Environmental and Technical Efficiency in Large Gold Mines in Developing Countries

Magambo, Isaiah and Dikgang, Johane and Gelo, Dambala and Tregenna, Fiona

University of Johannesburg

May 2021

Environmental and Technical Efficiency in Large Gold Mines in Developing Countries

Isaiah Hubert Magambo[1], Johane Dikgang, Dambala Gelo[2]
and Fiona Tregenna[3]

[1] Public and Environmental Economics Research Centre (PEERC), School of Economics, University of Johannesburg, Johannesburg, South Africa. Email: isaiah.mgy@gmail.com
[2] School of Economics and Finance, The University of the Witwatersrand, Johannesburg, South Africa.
[3] DST/NRF South African Research Chair (SARChI Chair) in Industrial Development, University of Johannesburg, Johannesburg, South Africa.

The authors gratefully acknowledge the financial support from the University Research Committee (URC) at the University of Johannesburg, National Research Foundation (NRF) though the DST/NRF South African Research Chair in Industrial Development, and the Public and Environmental Economics Research Centre (PEERC) at the University of Johannesburg.

Abstract

Given the increasing importance of the mining sector in developing countries, an understanding of their level of environmental efficiency is useful, both to the industry itself and to policymakers. Environmental problems introduced by the sector are attracting extensive attention, so comprehensive analysis of their environmental performance has become increasingly important. This study evaluates the environmental performance of large gold-mining operations by applying a by-production model that specifies emission-generating technology, while incorporating a four-way error approach that captures mine-size heterogeneity, transient and persistent technical efficiency, and random errors. We applied a true random effects model (TREM), and a simulated maximum likelihood estimator (SMLE) based on the generalised true random effects model (GTREM). The former approach was estimated as a benchmark, while the latter was employed to estimate a four-component panel data stochastic frontier model. The four-components estimate separates firm heterogeneity from persistent and time-varying inefficiencies, thus providing more robust efficiency estimates and policy insights. Firm-level data from 2009 to 2018 were used; the results show the presence of environmental and technical inefficiencies. Moreover, each inefficiency was decomposed into transient and persistent inefficiencies. The GTREM predicts the average inefficiency to amount to 34% environmental (interaction between 19% transient and 18% persistent) and 30% technical (interaction between 4.4% transient and 27% persistent). The transient component of technical efficiency does not change over time, which may imply that the mines' managerial approaches are static. The presence of technical inefficiency implies that more than the minimal amounts of inputs are used to produce a given level of desirable output, which could be due to moral hazards and asymmetric information such as principal-agent problems. The presence of (environmental) inefficiency in the by-production model means that more than a minimal amount of the undesirable output is produced. The overall environmental performance of the mines in developing countries is low (66%) compared to other sectors, which indicates that there could be structural rigidities, poor environmental policies and regulations, poor enforcement, or any combination of the three. We also found robust empirical evidence that between 2009 and 2018, on average, gold-mining firms neither strongly increased nor strongly decreased their transient or their persistent technical and environmental efficiency. Besides, firms with high technical efficiency simultaneously have high environmental efficiency, which suggests that promoting high environmental efficiency will also promote high technical efficiency.

Keywords: environmental efficiency, technical efficiency, persistent and transient efficiency, gold mine.
JEL Codes: D24, D25, Q55, Q58.

1. Introduction

The gold-mining industry is one of the main pillars in the economy of developing countries, and is often credited for making a significant contribution to economic growth. Over the past decades, improved living standards have largely been attributable to the mining boom (Kumah, 2006; Bjorn, 2000). The rapid expansion of the gold industry has had the potential to confer many benefits, especially for the people of the developing countries, by providing employment and foreign exchange. However, several challenges confront the gold mining industry (World Gold Council,2019; Bainton and Holcombe, 2018; Carvalho, 2017). Generally, it has a bad reputation, due to its highly polluting nature and the fact that its costs are usually externalised on the local communities that host its operations (Kumah, 2006; Rashidi and Saen, 2015). Across the developing world, gold mining contributes to acid mine drainage, water, air and soil contamination, loss of biodiversity, ecosystem damage and deforestation; all of which may result in health issues and lost productivity in local communities.

The key challenge is the trend of the increasing complexity of mining gold deposits, with decreasing gold grades and prices. Low-grade deposits (which implicitly means more resource inputs for lower output over time) are economically marginal, as the capital and operating cost requirements are relatively high, and the metal recovery volumes are sub-optimal (Neingo and Tholana, 2016; Pimentel, Gonzalez and Babosa, 2016, Shafiee and Topal, 2010). Alongside the context of challenging competition and a shifting social contract between business, government and civil society, critical social and environmental issues have increasingly become part of the mining-industry landscape (Bainton and Holcombe, 2018; Carvalho, 2017). The emphasis on addressing various environmental issues, including CO_2 emission and energy and water consumption, has recently become more pronounced due to concerns about climate change and drought (Gorain, Kondos and Lakshmanan, 2016; Smith and John, 2010).

The recent downturn in commodity prices and increase in environmental scrutiny has raised concerns about the sustainability of gold-mining companies. This challenge has also highlighted the importance of improving the sector's environmental efficiency. Consequently, many gold-mining companies regard raising efficiency and productivity as one of their main goals. In the context of the debate regarding sustainable development in mining, this paper examines the technical and environmental performance of the gold-mining industry in developing countries.

As the environmental sustainability of economic activities has become of increasing interest, firm-performance studies have evolved to include environmental concerns, and conventional efficiency measures have been extended to include both technical and environmental dimensions (Färe, Grosskopf, Lovell and Pasurka, 1989; Reinhard, Lovell and Thijssen, 1999; Oude Lansink and Van der Vlist, 2008; Serra, Chambers and Lansink, 2014; Kumbhakar and Tsionas, 2016). Little is known about the technical and environmental (in)efficiency of gold mining in developing countries. A better understanding of technical and environmental efficiency is key to improving regulations, monitoring, mitigation plans and enforcement, in order to achieve sustainable mining.

In analysing environmental efficiency, there is considerable debate on how to model pollution- and emission-generating technologies; however, the by-production model developed by Fernandez, Koop and Steel (2002), Førsund (2009), and Murty, Russell and Levkoff (2012) is viewed as the most appropriate model. Only a few studies (see Serra, Chambers and Lansink, 2014; Kumbhakar and Tsionas, 2016) have analysed technical and environmental efficiency using the by-production model in data envelopment analysis (DEA) and stochastic frontier analysis (SFA) settings. No study has considered separating firm heterogeneity from persistent and transient environmental efficiency, especially in mining firms. However, the four-component model developed by Kumbhakar, Lien and Hardaker (2014), Colombi et al. (2014) and Filippini and Greene (2016) accounts for firm heterogeneity and persistent and transient technical efficiency separately, and thus provides more robust efficiency estimates.

The four-component model is increasingly being used to estimate efficiency, in many studies (see Filippini, Geissmann and Greene, 2018; Heshmati, Kumbhakar and Kim, 2018; Colombi, Martini and Vittadini, 2017). To the best of our knowledge, this model has not been applied in assessing environmental performance. Separating firm-heterogeneity from time-varying inefficiency helps to quantify the magnitude of persistent inefficiency – which is important, as it captures the effects of inputs such as management as well as other unobserved inputs, which vary across firms but not over time. Thus, unless there is a change in management or in something that affects the management style of individual firms, such as a change in government policy towards the industry, changes in firm ownership and so on, it is very unlikely that the persistent inefficiency component will change (see Kumbhakar, Wang and Horncastle, 2015). The persistent inefficiency components explain the long-run inefficiencies.

By contrast, the transient component of inefficiency might change over time without any change in the operation of the firm. These are short-run inefficiencies, which may occur in one year but not the next year. Therefore, the distinction between the persistent and transient components of inefficiency is important, because they have different policy implications; the focus should be on the persistent inefficiencies, as they tend to affect performance in the long term. Moreover, failure to separate firm effects from persistent inefficiency in the model is likely to produce biased estimates of overall efficiency (see Kumbhakar, Wang & Horncastle, 2015). Applying the four-component model in the by-production specification to estimate persistent, transient and overall technical and environmental efficiency, in the context of gold mines in developing countries, represents a novel contribution.

The rest of the paper is structured as follows: Section 2 presents a literature review, and Section 3 the methodology and data. Section 4 presents estimation and results, and Section 5 our conclusions.

2. Existing literature

This section is divided into two main subsections; firstly, a theoretical literature review, which highlights the important debate on the evolution of efficiency-model specifications accounting for emission-generating technology. Secondly, an empirical literature review that reveals the evolution of the estimation approaches in which efficiency can be decomposed into estimation and the policy implications components, as far as improving efficiency is concerned.

