ECONOMIC EFFICIENCY OF RAILWAY
IN UKRAINE: RAILWAY ECONOMIES
OF SCALE, SCOPE, AND DENSITY

by

Dmytro Dushko

A thesis submitted in partial afulfillment of the
requirements for the degree of

MA in Economic Analysis

Kyiv School of Economics

2022

Thesis Supervisor:______________________ Professor Oleg Nivievskyi

Approved by ____
Head of the KSE Defense Committee, Professor

____

____

____

Date ____

Kyiv School of Economics

Abstract

ECONOMIC EFFICIENCY OF RAILWAY
IN UKRAINE: RAILWAY ECONOMIES
OF SCALE, SCOPE, AND DENSITY

by Dushko Dmytro

Thesis Supervisor: Professor Nivievskyi Oleg

Ukraine railways are on the way to reform from a natural monopoly to a competitive company and this process is still ongoing. At the same, time there are no empirical studies that were developed to understand the optimal structure of UZ and which type of two railways model is better for Ukrainian railways: European type (vertical separation) or American type (horizontal separation). With the unique data provided by UZ this research aims to investigate whether existing economies of scale, scope, and density on UZ, and by answering this question, we can suggest to policymakers which type of model is optimal for UZ. Also, by using Stochastic frontier analysis we want to investigate the efficiency of railways units (branches) and the overall efficiency of UZ.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION 1
CHAPTER 2. LITERATURE REVIEW 12
2.1 Parametric methods estimation of the econometric model 12
2.2 Stochastic frontier approach (SFA) 14
CHAPTER 3. METHODOLOGY 15
3.1 Estimation of the cost function 15
3.2 Stochastic frontier function 18
CHAPTER 4. DATA CONSTRUCTION AND DESCRIPTION 24
CHAPTER 5. EMPIRICAL RESULTS 26
5.1 Translog estimation 26
5.2 SFA estimation 27
CONCLUSIONS 27
WORKS CITED 28
APPENDIX A 35
APPENDIX B 36

LIST OF FIGURES

Number *Page*

Figure 1.Cargo origin points 4

Figure 2. Freight turnover by types of transport, 2015-2021.............................5

Figure 3. Railways utilization 6

Figure 4. Rolling stock productivity...8

Figure 5. Cost frontier and efficiency...20

Figure 5. Correlation plot of variables of interest...25

LIST OF TABLES

Number *Page*

Table 1. Descriptive statistics of independent and dependent variables 24

Table 5.1. Estimation results for the total cost function-first-order coefficients and indicators for economies of scale 28

Table 5.2. SFA estimation results 29

Table 5.3. Estimation of mean efficiency 31

ACKNOWLEDGMENTS

I would like to express deep gratitude to all Research Workshop professors and especially my thesis advisor Prof. Oleg Nivievskyi, without him, the idea of this work would not be put to life. I deeply appreciated their constant support, insightful suggestions, and motivation during the process of accomplishment of my thesis. I would also like to admit the contribution of Ivan Yuryk, without his help with the data, this thesis would not have been possible.

I would also like to thank all the KSE community, faculty members, and all classmates of EA'22 for helping me during my study at KSE. My special appreciation to Anna Shcherbiak, Daniel Krasovytskyi, and Mykyta Horovoi for their encouragement and care and who helped me to push my study through two years.

There are no words to express my gratitude and love to my mother Inessa, who constantly encouraged and helped me to remain resistant and optimistic in the face of all difficulties.

I would also like to express special gratitude to our armed forces, who are defending Ukraine and its borders at the cost of their lives so that we can live in our homes and work.

Chapter 1

INTRODUCTION

In general, all modes of transport are used to provide their services by different types of vehicles and infrastructure, so it can be varied from one transport mode to another, e.g. aircraft with terminals and routes for air, vessels, and ports for sea transport, and different types of vehicles and roads for auto-mobile transport modes. For railways, there is a unique combination of railway services and network infrastructure. If this combination fails in terms of efficiency, it leads to losses in the provision of freight and passenger services for society through the underutilization of resources. Lack of efficient railway transport leads to higher transportation costs and lower producers' profits or higher consumers' expenditures. Costs play an important role in determining the price of products and services. The problem of establishing economically justified costs for specific railway services arises when providing services for freight and passenger rail transport. In this study, we are interested in investigating the cost characteristics and cost performance of UZ to determine whether scale, scope, and density economies[1] exist in Ukraine's railways. Investigating different economies of scale allows for a discussion of efficiency UZ in terms of optimal structure scale and density. At the present there is a hot debate on types of railways models that UZ should be incorporated – European type model (vertical separation) or American type model (horizontal separation). The aim of this paper is to provide a strong reason for which type of reform is better for the Ukrainian railways in order to draw some conclusions for assisting the Ukrainian authorities involved in railway

[1] Economies of scale are average costs decline (cost savings) with more output (traffic). Economies of scope are cost advantages (savings) in producing more outputs than a single one. Economies of density refer is a decline in average cost resulting from an expanded network of a given size.

industry reforming and regulating.

UZ is organized as a single highly integrated company with 6 regional branches: Lviv, Odesa, Near-Dnipro, Southern, Southwestern, Donetsk railway, and 33 functional branches. The main purpose of UZ (Ukrzaliznytsia Integrated report 2020) is to meet the needs of Ukraine in railway transportation and logistics, achieve and maintain a strong financial position of the company, efficient operation and development of railway transport, creating conditions for increasing the competitiveness of the transport industry.

UZ is on the way to ownership unbundling from "quasi-monopoly" to competitive company. UZ envisions the separation of its operations into four lines of business (UZ Integrated Report 2020): infrastructure, cargo, passenger, and maintenance. The unbundling is expected to bring UZ more in line with the EU-type model (vertical separation). But existing economies of scope (vertical integration) looks pointing against vertical split reform. Because of a large share of fixed cost in UZ assumes the presence of substantial economies of scope or vertical integration. Accordingly, a tangible benefit of the Horizontal separation model is that it saves vertical integration of the operator operating both trains and tracks, and at the same time requires the vertically integrated company to provide access to competitors and thus create competition at the transportation stage. Similarly, the disadvantage of the model is that the integrated the company will be interested in providing more favorable access conditions - lower prices and simple procedural conditions - for its trains compared to independent rolling stock operators. Vertical separation (VC) allows competition between rolling stock operators through open access to infrastructure that maintains a monopoly position. The downside of VC is that it, by definition, destroys the economies of scope that both the infrastructure and rolling stock operator use at the same time. Although VC suffers from conflicts of interest and possible abuse of market power as well as small incentive to invest.

To date, a few scholars (Pittman, R. 2020, Sinclair, J. 2020) tried to study UZ restructuring, however, most papers can be regarded as descriptive: the authors analyze which mode of railways European or American type better suits UZ. There is no empirical paper that analyses the statistical relationship between the cost performance of UZ and optimal structure (model). Thus, it implies the objective of the thesis - to test empirically which type of model is better for Ukrainian railways. My work will be the first one that provides a systematic analysis of cost performance and the optimal structure of UZ. By the term "cost performance" we mean examining costs. In other words, how costs vary with output, the size of the network, and different types of outputs produced. By studying costs, we can investigate the nature of the production technology of firms. Production technology refers us to the question of whether or not output is produced with a constant return to scale. We are primarily concerned to measure at which extent UZ is working at economies of scale where economies of scale and scope are increasing.

