A Nonparametric Economic Analysis of the US Natural Gas Transmission Infrastructure: Efficiency, Trade-Offs and Emerging Industry Configurations

Corrado lo Storto [ID]

Department of Industrial Engineering, University of Naples Federico II, 80125 Naples, Italy;
corrado.lostorto@unina.it

Abstract: This paper presents a study aimed at measuring the efficiency of the transmission segment of the US natural gas industry from an economic perspective. The gas transmission infrastructure is modeled as an economic production function and a multi-stage modeling approach based on the implementation of Data Envelopment Analysis is employed to obtain an efficiency measure in a two-dimension performance space, i.e., cost and revenue-efficiency. This approach allows taking into account conflicting business goals. The study also performs cluster analysis to uncover homogeneous efficiency profiles relative to the gas transmission systems to explore determinants of efficiency rates, and trade-off situations. A sample containing 80 US gas transmission systems is used in the analysis. Results indicate that the transmission segment of the US gas industry has considerable inefficiencies, while average cost and revenue-efficiency scores are 0.324 and 0.301, and only three transmission systems achieve high scores on both efficiency dimensions. Cluster analysis identified seven configurations. In three of them there are no trade-off situations between cost and revenue efficiencies. However, only in one of them gas transmission systems have high efficiencies. The remaining four configurations exhibit trade-off situations having different intensity. Such trade-offs can be determined by the gas transmission infrastructure size.

Keywords: natural gas industry; United States; transmission systems; Data Envelopment Analysis; efficiency; trade-offs; configurations

1. Introduction

About one quarter of the United States energy needs depend on natural gas supply. According to estimates from the US Energy Information Administration, the total natural gas consumption in 2016 was 27,485,517 million cubic feet [1]. Natural gas is generally used as fuel in natural gas processing plants, fuel used by vehicles, and in private dwellings, including apartments, for heating, air-conditioning, cooking, water heating, and further household uses [2].

Since the beginning of the 2000s the supply of natural gas is playing an important role in the energy strategy of US, ensuring that the economy of the country relies more and more on diversified mix of energy sources. Indeed, between 2001 and 2015, the production of natural gas in US has increased by more than 40%, whereas its price (city-gate price) diminished by about 25% making natural gas a more competitive source in the energy market [3,4].

As in other countries, the transmission segment of the natural gas industry is regulated both at the federal and local levels. At the federal level, regulating entities are the Federal Energy Regulatory Commission (FERC), Pipeline and Hazardous Materials Safety Administration (PHMSA), Occupational Safety and Health Administration (OSHA), Transportation Safety Administration (TSA), and Environmental Protection Agency (EPA), whereas at the local level there are a number of public service or public utility commissions whose main task is to control that the local distributors choose

a good balance between investment in network maintenance, upgrade and expansion and tariffs to market consumers.

Natural gas is economically and efficiently transported through pipeline systems under pressure (ranging from 500 to 1400 psi). A typical transmission system is composed of pipelines, compression stations, valves, processing plants, metering stations, control systems and storage facilities. Pipelines are wide-diameter lines that move natural gas over long distances from the producing or the processing stations to storage facilities and distribution centers. Pipelines are routinely inspected and repaired when needed to keep them at the optimal level of efficiency and safety. Compression stations create pressure differentials allowing the gas flowing from an area of higher pressure to an area of lower pressure. They are located every 50–60 mile intervals along the transmission pipeline and are generally operated by using turbine compressors or electric motors. Valves are mechanical devices utilized to control the flow of gas along a pipeline and section the pipeline to carry on preventive or unplanned maintenance operations. Processing plants are necessary to separate natural gas from hydrocarbon gas liquids and water, and to remove impurities. Metering stations allow the transmission system operator to measure the flow of gas along the pipeline. SCADA systems are used to collect and process data in real time necessary to gas flow monitoring along the pipeline [5]. These control systems can remotely operate compressors and valves to adjust flow rates. Storage facilities are used to balance market demand and supply of natural gas.

Since the 1950s an extensive and interconnected transportation infrastructure has been developed to move natural gas from regions where it is produced to regions where it is consumed [6]. The following figures give evidence of the complexity of the US natural gas infrastructure:

- 210 interstate and intrastate gathering and transmission pipelines that extend for more than 320,000 miles around the country (including about 20,000 miles of gathering pipelines);
- more than 1400 compression stations, 11,000 delivery points, 1400 interconnection points, 5000 receiving points, 24 hubs, 400 natural gas storage facilities, and eight liquefied natural gas (LNG) facilities.

On average, the transmission infrastructure moves 70 billion cubic feet of natural gas to 1300 local distribution companies that sell this commodity to more than 71 million customers, e.g., households, commercial and industrial firms.

According to forecasts by the US Energy Information Administration demand for natural gas will double by the end of 2030. To support such a growth in gas consumption, the gas transmission infrastructure has developed by a factor of 100 in the last 50 years, and the positive trend will continue with a rate of about 7% per year over the next two decades. As about half of the natural gas transmission network was built between the 1950s and 1960s, additional investment is needed to keep the infrastructure in operation.

Network efficiency is an important goal to achieve both in the planning of new transmission infrastructure and the management of the existing one. However, while engineers are generally more interested to increase thermal, compressing and hydraulic pipeline efficiencies, improving the overall economic efficiency of the natural gas network should be a major concern of policy makers as a higher economic efficiency is usually related to lower cost and prices to customers [7].

This paper presents an efficiency study of the transmission segment of the US natural gas industry by assuming an economic perspective. The gas transmission infrastructure is modeled as an economic production function and its economic efficiency is defined as the ratio of weighted outputs to weighted inputs. The combined resources necessary to operate the transmission network (i.e., pipeline, compressors stations, people, etc.) should be combined in the most efficient way to provide competitive and cost-effective movement of the natural gas from one location to another. Henceforth, more output per unit indicates greater efficiency.

The study implements Data Envelopment Analysis (DEA) and adopts a multi-stage modeling approach to generate an efficiency measurement in a two-dimension space for a sample of US natural

gas transmission systems taking in account conflicting business goals. Furthermore, the study performs cluster analysis to uncover homogeneous efficiency profiles relative to the gas transmission systems which help explore determinants of high-efficiency and trade-off situations. The paper is organized as follows: Section 2 presents a literature review focused on efficiency measurement in the gas industry by using DEA. Section 3 introduces the multi-stage modeling approach, while Section 4 presents the DEA method. Section 5 provides information relative to sample, data and variables of DEA models specification adopted to conduct the efficiency analysis. Results of the study are reported in Section 6. Finally, conclusions and limitations are discussed in the last section.

2. Literature Review

There is a huge amount of papers that consider the measurement of economic efficiency in the energy industry. Particularly, some scholars performed extensive literature reviews considering papers that used DEA to deal with different efficiency measurement issues in the field of energy generation and management [8–11]. However, there are relatively few papers that provided efficiency measurements in the specific gas industry adopting DEA, and most of the efficiency studies focused on the distribution segment of the industry. Hollas et al. [12] investigated the impact of the Natural Gas Policy Act of 1978 and policies of the US Federal Energy Regulatory Commission (FERC) that increased the level of competition on the industry efficiency. Scholars employed DEA to examine the economic efficiency of gas distributing companies between 1975 and 1994. Results showed that the reduction of scale economies did not modify the operators' economic efficiency. Hawdon [13] employed bootstrapped DEA to estimate efficiencies of the natural gas industry in 33 countries. The scholar finds that the national industry efficiencies are significantly affected by rising or falling of gas sales. Moreover, results support the assumption that the reforms of the energy market occurred in some countries (for instance, UK) have positively affected the efficiency of the gas industries, improved the utilization of labor and utilization of capital. Jamasb et al. [14] used non-parametric DEA and regression analysis to study the impact of US gas industry regulatory reform for a panel of US interstate companies in terms of static efficiency and productivity. Sample contains 39 companies observed from 1997 to 2004. Results indicate that regulation stimulated efficiency increase. Erbetta and Rappuoli [15] examined the nature of returns to scale in the Italian natural gas distribution industry by using DEA. Results show that scale inefficiencies negatively affect the overall efficiency of gas operators, whereas technology shows increasing returns only for the smallest operators suggesting that efficiency improvement can be achieved by intensify the merging process and concentration that have characterized the early years of the 2000s. Goncharuk [16] developed three DEA models to calculate the efficiency in the gas industry distribution segment In Ukraine and US comparing 54 Ukrainian and 20 US operators. The author analyzed factors that have an impact on efficiency, e.g., scale, regional location, ownership, etc. The benchmarking study allowed find that Ukrainian gas distribution companies are generally scarcely efficient and there is a potential 10% resource consumption that should be reduced to achieve industry efficiency in Ukraine. Zorić et al. [17] conducted a cross-country benchmarking study considering a sample including gas distribution utilities in Slovenia, the Netherlands and UK. This study showed that UK utilities perform better than Dutch and Slovenian utilities, and these latter are less efficient than Dutch utilities, even though they operate at optimal scale. According to scholars, such efficiency difference might be explained in terms of a more extensive regulation of the UK gas market. Amirteimoori and Kordrostami [18] proposed a Euclidean distance-based measure of efficiency to develop a DEA super-efficiency score is proposed to have a better discrimination of units. The super-efficiency model is utilized to estimate efficiency of 25 Iranian gas companies. Nieswand et al. [19] employ PCA-DEA to measure the efficiency of 37 US natural gas transmission companies in 2007. Particularly, they implement two model settings which include the same cost measurement but differ in the number of cost drivers under the assumption of variable returns to scale technology. The adoption of PCA-DEA allows reduce the number of efficient companies in the sample in comparison to conventional DEA. Ertürk and Türüt-Aşık [20] adopted DEA to evaluate the