2.1 Theoretical review

There has been considerable debate in the technical literature on the appropriate modelling of pollution in production technologies. One standard approach assumes that pollution and production are complementary outputs, therefore pollution is treated as though it were an input (De Koeijer et al., 2003; Reinhard, Lovell and Thijssen, 2000; Reinhard et al., 1999). This is based on the observed positive correlation between pollution and intended output. While analytically convenient, this reduced-form specification ignores both the physical reality and the requirements of material balance, which delimits smaller systems or processes where the material inputs should balance the output (Murty, Russell and Levkoff 2012; Halkos and Tzeremes, 2012; Førsund, 2008; Pethig, 2006). This model is referred to as a single-equation model. In this setting, the technology (usually a directional distance function) is specified by a single equation (which can be estimated using either the DEA[4] or the SFA approach) in which good and bad outputs, as well as good and bad (pollution-generating) inputs, enter as arguments. That is, bad outputs can be treated as inputs (Lee, Park and Kim, 2002; Hailu and Veeman, 2001; Reinhard and Thijssen, 2000; Reinhard et al., 1999; Baumol and Oates,1988); and since inputs are assumed to be freely disposable, so are bad outputs. Treating bad outputs as inputs violates the axioms of production theory. A competing approach in single-equation models treats pollution as a weakly disposable or unintended output, subject to null jointness (Färe et al., 1989; Piot-Lepetit and Vermersch, 1998; Wossink and Denaux, 2006).

Generally, the property of detrimental variables of weak disposability is well known, and has been used in several formulations (Oliveira et al., 2017; Godoy-Durán et al., 2017; Färe et al., 2004; Zofio and Prieto 2001; Chung, Färe and Grosskopf, 1997; Tyteca, 1997; 1996). This approach is widely accepted by environmental economists. Nonetheless, most previous studies have analysed the performance of the technical and environmental efficiency of the mining firms (see Oliveira, Camanho and Zanella, 2017) in a single-equation-model setting. This technically violates the production theory assumption, by assuming undesirable outputs as input. Moreover, the material balance assumption is also violated.

The violations steam from the fact that in many cases, production of good, intended outputs also generates some bad, unintended outputs. Pollution is a common example of a bad output. Because it is an unintended outcome, modelling production processes using standard tools may not be appropriate. In a case in which the production process is inefficient, the implications for modelling are two-fold. Firstly, how to model technical and environmental efficiency, and can they can be separated? A production process is said to be environmentally inefficient when the production of pollutants can be reduced without reducing the production of good outputs, given the technology and the input vector. Fernandez, Koop and Steel (2002) define environmental efficiency as the quantity of pollution that can be reduced without sacrificing good output, by adopting the best-practice technology. In general, if a firm is fully technically efficient, a decrease in a bad output is only possible if production of a good output is also reduced. This

[4] The non-parametric (DEA) approaches are limited as far as decomposition of efficiency into persistence and transience is concerned. Such decompositions require econometric approaches, which make SFA more relevant. In this regard, semi-parametric and non-parametric SFA approaches are worth consideration.

property is not automatically satisfied in a model where the same technology is used to produce good and bad outputs jointly – a single-equation-model case (Kumbhakar et al., 2015).

Given that the production of bad outputs increases with good outputs, it is often argued that the monotonic relationship between good and bad outputs is similar to the relationship between inputs and good outputs. Färe, Grosskopf and Weber (2005) criticised the idea of treating bad outputs as inputs. They argued that the treatment of bad outputs as inputs with strong disposability properties (Lee et al., 2002; Hailu and Veeman, 2001) would yield an unbounded output set, which is not physically possible if traditional inputs are a given. Good and bad outputs should satisfy the weak disposability condition. This is an important consideration, and implies that bad outputs cannot be treated as inputs (Kumbhakar, Wang and Horncastle, 2015).

To address several weaknesses identified in the single-equation model, Fernandez et al. (2002) proposed the two-equation model. This model employs the separable-distance function, which allows good and bad output to be modelled separately. Two recent groundbreaking proposals modifying the two-equation model are Coelli, Lauwers and Van Huylenbroeck (2007) and Murty, Russell and Levkoff (2012). Coelli et al. (2007) proposed the inclusion of pollution into conventional productive-efficiency measures by using the materials balance concept. Using the same line of argument, Murty et al. (2012) modelled polluting technologies as an amalgamation of two technologies: an intended-output and a residual-generation technology. Murty et al. (2012), building on the ideas of Frisch (1965), Murty and Russell (2002) and Førsund (2009), argued that analytically, pollution-generating technologies are best modelled as the intersection of two sub-technologies: an intended-production sub-technology and a residual-generation sub-technology. They referred to this structure as a 'by-production technology'.

The by-production approach models the technologies as intersections of two independent sub-technologies, reflecting the relations between goods in intended-output production designed by human engineers, and the emission-generating mechanism of nature governed by material-balance considerations. Moreover, the model assumes the production of a good or desired output (in this case, gold) also produces something undesirable (pollution); that is, a by-production of good and bad outputs. Good outputs are freely disposable, but bad outputs are not. Bad outputs cannot be substituted for good outputs. Bad outputs may be substitutable for some good inputs (Murty and Russell, 2018).

Kumbhakar and Tsionas (2016) followed this idea, considering a modelling approach in which the technology to produce good outputs is specified in terms of a standard transformation function with input-oriented technical inefficiency. Because bad outputs are viewed as by-products of good outputs, the technology for producing bad outputs is naturally separated from the technology for good outputs. This allows one to estimate technical and environmental efficiencies, defined in terms of the technologies to produce good and bad outputs respectively. Therefore, as far as modelling pollution is concerned, the by-production model is the most appropriate. The few studies that have adopted this approach include (but are not limited to) Serra, Chambers and Lansink (2014) in the DEA setting, and Kumbhakar and Tsionas (2016) using a stochastic frontier (SF) approach.

To the best of our knowledge, no previous mining-efficiency study has applied the by-production suggested by Fernandez et al. (2002); all therefore potentially suffer from positive correlation, violation of axioms of production theory, and violation of the material balance assumption, as discussed by Murty et al. (2012).

2.2 Empirical review

Parametric estimation of efficiency has always been split into cross-section and panel-data analysis. For a recent detailed survey of both cross-sectional and panel stochastic frontier models, see Greene (2010). Stochastic frontier models originated with Aigner, Lovell and Schmidt (1977) and Meeusen and Van den Broeck (1977). These models were intended for cross-sectional applications, and made strong assumptions about errors. The original stochastic frontier model included both components in the error term: $\varepsilon_i = \eta_i^+ + \alpha_i$. This was adopted in the pooled panel, and the two components in the error term were presented as $\varepsilon_{it} = u_{it}^+ + v_{it}$, Statistical noise was assumed to be normally distributed, while technical inefficiency was assumed to be distributed according to a specific one-sided distribution such as exponential or half-normal. Furthermore, statistical noise and technical inefficiency were assumed to be independent of each other and of the explanatory variables: the inputs (Lee and Schmidt, 1993).

Panel data addresses most of the limitations encountered in cross-sectional analysis, including accounting for some of the heterogeneity that may exist by introducing an ‘individual (unobservable) effect’; that is, time-invariant and individual-specific unobserved heterogeneity, which does not interact with other variables. In addition, it allows for the examination of whether inefficiency has been persistent over time or is time-varying (modelling the temporal behaviour of inefficiency).

The standard stochastic frontier panel data models have been extended in several directions. Estimation of some of these models included making fewer assumptions, at the same time using more flexible modelling approaches. For example, heterogeneous technologies have been the focus of much research, including random coefficient stochastic frontier models. Other examples include latent class or mixture models and Markov switching models. More recently, an important line of research has been the formulation and estimation of panel models, in which firm effects are separated from inefficiency. In the panel dataset the estimation techniques have thus evolved to account for firm heterogeneity, as well as time-varying and time-invariant inefficiencies. These evolutions can be summarised thus:

1. *Time-Invariant Technical Inefficiency Models*

   A time-invariant model is one whose behaviour (i.e. its response to inputs) does not change with time, which represents a lack of technological advances or gains. In a standard panel data model, the focus is mostly on controlling firm heterogeneity due to unobserved time-invariant covariates. The innovation in time-invariant stochastic frontier models (developed in the 1980s) was to make these firm effects one-sided, to give them an inefficiency interpretation. In some models, these inefficiency effects were treated as fixed parameters (see Schmidt and Sickles, 1984), while others treated them as random variables (see Pitt and Lee, 1981; Kumbhakar, 1987). Various estimation methods are available for this type of model, depending on whether the inefficiency effects are assumed to be fixed or random, and whether distributional assumptions are made regarding the inefficiency and noise components.

   Schmidt and Sickles (1984) introduced a model that assumes the inefficiency effects to be time-invariant and individual-specific.[5]

$$\begin{aligned} y_{it} &= \beta_0 + x_{it}'\beta + v_{it} - u_i^+ \\ &= \alpha_i + x_{it}'\beta + v_{it} \qquad (1) \end{aligned}$$

[5] The model is a standard panel data model, where α_i is the unobservable individual effect. Indeed, standard panel data are fixed, and random effects estimators are applied here to estimate the model parameters including α_i – the only difference is that we transform the estimated value of $\hat{\alpha}_i$ to obtain an estimated value of u_i, $\hat{u}_i$.

where $\alpha_i \equiv \beta_0 - u_i^{+}$[6]

The strong distributional assumptions that were necessary in the cross-sectional setting were replaced by a single assumption, that technical inefficiency is time-invariant (see Lee and Schmidt, 1993). This assumes technical inefficiency to be individual-specific and time-invariant; that is, the inefficiency levels may be different for different firms, but they do not change over time. This implies that an inefficient unit (e.g. a mine) does not learn over time. This might be the case in some situations, for example where inefficiency is associated with managerial ability, and there is no change in management for any of the firms during the period of the study; or if the time of the panel is particularly short. However, even this is unrealistic at times, particularly when considering the oligopolistic nature of the market[7]. To accommodate the notion of productivity and efficiency improvement, Cornwell, Schmidt and Sickles (1990), Kumbhakar (1990), Battese and Coelli (1992), Lee and Schmidt (1993) and Kumbhakar and Wang (2005) all considered models that allow inefficiency to change over time, referred to as time-varying technical inefficiency models.