For developing countries such as Ukraine, where a sound transportation system is one of the basic needs for efficient utilization of scarce resources. Railway system is an important element in the supply chain of Ukraine. UZ delivers raw materials, fuel and energy resources, components, equipment, etc. According to UZ data on the structure of the volume of freight transport for all types in 2020, most of the transported goods have a raw material origin, they account for 74% correspondingly. And another evidence that UZ serves mainly raw materials, is Figure 1 which illustrates "the points of cargo occurrence" indicating the nature of the goods sent within Ukraine in 2014 (without transit, which is less than 8% of the total volume of goods sent and has its own characteristics). The size of the circle corresponds to the volume of shipped goods in 2014. Black color is goods of raw origin, which include iron ore, coal, building materials (crushed stone, sand), grain, coke, wood. Red color represents goods with a higher added value, which can at least conditionally be called finished products - steel and cast iron, oil products, chemical and mineral fertilizers, cement, etc. (Vox Ukraine, 2015)

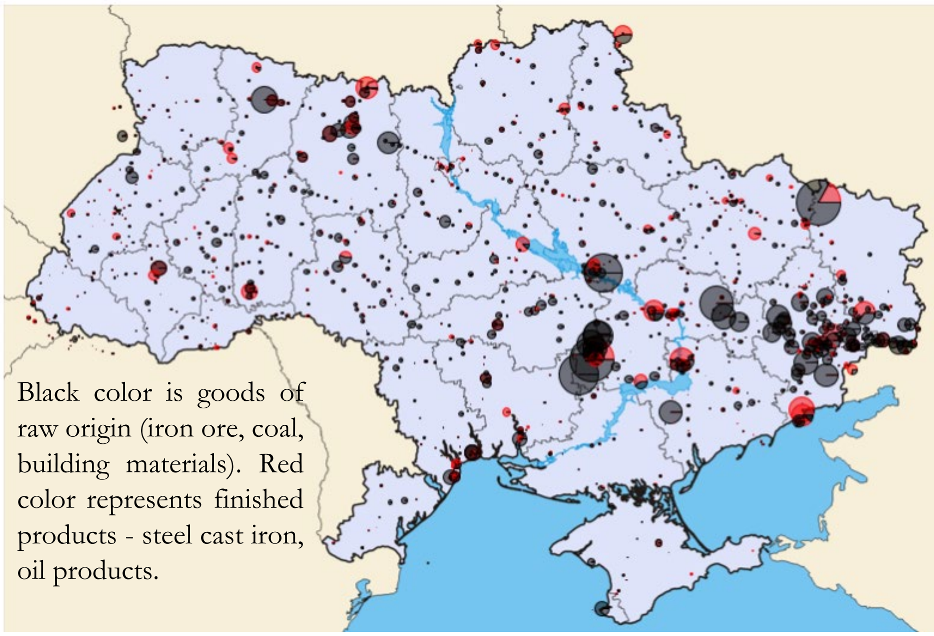

Figure 1.Cargo origin points

Source. Vox Ukraine.

The poor economic performance of UZ led to a decrease in the market share in favor of the other kinds of transport in both freight and passenger traffic. From Figure 1 we can observe that during the period 2015 - 2020 the UZ occupied the largest share from 61%-51%, meanwhile, this share is decreasing. In comparison road transport has shown significant growth over the past five years. For 2020 road transport has the highest share of 19% compared to previous years, while the railroad has 51%. Shares of other modes of transport were not shown significant changes. To summarize, we can say that if the situation doesn't change drastically there is a risk that in the future, a significant share of the railroad will be shifted to road transport. Although, the share decrease can be supported by the fact that UZ has lost "the trust" of domestic and foreign cargo owners (the transition of traditional UZ's railway cargo (transportation of agriculture and mining materials) to other modes of transport, implementation of private traction, outflow of qualified and loyal employees). By now the war has stopped this process, meanwhile, these problems will occur again after the war will end.

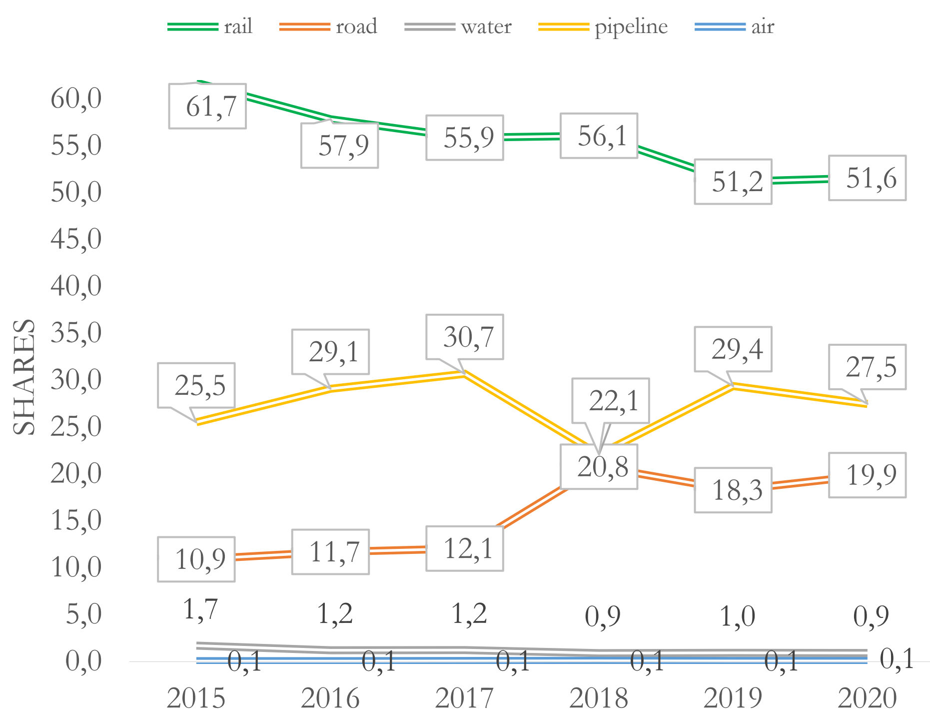

Figure 2. Freight turnover by types of transport, 2015-2021

Source. The author calculations based on UZ data

Another factor that influences of cost performance of UZ are the low utilization of the railroad network. From the Figure 3 we can observe that transportation volume has substantially declined during the last years, meanwhile the infrastructure volume has remained virtually the same around 20000 km. While the freight turnover in 2020 is around 80% of the 2005 volume and passenger turnover is only 20% of 2005 volume. This leads to important conclusion that low utilizations of rail network may lead to substantial increase of the fixed cost of railway transportation (IER, 2001). It means that rails infrastructure is an increasing burden for UZ. As the cost of idle infrastructure is included into railways tariffs, railways consumers have to pay for the railways nobody uses.