efficiency of 38 gas management companies in the distribution segment of the Turkish natural gas industry. Technical, allocative and cost efficiencies under the assumptions of both constant and variable returns to scale were calculated to identify reasons of low performance and trajectories of improvement. Scholars found that the high investment costs are a major cause of inefficiency. Additionally, public owned and large size operators utilize resources more efficiently. Sadjadi et al. [21] utilized a stochastic super-efficiency DEA model to assess the efficiency of a sample of 27 Iranian province gas companies and generate a ranking across them. The model is based on a robust optimization technique that may be an alternative to sensitivity analysis and stochastic programming. Marques et al. [22] implemented DEA under the assumption of variable returns to scale to evaluate the efficiency of the Portuguese gas distributors to identify targets for the regulatory period, 2010–2013. To avoid sample misspecifications, distributors were divided into three groups with different scale factors and exogenous factors were also included in the study. The cross-section analysis was crossed with a dynamic one using panel data methodology. Results suggest that factors influencing the efficiency of Portuguese gas distributors can be different and depend on the characteristics of the company and the operating context. lo Storto [23] evaluated the operational, density and scale efficiencies of the natural gas distribution industry in Italy by implementing DEA. The empirical analysis considered a sample of 32 natural gas distributing companies. Results indicate that the average operational inefficiency in the sample is about 25%, and scale inefficiency is a major cause of scarce performance. Yardimci and Karan [24] measured the efficiency and service quality of a sample of Turkish natural gas distribution companies. Both DEA and statistical methods were used to calculate efficiency, whereas the quality of service was used to rank companies analyzing the correlation between efficiency and service quality. Results are useful to obtain insights relatively to the effectiveness of market regulation and the adoption of reward/penalty schemes to choose industry tariffs. Goncharuk and lo Storto [25] performed a cross-country benchmarking study considering a mixed sample of natural gas distributing companies in Italy and Ukraine. They use a 2-stage DEA procedure to estimate efficiency of gas providers and find critical context factors and policy issues that impact on it. Results show that both countries are low performing with respect to concessionaire operational efficiency and size. However, while increasing efficiency is necessary to reduce cost and improve quality of service, experience indicates that other goals may be critical at different stages of the reform of the industry in both countries.

Scholars who used DEA to measure efficiency in the gas industry generally adopted a "black-box" modeling approach. However, the black-box approach is unable to provide robust efficiency measurements when the units to be evaluated are complex systems and there are conflicting business goals that influence the management decision-making and have an impact on the system performance. The aim of this study is to fill these gaps.

3. The Measurement of the Economic Efficiency of Gas Transmission Infrastructure

Implementing DEA to measure the gas transmission infrastructure efficiency requires that the corresponding production technology is modeled in terms of inputs and outputs. To have a more accurate efficiency measurement and accounting for both financial and operational issues, the gas transmission production process was split into the following three stages: (a) cost generation; (b) operations management; (c) revenue generation (Figure 1). The outputs of the first production stage (cost generation) are used as inputs of the second production stage (operations management), whereas the outputs of this stage are used as inputs of the third production stage (revenue generation).

Hence, this multi-stage production model of the gas transmission infrastructure is more effective than a conventional one-stage or black-box model to understand how different types of resources (i.e., financial and physical) are sequentially utilized and transformed to produce necessary outputs. Indeed, by adopting this approach the underlying production function is decomposed into three interlinked sub-production functions that capture the same number of different efficiency components, i.e., cost generation-efficiency (CE), operations management-efficiency (OE) and revenue-generation efficiency (RE).

At the first stage of this model (cost generation), costs are incurred to carry on the gas transmission business operations, preventive and unplanned maintenance when the infrastructure is utilized. The gas transmission infrastructure is efficient from the cost-generation view if it can be operated and maintained spending the minimum cost. At the next stage (operations management), the model is focused on the service provided by the infrastructure (i.e., gas transmission from one point to another). Efficiency increases when the same volume of natural gas can be delivered by utilizing a physical infrastructure having reduced capability (i.e., length, number of compression facilities). Hence, for a given volume of gas transmitted from one point to another, the lower the infrastructure capability, the higher the operations-management efficiency. The companies operating the transmission infrastructure are generally reluctant to make additional investment to avoid reducing profits. Finally, at the last stage (revenue generation), revenues are generated selling the gas transmission service to the distributing companies. Accordingly, the revenue-generation efficiency is measured as the ratio of the revenues generated selling the gas transmission service to the distributing companies to the volume of gas transmitted. The higher the revenues are for a given volume of gas, the higher the revenue-generation efficiency.
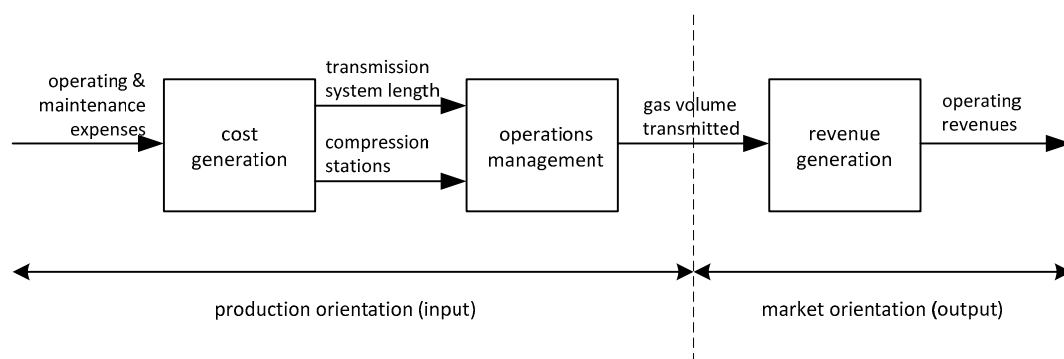

Figure 1. The multi-stage production model of the gas infrastructure system.

Whereas the first two efficiency components are measured in terms of input reduction, the third one is measured in terms of output increase. Therefore, to implement Data Envelopment Analysis (DEA) and compute efficiency consistently, the multi-stage model was reorganized into two main parts, one having an input orientation (production segment) and including the cost-generation and operations management stages, and one having an output orientation (market oriented segment), including the revenue-generation stage. Particularly, Network DEA (NDEA) was used to calculate efficiency in the production-orientated segment of the model and conventional DEA to calculate efficiency in the market-oriented segment. In the next section, both methods are illustrated.

4. Method

Data Envelopment Analysis (DEA) is a nonparametric method based on the adoption of linear programming techniques commonly used to evaluate the efficiencies of a set of units denominated DMUs (i.e., decision-making units). The efficient DMUs are identified from this set and combined to construct an efficient frontier used as a benchmark to measure the efficiency of inefficient units [26]. Efficiency is measured as the ratio of the weighted sum of output variables to the weighted sum of input variables. The method does not require any assumption about the functional form of the relationship necessary to convert inputs into outputs and the weights utilized to combine them. Hence, the production technology that transforms inputs into outputs is generally considered as a black-box [27].

In this study, gas infrastructure systems are considered as DMUs in the DEA model formulation. We assume there are n DMUs ($j = 1, \dots , n$) corresponding to the same number of gas transmission

systems that should be evaluated. Every DMU consumes varying amounts of m different inputs to produce r different outputs.

Particularly, the generic transmission system DMU_j consumes amounts x_j of inputs (x_{ij} with $I = 1,$ ..., m), whereas produces amounts y_j of outputs (y_{kj} with $k = 1, ..., r$). $x_o \equiv (x_{1o}, ..., x_{mo})$ and $y_o \equiv (y_{1o}, ..., y_{ro})$ indicate amounts of inputs and outputs of the gas transmission system identified by DMU_o that is under evaluation. $X = (x_{ij}) \in \Re^{m \times n}$ and $Y = (y_{kj}) \in \Re^{r \times n}$ with $X > 0$ and $Y > 0$ respectively denote the $m \times n$ input and the $r \times n$ output matrices for the n systems.

The slack-based-measure (SBM) efficiency index is used in the specification of the DEA model because it does not assume proportional changes of inputs or outputs and, consequently, provides more realistic efficiency measurements [28]. Inputs and outputs of DMU_o (x_o, y_o) can be described as follows:

$$
\begin{aligned}
x_o &= X\lambda + s^- \\
y_o &= Y\lambda - s^+, \ \lambda \geq 0
\end{aligned}
\tag{1}
$$

where s^- and s^+ are respectively input and output slack variables, and λ is a nonnegative vector in $\Re^n$. When output is increased by s^+ and/or input is decreased by s^- DMU_o can achieve full efficiency.

Further, we assume that the gas transmission infrastructure production function has variable returns-to-scale (VRS) at all production segments because of the great variance across the gas transmission infrastructure size in the sample.

Two DEA models are specified, one for the input-oriented segment and one for the output-oriented segment of the production model of the transmission infrastructure.

4.1. Production Oriented Segment

The weighted network slacks-based measure (NSBM) model proposed in the literature is used to evaluate efficiencies of gas transmission systems in the production-oriented segment [29]. We consider a multiple-stage production process consisting of T production stages ($t = 1, ..., T$) and assume there are m_t and r_t inputs and outputs to stage t. The link from stage t to stage h and the set of links are denoted by (t, h) and L, respectively. The observed measurements of inputs to DMU_j at stage t are $\{\mathbf{x}_j^t \in \Re_+^{m_t}\}$ ($j = 1, ..., n; t = 1, ..., T$) and the observed measurements of outputs from DMU_j at stage t are $\{\mathbf{y}_j^t \in \Re_+^{r_t}\}$ ($j = 1, ..., n; t = 1, ..., T$), respectively. The observed data that measure the linking intermediate products from stage t to stage h are $\{\mathbf{z}_j^{(t,h)} \in \Re_+^{g_{(t,h)}}\}$ ($j = 1, ..., n; (t, h) \in T$) where $g_{(t,h)}$ is the number of items in link (t, h). In addition, we assume that intermediate links are freely determined.