2. *Time-Varying Technical Inefficiency Models*

In contrast to the time-invariant model above, the behaviour (its response to inputs) of this model changes over time. Such a model allows inefficiency to be individual-specific but time-varying (i.e. the inefficiency of each cross-sectional unit evolves along a specific path, which can either be the same for all units or different for different cross-sectional units).

$$y_{it} = \beta_0 + x_{it}'\beta + v_{it} - u_{it}^{+} \qquad 2$$

Colombi et al. (2014) refer to time-varying inefficiency as short-run inefficiency, and mention that it can arise due to failure in allocating resources properly in the short run. They argue that (for example) a mine with excess capacity may increase its efficiency in the short run by reallocating the workforce across different activities. Thus, some of the engineers' and miners' daily working hours might be changed to include other mine activities. This is a short-run improvement in efficiency that may be independent of short-

[6] Where the superscript (+) indicates the non-negative value of the corresponding error component.

[7] Another potential issue with this model is the time-invariant assumption of inefficiency. If T is large, it seems implausible that the inefficiency of a firm could stay constant for an extended period, and that a firm with persistent inefficiency would survive in the market.

run inefficiency levels in the previous period, which may justify the assumption that v_{it} is independent and identically distributed (IID). However, this does not impact on the overall management of the mine, so v_{it} is independent of time-invariant inefficiency. This approach does not allow for individual effects (in the traditional sense) to exist alongside inefficiency effects. Thus, it allows for inefficiency and individual heterogeneity to be separated. Two approaches are used in estimating these models: fixed-effect (FE) and random-effect (RE). In the FE models the time-varying inefficiency term is non-stochastic (i.e. a parametric function of time), whereas in the RE model the inefficiency effect is composed of either a random term or a combination of a time-invariant stochastic term and a time-varying deterministic function.

3. *Models that separate firm heterogeneity from inefficiency*

The previous time approaches (Equation 1 above) had a drawback in that they fail to distinguish between individual heterogeneity and inefficiency, as they treat unobserved heterogeneity as a measure of inefficiency. In other words, all the time-invariant heterogeneity is confounded with inefficiency; and therefore, the inefficiency component might be picking up heterogeneity in addition to (or even instead of) inefficiency (Greene, 2005a; 2005b). Thus the 'true' random or fixed-effect models proposed by Greene (2005a; 2005b) separate firm effects from inefficiency. Greene's true random-effects model has the following error specification:

$$\varepsilon_{it} = \alpha_i + v_{it} + u_{it}^{+}. \tag{3}$$

In this model, the time-invariant component α_i is viewed as an individual heterogeneity that captures the effects of time-invariant covariates, and has nothing to do with inefficiency. If this is true, then the results from the time-invariant inefficiency models are incorrect. The popular models in this category have been the true fixed-effects model (TFE) and the true random-effects model (TRE) advocated by Greene. However, these models consider any producer-specific time-invariant component as unobserved heterogeneity. Thus although firm heterogeneity is now accounted for, it comes at the cost of ignoring long-term (persistent) inefficiency. In other words, long-run inefficiency is again confounded with latent heterogeneity.

4. *Models that separate persistent and time-varying inefficiency*

The model discussed above – though it separates firm heterogeneity from time-varying inefficiency (which is modelled either as the product of a time-invariant random variable and a deterministic function of covariates, or independent and identically distributed [IID] across firms and overtime) – does not account for persistent technical inefficiency, which is hidden within firm effects. Consequently, the model is mis-specified, and tends to produce a downward bias in the estimate of overall inefficiency, especially if persistent inefficiency exists when there is no change in the operation of the firm.

The models proposed by Kumbhakar (1991), Kumbhakar and Heshmati (1995) and Kumbhakar and Hjalmarsson (1993; 1995) treat firm effects as persistent inefficiency, and include another random component to capture time-varying technical inefficiency. The composite error term in the Kumbhakar-Heshmati (1995) and Kumbhakar-Hjalmarsson (1995) models is of the form[8]:

$$\varepsilon_{it} = \eta_i^+ + v_{it} + u_{it}^+. \qquad 4$$

Therefore, the distinction between the persistent and transient components of inefficiency is important, because they have different policy implications (Kumbhakar and Heshmati, 1995). Identifying the magnitude of persistent inefficiency is important, especially in short panels, because it reflects the effects of inputs such as management (Mundlak, 1961), as well as other unobserved inputs that vary across firms but not over time. These models account for persistent inefficiency by ignoring firm heterogeneity components.

So the question is: should one view the time-invariant component as persistent inefficiency (as per Kumbhakar and Heshmati, 1995; Kumbhakar and Hjalmarsson, 1998; 1993; Kumbhakar, 1991), or as individual heterogeneity that captures the effects of time-invariant covariates and has nothing to do with inefficiency (as per Green, 2005a; 2005b)? The answer lies somewhere in between. That is, part of the firm effects in Greene (2005a; 2005b) may be persistent inefficiency. Similarly, part of persistent inefficiency in the models proposed by Kumbhakar et al. may include unobserved firm effects. Since none of the assumptions used in the models cited above are fully satisfactory, we consider a generalised true random-effect (GTRE) model that decomposes the time-invariant firm

[8] Note that these specifications are no different from the models proposed by Green (2005a; 2005b) mentioned earlier. The difference is in the interpretation of the 'time-invariant term'.

effects into a random firm effect (to capture unobserved heterogeneity, as in Greene, 2005a; 2005b), and a persistent technical inefficiency effect (as in Pitt and Lee, 1981; Schmidt and Sickles, 1984; Kumbhakar, 1987).

Thus, the models in Equations (3) and (4) are generally referred to as three-component models, as they divide the error term into three components. The three-component model has some serious shortcomings; it does not separate firm effects from persistent inefficiency, which led to the four-component panel-data stochastic frontier model. This last distinguishes between long- and short-run inefficiency, and accounts for unobserved heterogeneity. In other words, the four-component model separates firm effects, persistent inefficiency and time-varying inefficiency. The study by Kumbhakar and Tsionas (2016), though it controls for endogeneity, does not control for unobserved firm effects, which must be separated from persistent inefficiency.

5. *Models that separate firm effects, persistent inefficiency and time-varying inefficiency*

A model that separates persistent and time-varying inefficiency views firm effects (fixed or random) as long-term (persistent) inefficiency, with an added second component to capture time-varying technical inefficiency. As such, this model confounds firm effects (that are not part of inefficiency) with persistent inefficiency. Consequently, this model is mis-specified, and is likely to produce an upward bias in inefficiency by treating firm-effects as inefficiency. The models that separate firm heterogeneity from inefficiency view firm effects (fixed or random) as something other than inefficiency; thus, these models fail to capture persistent inefficiency, which is compounded with firm effects. Consequently, these models are also mis-specified, and tend to produce a downward bias in the estimate of overall inefficiency.

The models by Kumbhakar, Lien and Hardaker (2014), Tsionas and Kumbhakar (2014) and Colombi et al. (2014) overcome some of the limitations of the earlier (TFE and TRE) models, and are commonly referred to as generalised true fixed-effect (GTFE) models and generalised true random-effect (GTRE) models, when fixed- or random-effect models respectively are specified. In these models, the error term is split into four components to consider different factors affecting output, given the inputs. The first component captures firms' latent heterogeneity (Greene, 2005a; 2005b), which has been disentangled from the inefficiency effects, while the second component captures short-run (time-varying) inefficiency. The third component captures persistent or time-invariant inefficiency, as in Kumbhakar and Hjalmarsson (1993) and Kumbhakar and Heshmati (1995), while the last component captures random shocks. The fully flexible error specification is:

$$\varepsilon_{it} = \alpha_i + v_{it} + \eta_i^+ + u_{it}^+. \qquad 5$$

The economic rationale for all these components is discussed in Kumbhakar, Lien and Hardaker (2014). Moreover, Colombi et al. (2014) give detailed justification for the use of a four-way error component model. If policymakers (regulators) are interested in eliminating persistent inefficiency that is often attributed to regulation, it is necessary to estimate it first. Estimating a model with only one inefficiency component (with or without controlling for firm effects) is likely to give incorrect estimates of inefficiency.

