CENTRAL EUROPEAN REVIEW OF ECONOMICS AND MANAGEMENT

Volume 2, Number 3
September 2018

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

Vol. 2, No. 3

Wydawnictwo: Wyższa Szkoła Bankowa we Wrocławiu, ul. Fabryczna 29-31, 53-609 Wrocław

Contact: Tomasz Rólczyński, redakcja@wsb.wroclaw.pl, +48 71 376 23 43, ul. Fabryczna 29-31, 53-609 Wrocław, Poland

The journal is reviewed according to the principle of double blind peer review, as well as in compliance with the standards of the Polish Ministry of Science and Higher Education. CEREM is a continuation of the WSB University in Wrocłąw Research Journal (Zeszyty Naukowe WSB we Wrocławiu – ISSN 1643-7772; eISSN 2392-1153), which is included in B list of the Polish Ministry of Science and Higher Education from 23.12.2015 (8 points).

CEREM is registered in the IC Journal Master List, and indexed in CEJSH, Index Copernicus and BazEkon

ISSN 2543-9472; eISSN 2544-0365
Cover and logo design: Sebprojekt.pl
Wydawnictwo: Wyższa Szkoła Bankowa we Wrocławiu, ul. Fabryczna 29-31, 53-609, Wrocław.

CONTENTS

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 7-41, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Accounts receivable overdue and market dynamics: a case study

Niek STAM, Wim WESTERMAN

Groningen University, The Netherlands

Abstract:

Aim: This study aims to find out which market dynamics are currently relevant for accounts receivable levels (specifically overdues), and how lessons learned can be used by credit management.

Design / Research methods: The unique research strategy is characterised as a single case study with design elements at two country units of a company to be named FEED. The classification and overview of relevant market dynamics provide valuable insights for determinants and intercompany differences in receivables, and whether these arise at the country or market level.

Conclusions / findings: The classification and overview of relevant market dynamics provide valuable insights for determinants and intercompany differences in receivables, and whether these arise at the country or market level.

Originality / value of the article: The findings suggest adjustments of the literature in that the interest rate is currently not a relevant factor. Moreover, instead of focussing on costs of capital, an emphasis on default risk is more applicable nowadays, and hence researchers should focus on overdue instead of on receivables in general.

Key words: Accounts Receivable, Overdue, Market Dynamics, Case Study
JEL: G30, Q13

1. Introduction

To be called FEED, is a market leader in feed solutions. Production companies such as FEED generally have relatively large amounts of net working capital. They hold large sums of cash, have large inventories in the form of raw commodities, and have large receivables due to credit sales. Working capital management (WCM)

Correspondence address: Wim Westerman, University of Groningen, Faculty of Economics and Business, Nettelbosje 2, 9747 AE Groningen, The Netherlands. E-mail: . w.westerman@rug.nl
Received: 30.08.2017, 20.02.2018, Revised: 24.04.2018, Accepted: 01.05.2018
Doi: http://dx.doi.org/10.29015/cerem.678

efficiency has been relevant for these types of companies for a long time, but the ongoing global economic uncertainty may urge them to optimize working capital practices right now (REL Consultancy 2016). This especially counts for companies whose receivables are largely dependent on the uncertain economic situation of their customers, such as FEED operating in the dairy cow sector, where low milk prices can lead to payment difficulties for dairy farmers.

Similar to other production companies, a large proportion of sales at FEED are done on the basis of trade credit, which creates large receivables. In 2015, about one fifth of the receivables were overdue, slightly more than the company's net income. Obviously, overdues larger than net income pose risks for the business and uncertainty for investors when customers default on debt, especially considering the economic uncertainty in the dairy cow sector. Hence, indicating risky customers to prioritize the reduction of overdue is a relevant topic for FEED. Companies in various commodity industries face similar problems, think of wood and cement producers whose receivables are likely dependent on the economic situation of customers and the market demand for furniture and construction.

Furthermore, FEED has substantial differences regarding receivables and (especially) overdues levels between country units. This is very notable in the dairy cow sector in the Dutch and East German units. Although both areas are geographically close to each other and adopt direct sales to farms, there are a lot of differences between the two because of market dynamics. The Business Dictionary defines this term as follows: *"interaction between forces of demand and supply and the pricing signals they generate. In most free (open) markets any significant part of market dynamics is beyond the control of any firm or group"* (WebFinance Inc. 2016). Hence, as to receivables, this paper considers market dynamics as external factors beyond the control of FEED that influence the demand for trade credit (and hence receivables and overdues). Obviously, more companies endure this problem when they operate across country markets, where varying market dynamics influence receivables and overdues.

FEED is unsure about how market dynamics influence its receivables and overdues, and how these give rise to differences between country units. Recent literature establishes a variety of factors that can give an indication as to how these arise, among which the interest rate. However, contrary to previous literature (cf. Biais, Gollier 1997; Nilsen 2002; Filbeck, Krueger 2005), the low Dutch and German interest rate might currently be less determining for receivables levels. Furthermore, in standard corporate finance literature, receivables are often regarded as an investment in clients (Hillier et al. 2016). The costs thereof are relatively low nowadays, with the current low interest rates. Combined with the ongoing economic difficulties in many industries (such as the dairy cow sector), default risk may currently be more stringent for trade credit providers than their costs of capital.

Previous literature indicates that because of the late-payment penalty on trade credit, companies are unlikely to generate overdues except when they lack sufficient funds (Petersen, Rajan 1997). Hence, overdues specifically indicate the proportion of receivables with a high default risk. Therefore, a focus on the market dynamics that influence overdues might be appropriate, an area hardly studied thus far. This can give indications of what market dynamics are relevant at times of low costs of capital and high default risk, and how credit management should adapt to them. This can lead to an adjustment of the previously stated factors and will shed new light on whether certain factors are still relevant for receivables and credit decisions today. Moreover, a single case comparison between country units can indicate whether differences in receivables and overdues are market-specific or country-specific.

The objective of this study is to investigate which market dynamics are currently relevant for accounts receivable levels (specifically overdues), and how lessons learned can be used by credit management. Hence, this paper adopts a unique approach by reporting on a case study in the dairy cow sector with an internationally stratified company, thereby examining intercompany differences between two country units of FEED, namely the Dutch and the East German units. Doing so can uncover the respective relevance of country-specific and market-specific factors that drive accounts receivable overdue and subsequent procedures to prioritize risky customers to reduce receivables and overdue. By filling in the gaps indicated, this paper can contribute to the understanding of WCM in general.

The remainder of this paper is structured as follows. First, relevant literature is reviewed. Next, method and data are elaborated upon. Following, relevant market dynamics are discussed and translated towards country unit differences in accounts receivable overdues. Also, on the basis of previous lessons learned, a recommendation framework is formulated and the findings are discussed. Finally, conclusions for practice and academics are drawn.

2. Literature review

This section will elaborate on literature related to the factors that determine trade credit demand to provide initial guidance for relevant market dynamics (external determinants) that determine receivables and overdue levels (internal processes).

2.1. Accounts receivable overdues

Hillier et al. (2016) consider accounts receivable (AR) as an investment in customers by means of trade credit. The way companies manage trade credit largely influences the amount of their receivables and hence also their overdues.

The terms of sale are decided internally and are the first credit management factor that influence AR. Defining the credit period (the payment term) is part of the terms, which is generally determined by the probability of default, size of the account, and whether the goods are perishable (Hillier et al. 2016). The length of the payment term inevitably influences AR and days sales outstanding (DSO), and consequently overdues. The DSO is calculated as the accounts receivable divided by the one day revenue (REL Consultancy 2016). More efficiency is generated with a shorter DSO due to faster collection of receivables.

Cash terms are another part of the terms of sale. These terms incentivize customers to pay earlier and discourage to generate overdue (by means of penalties). Also, Hermes et al. (2016) confirm enforcement mechanisms as relevant determinants of trade credit. Next to this, the credit instrument (usually an invoice) determines how the terms of sale are communicated and executed. The collection policy determines how the receivables are collected.

According to Howorth and Reber (2003), overdues follow from trade credit demand, and therefore overdues can be viewed as part of AR. Consistent with Petersen and Rajan (1997), Pike and Cheng (2001) indicate that there is a higher likelihood of generating overdue when customers’ liquidity is weak. Hence, overdues specifically indicate the customers within AR with a relatively high default risk. It is therefore that they warrant specific attention.

2.2. Market dynamics

Filbeck and Krueger (2005) find large differences in working capital levels across time. They argue that these differences might be due to external economic factors, i.c. market dynamics. The literature provides several suggestions for market dynamics influencing receivables.

First, the five C's of credit are often used in standard literature (Hillier et al. 2016), as a method to determine a customer's creditworthiness and the risk of default. Hence, these factors influence trade credit: 1) character, 2) capacity, 3) capital, 4) collateral and 5) conditions. Character defines the willingness of the customer, or "mentality" to repay trade credit. Capacity defines how able a customer is to repay trade credit and can be measured in terms of operating cash flows that directly affect the customer's ability to cover receivables. Capital is defined as the reserves the customer has and influences the ability to repay trade credit. Collateral indicates the customer's assets that can be liquidated to fulfill its obligations. Finally, economic conditions influence the customer's ability to repay trade credit. Trivially, it can be challenging to collect AR when debtors are experiencing economic hard times.

Second, Biais and Gollier (1997) found that small companies increase their trade credit as a consequence of an increase in the interest rate. Nilsen (2002) found similar results for both small and large companies. Hailemariam (2001) studied multiple cases in Eritrea. In order to finance operations, managers preferred internally generated funds (e.g. trade credit and retained earnings) instead of bank loans due to high interest rates, except when operational losses prohibited this. Moreover, Filbeck and Krueger (2005) argued that higher interest rates would make it less beneficial for customers to fulfill payments early, thereby stretching the AR of collectors. However, EU short term interest rates EU have been dropping from 5.11% in October 2008 to -0.29% in July 2016 (OECD Finance Department 2016). Therefore, it might be that this market dynamic currently is less influential on receivables.

Third, Petersen and Rajan (1997) found that companies resort to use more trade credit when banks do not provide funding. A tradeoff is visible between the flexibility of banks and the amount of trade credit (and hence receivables). Furthermore, Biais and Gollier (1997) found that companies that lack the connections with banks for receiving loans resort to use more trade credit. Moreover, Howorth and Reber (2003) found that habitual late payments correlate positively with the difficulty of getting credit from banks.

Fourth, Petersen and Rajan (1997) found that more trade credit is given to larger companies, hence indicating a positive relationship between customer size and receivables. Peel, Wilson and Howorth (2000) found that large companies generally endure more late payments from customers. They did not investigate customers' size as a variable affecting receivables. Ng, Smith and Smith (1999), and Wilson and Summers (2002), link the size of the creditor to credit decisions, but they also do not investigate how customer size affects AR of the creditor.

Finally, Hermes et al. (2016) found that competition levels influence trade credit for wholesalers. This factor was also suggested by Filbeck and Krueger (2005). High levels of competition can decrease market power for suppliers. With relatively high market power, customers may demand longer payment terms and are more likely to switch to the competition. However, Hermes et al. (2012) indicated that trade credit prevents customers from switching. Similarly, Petersen and Rajan (1997) found that companies offering trade credit can get the advantage of gaining more customer information, increasing dependency of the customer to the supplier, and using the assets of the customer as collateral for AR.

Based on the above, major market dynamics influencing AR are as follows: character, capacity, collateral, capital, conditions, competition, interest rate, banks, and customer size. These factors provide a categorization and overview of relevant market dynamics to be indicated by the case study. Taken together, lessons can be learned for credit management to improve AR (reduce overdues) with a procedure that incorporates relevant market dynamics by prioritizing customers with the highest likelihood of default and generating overdue.

3. Method and data

The overarching methodological framework of this paper is an explorative case study with design elements. Yin (2014: 16) describes a case study as follows: "*A case study is an empirical inquiry that investigates a contemporary phenomenon (the "case") in depth and within its real-world context, especially when the boundaries between phenomenon and context may not be clearly evident*". A design study is meant to offer a solution or correction to a present problem. Its first (diagnosis) stage can be completed by a case study, during which the detection of the problem and the gathering of the data is executed. The aim of this research particularly asks for a case study, since the improvement of AR overdue practices based on lessons from market dynamics is a practical and unclear delimited phenomenon that can be affected by various factors. FEED's involvement in the dairy cow sector is especially interesting because of substantial differences in AR overdues in view of market dynamics.

A standard component of a case study is a protocol, in which an overview, data collection procedures, questions for data collection and guidelines for the case study report are described (Yin 2014). Whilst the protocol was updated frequently from new knowledge, sticking to it increased the reliability of this research. In analyzing the case, this research strives towards analytical generalization to generalize the findings towards notions that can be applicable to other cases (Yin 2014). The second research stage offers an improvement to diminish the problem. It is labelled as the design stage, during which specific recommendations for FEED AR overdue practices based on previous lessons from market dynamics are formulated.

3.1. Multiple data sources

A field researcher had the possibility to combine multiple sources of data when being intern at FEED for about half a year. In this way, information from different sources could be balanced and verified. According to Yin (2014) this is a major benefit of a case study in comparison to other research methods. The practice of triangulation is followed in that interviews, literature and internal documentation are combined to obtain a deep and complete understanding of the case. The collection of information from different sources should aim to corroborate the same finding, a method that is labeled by Yin as *converging lines of inquiry.* This is supported by Jick (1979), who states that a greater accuracy can be achieved in this way.

3.2. Interview method

A technique of snowball sampling was adopted, whereby interview subjects refer potential future subjects. Thereby, the subject automatically becomes an assistant in the research (Biernacki, Waldorf 1981). The FEED Finance Director Netherlands provided the first subjects to start with. After these subjects were interviewed, the amount of referrals increased until the study was completed. The information gathered during previous sessions determined the choice of topic for subsequent interviews.

In order to preserve flexibility, interviews were conducted in an unstructured way. This was needed because of subjects' different expertise and the unknown relevance of certain information during the first research stages. Open questions enable to get a broad sense of the case, because subjects are free to elaborate on whatever they think is relevant and can provide clarification if needed. In order to benefit from triangulation, results from previous interviews were checked during the next interviews. Notes were taken and discussed with the subject at the end of each interview. Thereby the validity of the data was preserved as well as possible.

The key interview findings were summarized in a data matrix, which gave an overview of the subjects, the interviews, key topics, and key findings. Pre-interviews were held to specify the case. Some of these remained unused. Moreover, some findings were omitted because of irrelevance to the analysis, or when statements given were not corroborated. After the pre-interviews, the standard interviews were held. Evaluation interviews with specialized staff, directors and executives were held at multiple occasions to verify key findings. Individual interview reports were anonymized in order to protect the so-called innocent. Multiple interviews could be conducted with the same subject if desired and usually lasted 1 to 1.5 hours. In total, 19 subjects were interviewed over 34 unstructured interviews.

3.3. Diagnosis stage

In the first stage of this study, existing literature, internal documentation and unstructured interviews were conducted in order to get a deeper understanding of the market dynamics that influence AR overdues. These interviews were designed to diagnose how differences in AR overdues between country units arise, and to recognize potential opportunities of improving AR overdue practices on the basis of lessons learned. Moreover, the interviews were used to verify previous findings from other data sources and provide feedback for the research. In this way, information was triangulated. Some company-specific information in this paper was slightly adjusted for confidentiality reasons, without affecting the argumentation and analysis.

3.4. Design stage

During the second stage of this study, a recommendation framework was designed to improve FEED AR overdue practices by prioritizing customers likely to generate overdue. This was done on the basis of the lessons learned during the diagnosis stage. Furthermore, evaluation of the framework was conducted by comparing it to previous literature, and to what extend the included factors were country-specific or market-specific. Finally, the main findings of the case study were highlighted and where possible generalized for both practical and academic purposes, thereby providing suggestions for future research.

4. Market dynamics overview

In this part, the first section of the diagnosis phase, an overview of relevant market dynamics influencing AR overdues at FEED is developed. An initial list of drivers was indicated by previous literature and complemented by other documentation and unstructured interviews.

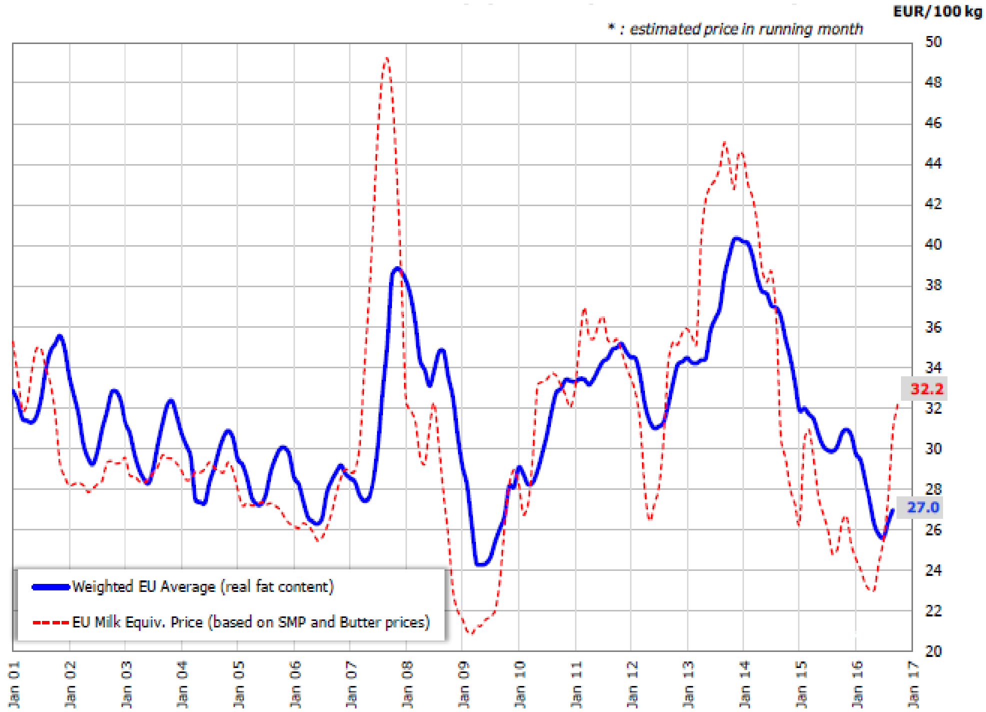
When investigating FEED AR overdues, the most relevant market dynamics can be classified as follows: 1) business cycle, 2) dairy farm financial structure, 3) farmer mentality, and 4) third party market participants. This categorization is as complete as possible, since no further categories that qualify to be considered as market dynamics were identified during the interviews. The categories and individual market dynamics were formulated iteratively by interviews and were verified during evaluating interviews with specialized senior staff and higher management. The market dynamics cannot be fully considered as independent determinants, since they can complement, reinforce or counteract each other.

4.1. Business cycle

Guided by previous literature, and based on multiple interviews, this study indicates the business cycle (more specifically the milk price) as the most relevant market dynamic for overdue in the FEED dairy cow sector. Milk prices reflect general economic conditions and are directly related to lower operating cash flows for dairy farmers, which is consistent with standard literature (Hillier et al. 2016) regarding conditions and capacity. The business cycle can be viewed as an agglutination of these factors, whereby a low milk price pressures and delays other expenses such as animal feed. The problem is illustrated by growing overdues, and a growing amount of customers asking for deferred payments during low milk price periods. Supported by multiple interviews, the business cycle is negatively related to FEED overdue. An anti-cyclical movement is confirmed (all else being equal) throughout.

Anticipating the abolishment of the milk quota in April 2015, milk prices for EU producers decreased significantly (Centraal Bureau Statistiek 2016), see figure 1. This is largely due to oversupply of milk in the Netherlands and Ireland. Together with a lagging demand of dairy products in China’s stagnant economy and a Russian boycott on EU agricultural products, this makes for an exceptional long period of low EU milk prices. According to specialized staff members, this is why many EU dairy farmers currently do not have the financial resilience to survive. This is supported by the claims in the sector, stating that the majority of dairy farmers did not allow for such an extreme situation in their liquidity planning. Hence, and supported by evaluating interviews, the business cycle is defined as the first market dynamic category.

Figure 1. Developments in the European raw milk price



Source: European Commission (2016).

4.2. Dairy farm financial structure

Guided by previous literature, based on various interviews, and supported during evaluating interviews, this paper now identifies a second relevant market dynamic category for overdues in the FEED dairy cow sector. This category comprises market dynamics related to the financial structure of farms and how this determines their abilities to fulfill FEED receivables. For convenience, two relevant subcategories are singled out: farm debt ratio and farm size.

Farm debt ratio. On the basis of multiple interviews, the debt ratio is identified as a relevant market dynamic related to financial structure, consistent with standard literature (Hillier et al. 2016). The debt ratio affects the capacity to repay trade credit, is proportional to the amount of collateral of a farm, and relates to its amount of capital reserves. First, debt-financed farms have relatively high fixed costs due to interest and rent expenses and have low capital reserves because of previous investments. This burdens other payments and generates overdue. Second, higher fixed costs raise the vulnerability for market fluctuations, e.g. a milk price decline. An EU study by Ernst & Young (2013) confirms this. Hence, the debt ratio market dynamic interacts with the business cycle, strengthening its influence on FEED overdue. It is also indicated that the degree of farms being financed by debt positively relates to overdues.

Farm size. Farm size in terms of number of animals is also viewed within FEED as a relevant market dynamic with respect to financial structure. This is confirmed by Petersen and Rajan (1997) regarding to size as a determinant of trade credit. Farm size interacts for a large part with farm debt ratio and the business cycle. First, large farms tend to be more growth oriented and are thus more likely to attract debt and endure liquidity shortages when confronted with lower milk prices. Generating overdue due to low milk prices is reinforced by heavy investments in dairy farms, increasing rent and interest expenses. Second, unlike smaller farms, large farms have more labor costs that cannot be reduced fast when needed, which adds to their financial inflexibility during economic troughs and increases the likelihood to generate overdue. Furthermore, trivially the number of animals is directly related to the amount of feed needed. FEED sells the majority of its feed on credit. Hence, the larger the farm, the higher the absolute amounts of overdue are generated. Thus, it can be stated that farm size positively relates to FEED overdue. It influences the financial structure of farms through the debt ratio and financial vulnerability to the business cycle.

Farmer mentality. Based on literature and various interviews, this study now indicates a third relevant market dynamics determinant for overdues in the FEED dairy cow sector. It refers to the farmers' personal thread of thinking and how this influences their way of fulfilling FEED receivables. Two subcategories are distinguished here: degree of input drive and payment mentality.

Degree of input drive. Multiple interviews identify input-oriented and output-oriented customers. Input driven customers mainly decide upon prices and payment terms. Output driven customers look beyond this and focus at bottom line output (milk quality and quantity per cow). The latter group makes use of sophisticated measurement techniques to calculate both the feed expenses and the according output. Hence, feed prices and payment terms can thus be compensated by a more than proportional output growth. It may be reasoned that due to their focus on payment terms, input driven customers extend their payments and are more likely to generate overdue. In principle, this market dynamic can be positively related to FEED overdue, but no clear further support for a relation between input drive and FEED overdue was identified.

Payment mentality. Customers differ in willingness to pay for animal feed, in line with the character factor (Hillier et al. 2016). The majority of customers have a good payment behavior. A small portion of the customers is responsible for the majority of payment problems. Collection is harder for this group, and these individuals are responsible for a continuous level of overdue. Furthermore, a worse payment mentality is linked to a higher DSO, and hence to higher overdue levels. Trivially, customer payment mentality negatively relates to FEED overdue.

4.3. Third party market participants

Guided by previous literature and based on various interviews, this paper identifies a fourth determinant for AR overdues in the FEED dairy cow sector. This category comprises market dynamics related to other market participants that influence the farmers' behavior as to FEED receivables. Three subcategories are delineated: FEED competitors, dairy factories and banks.

FEED competitors. In principle, competition levels affect the market power of FEED negatively. Higher levels of competition create more bargaining power for the customers to negotiate lower prices and longer payment terms. This is consistent with Hermes et al. (2012) and Hermes et al. (2016) and supported within FEED in that higher competition levels in general imply longer payment terms. A customer's tendency to switch to competitors is based upon the following aspects: 1) advisor quality and relationship, 2) feed quality, 3) other interactions (e.g. credit control), and 4) price and payment terms. In competitive markets, FEED needs to improve on these aspects in order to survive. FEED is thus forced to comply with competition standards in competitive markets with regards to payment terms and loosen AR overdue practices. Nevertheless, while this market dynamic can be positively related to FEED overdues and is therefore incorporated in the categorization, no clear further support for a real relationship between competition levels and FEED overdues could be identified.

Dairy factories. Farmers are inevitably dependent on the payment from dairy factories for the pickup of milk. Different factories adopt (slightly) different payment dates and frequencies, influencing the operating cash flows of the farmers and therefore their capacity to pay for other expenses such as animal feed. Hence, it can be reasoned that the longer it takes for dairy farmers to receive the "milk money" relative to the delivery of animal feed, the higher overdue (considering payment terms as equal overall). Whereas this market dynamic is positively related to FEED overdue and therefore is incorporated in the categorization, no clear further support for a real relationship between dairy factory payment date and FEED overdue could be identified.

Banks. FEED interviewees tell that banks are important third parties influencing the financial structure of farms. When banks are less flexible in their credit providence towards customers, the latter group is more likely to have liquidity problems and thus to generate overdue. This is in line with Petersen and Rajan (1997) and Howorth and Reber (2003). Banks are currently more precautious in credit providence towards dairy farms. First this is due to increased EU dairy cow sector regulation from 2007 onwards, leading to more financial uncertainty, as is generally confirmed in the sector. Second, the liquidity of dairy farms has fallen significantly due to the current low milk prices in the EU. The above mentioned mechanism is particularly prevalent during periods of economic difficulty, illustrating once again the interdependency of market dynamics as to financial structure, the business cycle and bank flexibility on overdue.

5. From market dynamics to AR differences

In this section, covering the second part of the diagnosis phase, differences in AR are elaborated first. Second, the most relevant market dynamics responsible for these differences are indicated in order to provide guidance for the recommendation framework.

5.1 Differences in accounts receivable

Interviews point at significant differences between the Dutch and East German unit with respect to AR (and especially overdues) which are particularly prevalent in the dairy cow sector, thereby supporting the relevance of this case study. The accounts receivable ratio for the Dutch unit is less than half of the accounts receivable ratio for the East German unit. The DSO for the Dutch unit was also much lower than for the East German unit, indicating a considerable discrepancy. The same counts for the overdue ratio. The data shows a clear reduction in overdues for the East German unit though. It is indicated that high 2013 and 2014 milk prices partially explain the reduction of overdue in these years extending towards most of 2015, and lower milk prices in 2015 and 2016 led to an overdue increase. These "lags" are caused by gradual improvement (deterioration) of private financial buffers of farmers, following an increase (decrease) of the milk price and influencing their payment abilities.

Obviously, overdues would be poor indicators if payment terms differ significantly. However, there are no significant differences. Since the East German unit orders relatively more raw material feed due to the larger size of farms in this area (large farms can mix the raw material into full feed themselves), this can partially contribute to the larger DSO for this country unit. All else being equal, these slight differences in payment terms cannot be the full explanation of the major discrepancies in DSO and overdue.

A higher DSO and overdue can actually contribute to FEED's performance in case it contributes to customer retention. This is consistent with findings from Hermes et al. (2013). However, the healthy financial lever in the East German unit is already surpassed, indicating that customers use outstanding trade credit to finance other expenses. Hence, FEED strives to further reduce the East German overdue significantly in the near future.

5.2. Market dynamics at play

5.2.1. Business cycle

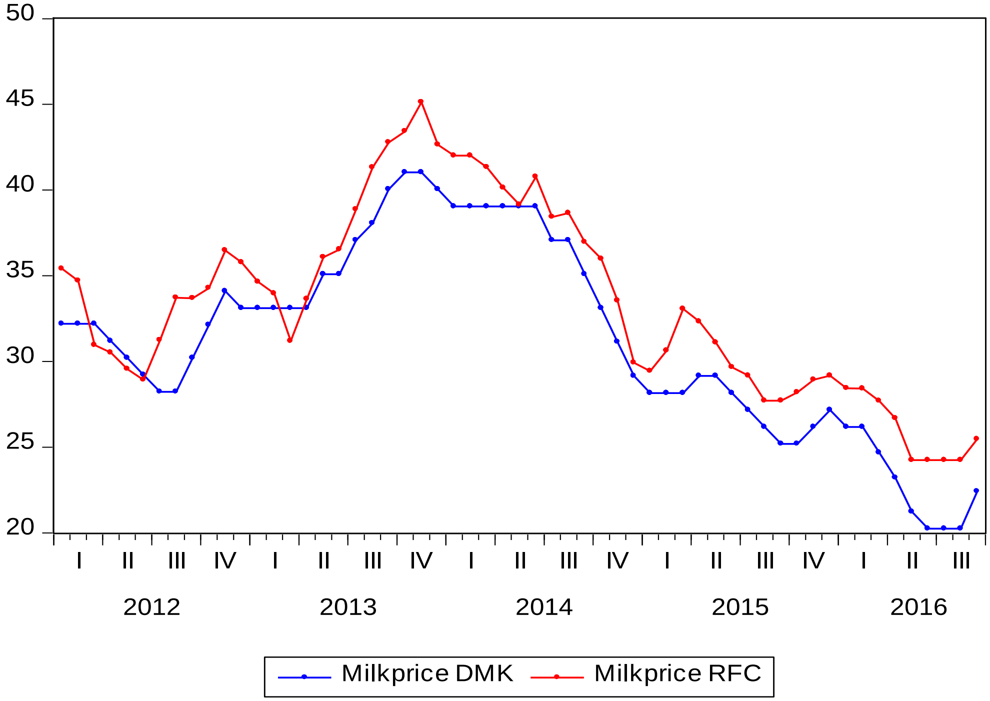
Based on interviews and supported by a basic regression analysis, the major cause for the differences in AR overdues between the Dutch and the East German unit is the milk price, which is significantly lower in the latter area. This is shown in figure 2 along standardized milk prices paid by the major dairy company in East Germany, Deutsches Milchkontor GmbH (DMK), and the major dairy company in the Dutch area, Royal FrieslandCampina N.V. (RFC). The main reason for this discrepancy is the ability of dairy factories to add value to raw milk. RFC can turn raw milk into high-value-added products like baby milk powder. Hence, it is able to provide milk suppliers with relatively high milk price. Milk in East Germany is more locally consumed and less transformed in the process. DMK and other dairy companies are therefore adding less value to their products here and provide relatively low milk prices. A few dairy companies in East Germany do possess the ability to add more value to milk, e.g. Müllermilch. Yet, these companies do not (have to) pay higher milk prices to dairy farms, since their low-value-adding competitors are not doing this either.

Internal FEED data from 2012 until 2016 show that overdues for the East German unit changed almost immediately after a downward move of the milk price in that area. Dutch dairy farms however, appeared to generate overdues only after a few months. This might indicate that East German dairy farms are more sensitive to milk price changes compared to the Dutch area due to their relatively poor financial structure and a lack of financial buffers to absorb negative cash flows. The differing response of overdues to business cycle changes was also found during multiple interviews.

A force potentially counteracting the above discrepancy is the size of dairy farms. Dairy factories usually provide a premium upon the milk price for larger suppliers because these suppliers can provide them with fully loaded milk trucks and the ability to transport milk over longer distances, providing additional advantages. Since the size of dairy farms in East Germany (measured by the number of cows) is generally a multitude of Dutch dairy farms, this counteracts the above mentioned forces and decreases the discrepancy in milk prices.

However, the relatively low milk prices in East Germany remain the major factor contributing to the differences in AR overdues. It must be noted though, that the discrepancy in overdues is likely to increase in the near future due to planned phosphate regulation in the Netherlands. Confirmed by LTO Nederland (2016), this reduces the amount of allowed manure and is likely to reduce the number of animals in total. Therefore the milk production is likely to decrease, adding an upward force to milk prices in the Netherlands.

Figure 2. Standardized milk prices; 2012 until September 2016



Source: LTO Nederland (2016).

On the basis of the argumentation above, the business cycle market dynamic is a relevant market dynamics for explaining the country unit differences in AR overdues and will therefore be included as a foundation for the recommendation framework in the next section.

5.2.2. Dairy farm financial structure

Farm debt ratio. Based on multiple interviews, the second factor causing differences in AR overdues refers to farms in East Germany being usually more debt-financed than farms in the Dutch area. After the old communist system collapse in East Germany, the buildup of new private farms often required debt funding. The debt-focused approach is reinforced recently by strong increases in land values in East Germany, which have more than doubled in the past decade. This increased collateral for dairy farmers, which they used to attract more loans from banks and benefit from economies of scale. Yet, the ratio of liabilities to land value has worsened as well, making dairy farmers even more vulnerable to economic troughs. This is in contrast to the situation in the Dutch area, where many Dutch family-owned farms were inherited and are relatively small, and therefore more equity-financed. Farmers in the Netherlands also have invested heavily in new barns and equipment in the recent years before the abolishment of the quota, even doubling their investments per kg of milk in the past ten years (Rabobank 2016). Nevertheless, Dutch dairy farms are still relatively more equity-financed.

Hence, farms in East Germany have higher relative fixed costs (rent and interest expenses) compared to their Dutch counterparts. This makes them particularly more susceptible to the business cycle. They are more likely to defer payments and generate overdue during low milk price periods. This confirms the link between the business cycle and the debt ratio established earlier. Moreover, a combination of low milk prices and a deteriorated financial structure leads to more farm bankruptcies in East Germany nowadays.

On the basis of the argumentation above, the farm debt ratio market dynamic is a relevant variable for explaining the country unit differences in AR overdues and will therefore be included as a foundation for the recommendation framework in the next section.

Farm size. Based on various interviews, the discrepancy in financial structure is reinforced with the average farms in East Germany being of considerable larger size. A larger proportion of customers have more than 500 animals (defined as the XXL customer segment) in East Germany than in the Dutch area. East German customers are more susceptible to liquidity shortages during periods of low milk prices due to high fixed costs and inflexible labor costs, thereby having a negative impact on AR overdues at FEED. However, the size discrepancy between the areas is likely to decrease. Dutch farms are gradually increasing economics of scale due to mergers and the inability to find successors.

On the basis of the argumentation above, the dairy farm size market dynamic is a relevant variable for explaining the current country unit differences in AR overdues and will therefore be included as a foundation for the recommendation framework in the next section.

5.2.3. Farmer mentality

Input vs. output drive. Based on interviews, East Germany is more input oriented in general, whereas the Dutch sector is more output oriented. This discrepancy is illustrated with a high adoption rate of a software platform developed by FEED to give customers more technical and economic insights into their farms and maximize the quality and quantity of milk per animal (FEED 2016). However, less East German customers make use of the platform. There are two reasons for this discrepancy. First, historical usage of the platform has been less given the short history of FEED in East Germany (since 2012). Second, East German farms use more raw materials as feed, and these are low value-added products for which it is hard to gain a competitive advantage. For FEED, differentiating from competitors on a technical level is hard in this product segment. Since customers in East Germany are thus more input oriented, they tend to be more price and payment term oriented.

Interviews tell that the above contrasts result in relatively more switching behavior, and more price and payment term drive in East Germany. This may be one explanation for the slightly longer average payment terms in East Germany and hence a longer DSO, but does not grasp the discrepancy in overdues. Also, as stated earlier, no clear further support for a positive relationship between input drive and FEED overdues could be identified.

On the basis of the argumentation above, the input/output drive market dynamic does not provide substantial explanation for country unit differences in AR overdues and will not be included as a foundation for the recommendation framework in the next section.

Payment mentality. There is not much country unit difference in willingness of customers to fulfill payments. The payment mentality for the both areas is good: customers pay when they are able to. This factor can be merely responsible for individual differences within each area. However, interviews urged to include a measure of trustworthiness or payment behavior in the recommendation framework. On the basis of the argumentation above, the payment mentality market dynamic does not provide substantial explanation for country unit differences in AR overdues but will be included as a foundation for the recommendation framework in the next section.

5.2.4. Third party market participants

Competition. The market in the whole of Germany is more fragmented compared to the Netherlands: there are more competitors for FEED and the market share is lower. In the Netherlands, FEED has a relatively high market share, which is likely due to the long term presence in the country and several acquisitions. In East Germany the market share is much lower and the number of competitors is higher. The lower market share may be partially due to the relative short presence of FEED in the area, only acquiring the East German unit in 2012. This leads to a more leading position for FEED in the Netherlands, whereas in East Germany it has to follow the competition in prices and payment terms in order to prevent customers from switching to competitors. Therefore this factor may partly explain why payment terms in East Germany are longer, and thus result in a higher DSO. However it does not clearly capture why overdue in East Germany is considerably higher compared to the Dutch area.

On the basis of the argumentation above, the competition market dynamic does not provide substantial explanation for country unit differences in AR overdues and will not be included as a foundation for the recommendation framework in the next section.

Dairy factories. In the Dutch area, the majority of dairy farmers is paid by RFC on the 14^{th} every month, whereas in East Germany dairy factories like DMK pay dairy farmers on the 20^{th} each month. However, the payment terms adopted in both regions are suited towards these payment dates. Customers that request to pay after the receipt of milk money, instead of paying after 7 or 21 days of delivery, are billed on the date in accordance with the respective milk factory. Therefore it can be assumed that this issue is unlikely to result in differences with regards to overdues between both areas. Furthermore, no broad support for a significant positive relationship between dairy factory payment dates and FEED overdues could be identified.

On the basis of the argumentation above, the dairy factory market dynamic does not provide substantial explanation for country unit differences in overdue and will therefore not be included as a foundation for the recommendation framework in the next section.

Banks. Based on interviews, banks in East Germany usually employ less flexible credit conditions for dairy farmers compared to banks in the Dutch area. First, this is due to a more formal culture in Germany where a contract usually is not renegotiated after being signed, whereas in the Netherlands this is more accepted. Second, the German banking sector is relatively more regulated, historically more focused on agriculture and less risk diversified than the main Dutch agricultural banks. Therefore, German banks are currently more constrained and have less flexibility in providing credit to farmers. Third, East German “Landesbanken” are less willing to provide credit, since farmers in this area have relatively more financial distress. This leads to a higher cost of capital, less credit availability and shorter repayment periods, eventually raising fixed costs for dairy farmers even further and increasing the likelihood of generating overdues. Combined with the business cycle, debt ratio and farm size, this adds to the financial inflexibility of East German dairy farmers.

On the basis of the argumentation above, the bank flexibility market dynamic does provide substantial explanation for country unit differences in AR overdues. However, it amply follows from the farmers’ financial structure and it would be very difficult to alter bank flexibility. Furthermore, due to uncertainty regarding the implementation of phosphate laws in the Netherlands, Dutch banks are increasingly strict in their credit providence. Hence, this market dynamic will not be included as a foundation for the recommendation framework.

5.3. Foundation for recommendations

A summary of the findings in the diagnosis stage is given by table 4. The columns specify the respective market dynamic categorization belonging to each market dynamic subcategory and its perceived influence (positive or negative) on AR overdues. The remaining columns specify the situation in the Dutch and East

German units, and the effect of the respective market dynamic on the differences in AR overdues between both country units.

Table 1. Overview of market dynamics influencing overdue

Market dynamic Category	Subcategory, specific for dairy cow sector	Effect on overdue	NL	GER	Effect on AR differences
Business cycle	Standardized milk prices	-	High	Low	Confirmed
Financial structure	Farm debt ratio	+	Medium	High	Confirmed
Financial structure	Farms size by number of animals	+	Low	High	Confirmed
Mentality	Degree of input Orientation	+	Low	High	Not substantial
Mentality	Degree of payment mentality	-	High	High	Not substantial
3rd parties	Competition level	+	Low	High	Not substantial
3rd parties	Dairy Factory payment date	+	14th	20th	Not substantial
3rd parties	Degree of bank flexibility	-	Medium	Low	Confirmed

Source: Authors' own elaboration

Some market dynamics appear to have no substantial influence on AR overdue differences. The most relevant and usable market dynamics are used to build a recommendation framework upon: the business cycle, dairy farm financial structure and payment mentality.

6. From market dynamics to recommendation framework

In the design part, the lessons on market dynamics in the previous sections will be utilized to build a recommendation framework for improving AR and reducing overdue by prioritizing customers likely to generate overdue. First, general framework requirements are reviewed. Next, a dynamic risk segmentation procedure, its extensions, and its limitations are discussed.

6.1 General framework requirements

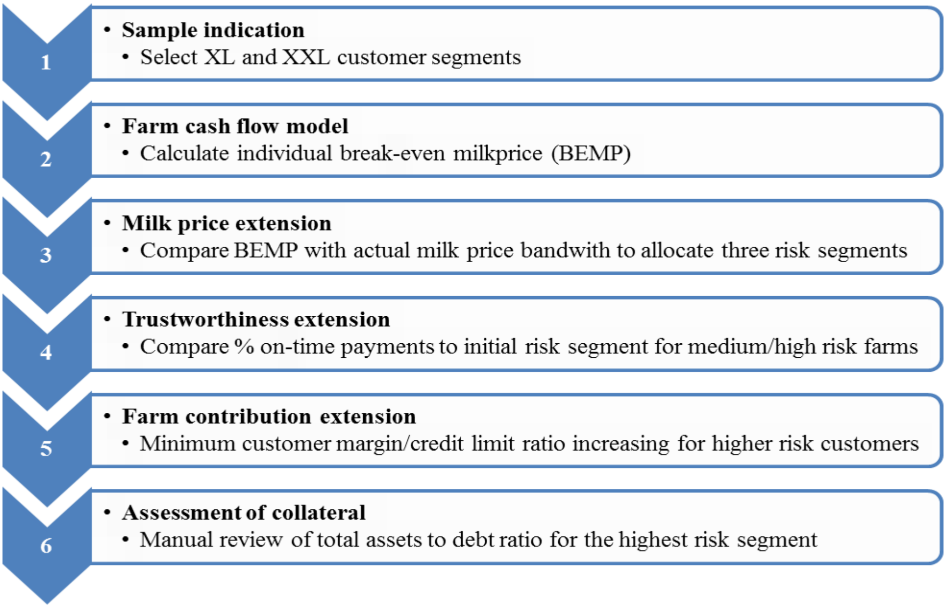
Based on multiple interviews, the AR collection policy can be more proactive when adopting the business cycle as a forward looking indicator. FEED needs to anticipate the customers' financial structure based on the current movement of the milk price and adjust collecting policy accordingly, by being stricter when the milk price starts to rise after which farmers will gradually build financial resilience and to loosen as soon as the milk price falls. The few-months' lag between the rise (fall) of the milk price and improvement (deterioration) of the financial structure of the farmers allows for a suitable forward-looking indicator. In line with business cycle market dynamics, a dynamic procedure should assess a customer's riskiness to generate overdue by anticipating on the farmers' financial structure based on the milk price.