The input-oriented efficiency θ_o^* of DMU_o can be evaluated by solving the following linear program:

$$
\begin{aligned}
\theta_o^* = &\min_{\lambda^t, s^{t-}} \sum_{t=1}^{T} w^t \left[1 - \frac{1}{m_t} \left(\sum_{i=1}^{m_t} \frac{s_i^{t-}}{x_{io}^t} \right) \right] \\
&\text{subject to} \\
&\mathbf{x}_o^t = \mathbf{X}^t \boldsymbol{\lambda}^t + \mathbf{s}^{t-}, \quad t = 1, ..., T \\
&\mathbf{y}_o^t = \mathbf{Y}^t \boldsymbol{\lambda}^t - \mathbf{s}^{t+}, \quad t = 1, ..., T \\
&\mathbf{e}\boldsymbol{\lambda}^t = 1, \quad t = 1, ..., T \\
&\sum_{j=1}^{n} \lambda_j^t = 1 (\forall t), \quad \lambda_j^t \geq 0 \quad (\forall j, t) \\
&\boldsymbol{\lambda}^t \geq \mathbf{0}, \ \mathbf{s}^{t-} \geq \mathbf{0}, \ \mathbf{s}^{t+} \geq \mathbf{0}, \ t = 1, ..., T \\
&\mathbf{Z}^{(t,h)} \boldsymbol{\lambda}^h = \mathbf{Z}^{(t,h)} \boldsymbol{\lambda}^t, \ (\forall j, t)
\end{aligned}
\tag{2}
$$

where:

w^t is the relative weight of production stage t, $\sum_{t=1}^{T} w^t = 1$, $w^t \geq 0$ $(\forall t)$

$\lambda^k \in \Re_+^n$ is an intensity vector related to production stage t ($t = 1, ..., T$)

$\mathbf{Z}^{(t,h)} = \left(\mathbf{z}_1^{(t,h)}, ..., \mathbf{z}_n^{(t,h)} \right) \in \Re^{g_{(t,h)} \times n}$

4.2. Market Oriented Segment

Efficiencies of gas transmission systems in the output-oriented segment of the multi-stage production model are calculated by implementing a conventional DEA model, under the assumption of output orientation (maximization) [30].

Under the assumption of output orientation and variable returns to scale, in the SBM-model the efficiency of a $DMU_o(x_o, y_o)$ can be measured by solving the following fractional program [28]:

$$\min \frac{1}{\rho} = 1 + \frac{1}{r} \sum_{k=1}^{r} \frac{s_k^+}{y_{ko}}$$
$$s.t. \ X\lambda + s^- \leq x_o$$
$$Y\lambda - s^+ = y_o$$
$$\sum_{j=1}^{n} \lambda_j = 1 \qquad (3)$$
$$\lambda \geq 0, \ s^- \geq 0, \ s^+ \geq 0,$$
$$k = 1, 2, ..., r$$

Variables s^- and s^+ measure the distance of DMU inputs and outputs from inputs $X\lambda$ and outputs $Y\lambda$ of a virtual unit. When $s_o^+ = s_o^- = 0$, $1/\rho = 1$ and DMU_o is efficient.

5. Data, Variables and DEA Model Specifications

This study utilized data relative to US natural gas transmission systems in 2012 that are available in the *Oil & Gas Journal* [31,32]. The research sample includes 80 gas transmission systems chosen on the base of data availability for all variables used in the efficiency analysis that is about 49% of the total number of transmission systems reported by [31].

Table 1 displays variable main statistics for the production-oriented segment (model 1) and the market-oriented segment (model 2) of the multi-stage production model of the transmission infrastructure. Figures show that there is a considerable variance among transmission systems in the sample. For instance, the length of the transmission system varies from 28 miles to 14,807 miles, while the number of compression stations ranges between 0 and 121. Three types of variables were used, "pure" inputs, "pure" outputs and "mixed" (or intermediate) variables. The first two types include variables which are used as inputs or outputs, while the third type includes variables used either as inputs or outputs, depending on the DEA model specification and stage considered.

Table 1. Inputs, outputs and intermediate (input/output) links.

| Variable | Type | | | Description | Unit | Mean | St.dev. | Max | Min |
| | Model 1 | | Model 2 | | | | | | |
	Stage 1	Stage 2							
X_1	input	-	-	operating and maintenance expenses	US$ ($\times$ 1000)	80,134	118,987	568,156	1421
Z_1	output	input	-	transmission system length	miles	2379	3294	14,807	28
Z_2	output	input	-	no. of compression stations	number	18	23	121	0
Z_3	-	output	input	gas volume trans. for others	MMcf	514,760	631,797	4,043,156	11,088
Y_1	-	-	output	operating revenue	US$ ($\times$ 1000)	243,440	289,840	1,367,655	2010

Legend: model 1 = production-oriented segment; model 2 = market-oriented segment; MMcf = Million Cubic Feet.

In model 1 there are one input, one output and two mixed variables shared between stage 1 and stage 2. Particularly, the amount of operating and maintenance expenses (X_1) and the gas volume transmitted (Z_3) were introduced in the analysis respectively as input in stage 1 and output in stage 2 of model 1, while the transmission system length (Z_1) and the compression stations (Z_2) were used as outputs in stage 1 and inputs in stage 2 of the same model. Model 2 contains one input and one output, respectively the gas volume transmitted (Z_3) and the operating revenue (Y_1). Observed data relative to each variable were normalized by dividing them by their means.

Additionally, as variables Z_1 and Z_2 showed high correlation, the method proposed by [33] was adopted to merge them performing Principal Component Analysis (PCA). The new composite variable was used as a proxy of the infrastructure capacity [34].

Table 2 summarizes information relative to the two DEA models adopted in the study. Variable $Z[1, 2]$ was constructed as the weighted average of Z_1 and Z_2 by utilizing the PCA eigenvectors as weights.

Table 2. DEA model specifications.

| Model | Production Model Segment | Method | Variables | | | Orientation |
			Input	Output	Intermediate	
model 1	production-oriented	NSBM DEA	X_1	Z_3	$Z[1, 2]$	input
model 2	market-oriented	SBM DEA	Z_3	Y_1	-	output

Legend: Variable $Z[1, 2]$ was obtained merging variables Z_1 and Z_2.

6. Results

6.1. Efficiency Measurement

The results of the calculation of the efficiencies of the gas transmission infrastructure systems by implementing model 1 and model 2 are displayed in Table 3. The efficiency scores obtained by performing model 1 are reported in column named "costEff" and the efficiency scores obtained performing model 2 appear in column "revEff" (model 2), respectively. Particularly, the index denominated costEff provides an aggregate measurement of the efficiencies at stages "cost generation" and "operations management" of the multi-stage model (relative to the overall production-orientated segment).

Natural gas infrastructure systems that are on the efficiency frontier enveloped by model 1 are considered 100% cost-efficient, while those that are on the frontier enveloped by model 2 are identified as 100% revenue-efficient. Both DEA models have high discriminating capability as a relatively small number of gas transmission systems are fully efficient.

Six systems are 100% cost-efficient and four systems are 100% revenue-efficient. Particularly, Transcontinental Gas Pipe Line Co. LLC (Houston, TX, USA) achieves the maximum efficiency in both models, while Tennessee Gas Pipeline Co. is placed on the efficient frontier in model 1 and is very close to the efficient frontier in model 2. These results suggest that cost-efficiency and revenue-efficiency can be two compatible business goals to achieve at the same time.

In both segments of the multi-stage production model of the gas infrastructure systems, efficiency scores are extremely variable. Indeed, the standard deviation values displayed in Table 3 are relatively high and very close to means. However, average efficiencies are very low, respectively 0.324 (cost-efficiency) and 0.301 (revenue-efficiency), whereas minimum efficiencies are 0.018 and 0.007 emphasizing the low efficiency of the transmission segment of the US natural gas industry. Even though the cost and revenue efficiencies have similar means and standard deviations values, the behaviors of the two indexes are very different and there is no correlation between them as their plot in Figure 2 shows. Indeed, this plot supports what emerged from Table 3. Three gas infrastructure systems having both high cost and revenue-efficiencies can be easily identified: Transcontinental Gas Pipe Line Co. LLC, Tennessee Gas Pipeline Co., and Tennessee Gas Pipeline Co. However, the remaining transmission systems are scattered in the plane and there is no evident association between the efficiency indexes.

Table 3. Efficiencies for the two DEA models.