Figure 3. Railways utilization[2]

Source. The author calculations based on UZ data

The current state of the railway complex of Ukraine is deplorable. Depreciation

[2] Freight turnover is the product of a certain quantity of cargo (in tons) and the distance of the transport (in km). It is measured in ton-km. Passenger turnover is calculated by multiplying the number of passengers transported to transportation distance.

of infrastructure, locomotive and wagon fleet is 70-98%. The structure of UZ Wear and Tear of Fixed Assets are presented in Appendix A. The fleet of infrastructure, rolling stock has not been updated for years, it implies that UZ does not have enough funds for modernization and renovation of own assets. The consequence of funds imbalances is the systemic under fulfillment of capital investment plans: in 2019, the capital investment plan was only 57% fulfilled (10.4 out of the planned UAH 18.25 bln.), in 2018 – by 67% (16.9 out of the planned) UAH 25.1 bln. (Green book - "Railway freight transportation")

Signs of overcapacity are present also in rolling stock utilization. Figure 4 shows that average wagon productivity has declined together with the number of wagons and it implies the deficit of traction of UZ, which many railway company customers experienced already. Wagon productivity influence high level of wear, which does not allow the use of the entire fleet of wagons. At normal railways companies the relationship between wagon productivity and the number of wagons should be inverse. Which means that with higher productivity the number of wagons should also increase. This observation is supported by the fact that railcar turnover has decreased by up to 80% since 2005. This means that UZ is short of railcars..

These features imply inefficient allocation of resources and discourage cost transparency, reducing room for competition in the railway industry. Under these circumstances the measurement of the economic efficiency of railways becomes interesting and challenging, therefore it creates excellent opportunities to investigate the impacts and effects of current policy choices on economic efficiencies.

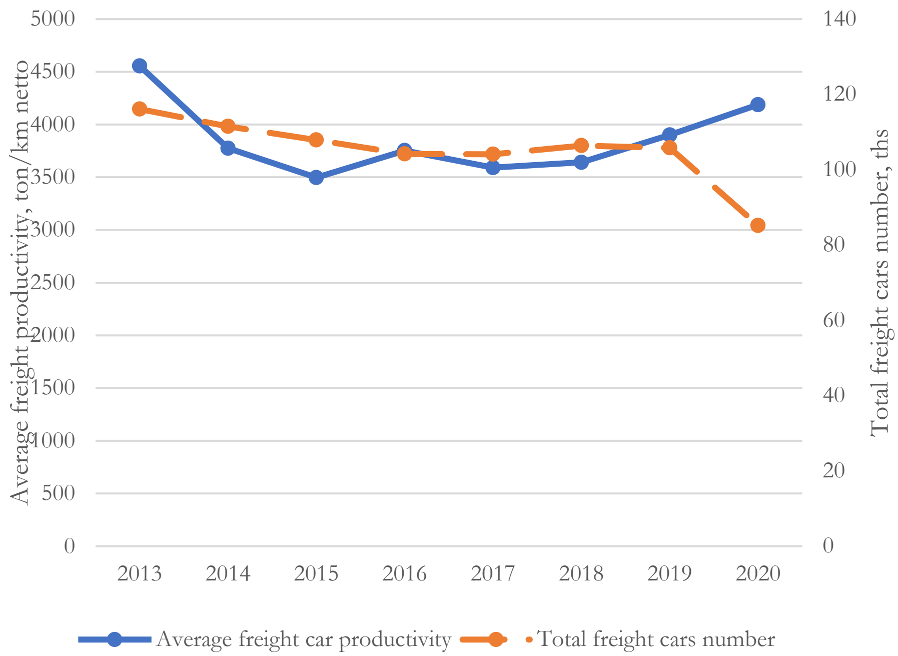

Figure 4. Rolling stock productivity[3]

Source. The author calculations based on UZ data

Over the past decade, there are a lot of research papers that investigate the state and characteristics of reforms in Ukrainian railways, but no one of them is based on cost performance. The unique contribution of this paper is the policy implication in terms of railroad cost analysis, economic development and estimating governance performance on management UZ.

This paper is organized as follows: Chapter 2 reviews the literature on an estimation of a translog cost function and stochastic frontier approach (SFA); Chapter 3 discusses the methodology of the analysis and model specifications, some theoretical concerns with SFA, specification of the regression model.

---

[3] Freight turnover is the product of a certain quantity of cargo (in tons) and the distance of the transport (in km). It is measured in ton-km.
Passenger turnover is calculated by multiplying the number of passengers transported during report period to the tariff distance of their transportation.

Chapter 4 provides data description. Chapter 5 presents the main empirical results. Chapter 6 concludes all key findings of the paper.

Chapter 2

LITERATURE REVIEW

This chapter discusses the overview of relevant literature for my thesis. Among recent studies on the measurement of economic efficiency of railways and different economies of scale, a substantial number of researchers thrives on understanding how to improve efficiency, increase freight and passenger traffic and reduce subsidies. Firstly, I discuss the estimation of the railway cost function, and secondly, I discuss the stochastic frontier approach (SFA).

The best way of assessment of different economies of scale is to estimate translog cost function. Estimation of translog railway cost function has been a popular topic in the railroad economics literature for the past two decades. The main finding of these studies is that railroad companies demonstrated returns to density. But the results are quite diverse, especially if talk about government ownership vs private railroads.

Bitzan and Wilson (2007) published a very detailed paper, where they estimated a hedonic cost function for multi-product railroad firms. On the basis of previous cases, authors estimated costs for specific types of railroad services, where shipment attributes are allowed to vary. Bitzan and Wilson measured railroad costs as a translog function of input prices, technological characteristics and ton-miles of unit train services, and way and through train services. However, in this kind of analysis, researchers face main difficulty is multiplicity of railways outputs. The author innovation is to considered differences in the characteristic of different outputs (shipment size and length of haul) and their impacts on costs. While in previous studies specification of output is varied: car-miles in bulk, general traffic as output, unit train ton-km, way train ton-miles. Representing traffic of passenger and freight as a single unit of measurement may lead to biased efficiency estimates. In their approach closely follows Chiang and Friedlander's (1984) study, they used multiple outputs that were adjusted by shipment characteristic. Although, the author's findings suggest significant economies of density. Specifically, they find a large difference in the elasticities of cost with respect to different outputs and substantial impacts of the hedonic variables on marginal costs. They illustrated how the cost estimated could be varied with the characteristic of the individual shipments.

Filippini and Maggi (1993) studied the efficiency of the Swiss Private Railways through regulatory and economic context. To achieve their goal, they investigated the overall and scale efficiency of 48 Swiss private railway companies. To measure the scale and density economies authors estimated a translog cost function for four years of panel data by adding compound indicators such as network size and structure. The authors provide evidence for a discussion of efficiency in terms of optimal scale and density. The researchers reported results that most of the Swiss private railway companies conduct at an improperly low scale and density. Another interesting aspect of Filippini and Maggi the companies are homogenous in terms of overall cost efficiency it is due to the influence of regulation in terms of the subsidy structure.