There is a need for flexible collection policy targeted towards the customer's financial structure, in line with the second relevant market dynamic. Overdue procedures are conducted "manually" after overdues occur, which sometimes means the credit limit (the maximum amount of credit outstanding to a single customer) has already been surpassed. Part of this "ex-post" management involves looking at the customer financial data and bringing in a customized solution whenever overdue is problematic. This costly and time-consuming process should rather be substituted by proactive action. Also, policies do not structurally differ upon the financial structure of the customer. Furthermore, there is a need for more insight in customer financials. Therefore, in line with previous lessons from the customer's financial structure market dynamic, the procedure needs to include an additional set of variables regarding the financial health and collateral of the customer and adjust their perceived riskiness accordingly.

6.2. Dynamic risk segmentation procedure

To incorporate the above requirements, this study proposes a dynamic risk segmentation procedure that categorizes customers into risk segments (low, medium, high). The procedure was designed iteratively to generate a funnel effect, filtering and narrowing down the number of customers likely to generate overdue along the steps, eventually leaving the top priority customers in the highest segment. The six steps and their sequence are shown in figure 3. General framework requirements are met by proactively categorizing key customers on the basis of the milk price, combined with an analysis on financial structure, which allows FEED to adjust overdue procedures based on risk segments, thereby prioritizing the process of collecting AR. Finally, the procedure forms a structural tool to easily communicate decisions towards upper management and to gain insights into customer payment behavior.

Figure 3. Steps in the dynamic risk segmentation procedure



Source: Authors' own elaboration

Step 1. Sample indication:

First, based on interviews (with one staff member disagreeing), yearly information collection and analysis is most beneficial with the largest customers. The procedure is currently tested for the XL (200-500 animals) and XXL (more than 500 animals) customer segments. This is consistent with literature regarding customer size and case study findings regarding the customer size market dynamic. Information is available from the FEED customer relationship management (CRM) system. In this step, customers are not allocated to risk segments yet.

Step 2. Farm cash flow model:

Second, individual customer cash flows are investigated, in line with literature regarding capacity and connects to the business cycle and debt ratio market dynamics. Based on advices during interviews, this study therefore uses a quick scan (Farm Credit East 2016) to indicate operational cash flows with relatively little information. The most important variables for the quick scan are as follows: 1) sold milk quantity per cow in kg, 2) farm size in number of animals, 3) costs of labor per kg milk, 4) costs of feed per kg milk, 5) fixed costs per kg milk, 6) other expenses per kg milk, 7) family living expenses per kg milk, and 8) non-milk income per kg milk. The first and second variables are merely needed to calculate the others. The third, fourth, fifth and sixth variable need to be calculated to gain an overview of the farm's cash flow expenses. The seventh variable is added to give an indication of the total cost price per kg milk. Finally, the eighth variable is subtracted to provide an indication of net costs of production (i.e. the break-even milk price). A German example calculation is given in table 5.

Table 5. Example calculation of farm costs per kg milk

	Average
	7/14 - 6/15
Costs per kg sold milk	Cent/kg
Raw material feed	7.9
Total feed	9.5
Labor costs	7.1
Other expenses	10.1
Fixed costs	6.1
Total cash flow expenses	**40.7**
Family living expenses	2.5
Total cost price	**43.2**
Non-milk income	3
Break-even milk price	**40.2**

Source: Koesling Anderson (2015).

The analysis needs to be conducted yearly with a combination of 1) available data from the CRM system, 2) farm financial statements, 3) publicly available reports, and 4) sales department representatives. In this step, customers are not allocated into risk segments yet.

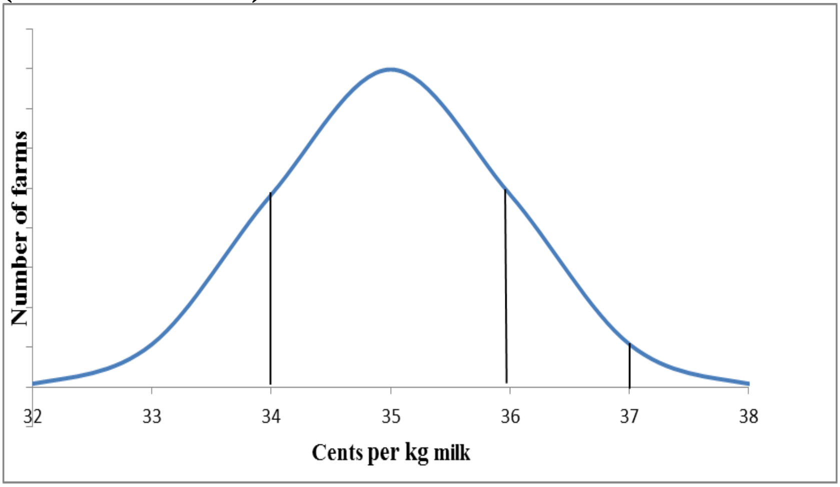
Step 3. Milk price extension:

Third, based on literature regarding conditions, the market dynamics analysis, and the framework requirements, the business cycle needs to be included as a forward looking indicator to anticipate customers' future cash flows. FEED staff points out to initially allocate customers in risk segments after this step is conducted. This can be done by utilizing the break-even milk price from the previous step and comparing it with the actual milk price, thereby generating a robust indication of the customer's cash margin and its likelihood to generate overdue. Therefore, a monthly review of the actual milk price corresponds with an immediate potential shift to a higher (or lower) risk segment of individual customers when the milk price passes the predefined bandwidth from the individual break-even milk price (BEMP). The initial bandwidth is defined as the actual milk price subtracted by the lowest BEMP in the total sample per region. This bandwidth is divided into three sections, one for each risk segment. For every region, the width of these sections is defined as a fixed number of cents per kg milk. Using data from the CRM system, customers are allocated in risk segments matching the section their BEMP belongs to (cf. figure 4).

Step 4. Farmer mentality extension:

Fourth, based on the market dynamics analysis and advices from interviews, a yearly revise of previous payment behavior should be included in the procedure. This is consistent with the literature regarding character and the mentality market dynamic. Based on combined staff member advices, this measure of trustworthiness is defined as the Euro value of on-time payments as a percentage of the total Euro value of payments per customer. To correct for occasional payment delays (such as international bank transfer delays), five days are added to the due date to calculate the definitive on-time payment threshold. On-time payments percentages are compared with the average per region and consequently divided into two groups.

Figure 4. Example of risk segment allocation with milk price of 37 cents per kg milk, maximum BEMP deviation of 1 cent (high risk farms) and 3 cents (medium risk farms)



Source: Authors' own elaboration

Customers whose mentality (e.g. untrustworthy; below average on-time payments) differs substantially from their initial risk segment (e.g. medium risk) need to shift to another risk segment (e.g. high risk). One staff member disagreed with this procedure and proposed a manual review of trustworthiness in this step, but no further support for this notion was identified. Information is available from FEED's credit management system.

Step 5. Customer contribution extension:

Fifth, interviews point out that the procedure also should include an earnings element. Hence, yearly gross customer contribution margins (which differ between total feed and raw material feed customers) are compared with the credit limit to give indications. The minimum ratio should be increasing for higher risk segments, thereby exerting less tolerance towards higher risk segments. Below the respective ratio of a risk segment, customers are automatically allocated to a higher risk segment. The data needed are available in the FEED CRM system.

Step 6. Assessment of collateral:

Sixth, interviews indicate that a final assessment needs to be conducted to include the amount of collateral a farmer has. This is consistent with literature regarding size and collateral and the size and debt ratio market dynamics. Hence, farm assets and debt ratio need to be reviewed yearly. Based on advices, this can be done manually and conducted merely for customers in the highest risk segment. Data can again be drawn from the CRM system.

Additional remarks:

The procedure needs to be conducted per owner, instead of per legal entity. Otherwise, farmers with multiple business entities can form an unseen risk. As mentioned above, all steps are followed on a yearly basis with exception of step three, which is reviewed monthly as a forward looking indicator. Regardless of this frequency, one may run the entire procedure in the case of major macroeconomic events. Thus, the flexibility of the procedure is preserved. The risk segments generated from this procedure allow FEED to prioritize customers based on likelihood to generate overdue and to match collection policy accordingly. FEED can then act proactively to agree on payment plans with customers likely to generate overdue, or adjust payment terms based on the average BEMP. Furthermore, high risk customers can be denied extra trade credit and should fulfill their payments before an additional delivery of feed.

6.3 Framework limitations

Staff members indicate a need for customer information. Regarding future customers, financial statements requirements should be included in FEED's terms and conditions. Yet, a small proportion of the existing customers might be unwilling to provide this information. This unwillingness can be translated into risks and hence the customer may be allocated automatically in the highest risk segment. Moreover, information availability of farms has recently improved, since banks are increasingly strict regarding farm financial statements. Also, the data availability at large farms (which are focal in the procedure) is relatively good.

6.4 Framework evaluation

The framework provides several interesting contributions to the existing literature. Large farms are selected in procedure step 1, after which their operating cash flows are calculated in procedure step 2 and compared to the business cycle in procedure step 3. This is consistent with the literature regarding size, capacity and condition factors. In the analysis, the business cycle was viewed as an agglutination of capacity and conditions factors. These market dynamics were incorporated in the framework due to their relevance for accounts receivable and overdues and their contribution to country unit differences. Interviews with staff members pointed out that West- and East Germany can be seen as separate markets here, since the latter has considerably larger farms and substantially lower milk prices. This leads to considerable intra-country differences regarding AR overdues, and between the Netherlands and East Germany, whereas the Netherlands and West Germany are more similar. Hence, although previously labelled as country differences, actually rather market differences were found.

In procedure step 4, payment mentality was reviewed and found to be consistent with the literature on the character factor and the payment mentality findings. Due to its relevance for AR and overdues, it was incorporated in the framework. However, it appeared not to contribute to country unit differences, but to individual differences within each market. Hence, differences were not found on a country or market level, but on an individual level.

In procedure step 5, customer contribution was reviewed. It is rather an internal earnings than a market dynamics requirement that was not found in the case study or in the literature as a factor influencing AR or overdues. However, the gross contribution amount can be linked to customer size, which is consistent with literature on the size factor. Due to the link between customer size and the demand for feed type, East German farms generally have a lower contribution ratio (but a higher gross contribution amount) compared to West German farms since the contribution ratio on raw material feed is lower. This indicates intra-country differences and between the Netherlands and East Germany, mainly due to the size market dynamic. Hence, differences were not found on a country level, but rather on a market level.

In procedure step 6, collateral was reviewed. This is consistent with the literature regarding collateral and capital factors, and the findings where the debt ratio can be viewed as an agglutination of collateral-, capacity- and capital factors. It was relevant for AR and overdues, and country unit differences. West- and East Germany can be seen as separate markets, since the latter has farms with generally higher debt ratios, and the first is similar to the Netherlands. Thus, differences were rather found on market levels than on country levels.

Both the input drive and dairy factories were not incorporated in the framework, which is consistent with the literature and the case study where they were not found to influence AR and overdues. Input drive differs between the Netherlands, and the whole of Germany. Hence this can be seen as a country level difference. However, dairy factory payment dates differ between East- and West Germany. Thus, this difference rather acts on the market level.

Bank flexibility was not included in the procedure due to implementation difficulties. However, consistent with the literature, it was found to be relevant for AR overdues and country unit differences. Based on previous sections, differences in this aspect were found on the market level, which is largely due to differences in size, debt ratio and the business cycle between the Netherlands and West Germany on the one hand, and East Germany on the other.

Competition levels were not included in the procedure due to its irrelevance for the case at hand. This is likely due to FEED’s policy on limited adjustment to competitor payment terms and its unclear relationship with overdue. Differences in this aspect were found on a country level, since competition is much fiercer in the whole of Germany than in the Netherlands.

Finally, in sharp contrast to previous literature, the interest rate never came forward in the case study and was not incorporated in the framework. This is likely due to the current low interest rate, making it irrelevant as a market dynamic for AR overdues. This is reinforced by the current low costs of capital in general, making default risk relatively more important for credit decisions. Hence, the framework was tailored towards factors that determine overdues (which have more emphasis on default risk) instead of receivables in general.

7. Conclusion

This research study has created an overview and categorization of relevant market dynamics and designs a recommendation framework for FEED. Based on the literature, interviews and other documentation, an overview and categorization of relevant market dynamics influencing AR overdues for FEED in the dairy cow sector is given. With the help of interviews and various documents, AR overdue differences between the Dutch and East German units are discussed and explanatory market dynamics for these discrepancies are defined, resulting in the foundation of a recommendation framework. The framework is designed and evaluated on the basis of previous literature, ultimately developing extensive procedures for prioritizing customers and improving AR overdue practices in the dairy cow sector.

Regarding practical contributions for FEED, this study provides in-depth insights into market dynamics giving rise to AR overdues and subsequent differences between country units. This knowledge can be utilized for improving evaluation of country unit data. Second, it can be used for guiding and structuring communication towards investors, giving them deeper understanding of risks. Third, the recommendation framework makes credit decisions more insightful and customer-specific by prioritizing customers. Thus, procedures, guidelines and other credit decisions can be adjusted to improve FEED overdue. Fourth, the dynamic risk segmentation procedure provides a structural tool to communicate credit decisions from the credit control department towards senior management, thereby improving mutual alignment, and transparency, while saving time. Fifth, after some modifications, the framework can be extended towards other sectors, such as the swine and poultry sector with the price of piglets, broilers or eggs to be used as a business cycle variable. Sixth, the recommendation framework allows for a better customer service towards low-risk segments; these customers will be less bothered by FEED credit control since priorities shift towards high-risk segments. Finally, after modifications, the framework can be used by FEED’s sales department when evaluating prospective customers’ potential payment behavior and likelihood to generate overdues.

In addition, this research contributes to the literature by examining a relatively uncharted area of WCM. The overview and categorization are partially consistent with previous literature and provide unique insights into market dynamics determining AR overdues, thereby guiding future research into what extent these factors are market-specific or country-specific. Second, this case study is unique in that it uncovers how discrepancies in AR overdues arise between country units, thereby supporting future research regarding intra-company AR differences. Third, whereas the literature has indicated competition levels as a relevant market dynamic for trade credit, this study shows that its relationship with AR overdues is unclear and further research is required. Moreover, competition was not indicated as a relevant market dynamic influencing country unit differences, suggesting that company policy regarding adjustment to competitor payment terms was of influence. Fourth, in sharp contrast to previous literature, the interest rate was not indicated as a relevant market dynamic for AR overdues and country unit differences. Also, today's low interest rate and economic uncertainty shifts the focus of credit decisions from costs of capital towards default risk, and hence a focus on what factors specifically determine overdues instead of receivables as a whole is asked for. This is shown by a recommendation framework with market dynamics determining likelihood of customers generating overdue standing central in most of its steps and interest rates being excluded.

A view of receivables as an investment in clients should therefore focus on customer relationships, instead of the costs of providing credit. Hence, this case study emphasizes a relationship between WCM and trust, and shows that a trust-based view of inter-company relationships (cf. Nooteboom 1996) is useful. Moreover, WCM as part of risk management shifts partially away from financial risk (particularly interest rate risk), and nowadays focusses more on credit risk, where evaluating and accepting (or rejecting) customers is key. Therefore, this phenomenon provokes a cascade towards operational risk, and consistent with Heijes (2016), a focus from financial risk towards operational risk can be depicted.

Trivially, this study has limitations that can guide future research. Although it investigates intercompany differences in AR overdues, it does not cover how the differences translate into country unit profitability. The literature highlights how shorter collection time of receivables is correlated with higher profitability, but aggregates this relationship across companies. Thus, company-specific research is urged to investigate potential differences between country units. Also, this paper investigates the market dynamics influencing AR overdues for a single company within a single industry. Hence, investigation of other companies, industries and/or countries is helpful. Such a study can indicate the respective relevance of competition levels for overdues, which remain inconclusive after this research. Moreover, future research into the influence of the interest rate on credit decisions at similar companies may uncover the respective relevance of interest rates for working capital management at other companies, such as in specific highly levered industries or in regions outside of Western Europe. Finally, after research-based adjustments, the market dynamics overview and categorization, and the subsequent recommendation framework can be applied in other sectors and/or companies.

References

Biais B., Gollier C. (1997), Trade credit and credit rationing, „Review of Financial Studies", vol. 10 no. 4, pp. 903-937.

Biernacki, P., Waldorf, D. (1981), Snowball sampling. Problems and techniques of chain referral sampling, „Sociological Methods and Research", vol. 10 no. 2, pp. 141-163.

Centraal Bureau Statistiek (2016), Consument merkt weinig van lagere melkprijs boeren, Centraal Bureau Statistiek, The Hague, https://www.cbs.nl/nl-nl/nieuws/2016/40/consument-merkt-weinig-van-lagere-melkprijs-boeren [26.08.2018].

Ernst & Young (2013), Analysis on future developments in the milk sector, European Commission, Brussels, https://ec.europa.eu/agriculture/sites/agriculture/files/events/2013/milk-conference/ey-independent-experts-analysis-presentation_en.pdf [26.08.2018].

European Commission (2016), Milk market situation, European Commission, Brussels.

Farm Credit East (2016), Northeast dairy farm summary, Farm Credit East, New York.

Filbeck G., Krueger T. (2005), An analysis of working capital management results across industries, „American Journal of Business", vol. 20 no. 2, pp. 11-20.

Hailemariam S. (2001), Corporate value creation, governance and privatisation, unpublished PhD thesis, University of Groningen.

Heijes C.P.A. (2016), Accounting of working capital. Evidence from the work floor. A case study on the use of financial and operational ratios, University of Groningen [unpublished].

Hermes N., Kihanga E., Lensink R., Lutz C. (2012), The impact of trade credit on customer switching behaviour. Evidence from the Tanzanian rice market, „Journal of Development Studies", vol. 48 no. 3, pp. 363-376.

Hermes N., Lensink R., Lutz C., Thu U. (2016), Trade credit use and competition in the value chain, „Economics of Transition", vol. 24 no. 4, pp. 765-795.

Hillier D., Ross S., Westerfield R., Jaffe J., Jordan B. (2016), Corporate Finance, 3rd edition, McGraw-Hill Education, London.

Howorth C., Reber B. (2003), Habitual late payment of trade credit. An empirical examination of UK small firms, „Managerial and Decision Economics", vol. 24 no. 6-7, pp. 471-482.

Jick T.D. (1979), Mixing qualitative and quantitative methods. Triangulation in action, „Administrative Science Quarterly", vol. 24 no. 4, pp. 602-611.

Koesling Anderson (2015), Milch-manager Auswertung, Koesling Anderson, Dahlenwarsleben.

LTO Nederland (2016), Standardized milk price calculations, LTO Nederland, The Hague.

Nederlandse Zuivel Organisatie (2016), Factsheet prijsschommelingen melk, Nederlandse Zuivel Organisatie, The Hague.

Ng C.K., Smith J.K., Smith R.L. (1999), Evidence on the determinants of credit terms used in interfirm trade, „The Journal of Finance", vol. 54 no. 3, pp. 1109-1129.

Nilsen J.H. (2002), Trade credit and the bank lending channel, „Journal of Money, Credit and Banking", vol. 34 no. 1, pp. 226-253.

Nooteboom B. (1996), Trust, opportunism and governance. A process and control model, „Organization Studies", vol. 17 no. 6, pp. 985-1010.

OECD Finance Department (2016), Short-term interest rates, https://data.oecd.org/interest/short-term-interest-rates.htm [26.08.2018].

Peel M.J., Wilson N., Howorth C. (2000), Late payment and credit management in the small firm sector. Some empirical evidence, „International Small Business Journal", vol. 18 no. 2, pp. 17-37.

Petersen M.A., Rajan R.G. (1997), Trade credit. Theories and evidence, „Review of Financial Studies", vol. 10 no. 3, pp. 661-691.

Pike R., Cheng N.S. (2001), Credit management. An examination of policy choices, practices and late payment in UK companies, „Journal of Business Finance & Accounting", vol. 28 no. 7-8, pp. 1013-1042.

Rabobank (2016), Cijfers & Trends. Banche informatie melkveehouderij, Rabobank, Utrecht.

REL Consultancy (2016), Europe working capital survey, The Hackett Group, Inc, London – Sydney.

WebFinance Inc. (2016), Definition of market dynamics, http://www.businessdictionary.com/definition/market-dynamics.html [26.08.2018].

Wilson N., Summers B. (2002), Trade credit terms offered by small firms. Survey evidence and empirical analysis, „Journal of Business Finance & Accounting”, vol. 29, pp. 317-351.

Yin R.K. (2014), Case study research. Design and methods, Sage Publications, London.

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 43-54, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

A comparative analysis of operational efficiency between Chinese and Indian commercial banks

Nan ZHU, Huajie ZHANG

Southwestern University of Finance and Economics, Chengdu, China

Abstract:

Aim: The objective of this paper is to make comparative analysis on operational efficiency between Chinese and Indian commercial banks (CBs).

Design / Research methods: Following the previous scholars' study, two models with different sets of input and output variables have been used to show how efficiency scores vary with change in inputs and outputs. The efficiency scores are measured by using data envelopment analysis (DEA) approach.

Conclusions / findings: The mean technical efficiency score of Chinese CBs is always relatively higher than the corresponding score of Indian CBs in 2012-2013, respectively. In terms of technical efficiency and pure technical efficiency, the performance of foreign banks in China is always relatively lower than that of foreign banks in India.

Originality / value of the article: While many similar studies have evaluated the performance of banking industries in different countries, very few studies have evaluated the performance of banking sectors between Chinese and Indian economies. The paper would be of interest for OR scholars and practitioners in financial industry.

Implications of the research (if applicable): The next step of this study could collect more samples and use Malmquist index method to conduct further study on efficiency, efficiency changing and productivity, in order to conduct further competitive power analysis on both of banking industries of China and India.

Key words: Data envelopment analysis, Commercial banks, China, India.

Correspondence address: Nan ZHU, Huajie ZHANG, Southwestern University of Finance and Economics, Chengdu 610074, China E-mail: . zhunan@swufe.edu.cn, 111020209001@2011.swufe.edu.cn
Received: 30.08.2017, Revised: 07.11.2017, Revised: 04.07.2018, Accepted: 04.07.2018
Doi: http://dx.doi.org/10.29015/cerem.535

1. Introduction

Organization management gets people together for organizational strategic objectives and enables the optimal use of scarce resources through planning, organizing, leading and control at the workplace. Usually, a commercial bank (CB), which is a special service organization, is a type of financial intermediary and of bank that provides services such as accepting deposits, making business loans, and offering basic investment products. Banks are vital organizations in any country as they significantly contribute to the development of an economy through serving customers, and play the major role in economic development.

The objective of this paper is to make comparative analysis of operational efficiency between Chinese and Indian CBs by using data envelopment analysis (DEA) approach introduced by Charnes, Cooper and Rhodes (1978). While many similar studies have evaluated the performance of banking industries in different countries, very few studies have evaluated the performance of banking sectors between Chinese and Indian economies.

Both of China and India belong to developing countries. They are the two most populous countries and fastest growing major economies in the world. In this paper, following the previous scholars' study and using DEA method, with available published data and by setting up two models, we make comparative analysis of operational efficiency (including technical efficiency, pure technical efficiency and scale efficiency) between Chinese and Indian CBs for the span of two years, 2012 and 2013, respectively.

The paper is organized as follows. A brief review of the current state of the Chinese and Indian banking sector is provided in Section 2. In Section 3 methodology is discussed. Section 4 presents the results and Section 5 concludes this paper.

2. A brief overview of the Chinese and Indian banking industry

China and India are separated by the geographical obstacles of the Himalayas. The China's population of in 2013 is about 1.36 billion. India is the second-most populous country over 1.2 billion people. Historically, China and India have had relations for more than 2,000 years. On 1 January 1950, the People's Republic of China established diplomatic relations with the Republic of India. Since then the bilateral economic relationship has been increased significantly.

In China, the CBs are those enterprise legal persons which are established to absorb public deposits, make loans, arrange settlement of accounts and engage in other businesses. CBs shall work under the principles of safety, liquidity and efficiency, with full autonomy and assume sole responsibility for their own risks, profits and losses, and with self-restraint. At the end of 2013, the Chinese banking industry had 3,949 financial institutions with 3.55 million employees. The banks include: 5 large and state-owned CBs (Industrial & Commercial Bank of China, China Construction Bank, Agricultural Bank of China, Bank of China and Bank of Communications), 12 joint-stock CBs, 145 city CBs, 468 rural CBs, 122 rural cooperative banks, 1,803 rural credit cooperatives, one postal savings bank and 42 foreign financial institutions, etc. (China Banking Regulatory Commission 2014).

Since July 2013, the Chinese banks have been free to set their own lending rates. In comparison to their counterparts, the 5 state-owned CBs exhibit strong capabilities and competitiveness compared to either in terms of financial indicators: such as asset scale and profitability. E.g., at the end of 2013, the total sum of assets of 5 big banks is RMB 11.254 trillion (US$ 1.844 trillion), hold 43.34% of total financial asset of the Chinese banking financial institutions (China Banking Regulatory Commission 2014).

The Indian banking industry is broadly classified into scheduled banks and non-scheduled banks. The scheduled banks are further classified into: State Bank of India and its associates; nationalized banks; Indian private sector banks; foreign banks; and regional rural banks. Generally banking in India is fairly mature in terms of supply, product range and reach-even though reach in rural India. The term CBs in India refers to both scheduled and non-scheduled CBs. The CBs are consisted of public sector CBs, private sector CBs and foreign CBs. Public sector CBs are owned and operated by the government as the government holds a major share in them. A well-operated public sector CB can help state and local governments in getting through cash crunches. The Indian government presently hires the CBs for different purposes like tax collection and refunds, payment of pensions, etc. (Reserve Bank of India 2014).

By 2013 the Indian Banking Industry employed 1.18 million employees and had a total of 109,811 branches in India and 171 branches abroad and manages an aggregate deposit of ₹67,504.54 billion (US$1.1 trillion) and bank credit of ₹52,604.59 billion (US$820 billion). During the financial year Mar 2013-Mar 2014, there were 27 public sector CBs in India out of which 6 were State Bank of India and its 5 associates banks (State Bank of Bikaner and Jaipur, State Bank of Hyderabad, State Bank of Mysore, State Bank of Patiala and State Bank of Travancore), and 21 were nationalized CBs. At the same time, there were 20 private sector CBs, 43 foreign CBs, regional rural banks, cooperative banks, other type banks and financial institutions in India. On the performance of Indian scheduled CBs, in terms of consolidated operations, the consolidated balance sheet of the CBs in 2013-2014 registered a decline in growth in total assets and credit for the fourth consecutive year. With both credit and deposit growth more or less same, the outstanding credit to deposit ratio at the aggregate level remained unchanged at around 79% (Reserve Bank of India 2014).

3. Methodology

3.1. Data envelopment analysis

Charnes, Cooper and Rhodes (1978) introduced DEA as non-parametric efficiency analysis for measuring the efficiency of Decision Making Units (DMUs). Consider a set of J decision-making units (DMUs) with n input and m output in T (t=1,…,T) periods. Assume in time period t, decision-makers are using inputs $x^t \in R_+^n$, to produce outputs $y^t \in R_+^m$. Define the input requirement set in period t, which is:

$L^t(y^t) = \{ x^t: x^t \text{ can produce } y^t\}$.

Assume that $L^t(y^t)$ is non-empty, closed, convex, bounded and satisfies strong disposability property of inputs and outputs. $L^t(y^t)$ is bounded from below by the input isoquant (a constant returns to scale (CRS) production boundary), that is:

$$\text{Isoq}L^t\left(y^t\right)=\left\{x^t : x^t \in L^t\left(y^t\right),\ \lambda x^t \notin L^t\left(y^t\right) \text{ for } \lambda<1\right\}.$$

Define the input distance function of period t as following:

$$D^t\left(y^t, x^t\right)=\sup_{\theta}\left\{\theta:\left(x^t/\theta\right)\in L^t\left(y^t\right),\ \theta>0\right\}.$$

Hence, define the technical (or productive) efficiency (TE) in period t as following:

$$\text{TE}^t\left(y^t, x^t\right)=1/D^t\left(y^t, x^t\right). \qquad (1)$$

In general, TE<1, indicates that the DMU under assessment, comparing with other DMUs, is productively inefficient since its production is based on excessive input usage. TE=1, indicates the DMU is fully productively efficient.

It is well known that TE can be further decomposed into the pure technical efficiency (PTE) and scale efficiency (SE) (Banker et al. 1984):

$$\text{TE}=\text{PTE}\times\text{SE}. \qquad (2)$$

In general, as TE, PTE or SE<1, indicates that the DMU under assessment, comparing with other DMUs, is pure technically inefficient or scale inefficient.

Following the above DEA models, many theoretical studies as well as applications are surveyed (Emrouznejad, De Witte 2010; Emrouznejad, Yang 2018). At present, the DEA models and development with applications in banking and finance areas can be seen. See, for examples, Emrouznejad and Anouze (2010), Hada and Tamang (2014), Wanke et al. (2016, 2017), and Zhu et al. (2017).

3.2. Two input-output models and solving

In the banking sector, Avkiran (1999), Sathye (2003) measured the productive efficiency (PE, i.e. TE) of banks in Australia and India by using DEA approach, respectively. Two input-output models, i.e., Model A and Model B, in their studies, have been constructed and used to show how efficiency scores vary with change in inputs and outputs. Following the same study direction, Zhu et al. (2004, 2012) studied the TE of Chinese main CBs by using the similar input-output DEA models, respectively; Recently, Hada et al. (2017) conducted a study on the productive efficiency between Nepal and China banking industry in year 2012 and 2013.

In this paper, following the previous scholars' work, two models, i.e., Model A and Model B, are provided and used:

	Model A	**Model B**
Inputs	Interest expense Non-interest expense	Deposits Staff numbers
Outputs	Net interest income Non-interest income	Net loans Non-interest income

Data used in this study is gathered from Bankscope database and annual reports of the banks from 2012 to 2013. Through data cleansing, we have got the samples of 100 Chinese CBs and 53 Indian CBs in 2012 and 2013. Chinese samples consist of 5 state-owned CBs, 12 joint-stock CBs, 54 city CBs, 15 rural CBs and 14 foreign CBs in China. Indian samples consist of State Bank of India and its 5 associates, 19 nationalized CBs, 19 private sector CBs, 4 foreign CBs in India (Citibank, HSBC, Standard Chart Bank and Bank of America) and 5 other type CBs in India.

The DEA problems are solved in the paper using the computer software DEA-Solver. The operational efficiency given is calculated in the input-oriented measure.

4. Results

The DEA results of the analysis are discussed in the following. Table 1 shows that by using the two DEA models, the mean operational efficiency score of all 153 sample CBs in 2012 and 2013, respectively.

Through Table 1, we see that the mean technical efficiency (TE) scores of the whole 153 banking samples collected from both of China and India, obtained by using both Model A and Model B, are slightly increased from 2012-2013. The mean scale efficiency (SE) scores are always relatively higher than the mean pure technical efficiency (PTE) scores.

Table 1. Mean operational efficiency score

Model A	**2013 TE**	**2013 PTE**	**2013 SE**	**2012 TE**	**2012 PTE**	**2012 SE**
All 153	0.6609	0.7277	0.9146	0.6502	0.7487	0.8712
China All 100	0.7465	0.8052	0.9297	0.7323	0.8208	0.8968
India All 53	0.4993	0.5816	0.8862	0.4953	0.6126	0.8229
Model B	**2013 TE**	**2013 PTE**	**2013 SE**	**2012 TE**	**2012 PTE**	**2012 SE**
All 153	0.6823	0.7526	0.9150	0.6719	0.7431	0.9115
China All 100	0.6914	0.7521	0.9272	0.7057	0.7519	0.9448
India All 53	0.6651	0.7536	0.8921	0.6081	0.7265	0.8488

Source: authors' own elaboration

Comparative analysis could be made. Mean TE score of Chinese CBs is relatively higher than the corresponding score of Indian CBs except PTE score of Model B in 2013 ($0.7521<0.7536$). Using Formula (2): $TE=PTE\times SE$, we can also make factor analysis on TE. In Table 1, that $PTE<SE$ is always true. Thus, the low PTE score brings the low TE score.

In detail, we have Tables 2-4 by using two DEA models. We can make similar comparative analysis through these tables. In Tables 2 and 3, "CH" means China, "CH 5 State" means 5 Chinese state-owned banks, "Joint" means joint-stock bank, "City" means city bank, "Rural" means rural bank, and "Foreign" means foreign bank in China. "IN" means India, "IN 6 State" means State Bank of India and its 5 associates, "National" means nationalized bank, "IN 24 General" means 19 private sector banks and 5 other type CBs in India, and "Foreign" means foreign bank in India. In Table 4, "IN 25 Public" means State Bank of India and its 5 associates, and 19 nationalized banks, "Private" means private sector banks, and "Others" means other type CBs in India.

Through Tables 2-3, we see that, in terms of TE and PTE, the performance of China's 5 state-owned banks is relatively higher than that of State Bank of India and its 5 associates, and China's other CBs; however, in term of SE, the performance of China's 5 state-owned banks is always relatively lower than that of State Bank of India and its 5 associates, and China's other CBs, respectively. However, in terms of TE and PTE, the performance of foreign banks in China is always relatively lower than that of foreign banks in India.

Table 2. Mean operational efficiency score of Model A

Model A	2013 TE	2013 PTE	2013 SE	2012 TE	2012 PTE	2012 SE
CH 5 State	0.8409	0.9741	0.8632	0.7734	0.9816	0.7872
CH 95 Others	0.7415	0.7963	0.9332	0.7301	0.8123	0.9025
CH 12 Joint	0.7252	0.8597	0.8460	0.6460	0.8630	0.7504
CH 54 City	0.7805	0.8155	0.9582	0.7481	0.8120	0.9236
CH 15 Rural	0.7419	0.7893	0.9414	0.7974	0.8652	0.9233
CH 14 Foreign	0.6044	0.6756	0.9023	0.6607	0.7131	0.9293
IN 6 State	0.4297	0.4542	0.9525	0.4145	0.5163	0.8091
IN 47 Others	0.5082	0.5979	0.8777	0.5056	0.6249	0.8247
IN 19 National	0.4732	0.5143	0.9275	0.4531	0.5805	0.7878
IN 24 General	0.4847	0.6157	0.8314	0.4881	0.6076	0.8392
IN 4 Foreign	0.8160	0.8880	0.9189	0.8591	0.9403	0.9127

Source: authors' own elaboration

For Chinese CBs, by using Model B, Zhu et al. (2004) discussed two groups of Chinese CBs for the years 2000-2001, that is, state-owned banks and joint-stock banks, and obtained that the mean TE score of state-owned banks is relatively lower than that of joint-stock banks in 2000 and 2001, respectively. For the years 2012-2013, through Table 3 that is the result by using Model B, we can see that the mean TE score of state-owned banks is still relatively lower than that of joint-stock banks, respectively. However, through Table 2 that is the result by using Model A, we can see that the mean TE score of state-owned banks is relatively higher than that of joint-stock banks in 2012-2013, respectively, that are the opposite results.

Table 3. Mean operational efficiency score of Model B

Model B	2013 TE	2013 PTE	2013 SE	2012 TE	2012 PTE	2012 SE
CH 5 State	0.7342	0.9537	0.7655	0.7234	0.9542	0.7545
CH 95 Others	0.6891	0.7415	0.9357	0.7047	0.7412	0.9548
CH 12 Joint	0.8414	0.9269	0.9086	0.8672	0.9377	0.9257
CH 54 City	0.6447	0.6814	0.9533	0.6562	0.6810	0.9677
CH 15 Rural	0.7148	0.7582	0.9392	0.7150	0.7632	0.9374
CH 14 Foreign	0.7025	0.7962	0.8874	0.7415	0.7813	0.9488
IN 6 State	0.6790	0.7406	0.9316	0.6101	0.7166	0.8714
IN 47 Others	0.6633	0.7553	0.8871	0.6078	0.7278	0.8459
IN 19 National	0.6455	0.7467	0.8718	0.5936	0.7488	0.7983
IN 24 General	0.6641	0.7392	0.9076	0.5917	0.6794	0.8859
IN 4 Foreign	0.7427	0.8928	0.8364	0.7723	0.9178	0.8323

Source: authors' own elaboration

For Indian CBs, by using Model A and Model B, Sathye (2003) discussed three groups of Indian banks for the year 1997, that is, publicly owned, privately owned and foreign owned, and obtained that the mean efficiency score of Indian banks compares well with the world mean efficiency score and the efficiency of private sector banks as a group is, paradoxically lower than that of public sector banks and foreign banks in India. However, through Table 4 in this paper, we can see that the TE score of private sector banks in India as a group is always higher than that of public sector banks; however, always lower than foreign banks in India in 2012 and 2013, respectively.

Table 4. Mean operational efficiency score of Indian banking industry

Model A	2013 TE	2013 PTE	2013 SE	2012 TE	2012 PTE	2012 SE
IN 25 Public	0.4627	0.4999	0.9335	0.4439	0.5651	0.7929
IN 19 Private	0.4826	0.5560	0.8913	0.4990	0.5733	0.8913
IN 5 Others	0.4927	0.8423	0.6037	0.4468	0.7378	0.6413
IN 4 Foreign	0.8160	0.8880	0.9189	0.8591	0.9403	0.9127
India All 53	0.4993	0.5816	0.8862	0.4953	0.6126	0.8229
Model B	**2013 TE**	**2013 PTE**	**2013 SE**	**2012 TE**	**2012 PTE**	**2012 SE**
IN 25 Public	0.6536	0.7452	0.8861	0.5976	0.7411	0.8159
IN 19 Private	0.6944	0.7261	0.9587	0.6256	0.6679	0.9435
IN 5 Others	0.5489	0.7889	0.7133	0.4627	0.7231	0.6670
IN 4 Foreign	0.7427	0.8928	0.8364	0.7723	0.9178	0.8323
India All 53	0.6651	0.7536	0.8921	0.6081	0.7265	0.8488

Source: authors' own elaboration

5. Conclusion

China and India are two of the world's oldest civilizations and have co-existed in peace for millennia. In this paper, we make comparative analysis of operational efficiency between Chinese and Indian CBs in 2012 and 2013 by using DEA approach. Two DEA output-input models, i.e. Model A and Model B, have been used to show how efficiency scores vary with change in inputs and outputs.

We have that mean technical efficiency score of Chinese CBs is always relatively higher than the corresponding score of Indian CBs in 2012-2013, respectively. In terms of technical efficiency and pure technical efficiency, the performance of China's 5 state-owned banks is higher than that of India's State Bank of India and its 5 associates, and China's other CBs, respectively; however, in term of scale efficiency, the performance of China's 5 state-owned banks is relatively lower than that of State Bank of India and its 5 associates, and China's other CBs, respectively. In terms of technical efficiency and pure technical efficiency, the performance of foreign banks in China is always relatively lower than that of foreign banks in India. The performance of private sector banks in India as a group is always relatively higher than that of public sector banks in India; however, lower than that of foreign banks in India.

The next step of this study could collect more samples and use Malmquist index method to conduct further study on efficiency, efficiency changing and productivity, in order to conduct further competitive power analysis on both of banking industries of China and India.

Acknowledgments: We would like to thank to two anonymous reviewers for their insightful comments and suggestions on this paper.

References

Avkiran N.K. (1999), The evidence of efficiency gains. The role of mergers and the benefits to the public, „Journal of Banking and Finance", vol. 23 no. 7, pp. 991-1013.

Banker R.D., Charnes A., Cooper W.W. (1984), Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis, „Management Science", vol. 30 no. 9, pp. 1078-1092.

Charnes, A., Cooper W.W., Rhodes E. (1978), Measuring the efficiency of decision making units, „European Journal of Operational Research", vol. 2 no. 6, pp. 429-444.

Cooper W.W., Seiford L.M., Tone K. (1999), Data Envelopment Analysis, a comprehensive text with models, applications, references and DEA-solver software, Kluwer Academic Publishers, Boston.

China Banking Regulatory Commission (2014), Annual Report 2013, China CITIC Press, Beijing.

Emrouznejad A., Anouze A.L. (2010), Data Envelopment Analysis with classification and regression tree. A case of banking efficiency, „Expert Systems”, vol. 27 no. 4, pp. 231-246.

Emrouznejad A., De Witte K. (2010), COOPER-framework. A unified process for non-parametric projects, „European Journal of Operational Research”, vol. 207 no. 3, pp. 1573-1586.

Emrouznejad A., Yang G. (2018), A survey and analysis of the first 40 years of scholarly literature in DEA: 1978-2016, „Journal of Socio-Economic Planning Sciences”, vol. 61 no. 1, pp. 1-5.

Hada S.S., Tamang G. (2014), Frontier analysis of Nepalese commercial banks. DEA Approach, „Journal of Mathematics and System Science”, no. 4, pp. 675-682.

Hada S.S., Zhu N., Tamang G. (2017), Frontier analysis of Nepalese and Chinese commercial banks. DEA approach, „International Journal of Operational Research – Nepal”, no. 1, pp. 39-46.

Sathye M. (2003), Efficiency of banks in a developing economy. The case of India, „European Journal of Operational Research”, vol. 148 no. 3, pp. 662-671.

Reserve Bank of India (2014), Financial stability report (Including trend and progress of banking in India 2013-2014), https://rbidocs.rbi.org.in/rdocs/Publications/PDFs/0FLTP577BF4E172064685A26A73A6BC9210EC.PDF [03.09.2018].

Wanke P.F., Barros C.P., Emrouznejad A. (2016), Assessing productive efficiency of banks using integrated Fuzzy-DEA and bootstrapping. A case of Mozambican banks, „European Journal of Operational Research”, vol. 249 no. 1, pp. 378-389.

Wanke P.F., Barros C.P., Emrouznejad A. (2017), A comparison between stochastic DEA and Fuzzy DEA approaches. Revisiting efficiency in Angolan banks, „RAIRO-Operations Research”, vol. 52 no. 1, pp. 285-303.

