., J. Pastor, and L. Serrano. 2002. Cost and revenue inefficiencies in the European railways. *International Journal of Transport Economics* 29: 279–308.

Caves, D. W., L. R. Christensen, and J. A. Swanson. 1980. Productivity in U.S. railways 1951–1974. *Bell Journal of Economics* 11: 166–181.

Caves, D. W., L. R. Christensen, and J. A. Swanson. 1981a. Productivity growth, scale economies, and capacity utilization in the U.S. railways. *American Economic Review* 71: 994–1002.

Caves, D. W., L. R. Christensen, and J. A. Swanson. 1981b. Economic performance in regulated and unregulated environments: A comparison of U.S. and Canadian railways. *The Quarterly Journal of Economics* 96: 559–581.

Caves, D. W., L. R. Christensen, M. W. Tretheway, and R. J. Windle. 1985. Network effects and the measurement of returns to scale and density in U.S. Railways, in Daugherty, A.F., ed., *Analytical Studies in Transport Economics*. Cambridge: Cambridge University Press.

Chatterjee S, A. S. Hadi, and B. Price. 2000. *Regression analysis by example*, 3rd ed. New York: John Wiley & Sons.

Cowie, J., and G. Riddington. 1996. Measuring the efficiency of European railways. *Applied Economics* 28: 1027–1035.

Dodgson, J. S. 1985. A survey of recent developments in the measurement of rail total factor productivity, in Button, K. J., and Pitfield, D. E., eds., *International Railway Economics*. Aldershot: Gower.

Freeman, K.D., T. H. Oum, M. W. Tretheway, and W. G. Waters II. 1985. The total factor productivity of the Canadian Class I railways: 1956–1981. *Logistics and Transportation Review* 21: 249–276.

Graham, D. J., A. Couto, W. E. Adeney, and S. Glaister. 2003. Economies of scale and density in urban rail transport: Effects on productivity. *Transportation Research* E 39: 443–458.

McGeehan, H. 1993. Railway costs and productivity growth: The case of the Republic of Ireland, 1973–1983. *Journal of Transport Economics and Policy* 27: 19–32.

Oum, T. H., and C. Yu. 1994. Economic efficiency of railways and implications for public policy: A comparative study of OECD countries' railways. *Journal of Transport Economics and Policy* 28: 121–138.

Oum, T. H., W. G. Waters II, and C. Yu. 1999. A survey of productivity measurement and efficiency in rail transport. *Journal of Transport Economics and Policy* 33: 9–42.

Park, R. E. 1966. Estimation with heteroskedasticity error terms. *Econometrica* 44 (3): 461–465.

Ramsey, J. B. 1969. Tests for specification errors in classical linear least squares' regression analysis. *Journal of the Royal Statistical Society*, series B 31: 350–371.

Tretheway, M. W., W. G. Waters II, and A. K. Fok. 1997. The total factor productivity of the Canadian railways, 1956–91. *Journal of Transport Economics and Policy* 31: 93–113.

Wunsch, P. 1996. Cost and productivity of major urban transit systems in Europe. *Journal of Transport Economics and Policy* 30: 171–186.

About the Authors

Mohammed A. Quddus (*m.a.quddus@lboro.ac.uk*) is a lecturer in transport studies at Loughborough University in the UK. He received his Ph.D. from Imperial College London in 2006 in the area of map matching algorithms for transport telematics applications. His main research interests include transport planning and policy, transport risk and safety, intelligent transport systems, and geographic information science (GIScience).

Nigel Harris (*nigel.harris@railcons.com*) is a railway planner with research interests in fares, network modeling, service planning, demand forecasting, scheme appraisal, and railway business planning. He has coauthored/edited key texts on the privatization of British Rail and on planning passenger and freight railways, as well as having published more than 50 papers. He is a visiting lecturer at the Universities of Newcastle and Birmingham.

Daniel Graham (*d.j.graham@imperial.ac.uk*) is Senior Research Fellow in the Centre for Transport at Imperial College London. He was previously at the London School of Economics where he received his Ph.D. in 1996. He currently works on a range of themes in transport economics and policy and in urban and regional economics.