Bereskin (2009) developed an econometric cost model using publicly available data and the methodology developed for studying average and marginal costs in the railroad industry. The authors structured his model into individual elasticity estimates for operating parameters to test economies of scope.

The study by Bitzan and Karanki, (2021) published a very detailed report about the current cost structure of the U.S. railroad industry and differential pricing. They examined the extent of economies of scale in the railroad industry and its implications for differential pricing. They also provided a non-technical explanation of cost concepts and discussed the role of cost in the pricing decision of firms. And finally, they estimated the generalized cost function by using the translog functional form. Their findings conclude that the extent of economies of scale (density) has slightly decreased over time, but still, the U.S. railroad industry examines substantial economies of scale. This persistence of scale economies suggests differential pricing for the railroad industry, thus the policymakers should be cautious in implementing policies that limit differential pricing in the industry.

Daniel, Pels, and Rietveld (2010) analyzed the rail cost function of the Dutch National Railway Company (NS) to investigate the relationship of the effect of firm size on cost. In the study, the researcher used a translog variable specification to estimate the cost function for NS. The reported results indicate returns to density economies parameters, so they cannot reject the hypothesis of constant density economies. It provides evidence that "size matters" and E.U. policymakers should stimulate completion on track.

Researchers started to apply stochastic frontier analysis to the measurement of rail performance. Among them:

Cesar Rivera-Trujillo (2005) investigated the technical efficiency performance of the rail sector and particularly the measurement of performance at the international level. The researcher presented a review of the main technical efficiency-measurement methods and made a comparison among them. The author emphasized the distance function method because it allows multi-output considerations, separates inefficiency from stochastic noise, allows hypothesis testing, and does not require price data or assume a specific behavior (e.g. cost minimization). The stochastic frontier methods are most appropriate for the measurement of international railways in terms of errors and exogenous factors such as statistical noise, bad luck, strikes, etc., these factors are expected to play a significant role in the measurement of technical performance. The reported results indicate that the U.S. and Canada are the most technically efficient railways, due to the partial productivity indicators, and less efficient are Brazil, Mexico, and Chile. Also, the author provides evidence that contributes to productivity due to technical change and found a positive and significant and significant change in technical performance has resulted in privatization in Canada and Latin America.

Coelli and Perelman (2000) measured and compared the performance of European railways. For estimation, the multioutput distance function researcher used corrected ordinary least squares (COLS). Also, the authors have compared results with single-output production functions. The reported results indicate a substantial difference in parameter estimates and technical efficiency rankings, which undermines the reliability of single-output models. Coelli and Perelman provided evidence the technical efficiency of European railways differs substantially from country to country. Another interesting aspect of the author's study is a significant improvement in the performance of European railways during 1980 and the performance was mainly driven by substantial cuts in labor usage and rolling stock. It is a consequence of stricter government budgetary restrictions.

Cantos and Maudos (2001) estimated a cost function on the basis of a stochastic frontier function. The authors defined impropriety as the loss of revenue by companies compared with the maximum levels that they could achieve. The aim of their study is to determine whether the losses of operating revenue for companies were important. In other words, they tested inefficiency on the revenue side, particularly whether or not companies operated at the efficiency frontier in order to maximize their levels of revenue given the productive structure of the companies. The authors also analyzed the hypothesis that there is a significant correlation between cost and revenue efficiency. The reported results suggest a negative correlation between cost efficiency and revenue efficiency. Indicating that that concentrating only on the cost side gives only a partial view of the problem.

Cantos and Villarrolla (2000) in their study authors estimated the levels of productivity, efficiency, and technical change for European railway companies. The study is based on an analysis of the stochastic frontier cost function. The reported result shows that the source of productivity growth is technical progress, followed by gains from efficiency. And the most efficient railroads are those with a higher degree of financial and management impendence.

To our knowledge, in Ukraine, no empirical study has explicitly investigated economies of scale, scope, and density.

C h a p t e r 3

METHODOLOGY

The research aims to answer question whether UZ of whether economies of scale, scope, and density in UZ exist. For this reason, we want to estimate translog cost function and derive from its coefficients that will help to determine different economies of scale. We want to test hypothesis that integrated company UZ realizes economies of scale, scope, and density and produces railway services with a higher level of efficiency against the alternative hypothesis that UZ has large operational network thus it may experience diseconomies.

3.1 Estimation of the translog cost function

In examining the economies of scale, scope, and density of the UZ, it is important to distinguish between three different types of economies that may exist. Economies of scale refer us to average costs decline (cost savings) with more output (traffic). Economies of scope refer us to cost advantage (savings) in producing more outputs than a single one. Economies of density refer us to a decline in average cost resulting from an expanded network of a given size e.g. economies serving larger markets.

This section describes the methodology used to estimate the translog cost function. We are now in a position to examine UZ's overall cost structure. The sense of total cost function that it provides information about the output-cost relationship.

The total cost function shows that, for any set of input costs (w – vector of input prices) and given output level (y – vector of output), and the given technology (t

– vector of technological attributes), the minimum total cost incurred by UZ is

$$C = C\,(w, y, t) \qquad (3.1)$$

The cost function is twice continuously differentiable C^2, non-decreasing in input prices and output $f'(C) > 0$, by assumption, and concave in input prices and homogenous of degree one which is a general feature of all cost function

Vector input prices include four-factor prices: the price of labor, the price of capital (rail cars and locomotives), the price of fuel, and the price of materials. As costs are expected to vary, not only with the output size, we should add variables of technology characteristics in the equation, in order to find differences in the nature of services provided such as network length and time trend.

For estimation of the cost function, we must employ a specific functional form of the cost function. This form can be interpreted as a second-order approximation to the twice-differentiable cost function. There are several functional forms that satisfy this requirement – the generalized Leontief, generalized Cobb-Douglas, and translog cost function. In our estimation, we choose the translog functional form, because the translog function is quite flexible for almost all cases implying a set of derived demand equations and accommodating various cost function flexibility. Derived demand equations are linear in the parameters and represent a very general cost structure. In our study, we follow the Filippini and Maggi approach (1993) of estimation translog cost function. It is the standard approach of estimation translog costs function, where total operating costs is the dependent variable and independent variables are four input prices and attributable characteristic of railways.