Zhu N., Li J., Wu Q., Cheng W.L. (2012), On productive efficiency and TFP changes of Chinese commercial banks, „Economist (Jing Ji Xue Jia)”, no. 9, pp. 56-61 (in Chinese).

Zhu N., Liu Y., Emrouznejad A., Huang Q. (2017), An allocation Malmquist index with an application in the China securities industry, „Operational Research – An International Journal”, vol. 17 no. 2, pp. 669-691.

Zhu N., Zhuo X, Dong Y. (2004), A case study of the efficiency of China’s national commercial banks and the reform strategy, „Management World”, no. 2, pp. 18-26 (in Chinese).

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 55-79, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Is there improvement in total factor productivity growth of the Indian pharmaceutical industry after TRIPS agreement? Evidence from Biennial Malmquist Index

Dipyaman PAL

Bethune College, Kolkata, West Bengal, India

Chandrima CHAKRABORTY

Vidyasagar University, Midnapore, WestBengal, India

Arpita GHOSE

Jadavpur University, Kolkata, India

Abstract:

Aim: Indian Pharmaceutical Industry (IPI) has undergone a massive makeover–from a modest beginning of "process patents regime" in the seventies to a modern and WTO compatible regime under the Trade Related Intellectual Property Rights System (TRIPS) in 2005. This paper estimates Total Factor Productivity Growth (TFPG) of Indian Pharmaceutical Industry (IPI) using firm level data from 2000 to 2013.

Design / Research methods: We have used nonparametric approach of Data Envelopment Analysis (DEA) using Biennial Malmquist Index.

Conclusions / findings: The results of estimation suggest an increase in overall TFPG of IPI after TRIPS agreement and also those vertically integrated firms involved in both bulk drugs production and formulation activities are less productive compared to firms that are involved in production of only bulk drug or formulation activity.

Originality / value of the article: This paper examines whether productivity of IPI has increased after 2005 i.e. after the period of TRIPS, by estimating TFPG for two sub-periods, i.e., from 2000 to 2005 and 2006 to 2013.

Implications of the research: The decomposition of TFPG suggests that for overall period 2000-2013, scale changes are the most important factor causing the productivity changes and among the other two alternative sources of TFPG, efficiency change dominates over technical changes. For the sub-period 2006-2013, the improvement in the scale efficiency may push the firms to a higher TFPG, whereas for

Correspondence address: Dr. Dipyaman PAL, Department of Economics, Bethune College, Kolkata, West Bengal, India, E-mail: dipyaman.pal@gmail.com, chandrimaeco@gmail.com, arpitaju@gmail.com
Received: 14.09.2017, Revised: 21.03.2018, Revised: 28.06.2018, Accepted: 28.06.2018
doi: http://dx.doi.org/10.29015/cerem.546

2000-2005 the better utilization of factors of production is the main driver of TFPG. A second stage panel regression suggests that R&D expenditure, Marketing expenditure, Market size, Capital-Labour ratio, import intensity and export intensity have positive and significant influence on TFPG.

Key words: Total Factor Productivity Growth (TFPG), TRIPS, Indian Pharmaceutical Industry, Data Envelopment Analysis (DEA), Biennial Malmquist Index.

JEL: L65, C14, C33

1. Introduction

The growth of Indian Pharmaceutical industry (IPI) can broadly be classified into three phases. *The first phase* corresponds to the period 1900-1970 which signifies the dominance of multinationals (MNCs) (the market share being 68% in 1970), which prevented the indigenous companies from producing new drugs, using the then existing patent law. Indigenous companies themselves were keener to process imported bulk drugs rather than developing the industry from basic stages. The size of the pharmaceutical industry was then very small as compared to its present status. *The second phase* corresponds to the period 1970-1990 witnessing the amendment of the Patent Act of 1911 which came into force in 1972. This change brought a renaissance to the IPI. After the changes in the patent law, large scale production of bulk drugs was started by the indigenous sector in the late 1970's, particularly in the 1980's, as a result of which first, imports were replaced and secondly, consumption increased significantly leading to the unprecedented growth in formulation activity. Exports started increasing steadily. Till 1987-1988, imports were larger than exports except for a few years but with steady increase in exports the country has become a net exporter since 1988-1989. The net results of this are that MNCs lost their market domination. The market share of the MNC's declined from around 60% in the late 1970's to around 40% in the early 1990's. The favourable environment attracted the entry of a number of new firms. *The third phase* corresponds to the period after 1990s when significant changes occurred in Pharmaceutical sector with introduction of trade liberalization measures. During the period 1990s some significant changes occurred in the Pharmaceutical sector with the introduction of trade liberalization measures like amendment of FERA and MRTP Acts and delicensing of the drugs, reserved for production by the public sector. During this period Government of India signed the Trade Related Intellectual

Property Rights System (TRIPS) agreement which came into existence with World Trade Organisation (WTO) established on 1 Jan1995 replacing The General Agreement on Tariff and Trade (GATT). The private sector growed rapidly along with increase in the competition among the domestic firms and foreign companies. As a result production of IPI increased manifoldalong with a sharp and steady increase in export. Net export as a percentage of total exports have also increased (Chaudhuri 2005). All those drugs which were reserved for the production by the public sector were delicenced in two stages. One immediate impact of this delicensing of the drugs was that production increased manifold besides increase in the competition among the domestic firms and foreign companies in 1990s. Both production and export have grown remarkably fast. There was sharp and steady increase of production and also of bulk drug production. As a result net exports as a percentage of exports have increased (Chaudhuri 2005).

Indian Pharmaceutical Industry (IPI) has undergone a massive makeover–from a modest beginning of "process patents regime" in the seventies to a modern and WTO compatible regime under the TRIPs Agreement in 2005.It ranked 3rd in volume and 14th in value in the global pharmaceutical market (Kalani, 2011). Since 2005, India has started full-fledged product patent regime in pharmaceuticals and are to develop new drugs themselves or to collaborate with the MNCs as manufacturing or marketing partners for the new drugs developed by the MNCs. (Chaudhuri 2005).

At this onset naturally the question arises that what happens to the total factor productivity growth (TFPG) in IPI especially after the TRIPs Agreement in 2005? The estimation of TFPG is thus essential, given the changed scenario of IPI in 2005. Following trade liberalization measures, improved performance of the industrial firms is now being called for an increase in productivity of a unit is now supposed to be a prerequisite for growth or even mere survival. In fact, government policies particularly after 2005, have gradually turned out to be less friendly to less productive firms. Such an analysis will definitely be helpful for framing appropriate policies for the development of IPI. The perusal of the literature on IPI signifies dearth in the studies dealing with these issues although some econometric studies are available on IPI (Lalitha 2002; Kumar 2001; Madanmohan 1997; Nagarajan, Barthwal 1990; Singh 1989; Chandrasekhar, Purkayastha 1982). The present paper adds the literature in this direction.

This paper uses non-parametric approach of DEA to estimation of TFPG. Analysis of TFPG as well as finding out their determinants is of greater research interest as such studies may help policy makers and managers to devise and implement policies that may enhance TFPG in this dynamic and globally competitive industry. Studies on TFPG related to IPI are few in number like Saranga and Banker (2010), Pannu, Kumar and Farooquie (2010), Kamiike, Sato and Aggarwal (2012), Ghose and Chakraborty (2012) among others.

Sarangaand Banker (2010) studied the productivity change and factors behind from 1994 to 2003 using DEA. They found that few innovative firms have pushed the production frontier thereby increasing technical and productivity gains. They argued that higher technical and R&D capabilities and wider new product portfolios of multinational-companies have contributed to positive technical and productivity changes. Whereas Pannu, Kumar and Farooquie (2010) using DEA analysed the impact of R&D and innovation on relative efficiency, productivity change and firm performance between 1998 and 2007. They found a positive impact of innovation and patents on productivity, market share, exports and ability to attract contract manufacturing. Study by Kamiike, Sato and Aggarwal (2012) using unit-level panel database analysed the impact of industry dynamics on TFPG across regions from 2000-01 to 2005-2006. They found that productivity growth is relatively higher in agglomerated region and effects of plant dynamics on productivity growth differ. Study by Ghose and Chakraborty (2012) estimated TFPG by estimating production function from 1973-1974 to 2003-2004, adjusted for stationarity after ADF-unit-root test. Translog form gave the better fit. Variation in TFPG is also explained.

Given this background, the **objectives** of the present paper are **first of all** to find out the TFPG of IPI by using Biennial Malmquist Index (BMI) of non-parametric method of Data Envelopment Analysis (DEA) for the period 2000 to 2013. This study also tried to find out the changed behavior of TFPG for IPI after 2005. After finding out the extent of TFPG, the **second objective** is to decompose TFPG into its different components: technical changes, efficiency changes and scale efficiency changes to check which component dominates over the other while finding out the major sources of TFPG. **Thirdly**, this paper tries to explain the factors behind the variation in TFPG of IPI.

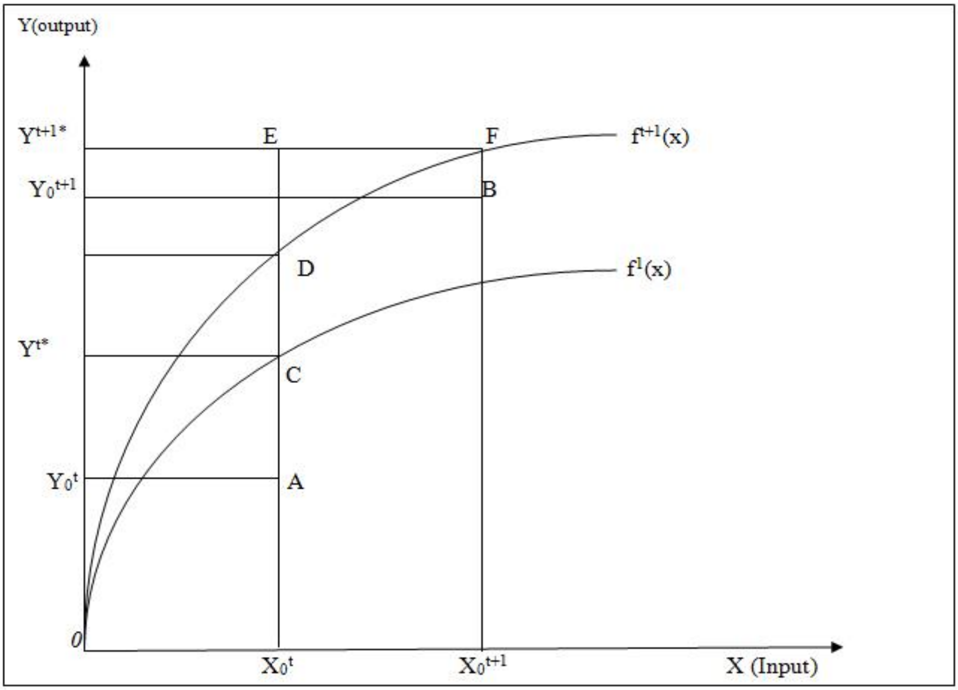
The paper unfolds as follows: Section 2 describes the methodology and data sources. The results of estimation are present in section 3. Section 4 concludes.

2. Methodology and Data Source

2.1 Methodology of measuring Biennial Malmquist Index (BMI)

In this paper we adopt the non-parametric (primal) approach to measure total factor productivity change. In the non-parametric approach, productivity index is used to measure productivity change.

Figure 1. Measurement and decomposition of productivity index



Source: Authors' own elaboration.

Figure 1 illustrates the measurement of productivity index and decomposition of it into above mentioned three components for a single input-single output case.

If in period t a firm produces output Y_0^t (Point A) from input X_0^t its productivity is

$$\pi_t = \frac{Y_0^t}{X_0^t} = \left(\frac{AX_0^t}{OX_0^t}\right) \qquad \ldots 1$$

Similarly, in period t+1, when output Y_0^{t+1} (Point B) is produced from input X_0^{t+1}, the productivity is

$$\pi_{t+1} = \frac{Y_0^{t+1}}{X_0^{t+1}} = \left(\frac{BX_0^{t+1}}{OX_0^{t+1}}\right) \ldots 2$$

The productivity change in the period t+1, with period t as the base is measured by

$$\pi_{t+1|t} = \frac{\pi_{t+1}}{\pi_t} = \frac{\left(\frac{BX_0^{t+1}}{OX_0^{t+1}}\right)}{\left(\frac{AX_0^t}{OX_0^t}\right)} \qquad \ldots 3$$

Now suppose that the production function is $Y^{t*} = f^t(X^t)$ in period t and$Y^{t+1*} = f^{t+1}(X^{t+1})$ in period t+1. Because each observed input-output bundle is by definition feasible in the relevant period, $f^t(X^t) \geq Y_0^t$and $f^{t+1}(X^{t+1}) \geq Y_0^{t+1}$. Thus the productivity index, as defined in (3), can be rewritten and decomposed as

$$\pi_{t+1|t} = \frac{\pi_{t+1}}{\pi_t} = \frac{\left(\frac{BX_0^{t+1}}{OX_0^{t+1}}\right)}{\left(\frac{AX_0^t}{OX_0^t}\right)}$$

$$= \frac{\left(\frac{BX_0^{t+1}}{FX_0^{t+1}}\right)}{\left(\frac{AX_0^t}{CX_0^t}\right)} X \frac{\left(\frac{FX_0^{t+1}}{OX_0^{t+1}}\right)}{\left(\frac{CX_0^t}{OX_0^t}\right)}$$

$$= \frac{\left(\frac{BX_0^{t+1}}{FX_0^{t+1}}\right)}{\left(\frac{AX_0^t}{CX_0^t}\right)} \left(\frac{DX_0^t}{CX_0^t}\right) \frac{\left(\frac{FX_0^{t+1}}{OX_0^{t+1}}\right)}{\left(\frac{DX_0^t}{OX_0^t}\right)}$$

$$= \left[\frac{\frac{Y_0^{t+1}}{f^{t+1}(X^{t+1})}}{\frac{Y_0^t}{f^t(X^t)}}\right] X \left[\frac{f^{t+1}(X^t)}{f^t(X^t)}\right] X \left[\frac{\frac{f^{t+1}(X^{t+1})}{X^{t+1}}}{\frac{f^{t+1}(X^t)}{X^t}}\right]$$

=TEC *X* TC *X* SEC …4

The first component in this expression (*TEC*) is the ratio of the technical efficiencies of the firm in two periods and captures the contribution of technical efficiency change over time. The second term (*TC*) shows how the maximum producible output from input X_0^t changes between period *t* and *t* +1 and captures the autonomous shift in the production function due to technical change. Finally the last term (*SEC*) identifies the returns to scale effect over time.

The Malmquist Productivity Index, introduced by Caves, Christensen, and Diewert (1982) and operationalized by Färe, Grosskopf, Lindgren and Roos (1992) (FGLR) to measure productivity change, is a normative measure based on a reference technology underlying observed input output data. Färe et al. (1992) (FGLR) decomposed the Malmquist Productivity Index (MPI) into technical change (TC) and technical efficiency change (TEC) considering the constant return to scale (CRS) frontier as the benchmark. However, assumption of global constant return to scale is not always a meaningful assumption about the underlying technology, so the FGLR decomposition is not particularly meaningful when CRS does not hold. In their paper Färe, Grosskopf, Norris, and Zhang (1994) re-modified and extended decomposition by considering variable returns to scale and isolate specific contributions of technical efficiency change (TEC), technical change (TC), and scale efficiency change (SEC) towards the overall productivity change. According to Ray and Desli (1997), this decomposition raises a problem of internal consistency because it uses CRS and variable returns to scale (VRS) within the same decomposition. They provide a modified decomposition by using the variable returns to scale frontier as a benchmark. In that decomposition, scale efficiency change is obtained by considering both the constant returns to scale technology and the variable returns to scale technology. However, when one estimate cross-period efficiency scores (which is measured by comparing actual output of a firm in period *t* with the maximum producible output from period *t* +1 input set) under a VRS technology, it may result in linear programming infeasibilities for some observations.

In 2011, Pastor, Asmild and Lovell provides a new Malmquist Index which is known as the Biennial Malmquist Index (BMI) which used the same decomposition as provided by Ray and Desli but it solved the infeasibility problem associated with the Ray-Desli decomposition of the Malmquist Index. Instead of using a contemporaneous production possibility frontier, they estimated the technical efficiency of a production unit with reference to a biennial production possibility frontier.

2.2 Non Parametric Estimation of Productivity Index

This study considers the non-parametric method of Data Envelopment Analysis (DEA) introduced by Charnes, Cooper and Rhodes (1978) and further generalized for variable returns to scale technology by Banker, Charnes and Cooper (1984) in order to measure and decompose the Malmquist index of total factor productivity.

The major advantage of using DEA is that, unlike in the parametric approach, there is no need to specify any explicit functional form for the production function (e.g., Cobb-Douglas or Translog) and mathematical programming techniques can be used to get point-wise estimates of the production function. In fact, DEA allows one to construct the production possibility set from observed input-output bundles on the basis of the following four assumptions:

a) All observed input-output combinations are feasible;
b) The production possibility set is convex;
c) Inputs are freely disposable; and
d) Outputs are freely disposable.

Now, consider an industry producing one output y^t from one input x^t in period *t*. The input output bundle (x^t, y^t) is considered as feasible if the output y^t can be produced from the input x^t . Let (x_j^t, y_j^t) represent the input-output bundle of firm *j*; and suppose that input-output data are observed for *n* firms. Then, based on the above assumptions, in period *t*, the production possibility set showing a variable returns to scale (VRS) technology is

$$T_v^t = \left\{ (x, y) : x \geq \sum_{j=1}^{n} \lambda_j x_j^t ; y \leq \sum_{j=1}^{n} \lambda_j y_j^t ; \sum_{j=1}^{n} \lambda_j = 1; \ \lambda_j \geq 0; (j = 1,2,3, \dots n) \right\}$$

Under the constant returns to scale (CRS) assumption, if any (*x*, *y*) is feasible, so is the bundle (*kx*,*ky*) for any *k* > 0. The production possibility set then becomes

$$\boldsymbol{T_c^t} = \left\{(\boldsymbol{x},\boldsymbol{y}): \boldsymbol{x} \geq \sum_{j=1}^{n} \lambda_j x_j^t; y \leq \sum_{j=1}^{n} \lambda_j y_j^t; \lambda_j \geq 0; (j = 1,2,3, \ldots n)\right\}$$

One can measure the output-oriented technical efficiency $TE^t(x_s^t, y_s^t)$ of a firm *s* in period *t* by comparing its actual output y_s^t with the maximum producible quantity from its observed input x_s^t. Therefore, the output-oriented technical efficiency of firm *s* in period *t* is

$TE^t(x_s^t, y_s^t) = \frac{1}{\theta_s^*}$; where $\theta_s^* = max \quad \theta : (x_s^t, \theta y_s^t) \in T^t$ and T^t is the period t production possibility set.

An alternative characterization of technical efficiency in terms of the Shephard Distance Function is $D^t(x_s^t, y_s^t) = \min \lambda : \left(x_s^t, \frac{1}{\lambda} y_s^t\right) \in T^t$. It can be seen that $\lambda = \frac{1}{\theta_s^*}$

Caves et al. (1982) defined the Malmquist Productivity Index as the ratio of the period *t* and period *t* +1 output-oriented Shephard distance functions pertaining to a certain benchmark technology. Equivalently, the Malmquist Index of total factor productivity of the firm *s* is

$$\boldsymbol{M_c}(x_s^t, y_s^t;\ x_s^{t+1}, y_s^{t+1}) = \left[\frac{TE_c^{t+1}(x_s^{t+1}, y_s^{t+1})}{TE_c^{t+1}(x_s^t, y_s^t)} \frac{TE_c^t(x_s^{t+1}, y_s^{t+1})}{TE_c^t(x_s^t, y_s^t)}\right]^{\frac{1}{2}} \qquad \ldots 5$$

The standard non-parametric DEA model used to estimate the period *t* output-oriented technical efficiency of a firm *s*, relative to contemporaneous CRS frontier is

$\theta_s^* = max\ \theta$

Subject to $\sum_{j=1}^{n} \lambda_j y_j^t \geq \theta y_s^t$;

$$\sum_{j=1}^{n} \lambda_j x_j^t \leq x_s^t;$$

$\lambda_j \geq 0; (j = 1,2,3, \ldots n)$;

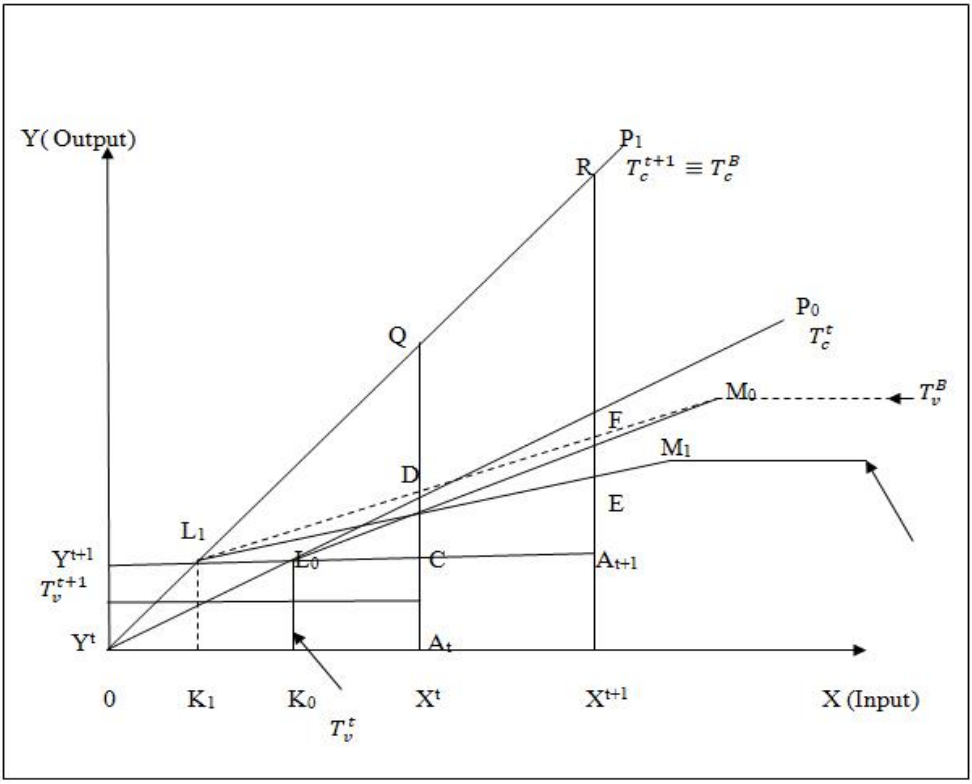
And $TE^t(x_s^t, y_s^t) = \frac{1}{\theta_s^*}$

By imposing the additional restriction $\sum_{j=1}^{n} \lambda_j = 1$ in this DEA model, period t out-put oriented technical efficiency under VRS technology of a firm s can be estimated as $TE_v^t(x_s^t, y_s^t)$.

It has been already mentioned that the Biennial Malmquist Index introduced by Pastor, Asmild, and Lovell (2011) provides the same decomposition and avoids the infeasibility problem associated with the Ray-Desli decomposition of the Malmquist Index.

Instead of using a contemporaneous production possibility frontier, they estimated the technical efficiency of a production unit with reference to a biennial production possibility frontier. So in order to understand the Biennial Malmquist Index one has to first construct the Biennial Production Possibility Frontier.

Figure2. Simple graphical illustration of biennial production possibility frontier



Source: Authors' own elaboration.

Figure 2 provides an illustration of the biennial production possibility frontier and measure of output-oriented technical efficiency with reference to it for a firm, producing a single output from a single input, observed in two time periods *t* and *t+1* (point A and B respectively). The VRS frontiers for period *t* and *t* +1 are indicated by $K_0L_0M_0$- extension and K_1 L_1M_1- extension respectively. The rays through origin $OP_0$and OP_1represent the CRS frontiers for period *t* and period *t* +1respectively. The biennial VRS frontier is indicated by the broken lineK_1 L_1 DFM_0- extension and the biennial CRS frontier in this case coincides with that of period *t*+1. Output-oriented technical efficiency of the firm with reference to CRS biennial frontier in period *t* is

$TE_c^B(x^t,y^t) = \left(\frac{A_tX^t}{QX^t}\right)$and that for period t+1 is $TE_c^B(x^{t+1},y^{t+1}) = \left(\frac{A_{t+1}X^{t+1}}{RX^{t+1}}\right)$.

Similarly with reference to the VRS biennial frontier,$TE_v^B(x^t,y^t) = \left(\frac{A_tX^t}{DX^t}\right)$ and $TE_v^B(x^{t+1},y^{t+1}) = \left(\frac{A_{t+1}X^{t+1}}{FX^{t+1}}\right)$ show the levels of technical efficiency for the period t and t+1 respectively. The reference technology set $T^B$is defined as the convex hull of pooled data from both period *t* and *t* +1.

Using the output-oriented technical efficiency scores with reference to a CRS biennial frontier, the Biennial Malmquist Productivity Index of the firm *s* producing a single output from multiple inputs is measured as (Since the Biennial Malmquist Index of productivity uses the biennial CRS production possibility set, which includes the period *t* and *t+1* sets, one need not to calculate a "geometric mean" of two productivity indexes while measuring it)

$$M_c^B(x_s^t,y_s^t;\ x_s^{t+1},y_s^{t+1}) = \frac{TE_c^B(x_s^{t+1},y_s^{t+1})}{TE_c^B(x_s^t,y_s^t)} \qquad \ldots 6$$

The decomposition of this Biennial Malmquist productivity index is

$$M_c^B(x_s^t,y_s^t;\ x_s^{t+1},y_s^{t+1}) = TEC\ X\ TC\ X\ SEC \qquad \ldots 7$$

Where

$$\text{TEC}= \frac{TE_v^{t+1}(x_s^{t+1},y_s^{t+1})}{TE_v^t(x_s^t,y_s^t)}, \qquad \ldots 8$$

$$\text{TC} = \frac{TE_v^B(x_s^{t+1},y_s^{t+1})/TE_v^{t+1}(x_s^{t+1},y_s^{t+1})}{TE_v^B(x_s^t,y_s^t)/TE_v^t(x_s^t,y_s^t)}, \text{ and} \qquad \ldots 9$$

$$\text{SEC}= \frac{TE_c^B(x_s^{t+1},y_s^{t+1})/TE_v^B(x_s^{t+1},y_s^{t+1})}{TE_c^B(x_s^t,y_s^t)/TE_v^B(x_s^t,y_s^t)} \qquad \ldots 10$$

Now from figure 2 one can define Biennial Malmquist productivity index as

$$M_c^B(x_s^t,y_s^t;\ x_s^{t+1},y_s^{t+1}) = \frac{\left(\frac{A_{t+1}X^{t+1}}{RX^{t+1}}\right)}{\left(\frac{A_tX^t}{QX^t}\right)}$$

The decomposition of this Malmquist productivity index is

$$M_c^B(x_s^t, y_s^t;\ x_s^{t+1}, y_s^{t+1}) = \frac{(A_{t+1}X^{t+1}/EX^{t+1})}{(A_tX^t/CX^t)} \text{ x } \frac{(A_{t+1}X^{t+1}/FX^{t+1})/(A_{t+1}X^{t+1}/EX^{t+1})}{(A_tX^t/DX^t)/(A_tX^t/CX^t)} \text{ x } \frac{(A_{t+1}X^{t+1}/RX^{t+1})/(A_{t+1}X^{t+1}/FX^{t+1})}{(A_tX^t/QX^t)/(A_tX^t/DX^t)}$$

Where

$$\text{TEC}=\frac{(A_{t+1}X^{t+1}/EX^{t+1})}{(A_tX^t/CX^t)}$$

$$\text{TC} = \frac{(A_{t+1}X^{t+1}/FX^{t+1})/(A_{t+1}X^{t+1}/EX^{t+1})}{(A_tX^t/DX^t)/(A_tX^t/CX^t)}$$

$$\text{SEC} = \frac{(A_{t+1}X^{t+1}/RX^{t+1})/(A_{t+1}X^{t+1}/FX^{t+1})}{(A_tX^t/QX^t)/(A_tX^t/DX^t)}$$

The appropriate DEA model to estimate period *t* output-oriented technical efficiency $TE_c^B(x_s^t, y_s^t)$ of firm *s*, with reference to a CRS biennial production possibility set is

$$\varphi_s^* = \max \varphi$$

Subject to $\sum_{k=t,t+1}\sum_{j=1}^{n_k} \lambda_j^k y_j^k \geq \varphi y_s^t;$

$$\sum_{k=t,t+1}\sum_{j=1}^{n_k} \lambda_j^k x_j^k \leq x_s^t;$$

$$\lambda_j^k \geq 0;$$

Where n_k is the number of observed firm in the period k and $TE_c^B(x_s^t, y_s^t) = \frac{1}{\varphi_s^*}$

Period t output oriented technical efficiency $TE_v^B(x_s^t, y_s^t)$ of firm *s*, with reference to a VRS biennial production possibility set is

$$\varphi_s^* = \max \varphi$$

Subject to $\sum_{k=t,t+1}\sum_{j=1}^{n_k} \lambda_j^k y_j^k \geq \varphi y_s^t;$

$$\sum_{k=t,t+1}\sum_{j=1}^{n_k} \lambda_j^k x_j^k \leq x_s^t;$$

$$\sum_{k=t,t+1}\sum_{j=1}^{n_k} \lambda_j^k = 1;$$

$$\lambda_j^k \geq 0;$$

Where n_k is the number of observed firm in the period k and $TE_c^B(x_s^t, y_s^t) = \frac{1}{\varphi_s^*}$

2.3 Determinants of TFPG

The variables used as possible determinants of TFPG are export intensity (E), import intensity (I), Market size (MS), R&D expenditure (RD), marketing expenditure (M) and Capital-Labour ratio (K/L).

Export intensity (E): One of the important characteristics of IPI is that they re-engineer the imported technology and then re-export the product. Export plays a very important role for the growth in the pharmaceutical sector. [Theoretical and empirical literature supporting positive role of exports Goldar et. al. (2004), Ray (2006)]. The findings of Indian studies are mixed and industry specific, even during post-reform period. IPI exports a lot (Chaudhuri 2005). There is a common opinion that international export enhances economic growth of involved firms (see Balassa 1988). Economic policies under export-led growth strategy have been widely supported on the argument that exposure to international market through export helps to increase growth of exporters. Similarly, advocates of endogenous growth theory believe that export plays a crucial role by improving productivity and hence growth through innovation (Grossman, Helpman 1991) and technology transfer (Barro, Sala-i-Martin 1995). Through participation in export, growth can occur as a result of many factors such as capital accumulation, adoption of new technologies, research and development, changes in the organization of firms, etc. Export intensity is defined by Export as a ratio to total sales

Import intensity (I): IPI also imports a lot of goods especially intermediate goods. The imported intermediary good is an important channel through which technological diffusion takes place (see Tybout 2000); this may affect productivity and growth favorably. Imports allow countries to take advantage of other countries technology embodied in imported inputs. Suffice here to mention theories of import-led growth due to Grossman and Helpman (1991). The removal of quantitative restrictions on imports and lowering of customs duties in the post liberalization era of the Indian economy should have improved access of imported raw materials and capital goods. Imports of materials embodying latest technologies should foster productivity, efficiency and the growth of the firms. Goldar et al. (2004) and Mazumder et al. (2012) reported positive relation between efficiency and imports. Import intensity is defined by import as a ratio to total sales

Market size (MS): It can be argued that higher the Market Size (MS) less is the competition. MS captures the effect of market structure on TFPG. Some studies argued that a negative relation between MS and TFP growth may occur because as MS falls, competition increases which may lead to cost-consciousness and drive for technological advancement. Others may point out that the advantages of big size, secured market and expect a positive association between MS and TFPG because as MS rises, competition falls, larger units are becoming more productive may be due to the advantages of big size and secured market. The conclusion from the empirical literature also varies and does not provide us a single answer (Katz 1969; Kendrick 1973). MS is obtained for each firm considering the ratio of total sales of each firm to total sales of Pharmaceutical industry.

R&D Expenditure (RD): In recent years theoretical models related to endogenous growth give emphasis to that R&D expenditures of individual firms contribute to unremitting long run growth of an economy through their industry-wide spillover effect (Grossman, Helpman 1990a, 1990b) because as individual firms invest in R&D for private knowledge that enhances their productivity and profit. Private knowledge of individual firms then spills over to the rest of the industry and becomes social knowledge which acts as external effects in enhancing the productivity of the firms. With this positive spill-over effect of R&D, a constant or decreasing returns to scale aggregate production function may exhibit increasing returns to scale and thus may lead to sustained long run growth (Raut, Srinivasan 1993). Again, Cohen and Levinthal (1989) among others pointed out that even as knowledge from private R&D capital spills over to create social or public domain knowledge, a firm must invest in R&D to obtain the technical capability needed to make use of the public domain knowledge to improve its productivity and efficiency. One explanation of this later view is that industry-wide knowledge will not contribute to productivity gains unless the firm invests in R&D. The technological capabilities approach also pointed out that the firm level technological capabilities in developing countries are formed through slight innovations which include incremental modifications in the plants and machineries, efficiently using technologies, imitation, absorption and adaptation of imported technologies etc. These small modifications are largely generated by firms' in-house R&D efforts and the development of human resources and skills, notably on job training (Lall 2000). Thus question can be raised to what extent increase in R&D helps to promote productivity in this IPI sector? The present article tests this hypothesis empirically, where R&D activity is measured by R&D expenditure which is defined by R&D expenditure as a ratio to total sales.

Marketing Expenditure (M): This variable is measured by Marketing Expenditure as a ratio to total sales and it also serves as a proxy for product differentiation. Sheth and Sisodia (2002) argued that low productivity is due to the descending of marketing efficiency. Their study point out that some changes are needed at the corporate level and the most fundamental one is that corporations should treat marketing as an investment rather than an expense. Kao et al. (2006) evaluates Technical and Allocative Efficiency in Marketing and explains the positive relation between return and marketing expenditure, which is defined as a kind of investment. The return can be in the form of increased sales, or customers, or some form of infrastructure that makes acquiring these items easier.

Degree of mechanization (K/L): Degree of mechanization is captured by Capital-labor ratio (K/L) which serves as a technological variable[1]. Generally, positive relationship between K/L and TFP growth is expected with the argument that capital-intensive technology or sophisticated, advanced technology will facilitate productivity growth by encouraging learning by doing. Thus it is interesting to test the hypothesis that whether more the firm is capital intensive higher may be TFPG. Ray (1997) found a positive relationship between these two. Whereas Ahluwalia (1991) find negative association between these two variables and argued that the industries with higher capital-labour ratio were the heavy industries under the public sector which places constraints on the operation of these industries with it's adverse impact on productivity.

Dummy Variables for formulation (D_F) and both formulation & bulk drugs (D_{FB}): Some of the studies like Chaudhuri (2012) argued that Imports of high priced finished formulations are expanding rapidly with manufacturing investments lagging behind. The aggregate market share of the MNCs in the formulations market has gone up dramatically with the taking over of some Indian companies by the MNCs.

---

[1] K/L is considered as determinant of TFP growth by Ray (1997) and Bandyopadhyay (2000).

Thus it is necessary to check whether vertically integrated firm involved in both bulk drugs production and formulation activities are less productive or not as compared to firms that produce only bulk drugs or the firms doing formulation. For finding out whether vertically merged firms are doing well in terms of TFPG compared to the firms not merged vertically, two dummy variables have been defined, one for firms engaged in formulation (D_F) and another for firms producing both formulation & bulk drugs (D_{FB}) taking firms engaged in bulk drugs production as the reference category. So D_F=1, for firms engaged in formulation and 0 otherwise; similarly D_{FB}=1 for firms producing both formulation & bulk drugs and 0 otherwise.

Time Dummy (D_T): Also time dummy variables D_T is defined as D_T =1 for 2006 to 2013 and 0 otherwise to capture the effect of TRIPS on TFPG.

It is expected that firms which are incurring more RD can increase their TFPG by expanding their information set. Again by spending on M a firm can increase its market share. Hence sales increases and the firm may try to increase its production with more efficient technologies. Also K/L and TFPG may found to be positively linked which may imply that the industry perhaps is conducive for capital-intensive production process. MS has positive role on the TFPG which implies that an increase in market size will improve TFPG of IPI may be due to easier access of quality inputs and getting advantage of scale economies. It may be quite evident that units which are enjoying more export per unit of output are more productive than others as they are learning suitable measures to improve their productivity level. Also one may expect positively affect TFPG as one of the purposes of doing import in IPI is to carry out the re-export process.

For finding out the determinants of TFPG, a panel regression analysis has been done using a seemingly unrelated regression (SUR) framework where each regression was adjusted for contemporaneous correlation (across units) and cross section heteroscedasticity and test for better model-fixed/random with Hausman test is done. In this paper SUR framework has been considered because since we are considering export intensity (E), import intensity (I), R&D expenditure (RD), marketing expenditure (M) and Capital-Labour ratio (K/L) among the determinants of TFPG, it is quite possible that the for IPI, decision to undertake export intensity

(E), import intensity (I), R&D expenditure (RD), marketing expenditure (M) and Capital-Labour ratio (K/L) by one firm may influence the decision of the same by other firms and hence the error term explaining the TFPG of one firm may be correlated with the error explaining the TFPG for the others.

2.4 The Data sources

The present study uses CMIE Prowess data base and those firms are selected for which all the data of inputs and outputs and the determinants are positive. On the basis of this fact, 90 firms have been selected. The time period is 2000 to 2013.

3. Results of estimation

3.1 Results of TFPG

The TFPG for each of the years and also each firms are estimated. The results are then summarized to generate the information regarding the changes of TFPG for each year. Such estimation results are presented in Table 1. To capture the effect of TRIPS this paper divides total sample period in to two sub-periods, i.e., from 2000 to 2005 and 2006 to 2013 and compare the estimated values of TFPG for these periods. These results are also presented in Table 1.

From Table 1 it can be concluded that there has been an increased in the TFPG over the total sample period. Not only that values of TFPG increased in the second half, i.e. after TRIPS as compare to 2000 to 2005. So, it can be concluded that an increase in overall TFPG of IPI after TRIPS agreement is evident.

3.2 Results of Decomposition of TFPG

The estimated results of TFPG are then decomposed into Efficiency Changes, Scale Efficiency Changes and Technical Changes following the formula 7 to 10. For each of the year, the overall changes in the decomposition of TFPG as well as changes over the period from 2000-2005 and 2006-2013 are estimated.

Table 1. Results of TFPG of IPI

YEAR	TEC	TC	SEC	MI
2000	0.995147	1	1.002609	0.997743
2001	1	1	1.014529	1.014529
2002	1	1	0.998382	0.998382
2003	1	1	1	1
2004	1.097111	1.002249	0.989028	1.087514
2005	1	1	1.000529	1.000529
AVERAGE (2000 TO 2005)	**1.015376**	**1.000375**	**1.000846**	**1.01645**
2006	1.032924	1	1.040142	1.074387
2007	1	1	1.002688	1.002688
2008	1	1	1.071032	1.071032
2009	1	1	1.000456	1.000456
2010	1.0178	1.000506	1.020663	1.039357
2011	1	1	1.000697	1.000697
2012	1.005563	1.00018	1.008582	1.014375
2013	1.006945	1.000172	1.007028	1.014196
AVERAGE (2006 TO 2013)	**1.007904**	**1.000107**	**1.018911**	**1.027149**
OVER ALL AVERAGE	**1.011106**	**1.000222**	**1.011169**	**1.022563**

Source: Authors' own elaboration.

All the results are presented in Table-1. Entries in column TEC show average annual changes in the level of technical efficiency over time, a value greater than unity for this component implies that, for that particular year IPI has experienced improvement in technical efficiency over the previous period. Similarly, an entry with value greater (less) than unity in column TC reflects technological progress (regress) over time. The change in scale efficiency over time is reported in column SEC, with a value exceeding one again signaling an improvement in scale efficiency. From the results of Table 1, it can be concluded that productivity growth is mostly driven by the change in the scale efficiency for the entire sample period. The second important factor behind the changes in TFPG is the change in the technical efficiency. The change in the technology has the lowest impact on the increase in the productivity. So it can be concluded that changes in the scale and better utilization of factors of production may pushed IPI to be on higher TFPG for the period 2000 to 2013.

Now if one consider for the period 2000 to 2005 it can be concluded that the change in the technical efficiency is major factor behind the increase in TFPG. Productivity growth is also driven by the change in the scale efficiency for this period. Again, change in the technology has the lowest impact on the increase in the productivity. For the period 2006 to 2013 productivity growth has increased mostly for the change in the scale efficiency followed by the change in the technical efficiency. The change in the technology has again the lowest impact on the increase in the productivity.

Thus in conclusion it can be said that scale changes are the most important factor causing the productivity changes for IPI. Among the two other alternative sources of TFPG, an efficiency change dominates over technical changes. Thus, in case of IPI, the improvement in the technical efficiency may push the firms to a higher TFPG for the period 2006-2013. In case of 2000-2005 the better utilization of factors of production is the main factor behind the improvement in TFPG.

3.3 Results of Determinants of TFPG

All the results of a second stage panel regression are presented in Table 2. The estimated model also reports Adjusted R^2 which represents the overall fit of the model, which is based on the difference between residual sum of squares from the estimated model and the sum of square from a single constant only specification, not from a fixed effect only specification. High value of Adjusted R^2 shows that the fitted models are reasonably good.

From the results of estimation of growth equation it can be concluded that there exists an inverted U shape relationship between TFPG and export intensity, capital-labour ratio and market size implying that there exists a threshold limit beyond these variables may affect the TFPG in reverse way. The overall marginal effects of all the determinants are positive implying that on a whole these determinants may increase the TFPG. So, it can be concluded that Capital-Labour ratio, market size and export intensity have positive and significant influence on the TFPG. The positive linkage between Capital-Labour ratio and TFPG may imply that this industry perhaps is conducive for capital-intensive production process. The relation between export intensity and TFPG is obtained to be positive suggesting that those units which are enjoying more export per unit of output are more efficient than others. In IPI one of the purposes of doing import is to carry out the re-export process. The effect of import intensity on TFPG of IPI is positive and significant. Market size has positive role on the TFPG which may imply that big firms are grown faster than the large firm. The effect of R&D expenditure and Marketing expenditure on TFPG is positive and statistically significant although these relationships are not linear in nature.