The decomposition in Equation 5 above may be desirable for policy purposes, especially in regulated industries. Since η_i^+ does not change over time, if a regulator wants to improve efficiency then some fundamental change in management or policy must occur. In a regulated industry, all the firms may be operating under excess capacity, which might be reflected in high values of η_i^+ ; but as long as η_i^+ is similar among all firms, relative persistent inefficiency among firms will be low. In such a case, the rankings of firms based on relative values of η_i^+ will be similar, and the regulator cannot punish some firms because all firms have high values of η_i^+. However, the estimates of η_i^+ provide useful information about the firms in the industry, because high values of η_i^+ are indicators of non-competitive market conditions. This is because in a competitive market, there is no persistent inefficiency; i.e. persistently inefficient firms will go out of business. The short-run inefficiency u_{it}^+ can be adjusted over time without a major policy change. Thus, for example, if the short-run inefficiency component for a firm is relatively large in a particular year, then it may be argued that the inefficiency is caused by something unlikely to be repeated in the following year. On the other hand, if a firm's persistent inefficiency component is large, then it is expected to operate with a relatively high level of inefficiency over time, unless some changes in policy and/or management take place. Thus, a high value of η_i^+ is of more concern from a long-term point of view, because of its persistent nature, than a high value of u_{it}^+.

As discussed in the introduction, the pollution that gold mines produce is significant, and is becoming an important policy issue. To address some of the key policy questions of interest in this paper, it is important to assess the technical and environmental performance of the gold mines using the four-component model. To the best of our knowledge there are no studies in developing countries that have used firm-level data to analyse technical and environmental efficiency in the by-production technology model settings. Thus, this study builds on Kumbhakar and Tsionas (2016) by further decomposing the error term into four components, and provides policy implications in the mining sector. By using firm-level analysis, this study will generate valuable insights for mining companies and for policymakers, which should enable them to make better-informed decisions on how to improve their environmental efficiency.

3. Methodology

In this section we present the mixed methodological approach adopted in this study. The two-equation model accounting for technical and environmental efficiency is described, as well as the four-component panel-data stochastic frontier model. The latter model disentangles unobserved firm effects (firm heterogeneity) from persistent (time-invariant/long-term) and transient (time-varying/short-term) technical or environmental inefficiency.

3.1 Model Specification

As in Kumbhakar and Tsionas (2016), we follow the by-production specification approach of FKS[9] (2002), Førsund (2009) and MRL[10] (2012), which uses two separate technologies to model good and bad outputs. The former describes the textbook-type production process (i.e. inputs – good and bad) in which the transformed, desirable outputs do not depend on bad outputs. Further, it satisfies all the standard properties, most importantly the free-disposability property (for the derivation of properties, see Kumbhakar and Tsionas, 2016). The latter can be viewed as a residual generation technology which models the production of bad outputs as a function of good outputs (FKS, 2002), bad inputs (MRL, 2012) or both bad and good inputs (Førsund, 2009). The positive relationship between bad and good outputs in FKS (2002) follows from this residual technology, which embeds the relationship explicitly. Further, inefficiency is allowed in each technology, so technical inefficiency can be distinguished from environmental inefficiency.

[9] Fernandez, Koop and Steel (2002).
[10] Murty, Russell and Levkoff, (2012).

3.1.1 Good output (technical efficiency) technology

Using a transformation function (TF) representation of the underlying technology with input orientation (IO) inefficiency to produce good outputs, we assume that good outputs Y are exogenously given (in our application the 'good output' is gold produced, which is demand-determined and therefore exogenous to the firm). This justifies an input distance function (IDF[11]), represented as:

$$F(Y, \theta X^g, X^b, t) = 1 \qquad 6$$

The transformation function F (▪) is assumed to satisfy all the standard monotonicity properties (for a discussion of this property, see Kumbhakar and Tsionas, 2016). Using the linear homogeneity restrictions (in θX^g), the transformation function in Equation 6 above can be expressed as:

$$\left(\theta X_1^g\right)^{-1} = \psi^g\left(Y, \tilde{X}^g, X^b, t\right) \qquad 7$$

where X_1^g = labour, Y = gold output, $\tilde{X}^g$ = capital, energy, water; X^b = fuel, and t = time trend.

Linear homogeneity was imposed a priori by normalising gold output and input with respect to the constant 'labour' (we chose labour as the numeraire, and define the other variables). In a translog (TL) form of the transformation function – TF(F) – we can represent Equation 6 above as:

$$x_{1,it}^g \equiv TL\left(\tilde{x}_{j,it} y_{m,it}, t\right) + v_{it} + u_{it} \qquad 8$$

$$x_{1,it}^g = \alpha_0 + S'_{it} a + \frac{1}{2} S'_{it} A S_{it} + v_{it} + u_{it} \qquad 9$$

where input-oriented technical inefficiency $u_{it} = -ln\theta > 0$ is technically inefficient, which is the percentage overuse of inputs due to inefficiency, while v_{it} is statistical noise.

A three-component model (separating firm heterogeneity from efficiency) may be specified as:

$$x_{1,it}^g = \alpha_0 + S'_{it} a + \frac{1}{2} S'_{it} A S_{it} + \alpha_i + v_{it} + u_{it}^+ \qquad 10$$

[11] If outputs are exogenously given, and the objective is to minimise input, then the natural choice is to use an input distance function (IDF) and estimate the efficiency component(s). Note that in an IDF, inputs are endogenous (Kumbakhar et al., 2015). If the technology is not known (which is the case in reality) and it must be estimated econometrically, then the issue of endogeneity cannot be avoided.

The proposed alternative approach using a four-component model is defined as:

$$x_{1,it}^{g} = \alpha_0 + S_{it}'\boldsymbol{a} + \frac{1}{2}S_{it}'\boldsymbol{A}S_{it} + \alpha_i + \eta_i^{+} + u_{it}^{+} + v_{it} \qquad 11$$

where $\tilde{x}_{j,it}^{g} = x_{j,it}^{g} - x_{1,it}^{g}, j = 2, \ldots, J$. $S_{it}' = (y'_{it}, x_{it}^{b'}, \tilde{x}_{it}^{g'}, t)$. **A, a** are the vector and matrix of the relevant parameters. In this model, the error term is split into four components to consider different factors affecting output, given the inputs. The first component α_i captures firms' latent heterogeneity, which must be disentangled from inefficiency; it is a random mine effect which captures mines' heterogeneity (Green, 2005a; 2005b). The second component η_i^{+}captures the persistent (long-run) technical inefficiency component. The third component u_{it}^{+} captures short-run or transient technical inefficiency; while the last component v_{it} captures random shocks, which is similar to the noise component in a standard regression model. In this model, the overall technical efficiency (OTE) obtained is the product (interaction) of persistent technical efficiency (PTE) and transient technical efficiency (RTE); that is, OTE= PTE x RTE.

3.1.2 Bad output (environmental efficiency) technology

We specify the technology to produce single bad outputs (carbon dioxide, CO_2) as follows:

$$H(Y, \lambda Z, X_b, t) = 1 \qquad 12$$

where $\lambda \leq 1$ is an environmental inefficiency in the production of Z. More specifically, this radial measure shows that $(1-\lambda_q)$ 100% is the rate at which the production of bad output Z can be reduced without reducing good outputs and bad inputs. Note that the technology to produce bad outputs is assumed to be homogeneous with degree 1 in Z. It is also assumed the ratios of bad outputs are predetermined/exogenous, and do not correlate with the error components. Linear homogeneity was imposed a priori by normalising CO_2 emitted and abad input (fuel) with respect to the constant (gold) output.

The transformation function H (▪) is assumed to satisfy all the standard monotonicity properties. Using the linear homogeneity restrictions and IDF specification, the transformation function in Equation 12 above can be expressed as:

$$(y_{it})^{-1} = g(z_{it}, x_{it}^{b}, t) \qquad 13$$

where $Y = \text{gold output}, z_{it} = \text{ bad output } (\text{CO}_2), X^b = \text{fuel, and } t = \text{time trend.}$

For environmental efficiency we also assume the translog functional form on g (▪), for the transformation function-TF (H). Thus, Equation 12 above is re-written as:

$$y_{it} \equiv TL(z_{it}, x_{it}^{b}, t) + v_{it} + u_{it} \qquad 14$$

$$y_{it} = a_0 + P_{it}'\delta + \frac{1}{2}P_{it}'\Delta P_{it} + \zeta_{it} + \tau_{it} \quad 15$$

τ_{it} = ln λ_{it}≥0 is environmental efficiency and ζ_{it} is an error term. Further, δ and Δ are vector and matrix of relevant parameters in the translog function, representing the production of bad outputs. In a three-component model (separating firm heterogeneity from efficiency):

$$y_{it} = a_0 + P_{it}'\delta + \frac{1}{2}P_{it}'\Delta P_{it} + \beta_i + \zeta_{it} + \tau_{it}^{+} \quad 16$$

In a four-component model, this can be written as:

$$y_{it} = a_0 + P_{it}'\delta + \frac{1}{2}P_{it}'\Delta P_{it} + \beta_i + \phi_i^{+} + \tau_{it}^{+} + \zeta_{it} \quad 17$$

where $S_{it}' = (z'_{it}, x_{it}^{b'}, t)$. The error term is split into four components to consider different factors affecting output, given the inputs. The first component β_i captures firms' latent heterogeneity, which must be disentangled from inefficiency; the second component τ_{it}^{+} captures short-run (time-varying) environmental inefficiency. The third component ϕ_i^{+} captures persistent or time-invariant inefficiency, while the last component ζ_{it} captures random shocks. In this model the overall environmental efficiency (OEE) is then obtained from the product of persistent environmental efficiency (PEE) and transient environmental efficiency (TEE); that is, OEE= PEE x TEE.