For our four-input model, we write the cost function with symmetry and returns to scale imposed as

$$\ln C = a_0 + \sum_i a_i \ln w_i + a_y \ln y + a_N \ln N + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij} \ln w_i \ln w_j$$
$$+ \frac{1}{2}\sum\nolimits_i a_{yy} (\ln y)^2 \qquad (3.2)$$
$$+ \frac{1}{2}\sum\nolimits_i a_{NN} (\ln N)^2 + \frac{1}{2}\sum\nolimits_i a_{Ni} \ln\ w_i \ln N + \frac{1}{2}\sum\nolimits_i a_{iy} \ln\ w_i \ln y \ + \epsilon$$

Where $\ln C$ is the Total cost of UZ, w_i is a vector of prices, y is output (Ton-km), N is the Network length and ϵ is a disturbance term. According to the cost minimization assumption, we can apply Shephard's lemma to find the shares of the inputs in the total cost of producing Y. Which is elasticity with respect to factor price or input demand equations.

$$x_i = \frac{dC}{dw_i} \qquad (3.3)$$

In the form of a natural logarithm, we can rewrite the share of the total cost in the form.

$$s_i = \frac{d\ \ln C}{d\ \ln w_i} = a_i + \sum_j a_{ij} \ln w_j + \sum_m y_m \ln y_m + \delta \ln N$$
$$+ \ \epsilon_i \qquad (3.4)$$

For the translog function, we require the following restriction:

$$\sum_i a_i = 1, \sum_i a_{ij} = \sum_i a_{Ni} = 0 \ , a_{ij} = a_{ji} \qquad (3.5)$$

Under these restrictions, the translog cost function is homogeneous of degree one in factor prices. This implies that by the symmetry of cross-term prices the sum of them is equal to zero. Another implication that we can derive is that a doubling in factor prices leads to a precisely double in costs of producing any given output level.

As has become standard practice, in our estimation procedure we have chosen Zellner's (1962) seemingly unrelated regression (SUR) technique for simultaneous estimation of the cost and share equations. The equations are estimated as a system since individual errors will be correlated across the equations because prices can be not completely exogenous and can correlate with individual errors. From this, it appears that we can estimate cost function by OLS and our estimates will be unbiased, but no longer BLUE, and variance-covariance are no longer diagonal because of cross equations' errors and inefficiency.

Economies of density are calculated as inverse elasticity with respect to output. We can define ED_C as a proportional growth in total cost resulting from a proportional growth in output, fixing at the constant level all input prices.

$$ED_C = \frac{1}{\frac{d\, In\, C}{d\, In\, y}} \qquad (3.6)$$

If ED_C is greater than 1 we identify economies of density and vice versa if we identified that ED_C is below 1.

Economics of scale is calculated as inverse elasticity with respect to output plus elasticity with respect to network size. We can define ES_C as a proportional growth in total cost resulting from a proportional growth in output plus the indicator of network size and structure, fixing at the constant level all other factors.

$$ES_C = \frac{1}{\frac{d\,In\,C}{d\,In\,y} + \frac{d\,In\,C}{d\,In\,t(N)}} \qquad (3.7)$$

If ES_C is greater than 1 we identify economies of scale and vice versa if we identified that ES_C is below 1 we will speak about diseconomies.

$$EScale_C = 1 - \frac{1}{\frac{d\,In\,C}{d\,In\,y}} \qquad (3.8)$$

Stochastic Frontier Approach (SFA) seems to be the most popular in the field of railroad research. If we functional form we are supposed to get a precise result, but in fact, distributional assumptions to be made in most cases are somewhat arbitrary. (Cantos 2001).

Efficiency is a comparison between observed (real) output and inputs with optimal values of inputs and outputs used in particular production technology. Technical efficiency refers us to production at the production frontier. In other words, it is the maximum level of output that can be achieved given the level of input. For indicating inefficiency, we refer to deviations from the cost frontier. Meaning that the units on the frontier are efficient, whereas belonging to the area above the frontier are inefficient (Figure 5). And the area below the frontier is not appropriate, since the most cost-efficient unit is located on the frontier.

In the stochastic model of the frontier cost function, it is assumed that any deviation of the observed cost from theoretical cost function is random component such as measurement errors, disturbances and inefficiencies. In order to capture this effect, we can add in the frontier function statistical noise – stochastic elements outside the control of the producer. We assume that they are independent and identically distributed.

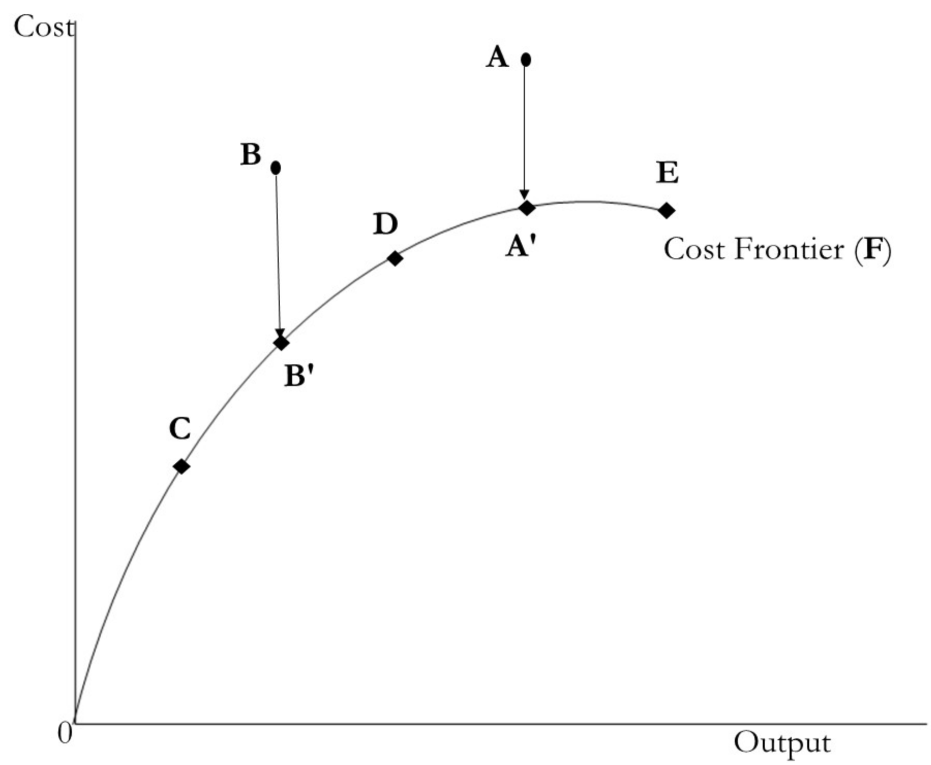

Figure 5. Cost frontier and efficiency

Stochastic frontier is estimated by using the translog function. We use the same specification as in previous case plus cost inefficiency u and statistical noise v. Also, we assume that UZ has more control over the inputs rather than the output produced, so we use an input orientation approach.

The general specification of a stochastic cost frontier is:

$$\ln C = a_0 + \sum_i a_i \ln w_i + a_y \ln y + a_N \ln N + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij} \ln w_i \ln w_j$$
$$+ \frac{1}{2}\sum\nolimits_i a_{yy} (\ln y)^2$$
$$+ \frac{1}{2}\sum\nolimits_i a_{NN} (\ln N)^2 + \frac{1}{2}\sum\nolimits_i a_{Ni} \ln\ w_i \ln N + \frac{1}{2}\sum\nolimits_i a_{iy} \ln\ w_i \ln y + u\,\imath$$
$$+ v \tag{3.9}$$

where $u \geq 0$ accounts for cost ineffciency and v accounts for statistical noise. This model can be rewritten as:

$$c = c(w, y)\, e^u e^v \tag{3.10}$$

The cost effciency according to Shepard is:

$$CE = \frac{c}{c(w,y)\, e^v} = \frac{f(x)\, e^u e^v}{c(w,y)\, e^v} = e^u \tag{3.11}$$

Chapter 4

DATA CONSTRUCTION AND DESCRIPTION

One of the main problems that researchers in this field face are the lack of complete and reliable rail data. Another problem is the difficulty of identifying the inputs and outputs used in the cost function. Once you have identified the variables of interest it can be very hard to find data on the prices faced by railways. Various authors have proposed different bundles of inputs and outputs, and provided different ways of estimation their prices. In case of railways, we don't observe the prices of inputs, but we observe how operating costs are split up between different categories.