Table 2. Estimated results of Second Stage Panel Regression

Dependent Variable: BMI				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	4.4912*	0.6555	6.8518	0.0000
E	0.1673*	0.0391	4.2783	0.0000
I	0.0170*	0.0027	6.3775	0.0000
RD	0.2654*	0.0603	4.4022	0.0000
M	2.3586*	0.9084	2.5964	0.0096
KBYL	0.1052*	0.0257	4.0977	0.0000
MS	7.5635*	3.1272	2.4186	0.0053
E2	-0.0024**	0.0012	-1.9787	0.0481
KBYL*KBYL	-0.0009*	0.0004	-2.4087	0.0061
MS2	-4.3360	2.8923	-1.4992	0.1810
DF	1.7439*	0.1711	10.1929	0.0000
DBOTH	-1.0458*	0.4618	-2.2645	0.0238
DT	0.8552*	0.4149	2.0612	0.0396
Adjusted R-squared	0.7515			
F-statistic	56.088692*			
Prob(F-statistic)	0			

*Significant at 1%; ** significant at 5%; *** Significant at 10%

Table 3. Marginal effects

E	0.1362
I	0.0170
RD	0.2654
M	2.3586
KBYL	0.3157
MS	7.2025

Source: Authors' own elaboration.

The dummy for vertically integrated firms involved in both bulk drugs production and formulation activities is negative and significant whereas the dummy for firms involved in formulation activity is positive and significant implying that those vertically integrated firms involved in both bulk drugs production and formulation activities are less productive compared to firms that produces only bulk drug or are involved in formulation activity. The coefficient of time dummy is positive and statistically significant implying that for the period 2006-2013 TFPG has increased as compared to the period 2000-2005.

4. Conclusion

Indian Pharmaceutical Industry (IPI) is one of the few industries which has been affected in a major way due to Trade Related Intellectual Property Rights System (TRIPS) agreement as from the year 2005 the existing Process Patent regime gave way to the Product Patent regime although the process of establishing a new patent regime in India started since 1995. In such an environment it will be interesting to examine whether there has been any improvement in the productivity of IPI after 2005 i.e. after the period of TRIPS. So, the paper estimates Total Factor Productivity Growth (TFPG) of Indian Pharmaceutical Industry (IPI) using firm level data from 2000 to 2013. TFPG is estimated by nonparametric approach of Data Envelopment Analysis (DEA) using Biennial Malmquist Index. To capture the effect of TRIPS this paper divides total sample period in to two sub-periods, i.e., from 2000 to 2005 and 2006 to 2013. An increase in overall TFPG of IPI after TRIPS agreement is evident. The decomposition analysis of TFPG suggests that scale changes are the most important factor causing the productivity changes for IPI. Among the two other alternative sources of TFPG, an efficiency change dominates over technical changes. In case of IPI, the improvement in the scale efficiency may push the firms to a higher TFPG for the period 2006-2013. In case of 2000-2005 the better utilization of factors of production is the main factor behind the improvement in TFPG.

This study pointed out those vertically integrated firms involved in both bulk drugs production and formulation activities are less productive compared to firms that produces only bulk drug or are involved in formulation activity. Also, for the period 2006-2013 TFPG has increased as compare to the period 2000-2005.

A second stage panel regression suggests that the determinants R&D expenditure, Marketing expenditure, Market size, Capital-Labour ratio, import intensity and export intensity have positive and significant influence on the TFPG implying that an increase in either of these variables can boost up TFPG of Indian Pharmaceutical Industry.

Thus this result reveals that although the TRIPS agreement may push the TFPG of IPI in a higher level but also in order to encourage total factor productivity growth, any policy changes that will lead to increase in the export intensity, import intensity, Market size, R&D expenditure, marketing expenditure and Degree of mechanization should be emphasized.

References

Ahluwalia I.J. (1991), Productivity growth in Indian manufacturing, Oxford University Press, New Delhi.

Balassa B. (1988), Outward orientation, in: Handbook of development economics, Chenery H., Srinivasan T.N. (ed.), vol. 2, Elsevier, New York, pp. 1645-1689.

Bandyopadhyay S. (2000), Impact of efficiency indicators on the growth of productivity. A survey and empirical evidence from India, „Margin", vol. 33 no. 1, pp. 84-98.

Banker R.D., Charnes A., Cooper W.W. (1984), Some models for estimating technical and scale efficiencies in Data Envelopment Analysis, „Management Science", vol. 30 no. 9, pp. 1078-1092.

Barro R.J., Sala-i-Martin X. (1995), Economic growth, McGraw-Hill, New York.

Caves D., Laurits W., Christensen R., Diewert W.E. (1982), The economic theory of index numbers and the measurement of input, output and productivity, „Econometrica", vol. 50 no. 6, pp. 1393-1414.

Chandrasekhar C.P., Purkayastha P. (1982), Transfer pricing in the Indian drug industry. An estimate and it's implications, „Social Scientist", vol. 10 no. 1, pp. 3-10.

Charnes A., Cooper W.W., Rhodes E.L. (1978), Measuring the efficiency of decision-making units, „European Journal of Operational Research", vol. 2 no. 6, pp. 429-444.

Charnes A., Cooper W.W., Rhodes E.L. (1981), Evaluating program and managerial efficiency – an application of Data Envelopment Analysis to Program Follow Through, „Management Science", vol. 27 no.6, pp. 668-697.

Chaudhuri S. (2005), The WTO and India's pharmaceuticals industry. Patent protections, TRIPS and developing countries, Oxford University Press, New Delhi.

Chaudhuri S. (2012), Multinationals and monopolies. Pharmaceutical industry in India after TRIPS, „Economic & Political Weekly", vol. 47 no. 12, pp. 46-54.

Cohen W.M., Levinthal D.A. (1989), Innovation and learning. Two faces of R&D, „Economic Journal", no. 99, pp. 569-596.

Färe R., Grosskopf S., Lindgren B., Roos P. (1992), Productivity changes in Swedish pharmacies 1980-89. A non-parametric Malmquist approach, „Journal of Productivity Analysis", vol. 3 no. 1/2, pp. 85-101.

Färe R., Grosskopf S., Norris M., Zhang P. (1994), Productivity growth, technical progress and efficiency changes in industrial countries, „American Economic Review", vol. 84 no. 1, pp. 66-83.

Ghose A., Chakraborty C. (2012), Total factor productivity growth in pharmaceutical industry. A look using modern time series approach with Indian data, „The Journal Of Industrial Statistics", vol. 1 no. 2, pp. 250-268.

Goldar B., Ranganathan V.S., Banga R. (2004), Ownership and efficiency in engineering firms, 1990-91 to 1999-2000, „Economic and Political Weekly", vol. 39 no. 5, pp. 441-447.

Grossman G., Helpman E. (1990a), Comparative advantage and long run growth, „American Economic Review", vol. 80 no. 4, pp. 796-815.

Grossman G., Helpman E. (1990b), Trade, innovation and growth, „American Economic Review", vol. 80 no. 2, pp. 86-91.

Grossman G., Helpman E. (1991), Innovation and growth in the global economy, The MIT Press, Cambridge.

Kalani S. (2011), Emerging role of IT in Indian pharmaceutical industry and its implications for the education system, „Pharma Times", vol. 43 no. 5, pp. 25-27.

Kamiike A., Sato T., Aggarwal A. (2012), Productivity dynamics in the Indian pharmaceutical industry. Evidences from plant-level panel data, „Science, Technology and Society", vol. 17 no. 3.

Kao L.-J., Chiu Ch.-Ch., Gilbride T.J., Greg T.O., Allenby M. (2006), A direct approach to evaluating technical and allocative efficiency in marketing, Fisher College of Business, Ohio State University, www.stat.osu.edu/~amd/papers/Efficiency.pdf [08.09.2018].

Katz J.M. (1969), Production function, foreign investment and growth, North-Holland Publishing Company, Amsterdam.

Kendrick J.W. (1973), Post war productivity trends in the United States, 1947-1969, National Bureau of Economic Research, New York.

Kumar N. (2001), WTO regime, host country policies and global patterns of MNE activity. Recent quantitative studies and India's strategic response, „Economic and Political Weekly", vol.36 no.1, pp 39-50.

Lalitha N. (2002), Indian pharmaceutical industry in WTO regime. A SWOT analysis, „Economic and Political Weekly", vol. 37 no. 34, pp. 3542-3555.

Lall S. (2000), Technological change and industrialization in the Asian newly industrializing economies, in: Technology, learning, and innovation. Experiences of newly industrializing economies, Kim L., Nelson R.R. (ed.), Cambridge University Press, Cambridge, pp. 13-68.

Madanmohan T.R. (1997), Exit strategies. Experience of Indian pharmaceutical firms, „Economic and Political Weekly", vol. 32 no.48, pp. 107-110.

Mazumder M., Rajeev M., Ray S.C. (2012), Sources of heterogeneity in the efficiency of Indian pharmaceutical firms, „Indian Economic Review", vol. 47 no. 2, pp. 191-221.

Nagarajan M., Barthwal R.R. (1990), Profitability and structure. A firm level study of the Indian pharmaceutical industry, „Indian Economic Journal", vol. 38 no. 2, pp. 70-81.

Pannu H.S., Kumar U.D., Farooquie J.A. (2010), Impact of innovation on the performance of Indian pharmaceutical industry using Data Envelopment Analysis, IIM Bangalore Research Paper no. 302.

Pastor J.T., Asmild M., Lovell C. (2011), The biennial Malmquist productivity change index, „Socio-Economic Planning Sciences", vol. 45 no. 1, pp. 10-15.

Raut L.K., Srinivasan T.N. (1993), Theories of economic growth, old and new, in: Capital investment and development, Basu K., Majumdar M., Mitra T. (ed.), Basil Blackwell, New York.

Ray S.C. (1997), Regional variation in productivity growth in Indian manufacturing. A nonparametric analysis, „Journal of Quantitative Economics", vol. B no.1, pp. 73-94.

Ray S.C. (2006), The changing role of technological factors in explaining efficiency in Indian firms, „Journal of Developing Areas", vol. 40 no. 1, pp. 127-140.

Ray S.C., Desli E. (1997), Productivity growth, technical progress, and efficiency change in industrialized countries. Comment, „The American Economic Review", vol. 87 no. 5, pp. 1033-1039.

Saranga H., Banker R.D. (2010), Productivity and technical changes in the Indian pharmaceutical industry, „Journal of the Operational Research Society", vol. 61 no. 12, pp. 1777-1788.

Sheth J.N., Sisodia R.S. (2002), Marketing productivity. Issues and analysis, „Journal of Business Research", vol. 55 no. 5, pp. 349-362.

Singh N. (1989), Trade behaviour of firms. An empirical study of the size and foreign collaboration effects, „Indian Economic Review", vol. 24 no. 1, pp. 45-65.

Tybout J.R. (2000), Manufacturing firms in developing countries. How well do they do and why? „Journal of Economic Literature", vol. 38 no. 1, pp. 11-44.

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 81-99, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Malmquist Data Envelopment Analysis as a tool to evaluate the productivity levels of container ports in developing countries located in east and southern Africa

Adeola OSUNDIRAN, Felix OKONTA

University of Johannesburg, South Africa

Abstract:

Aim: The purpose of this paper is to examine the productivity of 12 container ports located in East and Southern African developing nations for the period of 2014-2016. Furthermore, to investigate the sources of productivity change over the time period.

Design / Research methods: This research collects data on the 12 container ports. The productivity of these ports is analyzed using the Data Envelopment Analysis based Malmquist productivity index. This is decomposed into technological changes and technical efficiency. The sources of productivity change are identified.

Conclusions /findings: The major finding of this study is the trend in the port efficiency level over the three year period of analysis. Therefore assisting maritime policymakers and port authorities on what aspect of the port production need enhancement.

Originality/value of the article: Evaluation of ports in developing nations in Africa is not common. Also, the year under examination is less than five years. Therefore the result is relevant to port authorities as well as to the African nations.

Implications of the research: 90% of import and exports into developing African nations are done by sea. The implication of this is that an efficient or inefficient port will have a multiplier effect on the nation's economy. Great improvement in port productivity will enhance economic growth and development.

Limitations of the research: Port efficiency should be evaluated on a yearly basis to serve as a major determinant of port productivity. However, this evaluation is based on availability of data.

Key words: Ports, Data Envelopment Analysis, African Nations, Developing, Malmquist.

JEL: C33, L9

Correspondence address: Adeola OSUNDIRAN, Felix OKONTA,. E-mail: 216088926@student.uj.ac.za, fnokonta@uj.ac.za
Received: 02.10.2017, Revised: 08.02.2018, Revised: 01.03.2018, Accepted: 17.03.2018
doi: http://dx.doi.org/10.29015/cerem.573

1. Introduction

The term port comes from the Latin portus, which means gate or gateway (Rodrigue, Notteboom 2017). A seaport is an area of land and water with related equipment to permit the reception of ships, their loading, and unloading and the receipt storage and delivery of their goods (Coyle et al. 2011: 115). Port terminals play an integral role in the logistics chain by providing cargo-handling services to a wide spectrum of customers, including shipping lines, freight forwarders and various types of organizations. This paper focuses on the port productivity over 2014, 2015 and 2016 for the following ports located in the selected developing nations. The ports include; the Port of Mombasa in Kenya, the Ports of Richards Bay, Durban, East London, Coega, Port Elizabeth and Cape Town in South Africa, the Port of Nacala in Mozambique, the Port of Dar es Salaam in Tanzania, the Doraleh Container Terminal and Djibouti Port in Djibouti and the Port of Port Louis in Mauritius. These ports represent countries in the Southern and Eastern Part of Africa.

Measuring efficiency and productivity is an integral part of any productivity improvement goal (Cabanda, Emrouznejad 2014). The purpose of this paper is to evaluate the performance of selected Africa Ports based on its productivity and efficiency over a period of time. The focus is on the performance of container ports which converts inputs into outputs. Service organizations use and apply benchmarking techniques for measurement of service efficiency (Cabanda, Emrouznejad 2014). The Data Envelopment Analysis model is based on a linear programming technique that evaluates the efficiency of entities relative to best practice observations (Charnes et al. 1978). Fare, Grosskopf, Lindgren and Roos (1994) used DEA to compute a Malmquist Production Index (MPI) which measures a unit's overall productivity change. The DEA based Malmquist Production index captures productivity change in terms of quantities without reference to input prices or output values.

Since 1978 little research has made use of Malmquist Production Index to evaluate ports in Southern and Eastern Africa. A decomposition of calculated Malmquist indices make it possible to identify what factors whether technical efficiency or technological progress determines the changes in seaports productivity in 2014-2016. This paper, therefore, makes use of DEA based MPI to evaluate the performance of 12 selected Southern and Eastern Africa ports over a three year period. Port evaluation is critical for these regions because of the famine, drought and other natural disasters that have affected the horn of Africa. The ports located in East Africa have a pivotal role to play in the distribution of humanitarian relief shipment. Therefore, the efficiency of these ports are critical. This paper is categorized into six segments. Section 2 will focus on concepts and objectives; section 3 focuses on the variables, data and method, Section 4 deals with the results and discussion, whilst Section 5 focuses on the conclusion and recommendation.

2. Concepts and objectives

For the purpose of evaluation, the decision-making units are ports. This section focuses on the definition of terms and concepts.

Container ports serve as an important node in facilitating the efficient flow of containerized cargoes (Notteboom, Yap 2012). The container port is further differentiated by its functions, which consists of serving primarily as a gateway port that acts as an interface between hinterland and deep-sea routings of containerized cargoes, or of serving primarily as a transhipment port that acts as an interface for interchange between deep-sea routings of containerized cargoes (Notteboom, Yap 2012).

A container terminal can be defined as any location where freight and passengers either originates, terminates or are handled in the transportation process (Rodrigue, Slack 2017). Terminals require specific facilities and equipment to accommodate the traffic that they handle (Rodrigue, Slack 2017). Terminal operators want to maximize operational productivity and land space as containers are handled at the berth and in marshaling yards. Container handling productivity is directly related to the transfer functions of a container terminal, including the number and movement rate of quayside container cranes, the use of yard equipment and the productivity of workers employed in waterside, landside and gate operations (Le-Griffin, Murphy 2006).

2.1. Brief background of the ports

This section gives a brief overview of the ports that are being evaluated in both the East and Southern Africa. Ports in the East Africa sub-region, includes the Port of Mombasa, Port of Djibouti, Doraleh Container Terminal, and the Port of Dar es Salaam. In the Southern Africa sub-region, the ports covered includes the Ports of Richards Bay, Durban, East London, Ngqura, Cape Town, and Nacala. The Port of Port Louis is an Indian Ocean island nation.

Port of Mombasa

The Port of Mombasa is a critical gateway for Central Africa's landlocked countries. Developments in the port are, therefore, of great significance (Foster, Briceno-Garmendia 2010). The Port of Mombasa is the busiest port in East Africa. It serves countries such as Uganda, Rwanda, Burundi, South Sudan and the eastern gateway of the Democratic Republic of Congo (DRC) (African Development Bank 2010: 60). The port handles containers, general cargo, dry bulk and liquid bulk.

Port of Djibouti and Doraleh Container Terminal

These two ports are located in Djibouti. Djibouti is positioned at the Horn of Africa. Its strategic location makes it efficiency key to Ethiopia. According to African Development Bank (2010: 60), the Djibouti terminal offers the most modern facilities but needs more investment to meet the high transit demand from Ethiopia. For decades, Ethiopia as a developing nation has suffered from famine.

Port of Dar es Salaam

The Port of Dar es Salaam is located in Tanzania. This port's efficiency is critical to other countries such as Zambia, Malawi, DRC, Burundi and Rwanda that makes use of its services.

Ports in South Africa

Transnet Port Terminals oversee the Ports of Richards Bay, Durban, East London, Ngqura, and Cape Town in South Africa. These ports handle most of Southern African imports and exports. Furthermore, these ports play an important role for the landlocked economies of the sub-region including Botswana, Lesotho, Swaziland, Malawi, Zimbabwe and Zambia (African Development Bank 2010: 65).

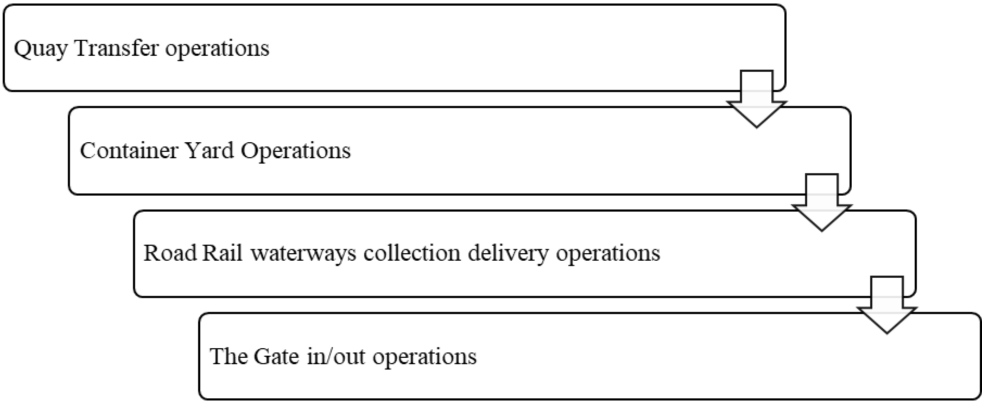
Port of Nacala

The Port of Nacala is located in Mozambique. Though Mozambique has other ports such as the Port of Maputo and Beira, the focus is on the Port of Nacala. The Port of Nacala has a rail connection to Malawi (African Development Bank 2010: 62). Malawi is land logged.

2.2. Container port production process

The ports are categorized as decision making units and homogenous because of the similarity involved in the production processes. In maritime transportation, port throughput is the total number of tons loaded and unloaded within a certain period. In statistical records or handbooks, this data consists of both imports and exports. Throughput is, therefore, the sum of import and export cargoes (Tetteh et al. 2016). At the container Ports, four major operations take place. See Figure 1.

Figure 1. Container production process



Source: authors’ own elaboration.

- Quay transfer operations

The container vessel arrives at the Port. At the port, the container is loaded/unloaded from the ship with the use of a ship-to-shore crane and placed in the port's apron, the staging location. These are operations that refer to the transfer of containers from the quayside to the stacking areas or vice versa.

- Container yard operations

These are operations that involve the positioning of the container into a container stacking yard before being loaded onto the vessel as export, or before being moved out of port as imports or being loaded onto another vessel as transshipment.

- Road-rail-waterways collection delivery operations

These are all the necessary actions that allow the container to be loaded or unloaded onto a truck (road transport) or train or water barge. Container ports that have on-site rail services also have rail entry and departure gates for trains transporting containers to and from the port.

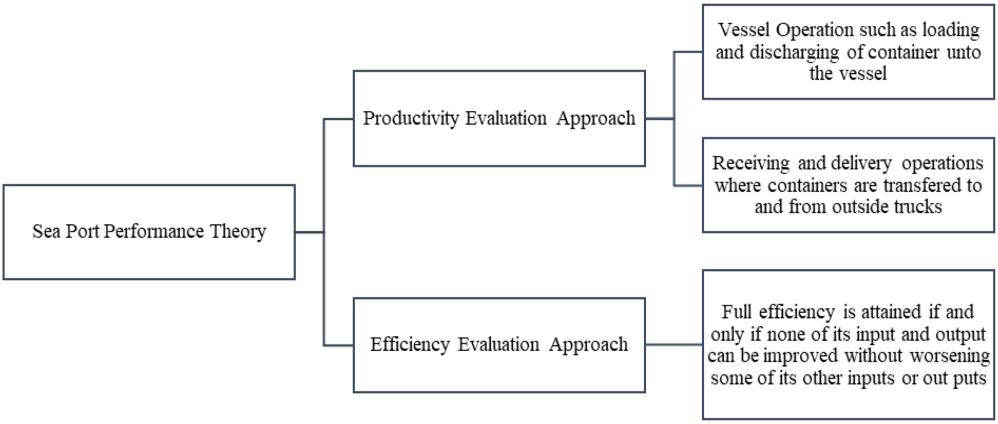
- The gate in/out operations

This mostly relates to road transport. It refers to all the documentation necessary for the container to be loaded onto a vessel for export, or loaded onto a truck as an import. The inland interchange gate allows for the entry and departure of containers by land (or inland waterways) to and from the ports. Imported or exported containers are subject to inspection for proper documentation and security requirements. These gates consist of entry and departure gates. For instance, a truck may arrive at the entry gate with a chassis loaded with a container. At the truck entry gate, relevant information regarding truck movements is recorded. For example, information on the containers being hauled, the ships on which the containers will be loaded and the trucks and their chassis hauling the containers.

2.3. Seaport performance

There are two main research lines on the performance of seaports. These are the productivity evaluation and the efficiency evaluation (Baran, Gorecka 2015). Please see Figure 2.

Figure 2. Sea port performance theory



Source: authors' elaboration based on Baran, Gorecka (2015)

2.3.1. Technical efficiency

The first component of economic efficiency is technical, or productive, efficiency, which has been defined in several different manners in the literature. To Koopmans (1951: 33) a producer is technically efficient if an increase in any output or a decrease in any input requires a decrease in at least one other output or an increase in at least one input. Thus, for each technology for which isoquant and efficient subset diverge, there is a potential conflict between both technical efficiency concepts (Infante et al. 2013). In DEA, this is the output-oriented analysis.

2.3.2. Allocative efficiency

Allocative efficiency is achieved when resources are not wasted (Infante et al. 2013). Allocative efficiency in input selection involves selecting that mix of inputs (e.g labor and capital) that produces a given quantity at minimum cost. (Coelli et al.

2005).There are three conditions to be met for efficient allocation. This is economic efficiency, which involves technological efficiency as well as the use of production factors in such proportions in which costs are minimized (Infante et al. 2013). Consumer efficiency when consumers fail to improve after reassigning their budgets. Thirdly where the cost of producing additional product units equal to the benefits. The MC=MSB, where MC is marginal cost and MSB is a marginal social benefit. (Infante et al. 2013). Allocative and Technical efficiency combine to provide an overall economic efficiency measure (Coelli et al. 2005)

2.4. Review of literature on Data Envelopment Analysis based Malmquist Production Index

The researcher consulted several peer-reviewed journal to have a better understanding of how DEA based MPI has been used to evaluate container ports over the years. Several of the journals focused on Latin American, European and Asian Ports. Only one peer-reviewed article focused on an African nation. Baran & Gorecka (2015) used the Malmquist DEA to evaluate the seaport efficiency and productivity of the 18th world leading container ports. The paper used the Malmquist productivity to determine and analyze the productivity change and its decomposition of four container terminals during 1996-2012. Diaz-Hernandez (2008), used Malmquist DEA to measure productivity changes in cargo handling operations in Spanish ports for a period of 1994-1998. Bo-xin and Guo (2009) also investigated the long-term operating efficiencies of 10 leading container ports in China from 2001-2006. Nwanosike, Tipi and Warnock-Smith (2016) used the Malmquist productivity index decomposition approach to benchmark pre and post-reform total factor productivity growth of six major Nigerian seaports. These are Apapa, Calabar, Onne, Port Harcourt, Tincan Island and Warri Ports for six years before the reform 2000-2005 and six years after the reform 2006-2011. Núñez-Sánchez and Coto-Millán (2012) analyzed the evolution of total factor productivity and its decomposition between 1986 and 2005 in the Spanish ports. Estache, Tovar and Trujillio (2004) use the MPI to measure the changes in, and sources of efficiency since the Mexican reform. 11 main Mexican ports were evaluated from 1996-1999. Cheon, Dowall and Song (2010) evaluated how port institutional reforms influenced efficiency gains between 1991 and 2004. 98 major world ports were used and the MPI model was used.

3. Variables, data and method

This section examines the variables, data and methodology used for this research.

3.1. Variables

The variables used for this research includes the countries earmarked in the African map shown in Figure 3. The map captures six African nations where the 12 container ports are located in: the Port of Mombasa in Kenya, the Port of Richards Bay, the Port of Durban, Port of East London, the Port of Coega, Port of Port Elizabeth, the Port of Cape Town in South Africa, the Port of Nacala in Mozambique, the Port of Dar es Salaam in Tanzania, the Port of Djibouti and the Doraleh Container Terminal in Djibouti and the Port of Port Louis in Mauritius.

For the purpose of analysis, four inputs were used and one output. See Table 1 that shows the inputs and outputs used for this analysis. The inputs and outputs are of immense significance to the container ports. The number of container berths determines the volume of containers that the container port can handle. The cargo handling equipment such as cranes determines and enhances container offloading and on loading in the vessels. The number of cranes at a container terminal has a direct effect on how fast or slow a particular ship is worked on at the terminal because when there are more cranes at the terminal, it increases the number of containers handled per-ship-hour. When there are more ship cranes at a port, the terminal is able to handle more ships at the same time and this increases the scalability of the port (Tetteh et al. 2016). Tugs are critical to guide the movement of the containerships when the ply the unchartered African seashores. The length of the quay determines the ability of the vessel to turn around. The output of a container terminal is seen in the number of TEUs it is able to clear, transshipped or handled. The number of TEUs that a terminal handles determine its productivity (Turner et al. 2003).

Figure 3. Selected east and southern African nations

Source: authors' fieldwork, 2017.

Table 1. Input and output table

Inputs	**Outputs**
Number of Container Berths	Container Throughput
Number of Cranes	
Number of Tugs	
Length of Quay	

Source: authors' fieldwork, 2017.

3.2. Data

Data were obtained and then collated from various ports website. For data not on the website, the researcher approached the ports representatives and regional association representatives via email and they responded with the correct data for the period of analysis. This section focuses on the inputs, output used for the MPI analysis. See Table 2.

Table 2. Port features 2014-2016

Year	DMU	Number of Berths	Number of Tugs	Number of Cranes	Quay Length	Container Throughput
2014	Port of Mombasa	6	8	4	1204	1012002
2014	Port of Richardsbay	3	0	36	350	24189
2014	Port of Durban	7	23	58	2550	2664330
2014	Port of East London	3	0	4	300	41957
2014	Port of Coega	2	6	23	700	705377
2014	Port of Port Elizabeth	2	0	15	600	259917
2014	Port of Cape Town	4	8	227	1151	892557
2014	Port of Nacala	2	0	2	372	97081
2014	Port of Dar es Salaam	4	5	4	720	414059
2014	Doraleh Container Terminal	2	32	5	1050	793317
2014	Djibouti Port	2	4	2	400	70710
2014	Port Louis	2	5	1	560	403001
2015	Port of Mombasa	6	8	4	1204	1076118
2015	Port of Richardsbay	3	0	18	350	19011
2015	Port of Durban	7	23	60	2550	2770335
2015	Port of East London	3	0	1	300	66293
2015	Port of Coega	4	10	14	700	636663
2015	Port of Port Elizabeth	2	0	10	600	216629
2015	Port of Cape Town	4	8	150	1151	888976
2015	Port of Nacala	2	0	2	372	79417
2015	Port of Dar es Salaam	4	5	4	720	644619
2015	Doraleh Container Terminal	2	32	5	1050	836800
2015	Djibouti Port	2	4	2	400	73365
2015	Port Louis	2	5	1	560	361109

Table 2. Cont.

Year	DMU	Number of Berths	Number of Tugs	Number of Cranes	Quay Length	Container Throughput
2016	Port of Richardsbay	3	0	18	350	12302
2016	Port of Durban	7	23	60	2550	2620026
2016	Port of East London	3	0	1	300	71901
2016	Port of Coega	4	10	14	700	572021
2016	Port of Port Elizabeth	2	0	10	600	152455
2016	Port of Cape Town	4	8	150	1151	926611
2016	Port of Nacala	2	0	2	372	71142
2016	Port of Dar es Salaam	4	5	4	720	603290
2016	Doraleh Container Terminal	2	32	5	1050	914017
2016	Djibouti Port	2	4	2	400	73172
2016	Port Louis	2	5	1	560	388514

Source: fieldwork, 2017.

3.3. Methods

The Malmquist total factor productivity index was first introduced by Malmquist (1953). Malmquist production index is considered as a standard approach to measuring the efficiency in the light of time changes (Rahsidi et al. 2014; Huguenin 2012). Malmquist model captures the variations in the port performances in the selected ports over a period of time.

$$M(Y_{t+1,}X_{t+1,}Y_{t,}X_{t,}) = \underbrace{\frac{D^t(Y_{t+1,}X_{t+1,})}{D^t(Y_{t,}X_{t,})}}\left[\underbrace{\frac{D^t(Y_{t+1,}X_{t+1,})}{D^{t+1}(Y_{t+1,}X_{t+1,})} X \frac{D^t(Y_{t,}X_{t,})}{D^{t+1}(Y_{t,}X_{t,})}}\right]^{1/2}$$

Efficiency change Technological Change

Where:

X_t and X_{t+1} input vectors of dimension at time t and t+1

Y_t and Y_{t+1} corresponding k-output vectors

D_t and D_{t+1} denote an input

$D(x,y) = \max(\rho : (x/\rho x \in L(y)$ (2)

Where *L(y)* represents the number of all input vectors with which a certain output vector y can be produced, that is *L(y)* = (*x:y* can be produced with *x*).

P in equation (2) can be understood as a reciprocal value of the factor by with the total inputs could be maximally reduced without reducing output.

M= measures the productivity change between periods *t* and *t*+1. Productivity declines if $M<1$, remains unchanged if $M=1$ and improves if $M>1$.

Computation experiments have been carried out with the application of DEA Malmquist method implemented in the specialized software called PIM-DEA.

4. Results and discussion

Malmquist indices measure the productivity change over time at DMU level. The framework describes the interlinking between the inputs and the outputs that were used for evaluation. As the main activity of container ports are handling containers, one output and four inputs were used.

- ✓ Input X_{1-} Number of Berths
- ✓ Input X_2 Number of Tugs
- ✓ Input X_3 Number of Cranes
- ✓ Input X_4 Quay Length
- ✓ Output Y_1 Container Throughput

4.1. Technical change

In the computation of DEA MPI, two major issues are emphasized, firstly it is the technical efficiency change which can also be known as the catch-up effect. The boundary shift technical change, which is also known as the technology change. The efficiency catch up captures the change in technical efficiency between 2014 and 2015; 2015 and 2016. Table 3 and Table 4 shows the technical efficiency and technology change between 2014-2015; 2015-2016.

Table 3 indicates that the Ports of Mombasa, Durban, East London, Port Elizabeth, Nacala, Doraleh Container terminal and Port Louis show Technical Efficiency-TE=1. This implies that there has been not much change in the technical efficiency level of these ports over the three year period. Other ports such as

Richards Bay, Coega, Cape Town, Dar es Salaam and Djibouti all show TE<1. The implication for these ports is that there is need for improvement in their technical efficiency levels. Overall, all the 12 ports need to improve on its technical efficiency levels.

Table 3. Technical efficiency

	Technical Efficiency	
DMU	2014-2015	2015-2016
Port of Mombasa	1	1
Port of Richardsbay	0.1504	0.1383
Port of Durban	1	1
Port of East London	1	1
Port of Coega	0.8639	0.812
Port of Port Elizabeth	1	1
Port of Cape Town	0.8395	0.9215
Port of Nacala	1	1
Port of Dar es Salaam	0.9784	0.9109
Doraleh Container Terminal	1	1
Djibouti Port	0.2016	0.1996
Port Louis	1	1

Source: authors' calculations, 2017.

4.2. Technological change

Table 4 shows that the Port of Mombasa maintained a Technological change - TC=1 for the years examined. The Port of Mombasa needs to improve on its technology change. The Port of Richards bay has a TC<1 which indicates a need to improve its technological level. There is an increase of 58% for the port of East London, however, there is still need for much improvement in terms of technology. The port of Coega shows a slight decline in technology. The Ports of Port Elizabeth, Nacala, Doraleh Container Terminal and Port Louis had a TC=1. This means that there is no progressive shift in technology. Other ports like the Ports of Cape Town, Dar es Salaam and Djibouti had TC level of less than 1, which implies a decline in technology.

Table 4. Technological change

	Technology Change	
DMU	2014-2015	2015-2016
Port of Mombasa	1	1
Port of Richardsbay	0.1595	0.1504
Port of Durban	1	1
Port of East London	0.4228	1
Port of Coega	0.9933	0.8639
Port of Port Elizabeth	1	1
Port of Cape Town	0.8573	0.8395
Port of Nacala	1	1
Port of Dar es Salaam	0.6656	0.9784
Doraleh Container Terminal	1	1
Djibouti Port	0.206	0.2016
Port Louis	1	1

Source: authors' own elaboration.

4.3. MPI change

The MPI is the combination of the Technical efficiency change and the Technological change. For the year 2014 and 2015, the Ports of Mombasa, Durban, Port Elizabeth, Nacala, Doraleh Container Terminal and Port Louis all maintained a MPI=1. An indication of no improvement. Other ports such as Ports of Richards Bay, East London, Coega, Cape Town, Dar es Salaam and Djibouti had MPI<1 which shows that there is room for improvement. Please see Table 5.

For the year 2015-2016, please see Table 6, the Ports of Mombasa, Durban, East London, Port Elizabeth, Nacala, Doraleh Container Terminal and Port Louis maintained MPI=1. This means no improvement. Other ports such as Port of Richards Bay, Coega, Cape Town and Dar es Salaam and Djibouti had MPI<1. This implies the need for improvement.

Table 5. MPI 2014-2015

	Technical Efficiency	Technology Change	MPI
Port of Mombasa	1	1	1
Port of Richardsbay	0.1504	0.1595	0.023989
Port of Durban	1	1	1
Port of East London	1	0.4228	0.4228
Port of Coega	0.8639	0.9933	0.858112
Port of Port Elizabeth	1	1	1
Port of Cape Town	0.8395	0.8573	0.719703
Port of Nacala	1	1	1
Port of Dar es Salaam	0.9784	0.6656	0.651223
Doraleh Container Terminal	1	1	1
Djibouti Port	0.2016	0.206	0.04153
Port Louis	1	1	1

Source: authors' calculation, 2018.

Table 6. MPI 2015-2016

	Technical Efficiency	Technology change	MPI
Port of Mombasa	1	1	1
Port of Richardsbay	0.1383	0.1504	0.0208
Port of Durban	1	1	1
Port of East London	1	1	1
Port of Coega	0.812	0.8639	0.701487
Port of Port Elizabeth	1	1	1
Port of Cape Town	0.9215	0.8395	0.773599
Port of Nacala	1	1	1
Port of Dar es Salaam	0.9109	0.9784	0.891225
Doraleh Container Terminal	1	1	1
Djibouti Port	0.1996	0.2016	0.040239
Port Louis	1	1	1

Source: authors' calculation, 2018.

5. Conclusion and recommendation

The purpose of this paper is to evaluate the efficiency of the 12 African container ports. The panel data is over the period of 2014-2016. The productivity change was decomposed in to efficiency change and technological change. This was based on the 4 inputs and 1 output used. For the Ports that have M<1 there is productivity decline, while ports that have M=1 indicates that there is productivity stagnancy. The changes in port productivity was as a result of the changes in technology and technical efficiency. Overall, six ports between 2014 and 2015 had MPI=1, whilst the number of ports increased to seven between 2015-2016. For ports that shows decline in Malmquist Production index, port authorities should focus on improving the technical efficiency and technological change.

References

African Development Bank (2010), African Development Report 2010. Port, logistics and trade in Africa, Oxford University Press, New York.

Baran J., Gorecka A. (2015), Seaport efficiency and productivity based on Data Envelopment Analysis and Malmquist Productivity Index, „Logistics & Sustainable Transport", vol. 6 no. 1, pp. 25-33.

Bo-xin F., Xiang-qun S., Zi-jian G. (2009), DEA-based Malmquist Productivity Index measure of operating efficiencies. New insights with an application to container ports, „Journal of Shanghai Jiaotong University (Science)", vol. 14 no. 4, pp. 490-496.

Charnes A., Cooper W.W., Rhodes E. (1978), Measuring the efficiency of decision making units, „European Journal of Operational Research", vol. 2 no. 6, pp. 429-444.

Cheon S., Dowall D.E., Song D. (2010), Evaluating impacts of institutional reforms on port efficiency changes. Ownership, corporate structure, and total factor productivity changes of world container ports, „Transportation Research. Part E", vol. 46 no. 4, pp. 546-561.

Coelli T.J., Prasada Rao D.S., O'Donnell C.J., Battese G.E. (2005), An introduction to efficiency and productivity analysis, 2nd ed., Springer, Boston.

Coyle J.J., Novack R.A., Gibson B.J., Bardi E.J. (2011), Transportation. A supply chain perspective, 7th ed., South-Western Cengage Learning, Mason OH.

Diaz-Hernandez J.J., Martinez-Budria E., Jara-Diaz S.R. (2008), Productivity in cargo handling in Spanish ports during a period of regulatory reforms, „Network and Spatial Economics", vol. 8 no. 2, pp. 287-295.

Cabanda E., Emrouznejad A. (ed.) (2014), Managing service productivity. Using frontier efficiency methodologies and multicriteria decision making for improving service performance, Springer, New York.

Estache A., Tovar de la Fe B., Trujillo L. (2004), Sources of efficiency gains in port reforms. A DEA decomposition of a Malmquist TFP index for Mexico, „Utilities Policy", vol. 12 no. 4, pp. 221-230.

Fare R., Grosskopf S., Lindgren B., Roos P. (1994), Productivity developments in Swedish hospitals. A Malmquist output index approach, in: Data Envelopment Analysis. Theory, methodology and applications, Charnes A., Cooper W.W. (ed.), Springer, Dordrecht.

Huguenin J. (2012),. Data Envelopment Analysis (DEA). A pedagogical guide for decision makers in the public sector, Swiss Graduate School of Public Administration, Lausanne.

Infante Jiménez Z., Gutierrez A., Ortiz-Paniagua C.-F. (2013), Port efficiency in APEC, in: Proceedings of APEC Study Centre Consortium Conference 2013, https://www.researchgate.net/publication/262637197_Port_Efficiency_in_APEC [08.09.2018].

Koopman T.C. (1951), An analysis of production as an efficient combinations of activities, John Wiley and Sons Inc, New York.

Le-Griffin H.D., Murphy M. (2006), Container terminal productivity. Experiences at the port of Long Beach and Los Angeles, in: Container terminal productivity. Conference proceedings, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.410.615&rep=rep1&type=pdf [08.09.2018].

Malmquist S. (1953), Index numbers and indifference surfaces, „Trabjos de Estatistica", vol. 4 no. 2, pp. 209-242.

Nooteboom T., Yap Y.Y. (2012), Port competition and competitiveness, in: The Blackwell Companion to Maritime Economics, Talley W.K. (ed.), Blackwell, Hoboken.

Nunez-Sanchez R., Coto-Millan P. (2012), The impact of public reforms on the productivity of Spanish ports. A parametric distance function approach, „Transport Policy", vol. 24C, pp. 99-108.

Nwanosike F.O., Tipi N.S., Warnock-Smith D. (2016), Productivity change in Nigerian seaports after reform. A Malmquist productivity index decomposition approach, „Maritime Policy & Management", vol. 43 no. 7, pp. 798-811.

Rashidi S.F., Barati R., Kamali Z. (2014), Measurement efficiency by Malmquist method in Data Envelopment Analysis, „Trends in Applied Sciences Research", vol. 9 no. 1, pp. 54-60.

Rodrigue, J.P., Notteboom T. (2017), Port terminals, in: The geography of transport systems, Rodrigue J.P., Slack B. (ed.), 4th ed., Routledge, New York, https://transportgeography.org/?page_id=3235 [08.09.2018].

Rodrigue J.P., Slack B. (ed.) (2017), The geography of transport systems, 4th ed., Routledge, New York, https://transportgeography.org/ [08.09.2018].

Roll Y., Hayuth Y. (1993), Port performance comparison applying Data Envelopment Analysis (DEA), „Maritime Policy and Management", vol. 20 no. 2, pp. 153-161.

Tetteh E.A., Yang H.L., Mama F.G. (2016), Container ports throughput analysis. A comparative evaluation of China and five West African countries seaports efficiencies, „International Journal of Engineering Research in Africa", vol. 22, pp. 162-173.

Turner H., Windle R., Dresner M. (2003), Northern American container port productivity: 1984-1997, „Transportation Research. Part E", vol. 40 no. 4, pp. 339-356.

Foster V., Briceno-Garmendia C. (ed.). (2010), Africa's infrastructure. A time for transformation, The World Bank, Washington, https://siteresources.worldbank.org/INTAFRICA/Resources/aicd_overview_english_no-embargo.pdf [08.09.20818].