3.2 Model Estimation

The four-component errors, particularly in a generalised true random-effects model (GTREM), are usually estimated using one of several methods, in this case the three-step method-of-moment estimator (MME) by Kumbhakar et al. (2014). In this setting, given the structure of the four separate errors, deriving the likelihood function was previously seen as infeasible. This approach estimated each component separately, making implementing its procedures straightforward.

The second method, proposed by Colombi et al. (2014), found a tractable likelihood function which fits all four components. Results are drawn from skew-normal and closed skew-normal (CSN) distributions. Assuming v_{it} is i.i.d. normal and u_{it} is i.i.d. half normal, the sum of these two distributions K_{it} has a skew-normal distribution. Using the same argument, if η_i is i.i.d. half normal and ν_i is i.i.d. normal, the sum of these two distributions V_i is a skew-normal distribution. Thus, the overall distribution will be the sum of two skew-normal distributions

$(K_{it}\ and\ V_i)$, and is a closed skew-normal (CSN) distribution. This innovative approach allows for the estimation of the four-component model to be undertaken using a single-stage maximum likelihood estimator (MLE) method (also referred to as a full maximum likelihood method) based on CSN distribution (see Table 1 below for a breakdown of the distribution assumptions).

The third approach, proposed by Tsionas and Kumbhakar (2014), used similar assumptions to that proposed by Colombi et al. (2014); but instead of using the MLE, they used a Bayesian Markov chain Monte Carlo (MCMC) approach to estimate the GTRE model. The fourth method, proposed by Filippini and Greene (2016), argued that although the CSN framework gives a closed-form expression of the log-likelihood function, implementing it in practice is a challenging task. Thus they proposed a simulated maximum likelihood estimator (SMLE), which overcomes many of the challenges associated with MLE. Using the insights of Butler and Moffitt (1982), Filippini and Greene (2016) noted that the density in Colombi et al. (2014) can be greatly simplified by employing conditioning on η_i and ν_i. In this case, the conditional density is simply the product over time of Ti univariate skew-normal densities. Thus, only a single integral (as opposed to Ti integrals) must be evaluated.

In this study, as discussed above, we employ the TREM by Green (2005b) and the SMLE GTREM by Filippini and Greene (2016). The TREM results are from the three-component model, in which firm heterogeneity is separated from inefficiency. This model serves as the benchmark to which the SMLE GTREM is compared. We chose the SMLE GTREM since it is a more advanced model for estimating the four-component model.

Table 1: Development of distribution assumptions

		TREM Green (2005a, 2005b)	**MME** Kumbhakar et al. (2014)	**MLE and SMLE** Colombi et al. (2014), Tsionas and Kumbhakar (2014), Filippini and Greene (2016)		
	Component	**Distri-bution**	**Distri-bution**	**Distri-bution**	**Distri-bution**	**Assumption**
1	Random effect: ν_i	Normal	Normal	Skew Normal		Homoskedastic

2	Persistent inefficiency: η_i	—	Half-Normal		Closed Skew Normal (CSN)	Homoskedastic
3	Random noise: υ_{it}	Normal	Normal	Skew Normal		Homoskedastic
4	Transient inefficiency: μ_{it}	Half-Normal	Half-Normal			Homoskedastic

Table 1 above presents the development of distribution assumptions and their respective estimation methods. Column 1 shows the number of error components. Column 2 details the three-component model (TREM), which estimates the components by maximum likelihood estimator (MLE). Column 3 details the MME GTREM model, in which the method of moment estimator (MME) is used to estimate the four components separately. Column 4 presents the models which use the CSN assumption (see Colombi et al., 2014; Tsionas and Kumbhakar, 2014; Filippini and Greene, 2016).

The four-component model is an extension of the TFE or TRE models proposed by Green (2005a; 2005b respectively). This model can be estimated by assuming that either the inefficiency component (μ_{it}) is a fixed parameter that directly influences the dependent variable (the fixed-effect model), or the inefficiency component (μ_{it}) is a random variable that correlates with the independent variable (the random-effect model). This model is known as the Generalised True Fixed Effect (GTFE) model, in cases of Fixed Effect or Generalised True Random-Effect (GTRE) model is considered for a random-effect model (Tsionas, Malikov and Kumbhakar, 2020). We assume that none of the covariates (good and bad outputs, bad inputs, and the ratios of good inputs) are correlated with either the inefficiency or the noise term.

3.3 Data

As indicated in previous sections, this study uses firm-level data on energy consumption, labour, capital, fuel and water consumption as inputs, while volume of gold produced is considered a desirable output. CO_2 on the other hand is an undesirable or unintended output. The variables collected from financial statements were measured as follows: volume of gold produced in a year, measured in ounces (Oz); capital spent in a year, measured in millions of US dollars (M$); and labour, measured by total number of employees. The other variables of interest are energy, which captures the total energy used by the mine, measured in gigajoules

(GJ), and sustainability indicators, which were compiled from mining company sustainability reports: carbon dioxide emissions (CO_2), measured by the total kilotons (kt) of carbon dioxide emitted in a year; quantities of water used in the production process, measured in cubic metres (m^3); and fuel, measured by total fuels used by the mine in kilolitres (Kl). In the few cases where coal was used, we converted coal amounts into fuel equivalents. The choice of variables was made for consistency with those in similar studies (see Oliveira et al., 2017; Arabi, Munisamy, Emrouznejad and Shadman, 2014; Zhou, Ang and Wang, 2012; Hua, Bian and Liang, 2007).

The mine-level data set for this analysis covers the years 2009 to 2018 and was derived from a variety of sources, such as annual financial statements, sustainability reports and websites. These data are in the public domain, and are published voluntarily by mining companies on an annual basis. We have data for 34 large gold-mining companies (see Appendix 1 Table 1 for a description of the mining firms used in the analysis), observed over 10 years. The choice of firms was based on the availability of complete information and firm-specific data for the variables required for the analysis. The data were analysed using Stata for the TRE model, while R was employed for estimating an SMLE.

4. Results

4.1 Descriptive Statistics

Table 2 below presents results for each of the variables of interest. The descriptive statistics represent the variables of the production functions in Equations 11 and 17 above. The statistics are based on the full sample of observations.

Table 2: Descriptive statistics of inputs and outputs

Variable	Obs	Mean	Std.Dev.	Min	Max
Gold output: y [kilo-ounce]	340	185000	136000	3000	719000
Capital: k [million US dollars]	340	199.273	354.616	0.21	2387.384
Labour: l [Total number of employees]	340	2763.344	1809.99	117	9020
Energy: e [Gigajoules]	340	1425.852	1432.674	128.908	9250
Water: w [kilolitres]	340	2476.041	2057.005	137	11191
Oil-fuel consumption: f [kilolitres]	340	7253.68	15081.33	66.095	88454
CO_2 emitted: c [kilotons]	340	333.082	420.554	12	2178.667

Time trend: t	340	5.5	2.877	1	10

The standard deviations are relatively dispersed around their mean, which could be associated with a high range between the minimum and maximum values for all the variables. The variation in the range is because the mine operations are not homogenous, even though they are all gold-mining companies, as their product mix, location and exploration sites may be different. For example, a firm could have an open pit mine as opposed to underground mines, while the grade of ore can differ considerably across mines; and some have different milling processes, ranging from heap leaching to alternative leaching technologies. They also face different challenges regarding natural conditions, infrastructure, and the economic context of their operation.

4.2 Estimation of Technical Efficiency

The estimated technical efficiency results from the three frontier models are presented in Table 3 below. This is followed by a presentation of the technical efficiency scores, in Table 4 below. The results in Table 3 are grouped into two columns; the first shows benchmark results obtained from TREM estimation, while the second presents result from the multiple-equation SMLE GTREM model.

Table 3: Estimation of technical efficiency based on TREM and SMLE GTREM specifications

	TREM		SMLE GTREM	
Ln Labour llneg	Coef.	St.Err	Coef.	St.Err.
Ln Gold output; ly	1.325***	(0.344)	2.246***	(0.320)
ly^2	-0.159***	(0.027)	-0.224***	(0.026)
Ln (capital/labour); tk	-0.036	(0.123)	-0.394***	(0.088)
Ln (energy/labour); te	0.168	(0.185)	0.460**	(0.169)
Ln (fuel/labour); tf	0.099***	(0.020)	0.085***	(0.016)
Ln (water/labour); tf	0.007	(0.014)	0.026	(0.015)
Time trend; t	0.004*	(0.002)	0.006*	(0.003)
lytk	0.065***	(0.010	0.084***	(0.007)
lyte	-0.006	(0.016)	-0.025	(0.014)
tfte	0.0101	(0.009)	0.015	(0.008)

Constant	-9.382***	(2.229)	-16.178***	(1.953)
Number of observations	340		340	340
Firms' latent heterogeneity ν_i	-3.126***	(0.117)	-3.646	(8.148)
Persistent inefficiency η_i			-1.331***	(0.156)
Transient technical inefficiency υ_{it}	-7.207***	(0.862)	-2.057***	(0.082)
Random noise component μ_{it}	0.135	(0.257)	-0.415***	(0.024)
Log likelihood	101.511		107.135	

Note: Table 3 above presents the estimation results when applying TREM (Input Directional Function-IDF Equation10 above) and SMLE GRTEM frontier models to the IDF Equation 11. The SMLE GTREM results are based on 500 draws each. Asterisks: *** indicates significance at 1% level, ** at 5% level and * at 10% level. Standard errors are reported in parenthesis.