| Name	CostEff	RevEff	Name	CostEff	RevEff
Algonquin GAS Transmission LLC	0.216	0.309	Millenium Pipeline Co. LLC	0.408	0.246
Alliance Pipeline LP	0.176	0.315	Mojave Pipeline Co.	0.599	0.041
American Midstream (Ala Tenn) LLC	0.618	0.019	National Gas Fuel Supply Corp.	0.081	0.267
ANR Pipeline Co.	0.420	0.430	Natural Gas Pipeline Co. of America	0.328	0.544
Big Sandy Pipeline LLC	0.274	0.121	North Baja Pipeline LLC	0.487	0.092
Carolina Gas Transmission Corp.	0.125	0.140	Northern Border Pipeline Co.	0.823	0.259
Cheyenne Plains Gas Pipeline Co. LLC	0.220	0.227	Northern Natural Gas Co.	0.198	0.537
Cimarron River Pipeline LLC	0.079	0.074	Northwest Pipeline LLC	0.115	0.524
Colorado Interstate Gas Co.	0.135	0.387	Ok Tex Pipeline Co. LLC	0.955	0.007
Columbia Gas Transmission LLC	0.178	0.871	Ozark Gas Transmission LLC	0.178	0.049
Columbia Gulf Transmission Co	0.225	0.163	Paiute Pipeline Co.	0.096	0.090
Crossroads Pipeline Co	1.000	0.043	Panhandle Eastern Pipe Line Co. LP	0.073	0.444
Dauphin Islands Gathering Partners	0.154	0.265	Petal Gas Storage LLC	0.259	0.143
Destin Pipeline Co. LLC	0.213	0.074	Questar Overthrust Pipeline Co.	0.805	0.093
Discovery Gas Transmission LLC	0.268	0.028	Questar Pipeline Co.	0.081	0.281
Dominion Cove Point LNG	0.018	0.791	Questar Southern Trails Pipeline Co.	0.184	0.040
Dominion Transmission Inc.	0.061	1.000	Rockies Express Pipeline LLC	0.133	0.805
East Tennessee Natural Gas LLC	0.091	0.268	Ruby Pipeline LLC	0.172	0.566
El Paso Natural Gas Co	0.359	0.463	Sabine Pipe Line LLC	0.131	0.016
Elba Express Co. LLC	0.562	0.213	Sea Robine Pipeline Co. LLC	0.115	0.114
Empire Pipeline Inc	0.349	0.156	Southeast Supply Header LLC	0.426	0.175
Enable Gas Transmission	0.178	0.401	Southern Natural Gas Co.	0.211	0.530
Enable Mississippi River Transmission LLC	0.069	0.155	Southern Star Central Gas Pipeline Inc.	0.050	0.356
Equitrans LP	0.134	0.258	Tallgrass Interstate Gas Transmission LLC	0.048	0.246
ETC Tiger Pipeline LLC	0.347	0.421	TC Offshore LLC	0.146	0.060
Fayetteville Express Pipeline LLC	1.000	0.230	Tennessee Gas Pipeline Co.	1.000	0.841
Florida Gas Transmission Co. LLC	0.197	0.738	Texas Eastern Transmission LP.	0.529	0.793
Garden Banks Gas Pipeline LLC	0.428	0.024	Texas Gas Transmission LLC	0.478	0.364
Gas Transmission Northwest Corp	0.546	0.193	Trailblazer Pipeline Co.	0.220	0.042
Great Lakes Gas Transmission LP	0.212	0.158	TransColorado Gas Transmission Co.	0.311	0.110
Guardian Pipeline LLC	0.178	0.160	Transcontinental Gas Pipe Line Co. LLC	1.000	1.000
Gulf Crossing Pipeline LLC	0.070	0.351	Transwestern Pipeline Co. LLC	0.093	0.308
Gulf South Pipeline Co. LP	0.255	0.383	Trunkine Gas Co. LLC	0.390	0.179
Gulfstream Natural Gas System LLC	0.332	0.392	Tuscarora Gas Tranmission Co.	0.322	0.076
Horizon Pipeline Co. LLC	0.518	0.040	Vector Pipeline LP	0.794	0.119
Iroquois Gas Transmission Systems LP	0.190	0.313	Viking Gas Transmission Co.	0.204	0.073
Kern River Gas Transmission Co.	0.041	1.000	WBI Energy Transmission Inc.	0.086	0.182
KPC Pipeline LLC	0.210	1.000	Wyoming Interstate Co Ltd.	1.000	0.161
Maritimes & Northeast Pipeline LLC	0.124	0.394	mean	0.324	0.301
Midcontinent Express Pipeline LLC	1.000	0.233	st.dev	0.275	0.267
Midwestern Gas Transmission Co.	0.318	0.055	max	1.000	1.000
MIGC Inc	0.326	0.049	min	0.018	0.007
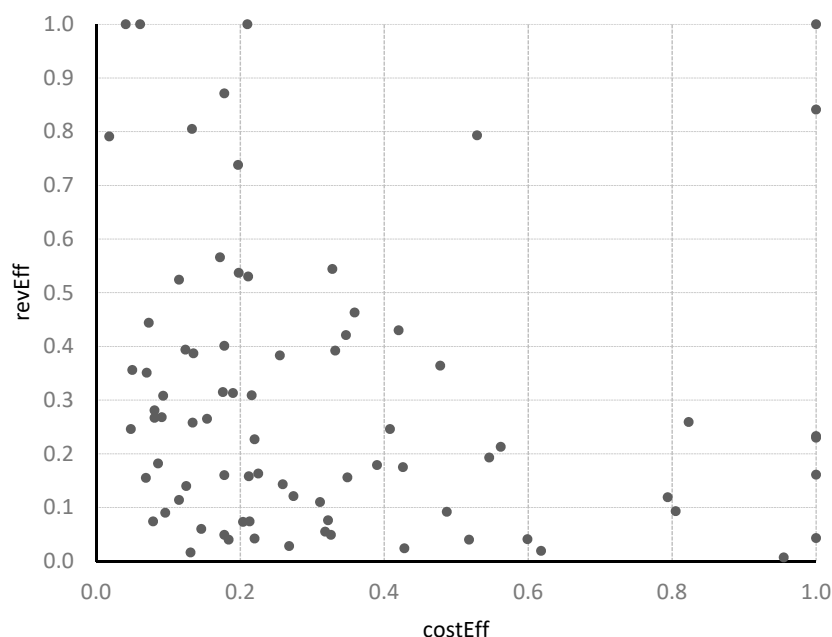
Figure 2. Plot of costEff and revEff scores.

6.2. Relationship between Cost and Revenue-Efficiencies and Trade-Off Analysis

An inductive approach was implemented to conduct a more in-depth data analysis [35]. Particularly, the clustering method was used as a tool to explore the relationship between the cost and revenue efficiency scores by finding similarities between the natural gas transmission systems regarding their efficiency measurements and finally extracting homogeneous configurations from the sample. In this study, configurations are defined as groups of gas transmission systems that share a common bi-dimensional efficiency profile. The analysis of configurations will help identifying the characteristics of the transmission infrastructure that affect efficiency performance and, particularly, efficiency trade-offs.

The cost-efficiency and revenue-efficiency scores were used as clustering variables, while the k-means clustering algorithm was chosen to identify groups because it is very efficient and easy to implement [36]. However, because the algorithm converges to an arbitrary local optimum and does not provide indications about the correct number of groups, the following "robust" clustering procedure was adopted. Several clustering strategies were employed to evaluate the stability of the results. Firstly, for every clustering iteration k, the k-means procedure was run three times by choosing initial centroids differently, i.e., the first k, the last k, k randomly selected observations in sample. Secondly, the order of the gas transmission systems in the dataset was changed randomly and the clustering procedure was re-run to identify potential outliers that might influence the results. Finally, the VRC index proposed by Calinski and Harabasz [37] was adopted to select the correct number of groups (Appendix A). To this aim, the k-means algorithm was iteratively performed nine times ($k = 2, \ldots, 10$) to have a wide range of clustering solutions from which to choose the best one.

The clustering procedure identified seven configurations. Table A1 in Appendix B presents the list of gas transmission systems classified by group. Figure 3 illustrates the output of the VRC analysis. At $k = 7$ VRC is 121.17 and ω is -57.39 which are the higher VRC and the lower ω values obtained by iterating the clustering procedure for different k. Table 4 provides summary statistics for all cases, while Table 5 reports main statistics for each configuration. The big size of the F statistics in the analysis of variance shows that the two clustering variables are statistically significant and relevant to identify homogeneous groups of gas transmission systems. Configurations differ with respect to the number of components and the efficiency profile. The smallest configuration—E—includes only 3 gas transmission systems that achieve high efficiency scores in both production and market orientation perspectives. That is consistent with the graphic plot in Figure 2, supporting the idea that "no trade-off situations" in which different performance goals are indeed compatible can actually be found in the US gas transmission market, although they are rare. This group includes only 3.75% of the total number of sample systems. These gas transmission systems can be considered excellent. Nonetheless, even important "trade-off situations" are not frequent. Indeed, figures in Table 5 indicate that only in configurations C and G there are critical trade-offs between cost and revenue-efficiency scores. These groups contain 8 and 7 gas transmission systems, respectively 10% and 8.75% of sample. The largest number of gas transmission systems in the sample achieves low efficiency rates in both segments of the multi-stage production model, as in the case of configurations A and B totally including half of the sample. As these transmission systems are low performing in comparison to other systems in the sample, the operating companies should carry on a more in-depth technical and business analysis to identify determinants of scarce efficiency. Finally, there are two configurations—D and F—in which a partial trade-off exists between the two efficiency measurements. In Figure 4, the plot of costEff and revEff score centroids relatively to each configuration clearly emphasizes the different situations.
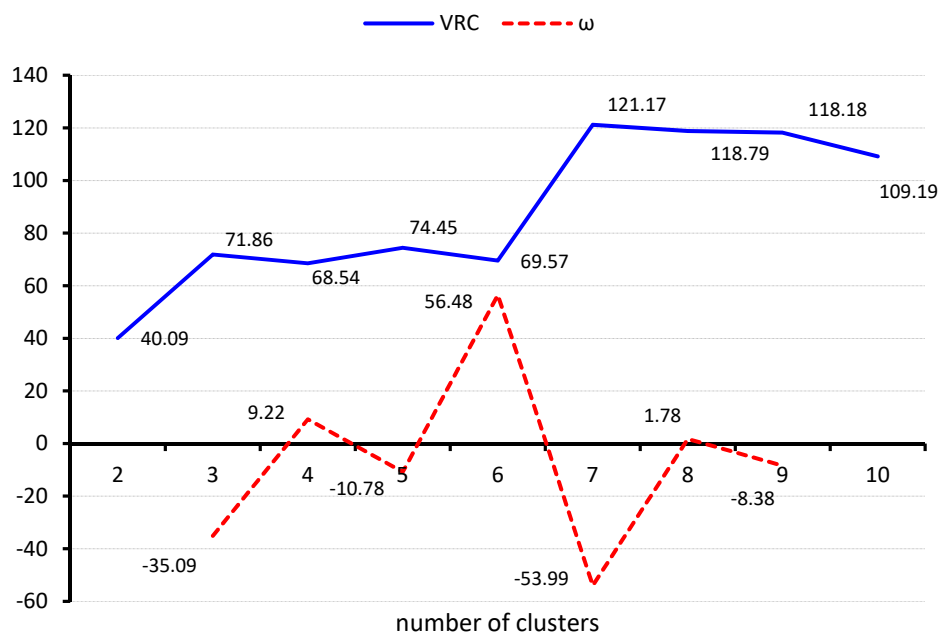
Figure 3. *VRC* and ω analysis.

Table 4. Analysis of variance.

	Between Sum of Squares	df	Within Sum of Squares	df	F
costEff	5.315	6	0.645	73	100.226
revEff	5.222	6	0.413	73	154.017
total	10.538	12	1.058	146	

Table 5. Statistics of clustering variables relative to pipeline infrastructure configurations.