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 101-124, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Evaluation of cost efficiency in tomato greenhouses: the case of seven agrobussines in México, 2016

Marisol Arvizu ARMENTA

Centro de Investigación en Alimentación y Desarrollo AC, México

Abstract:

Aim: The main objective of this research was to evaluate the efficiency of economic costs in seven units of greenhouse tomato production, during the production cycles of 2016, through the application of the stochastic frontier, depending on the type of packaging they handle and the cost structure that governs them.

Design / Research methods: The stochastic frontier model includes the analysis of the non-systematic random component, which assumes an extremely critical role in the analysis during the interpretations. With the calculation of the stochastic cost frontier we construct the cost inefficiency index represented by C_it, delimited below 0. The index shows the percentage in which the cost is exceeded and, therefore, the degree of inefficiency.

Conclusions / findings: The elaboration of the stochastic frontier finds its justification in the argument that the less efficient competitor is the one that receives the greater effects of the competition. In this sense, the location of the companies analyzed with respect to their own line of efficiency is essential for the design of the strategies of each company. The production units analyzed showed that, on some occasions, externalities are the cause of inefficiency, but contrary to what is established in theory, there are some units that show that the inefficiency with which they count is diminished by The influence of uncontrolled variables.

Originality / value of the article: The contribution of this research lies in the use of efficiency models in the primary sector, specifically in tomato´s greenhouses..

Key words: Stochastic Frontier Model, Efficiency in Costs, tomato´s Greenhouses

JEL: D14, Q13

Correspondence address: Marisol Arvizu ARMENTA, Centro de Investigación en Alimentación y Desarrollo AC, Carretera a La Victoria KM 0.6 CP. 83304, Hermosillo, Sonora, Mexico. E-mail: marisol.a.armenta@gmail.com
Received: 21.12.2017, Revised: 11.02.2018, Revised: 30.05.2018, Revised: 30.06.2018, Accepted: 30.06.2018
doi: http://dx.doi.org/10.29015/cerem.576

1. Introduction

The tomato (Lycopersicon esculentum mill), is one of the most commercialized vegetables in the world, with more than 177 million tons during 2016. In tomato production the top ten countries are: China (31.8%), India (10.3%), United States of America (7.3%), Turkey (7.1%), Egypt (4.4%), Italy (3.6%), Islamic Republic of Iran (3.5%), Spain (2.6%), Brazil (2.3%) and Mexico (2.2%) (FAO 2018).

The production of this vegetable configures a value chain that involves a series of links among which are: consumers, marketers, suppliers, governments and producers. In 2016, Mexico is positioned as the leading tomato exporter worldwide, the exported value of US$2.1 billion it was equivalent 53.3% of the national production of this vegetable, and 99.3% of sales of Mexican tomatoes went to the United States (CIA 2017). Tomato production is highly concentrated, the 54.1% of the national total in 2015 was produces in five entities; Sinaloa (27.4%), Michoacán (7.2%), San Luis Potosí (7.2%), Baja California (7.1%) and Jalisco (5.2%). (FIRA, 2016).

In tomato production, as in agricultural systems in general, the incorporation of technology has influenced in the increase of productivity and the efficiency of the value chain. Creating an intricate network of marketers-producers who have as their goal compliance with quality standards and just-in-time delivery processes such as those in automotive and aeronautical industries.

The technologies to tomato production can be classified into two broad areas: open field and protected agriculture. The first ones are involved traditional activities where production takes place outdoors at the mercy of insect’s attack and climatic effects. The second form use protector infrastructures (greenhouses and shade screens) that cover the crop from inclement weather, pests and diseases.

Therefore, the technology in greenhouses allows stepped production in harvest times to complement traditional production, since these closed and transparent structures allow the construction of the ideal artificial microclimate to grow plants out of season in good conditions, allowing continuity in the production and good prices (Henao 2001).

In México, the area sown in a conventional manner (open field) was reduced to an average annual rate of 6.7 percent between 2005 and 2015, going from 73,960 to 36,848 hectares. The decrease of the cultivated area in this cultivation modality has been greater in some entities such as Sinaloa, Baja California and Jalisco. On the other hand, the area established with protected agriculture (greenhouse and shade screens) increased from 395 to 13,747 hectares in the mentioned period, that is, it grew at an annual average rate of 42.6 percent. Greenhouse cultivations is concentrated in Sinaloa, Baja California and Jalisco, although it has also acquired greater importance in other entities such as Colima, State of Mexico, Hidalgo, Michoacán, Querétaro, San Luis Potosí, Sonora and Zacatecas. The increase in the surface area with protected agriculture infrastructure is attributed mainly to the success in the harvest of tomato quality export that is intended to the United States market (FIRA 2016). In this article, only the producers who wanted to share the necessary information for the study were consulted, the main tomato producing states were consulted in the greenhouse and the results of those who agreed to participate were presented.

Authors such as (Calvin, Cook 2005; Cook, Calvin 2005); analyze from the economic perspective tomato production in greenhouse, and have focused on marketing channels, production lines, cost structure and governance relationships between sellers, producers and buyers and they evaluated the generated employment and the economic multiplier effects generated.

Other studies such as (Engindeniz, Tuzel 2006), make an economic analysis of a greenhouse in Turkey, from its installation and its operation emphasizes the economic feasibility associated with the expansion of these greenhouses. On the other hand (Dodson 2002), studies the diversification of production from organic tomato production technologies (Mysore, Weng-Fei 1999), focuses on the economic dimensions of greenhouses in the United States, analyzing the multiplying effects of the production.

Somewhere research to evaluate the economic efficiency in tomato production show the main approaches to approach through which the issue of tomato economic efficiency has been investigated. From the perspective of profitability benefit-cost engaged in making comparisons using measures such as the ratios of physical productivity (divided product inputs) or average costs (cost divided product) (Sánchez López et al. 2004; Torres Lima et al. 2004; Rubocoa et al. 2016); use of Cobb Douglas functions (Ibitoye et al. 2015); estimation of shepherd-future coefficient and exponential model of combined profit function (Ayoola 2014), among others.

Since this is an economic activity that involves international competition, producers must conceive their investment project considering all the elements that demand efficiency in production and the search for profitability in a competitive environment where prices are the indicators that mark the fluctuations of supply and demand.

In this sense, there are aspects that are not considered in these analyzes of profit, such as: sector weaknesses, high capital costs, technical and management inexperience, as well as the shortage of suppliers of specialized inputs and/or services, infrastructure and technology, etc.

In the international field, following Laurinavičius (2017), there is a profuse literature that addresses the issue, only to cite some authors we list some of the research products Productive efficiency of agricultural sector is extensively analyzed (Gorton, Davidova 2004). A number of studies have been attempted to investigate the issues of efficiency by using widely applied frontier methods. Asmild nad Hougaard (2006) analyzed the influence of environmental improvement potential to efficiency of Danish pig farms. Davidova and Latruffe (2007) related the Czech farm efficiency to financial management. Vasiliev et al. (2008) employed the DEA method to analyze the efficiency of Estonian grain farms after Estonia's transition to the market economy and during the accession period to the European Union (EU). Rasmussen (2010) used SFA in the form of input distance functions to estimate efficiency of Danish crop, dairy and pig farms. Bojnec and Latruffe (2011) analyzed the relationships between size and efficiency of Slovenian farms.

However, there is an unfilled gap in the research on the analysis of the efficiency in production costs of Mexican protected agriculture from the perspective of stochastic frontier analysis of production and costs.

The present research aims to perform a stochastic frontier analysis in costs in tomato´s greenhouses: seven agrobusiness in México, 2016. Using for this purpose the models originally proposed by Aigner et al. (1977) and Meeusen and Van Den Broeck (1977), and adapted by Stevenson (1980), and which includes the non-systematic random component in substitution of those variables that are omitted and affect profitability.

This document is divided into 6 sections: The structure of costs in greenhouses; the stochastic frontier model; Characterization of companies and description of the variables; Packaging costs, variables not controlled, controlled and determinant; Monte Carlo simulation-application model; and conclusions.

2. The structure of costs in greenhouses

In the analysis of the economic dimensions and the profitability of the greenhouses, the cost structure and the evaluation of the efficiency of the same stand out. The economic cost is defined as:

"The economic cost analyzes the company thinking about the future, the allocation of scarce resources waiting to know what the cost will be in the future and how the company could reorganize its resources to reduce it and improve its profitability, therefore, the economic cost is equal to the cost of lost opportunities where there are costs that the company can and can't control" (Pindyck 2009: 208).

At International level cost efficiency has been analyzed with different models, some authors have applied the stochastic frontier model for agriculture, such as the Taiwanese case studied by Hung et al. (2008). This author applied the cost stochastic frontier model in a pure way to estimate the cost frontier and the efficiency of each company, to make the location of these in relation to the cost frontier.

Kvaløy and Tveteras (2008) studies the cost structure and vertical integration having as main contribution in the analysis of the curve of the average costs and the relation that they have as the scale of production.

Bateman et al. (2006) investigates the benefits and costs of agriculture in the framework of a strategy implemented by the European Union to give relation to the management and the cost of water in this primary activity, to analyze the structures of costs before and after the application of such strategy.

At the Latin American level Benach (2005) studies agricultural and industrial production cost models, analyzes the cost models used in rice production in Costa Rica, and designs proposals for new cost-of-production models. Reyes (1995) uses an econometric model of linear programming for different combinations of research and development, interest rates and agricultural prices, obtaining an efficient production structure and costs.

At the national and local level, the identified studies have focused on two aspects, the first one is that proposed by Kido (2007), who makes a comparative analysis of costs, analyzes the efficient cost and the opportunity cost having as two scenarios the planting of maize or the reforestation of the area in question. The second is that of Sánchez et al. (2004), who calculates the average cost of production of cotton to reach a point of equilibrium and characterizing the structure of costs and production of the company.

3. The stochastic frontier model

The stochastic frontier model includes the analysis of perturbations or non-systematic random component that substitutes or represents those variables that are omitted or ignored and that affect the product but which a deterministic or statistical model was not included in the analysis.

Within this logic, the error term replaces all variables that are not included in the analysis model for which there are different meanings. The non-systematic random component assumes an extremely critical role in the analysis of the models of stochastic frontier function that is seen during the interpretations that can be given to the model and hence the importance of using the stochastic frontier function model.

This model, proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977) where stochastic efficiency is assumed to follow a normal distribution of means. More flexible assumptions with respect to efficiency distribution were developed in the literature when including the truncated normal distribution of

Stevenson (1980) which allows a vector to be truncated positively so that the efficiency depends on specific variables.

The general model is:

$$lnC_{it} = lnC_{it}^{DK} + v_{it} + u_{it} \quad (1)$$

$$lnC_{it}^{DK} = f\ (X_{it} + Z_{it};\ \beta, \delta) \quad (2)$$

$$v_{it} \sim N(0, \sigma_v^2), \quad (3)$$

$$u_{it} \sim N\ (u_{it}, \sigma^2), u_{it} \geq 0, \quad (4)$$

$$u_{it} = z_{it}\gamma. \quad (5)$$

Where:

lnC_{it} : is the logarithm for total costs.

lnC_{it}^{DK} : The kernel is the determinant of the production frontier that is defined by the function (f). The kernel determinant is a function of two vectors of variables X_{it} y Z_{it}, and their corresponding coefficients of vectors β y δ which is based on a standardized logarithmic cost function, where X_{it} contains the logarithms of the products as quantities y_{it} and prices as inputs and the terms of interaction between them.

v_{it} is a random variable of mean 0 and with normal distribution. The importance of this distribution is that it allows modeling numerous natural, social and psychological phenomena. While the mechanisms underlying much of this type of phenomena are unknown, because of the sheer number of uncontrollable variables involved in them, the use of the normal model can be justified by assuming that each observation is obtained as the sum of a few independent causes, the normal distribution is important because of its relationship with the estimation by ordinary least squares.

u_{it} is the variable that captures the effect of cost inefficiency which is a measure of the additional cost as a percentage of the minimum cost. It is assumed that the random variable follows a normal distribution.

$Z_{it}\gamma$ this part a positive coefficient indicates that the growth in an exogenous variable cause that the inefficiency in the cost increases (Battese, Coelli 1995). As indicated in equation (2) the vector Z_{it} is included in the minimum cost function, which means that within of the exogenous variable in Z_{it} not only changes the distance between the current cost of the minimum, but can even shift the cost frontier lnC_{it}^{DK}.

4. Characterization of companies and description of the variables

The method followed by this investigation requires the description of companies and the variables that will be used to measure the efficiency of the unit, in addition to explaining each of the steps that are necessary for the application of the stochastic frontier model of cost. The companies that participate in this evaluation are tomato-producing units in greenhouses that have medium and high technology (Table 1).

The stochastic frontier of cost makes a count of the distance that has the current cost of the company and the frontier given by the established conditions and the variables used for the construction of this one, is due to this that for the interpretation of the indicators resulting from the model will be interpreted in a suitable way placing the production unit in the context of its productive indicators.

Table 1. Productive indicators

Indicators 2016	Monterrey	Saltillo	Parral	Cuauhtémoc	Sonora	Guanajuato	Sinaloa
Size of the company (M^2 total)	45,240	50,000	40,000	20,000	50,400	50,000	200,000
Unitary Performance (Kg/ M^2)	40.00	27.00	56.00	32.00	19.50	19.42	14.00
Average Selling Price ($/Kg)	16.20	16.30	9.20	8.29	15.00	14.00	15.00
Technology	High	High	High	Medium	Medium	Medium	Medium

Source: information provided by producers of 7 tomato producing greenhouses.
Notes: Kilograms (Kg); Square meters (M^2); Mexican pesos ($).

For the application of the model it is necessary to employ two types of variables, controlled and not controlled by each production unit. In the first instance, we have the uncontrolled variables, for this investigation we consider 3, the market price, the exchange rate and the price of natural gas in 2016. This type of variable shows the influence that the exterior has inside the structure of cost of the companies.

5. Packaging costs, variables not controlled, controlled and determinant

The controlled variables are represented by the cost structure of the producing units involved in the analysis. The companies analyzed do not necessarily have homogeneous accounting entries, making the controlled variables that are necessary for the application of the model incomparable. With this background, the first necessary step for the correct application of the model was the homologation of the cost structures remaining as shown in Table 2.

Table 2. Structure of cost of production of greenhouse tomato

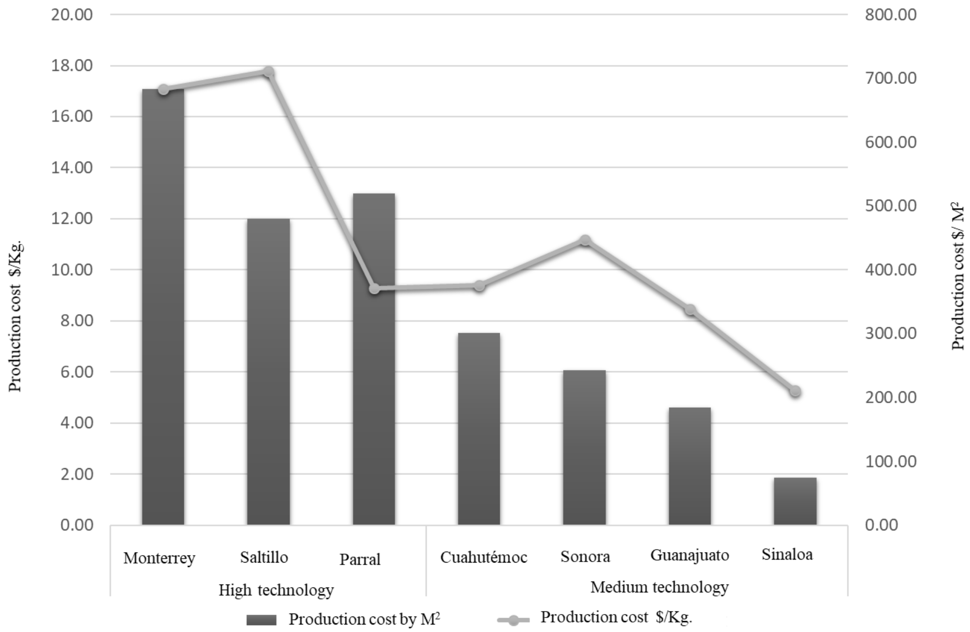
1. Production Indicators
1.1. Size of the Company (M^2 totales), 1.2 Unitary Perfomance (Kg/ M^2), 1.3. Average Selling Price ($/Kg), 1.3.1. Average exportation selling price ($/Kg), 1.3.2. Average national sellin price ($/Kg), 1.4. Technology
2. Income per Hectare
2.1. Main Product (Fresh Tomato), 2.2. Packaging maquila, 2.3. Government Payments
3. Operation costs by hectare
3.1. Variability costs, 3.1.1. Seeds (Vegetative material), 3.1.2. Growing substrate (soil treatment), 3.1.3. Fertilizers, 3.1.4. Packaging, 3.1.5. Electric Energy, 3.1.6. Gas CO_2, Fuel, 3.1.7. Chem/Bio Supplies, 3.1.8. Water, 3.1.9. Workforce, 3.1.10. Freight, 3.1.11. Comercialization, 3.2 Other variable costs, 3.2 Fixed Costs, 3.2.1. Company Admin, 3.2.2. Depreciation of assets, 3.2.3. Other fixed costs

Source: author's elaboration based on information of 7 greenhouses producers of tomato, homologation suggested by FIRA (2016).

Once the structure is homologated, the variables to be applied within the model should be calculated per square meter and per kilogram to calculate the economic indicators of the company (break-even point and operating profit of kilogram per square meter).

The cost of production of Monterrey, which should be mentioned is a company that has a high technology, almost 700 pesos per square meter and the Sinaloa, medium technology, with less than 100 pesos, other companies have its cost in a range of less than 600 and more than 200 pesos depending on the level of technology in which they are. Then the percentage of the main costs are presented, these are fertilizer, CO_2, labor, packaging and freight, with the highest cost being the packaging.

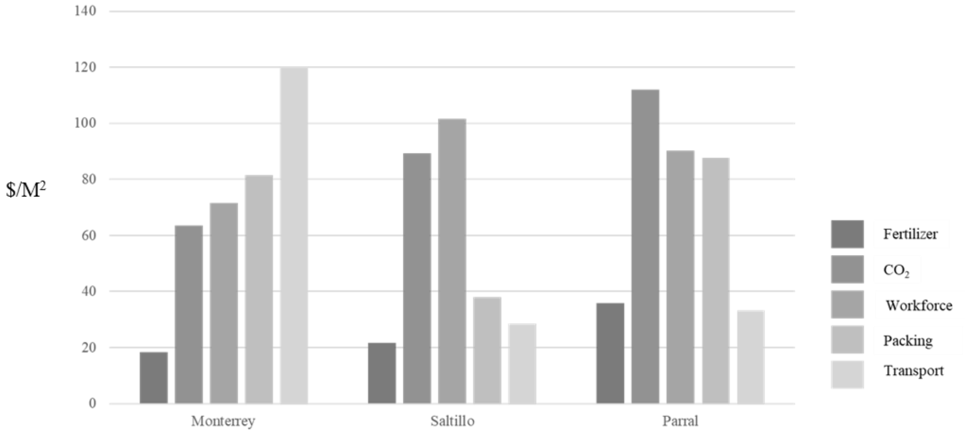
Graphic 1. Comparison of greenhouse tomato production costs in 7 Mexican companies, for 2016



Source: author's elaboration based on information of 7 greenhouses that produce tomato.

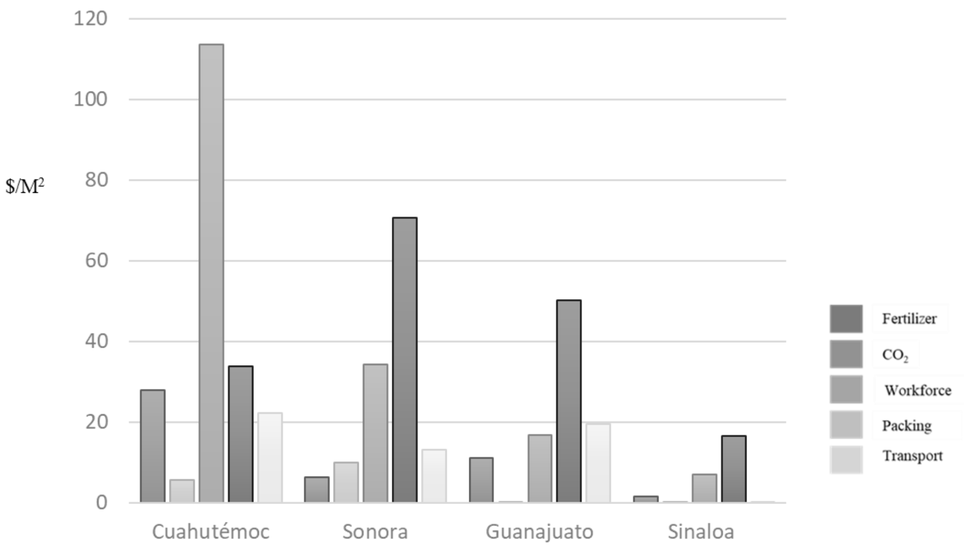
As can be seen in the previous section, the packaging plays a dominant role over other accounting entries reported by all companies, representing in some companies up to 92% of total costs, becoming a determinant variable in the total cost of the company, see Table 3.

Graphic 2. Composition of variable production costs of high tech greenhouse tomato on 2016



Source: author's elaboration based on information of 7 greenhouses that produce tomato.

Graphic 3. Composition of variable costs of production on greenhouses with medium tech that produce tomato



Source: author's elaboration based on information of 7 greenhouses that produce tomato.

Table 1. Average of packing above total cost

Technology level	City/State	Packing	Non Packing
		Average	Average
High	Monterrey	88.08%	11.92%
	Saltillo	92.08%	7.92%
	Parral	83.15%	16.85%
Medium	Cuauhtémoc	88.73%	11.27%
	Sonora	70.80%	29.20%
	Guanajuato	72.81%	27.19%
	Sinaloa	77.59%	22.41%

Source: author's elaboration based on information from 7 greenhouses that produce tomato.

Table 4. Definition of variabilities

Fixed cost	Market price by: Type of packaging	Variability cost	Packaging price by Kg
	Type of change	Total cost	
	Natural Gas price	Clearing Price	

Source: author's elaboration based on information of 7 greenhouses that produce tomato.

In Table 4 the variables defined for the application of the simulation are presented:

- Fixed cost: fixed costs reported by companies.
- Market price by type of packaging: these represent the daily costs per type of tomato packaging in two high and low price scenarios, the most used packages in the market are: 5kg Carton, 5kg Flats, 10lbs, 11lbs, 15lbs and 25 pounds; the simulated market price was the average monthly price per kilogram reported by USDA for the year 2016.
- Exchange rate: Daily peso-dollar exchange rate reported by Bank of Mexico for the year 2016.Natural gas price: quarterly natural gas price reported by Bank of Mexico for the year 2016.
- Total cost and variable cost: costs reported by companies.
- Packaging cost per kg: cost reported by the company.

6. Monte Carlo simulation and application of the model

The variables to be simulated will be weighted on their participation in the total cost to obtain a closer approximation to reality with the Monte Carlo simulation performed. The simulation is done by Excel spreadsheet in which 1,000 tests are carried out with controlled simulations applied to the uncontrolled and controlled variables. The decision criterion for choosing the variables to be applied in the stochastic frontier model is the profitability of the variables.

The stochastic cost-regression regression yields the results as shown in Figure 1, applied in the STATA software 14. In the first instance, we have the regression where the dependent and independent variables interact, the model has the property of separating the statistical error of the stochastic error for which the variables differ.

The interpretation is based on two components of the regression, the first is the sign and the second the coefficient. The sign shows whether the inefficiency is presented positively or negatively. In the case of the coefficient reflected in which percentage is increased or decreased inefficiency.

Figure 1. Stochastic cost-regression regression yields

.frontier ctm2 price cxe, uhet (price) vhet(cxe) cost nolog iterate (100)							
Stoc. Frontier normal/half-normal model					Number of obs	=	916
Log lokelihood		=	4105.218		Wald chi2 (2)	=	7.79E+07
					Prob > chi2	=	0.0000
		Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
ctm2							
	price	-0.0001668	0.0001473	-1.13	0.257	-0.0004555	0.0001219
	cxe	0.9724754	0.000114	8528.85	0.000	0.9722519	0.9726989
	_cons	3.559906	0.0003809	9344.82	0.000	3.55916	3.560653
lnsig2v							
	cxe	-7.911996	0.17225	-45.93	0.000	-8.2496	-7.574392
	_cons	2.764299	0.3099484	8.92	0.000	2.156811	3.371787
lnsig2u							
	price	5.088877	0.6825449	7.46	0.000	3.751113	6.42664
	_cons	-19.22836	0.8975534	-21.42	0.000	-20.98753	-17.46919

Source: Author's own elaboration.

6.1 Cost stochastic frontier

In the application of the model the distribution of stochastic error was obtained in two different scenarios with each one of the uncontrolled variables proposed and by the types of scenario that the database provides by making classification highly and inefficient.

Random variable: market price by type of packaging and scenario

In the case of the market price (Table 5), in the high price scenario the Cuauhtémoc company presents a high inefficiency because its coefficients reach 9.416%, in the case of the 11-pound package which is the one that represents the greatest inefficiency for the company followed by the 15-pound package in which 6.017% inefficiency increases as well as the 5-kg Carton package, then there is the 25-pound packaging that increases inefficiency by 2.788% for the company.

The packaging that increases to a lesser extent the inefficiency of Cuauhtémoc's company is 10 pounds, since its inefficiency would increase by 0.142%, all given the conditions presented in the analyzed year, since this distribution is conditional on an average annual exchange rate of 2016, reported by the Bank of Mexico and an average annual price of natural gas of 6,99.

Table 2. Distribution of stochastic error by the type of packaging of high price

States with inefficiency		price by type of packaging					
		5 Kg Carton	5 Kg Flats	10 Libras	11 Libras	15 Libras	25 Libras
Highly inefficient	Cuauhtémoc	6.008	1.394	0.142	9.416	6.017	2.788
	Sonora	1.304	4.713	0.083	13.654	5.899	n.a.
Little Inefficient	Guanajuato	-0.023	-0.261	0.056	-0.352	-0.041	0.587
	Saltillo	0.023	-0.104	-0.506	-0.036	-0.430	4.820
	Sinaloa	-0.195	0.162	0.000	-0.359	0.142	0.274

Source: author's elaboration based on information provided by producers.

The company located in Sonora, within the scenario of high prices, presents significant levels of inefficiency, especially in the 11-pound package as it increases by 13.654%, without forgetting that in all types of packaging analyzed has an increase of inefficiency ranging from 5.899% to 0.083%. In contrast to observing the behavior of stochastic error that has the company of Sinaloa can be noted that the levels of inefficiency it handles are minimal and fluctuate between 0.274% and -0.359%.

For the low price scenario (Table 6), with the aforementioned conditions of the exchange rate and the price of natural gas, there is a distribution of the error completely different from that obtained in the scenario of the high price, since for four types of packaging was not found convergence, which means that the cost structure is adequate for this scenario, with these types of packaging being 5 kg Flats, 11, 15 and 25 pounds, while for the Carton 5 kg package and 10 pounds the inefficiency increases considerably in 15% and 2% respectively.

As for the situation shown by the Sinaloa company, its cost structure is not affected to an important extent by changes in the random variable; shows small coefficients of increase in inefficiency, as shown by the 25-pound pack which is 0.18%, for other packaging inefficiency decreases from -0.003% to -1.49%.

Table 6. Stochastic error distribution by type of package low price scenario

		price by type of packaging					
States with inefficiency		5 Kg Carton	5 Kg Flats	10 Libras	11 Libras	15 Libras	25 Libras
Highly inefficient	Cuauhtémoc	15.820	n.a	2.572	n.a	n.a	n.a
	Sonora	3.272	0.311	-0.199	15.678	0.088	n.a
Little Inefficient	Guanajuato	0.992	-0.142	-0.070	0.289	-0.164	0.698
	Saltillo	-0.100	-0.111	-0.416	-0.581	2.443	2.068
	Sinaloa	-0.074	-0.004	-0.056	-1.496	-0.178	0.184

Source: author's elaboration based on information provided by producers.

In the low-price scenario, the Sonora company resulted in a stochastic error distribution reaching its maximum in the 11-pound package by increasing this inefficiency by 15%, this percentage being the worst scenario in the application of the stochastic frontier.

It is worth mentioning the Saltillo company because being a company with a high technology would have to comply with the assumption of being efficient in each of its cost components, however, the result was that, even though its inefficiency is small, the model finds a degree of convergence in the interaction of cost and the random variable.

6.2 Random variable: exchange rate

The distribution of the stochastic error when the uncontrolled variable was the exchange rate (Table 7), changed the situation that was presented with the market price, in this case the company of Sonora for example, in the scenario of the high price, step from being highly inefficient to little inefficient, and the Saltillo company faces the other way. This situation that is presented within the companies by the exchange rate is mainly due to the fact that they are companies that export tomatoes and the exchange rate is a variable that ultimately affects their efficiency.

As with the market price variable, the exchange rate reflects that the unit located in Cuauhtémoc continues to be the most inefficient, however it must be clarified that the exchange rate and market price coefficients are not equal, recalling that the coefficient is the one that determines the percentage in which the company is or is not inefficient, because the average of the coefficient is smaller in the exchange rate than in the market price. Even with this clarification, it is important to highlight the Cuauhtémoc case, since it presents a coefficient of increase in the inefficiency, 1.86% and 1.14% in the packages of 5 kg Carton and Flats, respectively, in the case of 10 and 15 pound packages the situation observed is different, since the variable contributes to the reduction of inefficiency, for the rest of the packages there is no convergence.

In the case of the Saltillo company it is observed that, although the coefficient is not as high as in Cuauhtémoc, it manages a certain level of inefficiency that ranges from 1.4% to 0.75% in all packages except the 25 pound packaging that collaborates to reduce this coefficient, this case is very important to highlight it as it is, as already mentioned above, a company with high technology that is not being efficient in the management of the components of the cost that it owns, which is incurring in levels of inefficiency, equal or superior in some cases to those of companies with medium technology.

The low price scenario (Table 8) in terms of the exchange rate variable shows again changes in companies that are highly inefficient and inefficient; this time it is worth mentioning the case of the Sonoran company as it returns to its highly inefficient position shown when the random variable was the market price, even surpassing the Cuauhtémoc company since it has indicators from 1.57% to 0.35%, these indicators still when they are not comparable with the percentages obtained with the distribution of the stochastic error with the market price reflect inefficiency in the company and a negative influence on the cost components analyzed, since it increases their inefficiency.

Table 7. Stochastic error distribution high price scenario

States with inefficiency		Exchange Rate					
		5 Kg Carton	5 Kg Flats	10 Libras	11 Libras	15 Libras	25 Libras
Highly inefficient	Cuauhtémoc	1.862	1.141	-0.614	n.a.	-0.489	n.a.
	Saltillo	0.875	1.493	0.760	1.355	1.464	-0.512
Little Inefficient	Guanajuato	0.194	0.194	0.194	0.194	0.194	0.990
	Sonora	0.009	-0.307	-1.751	0.450	-1.751	n.a.
	Sinaloa	0.742	0.742	0.742	0.742	0.742	0.742

Source: author's elaboration based on information provided by producers.

Table 8. Distribution of stochastic error exchange rate scenario low price

States with inefficiency		Exchange Rate					
		5 Kg Carton	5 Kg Flats	10 Libras	11 Libras	15 Libras	25 Libras
Highly inefficient	Saltillo	-0.338	1.044	-0.649	-0.379	0.352	n.a
	Sonora	1.262	1.044	1.044	1.599	1.502	n.a
Little Inefficient	Guanajuato	0.753	0.753	0.753	0.753	0.753	1.869
	Cuauhtémoc	n.a	n.a	-0.623	n.a	n.a	n.a
	Sinaloa	0.651	0.651	0.651	0.651	0.651	0.651

Source: author's elaboration based on information provided by producers.

Saltillo, on the other hand, continues to show the trend of inefficiency shown in the scenarios discussed above, reflects levels of inefficiency in two of the six types of packaging analyzed.

The Sinaloa firm continues to be efficient in managing its costs and although in the two types of scenarios the inefficiency that shows the distribution of the random variable is positive the coefficient that presents in both types of scenario continues below 1%, it is say, although the impact is minimal there is no significant influence of the random variable on the cost components shown by this company.

6.3. Random variable: natural gas price

Within the high price scenario when the uncontrolled variable is the price of natural gas, there are no coefficients that show high or low inefficiency, with the results shown in Table 9 shows that the price of natural gas is the uncontrolled variable which represents a minor influence on the cost structure of the companies analyzed.

Table 3. Stochastic error distribution natural gas price scenario high price

States with inefficiency		Natural gas price					
		5 Kg Carton	5 Kg Flats	10 Libras	11 Libras	15 Libras	25 Libras
Highly inefficient	Sinaloa	0.099	0.099	0.099	0.099	0.099	0.099
	Saltillo	0.224	0.178	0.041	0.178	0.168	n.a.
Little Inefficient	Guanajuato	-0.115	-0.115	-0.141	-0.115	-0.115	-0.057
	Sonora	n.a.	-0.007	-0.053	-0.016	n.a.	n.a.
	Cuauhtémoc	-0.012	0.023	n.a.	n.a.	-0.088	n.a.

Source: author's elaboration based on information provided by producers.

However, interesting cases continue to occur with the error distribution; in the first instance it is shown that the Sinaloa company now occupies the place in highly inefficient companies, because although the coefficient of inefficiency is very small (0.09%), it is positive and greater than the coefficients shown by the other units. Saltillo continues to show convergence with random variables.

It should be noted on this occasion that the company is located in Cuauhtémoc as it goes to the area of inefficient units, thanks to the fact that in three of the six packages analyzed there is no convergence, while in those that if convergence is found the coefficient is less than 1% both in increase of inefficiency and in a decrease of inefficiency.

Within the low-price scenario, the distribution of the error does not show significant changes in the levels of the coefficients, but does show significant changes in the distribution and classification of the producing units. First, there is the change of the Guanajuato company that, for a single occasion, appears in the high levels of inefficiency. This is due to the fact that the inefficiency that it shows, although minimal in coefficient is positive, in contrast to other companies.

Table 4. Stochastic error distribution natural gas price low price scenario.

States with inefficiency		Natural gas price					
		5 Kg Carton	5 Kg Flats	10 Libras	11 Libras	15 Libras	25 Libras
Highly inefficient	Guanajuato	0.068	0.068	0.068	0.068	0.068	0.080
	Sonora	-0.088	-0.089	-0.097	n.a	-0.008	n.a
Little Inefficient	Cuauhtémoc	n.a	n.a	-0.087	n.a	n.a	n.a
	Sonora	-0.088	-0.089	-0.097	n.a	-0.008	n.a
	Sinaloa	-0.066	-0.066	-0.066	-0.066	-0.066	-0.066

Source: author’s elaboration based on information provided by producers.

The Sonora unit presents levels of negative inefficiency, which means that it decreases its levels, emphasizes the behavior of the company of Cuauhtémoc as it has the best scenario as it does not find convergence in 5 of the 6 types of packaging.

Cost stochastic border, the case of Monterrey and Parral

The units of analysis of this research included two companies that showed a highly efficient behavior, these units have high technology and a structure of costs able to withstand the fluctuations of both the exchange rate, the price of natural gas and the price of market in each of the packages analyzed in the scenarios that this research is located.

At the time of the application of the model when looking for the influence of the uncontrolled variables on the composition of the costs of these companies did not obtain convergence, that is to say, the model showed that these units are not affected by the external conditions to the company and have the technology and context necessary for the optimum production of tomatoes for export.

Cost stochastic frontier, the case of Guanajuato

Among the companies that work with medium level technology, it is worth noting the case of the Guanajuato unit, because it is a company that in each of the scenarios and with the three uncontrolled variables revealed an efficient behavior. The coefficients and signs that were presented within this unit were definitive because even though it showed that the inefficiency increased or decreased according to the conditions established for each analyzed scenario, the coefficient revealed that the influence of the exchange rate, the price of natural gas and the market price by type of packaging does not impact on the composition of the cost. The above is mainly due to the level of technology it manages, since in several scenarios analyzed this company proved to have a composition of stable cost and little affected by the external conditions.

7. Conclusions

With the application of the stochastic frontier model, the influence of external variables on the cost structure of the producing units analyzed, showing different scenarios, showed that, on some occasions, externalities are the cause of the possible inefficiency can present in them, but contrary to what is established in theory, there are some units that show that the inefficiency with which they count is diminished by the influence of uncontrolled variables.

Cuauhtémoc and Sonora proved to be vulnerable units to external conditions and with cost components that do not have sufficient strength to resist the impact that these variables exert on them. In contrast, the Monterrey and Parral units have a cost composition capable of absorbing the effects that the external variables have, this explains the level of technology they have and the performance they have.

Saltillo, even though it has high technology, deserves special mention because the uncontrolled variables have an impact on the cost composition, revealing that this unit is vulnerable to external conditions. Finally, the units in Guanajuato and Sinaloa have the most efficient cost structures; external conditions do not pose a danger when measuring inefficiency. This is supported by the coefficients of the stochastic error distribution, which in a few cases exceeded 1% both in increase and decrease in inefficiency, this is an indicator of the strength of these companies, especially the unit of Guanajuato which is the one that maintains a more stable behavior of the seven units analyzed in this investigation.

Sinaloa is supported by the importance of the production volume and production value generated by within the national scope, these are indicators that collaborate so that its cost structure is one of the strongest within the units analyzed; with all this the units must establish strategies that lead them to a better functioning.

8. Recommendations

Each type of packaging represents a different market price, which was analyzed individually for each producing unit, resulting in an efficient type of packaging for each of which the above-mentioned conclusions were derived and which resulted in the proposal of different strategies presented below.

The first strategy proposal is for each producing unit to adopt the packaging for which its cost structure is adequate, each of the units analyzed in this investigation resulted in a certain type of packaging making the operation efficient according to the structure of the costs, for which it would be convenient to use that type of packaging.

For production units to produce in a type of packaging they must know the characteristics of the market. One of these characteristics is the time it must remain in the market, during the analysis of each of the units it was concluded that the best package for all was the 25-pound one.

This type of packaging has the characteristic that it remains throughout the year in the market, it means the 52 weeks, not only by cycle as produced by the units analyzed, from which the first strategy based on cost leadership is derived.

The strategy is the organization of producers that allows to supply the 52 weeks of the year to the target market, that is, to the United States. The organization that the producers of the analyzed companies can reach and the decisions on the type of packaging to which they are produced can be definitive aspects in the improvement of the efficiency levels of the producing units.

The 10-pound package also represents a good option for all companies, even though it is not as efficient for all companies, their levels of inefficiency are very small in all the scenarios previously presented, which means that they adapt to the structures of costs of the companies that were analyzed and that even collaborates with the decrease of inefficiency in some of the units.

The great advantage of this type of packaging for the Mexican unit is that its presence in the market occurs during the months of February to May, November and December, months in which the producer can meet that demand.

The strategy suggested for this type of packaging is for the production units located in Guanajuato, Sonora and Saltillo, as the results from the stochastic frontier support that tomato production must be carried out in this type of packaging, since inefficiency decreases of the companies.

References

Aigner D.J., Lovell C.A.K., Schmidt P. (1977), Formulation and estimation of stochastic frontier production function models, „Journal of Econometrics", vol. 6 no. 1, pp. 21-37.

Asmild M., Hougaard J.L. (2006), Economic versus environmental improvement potentials of Danish pig farms, „Agricultural Economics", vol. 35 no. 2, pp. 171-181.

Ayoola J.B. (2014), Comparative economic analysis of tomato (Lycopersicum esculenta) under irrigation and rainfed systems in selected local government areas of Kogi and Benue, Nigeria. „Journal of Development and Agricultural Economics", vol. 6 no. 11, pp. 466-471.

Bateman I.J., Brouwer R., Davies H., Day B.H., Deflandre A., Di Falco S. et al. (2006), Analysing the agricultural costs and non-market benefits of implementing the Water Framework Directive, „Journal of Agricultural Economics", vol. 57 no. 2, pp. 221-237.

Batesse G.E., Coelli T.J. (1995), A model for technical inefficiency effects in a stochastic frontier production function for panel data, „Empirical Economics", vol. 20, pp. 325-332.

Benach A.L. (2005), Estudio para la definición de la Estructura de Costos de Producción Agrícola de Arroz en Costa Rica, Instituto de Investigaciones en Ciencias Económicas, Universidad de Costa Rica [unpublished work].

Bojnec S., Latruffe L. (2011), Farm size and efficiency during transition. Insights from Slovenian farms, „Transformations in Business and Economics", vol. 10 no. 3, pp. 104-116.

Calvin L., Cook R. (2005), North American greenhouse tomatoes emerge as a major market force, „Economic Research Service”, April, https://www.ers.usda.gov/amber-waves/2005/april/north-american-greenhouse-tomatoes-emerge-as-a-major-market-force/ [08.09.2018].

CIA (2017), The world factbook. Field listing: exports – commodities, https://www.cia.gov/library/publications/the-world-factbook/fields/2049.html [08.09.2018].

Cook R.L., Calvin L. (2005), Greenhouse tomatoes change the dynamics of the North American fresh tomato industry, Economic Research Report no. 2, United States Department of Agriculture, http://ucce.ucdavis.edu/files/datastore/234-447.pdf [08.09.2018].

Davidova S., Latruffe L. (2007), Relationships between technical efficiency and financial management for Czech Republic farms, „Journal of Agricultural Economics”, vol. 58 no. 2, pp. 269-288.

Dodson M., Bachmann J., Williams P. (2002), Organic greenhouse tomato production. appropriate technology transfer for rural areas, www.attra.ncat.org [08.09.2018].

Engindeniz S., Tuzel Y. (2006), Economic analysis of organic greenhouse lettuce production in Turkey, „Scientia Agricola”, vol. 63 no. 3, pp. 285-290.

FAO (2018), FaoStat, http://www.fao.org/faostat/en/#data/QC [08.09.2018].

FIRA (2016), Panorama agroalimentario tomate rojo, Dirección de Investigación Económica y Sectorial, https://www.gob.mx/cms/uploads/attachment/file/200635/Panorama_Agroalimentario_Tomate_Rojo_2016.pdf [08.09.2018].

