Stochastic Frontier Estimation of the Technical Efficiency of Seaports in the United States

Neela D. Manage
Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431

Abstract
The objective of this paper is to identify factors that determine efficiency of container terminals at major U.S. seaports using stochastic frontier estimation. Growth in global maritime trade and the recent Panama Canal expansion project have resulted in important infrastructure upgrades at U.S. seaports that handle containerized cargo. This study plans to assess how these changes in their capital structure and the resulting increase in capacity had an impact on the productivity and efficiency of U.S. seaports. Seaports have an important impact on local businesses and job creation and are thus critically linked to the economic development of metropolitan areas where they are located. The empirical analysis utilizes panel data for the top twenty-five container seaports in the United States during the period 2015 to 2018. The data provide detailed measures of port throughput (i.e., the amount of cargo a port handles in a given year) and a variety of factors that influence port capacity. This study estimates fixed-effects and random-effects panel stochastic frontier production models to obtain estimates of port efficiency.

Key Words: Panel Data; Stochastic Frontier Models; Technical Efficiency

1. Introduction

Recent advances in the volume of international trade imply that transportation systems are extremely important for most countries throughout the world. Maritime trade accounts for ninety percent of international trade and the efficient operation of seaports is thus of tremendous importance for global trade.[1]

In recent years, container ports in the U.S. faced exogenous shocks which had adverse effects on the volume of maritime trade and port activity. These shocks resulted from natural events such as hurricanes and economic events (related to the recent U.S.-China trade war and financial shocks that resulted in a global recession in 2008) which resulted in a decline in international trade flows.

This study utilizes recent data from the U.S. Department of Transportation that includes consistent information about the operations of the top 25 U.S. container

[1] UNCTAD (2019)

ports for each year from 2015 to 2018. These data are used to estimate a panel data stochastic frontier model of production and to obtain estimates of technical efficiency for these ports.

2. Model, Methodology, and Data

2.1 Model of Container Cargo

The throughput of a container port is the volume of container cargo it handles annually and is measured in twenty-foot equivalent units (TEU), each nominally equal to one 20-foot container. Loaded and empty containers occupy the same space and are equal in terms of TEU. Container flows ae characterized as “inbound” (including imports received from foreign origins, domestic cargo from U.S. origins, and inbound empty containers) and “outbound” (including exports to foreign destinations, domestic cargo shipped to other U.S. destinations, and outbound empty containers). Container ships are loaded and unloaded using large shore-side cranes. Containerized cargo is transported domestically mostly by truck or rail (some is moved by barge on the inland waterway network).

Consider the production function for a container port’s throughput denoted by
$y = f(x_1, x_2, \ldots, x_n)$

where

y = a container port’s total annual throughput defined by twenty-foot equivalent (TEU) units (loaded inbound, loaded outbound, and empty inbound and outbound)
$x_1, x_2, \ldots, x_n$ = the inputs in the production process such as features of a port’s infrastructure and equipment

2.2 Methodology

This paper adopts stochastic frontier analysis for estimating container throughput at the top 25 U.S. container seaports. The stochastic frontier model was developed independently by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). In the stochastic frontier model of production two error terms are added to the production function: one is an idiosyncratic error that is random and the other accounts for technical inefficiency. There are several panel data estimators for estimating stochastic frontier models of production or cost in the econometrics literature (see Sickles and Zelenyuk (2019) and Kumbhakar and Lovell (2000) for surveys). This paper utilizes a few models with varying underlying assumptions: random effects vs. fixed effects, models assuming time-invariant inefficiency vs. time-dependent inefficiency, models with only basic input variables in the production function vs. models that also include the influence of environmental factors (these are not technically inputs but are assumed to influence production) on the production outcome. It is commonly accepted in the literature that there is no one model that is the best model for estimating panel stochastic frontier models. A suggested approach is to incorporate a variety of models with alternative features to determine the performance of these models in explaining the inefficiency aspect that is of primary interest in a stochastic frontier production model. The panel stochastic frontier models used in this study are summarized below.