Configuration	CostEff				RevEff			
	Mean	St.dev	Max	Min	Mean	St.dev	Max	Min
A (# 18)	0.122	0.056	0.220	0.048	0.310	0.068	0.444	0.182
B (# 22)	0.202	0.081	0.326	0.069	0.090	0.047	0.163	0.016
C (# 8)	0.922	0.097	1.000	0.794	0.143	0.093	0.259	0.007
D (# 11)	0.485	0.090	0.618	0.349	0.125	0.084	0.246	0.019
E (# 3)	0.843	0.272	1.000	0.529	0.878	0.108	1.000	0.793
F (# 11)	0.292	0.111	0.478	0.115	0.469	0.074	0.566	0.364
G (# 7)	0.120	0.079	0.210	0.018	0.886	0.113	1.000	0.738

In order to understand why some configurations are more efficient than others and are not related to trade-off situations, additional investigation was carried on taking into account the DEA model variables and a set of key structure and performance indicators (KSPIs). Specifically, KSPIs were obtained as ratios by dividing the following model variables by the transmission system length: no. of compression stations, gas volume trans. for others, operating & maintenance expenses, and operating revenue. Measurements relative to each KSPI were normalized by dividing them by the maximum to have scores in the range 0–1. The utilization of such indicators allows having useful information about the structural and performance characteristics of the gas transmission infrastructure independently of the transmission network size. Table 6 indicates that, on average, the gas transmission systems belonging to configuration E have a larger structural and operational size. The average infrastructure length is 10,363 miles and the average number of compression stations is 66, whereas the average volume of gas transmitted is 2,885,485 MMcf. Table 7 shows that gas transmission systems belonging

to configuration E have the lower number of compression stations per mile. However, the other KSPIs values, as a whole, do not seem to indicate that infrastructure systems in this configuration achieve better operational performance than systems belonging to the other configurations.
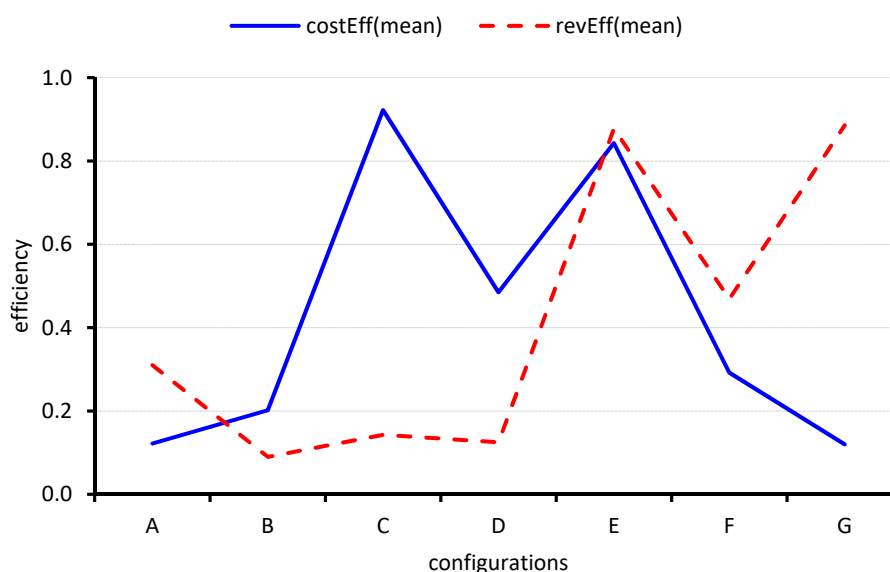

Figure 4. Plot of costEff and revEff centroids (means) relatively to the seven configurations.

Table 6. Statistics of DEA model variables relative to sample configurations.

Variable	Configuration						
	A	**B**	**C**	**D**	**E**	**F**	**G**
transmission system length	2346 (2088)	712 (822)	483 (442)	562 (862)	10,363 (1590)	6174 (4564)	3335 (3277)
no. of compression stations	21.28 (14.61)	6.09 (5.11)	5.63 (6.02)	4.00 (5.66)	66.00 (15.52)	35.09 (21.88)	39.00 (47.27)
gas volume trans. for others	415,109 (262,579)	163,719 (161,739)	558,976 (392,999)	261,480 (268,636)	2,885,485 (1,040,440)	917,942 (437,683)	572,162 (523,767)
operating&maintenance expenses	76,801 (57,448)	19,908 (17,029)	17,761 (17,436)	13,231 (15,512)	412,666 (142,212)	150,063 (112,134)	202,006 (198,205)
operating revenue	214,362 (112,622)	44,210 (35,571)	132,445 (107,131)	79,385 (70,067)	1,118,662 (221,028)	454,184 (122,353)	622,754 (395,967)

Legend: Table contains mean and standard deviation measurements relative to each configuration. Standard deviation values are in brackets.

As discussed above, configurations A and B exhibit low efficiency measurements. Both groups include gas transmission systems which, on average, have different size—2346 and 712 miles—respectively, that is lower than systems in group E. Nevertheless, they have a high number of compression stations per mile along the transmission network and greater operations and maintenance expenses per mile than systems belonging to other configurations. Configurations C and G for which Figure 4 clearly highlighted an important trade-off between the two efficiency dimensions contain infrastructure systems that on average have rather different lengths (483 and 3335 miles), too. Anyhow, both configurations possess a similar ratio of number of compression stations to gas transmission network length (0.65 vs. 0.73). Furthermore, figures in Table 7 emphasize the considerable difference relative to the remaining KSPIs consistently with the configurations efficiency indexes and trade-off typologies. Configurations D and F differ significantly with respect to the average length of their transmission systems, respectively 562 and 6174 miles. Howsoever, the KSPIs measurements in configuration F are on average lower than those in configuration D, especially the

first two. While in these configurations there is a partial efficiency trade-off, on average infrastructure systems in configuration F achieve higher efficiency scores whereas the trade-off between the two efficiency dimensions is decidedly smaller. Figure A1 in Appendix C summarizes information relative to individual configurations.

Table 7. Means of KSPIs measurements relative to sample configurations.

KSPI	Configuration						
	A	**B**	**C**	**D**	**E**	**F**	**G**
no. of compression stations per mile	0.79	1.00	0.65	0.68	0.40	0.46	0.73
gas volume trans. for others per mile	0.26	0.31	1.00	0.67	0.22	0.30	0.20
Operating & maintenance expenses per mile	0.28	0.25	0.18	0.24	0.23	0.17	1.00
operating revenue per mile	0.36	0.26	0.58	0.54	0.21	0.49	1.00

The previous analysis suggests that the transmission network size may have an important weight in the occurrence of trade-off situations, and, particularly increasing length may reduce potential trade-offs and improve performance at the same time.

7. Conclusions

The primary objective of the study presented in this paper was to measure efficiency of the transmission segment of the US natural gas industry. To address this research objective, a multi-stage modeling approach based on the implementation of DEA was adopted and the efficiency of a sample containing 80 US gas transmission systems was measured in a two-dimension performance space, i.e., cost-efficiency and revenue-efficiency. The proposed multi-stage model of the gas transmission system is more effective than the traditional black-box approach as it allows understanding more effectively how different types of inputs are sequentially utilized and converted into outputs and accounts for different technical and business goals—many times conflicting—that the operating companies have to deal with. Hence, the approach is especially apt to conduct benchmarking and efficiency studies as it allows analyzing the relationship between several efficiency measurements.

Utilizing cost and revenue efficiencies as grouping variables, cluster analysis was employed to identify homogeneous configurations of gas transmission systems having similar efficiency profiles and investigate the existence of efficiency trade-offs. Furthermore, key structure and performance indicators (KSPIs) were measured to have insights about the configurations extracted from sample.

Findings indicate that the transmission segment of the US natural gas industry has considerable inefficiencies, and average cost and revenue-efficiency scores are 0.324 and 0.301 respectively. Only 3 systems achieved high scores on both efficiency dimensions. Cluster analysis uncovered totally 7 configurations showing that in 3 of them there are no trade-off situations between cost and revenue efficiencies. However, only in one of such configurations gas transmission systems achieve high efficiency rates. The remaining four configurations exhibit trade-off situations having different intensity. Finally, the analysis of KSPIs suggests that trade-offs can be determined by the transmission infrastructure size.

Although this study is explorative in nature, it gives important contributions to literature on benchmarking and efficiency analysis in the energy industry, offers insights concerning performance trade-offs to the managers of the gas transmission operating companies, and useful information for policy and industry regulation. Particularly, empirical results have showed that the use of individual indicators rather than a comprehensive efficiency measurement may be misleading as provides not consistent indications. At the same time, the adoption of a multiple dimension efficiency modeling framework capable to deal with conflicting technical and business goals may help to identify trade-off situations.

The production model developed to measure the efficiency of the gas transmission infrastructure might benefit from the inclusion of additional variables. The efficiency score in the market oriented

segment of the gas transmission infrastructure production model is influenced by the natural gas price. Indeed, the revenues of the companies that operate the gas transmission infrastructure systems considerably depend on gas prices which can vary greatly in the US territory. Such differences are determined by the interaction of several factors, i.e., the distance from the area where the gas is produced to the area where it is utilized, the availability and capacity of the transmission pipeline to move gas from the producing areas, the physical storage capacity and trading hubs, state regulations, the level of direct and indirect competition, and volatility of gas consumption. Storage can play an important role in mitigating price volatility together with the adoption of contractual agreements and financial hedging instruments. In the evaluation of the infrastructure system efficiency, the physical storage of gas is as relevant as the pipeline and the compression facilities because it plays a key role as a mechanism for providing flexibility in the market, and finally, influencing short-term gas price volatility [38]. Both interstate and intrastate gas transmission companies rely on gas storage to maintain contractual balance, perform load balancing in order to preserve operational integrity of the pipeline, and regulate the level of gas supply over periods of fluctuating demand. Increasing the storing capacity usually requires large investment and storing gas may be relatively expensive and risky because of gas price volatility. Hence, the storage capacity of the gas transmission infrastructure system is an important variable that should be included in the efficiency analysis. Additional research might consider how the gas transmission infrastructure system efficiency is influenced by the interaction between price volatility and storage level. This study has not taken into account the storage capacity provided by operators of the gas infrastructure systems or independent companies as data were not available for such variable for most of the gas transmission systems included in the sample.