The findings show that in all three models, transient or time-varying technical inefficiency (υ_{it}) in the large gold mines is negative and statistically significant at a 5 per cent level of significance. Also, the persistent efficiency (only found in the GTRE models, in our case) was negative and statistically significant. The significant negative transient and persistent inefficiencies[12] imply that the mining firms may obtain the same level of output using fewer inputs; or they could adjust their short-run factors (such as management) or long-run factors (such as regulations). The high input usage could be attributed to several factors; among others, it may imply behaviour concerns such as moral hazards, or asymmetric information, which potentially could highlight the principal-agent problem in management.

The inclusion of the interaction terms (translog function relative to Cobb-Douglas function) was tested, the significant p-value (at 5 per cent level of significance) from the Wald test indicating that the coefficients of the interactions were not simultaneously equal to zero. Thus, the translog production function is preferred to the Cobb-Douglas function specification.

Table 4 below provides descriptive statistics of the estimated levels of technical efficiency for the TREM and SMLE GRTEM frontier models. The statistics are based on the full sample of observations. However, the TREM results had missing values (see the number of observations).

Table 4: Descriptive statistics of estimated technical efficiencies

[12] The presence of technical inefficiency implies that given inputs, less than the maximal possible amount of desirable output is produced. Alternatively, it means that more than the minimal amounts of inputs are used to produce a given level of desirable output.

	TREM	SMLE GTREM		
Variable	Transient	Transient	Persistent	Overall
Mean	0.821	0.956	0.731	0.699
Std. Dev.	0.130	0.001	0.057	0.054
Min	0.451	0.951	0.647	0.615
25% Pc.	0.708	0.955	0.677	0.647
Median	0.851	0.736	0.736	0.870
75% Pc.	0.928	0.784	0.785	0.750
Max	0.990	0.824	0.824	0.788
Obs.	260	340	340	340

The mean transient efficiency of the TREM (82%) is relatively smaller in magnitude than the mean transient result of the SMLE GTREM (96%0. The dispersion of the estimated transient efficiencies is slightly higher for the TREM than for the SMLE GTREM. As depicted in Figure 1 below, mean efficiency estimates within the four quartiles of the annual efficiency distributions are relatively constant over time, regardless of model specification (these claims were also verified by the insignificant-at-5-per-cent level of significant trend estimations provided in Appendix 1, Table 2). Hence, we find robust empirical evidence that on average, gold-mining firms neither strongly increased nor decreased their transient or persistent technical efficiency between 2009 and 2018. The SMLE GTREM is our preferred model, because it allows for simultaneous estimation of the level of persistent as well as transient technical efficiency. Moreover, the log-likelihood in the GTRE model is higher than that found in the benchmark model TRE. The predicted overall technical efficiency of this model amounts to 70% (96% transient and 73% persistent) on average.

Figure 1 below presents the development of estimated transient (since it is available in all three models) technical efficiencies under the TREM and SMLE specifications. For every individual year, firm-level transient technical efficiency estimates are separated into mean and quantiles – lower quantile (Pc 25), second quartile (median), and upper quantile (Pc75).

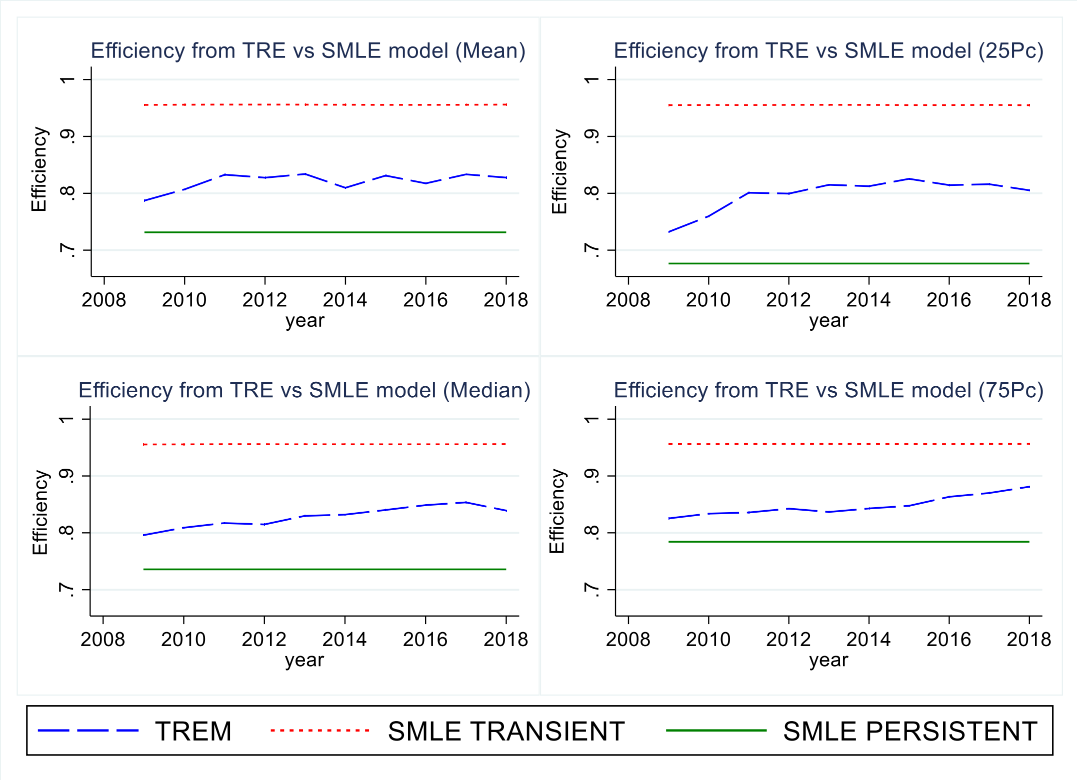

Figure 1: Development of estimated technical efficiencies over time

The TREM and the persistent efficiency component of the SMLE GTREM measure different kinds of technical efficiency, which explains why the correlation between these two estimated efficiency levels is low – and even negative, for SMLE (Table 5 below).

Table 5: Correlation matrix for components of technical efficiency

Variables	(1)		(2)		(3)
(1) TRE Transient	1.000				
(2) SMLE Persistent	-0.111	[-0.06]	1.000		
(3) SMLE Transient	0.104	[0.11*]	0.152	[0.18***]	1.000

Note: Table 5 presents the Pearson correlation coefficients between estimated technical efficiencies of TREM and SMLE GRTEM frontier models. Spearman[13] correlations are given [.] brackets. Asterisks: *** indicates significance at 1% level, ** at 5% level and * at 10% level.

In contrast, the correlation between the TREM technical efficiency and the transient efficiency of the SMLE GTREM is positive and comparatively high, as expected. In conclusion, the positive and statistically significant correlation between persistent and transient (SMLE

[13] The Spearman is preferred in this case, since it is based on the ranked value (efficiency scores are ranks) for each variable; rather than the Pearson, which evaluates the linear relationship between raw data.

GTREM) efficiencies suggest that firms showing a high degree of persistent technical efficiency are also simultaneously exhibiting production processes of a high degree of transient technical efficiency.

4.3 Estimation of environmental efficiency

The estimated environmental efficiency coefficients of the three frontier models, as well as their respective standard errors, are shown in Table 6 below.

Table 6: Estimation of Environmental Efficiency of the 34 gold mines

	TREM		SMLE GTREM	
Ln (gold output); lyneg	Coef.	St.Err.	Coef.	St.Err
Ln (CO_2 emitted); lC	-1.308***	(0.274)	-1.742***	0.052
lC^2	0.156***	(0.048)	0.138***	0.010
Ln (fuel/gold output); tZ	-0.292	(0.180)	-1.184***	0.0524
tZtZ	0.065*	(0.039)	0.066***	0.0134
tZlC	-0.067***	(0.016)	-0.213***	0.0061
tZt	-0.007	(0.005)	0.010***	0.0017
tCt	-0.004	(0.006)	0.015***	0.0018
t	-0.079	(0.049)	0.144***	0.015
Constant	-2.594**	(0.858)	-4.582***	0.1501
Number of observations	340		340	
Firms' latent heterogeneity ν_i	-1.945***	(0.408)	0.2275	0.2581
Persistent inefficiency η_i			-2.245***	0.2008
Transient technical inefficiency υ_{it}	-3.144***	(0.428)	-2.054***	0.2744
Random noise component μ_{it}	-0.496*	(0.256)	-1.457***	0.1064
Log-likelihood	-149.74629		232.525	

Note: Table 6 presents the estimation results when applying TREM (IDF Equation 16 above) and SMLE GRTEM frontier models to the IDF Equation 17 above. The SMLE GTREM results are based on 500 draws each. Asterisks: *** indicates significance at 1% level, ** at 5% level and * at 10% level. Standard errors are reported in parentheses.