2.2.1 Random Effects Model with Time-Invariant Inefficiency

Variations of this model have been used in several studies. The original model is the one utilized in Pitt and Lee (1981). The stochastic frontier model for production is given by

$$lny_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j lnx_{jit} + v_{it} - u_i \qquad (1)$$

Where y_{it} is output for firm *i* during period t, x_{jit} is the value of the *j*th input for firm *i* during period t, v_{it} is the idiosyncratic error term, and u_i is the technical inefficiency for the *i*th firm and is assumed to be time-invariant. Alternative distributional assumptions can be made regarding the error terms (such as normal-half normal, normal-truncated normal, normal-exponential). For example the normal-half normal model is

$$v_{it} \sim N(0, \sigma_v^2) \qquad (2)$$

$$u_i = |U_i|, U_i \sim N(0, \sigma_u^2) \qquad (3)$$

2.2.2 Random Effects Model with, Time-Invariant Inefficiency and Double Heteroskedasticity

For the model given above, equation (2) and equation (3) assumed that there was no heteroskedasticity and σ_v^2 and σ_u^2 were assumed to be constant.

$$lny_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j lnx_{jit} + v_{it} - u_i \qquad (1)$$

The double heteroskedasticity model combines equation (1) with the assumption that σ_v^2 and σ_u^2 are not constant and can be modeled as a function of variables affecting these variances Environmental factors can enter the model by assuming that they influence σ_v^2 and/or σ_u^2.

2.2.3 Random Effects Model with Time-Dependent Inefficiency (Time Decay)

This approach is based on models developed by Battese and Coelli (1992, 1995). The stochastic frontier model in equation (4) relaxes the assumption of time-invariant inefficiency.

$$lny_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j lnx_{jit} + v_{it} - u_{it} \qquad (4)$$

where

$u_{it} = g(z_{it})|U_i|$, where $g(z_{it}) = \exp\{-\eta(t-T_i)\}$ (5)

The degree of inefficiency for a firm decreases over time when $\eta > 0$ and increases over time when $\eta < 0$. Given that $t = T_i$ for the last period, the last period for firm *i* contains the base level of efficiency for the firm. If $\eta > 0$, the degree of inefficiency decays toward the base level.

2.2.4 Random Effects Model with Time Dependent Inefficiency (General Form)

Starting with equation (4) above,

$$lny_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j lnx_{jit} + v_{it} - u_{it} \tag{4}$$

the inefficiency term is given by $u_{it} = g(z_{it})|U_i|$ and the z variables are assumed to influence the firm inefficiency (Battese and Coelli 1992, 1995). The z variables can include the environmental factors likely to affect the production outcome.

2.2.5 Fixed Effects Model with Time-Invariant Inefficiency

Cornwall, Schmidt, and Sickles (1977) suggest a modification of a fixed effects linear regression:

$$lny_{it} = \alpha_i + \sum_{j=1}^{k} \beta_j lnx_{jit} + v_{it} \tag{6}$$

The estimated model is

$$lny_{it} = a_i + \sum_{j=1}^{k} b_j lnx_{jit} + v_{it} \tag{7}$$

$$= \max(a_i) + \sum_{j=1}^{k} b_j lnx_{jit} + v_{it} + [a_i - \max(a_i)]$$

$$= a + \sum_{j=1}^{k} b_j lnx_{jit} + v_{it} - u_i \tag{8}$$

Where $u_i = \max(a_i) - a_i > 0$

The most efficient firm has a technical efficiency equal to 1.

2.3 Data and Variable Definitions

The data for estimating panel stochastic frontier models of production for container ports in the U.S. are obtained from the Bureau of Transportation Statistics (BTS) of the U.S. Department of Transportation. The BTS identifies the top 25 U.S. container ports (in terms of TEUs) each year. The list of top 25 container ports by TEU remained relatively stable between 2015 and 2018 with some ports remaining on the list for one year and some remaining for two years. The final dataset was an unbalanced panel of 25 ports with 23 ports having data for all the four years (2015-2018) and two ports having data for only two years.