Future research should account for heterogeneities in the sample. This study neither distinguished between "intra-state" and "inter-state" gas transmission systems, nor considered the ownership structure of the operating companies. These heterogeneities can be important moderating factors of the efficiency-profile configuration-structure relationship. To deal with such heterogeneities a meta-frontier SBM DEA modeling approach might be adopted [39]. The efficiency analysis used input and output data relative to fiscal year 2012. The utilization of a dataset covering a wider time interval is necessary to confirm the results emerged from the configuration analysis because changes in the industry structure and demand in the energy market affect how inputs and outputs can be efficiently combined and, finally, influence the performance of the gas transmission infrastructure. Dynamic analysis also helps account for the influence of technical progress on efficiency [40].

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

The VRC index has been introduced by Calinski and Harabasz in 1974 to determine the correct number of groups in cluster analysis [37]. Milligan and Cooper [41] proved that the utilization of the VRC index to choose the optimal clustering solution works generally well.

If n is the number of data objects to be clustered and k is the number of groups obtained, the VRC_k index is given by:

$$VRC_k = \frac{\frac{SS_B}{(k-1)}}{\frac{SS_W}{(n-k)}} \tag{A1}$$

where SS_B and SS_W are the overall between-group sum of squares and within-group sum of squares.

To determine the correct number of groups, for each clustering solution k the additional index ω_k is computed as follows:

$$\omega_k = (VRC_{k+1} - VRC_k) - (VRC_k - VRC_{k-1}) \tag{A2}$$

The optimal number of groups is chosen finding a value of k that maximizes VRC_k and minimizes ω_k.

Appendix B

Table A1. List of natural gas infrastructure systems grouped by configuration.

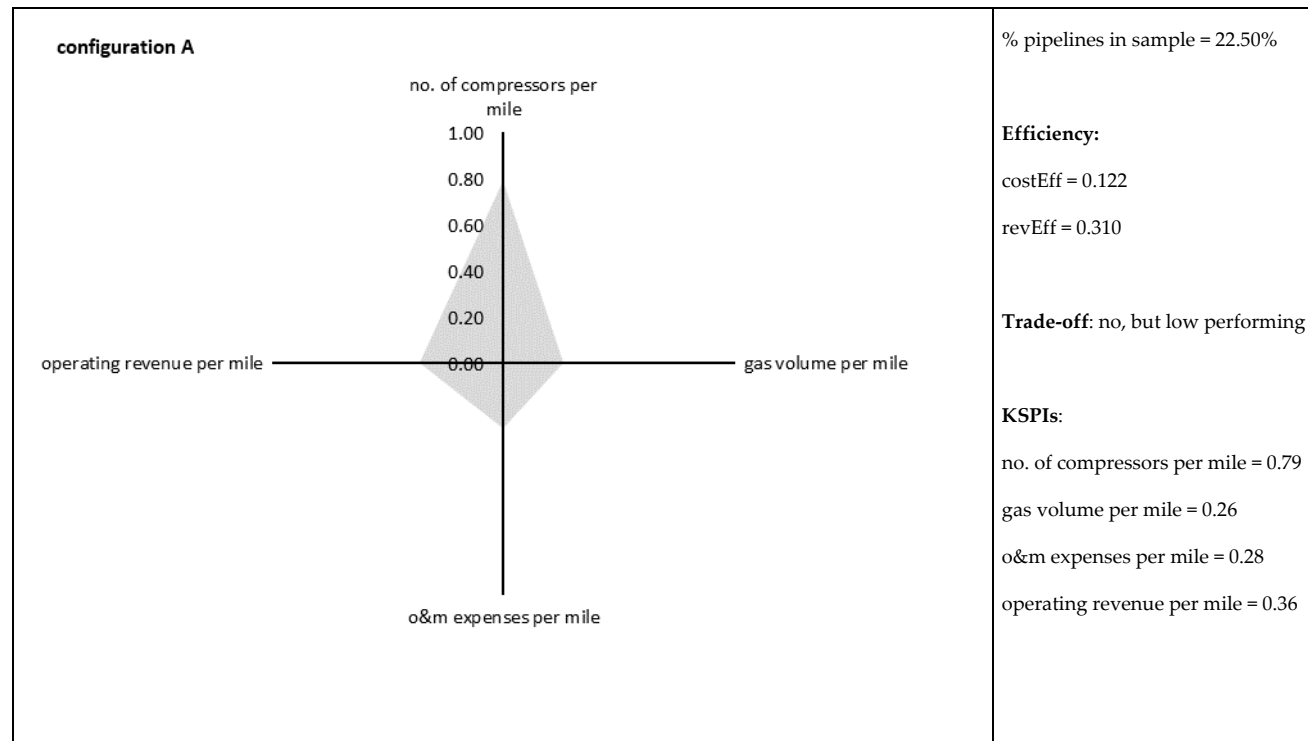| Name	Configuration	Name	Configuration
Algonquin GAS Transmission LLC	A	Crossroads Pipeline Co	C
Alliance Pipeline LP	A	Fayetteville Express Pipeline LLC	C
Cheyenne Plains Gas Pipeline Co. LLC	A	Midcontinent Express Pipeline LLC	C
Colorado Interstate Gas Co.	A	Northern Border Pipeline Co.	C
Dauphin Islands Gathering Partners	A	Ok Tex Pipeline Co. LLC	C
East Tennessee Natural Gas LLC	A	Questar Overthrust Pipeline Co.	C
Enable Gas Transmission	A	Vector Pipeline LP	C
Equitrans LP	A	Wyoming Interstate Co Ltd.	C
Gulf Crossing Pipeline LLC	A	American Midstream (Ala Tenn) LLC	D
Iroquois Gas Transmission Systems LP	A	Elba Express Co. LLC	D
Maritimes & Northeast Pipeline LLC	A	Empire Pipeline Inc	D
National Gas Fuel Supply Corp.	A	Garden Banks Gas Pipeline LLC	D
Panhandle Eastern Pipe Line Co. LP	A	Gas Transmission Northwest Corp	D
Questar Pipeline Co.	A	Horizon Pipeline Co. LLC	D
Southern Star Central Gas Pipeline Inc.	A	Millenium Pipeline Co. LLC	D
Tallgrass Interstate Gas Transmission LLC	A	Mojave Pipeline Co.	D
Transwestern Pipeline Co. LLC	A	North Baja Pipeline LLC	D
WBI Energy Transmission Inc.	A	Southeast Suppy Header LLC	D
Big Sandy Pipeline LLC	B	Trunkine Gas Co. LLC	D
Carolina Gas Transmission Corp.	B	Tennessee Gas Pipeline Co.	E
Cimarron River Pipeline LLC	B	Texas Eastern Transmission LP.	E
Columbia Gulf Transmission Co	B	Transcontinental Gas Pipe Line Co. LLC	E
Destin Pipeline Co. LLC	B	ANR Pipeline Co.	F
Discovery Gas Transmission LLC	B	El Paso Natural Gas Co	F
Enable Mississippi River Transmission LLC	B	ETC Tiger Pipeline LLC	F
Great Lakes Gas Transmission LP	B	Gulf South Pipeline Co. LP	F
Guardian Pipeline LLC	B	Gulfstream Natural Gas System LLC	F
Midwestern Gas Transmission Co.	B	Natural Gas Pipeline Co. of America	F
MIGC Inc	B	Northern Natural Gas Co.	F
Ozark Gas Transmission LLC	B	Northwest Pipeline LLC	F
Paiute Pipeline Co.	B	Ruby Pipeline LLC	F
Petal Gas Storage LLC	B	Southern Natural Gas Co.	F
Questar Southern Trails Pipeline Co.	B	Texas Gas Transmission LLC	F
Sabine Pipe Line LLC	B	Columbia Gas Transmission LLC	G
Sea Robine Pipeline Co. LLC	B	Dominion Cove Point LNG	G
TC Offshore LLC	B	Dominion Transmission Inc.	G
Trailblazer Pipeline Co.	B	Florida Gas Transmission Co. LLC	G
TransColorado Gas Transmission Co.	B	Kern River Gas Transmission Co.	G
Tuscarora Gas Tranmission Co.	B	KPC Pipeline LLC	G
Viking Gas Transmission Co.	B	Rockies Express Pipeline LLC	G