Gorton M., Davidova S. (2004), Farm productivity and efficiency in the CEE applicant countries. A synthesis of results, „Agricultural Economics”, vol. 30 no. 1, pp. 1-16.

Hung-Jen W., Ching-Chieng Ch., Po-Chi Ch. (2008), The cost effects of government subsidised credit. Evidence from farmers credit unions in Taiwan, „Journal of Agricultural Economics”, vol. 59 no. 1, pp. 132-149.

Henao M. (2001), Evaluación y caracterización morfológica del Lulo (Solanum quitoense). Tesis de ingeniero agronomo, Universidad Nacional de Colombia, Medellin [unpublished work].

Ibitoye S.J, Shaibu U.M., Omole B. (2015), Analysis of resource use efficiency in tomato (Solanum lycopersicum) production in Kogi State, Nigeria, „Asian Journal of Agricultural Extension, Economics & Sociology”, vol. 6 no. 4, pp. 220-229.

Kido-Cruz M., Kido A. (2007), Análisis comparativo de costos para el manejo y uso de suelo en la cuenca alta del río Caculta en Oaxaca, México, „Agrociencia”, vol. 41, pp. 355-352.

Kvaløy O., Tveteras R. (2008), Cost structure and vertical integration between farming and processing, „Journal of Agricultural Economics”, vol. 59 no. 2, pp. 296-311.

Laurinavičius E., Rimkuvienė D. (2017), The comparative efficiency analysis of EU members agriculture sectors, „Rural Sustainability Research”, vol. 37 no. 332, pp. 10-19.

Meeusen W., Van Den Broeck J., (1977), Efficiency estimation from Cobb-Douglas production functions with composed error, „International Economic Review", vol. 18 no. 2, pp. 435-444.

Mysore S., Wen-fei L.U. (2002). International competition in the greenhouse production of floriculture products – lessons for New York and India, Staff paper, Department of Agricultural, Resource and Managerial Economics, Cornell University, Ithaca NY, http://publications.dyson.cornell.edu/research/researchpdf/sp/2000/Cornell_Dyson_sp0004.pdf [08.09.2018].

Pindyck R., Rubinfeld D. (2009), Microeconomía, 7th ed., Pearson, Prentice Hall.

Rasmussen S. (2010), Scale efficiency in Danish agriculture. An input distance-function approach, „European Review of Agricultural Economics", vol. 37 no. 3, pp. 335-367.

Reyes H.M. (1995), Investigación agrícola, costo de la semilla mejorada y tasa de interés como instrumentos para el logro de la autosuficiencia alimentaria. El caso del maíz en Guatemala, 1975-90, Instituto de Ciencia y Tecnología Agrícolas, Guatemala.

Rucoba García A. et. al (2006), Análisis de la rentabilidad de un sistema de producción de tomate bajo invernadero en la región centro sur de Chihuahua, „Revista Mexicana de Agronegocios", vol. 10 no. 19, pp. 1-10.

Sánchez López E., Barrera Serrano M.J., Zavala J. (2004), Análisis del costo de producción del algodón 2001 en Mexicali Baja California, como un elemento de diseño de política de apoyo al subsector, „Revista Mexicana de Agronegocios", vol. 8 no. 14, pp. 198-210.

Stevenson Rodney E., (1980), Likelihood functions for generalized stochastic frontier estimation, „Journal of Econometrics", vol. 13 no. 1, pp. 57-66.

StataCorp (2015), Stata Statistical Software. Release 14, StataCorp LP, College Station, TX.

Torres Lima P. et al. (2004), Evaluación de la rentabilidad del desarrollo regional. El marco de la agricultura, „Región y sociedad", vol. 16 no. 29, pp. 109-144.

USDA (2016), Tomato report. Mexico continues to expand greenhouse tomato production, GAIN Report MX6021, https://gain.fas.usda.gov/Recent%20GAIN%20Publications/Tomato%20Annual_Mexico%20City_Mexico_6-1-2016.pdf [08.09.2018].

Vasiliev N., Astover A., Mõtte M., Noormets M., Reintam E., Roostalu H. (2008), Efficiency of Estonian grain farms in 2000-2004, „Agricultural and Food Science", vol.17 no. 1, pp. 1-40.

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 125-153, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Benchmarking technical efficiency of rice farms in Ghana: An empirical application of alternative production frontier approaches

Rebecca OWUSU COFFIE

The University of Western Australia, Australia, and The University of Cape Coast, Ghana

Abstract:

Aim: In spite of investments in new technologies to improve upon rice production in Ghana, productivity levels are still low. It is therefore important to assess the efficiency of farmers and identify sources of inefficiency to develop policies to reduce inefficiencies. This paper aims to investigate the extent and drivers of technical efficiency of rice farmers in Ghana.

Design/Research methods: bootstrap data envelopment and restricted single-stage stochastic frontier models are employed to examine the technical efficiency of farmers and its determinants. The data for empirical application come from a farm production survey comprising a total sample of 197 rice farmers in Ghana.

Conclusions and findings: The analyses revealed on average, farmers are about 65% technically efficient. This result indicates that there is a potential to improve upon technical efficiency of farmers by about 35% within the existing state of resources and technology. Furthermore, the drivers of technical efficiency were identified as food insecurity status and membership of farmer based organisation. Specifically, the results show an inverse relationship between food insecurity status and technical efficiency; where higher levels of food insecurity are associated with lower levels of technical efficiency. Also, membership of farmer based organisation increases technical efficiency of farmers. Contrary to previous studies, non-farm income and credit access were not identified as significant drivers of technical efficiency among the sampled farmers. On the basis of the findings, policies should aim at reducing food insecurity among farmers and encouraging membership of farmer based organisations.

Originality/value of the article: This paper provides evidence-based information on the extent of technical efficiency of rice farmers in Ghana and suggests measures for technical efficiency improvements.

Key words: Technical efficiency, rice production, food insecurity, new technologies, bootstrap data envelopment

JEL: Q1, Q12

Correspondence address: Rebecca Owusu Coffie, 35 Stirling Highway, Perth WA 6009 Australia. E-mail: rebecca.owusu@research.uwa.edu.au
Received: 30.09.2017, Revised: 11.07.2018, Accepted: 11.07.2018
doi: http://dx.doi.org/10.29015/cerem.577

1. Introduction

In Ghana, rice ranks second to maize as an important staple (Coffie et al. 2016; Ragasa et al. 2016) with consumption estimates around 30 kg/capita per year and this is predicted to reach 63kg/capita per year in 2018. In spite of the increasing demand of the crop, supply does not match up with demand and the deficit in supply is mainly accounted for through imports from other rice producing countries such as Vietnam (Angelucci et al. 2013). Currently, rice import constitutes 58 percent of total cereal imports in the country (Coalition for African Rice Development [CARD] 2010), a trend that is quite problematic because the international rice market is unstable and cannot be relied upon to supply the needed quantity of rice. Ravn (2014) asserted that population pressures and diminishing resources in many rice producing countries may affect the volume of rice traded by 2020, implying that there could be potential increases in prices at the world market, which may affect purchasing power at the domestic market, a phenomenon that would affect food security and general welfare of farmers.

Over the years, many Ghanaian governments have introduced policies with technologies to boost domestic rice production. Notable among these policies are the Medium Term Agricultural Development Programme (MTADP) in 1991-2000, Food and Agriculture Sector Development Policy One (FASDEP I) from 2002-2003, FASDEP 2 in 2007, Ghana Poverty Reduction Strategy 1 &2 (GPRS) from 2003-2009 and the Medium Term Agriculture Sector Investment Plan (METASIP) from 2009-2015. These policies have targeted various aspects of improving rice productivity such as the provision of extension services and improved seed varieties, adequate agronomic practices among others. Despite the investments made in the rice industry to boost productivity, production levels are still low. For instance, the average yield of rice per hectare in Ghana (1.9 tonnes/ha) is less than half of that of the world (4.3tonnes/ha). Given the importance of the crop in the country's food security agenda, there are suggestions for improvements in productivity. Performance (technical efficiency) measurement of rice production should, therefore, be considered as a main issue and an adequate policy based on efficiency improvements designed to ensure high production and food security.

The concept of technical efficiency measurement was initiated by the seminal article of Farrell (1957). In his paper on the measurement of productive efficiency, Farrell (1957) defined a simple measure of a firm's efficiency which accounts for multiple inputs within technical, allocative and economic efficiency. Since its introduction, there has been an expansion of the frontier methodology to investigate efficiency as an econometric and operational research method, with applications in the transport, financial and agricultural industries. Based on Farrell (1957) seminal work, a host of related frontier models classified as parametric and non-parametric were developed. The most common non-parametric frontier model is data envelopment analysis (DEA) proposed by Charnes et al. (1978). On the parametric front, later works led to the development of the stochastic frontier approach to efficiency estimation (Aigner, Chu 1968; Afriat 1972; Richmond 1974).

The revolution of the frontier methodology over time has been an interesting one: the original deterministic parametric frontier analysis has been replaced by DEA, which has increasingly become the most preferred measure of productive efficiency. DEA is mostly preferred because of its numerous advantages. First, it does not require assumptions about the underlying production technology and the error structure. Second, it has the capability to handle multiple inputs and outputs. On the other hand, the main attraction of the stochastic frontier approach (SFA) is its ability to account for noise in the data without necessarily attributing all deviations to inefficiency; the ability to conduct hypothesis testing and confidence interval construction. For extensive review and empirical applications of these two methodologies, see Emrouznejad et al. (2008), Seiford (1994), Bravo-Ureta and Pinheiro (1993), and Coelli (1995). However, in recent times, there have been criticisms of the DEA methodology due to its non-stochastic nature (Simar, Wilson 1998, 2007). Simar and Wilson (2007) highlighted the deficiency of the deterministic DEA method of sample estimates that exaggerate the level of efficiency within a sample. The authors proposed the use of the bootstrap approach to correct the estimates of technical efficiency. The introduction of the bootstrap procedure is meant to introduce stochasticity into the DEA model to account for noise and consequently allowing one to construct confidence intervals.

Empirically, the application of the frontier methodologies in developing countries has attracted much attention in the literature (Thiam et al. 2001; Bravo-Ureta, Pinheiro 1993; Battese 1992). Thiam et al. (2001) conducted a review of 32 studies on developing countries agriculture. Thiam et al. (2001) study revealed that Asia, in particular, India and Philippines have received the most attention from frontier researchers in developing countries. The study further reviewed an average technical efficiency value of 68% in developing countries agriculture. Brümmer (2001) also investigated the efficiency of private farms in Slovenia using the two approaches. The study concluded that there is a substantial degree of inefficiency by both methodologies, however, the DEA scores were found to be lower than SFA. In addition, Latruffe et al. (2004) conducted a study into the determinants of technical efficiency of crops and livestock farms in Poland and concluded that the DEA and SFA efficiency estimates are comparable. The study further identified two important determinants of efficiency: education and market integration. A comparative study of the Greek dairy farms was conducted by Theodoridis and Psychoudakis (2008). The results indicated not only the potential of improving the efficiency of the farmers but also the comparable nature of the two approaches of efficiency estimation: SFA and DEA. Although many studies have established the consistency of the results obtained from the two approaches, others have reported conflicting results (Fiorentino et al. 2006; Kumar, Arora 2010).

In the Ghanaian context, the frontier application to agriculture has not received much attention. Only a limited number of studies have examined the technical efficiency of farmers using the parametric frontier methods. For example, Al-Hassan (2008) investigated the technical efficiency of rice farmers in Northern Ghana and found an average technical efficiency of 53%. The study further found level of education, extension contact, farmers' age, family size as the drivers of technical efficiency. In another study, Al-hassan (2012) evaluated the technical efficiency of farmers in smallholder paddy farms in Ghana and obtained a mean technical efficiency of 64%. The determinants of technical efficiency according to the study were credit availability, family size and non-farm employment. Donkoh et al. (2013) studied the technical efficiency of rice production at the Tono irrigation scheme in the Northern region. Applying the stochastic frontier model to 85 farmers' data, they found a mean technical efficiency of 81%. The determinants of efficiency were identified as land, seed, fertilizer, crop expenditure, education and gender. Scope economies and technical efficiency of cocoa agroforestry systems in Ghana was studied by Ofori-Bah and Asafu-Adjaye (2011). The study objective was to examine the extent to which crop diversity affects farmer technical efficiency and whether cost complementarities exist from the sharing of farm inputs on the same plots within cocoa agroforestry industry in Ghana. The study employed the distance function approach of the stochastic frontier model to estimate the set objective. The study found the mean technical efficiency of multiple cropped cocoa farms as 86%. The determinants of technical efficiency by this study were the presence of shade trees, extent of crop diversity, age, education, gender of farmers and full time farming.

The following limitations can be identified from the previous technical efficiency studies in Ghana. Firstly, despite the importance of the bootstrapped DEA model, there have not been any empirical application in the Ghanaian agriculture. Secondly, none of the studies that applied the stochastic frontier model in the Ghanaian agriculture test for theoretical consistency, and impose the restrictions where applicable. Thirdly, previous researchers have not explored the possibility of drawing a comparison between the two approaches to technical efficiency in Ghana. Fourthly, the effects of determinants such as food insecurity and membership of farmer based organisation on technical efficiency have not been fully explored.

In this paper, the aforementioned limitations are addressed by examining the technical efficiency of rice farms in Ghana using restricted stochastic frontier and bootstrap DEA models. The restricted stochastic frontier model is based on the three–step approach introduced by Henningsen and Henning (2009) to impose theoretical restrictions on the stochastic frontier model. The restricted stochastic frontier approach is important to ensure that the estimated production function is theoretically consistent. On the other hand, the bootstrap DEA model follows the formulation of Simar and Wilson (2007). Beyond the estimation of the extent of technical inefficiency in farm production, the single stage stochastic frontier model and the Simar and Wilson (2007) second algorithm are employed to estimate the drivers of technical efficiency, with a specific emphasis on food insecurity and membership of farmer based organisation. Food insecurity status of farmers is very important because it tends to affect labour productivity with consequences on farm output. Food insecure farmers are likely to have low farm productivity compared to food secure farmers. Farmer based organisations (FBOs) are vital in delivering services to farmers (Addai et al. 2014). In Ghana for instance, FBOs have become a major policy objective in improving agricultural productivity among small-holder farmers. Specifically, in the area of rural service delivery and farm credit access. It is therefore important to provide evidence on the effects of FBOs on technical efficiency to help in policy formulation for productivity improvements.

The data for the empirical application come from a farm household production survey conducted in Ghana. Fitting restricted stochastic frontier and bootstrap DEA models to the data, an estimated average technical efficiency of 0.65 was obtained, suggesting that there is a potential to increase rice productivity in Ghana within the current state of inputs and technology. The results also showed that rice output is more responsive to intermediate input relative to land and labour inputs. In addition, it was observed that a majority of the sampled farmers are operating under increasing returns to scale, indicating that farm sizes are too small. Furthermore, food insecurity and membership of farmer based organisations were identified as the primary drivers of technical efficiency in Ghana. Therefore, to improve productivity in rice production, government must address the challenges associated with poor food security status of farmers and also encourage membership of farmer based organisations.

The rest of the paper is organised as follows: Section 2 presents production frontier estimation techniques; section 3 discusses model estimation procedure; section 4 describes data used in the empirical application, section 5 presents the empirical results and a comparison of the models, and finally; section 6, concludes the paper.

2. Production frontier estimation techniques

The frontier represents a best-practice technology among the farms and deviations are referred to as inefficiency. The frontier approach is used to measure productive efficiency. There are two main types of frontier methodology: parametric and non-parametric. The parametric frontiers (stochastic frontier approach-SFA) estimate efficiency using econometric techniques while the non-parametric frontiers (data envelopment analysis-DEA) measure efficiency using linear programming techniques. Also the parametric frontiers are stochastic while the non-parametric frontiers are deterministic. Whiles the deterministic approach assumes that any deviation from the frontier is inefficiency, the stochastic approach accounts for statistical noise. In the present study, both the parametric and non-parametric approaches are employed to examine the technical efficiency of rice farmers in Ghana and to derive its determinants. Following is a detailed specification of the stochastic frontier and the data envelopment models.

2.1. Stochastic frontier approach

The stochastic frontier approach (SFA) uses econometric techniques to specify the production, cost, revenue or profit function with a specific shape and makes assumptions about the distribution of the inefficiency and error terms (Eling, Luhnen 2008). The use of this approach may depend on the number of outputs (production function or distance function) and the type of data (cross sectional or panel data). Assuming a single output, the production frontier for x_i vector of *K* inputs may be specified as in (1):

$$y_i = f(x_i\!:\!\beta).TE_i \qquad (1)$$

where; y_i is the output, $f(x_i\!:\!\beta)$ is the production frontier which is deterministic, β is the vector of parameters to be estimated and TE_i is the output oriented technical efficiency. Reorganising (1) yields technical efficiency: $TE_i = \frac{y_i}{f(x_i:\beta)}$. When $TE_i = 1$, then the DMU is fully efficient, however, if $TE_i < 1$, then there is a deviation from the frontier. This deviation may entirely be attributed to inefficiency as is the case in deterministic frontiers or noise and inefficiency within the stochastic framework.

The stochastic frontier model incorporates a composed error structure with a two sided symmetric and a one sided component (Aigner et al. 1977; Van den Broeck et al. 1994). The one sided component reflects inefficiency whiles the two sided one captures the random effects outside the control of the production unit as well as measurement errors and other statistical noise typical of empirical relationships. SFA may be specified as in (2):

$$y_i = f(x_i:\beta) + v_i - u_i \qquad (2)$$

where; v_i is the stochastic random term (two sided component) indicating effects such as the environmental factors beyond the control of the farmer, measurement errors in the dependent variable and left-out explanatory variables, and u_i is the technical inefficiency term representing the factors that can be controlled by the farmer such as farm management factors. The distribution of the inefficiency term can be either half normal ($v_i \sim iiN(0, \sigma v^2)$), truncated normal ($v_i \sim iiN^+(\mu, \sigma v^2)$) or exponential ($v_i \sim iidexponential$) (Stevenson 1990; Aigner et al. 1997; Meeusen, Broeck 1977). SFA further requires a functional form specification, which according to Coelli et al. (2005) is based on the flexibility, linearity, regularity and parsimony of the functional form. Stressing on the flexibility, the normalised translog model provides second order approximation to the underlying technology (Abdulai, Huffman 1998; Coelli et al. 2005). For detailed discussion of the distributional assumptions and their proofs, see Coelli et al. (2005) and Kumbhakar (2003).

The estimation of technical efficiency alone is not enough; we have to identify the sources of inefficiency to derive policies to address those specific factors. In the stochastic frontier literature, the initial approach to accounting for the effects of environmental variables on the production frontier was the two stage approach. The two stage approach involves first stage estimation of the technical efficiency and a second stage regression of the environmental variables on technical efficiency (Battese, Coelli 1995). Over the years, the two stage approach has fallen out of favour in the empirical literature because of the potential biases that it introduces into the model estimations (Battese, Coelli 1995). Battese and Coelli (1995) introduced a single-stage estimation technique to address the challenges associated with the two stage approach. The single stage model involves a simultaneous estimation of the production frontier and the drivers of technical efficiency. The environmental variables (hereafter z variables) when introduced into stochastic frontier model in (2) results in the following specification:

$$y_i = f(x_i;\beta) + v_i - u_i + z_i \qquad (3)$$

where z_i is a vector of explanatory variables affecting efficiency of the farmers and all other variables are as earlier defined.

2.2. Data envelopment approach

The data envelopment approach (DEA) uses linear programming techniques to estimate the efficiency scores, which are measures of performance. The DEA gives a piece-wise linear frontier that envelopes the observed input and output data. The best practice production frontier for a sample of decision making units is constructed through a piecewise linear combination of actual inputs and output. All DMUs that lie on the frontier are referred to as efficient, whiles those that do not lie on the frontier are considered as inefficient. The first DEA model was introduced by Charnes et al. (1978) and ever since, many researchers have recognized it as an excellent tool for performance evaluation. Charnes et al. (1978) DEA model was based on the constant return to scale (CRS) assumption. However, the CRS is only suitable when all farms are operating at the optimal scale. In reality, most farms in developing countries may not be operating on an optimal scale, a situation that requires an alternative approach. Banker et al. (1984) introduced the variable return to scale (VRS) frontier to address the deficiencies of the CRS model.

The DEA model can either be input or output oriented depending on whether an input or output distance function is applied. In most empirical applications, the input oriented model is applied. In this paper, both the input oriented and the output oriented models are employed. The DEA approach assumes that all farms within a sample have access to the same technology for the transformation of a vector of *N* inputs denoted by *x*, into a vector of *M*, outputs, denoted as *y*. The input set under the variable return to scale is represented by:

$$L_v(Y) = \left\{\begin{matrix} X: \sum_{j=1}^{n} \lambda_j y_{kj} \geq y_k, & k = 1, \dots s; \\ \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_k, & i = 1, \dots m; \\ \sum_{j=1}^{n} \lambda_j = 1; & \\ \lambda_j \geq 0, & j = 1, \dots n \end{matrix}\right\} \quad (4)$$

The associated variable return to scale, piecewise linear output set and technology are represented as:

$$P_v(X) = \left\{\begin{matrix} Y: \sum_{j=1}^{n} \lambda_j y_{kj} \geq y_k, & k = 1, \dots s; \\ \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_k, ki = 1, \dots m; & \\ \sum_{j=1}^{n} \lambda_j = 1; & \\ \lambda_j \geq 0, & j = 1, \dots n \end{matrix}\right\} \quad (5)$$

and

$$T_v = \left\{\begin{matrix} (X, Y): \sum_{j=1}^{n} \lambda_j y_{kj} \geq y_k, & k = 1, \dots s; \\ \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_k, & i = 1, \dots m; \\ \sum_{j=1}^{n} \lambda_j = 1; & \\ \lambda_j \geq 0, & j = 1, \dots n \end{matrix}\right\} \quad (6)$$

The Farrell (1957) input oriented measure of technical efficiency (defined as minimizing input use in production to produce the same level of output) is obtained by solving the following linear programming problems *N* times:

$$F_v(Y_0,X_0) = \min \theta$$
subject to
$$\sum_{j=1}^{n} \lambda_j y_{kj} \geq y_{k0} \quad k = 1, \dots s;$$
$$\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{i0} \quad i = 1, \dots m; \qquad (7)$$
$$\sum_{j=1}^{n} \lambda_j = 1;$$
$$\lambda_j \geq 0, \qquad j = 1, \dots n$$

The output oriented model specified in (8) under the variable return to scale assumption is based on the premise of maximising output from existing resources. These two measures according to Coelli et al. (2005) are equivalent measures of the TE when constant return to scale exist.

$$F_v(Y_0,X_0) = Max\theta$$
subject to
$$y_{k0} \leq \sum_{j=1}^{n} \lambda_j y_{kj} \quad k = 1, \dots s;$$
$$\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{i0} \quad i = 1, \dots m;$$
$$\sum_{j=1}^{n} \lambda_j x_{ij} = \lambda_{ij} x_{ij}; \qquad (8)$$
$$\lambda_j = 1, \qquad j = 1, \dots n$$
$$\lambda_j \geq 0, \qquad j = 1, \dots n$$

where; θ is the input technical efficiency measure having a value of $0 \leq \theta \leq 1$. If $\theta = 1$, then the farmer is efficient. The vector λ is an $N \times 1$ vector of weights which defines the linear combination of the peers of the *i-th* farmer. Xλ and Yλ are efficient projections on the frontier. $N1'$ is an $N \times 1$ vector of ones.

The inability of the standard DEA model to account for measurement errors makes it less desirable for efficiency estimation particularly, in developing countries where data quality is a problem. This is because all deviations from the frontier are attributed to inefficiency. Some attempts have thus been made to introduce statistical properties to the DEA estimator. Simar and Wilson (2007) proposed the bootstrapping procedure to correct for the bias in the DEA estimator. Naïve bootstrapping method has been implemented by some authors (Lothgren 1999; Ferrier, Hirschberg 1997). However, Simar and Wilson (2007) criticise the naïve bootstrapping procedure as yielding inconsistent estimates and suggested that a smoothed bootstrap procedure is a more suitable alternative approach that yields consistent result.

This paper applies the Simar and Wilson (2007) model to correct for bias in the DEA estimator and construct confidence intervals. The central idea behind bootstrapping is to simulate a true sampling distribution by mimicking the data generating process (DGP). The smooth bootstrap procedure is based on the assumption that the distribution of the efficiency scores are normally distributed. Initially, the efficiency scores are estimated from the original data to produce a pseudo data where the output is fixed with an adjustment of the input vector by the estimated efficiency scores. Based on the generated pseudo data, new efficiency scores are calculated for samples in the data. Replicating the process yields the empirical distribution of the efficiency measures. For a detailed description of the bias corrected DEA model, see Simar and Wilson (2007).

2.2.1. Scale efficiency and return to scale

Technical efficiency measured by the VRS frontier corresponds to pure technical efficiency, which reflects the managerial performance of farmers to organise inputs into outputs. The pure technical efficiency measure is devoid of scale efficiency, which considers the size of operation. It is therefore important to consider the effect of size of operation on farmers' productivity. Scale efficiency (SE) is a measure of the optimal level of production. Specifically, SE is the ratio between the CRS technical efficiency and VRS technical efficiency. The measure of SE provides farmers with the ability to choose the optimum size of resources to attain the expected level of production. Inappropriate size of farm operation (that is either too small or big) may affect technical efficiency of farmers. The type of inefficiency of farm operations resulting in technical inefficiency is referred to as scale inefficiency. Based on scale inefficiency, we can compute the return to scale, which is a relative measure of the relationship between a constant return to scale frontier and a decreasing return to scale frontier. Return to scale can be classed as either increasing return to scale (IRS), decreasing return to scale (DRS) or constant return to scale (CRS). When a farm experiences IRS, then the farm size is too small to be efficient, and under DRS, the farms are too big. Farms operating under CRS indicates the farm has an optimal scale of production.

3. Data

The data for the empirical application come from a farm household production survey conducted between August and September 2014 in Northern and Upper East regions of Ghana for the 2013/2014 production season. The two regions were selected as the study sites because of the volume of rice produced and closeness to the major rice market, Ashanti region. A study by Angelucci et al. (2013) revealed that the two regions contribute almost 70% of the total volume of rice produced and therefore play a major role in the country's food self-sufficiency in rice production. Analysis of technical efficiency of rice production using these two regions is very important to identity regional variations in efficiency for a sound policy recommendation. A multistage sampling technique was employed in the data collection process. In the first stage, the stratified sampling method was used to categorise regions into districts and later communities. The simple random sampling technique was then used to select the farming households based on names provided by the Ministry of Food and Agriculture. Using face to face interview techniques, a total of 197 farming households were interviewed.

Based on previous studies, three inputs and a single output were considered in the production function estimation. Output was measured as an amount of paddy rice produced per hectare of rice farm. The inputs include farm size, labour and intermediate inputs. Farm size (X1) was measured as total area cultivated to rice in hectares. Labour (X2), total person-days committed to the production process by both family and hired labour. The family labour who counted are persons of the family unit that reside in the house and are actively involved in the production process, and intermediate input costs (X3): This was an aggregation of other production costs such as harrowing, seed cost, fertilizer cost, ploughing, and herbicides in Ghana Cedis (2.8 Ghana Cedis is equivalent to 1 USD).

With respect to the environmental variables, the following variables were considered: food insecurity, access to credit, membership of farmer based organisations and non-farm income activity. Food insecurity was measured using the Household Food insecurity Access Scale indicator developed by Coates et al. (2007). The raw food insecurity data obtained from the survey contained several variables and could not be used in the estimation directly. A factor analysis was conducted to reduce the number of variables. Specifically, tetrachoic factor analysis was employed to generate food insecurity score. The food insecurity score was used in the final model estimation, a value close to zero is indicative of higher food security and that close to 1, higher level of food insecurity. The non-farm income variable is a dummy representing whether a farmer is engaged in non-farm income activity or not. Similarly, membership of farmer based organisation measures whether a farmer is a member of farmer based organisation or not. The access to credit variable is also a dummy variable showing whether a farmer has access to credit.

Table 1. Descriptive statistics of data

Variable	Mean	SD	Min.	Max.
Inputs				
Farm size (x1) in ha	1.21	1.013	0.405	8.098
Labour (x2) in man-days	121.97	318.807	5	3628
Intermediate input (x3) cost	499.85	371.73	50	2510
Output				
Rice yield (y)	1761.52	1861	336	21000
Environmental variables				
Food insecurity	0.24	0.48	0	1
Access to credit (Yes=1)	0.157	0.365	0	1
Farmer group membership[1] (Yes=1)	0.645	0.479	0	1
Non-farm income activity (Yes=1)	0.289	0.455	0	1

Source: Author's own elaboration.

Sample descriptive statistics are presented in Table 1. It is evident from the table that farm sizes are mainly small with an average of about 2 hectares. This is indicative of the small-scale nature of rice production system and typical of production systems in developing countries (Coffie et al. 2016). In addition, it can also be observed from the table that on average, the sample are fairly food secure. The farmer group variable also shows that on average most farmers belong to existing farmer groups. The non-farm income activity variable indicates that most farmers are not engaged in non-farm income activities.

4. Model estimation

Five different models were estimated: a single stage restricted translog stochastic frontier model and four variants of the bootstrapped DEA model. Regarding the SFA model, first, a standard model was estimated and then tested for a violation of the monotonicity assumption. The test revealed a violation of the

[1] Note that farmer based organisation and farmer group is used interchangeably in this paper.

monotonicity assumption and this was corrected by estimating a restricted model by imposing the assumption. In the frontier literature, three approaches to imposing monotonicity restriction are identified: restricted maximum likelihood estimation (Bokusheva, Hockmann 2006), Bayesian inference approach (O'Donnell, Coelli 2005) and the three-step approach (Henningsen, Henning 2009). The three-step approach was adopted in this paper because it is easy to implement compared to the other approaches (Henningsen, Henning 2009). The three-step approach involves 1) estimating an unrestricted stochastic frontier model and extracting the unrestricted parameters of the frontier and the covariance matrix from the estimation 2) estimating a minimum distance function and extracting the restricted parameters 3) determining technical efficiency estimate of the farms and effects of variables explaining inefficiency based on the theoretically consistent frontier. For detailed information about the three-step approach, see Henningsen and Henning (2009). With respect to the bootstrapped DEA model, both the input and output oriented models were estimated under CRS, DRS and VRS using the second algorithm proposed by Simar and Wilson (2007): The approach involves a joint estimation of the frontier and the "z" variables that affect technical efficiency. The estimation of the models was implemented in R statistical programming platform.

5. Results and discussion

The results of the estimations are presented as follows. First, the production function estimates for the restricted stochastic frontier analysis model and bootstrapped DEA model estimates are presented. Second, scale efficiency, return to scale and determinants of technical efficiency are discussed.

5.1. SFA results

The estimation results from the unrestricted stochastic production function (first step) are presented in Table 2. It is observed from Table 2 that the coefficients of the production inputs are significant and all positive. By production theory, the production function should be monotonic. A test of the monotonicity condition of the production frontier showed that apart from the intermediate input, monotonicity is violated at 41.1 percent for the land input and 9.1 percent for the labour input. When monotonicity assumption is violated, the efficiency estimates are not easily interpretable. Given that monotonicity is violated in the data, a restricted model was estimated following the minimum distance approach of Henningsen and Henning (2009). Tables 3 and 4 report the estimates from the restricted model estimation. The last column in Table 3 shows the estimates from the restricted coefficients after adjusting the production frontier with the estimates in the final step. Results of the minimum distance function presented in Table 3 show that many of the coefficients have changed, however, similar to the unrestricted function (Table 2), production is more responsive to intermediate inputs. The estimated model parameters are considered as theoretically consistent.

Table 2. Unrestricted stochastic frontier estimation

Parameters	Estimate	Std. error
Production function		
β_0	0.416***	0.064
β_1	0.191**	0.059
β_2	0.273***	0.037
β_3	0.593***	0.055
β_{11}	0.217*	0.108
β_{12}	0.132*	0.054
β_{13}	-0.077	0.075
β_{22}	0.075.	0.041
β_{23}	0.131.	0.069
β_{33}	0.082	0.103
σ^2	0.385***	0.075
γ	0.891***	0.047
Inefficiency effects function		
δ_1	-0.257	0.187
δ_2	-0.481	0.337
δ_3	0.271***	0.040
δ_4	-0.447*	0.176

Source: Author's own elaboration.

NB: δ_1=Non-farm income, δ_2=credit access, δ_3=food insecurity, δ_4=farmer organisation membership

Considering the adjusted input elasticities (last column, Table 3), it is observed that output is more responsive to intermediate input, followed by labour input and land input. The input elasticity of intermediate input (0.575) implies that a unit change in intermediate input will result in 0.575 change in output. Similarly, a percentage change in labour and land inputs will cause a 0.294 and 0.136 change in output, respectively. Regarding return to scale, which is a summation of the first order input elasticities (1.005), we observe that farmers are operating under a slightly increasing return to scale. The cross-product of the input elasticities are relatively small, giving indication of a limited opportunity for input substitution.

Table 3. Minimum distance estimation

Parameters	Estimate	Diff.	Adj. estimates
β_0^0	0.485	0.069	0.484
β_1^1	0.136	-0.055	0.136
β_2^2	0.295	0.022	0.294
β_3^3	0.576	-0.017	0.575
β_{11}^0	0.041	-0.176	0.041
β_{12}^0	0.032	-0.1	0.031
β_{13}^0	-0.044	0.033	-0.044
β_{22}^0	0.082	0.007	0.082
β_{23}^0	0.088	-0.043	0.088
β_{33}^0	0.070	-0.012	0.070

Source: Author’s own elaboration.

The final model estimates in Table 4 show that the coefficient of the intercept is almost zero and the coefficient of the “frontier output” is closer to one, implying that the coefficients of the adjusted and the non-adjusted restricted production frontier estimates (Table 3) are almost identical. This result is comparable with the findings of Henningsen and Henning (2009). Unlike the findings of Henningsen and Henning (2009), there is a slight variation in the total error variance between the unrestricted (Table 2) and final model estimates (Table 4) after imposing the monotonicity restriction. Also, there is a decrease in the proportion of the variance of technical efficiency in the total error variance after imposing the restriction.

5.1.1. Partial production elasticity

The response of output to varying levels of each of the production inputs while holding the levels of the other inputs constant was examined. Results of the input elasticity calculated from the partial frontiers presented in Table 5 show that the elasticity values of all the inputs are positive, implying that output can be increased by increasing the level of individual inputs. The mean partial production elasticity of land increased from 0.077 in the unrestricted model to 0.117 in the restricted model. Similarly, the mean partial elasticity of the labour input increased from 0.171 to 0.223 after imposing the monotonicity restriction. Finally, the mean partial elasticity of the intermediate input increased slightly from 0.522 to 0.524 in the restricted model. A mean comparison test was calculated to test whether there is significant difference between the restricted and unrestricted parameters. The results show that there is a significant difference between the restricted and unrestricted estimates except for the intermediate inputs.

Table 4. Final stochastic frontier estimation

Parameters	Estimate	Std. error
α_0	0.005	0.046
α_1	0.997***	0.055
σ^2	0.382***	0.071
γ	0.879***	0.043
δ_1	-0.241	0.191
δ_2	-0.26	0.279
δ_3	0.268***	0.041
δ_4	-0.449*	0.180

Source: Author's own elaboration.
NB: δ_1 =Non-farm income, δ_2 =credit access, δ_3 =food insecurity, δ_4 =farmer organisation membership

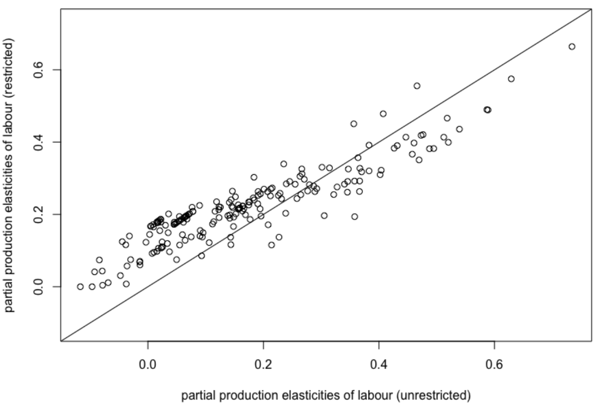
Table 5. Partial production elasticity

	Unrestricted model			Restricted model		
	X1	X2	X3	X1	X2	X3
Mean	0.077	0.171	0.522	0.117	0.223	0.524
Std.	0.189	0.163	0.138	0.042	0.108	0.098
Min	-0.437	-0.117	0.174	0.000	0.000	0.292
max	0.72	0.734	1.064	0.249	0.664	0.898

Source: Author's own elaboration.

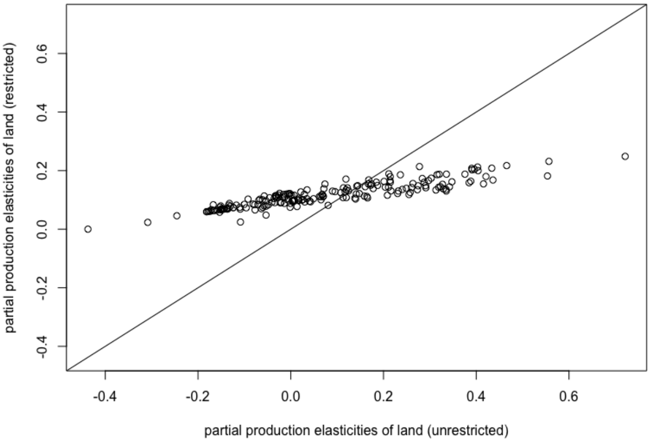
Next, a graphical correlation plot (Figures 1-3) is constructed to illustrate effects of imposing monotonicity restriction on the production input elasticities. The figures reveal that estimates based on the restricted and the unrestricted models are highly correlated with coefficient correlation of 0.91, 0.92 and 0.99 for land, labour and intermediate inputs, respectively. Although there is nearly perfect correlation between the restricted and the unrestricted input coefficients, one cannot conclude that models that do not impose theoretical restrictions on the production frontier are accurate. This is mainly because theoretically inconsistent production frontier would likely affect further estimations from the production frontier (O'Donnell, Coelli 2005).

Figure 1. Partial production elasticities of land input. Note the unfilled circles show observations with monotonicity violated in the unrestricted model



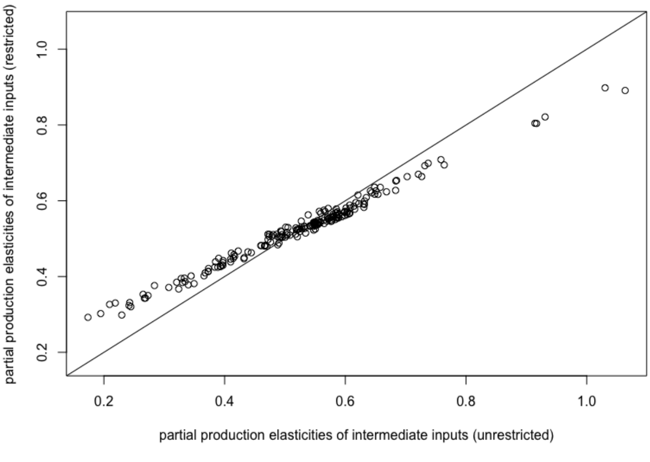
Source: Author's own elaboration.

Figure 2. Partial production elasticities of labour input. Note the unfilled circles show observations with monotonicity violated in the unrestricted model



Source: Author's own elaboration.

Figure 3. Partial production elasticities of intermediate input. Note the unfilled circles show observations with monotonicity violated in the unrestricted model



Source: Author's own elaboration.

5.2. DEA results

The input and output oriented DEA technical efficiency model estimates under variable return to scale (VRS), constant return to scale (CRS) and decreasing return to scale (DRS) are presented in Table 6. Columns 2-5 represent the input oriented model estimates, whiles the output oriented model estimates are presented in columns 6-10. The average technical efficiency under the VRS frontier for the standard DEA model for the input and output oriented specifications are 0.72 (Table 2) and 0.57, respectively. Correspondingly, the bias corrected estimates are 0.650 and 0.49 for the input and output oriented models, respectively. Based on the technical efficiency estimates, there is a potential to increase rice yield in Ghana with the available input and technology. The estimates obtained under the CRS and DRS models within the input oriented and output oriented models are similar and much lower than the VRS frontier estimates. These findings confirm previous studies outcome that the VRS frontier results in higher technical efficiency estimate (Matawie, Assaf 2010). The result further implies the standard DEA model has an upward bias (Balcombe et al. 2008).

Next, interval estimates of technical efficiency are discussed. As can be observed from the table, the average point estimates of technical efficiency have 0.1, 0.2 interval for the 95% confidence interval for the output and input oriented models under the CRS, VRS and DRS, respectively. However, the more efficient farms within the sample, the 95% confidence interval is significantly wider as can be illustrated by the minimum and maximum results for alternative orientations for the input and output oriented models, respectively.

Table 6. DEA input oriented technical efficiency estimates and determinants of technical efficiency

	Input oriented model				Output oriented model			
	Mean	Std.	Min.	Max.	Mean	Std.	Min.	Max.
TE_{vrs}	0.72	0.24	0.18	1	0.57	0.24	0.08	1
TE_{bc}	0.65	0.22	0.17	0.98	0.49	0.19	0.07	0.87
LB	0.56	0.21	0.15	0.98	0.39	0.17	0.07	0.75
UB	0.71	0.24	0.18	0.99	0.55	0.24	0.08	0.98
TE_{crs}	0.51	0.22	0.06	1	0.51	0.22	0.06	1
TE_{bc}	0.47	0.19	0.05	0.94	0.47	0.19	0.05	0.94
LB	0.41	0.17	0.05	0.88	0.41	0.17	0.05	0.88
UB	0.50	0.21	0.06	0.98	0.50	0.21	0.06	0.98
TE_{nirs}	0.51	0.22	0.06	1	0.52	0.22	0.08	1
TE_{bc}	0.45	0.19	0.05	0.82	0.47	0.19	0.07	0.89
LB	0.38	0.15	0.05	0.75	0.40	0.16	0.06	0.79
UP	0.50	0.22	0.06	0.97	0.51	0.21	0.08	0.97
Second stage model for the VRS frontier function								
	Mean		LB	UB		Mean	LB	UB
δ_0	-1.25	-	-43.37	-5.97	-9.82	-	-25.64	-0.73
δ_1	-0.37	-	-12.10	0.55	-0.25	-	-5.83	2.25
δ_2	0.07	-	0.01	0.75	0.23	-	0.002	0.52
δ_3	0.41	-	1.02	6.57	0.23	-	0.84	5.09
δ_4	-0.19	-	-6.79	2.39	-4.61	-	-10.88	-1.49

Source: Author's own elaboration.