Variable measuring the dependent variable (y variable):

teutotal = annual total container cargo in TEU units (total includes inbound loaded, outbound loaded, and empties- both in- and outbound)

Variables included as production inputs (x variables):

blength = total berth length, in feet
cr1 = number of Panamax cranes
cr2 = number of Post-Panamax cranes
cr3 = number of Super-Post-Panamax cranes

Environmental factors influencing production (z variables):

atlc = 1 if port is on the Atlantic Coast, = 0 otherwise
gulfc = 1 for a Gulf Coast/Mississippi River port, = 0 otherwise
odrail = 1 if rail intermodal container transfer facility is located within the terminal boundaries, = 0 otherwise

3. Econometric Results

3.1 Panel Stochastic Frontier Models

Prior to utilizing panel estimators described above, a pooled model is estimated using the stochastic frontier estimation framework. Model 1 is the pooled normal-half normal model based on the pooled sample of 96 observations. Model 2 is the normal-half normal model with time-invariant inefficiency. Model 3 is the time-dependent inefficiency model with time decay. Model 2 and Model 3 use panel estimation techniques. The signs of the estimated coefficients of the input variables are consistent with prior expectations. The infrastructure indicators such as berth length and three types of container cranes have a statistically significant and positive impact on the volume of throughput at container terminals (Table 1). The discussion of the results will mainly focus on the inefficiency component. A comparison of the error components from Model 1 and Model 2 suggests that σ_u has increased from 0.018 to 1.30 and σ_v has decreased from .398 to .099. This implies a large reallocation of the random components between noise and inefficiency resulting from the assumption of time invariance of the inefficiency in the process.

Table 1. Estimated Pooled and Panel Stochastic Frontier Models: Production of Container Port Throughput (TEUs)

Dependent Variable: *ln* teutotal

	Model 1: Pooled Model	Model 2: Random Effects (RE-TI)	Model 3: Random Effects (RE-TV1)
Constant	9.579***	12.777***	14.273***
lblength	0.391***	0.188**	0.126**
lcr1	0.139**	0.145***	0.069***
lcr2	0.172***	0.193***	0.099**
lcr3	0.398***	0.153***	0.066**
σ	0.39845	1.304	1.770
λ	0.0464	13.043	24.061
σ_u	0.01846	1.3002	1.768
σ_v	0.39802	0.099	0.073
η			0.0256

***, **, * denote significance at the 1%, 5%, 10% level

$\lambda = \sigma_u / \sigma_v$, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$, η is the time decay parameter

Model 2 RE-TI: Random Effects with Time Invariant Inefficiency

Model 3 RE-TV1: Random Effects with Time Dependent Inefficiency (time-decay model)

$u_{it} = g(z_{it})|U_i|$, where $g(z_{it}) = \exp\{-\eta(t-T_i)\}$

The estimated value of η (= 0.025) > 0 in Model 3 implies that the degree of inefficiency is decreasing in time for the firm. It was shown in the previous section that Model 3 becomes Model 2 when η = 0. The estimated values of the η parameter (which is close to zero) and the other coefficients imply that the two models are quite similar.

Estimates of a fixed effects model with time-invariant inefficiency are reported in Table 2. The estimated model coefficients of the x_{it} input variables are generally similar to the random-effects model with time-invariant inefficiency (Model 2). The estimated technical efficiency estimates are generally higher for the random effects model compared to the fixed effects model (see Figure 1 in the next section).

Table 2. Estimated Panel Stochastic Frontier Models: Random vs. Fixed Effects Model with Time-Invariant Inefficiency

Dependent Variable: *ln* teutotal

	Model 2 Random Effects (RE-TI)	Model 4 Fixed Effects (FE-TI)
Constant	12.777***	
lblength	0.188**	0.067
lcr1	0.145***	0.117***
lcr2	0.193***	0.105**
lcr3	0.153***	0.082**

***, **, * denote significance at the 1%, 5%, 10% level

3.2 Impact of Environmental Factors

The effects of environmental factors were analyzed by estimating variants of two models: the random effects model with time-invariant inefficiency (Model 2) and the random effects model with time-dependent inefficiency (Model 3).