Appendix C

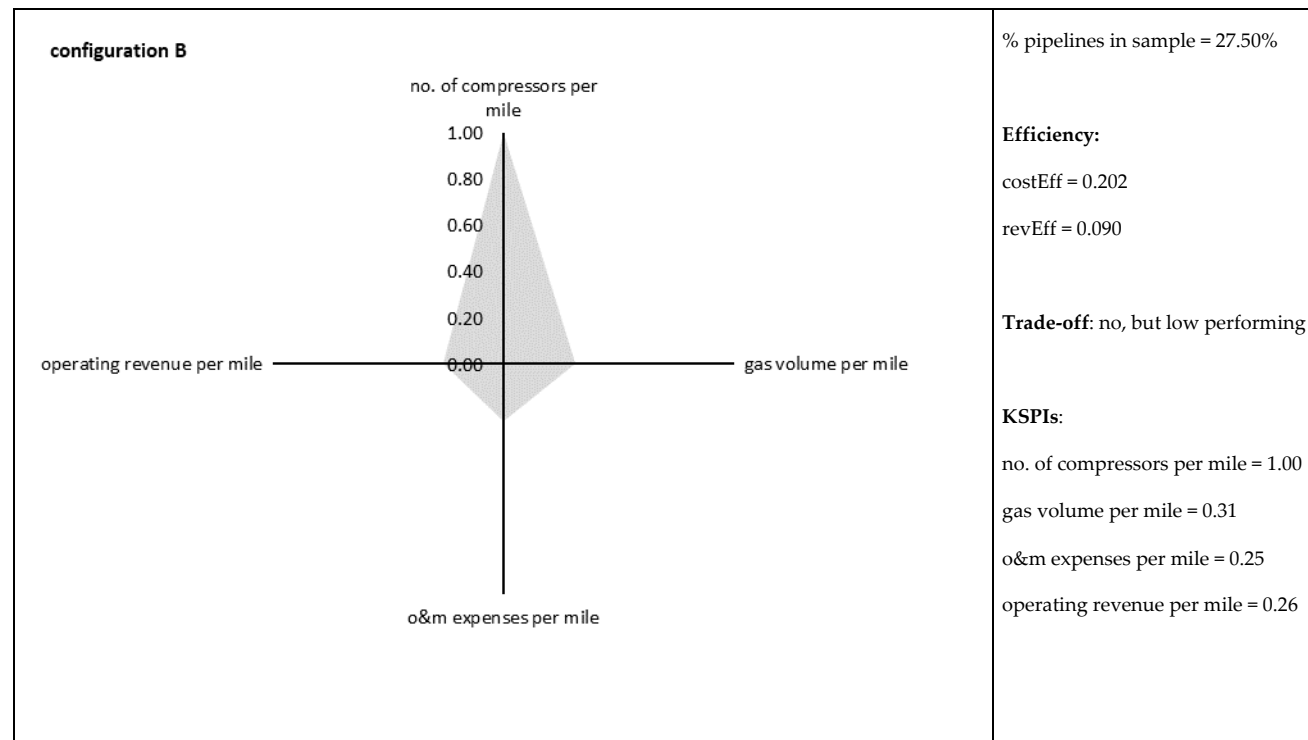**Figure A1.** Comparison of gas transmission configurations.
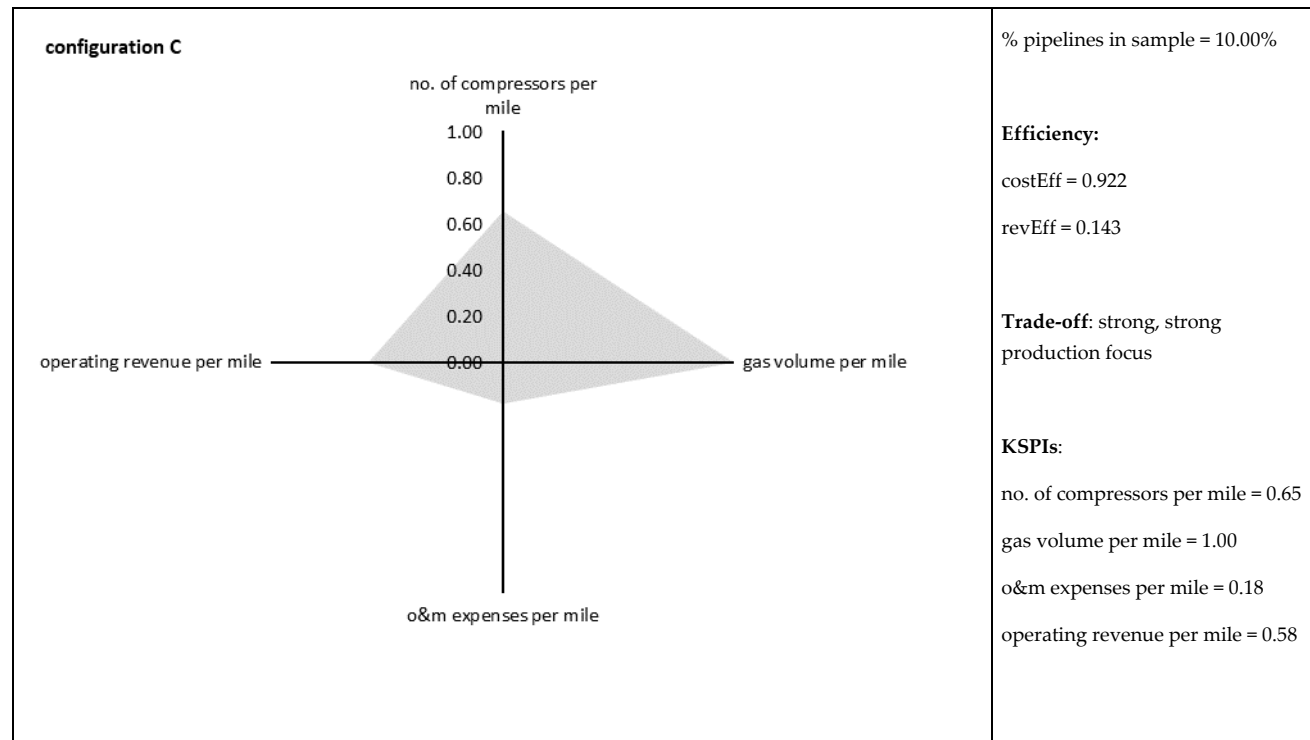
Figure A1. *Cont.*

Figure A1. *Cont.*
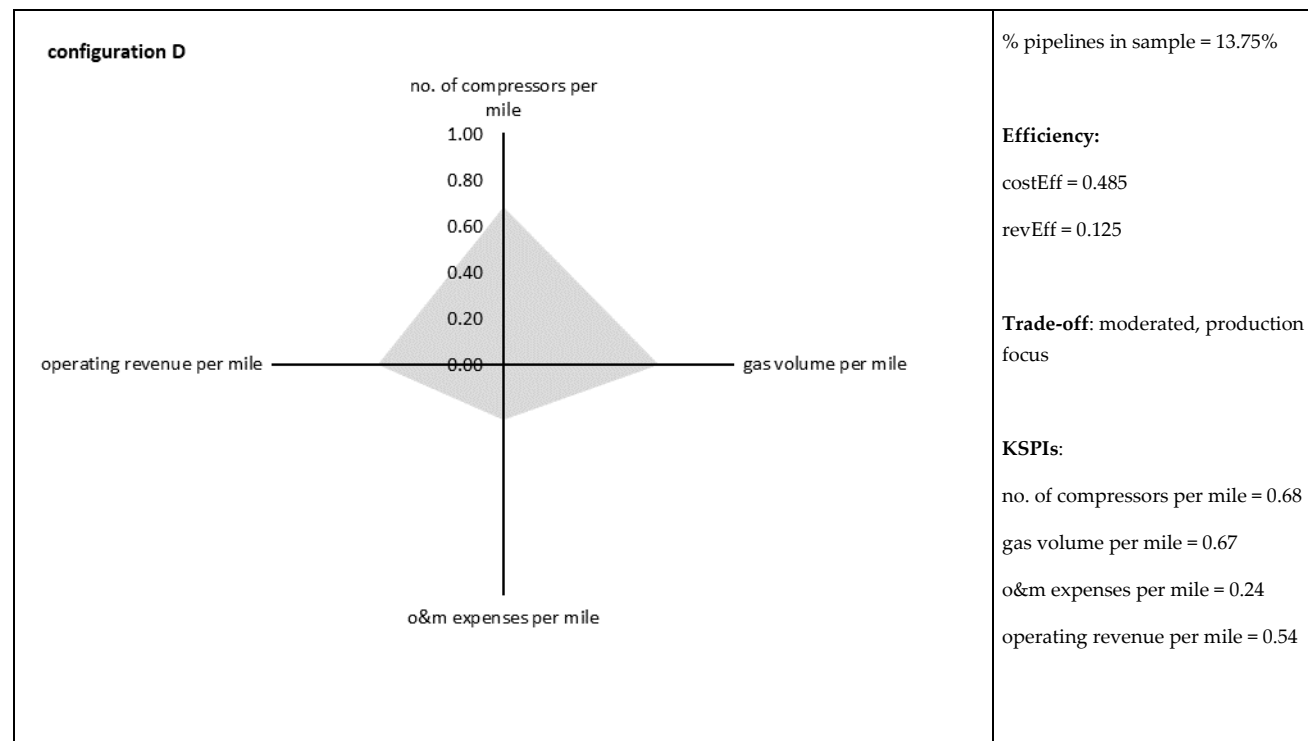
Figure A1. *Cont.*

configuration E

no. of compressors per mile

1.00
0.80
0.60
0.40
0.20
0.00

operating revenue per mile — gas volume per mile

o&m expenses per mile

% pipelines in sample = 3.75%

Efficiency:

costEff = 0.843

revEff = 0.878

Trade-off: no, high performing

KSPIs:

no. of compressors per mile = 0.40

gas volume per mile = 0.22

o&m expenses per mile = 0.23

operating revenue per mile = 0.21

Figure A1. *Cont.*

Figure contents:

configuration F

no. of compressors per mile

1.00
0.80
0.60
0.40
0.20
0.00

operating revenue per mile — gas volume per mile

o&m expenses per mile

% pipelines in sample = 13.75%

Efficiency:

costEff = 0.292

revEff = 0.469

Trade-off: very small, market orientation

KSPIs:

no. of compressors per mile = 0.46

gas volume per mile = 0.30

o&m expenses per mile = 0.17

operating revenue per mile = 0.49

Figure A1. *Cont.*

configuration G

no. of compressors per mile

1.00
0.80
0.60
0.40
0.20
0.00

operating revenue per mile

gas volume per mile

o&m expenses per mile

% pipelines in sample = 8.75%

Efficiency:

costEff = 0.120

revEff = 0.886

Trade-off: strong, market orientation

KSPIs:

no. of compressors per mile = 0.73

gas volume per mile = 0.20

o&m expenses per mile = 1.00
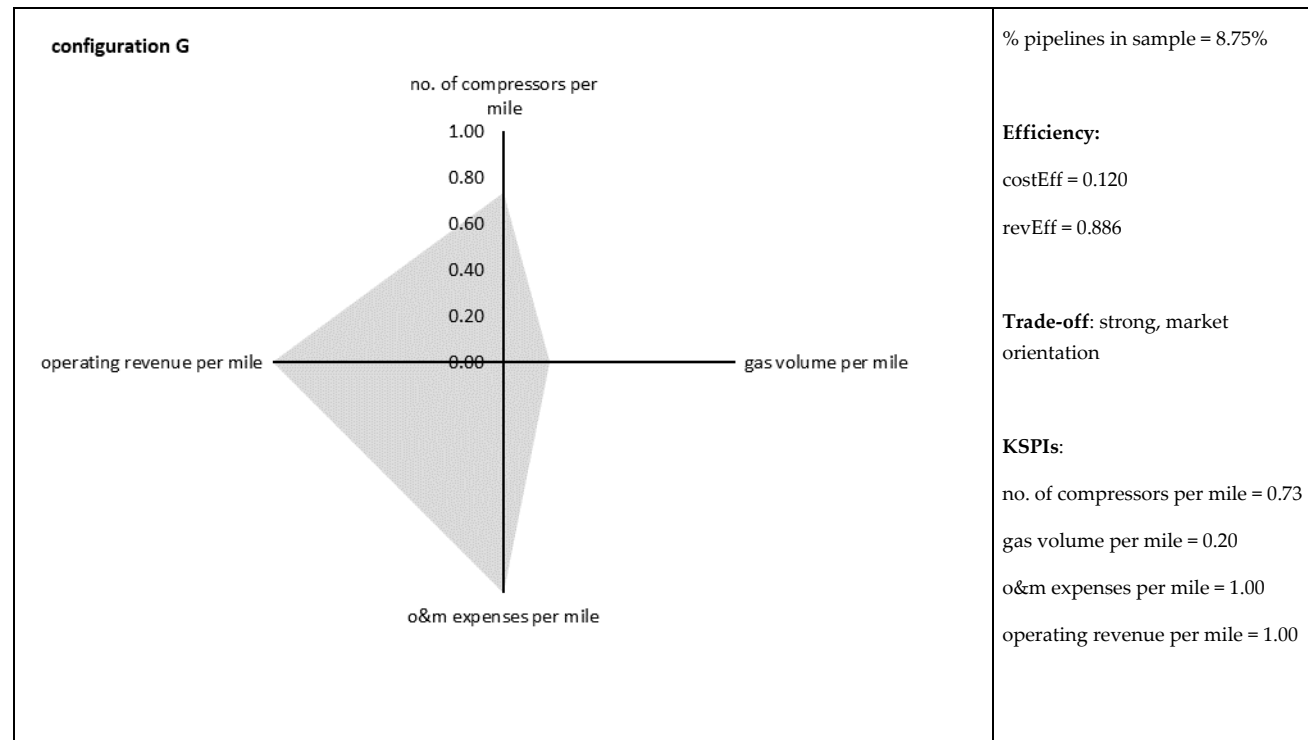
operating revenue per mile = 1.00

Figure A1. *Cont.*

References and Notes

1. Statistics & Analysis: U.S. Natural Gas Total Consumption. U.S. Energy Information Administration (EIA), 2017. Available online: https://www.eia.gov/dnav/ng/hist/n9140us2A.htm (accessed on 27 December 2017).

2. Statistics & Analysis: Natural Gas Consumption by End Use. U.S. Energy Information Administration (EIA), 2017. Available online: https://www.eia.gov/dnav/ng/NG_CONS_SUM_DCU_NUS_A.htm (accessed on 27 December 2017).

3. Statistics & Analysis: U.S. Natural Gas Marketed Production. U.S. Energy Information Administration (EIA), 2017. Available online: https://www.eia.gov/dnav/ng/hist/n9050us2a.htm (accessed on 27 December 2017).

4. Statistics & Analysis: U.S. Natural Gas Citygate Price. U.S. Energy Information Administration (EIA), 2017. Available online: https://www.eia.gov/dnav/ng/hist/n3050us3a.htm (accessed on 27 December 2017).

5. Supervisory Control and Data Acquisition.

6. The high interconnection and redundancy of the network makes the natural gas infrastructure particularly resilient and offers a great flexibility in moving gas.

7. National Action Plan for Energy Efficiency—Energy Efficiency as a Low-Cost Resource for Achieving Carbon Emissions Reductions. Prepared by William Prindle, ICF International, Inc.; 2009. Available online: https://www.epa.gov/sites/production/files/2015-08/documents/ee_and_carbon.pdf (accessed on 27 December 2017).

8. Lo Storto, C.; Ferruzzi, G. Benchmarking Economical Efficiency of Renewable Energy Power Plants: A Data Envelopment Analysis Approach. *Adv. Mater. Res.* **2013**, *772*, 699–704. [CrossRef]

9. Mardani, A.; Zavadskas, E.K.; Streimikiene, D.; Jusoh, A.; Khoshnoudi, M. A comprehensive review of data envelopment analysis (DEA) approach in energy efficiency. *Renew. Sustain. Energy Rev.* **2017**, *70*, 1298–1322. [CrossRef]

10. Sueyoshi, T.; Yuan, Y.; Goto, M. A literature study for DEA applied to energy and environment. *Energy Econ.* **2017**, *62*, 104–124. [CrossRef]

11. Zhou, P.; Ang, B.W.; Poh, K.L. A survey of data envelopment analysis in energy and environmental studies. *Eur. J. Oper. Res.* **2008**, *189*, 1–18. [CrossRef]

12. Hollas, D.R.; Macleod, K.R.; Stansell, S.R. A data envelopment analysis of gas utilities' efficiency. *J. Econ. Financ.* **2002**, *26*, 123–137. [CrossRef]

13. Hawdon, D. Efficiency, performance and regulation of the international gas industry—A bootstrap DEA approach. *Energy Policy* **2003**, *31*, 1167–1178. [CrossRef]

14. Jamasb, T.; Pollitt, M.; Triebs, T. Productivity and efficiency of US gas transmission companies: A European regulatory perspective. *Energy Policy* **2008**, *36*, 3398–3412. [CrossRef]

15. Erbetta, F.; Rappuoli, L. Optimal scale in the Italian gas distribution industry using data envelopment analysis. *Omega* **2008**, *36*, 325–336. [CrossRef]

16. Goncharuk, A.G. Performance Benchmarking in Gas Distribution Industry. *Benchmarking* **2008**, *15*, 548–559. [CrossRef]

17. Zorić, J.; Hrovatin, N.; Scarsi, G. Gas Distribution Benchmarking of Utilities from Slovenia, the Netherlands and the UK: An Application of Data Envelopment Analysis. *South East Eur. J. Econ. Bus.* **2009**, *4*, 113–124. [CrossRef]

18. Amirteimoori, A.; Kordrostami, S. A Euclidean distance-based measure of efficiency in data envelopment analysis. *Optimization* **2010**, *59*, 985–996. [CrossRef]

19. Nieswand, M.; Cullmann, A.; Neumann, A. Overcoming Data Limitations in Nonparametric Benchmarking: Applying PCA-DEA to Natural Gas Transmission. *Rev. Netw. Econ.* **2011**, *9*. [CrossRef]

20. Ertürk, M.; Türüt-Aşık, S. Efficiency analysis of Turkish natural gas distribution companies by using data envelopment analysis method. *Energy Policy* **2011**, *39*, 1426–1438. [CrossRef]

21. Sadjadi, S.J.; Omrani, H.; Abdollahzadeh, S.; Alinaghian, M.; Mohammadi, H. A robust super-efficiency data envelopment analysis model for ranking of provincial gas companies in Iran. *Expert Syst. Appl.* **2011**, *38*, 10875–10881. [CrossRef]

22. Marques, V.; Almeida, P.; Cunha, M.; Paço, M.; Rocha, M.; Trindade, A. What Drives Efficiency on the Portuguese Gas Distribution? In Proceedings of the 9th International Conference on the European Energy Market, Florence, Italy, 10–12 May 2012.

23. Lo Storto, C. Gas Distribution in Italy: A Non-Parametric Analysis of Companies Operational Efficiency. *Adv. Mater. Res.* **2014**, *838–841*, 1972–1978. [CrossRef]

24. Yardimci, O.; Karan, M.B. Efficiency and service quality analyses of the natural gas distribution companies: A case study of Turkey. In *Energy Technology and Valuation Issues*; Part III; Dorsman, A., Westerman, W., Simpson, J.L., Eds.; Springer: Cham, Switzerland, 2015; pp. 165–198.

25. Goncharuk, A.G.; lo Storto, C. Challenges and policy implications of gas reform in Italy and Ukraine: Evidence from a benchmarking analysis. *Energy Policy* **2017**, *101*, 456–466. [CrossRef]

26. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **1978**, *2*, 429–444. [CrossRef]

27. Cooper, W.W.; Seiford, L.M.; Tone, K. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*; Springer: New York, NY, USA, 2007.

28. Tone, K. A slacks-based measure of efficiency in data envelopment analysis. *Eur. J. Oper. Res.* **2001**, *130*, 498–509. [CrossRef]

29. Tone, K.; Tsutsui, M. Network DEA: A slacks-based measure approach. *Eur. J. Oper. Res.* **2009**, *197*, 243–252. [CrossRef]

30. Banker, R.D.; Charnes, A.; Cooper, W.W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag. Sci.* **1984**, *30*, 1078–1092. [CrossRef]

31. PennWell Corporation. In *Oil & Gas Journal*; PennWell Corporation: Houston, TX, USA, 2014; pp. 130–132.

32. Data were collected by reference [31] from FERC Forms 2 and 2A for major and non-major natural gas pipeline companies, reporting years 2012–13.

33. Daraio, C.; Simar, L. *Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications*; Springer: New York, NY, USA, 2007.

34. The eigenvalue inertia is 90.28% indicating that the new variable provides a good approximation of the original variables, as the higher the inertia is the higher the information amount of the original variables summarized by the factor.

35. Ketchen, D.J.; Shook, C.L. The application of cluster analysis in strategic management research: An analysis and critique. *Strateg. Manag. J.* **1996**, *17*, 441–458. [CrossRef]

36. Everitt, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*; John Wiley & Sons: Chichester, UK, 2011.

37. Calinski, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat. Theory Methods* **1974**, *3*, 1–27. [CrossRef]

38. Fang, H.; Ciatto, A.; Brock, F. U.S. Natural Gas Storage Capacity and Utilization Outlook. Technical Report Prepared by ICF International for Oak Ridge National Laboratory. ORNL/TM-2016/273; 2016. Available online: https://energy.gov/sites/prod/files/2017/02/f34/U.S.%20Natural%20Gas%20Storage%20Capacity%20and%20Utilization%20Outlook.pdf (accessed on 27 December 2017).

39. Lo Storto, C. Ownership structure and the technical, cost, and revenue efficiency of Italian airports. *Util. Policy* **2018**, *50*, 175–193. [CrossRef]

40. Lo Storto, C.; Capano, B. Productivity changes of the renewable energy installed capacity: An empirical study relating to 31 European countries between 2002 and 2011. *Energy Educ. Sci. Technol. Part A* **2014**, *32*, 3061–3072.

41. Milligan, G.W.; Cooper, M. An examination of procedures for determining the number of clusters in a data set. *Psychometrika* **1985**, *50*, 159–179. [CrossRef]