In this chapter, we present our data. Estimation of a translog cost function requires data from UZ's branches. In order to examine the extent of economies of scale, scope, and density in UZ, this study uses UZ financial data to estimate a cost function. We need to know the prices of the inputs used, total cost, output and the cost shares of the inputs used. The main sources for our analysis are the financial statements of the joint stock company "Ukrainian railways" and their corresponding six regional branches (Donetsk railway, Pridniprovska_railway, Southern_railways, South-Western railway, Odessa_railway, and Lviv_railway). The financial statements were prepared by the UZ representative. Financial statements contain a wide range of accounting data, including total costs, revenue and a breakdown of spending. While useful, this data does not include information on the prices of the inputs used in providing services by UZ, and so we have to create proxy prices based on what information is available. We used data from these financial statements in order to calculate input prices on a particular branch of the railway.

For estimation, panel data for six years (2015, 2016, 2017, 2018, 2019, and 2020) has been used. We have 144 observations (6 branches × 4 quarters × 6 years).

As it becomes standard in many studies of estimating railways cost function, the dependent variable is Total Operating Costs which are taken as the total of operating expenditure of the UZ. The output variable is measured in ton-kilometers.

We construct Labor Price per Hour for each branch of UZ as (Total Salary and Wages + Fringe Benefits - Labor Portion of Cap. Exp. Class. as Operating) / Labor Hours. Then we construct Materials price as total expenditures on materials divided over the total number of train kilometers. We compute price of the fuel as fuel expenses divided by the number of fuels used. Finally, we form the price of capital is calculated as the summation of two parts: the first part is Annual Depreciation and Capital expenditure per car and locomotive and the second part is Annual Deprecation and Capital expenditure per the length of network size.

The cost function also includes technological attributes which account for the differences in network structure. As an indicator of attributable characteristics, we use two different indicators. The first traditional indicator is the size of the network. Network size is the length of the railways and is measured in kilometers. The second indicator is the time trend.

Table 4.1 shows descriptive statistics of the variables included in the translog cost function

Table 4.1. Descriptive statistics of independent and dependent variables

	N	Mean	Std.Dev	Min	Median	Max
Network size	144	3731.83	705.20	2927.00	3637.50	4668.00
Price of Capital	144	4701.13	31851.76	2.87	1044.70	381239.05
Price of Fuel	144	26.56	8.00	14.62	25.22	62.22
Price of Labor	144	78.25	26.02	11.53	77.97	119.36
Price of Material	144	18.30	8.09	2.98	17.20	40.22
Share of fuel	144	0.49	0.17	0.07	0.47	0.87
Share of capital	144	0.34	0.17	0.03	0.34	0.79
Share of labor	144	0.95	0.29	0.10	0.91	1.54
Share of materials	144	0.49	0.26	-0.01	0.43	1.47
TC	144	2651804.94	731821.69	302122.00	2752164.50	4463755.00
Ton-km	144	12316.62	6530.94	3136.02	10032.35	27410.18

Also, to investigate how our data is correlated we use the additional tool such as complot, you can see the results below.

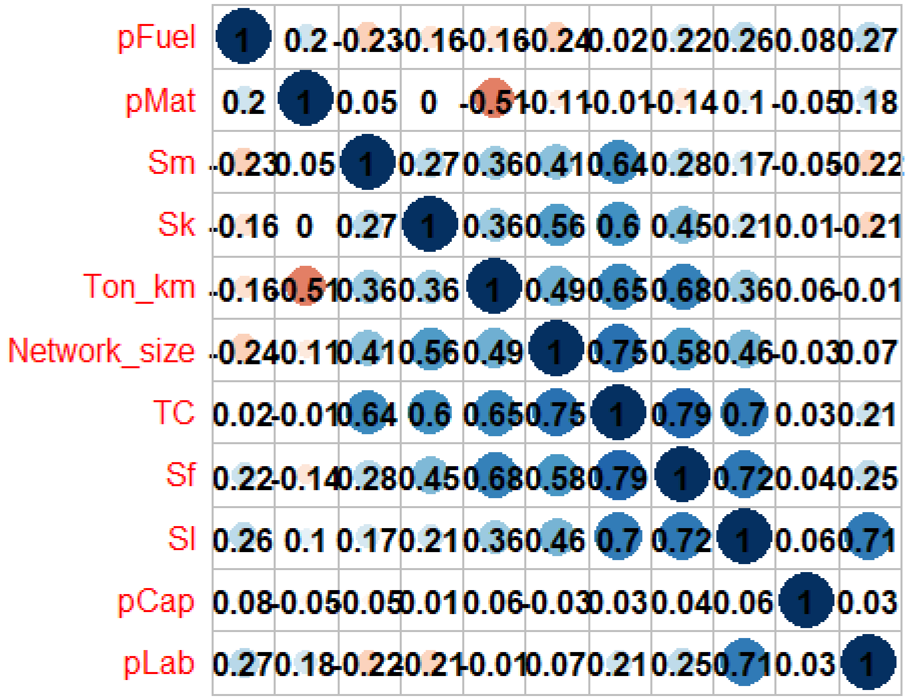

Figure 6. Correlation plot of variables of interest

Source: Author calculation

From Figure 5 we can observe that variables Average length of haul and Ton-km have a correlation of 0.87 which is high and could lead to biased estimates. So, in our analysis, we drop the Average length of haul variable.

Chapter 5

EMPIRICAL RESULTS

This chapter describes the estimation results of two main models stated in Chapter 3. We will proceed in 2 steps. First, showing the result for Translog and and examine whether the economies of scope, scale, and density exist., and on the second step is estimating the stochastic cost frontier, and showing the results of performances of UZ, using SFA approach.

5.1 Estimation results of the translog cost function

We firstly estimate through standard OLS (1) with homogeneous imposed for the translog cost function. Secondly, we obtain the cost of shares. Thirdly, as has become standard practice, the cost function and share equations are estimated simultaneously using Zellner's (1962) Seemingly Unrelated Regression (SUR) technique. Then we calculate economies of scale, scope and density[4].