The results in Table 6 above show that in all three models, transient or time-varying technical inefficiency (υ_{it}) in the large gold mines is negative and statistically significant at a 5 per cent level of significance. Persistent efficiency was also negative and statistically significant at a 5 per cent level of significance. The significant negative environmental inefficiencies[14] imply that the mines could reduce their bad output (CO_2) emissions without sacrificing good outputs (gold), by adopting the best-practice technology. Besides that, the Wald test between the translog and restricted Cobb-Douglas model shows that statistically, the interaction terms significantly improve the results compared to the restricted model.

Table 7 below provides descriptive statistics of the estimated levels of environmental efficiency of the TREM and SMLE GRTEM frontier models. The statistics are based on the full sample of observations.

Table 7: Descriptive statistics for the estimated environmental efficiencies

	TREM	SMLE GTREM		
Variable	Transient	Transient	Persistent	Overall
Mean	0.815	0.809	0.820	0.664
Std.Dev.	0.066	0.022	0.019	0.024
Min	0.374	0.659	0.790	0.531
25% Pc.	0.802	0.803	0.808	0.654
Median	0.829	0.811	0.814	0.661
75% Pc.	0.850	0.818	0.834	0.674
Max	0.910	0.886	0.874	0.7536
Obs.	340	340	340	340

The mean and median transient efficiency of the TREM (81.5% and 82.9% respectively) are relatively similar in magnitude to the mean and median transient result of the SMLE GTREM (81% and 81% respectively). The dispersion of the estimated transient efficiencies is slightly higher for the TREM than for the SMLE GTREM. As depicted in Figure 2 below, mean efficiency estimates within the first quartiles of the yearly efficiency distributions are relatively constant over time, regardless of model specification (see Appendix Table 2 for trend estimation results). Hence we find robust empirical evidence that on average, gold-mining firms neither strongly increased nor decreased their transient or persistent environmental

[14] The presence of (environmental) inefficiency in by-production therefore means that more than this minimal amount of the undesirable output is produced.

efficiency between 2009 and 2018. The predicted overall technical efficiency (SMLE GTREM) of this model amounts to 66% (81% transient and 82% persistent) on average.

As with the previous Figure 1, we present the development of estimated efficiencies under the TREM and the SMLE specifications – the difference (between Figure 2 and Figure 1) being that we are focusing on environmental efficiencies. Similarly, firm-level environmental efficiency estimates are separated into mean and quantiles: lower quantile (Pc 25), second quartile (median), and upper quantile (Pc75).

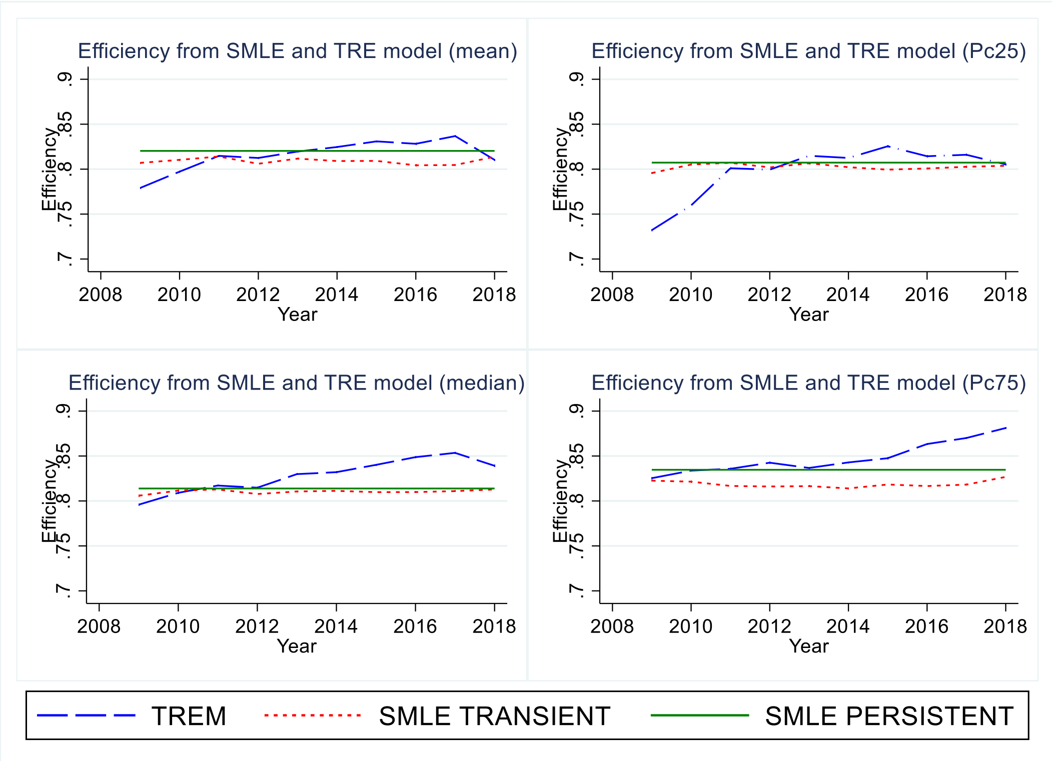

Figure 2: Development of estimated Environmental efficiencies over time

The TREM and the persistent environmental efficiency component of the SMLE GTREM measure different kinds of environmental efficiency; thus, the correlation between these two estimated efficiency levels is low (see Table below).

Table 8: Correlation matrix for components of environmental efficiency

Variables	(1)	(2)	(3)
(1) TRE Transient	1.000		
(2) SMLE Persistent	0.014 [0.017]	1.000	
(3) SMLE Transient	0.161 [0.14**]	0.103 [0.12**]	1.000

Note: Table presents the correlation coefficients between estimated environmental efficiencies of TREM and SMLE GTREM frontier models. Spearman correlations are given in [.] brackets. Asterisks: *** indicates significance at 1% level, ** at 5% level and * at 10% level.

As expected, the correlation between the TREM and the transient environmental efficiency of the SMLE GTREM is positive and comparatively high. In conclusion, the positive and statistically significant correlation between persistent and transient (SMLE GTREM) environmental efficiencies suggests that firms showing a high degree of persistent environmental efficiency are also simultaneously exhibiting production processes of a high degree of transient environmental efficiency.

4.4 Summary of both technical and environmental efficiencies under SMLE GTREM

Table 7 below provides an overall summary of the components of the technical and environmental efficiency components estimated for large gold mines in developing countries. The results show that on average, the mines' technical efficiency is 70%. These mining firms have high transient efficiency (96%), higher than their persistent efficiency (73%). These findings suggest the mines are technically efficient in the short run, while the long-run variables – such as regulations and structural rigidity (which may influence input usage) – pose large constraints on optimising the industry. However, the transient component of efficiency does not change over time, which may imply that the operations of the mines do not change over time.

Table 7: Descriptive statistics of both technical and environmental efficiencies

Variable	Obs	Mean	Std.Dev.	Min	Max
Persistent Technical Efficiency, PTE	340	0.731	0.056	0.647	0.824
Transient Technical Efficiency, TTE	340	0.956	0.001	0.951	0.962
Overall Technical Efficiency, TEO	340	0.699	0.054	0.615	0.788
Persistent Environmental Efficiency, PEE	340	0.82	0.018	0.79	0.874
Transient Environmental Efficiency, TEE	340	0.809	0.022	0.659	0.886
Overall Environmental Efficiency, OEE	340	0.664	0.024	0.531	0.754

On average, transient environmental efficiency has scored approximately 81%; persistent efficiency has scored slightly higher (82%), while overall environmental efficiency is recorded as around 66 per cent over the 10 years of the study (see **Error! Reference source not found.** below). Low environmental performance could be attributed to the fact that most developing countries face structural rigidity; infrastructure such as electrical supply may be limited, forcing them to use large amounts of fuel to generate enough power to run the mine. Poor environmental regulation and enforcement can also be a reason for poor environmental performance. A mine may not be able to adjust rigidity of this kind in the short run, which in turn can undermine persistent efficiency.

Error! Reference source not found. below presents the mean transient, persistent and overall firm-level technical and environmental efficiency estimates for the period of study.