First, the random effects model with time-dependent inefficiency (Model 3) is re-estimated to include environmental variables. The model includes the environmental factors (*atlc, gulfc, odrail)* as the z_{it} variables (see Model 5 column in Table 3). All the three environmental factors were statistically significant. The estimated value of σ_u in Model 5 has reduced to 0.669, compared to 1.768 in the basic time-dependent inefficiency model (Model 3 - without the environmental factors) and σ_v has increased very slightly to 0.1087 from .073 in Model 3.

Table 3. Estimated Panel Stochastic Frontier Models: Effects of Environmental Factors on Production of Container Port Throughput (TEU)

Dependent Variable: *ln* teutotal

	Model 5: Random Effects (RE-TV2)	Model 6: Random Effects (RETI-DH1)
Constant	13.566***	9.624***
lblength	0.153**	0.439***
lcr1	0.123***	0.059
lcr2	0.119**	0.179***
lsppx	0.102***	0.349***
Z_{it}* variables affecting *u		
atlc	0.432**	
gulfc	0.429***	
odrail	0.056***	
Variables affecting σ_v^2		
atlc		-2.322***
gulfc		-2.578***
Variables affecting σ_u^2		
odrail		1.840***
σ	0.675	0.638
λ	7.712	0.385
σ_u	0.669	0.229
σ_v	0.087	0.596

***, **, * denote significance at the 1%, 5%, 10% level

Model 5: Random Effects with Time Varying Inefficiency (General Form)
$u_{it} = g(z_{it})|U_i|$, where the variables in z = (*atlc, gulfc, odrail*)

Model 6: Random Effects with Time Invariant Inefficiency and Double Heteroskedasticity
$\{\ \sigma_v^2 = f(atlc, gulfc)\ ,\ \ \sigma_u^2 = g(odrail)\}$

Second, the random effects model with time-invariant inefficiency (Model 2) was modified to take into account the effects of environmental factors on the production outcome. These factors are assumed to affect either σ_v^2 (the variance of the idiosyncratic error) or σ_u^2 (the variance of the inefficiency component). The estimates for this double – heteroskedastic model (Model 6) are also reported in Table 3. The results suggest that the environmental factors are statistically significant. The estimated value of σ_u in Model 6 has reduced to 0.229, compared to 1.3002 in Model 2 (which did not include the environmental factors) and σ_v has increased slightly to 0.596 from .099 in Model 2.

3.3 Summary of Technical Efficiency Estimates

The estimates of technical efficiency for all ports were higher for the random effects-time-invariant inefficiency model with double heteroskedasticity (Model 6) than the basic random effects time-invariant inefficiency model without environmental factors (Model 2), both the random effects time-dependent/time-varying inefficiency models - with and without environmental factors (Model 3 and Model 5) and the time-invariant fixed effects model (Model 4). The mean, minimum and maximum values of the technical efficiency estimates from Models 2-6 are reported in Table 4.

Table 4: Summary Statistics for the Technical Efficiency Estimates for Models 2-6.

	mean	minimum	maximum
Model 2: RE-TI	0.384	0.094	0.966
Model 3: FE-TV1	0.267	0.044	0.953
Model 4: FE-TI	0.270	0.046	1.000
Model 5: RE:TV2	0.315	0.063	0.977
Model 6: RE-TI-DH1	0.693	0.261	0.969

Figures 1-3 chart the estimated technical efficiency for Models 2-6.