Estimation results and different economies of scale are presented in Table 5.1. This table shows only the first-order coefficients. You can refer to the complete regression result in Table A1 of Appendix A. Since total cost and other dependent variables are in natural logarithms and are being normalized, we can interpret the first-order coefficients as cost elasticizes. We can dismiss the

[4] A short reminder that economies of density are a proportional increase in total cost resulting from a proportional increase in output (Ton-km), holding other factors and characteristics fixed. Or in other words, economies of density are equivalent to the inverse of elasticity of total cost with respect to output. And if economies of density/scope are greater than 1 we identify economies of density/scope otherwise we observe no economies or diseconomies of density/scope exist. In another case when economies of density/scale are equal to 1 then we talk that no economies or diseconomies of density exist. One of the indicators that economies of density exist is if the average costs of UZ decrease as output increases through increasing traffic/trains on the existing tracks. Economies of scale don't exist if average costs are fixed when UZ crated additional output without changing the traffic.

interpretation of second-order coefficients as cost elasticities at the median point.

The results are satisfying as all coefficients are significant and carry the expected sign. At the median, the cost elasticities with respect to input prices are equivalent to the cost shares. The r-squaed is also satisfying with values around 0.9. The capital accounts of 3.2% of the railway costs while material accounts 26.6%., fuel accounts 36% and labor cost for remaining 65.8%. Interestingly enough, that figures correspond to financial statement of UZ we can see similar result regarding operating cost of shares. Labor costs are 57%, depreciation is 17%, fuel and energy accounts for 21% and cost of materials accounts only 4%.

As could be expected, the influence of the network length on cost is positive. The cost-elasticity is 63.2 in this case. The result presented in table 5.1 reveals that all the values of indicators for economies of scale and scale are great than 1, while economies of density below 1. Which means that economies on UZ exist and we fail to reject the null hypothesis that UZ operates under economies of scale. Economies of density below one also suggests that UZ is not highly density as could be expected.

Table 5.1. Estimation results for the total cost function-first-order coefficients and indicators for economies of scale

Variable	Coefficient	Variable	Coefficient
Intercept	188.259*** (23.886)	$\ln pFuel$	0.360*** (0.068)
$\ln pCap$	0.032* (0.013)	$\ln Y$	0.219** (0.067)
$\ln pMat$	0.266*** (0.046)	$\ln N$	0.636*** (0.160)
Observations	144		
Economies of density	0.396		
Economies of scope	1.54		
Economies of scale	1.55		
Note: *p <0.1; **p <0.05; ***p <0.01			

5.2 Estimation SFA for UZ

In this research we used the one-stage stochastic frontier approach to estimated cost efficiency scores and factors of cost inefficiency for 6 branches of UZ during 2015-2020. We apply standard SFA function in R to estimate the parameters of the translog cost frontier function. The results of the stochastic frontier estimation are shown in Tables 5.2 and 5.3.

Table 5.2 summarizes the estimation results obtained for the stochastic frontier approach. The coefficients of material and output (Ton-km) are statistically significant, 2.836 and 2.535, respectively. The positive signs indicate that an increase in output will lead to an increase in the total cost.

Gamma ratio that shows the variability for U and V can be used to estimate the relative inefficiency in a branch. This is an estimate of the amount of variation that occurs due to inefficiency compared to the sample noise. The values of gamma is 0.827 and significant at 1% level.

Table 5.3 shows the efficiency scores measure. The mean efficiency is 88.3 %. This value indicates that, to operate efficiently, branches could only reduce their input costs by 11.6 % without decreasing their outputs. In this study, the regional branches outputs are defined Ton-km. The score of the maximum branches efficiency is 91.36 % while the minimum efficiency score is 85.3%. The median efficiency is 88.32 % and the standard derivation is 1.97 %.

It is observed that mean efficiency doesn't vary a lot in 6 regional branches. We found that the mean efficiency is relatively high for all six branches because branches are similar in size and structure and also operate in one environment and legislation.

Table 5.2. SFA estimation results

VARIABLE	LOG(TC)	S.E.
LOG (PCAP)	0.210	(0.551)
LOG(PMAT)	2.836	(1.413)
LOG(PFUEL)	0.262***	(1.682)
LOG(TON_KM)	2.535***	(1.039)
LOG(NETWORK_SIZE)	30.072	(17.687)
TIME INDEX	-0.112*	(0.010)
0.5*LOG(PCAP)^2	-0.018	(0.069)
0.5*LOG(PMAT)^2	-0.038	(0.052)
0.5*LOG(PFUEL)^2	0.507	(0.127)
LOG(NETWORK_SIZE) ^2	-3.609	(2.168)
0.5*LOG(TON_KM) ^2	-0.282	(0.123)
$\ln(pCap) * \ln(pMat)$	0.078	(0.026)
$\ln(pCap) * \ln(pFuel)$	0.090	(0.026)
$\ln(pCap) * \ln(TON\_KM)$	0.014	(0.036)
$\ln(pCap) * \ln(NETWORK\_SIZE)$	-0.003	(0.089)
$\ln(pMat) * \ln(pFuel)$	-0.109	(0.105)
$\ln(pMat) * \ln(TON\_KM)$	-0.123	(0.086)
$\ln(pMat) * \ln(NETWORK\_SIZE)$	-0.233	(0.203)
$\ln(pFuel) * \ln(TON\_KM)$	-0.073	(0.097)

Table 5.2 – continued

$\ln(pFuel) * \ln(NETWORK\_SIZE)$	0.086	(0.256)
SIGMASQ	0.032***	(0.007)
GAMMA	0.827***	(0.006)
CONSTANT	99.747***	(76.265)
OBSERVATIONS	144	

Table 5.3. Estimation of mean efficiency

	Mean efficiency	Ranking
DONETSKA RAILWAY	0.8802611	4
PRIDNIPROVSKA RAILWAY	0.9130099	1
SOUTHERN RAILWAYS	0.852751	6
SOUTH WESTERN RAILWAY	0.8922895	2
ODESSA RAILWAY	0.8862124	3
LVIV RAILWAY	0.8779341	5

Chapter 6

CONCLUSIONS

Ongoing reform of restricting Ukrainian railways attracted attention of many reasecherchers, government representatives, railroad management as well as cargo carriers. They are particularly interested in consequences of the reform and how it will change the industry. In our study we want to contribute to current debates which type of railways models is better to UZ - European separation or Horizontal separation, by assessing economies of scale, scope and density of UZ. By assessing different economies, we can answer the question whether UZ realizes economies of scope or vertical integration. It is important to understand it in terms of future optimal structure of UZ.

A number of theoretical studies done in the field of comparison advantages and disadvantages of both types of models. My paper is the first work in the field, which provides analytical analysis of assessing different economies of scale. Following our expectations, Ukrainian railways were found to operate under economies of scale and scope, while not operating under economies. This evidence supported by our empirical results and the hypothesis that we tested. Results indicate that Ukrainian government should provide restructuration reform in way of Horizontal separation type model, rather than European type model in order to save economies of scope between operating units.