NB: LB-Lower bound, UB-Upper bound, δ_1=Non-farm income, δ_2=credit access, δ_3=food insecurity, δ_4=farmer organisation membership

5.2.1. Scale efficiency and nature of return to scale

Inefficiency in farm production may be attributable to either pure technical inefficiency or scale inefficiency. The subject of scale inefficiency is briefly discussed here. Scale efficiency and nature of return to scale were calculated from the DEA frontiers. Specifically, scale efficiency was calculated as the ratio of the CRS technical efficiency to the VRS technical efficiency, whiles the nature of return to scale was examined as the ratio of CRS to DRS technical efficiencies. The calculated mean scale efficiency and the nature of return to scale are presented in Table 7. From the table, the calculated scale efficiency value of 74% suggests that scale inefficient farms can reduce size by 26% without affecting the current output levels. With respect to nature of return to scale (Table 7), the result show that about two thirds of the sample farms operate under increasing return to scale, implying that these farms are too small to be efficient. They therefore have to expand operations to produce on the production frontier. Another 24 farms operate under constant return to scale and these farms are fully scale efficient. Finally, 14 farms operate under decreasing return to scale, suggesting that these farms are not fully technically efficient because they are too large.

Table 6. Scale and nature of return to scale estimates

Nature of return to scale	Number of farms	Percent
IRS	157	79.7
DRS	14	7.1
CRS	27	13.7
Scale efficiency		Percent
Mean	-	74
Std.	-	24
Min	-	21
Max.	-	100

Source: Author's own elaboration.
Note: IRS-increasing return to scale, DRS-decreasing return to scale, CRS-constant return to scale

5.3. Drivers of technical efficiency

The drivers of technical efficiency are now examined. The second block of Table 3 reports on the inefficiency effects model from the stochastic frontier approach, whiles that of the DEA model is reported in the second block in Table 6.

The simultaneous estimation approach was adopted in the SFA model estimation, whiles the second algorithm of Simar and Wilson (2007) bootstrap DEA function was employed in the DEA model. Following previous studies (Simar, Wilson 2007; Balcombe et al. 2008), a positive sign on an explanatory variable indicates an obstacle to technical efficiency, while a negative sign indicates a positive influence on technical efficiency. From the tables, it is observed that the food insecurity variable is negative and significant, suggesting that lower levels of food insecurity increases the technical efficiency of farmers. Also, the farmer group membership is negative and significant, indicating that farmer group membership increases technical efficiency of farmers. This finding is similar to previous studies (Bhatt, Bhatt 2014) outcome of the effects of farmer group membership on technical efficiency. Unlike previous studies (Villano, Fleming 2006), the non-farm income variable was not significant in explaining technical efficiency of farmers. The access to credit variable is also not a significant driver of technical efficiency in the sample.

5.4. Model comparison

A comparison is drawn between the restricted and unrestricted SFA estimates and also between the SFA and DEA model estimates. First, estimates from the two functions are similar, a finding that corroborates Henningsen and Henning (2009) study outcome. This is confirmed by the near perfect correlation of the technical efficiency estimates in Figures 4-5. Comparing the technical efficiency of the DEA model (output oriented) and that of the SFA model, it is observed that the DEA model technical efficiency estimate is slightly lower than the SFA model estimates, however, the technical efficiency estimates from the two models are highly correlated. Similar findings have been reported in Balcombe (2008) study on alternative frontier methodologies in Austrian dairy farms. The positive correlation between the SFA and the DEA model estimates show that either methodology could be employed in estimating the technical efficiency of farmers in developing countries, particularly when one accounts for biases in the DEA model.

6. Conclusions

The demand for rice in Ghana does not commensurate domestic supply. Over the years, various Ghanaian governments have introduced policies with new technologies to promote rice productivity in Ghana. With an increasing investment in rice production, it is becoming important for farmers to become more efficient in their ability to access and use available technologies to improve productivity. Improving productivity requires adequate assessment of farmers' efficiency and identification of the sources of inefficiency so that better policy and institutional innovations could be introduced to reduce inefficiencies in rice production. In this paper, the restricted single stage stochastic frontier (SFA) and the bootstrap data envelopment (DEA) models are applied to examine the technical efficiency of rice farms in Ghana.

Employing a total sample of 197 rice farms, the mean technical efficiency estimates were 0.65 for the SFA and 0.49 for the input oriented DEA model, suggesting that DEA potentially underestimates the technical efficiency of farmers. Based on the SFA estimate, it can be inferred that farmers can increase rice output by 35% within the existing state of inputs and technology. In addition, the results revealed that rice output is more responsive to intermediate input use. Generally, however, production is inelastic with respect to the inputs used in the production process. Furthermore, the results indicate that there is less avenue for input substitution. The mean scale efficiency was 74% with majority of the farmers exhibiting increasing return to scale, suggesting that generally, farmers have the capability to increase output with the current resources available to them.

Regarding the drivers of technical efficiency, food insecurity and membership of farmer based organisation were identified as the primary drivers of technical efficiency among the sampled farmers. Policy makers should therefore initiate measures to improve the food security status of farmers and encourage them to become members of farmer based organisations.

References

Abdulai A., Huffman W.E. (1998), An examination of profit inefficiency of rice farmers in Northern Ghana, „Economic Staff Paper Series”, Iowa State University, https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1295&context=econ_las_staffpapers [09.09.2018].

Addai K.N., Owusu V., Danso-Abbeam G. (2014), Effecrts of farmer based organisation on the technical efficiency of maize farmers across various agro-ecological zones of Ghana. „Journal of Economics and Development Studies”, vol. 2 no. 1, pp. 141-161.

Afriat S.N. (1972), Efficiency estimation of production functions, „International Economic Review”, vol. 13 no. 3, pp. 568-598.

Aigner D.J., Chu S.-F. (1968), On estimating the industry production function, „The American Economic Review”, vol. 58 no. 4, pp. 826-839.

Aigner D., Lovell C., Schmidt P. (1977), Formulation and estimation of stochastic frontier production function models. „Journal of Econometrics”, vol. 6 no. 1, pp. 21-37.

Al-Hassan S. (2008), Technical efficiency of rice farmers in Northern Ghana, Research paper, African Economic Research Consortium.

AL-Hassan S. (2012), Technical efficiency in smallholder paddy farms in Ghana. An analysis based on different farming systems and gender, „Journal of Economics and Sustainable Development”, vol. 3 no. 5, pp. 91-105.

Angelucci F., Asante-Poku A., Anaadumba P. (2013), Analysis of incentives and disincentives for rice in Ghana, „Monitoring African Food and Agricultural Policies”, FAO, Rome.

Balcombe K., Fraser I., Latruffe L. et al. (2008), An application of the DEA double bootstrap to examine sources of efficiency in Bangladesh rice farming, „Applied Economics”, vol. 40 no. 15, pp. 1919-1925.

Battese G.E. (1992), Frontier production functions and technical efficiency. A survey of empirical applications in agricultural economics, „Agricultural Economics”, vol. 7 no. 3-4, pp. 185-208.

Battese G.E., Coelli T.J. (1995), A model for technical inefficiency effects in a stochastic frontier production function for panel data, „Empirical Economics”, vol. 20 no. 2, pp. 325-332.

Bokusheva RA., Hockman H. (2006), Production risk and technical inefficiency in Russian agriculture, „European Review of Agricultural Economics”, vol. 33, no. 1, pp. 93-118.

Bravo-Ureta B.E., Pinheiro A.E. (1993), Efficiency analysis of developing country agriculture. A review of the frontier function literature, „Northeastern Agricultural and Resource Economics Association”, vol. 0 no. 1, pp. 1-14.

Brümmer B. (2001), Estimating confidence intervals for technical efficiency. The case of private farms in Slovenia, „European Review of Agricultural Economics”, vol. 28 no. 3, pp. 285-306.

Coalition for African Rice Development (CARD) (2010), Mapping of poverty reduction strategy papers (PRSPs), sector strategies and policies related to rice development in Ghana, https://riceforafrica.net/downloads/CARD%20DOCS/Ghana_Mapping_report_Final_version.pdf [09.09.2018].

Charnes A., Cooper W.W., Rhodes E. (1978), Measuring the efficiency of decision making units, „European Journal of Operational Research", vol. 2 no. 6, pp. 429-444.

Coelli T.J. (1995), Recent developments in frontier modelling and efficiency measurement, „Australian Journal of Agricultural and Resource Economics", vol. 39 no. 3, pp. 219-245.

Coelli T.J., Rao D.P., O'Donnell C.J., Battese G.E. (2005), An introduction to efficiency and productivity analysis, Kluwer Academic Publishers, Boston.

Coffie R.O., Burton M.P., Gibson F., Hailu A. (2016), Choice of rice production practices. A comparison between preference space and willingness to pay space, „Journal of Agricultural Economics", vol. 67 no. 3, pp. 799-819.

Donkoh S., Ayambila S., Abdulai S. (2013), Technical efficiency of rice production at the Tono Irrigation scheme in northern Ghana, „American Journal of Experimental Agriculture", vol. 3 no. 1, pp. 25-42.

Eling M., Luhnen M. (2008), Frontier efficiency methodologies to measure performance in the insurance industry. Overview and new empirical evidence, University of St. Gallen Working Papers on Risk Management and Insurance Paper, no. 56.

Emrouznejad A., Parker B.R., Tavares G. (2008), Evaluation of research in efficiency and productivity. A survey and analysis of the first 30 years of scholarly literature in DEA, „Socio-Economic Planning Sciences", vol. 42, pp. 151-157.

Farrell M.J. (1957), The measurement of productive efficiency, „Journal of the Royal Statistical Society. Series A (General)", vol. 120 no. 3, pp. 253-290.

Ferrier G.D., Hirschberg J.G. (1997), Bootstrapping confidence intervals for linear programming efficiency scores. With an illustration using Italian banking data, „Journal of Productivity Analysis", vol. 8 no. 1, pp. 19-33.

Fiorentino E., Karmann A., Koetter M. (2006), The cost efficiency of German banks. A comparison of SFA and DEA, „Deutsche Bundesbank Discussion Paper. Series 2: Banking and Financial Studies", no. 10, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.137.4435&rep=rep1&type=pdf [09.09.2018].

Henningsen A., Henning C.H.C.A. (2009), Imposing regional monotonicity on translog stochastic production frontiers with a simple three-step procedure, „Journal of Productivity Analysis", vol. 32, pp. 217-229.

JIRCAS (2012), Manual for improving rice production in Africa, https://www.jircas.go.jp/en/publication/research_results/2011_04 [09.09.2018].

Kumar S., Arora N. (2010), Analyzing regional variations in capacity utilization of Indian sugar industry using non-parametric frontier technique, „Eurasian Journal of Business & Economics", vol. 2 no. 4, pp. 1-26.

Kumbhakar S.C. (2003), Stochastic frontier analysis, Cambridge University Press, Cambridge.

Latruffe L., Balcombe K., Davidova S., Zawalinska K. (2004), Determinants of technical efficiency of crop and livestock farms in Poland, „Applied Economics", vol. 36 no. 12, pp. 1255-1263.

Lothgren M. (1999), Bootstrapping the Malmquist productivity index. A simulation study, „Applied Economics", vol. 6 no. 11, pp. 707-710.

MOFA (2009), National rice development strategy, https://www.jica.go.jp/english/our_work/thematic_issues/agricultural/pdf/ghana_en.pdf [09.09.2018].

O'Donnell C.J., Coelli T.J. (2005), A Bayesian approach to imposing curvature on distance functions, „Journal of Econometrics", vol. 126 no. 2, pp. 493-523.

Ofori-Bah A., Asafu-Adjaye J. (2011), Scope economies and technical efficiency of cocoa agroforestry systems in Ghana, „Ecological Economics", vol. 70 no. 8, pp. 1508-1518.

Ragasa C., Dankyi A., Acheampong P., Wiredu A.N., Chapoto A., Asamoah M., Tripp R. (2013), Patterns of adoption of improved rice technologies in Ghana, International Food Policy Research Institute Working Paper no. 35, http://www.ifpri.org/publication/patterns-adoption-improved-rice-technologies-ghana [09.09.2018].

Ravn K. (2014), Agriculture. The next frontier, „Nature", no. 514, pp. S64-S65.

Richmond J. (1974), Estimating the efficiency of production, „International Economic Review", vol. 15 no. 2, pp. 515-521.

Seiford L.M. (1994), A DEA bibliography (1978–1992), in: Data Envelopment Analysis. Theory, Methodology, And Applications, Charnes A., Cooper W.W., Lewin A.Y., Seiford L.M. (ed.), Springer, Dordrecht.

Simar L., Wilson P.W. (1998), Sensitivity analysis of efficiency scores. How to bootstrap in nonparametric frontier models, „Management Science", vol. 44 no. 1, pp. 49-61.

Simar L., Wilson P.W. (2007), Estimation and inference in two-stage, semi-parametric models of production processes, „Journal of Econometrics", vol. 136 no. 1, pp. 31-64.

Theodoridis A., Psychoudakis A. (2008), Efficiency measurement in Greek dairy farms. Stochastic frontier vs. data envelopment analysis, „International Journal of Economic Sciences and Applied Research", vol. 1 no. 2, pp. 53-66.

Thiam, A., Bravo-Ureta, B. E. and Rivas, T. E. 2001. Technical efficiency in developing country agriculture: a meta-analysis. „Agricultural Economics, 25, 235-243.

Van Den Broeck, J., Koop, G., Osiewalski, J. and Steel, M. F. 1994. Stochastic frontier models: A Bayesian perspective. „Journal of Econometrics, 61, 273-303.

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 155-174, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Measuring productivity and efficiency of seaports in India using DEA technique

Jeronimo Guilherme Remigio MONTEIRO

Goa University, Farmagudi, Ponda, India

Abstract:

Aim: In this paper we are looking at the seaports (in India called 'major ports') from the context of its trade and India's strategic importance in trade world after the initiation of economic reforms in 1991. It empirically estimates the levels of productivity and efficiency of seaports in India. This paper applies DEA technique to assess productivity and efficiency of seaports in India.

Design/Research methods: DEA technique is extensively used in the literature of economics to provide measures of firms' technical efficiency. These measures rank the firms by looking at their apparent performances over a period of time. DEA is a frontier model which is non-parametric since no functional specification or form is required to be mentioned.

Conclusions/Findings: The DEA results as discussed and reported in the paper have shown how Indian ports are performing over the years. This investigation alone is not sufficient to develop a benchmark in the port system of India. Rather it will do well to have a closer look at the Indian ports from the physical and financial performance point of view. This study made use of data envelopment analysis (DEA) to generate what we call an efficiency benchmarks and assessment of the Indian ports sector. With this modest attempt to investigate the port sector of India several issues are in the open one can further analyze and come to desired conclusions.

Originality/value of the paper: The main role of a port is to transfer goods between two transport modes. As far as Indian ports are concerned, there are few studies with regard to productivity and efficiency of the port sector. Since, there is an attempt in recent years to overhaul the infrastructure sectors of the Indian economy and especially seaports. There is a need to look at issues in port sector as well. Productivity and efficiency concerns should be the main aspect of the benchmarking of the performance of today's Indian ports.

Limitations of the research: Second stage DEA, distance function approach, Bayesian techniques, Carlo Monte techniques, can be alternatively used.

Key words: Productivity, Efficiency, Frontier, Parametric, Non-parametric, DEA.
JEL: C14; D24

Correspondence address: Jeronimo Guilherme Remigio Monteiro, PES, RSN College of Arts & Science, a.t.Goa University, Ponda Goa 403110, INDIA. E-mail: jgrm@rediffmail.com
Received: 27.08.2017, Revised: 15.11.2017, Revised: 08.03.2018, Revised: 04.07.2018, Accepted: 22.07.2018
doi: http://dx.doi.org/10.29015/cerem.529

1. Introduction

The transport sector of the Indian economy which comprises of railways, roads, ports and airways, have been the main focus for years from the policy makers of the country and especially more so since the ushering of economic reforms in 1991. Given the overall transport sector's contribution to the economic growth and development of this country, we need to relook at the policies and programs of the transport sector at large from the context of development of the economy. Ports are one such area in the entire transport sector of the economy which takes the major chunk of the budget of the Union Government. Investments in ports require huge amounts in terms of money and machinery. Table 1 shows investments in the transport sector of the Union Government during the first nine five year plans (1951-2002). Planning Commission is now replaced by NITI (National Institution for Transforming India) and such five year plans have been discontinued.

India has an extensive coastline of 7517 Kms, spread on the western and eastern shelves of the mainland and also along the Islands. Its coast is spread along nine States and four Union Territories. The nine coastal states of India namely; Gujarat, Maharasthra, Goa, Karnataka, Kerala (West Coast) and Tamil Nadu, Andhra Pradesh, Orissa, West Bengal (East Coast) have in all 12 major ports and around 200 minor and intermediate ports (often referred to as non-major ports).

The major ports of India are six on the west coast namely, Kandla, Mumbai, Jawaharlal Nehru (near Mumbai), Mormugao, New Mangalore and Cochin, and six on the east coast Tuticorin, Chennai, Vishakapatnam, Paradip, Kolkatta, and Haldia (though Haldia is a satellite port of Kolkatta). The Ennore Port Limited (EPL), being a newly constructed major port, is the first corporate major port registered under the Companies Act, 1956. It was commissioned in the year 2001.

The entries relating to the development of maritime ports are in the Seventh Schedule of the Constitution of India and therefore come under the purview of the Centre and the States as well. The twelve major ports of India are placed in the Union List of the Indian Constitution and, are such statutory bodies (trusts) administered by the Ministry of Shipping, Government of India under the provisions of the Major Port Trust Act, 1963 and Indian Ports Act, 1908.

Table 1. Planned investments in the transport sector in India 1951-2002

Sector	1st Plan (1951 -56)	2nd Plan (1956 -61)	3rd Plan (1961 -66)	Annual Plans (1966-69)	4th Plan (1969-74)	5th Plan (1974-79)	Annual Plans (1979-84)	6th Plan (1980-85)	7th Plan (1985-90)	Annual Plans (1990-92)	8th Plan (1992-97)	9th Plan (1997-2002)
Railways	217	723	1326	509	934	2063	714	6585	16549	10208	32302	46405
Roads & Bridges	147	242	440	309	862	1701	467	3887	6335	3656	16095	47600
Road Transport	-	-	27	55	128	503	143	1276	2151	986	3538	5933
Ports	28	33	93	53	249	488	57	725	1513	668	2302	5331
Shipping	19	53	40	32	155	469	147	468	720	939	3033	2909
IWT	-	-	4	6	11	16	6	63	188	57	152	280
LH&LS	-	-	4	2	6	9	2	*	*	4	25	58
Civil Aviation	23	49	49	66	177	294	132	957	1948	1055	7249	6599
Other Transport	-	-	-	-	-	-	-	-	72	118	244	1851
Transport total	434	1100	1983	1032	2522	5543	1668	13961	29476	17691	64940	117325
Transport as % of total plan	22.05	23.50	23.15	15.60	15.98	14.08	-	13.02	13.51	14.12	13.00	-

Source: Govt. of India, Planning Commission, Vision 2020 Transport, Mahesh Kapoor Report, 2002; * Included in the port sector

As far as Indian ports are concerned, there are few studies with regard to productivity and efficiency of the port sector. Studying Indian ports is crucial in today's competitive and globalized environment especially after the initiation of economic reforms in India in 1991. Accordingly, this study was necessitated to put things in right perspective for the seaports of India as regards to productivity and efficiency. Since, there is an attempt in recent years to overhaul the infrastructure sectors of the Indian economy and especially seaports. There is a need to look at issues in port sector as well. Productivity and efficiency concerns should be the main aspect of the benchmarking of the performance of today's Indian ports. Efficiency is indeed core in policy considerations and hence need to be quantified objectively in order to help monitor progress of the port sector.

With this background in mind, our objectives of this study clearly spell out our intention to carry forward our work on ports. The study had set two objectives- to undertake a comprehensive review of seaports in India and to empirically estimate the levels of productivity and efficiency of seaports in India. Both this objectives are being fulfilled with this study on seaports of India. The study employed the literature review and the author's own survey carried out for the study period of two years (2012 to 2014). The required data for our empirical analysis spans for eighteen years (1996 to 2014).

2. Literature survey

Ports are engines of growth and development for the economies they serve. They are thus the economic drivers of entire economies. A nation's international links and trade depends upon good port infrastructure and services. Ports form a vital link in the overall trading chain and, consequently, their level of performance and efficiency determines to a large extent a nation's international competitiveness (Tongzon, Ganesalingam 1994). Ninety five percent of the Indian overseas trade in volume terms and seventy five per cent in value terms are sea-borne.

Seaborne traffic depends on seaports for all its operations, since ports acts as interfaces between maritime and inland modes of transport (roads, railways or inland navigation system). Therefore, in order to have an efficient maritime transport system, seaports must work efficiently so as to benefit the users (shippers, exporters/importers, etc.) adequately. The basic objective of a seaport is to provide for goods and passengers a fast and safe transit through its facilities, so that generalized costs for passengers (fare + time) and for shippers (tariffs + storage time) are minimized. Another role that some large seaports play is to serve as hubs for connection and transshipment, allowing cargoes on different long-haul routes to be served more efficiently by several ships (Trujillo, Nombela 1999).

Port efficiency varies widely from country to country and, specially, from region to region. It is a well-known fact that some Asian ports (Singapore, Hong Kong) are the most efficient ports in the world, while some of the inefficient ports are in Africa (Ethiopia, Nigeria, Malawi) or in Latin America (Colombia, Venezuela, Ecuador) (Micco, Perez 2001). Of late, with port reforms in the world in the form of privatization and public-private partnership agreements, there are many investments in port sector under these agreements which are doing exceptionally well and have turned their terminals into profitable ventures.

Poitras et al. (1996) focuses on port efficiency and competitive environment in port industry. According to them ports efficiency is important for trade, economic development of the region and to face international competitive environment. DEA model was applied to measure port performance and efficiency of 23 ports across the world. The DEA empirical analysis uses two output measures: the number of twenty feet equivalent units (TEUs) containers handled per berth hour and the total number of containers handled per year both 20 and 40 foot. The input measures used are: container mix, average delays in commencing stevedoring (calculated as a difference between the berth time and gross working time), average quay crane productivity represented by the number of containers lifted per quay crane hour, the number of gantry crane present at the port, the frequency of container ship calls, and the average government port charges per container. The empirical findings of their DEA model are the set of two results for the CCR (Charnes, Cooper and Rhodes) and additive DEA models. Their results reveal that the CCR model identifies more substantially more inefficient ports (13 vs. 4) than the additive model, and ignoring a marginal increase for Port of Jakarta, attributes a higher level of inefficiency to those ports which are judged to be inefficient using both the models. Further commenting on their result depends upon the assumption they made about the production technology of ports. Ports that are judged to be inefficient with variable returns to scale (VRS) will also be inefficient with linear production relations, but not otherwise. Besides efficiency rankings their results also identify the efficient facet being used for comparison as well a combination of the inputs which are being inefficiently utilized and the deviation of specific outputs from the efficient level. The final conclusion of their findings is that port efficiency results depends upon the type of DEA model employed which, in turn depends upon the assumption made about the returns to scale properties of the port production function.

Tongzon (2001) examines the efficiency with respect to containerized cargoes across ports recognized for their high level of performance in Asia and Europe. Their study uses two outputs and six input measures of port performance for sixteen container ports for the year 1996. The output measures are cargo throughput and ship working rate, whereas port inputs are number of berths, number of cranes, and number of tugs, terminal area, delay time and labour (proxy to the number of port authorities' employees). Empirical results for CCR (Charnes, Cooper and Rhodes) models and additive DEA models ranked ports for efficiency. CCR model identifies six ports as slightly inefficient, while the additive model identifies three as inefficient ports. This is true as CCR model fits a linear production technology and the additive models features variable returns to scale, which require a larger number of ports to define the efficiency frontier. Again, ports that are judged to be inefficient with variable returns to scale will also be inefficient with linear production technology, but not the converse. Further, a ports' operational efficiency level does not depend solely on its size or its function (i.e., whether it is a hub or feeder port). Size of the port is not a determinant of port efficiency.

Barros (2003) on Portuguese port authorities set out an objective to measure and compares the efficiency and performance of a sample of Portuguese seaports to indirectly infer the role of incentives introduced by the Portuguese policy regulation. The activity they are studying of Portuguese seaports is the work carried out by the port authorities i.e. multi-output. The output variables comprise of ten indicators: ship, movement of freight, gross tonnage, market share, break-bulk cargo, containerized cargo, roll-on/roll-off (ro-ro) traffic, dry bulk, liquid bulk, and net income. Whereas the input variables consist of two indicators: labour measured by number of workers and capital measured by the book value of the assets. Their analysis show mixed results but, overall it can be said that the majority of the seaports are in the efficient frontier. Mean technical, allocative and economic efficiencies with constant returns to scale, declined for Portuguese seaports from 1999 to 2000, indicating incentive policy regulation has failed to drive the Portuguese seaports towards the efficient frontier. While in the case of variable returns to scale the mean value of all efficiencies is improving slightly during the period. Barros (2003) concludes that the results and the implications the reforms had on the Portuguese seaports. Especially the port of Aveiro is an exception to the efficiency results. In final the study proposes some managerial guidelines – the Portuguese Maritime Port Agency must upgrade inspection procedures regarding seaport activities in order to provide explicit incentives for increasing productive efficiency. Further scope of the data must be expanded to include contextual factors beyond managerial control. Data gathered must be published and there should be transparency in data information and dissemination to all stakeholders of the Portuguese seaports. And finally, benchmarks were provided for improving the operation of least performing seaports.

Park and De (2004) study on Korean ports for their performance and efficiency. A four stage DEA was applied, to overcome the limitations of basic DEA models: alternating the consideration of the variables as inputs and outputs to measure the productivity (stage 1), profitability (stage 2), marketability (stage 3), and the overall efficiency (stage 4). The outputs selected for estimation depends on what is being merchandise: total merchandise and number of ships (productivity); income (profitability); customer satisfaction (commercialization and global efficiency). The variables used as inputs are –docking capacity and cargo handling capacity (productivity and overall efficiency), cargo throughput and number of ship calls (profitability) and income (marketability). The efficiency result of CCR as well as BCC models for 11 Korean ports is ranked in order of productivity, profitability, marketability and overall efficiency. Efficiency results with reference to returns to scale was also calculated and the ports were classified as decreasing returns to scale, increasing returns to scale and constant returns to scale with both CCR and BCC models.

Cullinane et al. (2004) study on efficiency of container terminals of a sample of 25 ports comprised of a cross-sectional data 200 observations for the period 1992-1999(8 years). The product output that is measured by them is a container terminals in TEUs (twenty feet equivalent units), whereas the productive input used by them is the capital which also measures the work input (i.e. it incorporates input labour). Analysis says that efficiency of container terminals is not influenced by the size of the port. Most of the ports have constant returns to scale, which indicated that the scale of production is not the main source of inefficiency, which means port competition and competitiveness may have a major and direct impact on the measured levels of efficiency within container ports. There are other reasons of inefficiency in port production like differences in port ownership or governance, locational advantages and the form and level of competition faced.

Cullinane et al. (2005) study employs mathematical programming approach to measure container port production efficiency to the top 30 container ports of the world in 2001. The study uses alternative techniques of DEA and Free Disposal Hull (FDH) model. The results of these two techniques give interesting insight into current efficiency rankings and gives different variations. Use of appropriate variable definition of input and output factors is crucial element in meaningful applications of DEA and FDH. It is clear from this that a combination of DEA and FDH analysis can be of great significance and value to managerial decisions of ports and terminals and to the strategic decisions of port authorities.

Chudasama and Pandya (2008) study is the first one measuring efficiency of Indian ports by making use of Data Envelopment Analysis (DEA). Their main objective is to bring out the actual working and performance of the port sector in India. Port input variables used in their study are seven: No. of Cranes, No. of Berths, Storage area in Sq. mts, Average pre-berthing time in Days and Average turnaround time in days. The single output variable taken by them is Cargo Volume in million tonnes. The results of their study reveal a complete efficiency picture of Indian ports for the year 2005-06. DEA-BCC model yield a higher efficiency estimates than DEA-CCR model with average values of 0.98 and 0.86 respectively. Out of the 12 ports, 7 ports were identified as efficient and 5 ports turned out to be relatively inefficient when DEA-CCR model was applied. When DEA-BCC model was applied, all the ports except Paradip turned out to be efficient in the analysis. Empirical results also show that large scale of production is more likely to be associated with high efficiency scores. For instance correlation between efficiency score and port output is 0.84 for the DEA-CCR model. Another observation of their study is that port output is significantly correlated with number of vessels handled and the storage area. And lastly empirical results reveal that there exists waste in the production at 5 sample ports. The average efficiency of these 5 ports derived from DEA-CCR model amounts to 0.86. This shows that in theory the ports under study can on average increase their outputs to 1.16 (=1/0.86) times as much as their current level, by using the same level of inputs.

3. Methodology

Based on the review, we are now in a position to state the methodology to be use for this study. The Malmquist DEA technique is a non-parametric technique to compute technical efficiency when panel data is available. The Malmquist DEA technique was elaborated by Caves, Chriestensen and Diewert (1982) and Fare et al. (1994b). The Malmquist total factor productivity (TFP) index measures the TFP change between two data points by calculating the ratio of the distances of each data point relative to a common technology. This TFP change can be decompose into technical change and technical efficiency change. If the period *c* technology is used as the current period technology, the Malmquist (input-oriented) TFP change index between period *b* (the base period) and period *c* (the current period) can be written as:

$$m_i^c(q_b, x_b, q_c, x_c) = \frac{d_i^c(q_c, x_c)}{d_i^c(q_b, x_b)} \quad (1)$$

Note that in the above equation the notation $d_i^c(q_c, x_c)$ represents the distance from the period *c* observation to period *b* technology. A value of m_i greater than one indicates positive TFP growth from period *b* to period *c* while a value less than one indicates a TFP decline. According to Fare et al. (1994b) these two periods (*b* and *c*) indices are only equivalent if the technology is Hicks input neutral.

Following Fare et al. (1994b) specifies an input-oriented Malmquist productivity change index between period '*b*' (base period) and period '*c*' (current period) as:

$$m_i(y_b, x_b, y_c, x_c) = \left[\frac{d_i^c(y_c, x_c)}{d_i^b(y_b, x_b)} \text{X} \frac{d_i^b(y_c, x_c)}{d_i^c(y_b, x_b)} \right]^{1/2} \quad (2)$$

where:

y_c and y_b represent vector of outputs in period c and b respectively

x_c and x_b represent vector of inputs in period c and b respectively

An equivalent way of writing the Malmquist TFP index given in equation 2 is as follows:

$$m_i(y_b, x_b, y_c, x_c) = \frac{d_i^c(y_c, x_c)}{d_i^b(y_b, x_b)} \left[\frac{d_i^b(y_c, x_c)}{d_i^c(y_c, x_c)} \mathrm{X} \frac{d_i^b(y_b, x_b)}{d_i^c(y_b, x_b)} \right]^{1/2} \quad (3)$$

Where the ratio outside the square bracket measures the change in the input-oriented measure of Farrell technical efficiency between period '*c'* and '*b'* i.e. the efficiency change is equivalent to the ratio of technical efficiency in period '*b*' to the technical efficiency in period '*c*'. The remaining part of the equation 3 is a measure of technical change. It is the geometric mean of the shift in technology between the two periods evaluated at y_b and also at y_c. Thus the two terms in equation 3 are:

$$\text{Efficiency change} = \frac{d_i^c(y_c, x_c)}{d_i^b(y_b, x_b)} \quad (4)$$

and

$$\text{Technical change} = \left[\frac{d_i^b(y_c, x_c)}{d_i^c(y_c, x_c)} X \frac{d_i^b(y_b, x_b)}{d_i^c(y_b, x_b)} \right]^{1/2} \quad (5)$$

The DEA Malmquist techniques involves estimation of four distance functions in equation 2 which will involve four linear programming (henceforth referred as LP) problems, and subsequently, computation of TFP change using either equation 2 or 3. The four LPs are given below:

We begin by assuming CRS technology. The CRS input oriented LP is stated as below:

$$
\begin{aligned}
&\left[d_i^b\left(x_b, y_b\right)\right]^{-1} = \max_{\phi,\lambda,} \phi, \\
&st \\
&-\phi y_{jb} + y_b \lambda \geq 0, \\
&x_{jb} - X_b \lambda \geq 0, \\
&\lambda \geq 0,
\end{aligned}
\qquad (6)
$$

The remaining three LP problems are simple variants of this:

$$
\begin{aligned}
&\left[d_i^c\left(x_c, y_c\right)\right]^{-1} = \max_{\phi,\lambda,} \phi, \\
&st \\
&-\phi y_{jc} + y_c \lambda \geq 0, \\
&x_{j,c} - X_c \lambda \geq 0, \\
&\lambda \geq 0,
\end{aligned}
\qquad (7)
$$

$$
\begin{aligned}
&\left[d_i^b\left(x_c, y_c\right)\right]^{-1} = \max_{\phi,\lambda,} \phi, \\
&st \\
&-\phi y_{j,c} + y_b \lambda \geq 0, \\
&x_{j,c} - X_j \lambda \geq 0, \\
&\lambda \geq 0,
\end{aligned}
\qquad (8)
$$

$$\left[d_i^c\left(x_b, y_b\right)\right]^{-1} = \max_{\phi,\lambda,} \phi,$$
$$st$$
$$-\phi y_{jc} + y_c \lambda \geq 0,$$
$$x_{jc} - X_c \lambda \geq 0, \quad (9)$$
$$\lambda \geq 0,$$

Note that in LP's 7 and 8, where production points are compared with technologies from different time periods, the ϕ parameter need not be ≥ 1, as it must be when calculating Farrell efficiencies. The point could lie above the feasible production set. This will most likely occur in LP 7 where a production point from period *c* is compared with technology in period *b*. If technical progress has occurred, then a value of $\phi < 1$ is possible. It could also possibly occur in LP 9 if technical regress has occurred but this is less likely.

Also note that the φ and λ's likely to take different values in the above four LPs. All the above LPs must be solved for each firm in the sample. Thus in our case there are 12 firms (port trusts) and assume two time periods, 48 LPs must be solved. An extra time periods are added, one must solve an extra three LPs for each firm (to construct a chained index). If there are *T* time periods, then *(3T-2)* LPs must be solved for each firm in the sample. Hence, if there are *I* firms, then there are *I x (3T - 2)* LPs to be solved. In case of our study on ports *I=12* firms (port trusts) and *T=18* time periods, this would involve *12 x (3x18-2) =624* LPs.

The above approach can be extended by decomposing the (CRS) technical efficiency change into scale efficiency and a 'pure' (VRS) technical efficiency measure. This requires solution of two additional LPs (when comparing two production points). This would involve repeating LPs 8 and 9 with the convexity restriction *(N1'λ=1)* added to each. This provides estimates of distance functions relative to a variable returns to scale (VRS) technology. The number of LPs calculated accordingly increases from *N x (3T - 2)* to *N x (4T - 2).*

4. Data sources

The database used for this study is annual data relating to production (outputs) and productive factors (inputs) of twelve major ports of India. The data spans over 18 years period starting from 1996-97 and ending in 2013-14. This enabled us to get a fair idea about the functioning and progress of major ports of India in the last three years of last millennium and fifteen years of present millennium. This provides a panel of data of 216 observations which is adequate enough to adopt DEA (non parametric) method to estimate productivity and efficiency of major ports of India. In this study the latest port (Ennore Port Limited) could not be included because it was commissioned only in the year 2001. From the point of view of operational performance, Kolkatta and Haldia are taken as separate ports, otherwise, from the financial performance point of view it is always clubbed as a single major port. The 12 major ports covered by the study are: Chennai, Cochin, Haldia, Jawaharlal Nehru, Kandla, Kolkatta, New Mangalore, Mormugao, Mumbai, Paradip, Tuticorin and Visakhapatnam. These ports are spread across eastern and western coast of India. Six ports each being on both side of the coast. Four of the major ports Kolkatta, Mumbai, Chennai and Mormugao are the oldest being more than a century old.

The production (output) variable taken for non parametric estimation the two output variables taken are–volume of cargo traffic in million tonnes and number of vessels handled. We believe that there is some correlation between the volume of cargo handled and the number of vessels handled at the port. There are various measures of port’s inputs (productive factors) in this study. The seven inputs are land, labour, number of cranes, number of other equipments, number of berths, etc. Land is approximated by the storage area in square meters. It includes transit shed, warehouse container freight station, open area, etc. Labour is measured by the number of workers employed in each port. It consists of workers of all types – official, administrative, non-administrative and workers who load and unload ships (stevedores). The capital input is measured as number of berths in each port of study. Berth is a basic infrastructure for the ships. Whole lot of other infrastructure is used through the berth like number of cranes – which consists of Mobile, Wharf, Container Yard and Quay Cranes also, and number of other equipments used in cargo handling operation – which consists of Fork/Top lift Truck, Reach Stacker, Tractors, Trailers, Locomotives, Dozers, Excavators, etc. Taking all this into consideration berth is used a proxy for a capital input for this study. Also a good performance measure for a port is also the average pre-berthing time in days and average turnaround time in days, which is also included as inputs in this study. All these inputs are taken together for DEA, especially because DEA allows use of multiple outputs and inputs. These inputs are significant because they state the quality, quantity, capacity of infrastructure and operational performance of the ports. Further, capital investment to provide berth, cranes and other equipment incurred by the port authorities is huge running into crores of rupees and is spent over a particular range of period.

The required panel data for eighteen year period (1996-97 to 2013-14) was sourced from secondary sources. The major sources of our data collection comes from various issues of ‘Economic Intelligence Service, Centre for Monitoring Indian Economy, Mumbai, “Infrastructure”, “Basic Port Statistics of India”, published by Ministry of Shipping, Government of India, and “Major Ports of India: A profile”, published by Indian Ports Association, New Delhi. Besides, Annual Administration Reports of Major Ports (various issues) were also referred to compile the entire data set for this study. This data set is also supplemented by several visits to some of the port trust offices.

5. Empirical analysis and results

Performance evaluation and benchmarking are a widely used method to identify and adopt best practices as a means to improve the performance and increase productivity. Accordingly, this study has attempted to take stock of the Indian port scenario and provide possible benchmarks to policy makers and regulators in India. The non parametric estimation is carried out by using the software package called DEAP Version 2.1 (Coelli 1996b). The DEAP software employed in this study was developed at the Centre for Efficiency and Productivity Analysis, University of New England, Australia and is described by Tim Coelli (1996b).

We calculate Malmquist Productivity Index summary as well as the efficiency change, technical change and scale change components for each port in our sample and for all ports (at different time periods). First, we present malmquist indices for all the ports at different time periods. All the indices presented in the Table 2 are relative to the previous year and hence the indices begin with year 1996-97 as the base year.

Table 2. Malmquist productivity index – summary of annual means, 1996-97 to 2013-14

Year	Efficiency Change	Technical Change	Pure Efficiency Change	Scale Efficiency Change	Total Factor Productivity (TFP) Change
1996-97	1.000	1.000	1.000	1.000	1.000
1997-98	0.978	1.118	0.997	0.981	1.093
1998-99	0.983	1.123	0.988	0.996	1.104
1999-2000	0.987	1.132	0.994	0.993	1.118
2000-01	1.007	1.196	1.016	0.991	1.204
2001-02	0.956	1.057	0.985	0.971	1.011
2002-03	1.000	1.098	1.001	0.999	1.098
2003-04	0.987	1.167	0.956	1.033	1.152
2004-05	1.063	0.968	1.050	1.012	1.029
2005-06	1.012	1.005	1.013	0.999	1.017
2006-07	1.010	1.022	1.004	1.007	1.032
2007-08	0.955	1.881	0.952	1.004	1.796
2008-09	0.993	0.492	1.039	0.956	0.489
2009-10	1.059	1.021	1.003	1.056	1.082
2010-11	0.915	0.904	0.958	0.955	0.827
2011-12	1.069	0.959	1.053	1.015	1.025
2012-13	1.019	1.076	0.992	1.028	1.097
2013-14	0.997	1.081	1.008	0.989	1.078
Mean	0.999	1.047	1.000	0.999	1.046

Source: authors' own elaboration

Out of the 18 year period Indian ports exhibit scale efficiency for 7 years (2003-04, 2004-05, 2006-07, 2007-08, 2009-10, 2011-12 and 2012-13). That means they exhibits scale inefficiency for 11 years. Five indices are presented for all the ports in each year in the next five columns after the year column and for the different ports (for over all time periods) in Table 2. These five indices relate to efficiency change, technical change, pure efficiency change, scale efficiency change and Total Factor Productivity (TFP) change. From the Malmquist DEA analysis, TFP Index can be decomposed into technical efficiency change and technical change. Referring to

Table 1 the average total factor productivity change for all the 12 ports for the study period was 1.046, i.e. a growth of 4.6 percent over the sample period which can be best described as marginal considering the massive port structure. Further, there is hardly any improvement in the efficiency score and the technical change for the entire period was 4.7 per cent.

Next, we turn to a summary description of the average performance of each port over the entire 1996-2014 time period. All this information is presented in Table 3.