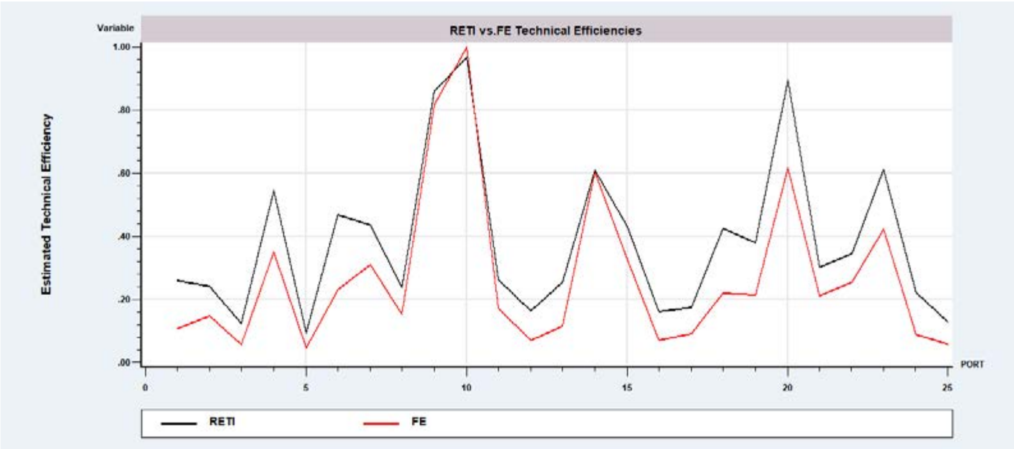

Figure 1. Estimated Technical Efficiency: RE-TI and FE-TI Models

Model 2: RE-TI Random Effects with Time Invariant Inefficiency
Model 4: FE-TI Fixed Effects with Time Invariant Inefficiency

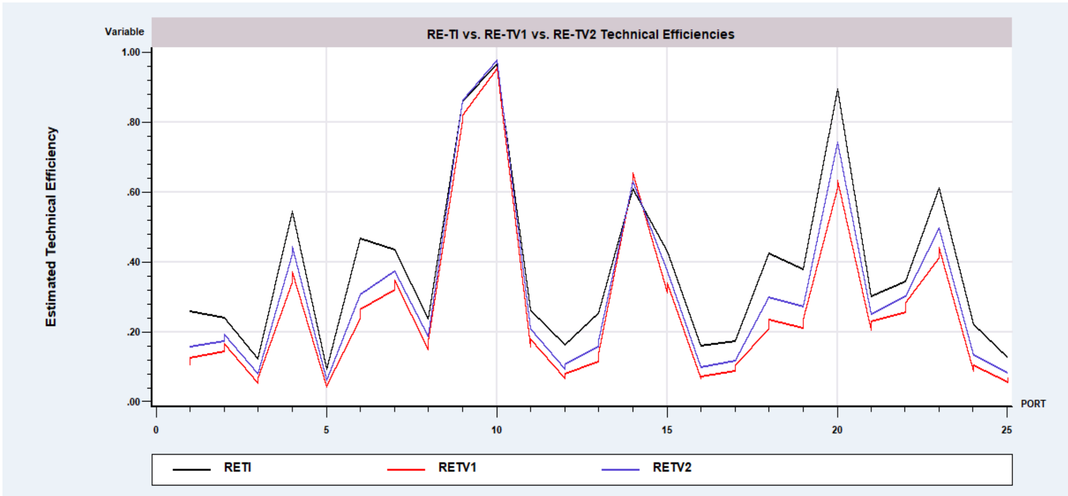

Figure 2: Estimated Technical Efficiency - RE-TI, RE-TV1, RE-TV2 Models

Model 2: RE-TI Random Effects with Time Invariant Inefficiency
Model 3: RE-TV1 Random Effects with Time Dependent Inefficiency (time decay)
$u_{it} = g(z_{it})|U_i|$, where $g(z_{it}) = \exp\{-\eta(t-T_i)\}$
Model 5: RE-TV2 Random Effects with Time Dependent Inefficiency (general form)
$u_{it} = g(z_{it})|U_i|$,, where the variables in z = (*atlc*, *gulfc*, *odrail*)