WORKS CITED

Bereskin, C. Gregory, 2009. "Railroad Economies of Scale, Scope, and Density Revisited," *Journal of the Transportation Research Forum, Transportation Research Forum*, vol. 48(2), pages 1-17. Available at: DOI: 10.22004/ag.econ.207130

Bitzan J. and Karanki F., 2021. "A Look at Railroad Costs, Scale Economies, and Differential Pricing". *U.S. Department of Agriculture, Agricultural Marketing Service*. Available at: https://www.ndsu.edu/fileadmin/challeyinstitute/Research_Briefs/USDA_Report-Railroad_Rates.pdf

Cantos, Pedro & Maudos, Joaqui-n, 2001. "Regulation and efficiency: the case of European railways," *Transportation Research Part A: Policy and Practice*, Elsevier, vol. 35(5), pages 459-472, June.

Cantos, Pedro & Villarrolla,2000 "Efficiency, technical change and productivity in the European rail sector: a stochastic frontier approach" International Journal of Transport Economics, Vol. XXVII, No. 1, pp. 55-76.

Coelli, Timothy and Perelman, Sergio, 2000, Technical efficiency of European railways: a distance function approach, *Applied Economics*, 32, issue 15, p. 1967-1976

Daniel, V. E., Pels, E., & Rietveld, P. 2010. Returns to density in operations of the Netherlands railways. *Rivista internazionale di economia dei trasporti*, XXXVII (2), p. 169-19

Filippini, M., Maggi, R. Efficiency and regulation in the case of the Swiss Private Railways. *J Regul Econ* **5,** p. 199–216 (1993). Available at https://doi.org/10.1007/BF01065366

Pittman, R. 2017. Reforming and restructuring Ukrzaliznytsia: a crucial task for Ukrainian reformers. *Science and Transport Progress*, (1(67), 34–50. https://doi.org/10.15802/stp2017/92775

Rivera-Trujillo, Cesar. 2005. Measuring technical efficiency in North and South American railways using a stochastic frontier model: an international comparison. p. 423-441. DOI:10.1016/B978-008044580-9/50117-3

R.W. Shephard. 1953. Cost and Production Functions, Princeton, N.J., *Princeton University Press*, vii p. 104.

Yelagin, Y., Brilyov D., and Zubarev. L. 2018, Reformation of Ukrzaliznytsya:

camp and features. *Bulletin of Economics, Transport, and Industry* 63: 75-82

APPENDIX A

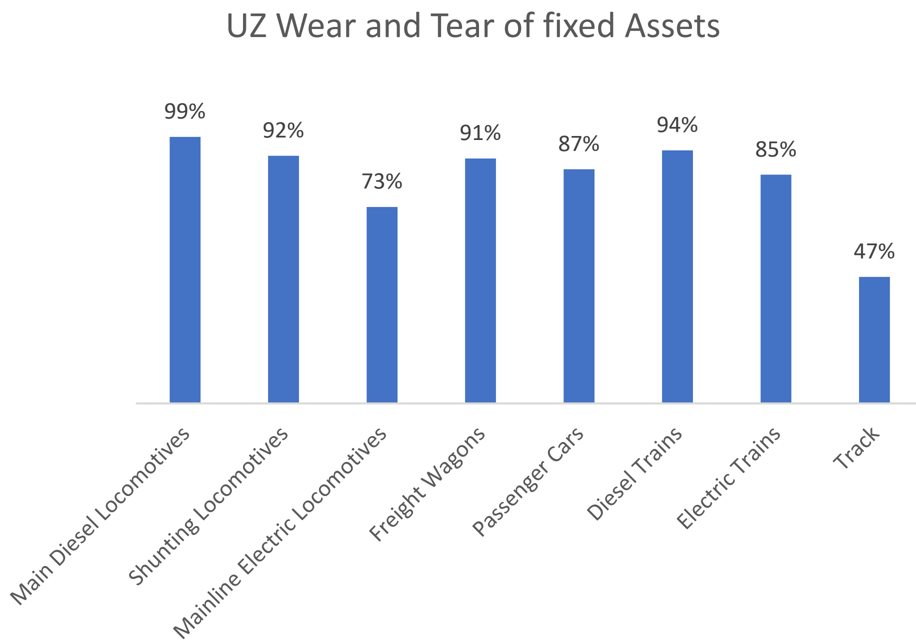

Figure A.1 UZ Wear and Teat of Fixed Assets.

APPENDIX B: TABLES

B1 Estimation Results for the Total Cost Function by SUR

Variable	Coefficient	Variable	Coefficient
Total cost equation (1)			
Intercept	-3.122***	$\ln pFuel$	4.734*
	(6.379)		(0.026)
$\ln pCap$	5.640**	$\ln Y$	7.739**
	(0.002)		(0.001)
$\ln pLab$	8.247**	$\ln N$	3.731
	(0.004)		(0.052)
$\ln pMat$	-1.061		
	(0.671)		
Quadratic Terms			
$\frac{1}{2}\ln(pLab)^2$	-4.8534	$\frac{1}{2}\ln(pFuel)^2$	-1.181
	(-1.234)		(-0.047)
$\frac{1}{2}\ln(pCap)^2$	-2.278	$\frac{1}{2}\ln(Y)^2$	-3.211**
	(1.845)		(-3.267)
$\frac{1}{2}\ln(pMat)^2$	1.166	$\frac{1}{2}\ln(N)^2$	-4.063
	(1.147)		(-1.802)
Cross-Interaction Terms			
$\ln(pCap) * \ln(pLab)$	-1.443*	$\ln(pLab) * \ln(pFuel)$	-5.679***
	(-2.096)		(0.345)
$\ln(pCap) * \ln(pMat)$	-7.434	$\ln(pLab) * \ln(Y)$	7.601
	(-1.296)		(-3.457)
$\ln(pCap) * \ln(pFuel)$	-8.690	$\ln(pLab) * \ln(N)$	-1.416
	(-0. 966)		(0.758)
$\ln(pCap) * \ln(Y)$	-1.111	$\ln(pMat) * \ln(pFuel)$	1.319
	(-1.69980)		(1.037)
$\ln(pCap) * \ln(N)$	3.008	$\ln(pMat) * \ln(Y)$	5.021
	(1.294)		(0.545)
$\ln(pLab) * \ln(pMat)$	3.641	$\ln(pMat) * \ln(N)$	-2.35
	(0.232)		(-0.978)
Capital share equation (2)			
Intercept	2.218***	$\ln pFuel$	1.192***
	(2.220)		(2.220)
$\ln pLab$	-1.102***	$\ln Y$	-7.292
	(2.220)		(0.086)
$\ln pMat$	1.993***	$\ln N$	1.173

	(1.906)		(0.296)

Table A1. - continued

Labor share equation (3)			
Intercept	2.871***	$\ln pFuel$	1.141***
	(2.220)		(2.220)
$\ln pCap$	-1.102***	$\ln Y$	-6.701***
	(2.220)		(2.220)
$\ln pMat$	-3.331***	$\ln N$	4.041***
	(2.220)		(2.220)
Material share equation (4)			
Intercept	-7.520***	$\ln pFuel$	-1.029***
	(2.220)		(2.220)
$\ln pCap$	1.192***	$\ln Y$	1.574***
	(2.220)		(2.220)
$\ln pLab$	-3.331***	$\ln N$	-1.347***
	(2.220)		(0.772)
Note: *p <0.1; **p <0.05; ***p <0.01			