Table 3. Malmquist productivity index – summary of firm means, 1996-97 to 2013-14

Ports	Efficiency Change	Technical Change	Pure Efficiency Change	Scale Efficiency Change	Total Factor Productivity Change
Chennai	1.003	1.089	1.000	1.003	1.093
Cochin	0.999	1.020	1.000	0.999	1.019
Haldia	1.000	1.077	1.000	1.000	1.077
JNPT	1.000	1.068	1.000	1.000	1.068
Kandla	1.000	1.058	1.000	1.000	1.058
Kolkatta	1.000	1.065	1.000	1.000	1.065
N. Mangalore	1.000	1.086	1.000	1.000	1.086
Mormugao	0.983	1,001	1.000	0.983	0.985
Mumbai	1.000	1.023	1.000	1.000	1.023
Paradip	1.000	1.039	1.000	1.000	1.039
Tuticorin	1.000	1.012	1.000	1.000	1.012
Visaz.	0.999	1.029	1.000	0.999	1.028
Mean	0.999	1.047	1.000	0.999	1.046

Source: authors' own elaboration

Note that if the value of the Malmquist index or any of its components is less than 1, which denotes regress or deterioration in performance, whereas values greater than 1 denote improvements in the relevant performance. Also, note that these measures capture performance relative to the best practice in the sample, where the best practice represents a 'world frontier', and the world in our case is defined as the ports in our sample. Now let us look at the Table 2, we see that, on average, productivity increased marginally over the 1996-97 to 2013-14 period for the ports in our sample – the average change in the Malmquist productivity index was 4.6 per cent as a whole. Moreover on an average, that growth was due to innovation (technical change) rather than improvements in efficiency (efficiency change). Moving to results across the ports, we note that Chennai has 9.3 per cent and New Mangalore has 8.6 per cent total factor productivity change on average, and that the entire change for both ports was due to innovation (technical change). In fact, for both ports technical change was the highest in the sample. (i.e. Chennai New Mangalore made use of modern technology). The total factor productivity change for Mormugao was lowest or best one could say was negligible at 0.985. Mormugao's total factor productivity change was far below than the sample average 4.6 percent. Of late this port on the west coast of India handles 40 per cent of the country's export of the cargo as iron ore and it is considered to be a main port for iron ore export in India. However, after 2012 the business of iron ore shrink due total ban imposed by the Supreme Court of India on account rampant illegal ore extraction is some states of India including Goa. Probably the dismal performance of Mormugao port may be due to this impact.

The DEA results as discussed and reported above have shown how Indian ports are performing over the years. This investigation alone is not sufficient to develop a benchmark in the port system of India. Rather it will do well to have a closer look at the Indian ports from the physical and financial performance point of view.

6. Conclusion

The main role of a port is to transfer goods between two transport modes. This activity requires coordination of a large number of activities that can be organized in many different ways. As pointed out by Friedrichsen (1999), the assessment of the performance of a port must, thus, be to address the efficiency of the overall port system. This concern we feel should be the main aspect of the benchmarking of the performance of today's Indian ports. Efficiency is indeed core in policy considerations and hence need to be quantified objectively in order to help monitor progress of the port sector. This study made use of data envelopment analysis (DEA) to generate what we call an efficiency benchmarks and assessment of the Indian ports sector. Alternatively this process can be also done in a number of ways like second stage DEA, distance function approach, Bayesian techniques, Carlo Monte techniques, etc. No doubt, with this attempt to investigate the port sector of India several issues are in the open one can further analyze and come to desired conclusions.

References

Banker R.D., Charnes A., Cooper W.W. (1984), Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis, „Management Science”, vol. 30 no. 9, pp. 1078-1092.

Barros C.P. (2003), Incentive regulation and efficiency of Portuguese port authorities, „Maritime Economics and Logistics”, vol. 5 no. 1, pp. 55-69.

Caves D.W., Christensen L.R., Diewert W.E. (1982), The economic theory of index numbers and the measurement of input, output and productivity, „Econometrica”, vol. 50 no. 6, pp. 1393-1414.

Charnes A., Cooper W.W., Rhodes E. (1978), Measuring the efficiency of decision making units, „European Journal of Operational Research”, vol. 2 no. 6, pp. 429-444.

Chudasama K., Pandya K. (2008), Measuring efficiency of Indian ports, Working Paper 3, Department of Research Methodology and Interdisciplinary Studies in Social Sciences, Veer Narmad South Gujarat University, Surat.

Coelli T.J. (1996a), A guide to FRONTIER version 4.1. A computer program for stochastic frontier production and cost function estimation, „CEPA Working Papers”, no. 7, Department of Econometrics, University of New England, Armidale.

Coelli T.J. (1996b), A guide to DEAP version 2.1. A Data Envelopment Analysis (computer) program, „CEPA Working Papers”, no. 8, Department of Econometrics, University of New England, Armidale.

Cullinane K., Song D.W., Ji P., Wang T. (2004), An application of DEA Windows analysis to container port efficiency, „Review of Network Economics”, vol. 3 no. 2, pp. 184-206.

Cullinane K., Song D.W., Wang T. (2005), The application of mathematical programming approaches to estimating container port production efficiency, „Journal Productivity Analysis”, vol. 24 no. 1, pp. 73-92.

Fare R., Grosskopf S., Lovell C.A.K. (1994a), Production frontiers, Cambridge University Press, Cambridge.

Fare R., Grosskopf S., Zhang Z. (1994b), Productivity growth, technical progress, and efficiency change in industrialized countries, „American Economic Review”, vol. 84 no. 1, pp. 66-83.

Farrell M.J. (1957), The measurement of productive efficiency, „Journal of the Royal Statistical Society. Series A (General)”, vol. 120 no. 3, pp. 253-290.

Friedrichsen C. (1999), Benchmarking of ports. Possibilities for increased efficiency of ports. Transport benchmarking. Methodologies, applications and data needs, European Conference of Ministers of Transport, European Commission, Paris.

Government of India (GOI) (1999), Corporatisation. Ennore Port & Jawaharlal Nehru Port, The Cornell Report, The Cornell Group Inc., New Delhi.

GOI (2002), Vision 2020 transport. Report prepared for Planning Commission by Mahesh Kapoor, New Delhi.

GOI (2003), Seventh Lok Sabha Estimates Committee (1981-82, 32nd Report), New Delhi.

GOI (2004), Draft policy for maritime sector, Ministry of Shipping, New Delhi.

GOI (2005), Basic port statistics of India, Ministry of Shipping, Road Transport and Highways, New Delhi.

GOI (2006), National Maritime Development Programme, Ministry of Shipping, Road Transport and Highways, New Delhi.

GOI (2006), Towards faster and more inclusive growth rate. An approach to Eleventh Five Year Plan, Planning Commission, New Delhi.

GOI (2007a), Report of working group for port sector for the Eleventh Five Year Plan (2007-2012), Ministry of Shipping, Road Transport and Highways, New Delhi.

GOI (2007b), Eleventh Five Year Plan, 2007-2012, vol. I-III, Planning Commission, New Delhi.

GOI (2007c), Annual report 2006-07, Ministry of Shipping, New Delhi.

GOI (2008), Economic survey 2007-08, Ministry of Finance, New Delhi.

GOI (2009), Major Ports Regulatory Authority Act (MPRAA), Ministry of Shipping, New Delhi.

Indian Port Association (IPA) (2002), Major ports of India. A profile – 2001-02, IPA, New Delhi.

IPA (2007), Coordination of business plan for major ports in India. Consolidated port development plan. Report prepared by Port of Rotterdam Authority, IPA, New Delhi.

Micco A., Perez N. (2001), Maritime transport costs and port efficiency, prepared for the seminar An alternative approach to efficiency measurement of seaports, „Towards Competitiveness. The Institutional Path", Santiago [unpublished work].

Park R., De P. (2004), An alternative approach to efficiency measurement of seaports, CMaritime Economics & Logistics", vol. 6 no. 1, pp. 53-69.

Poitras G., Tongzon J., Li H. (1996), Measuring port efficiency. An application of Data Envelopment Analysis, Discussion Paper 10/15/96, National University of Singapore, Singapore.

Tongzon J. (1995), Determinants of port performance and efficiency, „Transportation Research. Part A: Policy and Practise", vol. 29 no. 3, pp. 245-252.

Tongzon J. (2001), Efficiency measurement of selected Australian and other international ports using data envelopment analysis, „Transportation Research. Part A: Policy and Practise", vol. 35 no. 2, pp.107-122.

Tongzon J., Ganesalingam S. (1994), An evaluation of ASEAN port performance and efficiency, „Asian Economic Journal", vol. 8 no. 31, pp. 317-330.

Trujillo L., Nombela G. (1999), Privatization and regulation of seaport industry, Policy Research Working Paper 2181, World Bank, Washington.

UNCTAD (1969), Development of ports, UNCTAD Secretariat, New York.

UNCTAD (1976), Port performance indicators, UNCTAD Secretariat, New York.

UNCTAD (1976), Port productivity, p. I-III, UNCTAD Secretariat, Geneva.

World Bank (1995), India port sector strategy report, Report no. 14059-IN, World Bank, Washington DC.

Internet sources

Centre for Monitoring Indian Economy (CMIE), https://www.cmie.com/ [09.09.2018].

CENTRAL EUROPEAN REVIEW
OF ECONOMICS AND MANAGEMENT

ISSN 2543-9472; eISSN 2544-0365

Vol. 2, No. 3, 175-196, September 2018

www.cerem-review.eu
www.ojs.wsb.wroclaw.pl

Solution of the Traveler's Dilemma

Gabriela OSIECKA, Maciej JASIŃSKI

WSB University in Wrocłąw, Poland

Abstract:

Aim: The aim of the article is to show: 1) that the reasoning of perfectly rational players presented in 1994 by the author of the Traveler's Dilemma Kaushik Basu is incorrect and therefore leads to wrong conclusions, 2) how the reasoning of these players should look like and what solution it leads to.

Design / Research methods: Logical analysis.

Conclusions / findings: Perfectly rational Traveler's Dilemma players should use, according to game theory, so-called retrograde (iterative) induction. This is wrong, as in the set of Traveler's Dilemma games results the principle of transitivity is not met. We believe that perfectly rational players will achieve a better result when they make a random decision from a suitably limited set of decisions. After applying this method of decision making, perfectly rational players will achieve a result similar to those obtained by real players in experiments. Thus, the paradox described in the theory of games disappears, that perfectly rational players achieve worse results than real players

Originality / value of the article: A new way of making decisions in the Traveler's Dilemma game.

Implications of the research: A new way of making decisions in other games similar to the Traveler's Dilemma may allow to find new solutions in these games.

Limitations of the research: The described decision-making method can potentially be used in decision-making situations when the following five conditions are met: 1) the set of possible decisions of each player is greater than 2, 2) the winning matrix is known to both players and both know the purpose of their choices, 3) when it is played once with an unknown opponent, 4) when both players have to make their decision without knowing the opponent's choice, 5) when there is no decision, which is a stable balance point or when it is, but its choice means that the player does not achieve a satisfying result.

Key words: game theory, Traveler's dilemma, perfectly rational player, backward induction.

JEL: C70

Correspondence address: Maciej Jasiński, WSB University in Wrocław, ul. Fabryczna 29-31, 53-609 Wrocław, Poland. E-mail: maciej.jasinski@wsb.wroclaw.pl
Received: 19.06.2017, Revised: 05.04.2018, Revised:10.09.2018, Accepted: 10.09.2018
doi: http://dx.doi.org/10.29015/cerem.502

1. Introduction

In 1994 Kaushik Basu devised The Traveler's Dilemma (TD) (Kaushik Basu 1994: 391-395). It illustrates a situation when two travelers loose identical exotic souvenirs when returning from the holidays. Feeling guilty the airline wants to cover the loss. However, they do not know how much the souvenirs might have cost and suggest the following solution. Both travelers will, independently of each other, write down the value between 2 and 100 dollars on a piece of paper. If they both put down the same number then both of them will receive it. If the numbers are different, the traveler who wrote a lower number will get the amount he put down increased by a bonus of 2 dollars and the other will receive the amount decreased by a penalty of 2 dollars. If the first one writes 100 and the other 48, the first one will be paid 46 dollars (48-2=46) and the other will get 50 (48 +2= 50).

This paper addresses the problem to the players who:

1. are perfectly rational in striving for the best possible result for themselves,
2. both know that they are playing against with a perfectly rational player,
3. play the game only once with a particular player.

According to current game theory, rationally behaving players under these circumstances should write 2 dollars. This results from the following reasoning. If the other traveler puts down 100 on his piece of paper, it is best for the first one to write down 99 because then he will get 101 dollars. If the other one arrives at the identical conclusion and writes 99 instead 100 then the first players should write 98 as he will receive 100 dollars which is the best result provided the competitor put 99 on his piece of paper. If one player assumes that the other decides in favor of 98, it is best for him to bet on 97. If the other player makes up his mind for 97 then automatically 96 becomes the best for the first player.

The so-called backward induction makes both players write down 2 dollars each. This solution is the Nash equilibrium as it does not pay for the player to change his decision so both of them will stand by their decisions. This is called an equilibrium point.

2. Why cannot backward induction be applied in the Traveler's Dilemma?

Backward induction leads to correct conclusions when in a decision set of the game the rule of transitivity is met. Transitivity guarantees that when we compare decision A with B and we see that decision A gives a worse result than B (A is worse than B) then we compare decisions B and C stating that B is worse than C, then we may be sure that when we compare A and C we will arrive at the conclusion that A is always worse than C.

Backward induction in TD makes use of the transitivity rule. We should see it clearly when we describe the way of reasoning of both rational players. If the first player compares the choices of 100 and 99, regardless of the other player's decision, he will conclude that 99 is better than 100 for him as it always gives the same or better result – see Table 1.

Table 1. Matrix fragment of the first player's winnings

		Decisions of player 2						
		100	99	98	97	96	95	from 2 to 94
Decisions of player 1	100	100	97	96	95	94	93	from 0 to 92
	99	101	99	96	95	94	93	from 0 to 92
	98	100	100	98	95	94	93	from 0 to 92
	97	99	99	99	97	94	93	from 0 to 92
	96	98	98	98	98	96	93	from 0 to 92
	95	97	97	97	97	97	95	from 0 to 92

Source: Authors' own elaboration.
The numbers show the result to be achieved by the first player when the other one makes a specific decision.

If the first player rejects 100 and assumes that the other player will do the same then when comparing 99 and 98 he will conclude that 98 is better than 99 as no matter what his competitor will do, he will have the same or better results betting on 98 rather than 99. If we take 99 and 100 out of the decision set and compare 98 and 97 the we may say that for the first player 97 is the best which makes us reject 98 and so on. By doing so we arrive at the conclusion that the best decisions' set consists only of one element namely 2. This reasoning assumes that if 100 is worse than 99 and 99 is worse than 98 then we can remove 99 and 100 from the set of potentially best decisions so none of the players will get back to a rejected decision in a particular moment. If this reasoning is to be logical in TD, the transitivity rule must be met.

Let us check if this is the way in TD. When we were previously comparing two neighboring decisions we concluded that 100 was always worse than 99 then 99 was worse than 98, 98 was worse than 97[1]. Thus, we should arrive at the conclusion that 100 must always be worse than 97. Looking at Table 1 we notice that this is not the way it is. 100 is worse than 97 provided the other player does not make up his mind for 100. If we assume that 100 is not the best decision and that is why it will not be chosen then indeed 100 is always worse than the decision about 97. Without making such an assumption we cannot say that 100 is worse than 97. Such reasoning is logically flawed.

The above reasoning proves that in the set of TD decisions the transitivity rule is not complied with so backward induction assuming occurrence of the transitivity rule cannot be applied here.

Conclusion 1: Perfectly rational players in TD cannot apply backward induction as this thinking procedure is wrong. It results from the fact that in the results' set the transitivity rule is not met.

We already know how a perfectly rational player should not reason but we do not know how he should think.

3. How do perfectly rational players reason in TD?

As perfectly rational players share knowledge and are also aware that the game is played once, initial simple conclusions may be drawn.

---

[1] Looking at Table 1 we can see that 97 is better than 98 only in two cases namely when the other player chooses 98 or 97. In remaining 97 cases the decision about 98 is better or the same as 97.

Conclusion 2. There is at least one way of deciding in each game. A random selection is always available. We define it as one in which the probability of choosing each decision from a given set is the same. Thus, we may say that if in TD decisions are made non-randomly, we have at least two ways of making a decision.

The question arises: in what way do perfectly rational players choose a method of arriving at a decision? It does not concern the choice of a decision but the way of establishing which decision is the best for the player. The answer is simple. They compare alternative ways of making a decision and select the one which secures achievement of the best result in a particular game.

The fact that rational players will compare various ways of making a decision with a random selection in TD limits the area of selecting a decision. To understand why, let us first establish what result may be achieved in TD when both players randomly choose a decision out of 2-100 range. If both players randomly choose their decision, the probability of drawing each number from a given set must be the same. Then the effects of such a decision are best shown by the expected value of winning calculated with the same probability of choosing each number from the set of all possible decisions. Thus, we apply the Laplace criterion.

If both players randomly choose from the set of 2-100, the value of the expected winning will go up to 34.50168[2]. The expected value of winning does not guarantee that each player achieves such a result in a particular game but it informs us that if such a player played TD infinitely many times then the average winning in a single game would amount to 34.50168. When making a decision randomly the anticipated value is the best measure allowing us to choose the best solution.

This also means that it does not pay to make a decision non-randomly in favor of 32 or lower because then the maximum winning is 34 which is lower than in case of a random decision. If the two players arrive at such a conclusion, they limit their random choice to the range 33 to 100. Then the anticipated value of the winning must be higher than 34.50168. This in turn once again limits the set of permissible decision taken randomly.

---

[2] Calculations of the results are shown in appendix 1.

In brief, let us assume that previous reasoning made the players conclude that it did not pay to choose decisions lower than 90. Then their further reasoning will go according to the pattern shown in Table 2.

Table 2. Limiting the selection of decisions made randomly

Decision set	Anticipated value of the winnings based random selection from a particular set	Decisions which may be eliminated from the set
from 90 to 100	93.181818	90 and 91
from 92 to 100	94.518519	92
from 93 to 100	95.1875	93
from 94 to 100	95.85714	-

Source: Authors' own elaboration.

Let us assume that we have arrived at the conclusion that the selection must be limited to the set (90; … 100). The anticipated value of winnings based on random selection from this set amounts to 93.181818[3]. At the same time, players will figure out that it does not pay to make non-random decisions of 91 and lower as maximum winnings for 91 amounts to 93 which is less than with a random selection of the analyzed set. Thus, the selection area is limited to the set (92; … 100). The anticipated value of the winnings from this set of random decisions amounts to 94.51 which means that 92 may be removed from the set of acceptable decisions. These decisions allow the players to achieve 94 maximum which is less than in case of random selection. When the players limit their selection to the set (93; … 100) then on average they will achieve the result of 95.18 choosing randomly. Similarly, this will eliminate 93. It turns out that this is the last decision which may be eliminated thanks to applied reasoning. This is because the anticipated value of the winnings amounts to 95.85714 when selecting randomly from the set (94; … 100). This result does not let us go on with limiting decisions as this time non-random selection of 94 may result in 96 which is better than the anticipated value of the winnings from the set (94; … 100).

The above reasoning can be called: "Iterated Laplace Rationality by successively reducing the set of actions". We define them as follows:

[3] Calculations of the value are shown in the appendix.

a) Let the set of actions available to each player in a symmetric game be S_0 and set $i=0$.

b) At stage i calculate the expected reward of each player when both players pick each action from S_i with the same probability, call this expected payoff E_i.

c) S_{i+1} is obtained from S_i by removing all the actions which always give a payoff of less than E_i when the set of actions available is S_i.

d) If S_{i+1} is a strictly smaller set than S_i, then increase i by one and return to step b), otherwise S_i is the set of "Laplacian rational actions by successive reduction".

The above reasoning leads us to the following conclusion.

Conclusion 3. Regardless of the way the decision is made, both players will not go out of set (94; ...; 100). This is a stable point of equilibrium in a sense that if one of the players considers that it is best to limit his decision to set (94; ...; 100) and assumes that his opponent arrives at the same conclusion, it will not still pay to go out of this set and make a decision lower than 94. Thus, we can say that the limitation of the set of decisions to (94: ...: 100) is the Nash equilibrium point.

We already have a narrowed selection field of perfectly rational players and we still do not know in what way they are going to choose the best decision. For the purposes of narrowing the search area let us conduct the following reasoning using the proof by negation.

Let us assume that in TD there is one best non-random method of making a decision and it gives a better result than 95.85714[4]. Then we may be sure that both rational players will reject random selection and apply a non-random method. If both players apply the same method of making a decision, they will have to reach the same decisions. Symmetric pairs from (94; 94)[5] to (100; 100) will become a solution to the game for both players. If any decision of one player meets an identical one of the other one, it does not become his best decision as it is enough a particular player lowers his number by 1 and achieves a better result than

[4] This is an anticipated value of the winnings based on random selection made by both players out of 94 to 100 set.

[5] According to conclusion 3 we know that both players limit their selection by omitting 93 and lower numbers.

previously. Thus we arrived at the contradiction to the assumption we made. This proves that there is no one non-random way of making a decision which guarantees achievement of the best result lower than 95.85714.

The question arises what would happen if there was no best method of making a decision but there would be two or more of them. If there were two non-random ways of taking a decision and if they were equal they would have to lead to achievement of equally good results so both players would be made to arrive at the same decisions. However this is in contradiction to the statement that such a symmetric decision is the best one for a particular player. Thus we put forward the next conclusion.

Conclusion 4. In TD, perfectly rational players who have common knowledge when playing a single game will have to choose randomly from a set consisting of a maximum of 94 to 100 numbers.

Random selection in TD may denote two apparently different actions:

1. a player randomly chooses his decision out of a specific set of possibilities;
2. a player randomly assumes that his opponent will choose a particular decision and will non-randomly adjust his knowledge.

Both actions must be considered random though in the second one there is a random element which decides about the choice. It will be explained in detail in the further part of the paper when we will simulate reasoning of both players.

The fact that both players must randomly select a number from a given set means that we can use the Laplace criterion to evaluate these sets, because the probability of choosing each number from a given set is the same then.

As we already know that both players will have to make a random decision, let us ask a question whether we can narrow down the set (94; … 100) because TD has a rule that the higher the decisions, the higher anticipated value of the winnings. They reach the maximum for the set (99; 100). In random selection the anticipated value of winnings amounts to 99.25. Let us check if perfectly rational players will limit their choice to this set. Such a decision situation is presented in Table 3.

Table 3. Matrix of first player's winnings when both players limit their random choice to 99 and 100.

1st player's decision	2nd player's decision		
	random selection out of (100;99)	100	99
random selection out of (100;99)	99,25	100,5	98
100	98,5	100	97
99	100	101	99

Source: Authors' own elaboration.

Table 3 presents all possible decisions to be taken in the set (99;100). A player may randomly choose a decision out of this set or randomly assume that the opponent will make a specific decision and then non-randomly adjust his best decision. The analysis of the case presented in Table 3 indicates that regardless of the decision his opponent will make, it pays for the first player to non-randomly choose decision for 99. If the first player arrives at such conclusion then his opponent does the same by choosing 99 and his prior choice stops being the best as a he will achieve a better result when he responds with 98.

The assumption that the players will limit their choice to the set (99; 100) must be ruled out which means that both players are certain to extend their decision set to minimum (98; 99; 100).

At this point a doubt may arise why the players do not exclude 100 and 99 which turned out worse than 98 in the above reasoning. The answer is as follows. Those decisions could be excluded permanently if the result set in TD comprised the transitivity rule. However, it is not complied with so we must take into account 100 and 99 as they may turn out better in next situations allowing us to make decisions other than those in (99; 100) range.

Let us examine the situation when both players assume that the opponent selects a decision out of (100;99;98) set. Under these circumstances the player may randomly choose out of (100;99), (99;98), (100;98) or (100;99;98). The anticipated value of winnings when choosing randomly from the first set is the biggest [6] so in Table 4 we will only present this case and we will omit the other two as it does not affect the conclusions made. Apart from random selection we also consider the

[6] See appendix 2.

situation that each player may select each decision from the set. Table 4 presents a matrix of the winnings for the first player.

Table 4. Matrix of the first player's winnings when both players limit their choice to (98; 99;100) set.

1st player's decision	2nd player's decision			
	random selection out of (100;99)	100	99	98
random selection out of (100;99)	99,25	100,5	98	96
100	99	98	97	96
99	100	101	99	96
98	100	100	100	98

Source: Authors' own elaboration.

This time there is no best decision in this decision set for the first and the second player, just like it was in the first case, 98 is the best for the first player except for the situation when the second player bets on 100. The first player is unable to define which decision is best for him without assuming the other player's choice. This means that both players must make a random decision. This in turn denotes that one cannot define whether deciding in favor of 97 – going beyond the analyzed set- is profitable or not. Thus we cannot explicitly define if the players will be willing to extend the set up to (97; 98; 99; 100).

To make sure let us examine the matrix for (97; 98; 99; 100) set. Table 5 shows results of calculation. According to this table it pays the player to non-randomly decide on 97 only if the other one bet on 97 or 98. In remaining three cases 97 is not the best choice for the first player. As long as the first player is unable to define his best decision the other one does not know it either. This way we are certain that one cannot explicitly say whether it pays off to add 97 to (98; 99; 100) set.

Generalizing the aforementioned reasoning we can call the "Iterated Laplace Rationality by successively increasing the set of activities". This procedure can be defined as follows:

a) Let the set of actions corresponding to Pareto optimal payoff vectors in a symmetric game be S_0 and set i=0.

b) At stage *i* calculate the expected reward of each player when both players pick each action from S_i with the same probability, call this expected payoff E_i.

c) S_{i+1} are obtained from S_i by adding any action which gives a payoff of greater than E_i against some action from S_i.

d) If S_{i+1} is a strictly larger set than S_i, then increase *i* by one and return to step b), otherwise S_i is the set of "Laplacian rational actions by successive addition".

Table 5. Matrix of the first player's winnings when both players limit their choice to (97; 98; 99;100) set.

1st player's decision	**2nd player's decision**				
	random selection out of (100;99)	100	99	98	97
random selection out of (100;99)	99.25	100,5	98	96	95
100	98.5	100	97	96	95
99	100	101	99	96	95
98	100	100	100	98	95
97	99	99	99	99	97

Source: Authors' own elaboration.

Let us summarize our considerations about decisions which perfectly rational players will make.

Conclusion 5. We can clearly define that the decision set which the players are going to use to make a decision will randomly at maximum consist of (94;…;100) and minimum (98;…;100).

The reasoning we applied allowed use to significantly narrow the selection field of both players which does not mean that perfectly rational players will not try to narrow it down as the value of expected winnings goes up. However, there must be an additional criterion which will allow them to select from the indicated set of decisions and at the same time it will be a stable point of equilibrium. We did not find such a criterion, therefore the conclusion 5 is the last in the reasoning of perfectly rational players.

To sum up the main conclusions one may state as follows. If TD is once played by perfectly rational players having common knowledge then:

1. They cannot use backward induction in their reasoning as the winnings set does not comply with the transitivity rule.

2. They must make a random decision which in this case means that

a. they will randomly choose from a particular decision set or

b. they will randomly assume a decision to be taken by the other player and select the best decision for themselves.

3. We can clearly say that the choice will be limited to (94;…;100) set at a minimum and (98;…;100) set at a maximum

The decision is often defined as follows: "a conscious, non-random choice of one of the options of future action recognized and recognized as possible" (Koźmiński 2002: 85). Using this definition, we can say that the decision of perfectly rational players in TD will be: random selection of one of the numbers from the set not greater than (94; ...; 100) and not less than (98; ...: 100). In this case, the decision that I randomly choose a number from a given set is a conscious non-random choice of the player. Such a decision is a stable point of equilibrium (Nash equilibrium).

The above definition of the word decision, unfortunately, does not match the TD case, because the player cannot write on the card that randomly selects one of the numbers from the given set. Must enter a specific number. So if we cancel the previous definition of the term decision and specify that in TD the decision is the number entered on the sheet, then the balance ceases to be stable. No matter what number both players put on their cards from a given set, at least one of them always comes to the conclusion that a better result will be achieved when it changes its original decision.

If the player's decision in TD is the number he puts on the card, then there is only one pair of decisions in this game, which we will define as a stable equilibrium point. It is a pair (2; 2). Only then, each player knowing that the competitor entered number 2 on the card will conclude that the best choice for him is to enter on his sheet 2. Both players will then get the result 2. A perfectly rational player, however, has no imposed limitation that he must reach a stable equilibrium point. It has to achieve the best result. If an unstable game solution gives him a better score than 2, he must choose an unstable solution. The conscious decision of both players that each of them will randomly choose the number from the previously indicated set guarantees them that in the worst case they will achieve the result of 92. So, if players are perfectly rational, they must choose this decision.

The reasoning presented is that first the player limits the maximum set of best decisions so that it is a stable equilibrium point (Nash equilibrium) and then randomly chooses one decision from this set, it is most likely only meaningful if the minimum five conditions are met:

1. the set of possible decisions of each player is greater than 2,
2. the winning matrix is known to both players and both know the purpose of their choices,
3. when it is fought once with an unknown opponent,
4. when both players have to make their decision without knowing the opponent's choice,
5. when there is no decision, which is a stable balance point or if it is but its choice means that the player does not achieve the best result.

After fulfilling these conditions, it may turn out that the reasoning described in this article will make perfectly rational players make a decision that will give better results than previously indicated as the best. To illustrate the above theorem with examples deserves a separate article.

The TD solution presented here is not an example of limited rationality, i.e. when we choose the first decision that meets our minimum requirements for winning. The presented reasoning of perfectly rational players seeks to choose the optimal decision, which is, maximizing the player's winnings. In the course of the reasoning presented, none of the players assumed that he would choose a decision that would guarantee him a certain minimum score.

4. Perfectly rational players' decisions and real players' decisions

A lot of experiments have been carried out into TD. Most people chose decisions near 100 (Basu 2007: 75). According to a binding game theory 2 was a rational decision. Thus, the following conclusions were made:

1. Most people behave irrationally in TD.

2. Irrational behavior may produce better results than rational behavior (Basu 2007: 72-74).
3. In TD people are not guided by an egoistic desire to maximize their results but tend to cooperate and prefer solutions providing benefits to both players (Basu 2007: 74).

We have tried in this paper to prove that perfectly rational players will tend to choose decisions near 100 which must be considered rational. Thus, all three theses above become invalid.

Comparing perfectly rational players' decisions with real people's decisions we may formulate a new additional conclusion.

Conclusion 6: Under specific circumstances even people not knowing the method of rational decision are able to make one intuitively but it does not refer to spontaneous decisions but those made after consideration. In other words, in some cases people can sense what the best decision is and cannot logically explain why they consider it the best.

Validity of this thesis is proven by all experiments which were carried out before publishing this paper and in which TD players chose decisions near 100. The described way of reasoning in that the player as a result of the analysis first limits the scope of the best decisions to the smallest set and then cannot unambiguously determine which solution from this set is the best random selection seems to be quite commonly used by real people. This would explain why in real-world experiments players typed numbers close to 100 on cards.

Appendix 1

Let us agree on the amount of the winnings of the player when both players randomly make a decision out of the range 2 to 100. To follow the reasoning in a better way let us begin with a presentation of an abridged matrix of the first player's winnings – see Table 7.

Table 7. Matrix of the first player's winnings when both players randomly make a decision out of 2 to 100 range.

Decisions of player 1	Decisions of player 2									
	2	3	4	5	6	7	…	98	99	100
2	2	4	4	4	4	4		4	4	4
3	0	3	5	5	5	5		5	5	5
4	0	1	4	6	6	6		6	6	6
5	0	1	2	5	7	7		7	7	7
6	0	1	2	3	6	8		8	8	8
…										
97	0	1	2	3	4	5		99	99	99
98	0	1	2	3	4	5		98	100	100
99	0	1	2	3	4	5		96	99	101
100	0	1	2	3	4	5		96	97	100

Source: Authors' own elaboration.

Each player may make 99 decisions so the probability of obtaining a particular result at a particular decision amount to 1/99.

In the set of decisions of the first player we may observe occurrence of definite regularities depending on the decision taken by the opponent. Three cases may appear at maximum[7]:

1. if the opponent chooses a decision higher than the decision of the first player,
2. if the opponent chooses exactly the same decision as the first player,
3. if the opponent chooses a lower decision than the first player.

[7] 2 and 100 as extreme decisions are an exception. Only two cases appear then. For decision 2 there will be cases 1 and 2 and for decision 100 there will be 2 and 3.

In the first case player 1 will always attain the same result which equals k+2 where k stands for the decision of the first player. This result will repeat 100-k times. The repeating results are marked with the darkest gray in Table 7. When the first player decides on k=2, the repeating result will amount to k+2=4 and repeats itself 98 times as out of 99 decisions of the other player only one decision will not make the first player achieve the result of 4. For 98, the repeating result of the first player is k+2=100 and it will repeat 100-k which is twice.

In the second case the first player will always attain result k.

In the third situation the first player will achieve a result lower by 2 than the opponent's decision. These results are marked white in Table 7. Looking at the results of the first player we see that they constitute subsequent natural numbers in an arithmetic sequence beginning with zero and we always reach a result lower than decision k by 3. This sequence consists of k-3 elements (excluding 0 which does not affect the total of the sequence). When using a pattern for the total of an arithmetic sequence where:

$$S_n = \frac{(a_1 + a_n)n}{2} \tag{1}$$

$a_1 = 1$, a_n = k-3 and n = k-3[8] and after applying these values to pattern 1 we obtain:

$$S_n = \frac{(1+(k-3))(k-3)}{2}$$

Using the above patterns and previous considerations about the results which the first player will achieve in the first two cases, we may elaborate a pattern for the anticipated value of the first players' winnings when he makes decision k in TD:

$$E(k) = \frac{1}{99}\left(\frac{(1+(k-3))(k-3)}{2} + k + (100-k)(k+2)\right)$$

After transformations we arrive at the following:

$$E(k) = \frac{1}{99}\left(-0{,}5k^2 + 96{,}5k + 203\right). \tag{2}$$

---

[8] In this case a question may arise with regard to the field of number k. It turns out that we do not have to remove 2 (the lowest decision in TD) as after applying k=2 we receive the sum of 0 and it should be like that. The same sum will appear for k=3.

Applying pattern 2 we may calculate the amount of the anticipated value of the first player's winnings when the two players together randomly choose their decisions out of 2 to 100 set. This value will be marked E(2;…;100). As the player may make 99 decisions then:

$$E(2;\ldots;100) = \frac{1}{99}\left(E(2) + E(3) + \ldots + E(100)\right)$$

When applying patterns 2 to the above formula and after ordering it we obtain the following:

$$E(2;\ldots;100) = \frac{1}{99}\frac{1}{99}\left(-0{,}5\left(2^2+3^2+\ldots+100^2\right)+96{,}5(2+3+\ldots+100)+99\bullet 203\right) \quad (3)$$

The pattern for the sum of an arithmetic sequence consisting of subsequent natural numbers from 1 and squared is as follows:

$$S_n^2 = 1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

As we are interested in the sum comprising squares of subsequent natural numbers excluding one, we modify the above pattern by excluding 1 . After applying 100 to n we obtain:

$$S_n^2 = 1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

Using pattern 1 we may calculate the sum of the series of 2 to 100 numbers.

$$S_n = \frac{(a_1 + a_n)n}{2} = \frac{(2+100)99}{2} = 5049$$

Making use of two last numbers calculated and putting them into pattern 3 we may write down:

$$E(2;\ldots;100) = \frac{1}{9801}(-0{,}5 \times 338349 + 96{,}5 \times 5049 + 99 \times 203)$$
$$= 34{,}50168350168$$

This way we have calculated the anticipated value of the player's winnings when he together with the opponent randomly chooses his decision out of 2 to 100 set.

Appendix 2

If we are interested in the anticipated value of the winnings based on set 90 to 100 it is easier to present calculations in tables instead of transforming patterns elaborated in appendix 1.

Table 8. Matrix of the first player's winnings when both players randomly make a decision out of 90 to 100 range.

		Decisions of player 2											Anticipated value of the first player's winnings
		90	91	92	93	94	95	96	97	98	99	100	
Decisions of player 1	90	90	92	92	92	92	92	92	92	92	92	92	91.81818182
	91	88	91	93	93	93	93	93	93	93	93	93	92.36363636
	92	88	89	92	94	94	94	94	94	94	94	94	92.81818182
	93	88	89	90	93	95	95	95	95	95	95	95	93.18181818
	94	88	89	90	91	94	96	96	96	96	96	96	93.45454545
	95	88	89	90	91	92	95	97	97	97	97	97	93.63636364
	96	88	89	90	91	92	93	96	98	98	98	98	93.72727273
	97	88	89	90	91	92	93	94	97	99	99	99	93.72727273
	98	88	89	90	91	92	93	94	95	98	100	100	93.63636364
	99	88	89	90	91	92	93	94	95	96	99	101	93.45454545
	100	88	89	90	91	92	93	94	95	96	97	100	93.18181818
Anticipated value of the first player's winnings based on 90 to100 set													**93.18181818**

Source: Authors' own elaboration.

As you can see from Table 8, as a random selection produces an average result of 93.1818, it does not pay for the player to make non-random decision 90 and 91. In both cases the maximum winnings are lower that the anticipated value of winnings so a random selection from (90; …;100) produces a better result than a non-random selection of both decisions. Therefore both players will eliminate 90 and 91 out of the decision set.

Table 9. Matrix of the first player's winnings when both players randomly make a decision out of 92 to 100 range.

		Decisions of player 2									Anticipated value of the first player's winnings
		92	93	94	95	96	97	98	99	100	
Decisions of player 1	92	92	94	94	94	94	94	94	94	94	93.77777778
	93	90	93	95	95	95	95	95	95	95	94.22222222
	94	90	91	94	96	96	96	96	96	96	94.55555556
	95	90	91	92	95	97	97	97	97	97	94.77777778
	96	90	91	92	93	96	98	98	98	98	94.88888889
	97	90	91	92	93	94	97	99	99	99	94.88888889
	98	90	91	92	93	94	95	98	100	100	94.77777778
	99	90	91	92	93	94	95	96	99	101	94.55555556
	100	90	91	92	93	94	95	96	97	100	94.22222222
Anticipated value of the first player's winnings based on 90 to100 set											**94.51851852**

Source: Authors' own elaboration.

We can see that it does not pay the first player to non-randomly decide about 92 as the maximum winnings in this case are lower than the anticipated value with the selection from 92 to 100. That is why both players will eliminate the decision out of this set.

Table 10. Matrix of the first player's winnings when both players randomly make a decision out of 93 to 100 range.

		Decisions of player 2								Anticipated value of the first player's winnings
		93	94	95	96	97	98	99	100	
Decisions of player 1	93	93	95	95	95	95	95	95	95	94.75
	94	91	94	96	96	96	96	96	96	95.125
	95	91	92	95	97	97	97	97	97	95.375
	96	91	92	93	96	98	98	98	98	95.5
	97	91	92	93	94	97	99	99	99	95.5
	98	91	92	93	94	95	98	100	100	95.375
	99	91	92	93	94	95	96	99	101	95.125
	100	91	92	93	94	95	96	97	100	94.75
Anticipated value of the first player's winnings based on 90 to100 set										**95.1875**

Source: Authors' own elaboration.

As with random selection from (93;…;100) set the anticipated value of winnings amounts to 95.1875 it does not pay any of players to non-randomly choose 93 as the maximum winnings comes up to 95. That is why the decision about 93 will be removed from the decisions set.

Table 11. Matrix of the first player's winnings when both players randomly make a decision out of 94 to 100 range.

		Decisions of player 2							Anticipated value of the first player's winnings
		94	95	96	97	98	99	100	
Decisions of player 1	94	94	96	96	96	96	96	96	95.71428571
	95	92	95	97	97	97	97	97	96
	96	92	93	96	98	98	98	98	96.14285714
	97	92	93	94	97	99	99	99	96.14285714
	98	92	93	94	95	98	100	100	96
	99	92	93	94	95	96	99	101	95.71428571
	100	92	93	94	95	96	97	100	95.28571429
Anticipated value of the first player's winnings based on 90 to100 set									**95.85714286**

Source: Authors' own elaboration.

In this case we cannot eliminate 94 as the maximum winnings amounts to 96 which is more than the anticipated value when the decision is randomly made out if (94;…;100) set.

Appendix 3

In order to show that random selection is the best out of (100; 99) set as compared to random section out of (99; 98) and (100; 99; 98) we compare results presented in tables 12, 13 and 14.

Table 12. First player's winnings based on (100; 99) set

Decision of player 1	Decision of player 2	Winnings of player 1	Average winnings for player 1 based on his decision	Average winnings for player 1 based on all his decision
100	100	100	98.5	99.25
	99	97		
99	100	101	100	
	99	99		

Source: Authors' own elaboration.

Table 13. First player's winnings based on (100; 99; 98) set.

Decision of player 1	Decision of player 2	Winnings of player 1	Average winnings for player 1 based on his decision	Average winnings for player 1 based on all his decision
100	100	100	97.66666667	98.55555556
	99	97		
	98	96		
99	100	101	98.66666667	
	99	99		
	98	96		
98	100	100	99.33333333	
	99	100		
	98	98		

Source: Authors' own elaboration.

When comparing the results of the first player we may state that provided the player decides to randomly choose a decision from a particular set, he will obtain the best results when he limits hid choice to (100; 99) set.

Table 14. First player's winnings based on (98; 99) set.

Decision of player 1	Decision of player 2	Winnings of player 1	Average winnings for player 1 based on his decision	Average winnings for player 1 based on all his decision
99	99	99	98	98.375
	98	96		
98	99	100	98.75	
	98	98		

Source: Authors' own elaboration.

References

Basu K. On the Nonexistence of Rationality Definition for Extensive Games. "International Journal of Game Theory" 1990, vol 19. 33-44

Basu K. The Traveler Dilemma: Paradoxes of Rationality in Game Theory. "American Economic Review". 1994, 84(2), 391-395.

Basu K. Dylemat podróżnika. "Świat Nauki" (polish edition "Scientific American") VII 2007, 7(191). 72-77.

Becker T.,Carter M., Naeve J. Experts Playing the Traveler's Dilemma. Working Paper 2005, nr 252, Institute for Economics, Hohenheim University. ftp://ftp.repec.org/opt/ReDIF/RePEc/hoh/papers/252.pdf [13.09.2016]

Capra M., Goeree J., Gomez R., and Holt C.. Anomalous Behavior in a Traveler's Dilemma? "American Economic Review" 1999 89(3), 678-690.

Nash, J. F. Non-Cooperative Games. „Ann. Math." 54, 1951, 286-295,.

Rubinstein A. „Instinctive and Cognitive Reasoning: A Study of Response Times." http://www.arielrubinstein.tau.ac.il/papers/Response.pdf [10.07.2016]

Wolpert, D, Jamison J., Newth D., Harre M. Strategic Choice of Preferences: The Persona Model. "Santa Fe Institute Working Paper" 2011. http://www.santafe.edu/media/workingpapers/11-08-030.pdf [15.10.2016]

Zarządzanie. Teoria i praktyka - edited by: A. K. Koźmiński, W. Piotrowski; PWN Warszawa 2002