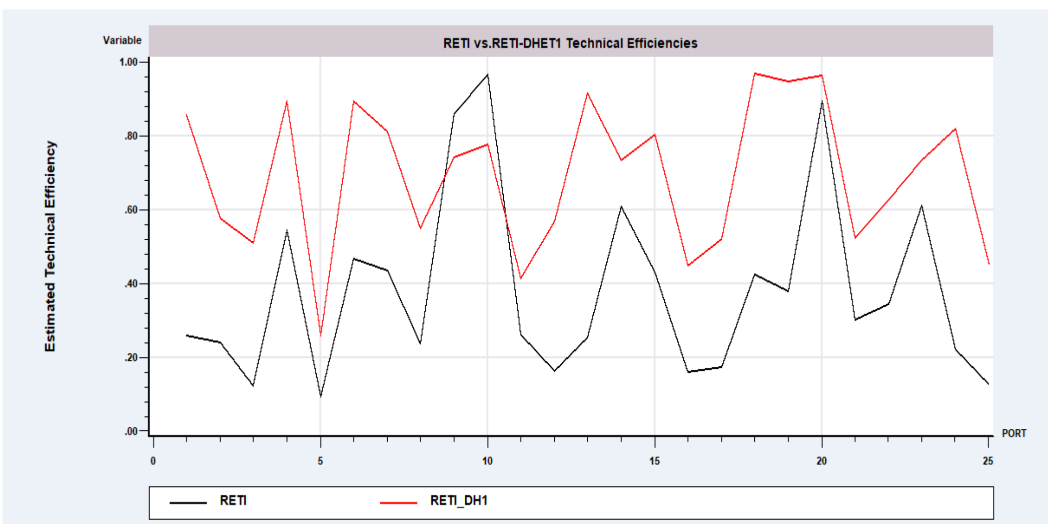

Figure 3: Estimated Technical Efficiency - RETI & RETI_DH1 Models

Model 2: RETI Random Effects with Time Invariant Inefficiency
Model 6: RETI-DH1 Random Effects with Time Invariant Inefficiency & Double Heteroskedasticity
$\{\ \sigma_v^2 = f(atlc, gulfc)\ ,\ \ \sigma_u^2 = g(odrail)\}$

4. Summary and Conclusions

The results of the panel stochastic frontier models reported and the resulting estimates of port technical efficiency suggest that the port infrastructure indicators used as input variables had a statistically significant impact on the volume of container cargo handled by these terminals. The environmental factors such as the presence of an on-dock rail transfer facility and location indicator variables for the Atlantic coast and Gulf coast were also found to exert a significant influence on the port container activity. There are two important implications of the results obtained in this study. First, port infrastructure investments in modern shoreside cranes are effective ways of increasing the port's capacity to load and unload cargo from the large container vessels commonly operating after the Panama Canal expansion project. Second, the regional variables indicating ports on the Atlantic and Gulf coasts imply that location is an additional determinant of port activity. These variables might partly represent random factors (e.g. hurricane impacts) but might also be linked to the inefficiency component of the port's activity.

References

Aigner, D., C.A.K. Lovell, and P. Schmidt (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics*, 6(1), 21-37

Battese, G. and T. Coelli (1992). "Frontier Production Functions, Technical Efficiency and Panel Data: With Applications to Paddy Farmers in India," *Journal of Productivity Analysis*, 3(1), 153-169.

Battese, G. and T. Coelli (1995). "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data," *Empirical Economics*, 20(2), 325-332.

Cornwell, C., P. Schmidt and R. Sickles (1990). Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels," *Journal of Econometrics*, 46(1), 185-200.

Bureau of Transportation Statistics, U.S. Department of Transportation.

Greene, W. (2005a). "Fixed and Random Effects in Stochastic Frontier Models," *Journal of Productivity Analysis*, 23(1), 7-32.

Greene, W. (2005b). "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model," *Journal of Econometrics*, 126(2), 269-303.

Cullinane, K. and T. Wang (2010). "The Efficiency Analysis of Container Port Production using DEA Panel Data Approaches," *OR Spectrum*, 32(3), 717-738.

Kumbhakar, S and C. A. K. Lovell (2000). *Stochastic Frontier Analysis*. New York, NY: Cambridge University Press.

Medda, F. and Q. Liu (2012) "Determinants and Strategies for the Development of Container Terminals," *Journal of Productivity Analysis*, 40, 83-98.

Meeusen, W. and van den Broeck {1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Composite Error," *International Economic Review*, 18(2), 435-444.

Pitt, M. and L. Lee (1981). "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," *Journal of Development Economics*, 9(1), 43-64.

Schmidt, P. and R. Sickles (1984). “Production Frontiers and Panel Data,” *Journal of Business and Economic Statistics*, 2(4), pp. 367-374.

Sickles, R. and V. Zelenyuk (2019). *Measurement of Productivity and Efficiency: Theory and Practice*. New York, NY: Cambridge University Press.

UNCTAD (2019). Review of Maritime Transport, United Nations.