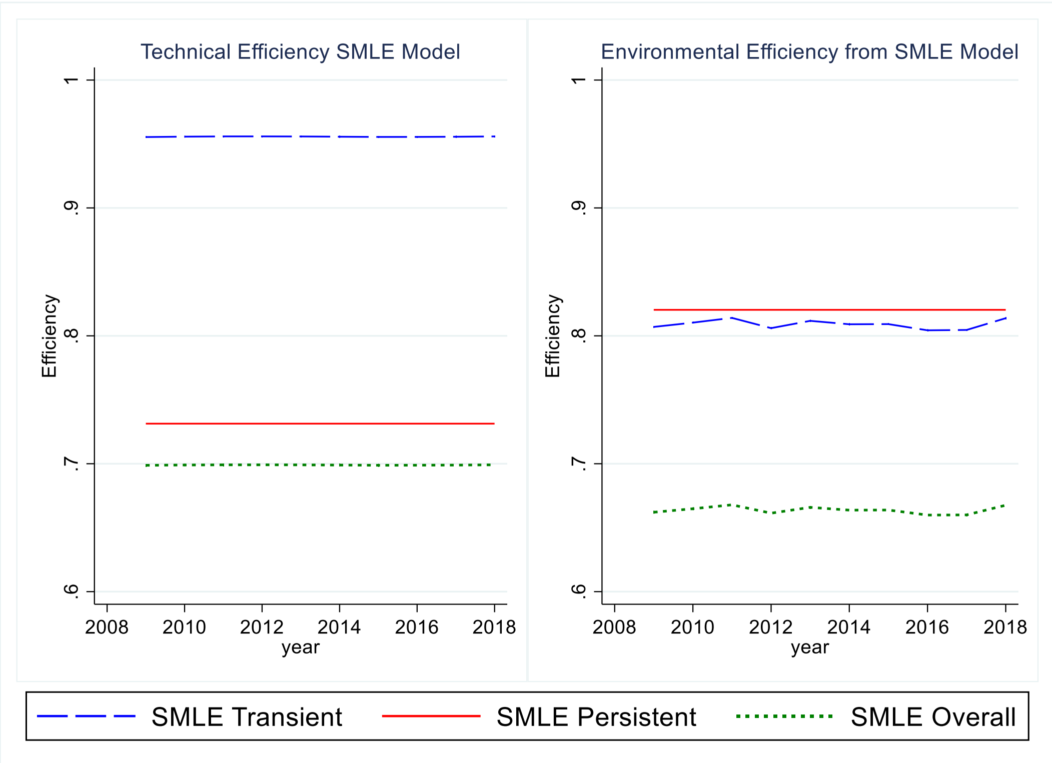

Figure 3: Technical and environmental efficiency components

The relationship between environmental and technical efficiency and their components is represented by the correlation coefficients in Table 8 below.

Table 8: Matrix of correlations of components for both technical and environmental efficiencies

Variables	(1)	(2)	(3)	(4)	(5)
(1) PTE	1.000				

(2) TTE	0.122	[0.15***]	1.000							
(3) PEE	-0.07	[0.11**]	0.001	[-0.0031]	1.000					
(4) TEE	0.031	[-0.03]	0.022	[-0.13**]	0.103	[0.12**]	1.000			
(5) OTE	0.99	[0.99***]	0.137	[0.19***]	-0.07	[0.103*]	0.032	[-0.033]	1.000	
(6) OEE	-0.02	[0.006]	0.016	[-0.0716]	0.69	[0.7***]	0.791	[0.7***]	-0.02	[0.001]

Note: Table 8 presents the correlation coefficients of the components for the estimated environmental and technical efficiencies of SMLE GRTEM frontier models. PTE indicates Persistent Technical Efficiency; TTE indicates Transient Technical Efficiency; OTE indicates Overall Technical Efficiency; PEE indicates Persistent Environmental Efficiency; TEE indicates Transient Environmental Efficiency; OEE indicates Overall Environmental Efficiency. Spearman correlations are given in [.] brackets. Asterisks: *** indicates significance at 1% level, ** at 5% level and * at 10% level.

The Spearman correlation results show that there is a positive and statistically significant correlation between technical and environmental persistent efficiencies, and a negative and significant correlation between technical and environmental transient efficiencies. While there is a positive but not significant correlation between overall efficiencies, similar results were found by Tamini, Larue and West (2012), except that the study was on agriculture farms in Quebec and used TRE model. Moreover, their study found a positive and significant correlation between overall efficiencies. However, the results are similar if we employ the same method (TRE Model), for which the positive and significant at the 1 per cent level correlation between environmental and technical efficiency is 0.2812 or [0.3011] person and spearmen correlations respectively. These results provide strong empirical evidence that firms with high technical efficiency simultaneously have high environmental efficiency.

5. Conclusion

5.1 Summary of the Findings

The goal of this paper was to estimate the transient, persistent and overall technical and environmental efficiency performance of large gold mines in developing countries. We employed a by-production model to specify the emission-generating technology, by applying two distinct frameworks: a TRE model, which is used as a benchmark, and the four-component estimation method using SMLE. We have contributed to the literature in two ways. First, we have provided the only estimates of firm-specific technical and environmental efficiency for mining companies in developing countries. Second, we have made a methodological contribution, by applying the three- (TRE) and four-component models (SMLE) in the by-production specification to provide robust estimates of persistent, transient and overall technical and environmental efficiency in the context of gold mining in developing countries.

The results show a strong correlation between the results for the TRE and SMLE models; which was in line with our expectation, since the estimation methods are similar. The results from both models show no significant trend in terms of environmental and technical efficiency, except that the TRE shows significant positive environmental efficiency. However, once the persistent component is introduced to the model, the trend disappears. Thus we have found robust empirical evidence that on average, gold-mining firms neither strongly increased nor decreased their transient or persistent technical and environmental efficiencies between 2009 and 2018.

On average, the technical efficiency of large gold mines in developing countries is 70%. The mining firms have higher transient efficiency than persistent efficiency. These findings suggest the mines are technically efficient in the short run, while the long-run variables – such as regulations and structural rigidity (which may influence input usage) – create large constraints on optimising the industry. Persistent high technical efficiency suggests the industry is highly competitive (demonstrating oligopolistic competition). However, the transient component of efficiency does not change over time, which may imply that the mines' managerial approach is static. The presence of technical inefficiency implies that more than the minimal amounts of inputs are used to produce a given level of desirable outputs, which could be due to moral hazards (due to low labour motivation) and asymmetric information, such as principal-agent problems. The moderate overall technical efficiency could be attributed mainly to the increasing complexity of mining low-grade gold deposits, which undermines the total output figures.

The overall environmental performance of the mines in the developing countries is low, which indicates that there may be poor environmental policies and regulations, or poor enforcement, or both.

The study found that there is a positive and significant correlation between transient technical and environmental efficiency. Similarly, a positive and significant correlation between persistence technical and environmental efficiency was also found. Moreover, there is a positive and significant correlation between overall technical and environmental efficiency. These results provide strong empirical evidence that firms with high technical efficiency simultaneously have high environmental efficiency, which suggests that promoting high environmental efficiency will also promote high technical efficiency.

5.2 Policy recommendations

Based on the findings in this study, mines should be able to improve their operations over time, since their transient efficiency is rigid. Methods may include (but are not limited to) increasing managerial ability, and adjusting management approach and composition to ensure more efficient allocation of resources. In addition, technology is available that uses less energy and water (and other inputs), while producing higher outputs. To encourage the use of hi-tech solutions, the industry should mechanised the incentive approach to the firms with high efficiency scores. These methods will overcome the challenges of mining lower grades of gold and increasing ore complexity.

In terms of improving long-term environmental efficiency, this study recommends strengthening environmental regulations, and enforcement and adjustment of the structural rigidities such as green energy supply. The regulators should institutionalise incentive-based solutions such as tax credits for lower emissions, which would discourage the use of bad inputs and incentivise firms to acquire better technology. In addition, tradable emission permits should be used to incentivise polluters to internalise the externality. The more efficient firms should benefit more from these incentives, motivating the less efficient firms to become more efficient. As far as structural rigidity and institutional capacity is concerned, governments should work on unlocking the structural barriers and promoting institutional transformation by building the necessary infrastructure.

An examination of the firm-specific determinants of transient, persistence and overall efficiencies for both technical and environmental efficiency is worthwhile, with an eye to finding approximation results in large samples with a large number of inputs and outputs.

References

Aigner, D., Lovell, C. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*(1), 21-37.

Arabi, B., Munisamy, S., Emrouznejad, A., & Shadman, F. (2014). Power industry restructuring and eco-efficiency changes: A new slacks-based model in Malmquist–Luenberger Index measurement. *Energy Policy*, *68*, 132-145.

Badunenko, O., & Kumbhakar, S. C. (2016). When, where and how to estimate persistent and transient efficiency in stochastic frontier panel data models. *European Journal of Operational Research*, *255*(1), 272-287.

Bainton, N., & Holcombe, S. (2018). A critical review of the social aspects of mine closure. *Resources Policy*, *59*, 468-478.

Battese, G. E., &and Coelli, T. J. (1992). , "Frontier pProduction fFunctions, tTechnical eEfficiency and pPanel dData: With aApplication to pPaddy fFarmers in India.," *Journal of Productivity Analysis*, 3, 153-1–69.

Baumol, W. J., Baumol, W. J., Oates, W. E., Baumol, W. J., Bawa, V. S., Bawa, W. S., & Bradford, D. F. (1988). *The Theory of Environmental Policy*. Cambridge Uuniversity Ppress.

Bjorn, S. (2000). Eco-efficiency: Creating more value with less impact. *Conches-Geneva: WBCSD*, 5-36.

Botin, J. A. (2009). *Sustainable Management of Mining Operations.* Society for Mining, Metallurgy an Exploration,. Inc. Englewood, CO.

Bravo-Ureta, B. E., Solís, D., López, V. H. M., Maripani, J. F., Thiam, A., & Rivas, T. (2007). Technical efficiency in farming: Aa meta-regression analysis. *Journal of Productivity Analysis*, *27*(1), 57-72.

Carvalho, F. P. (2017). Mining industry and sustainable development: Ttime for change. *Food and Energy Security*, *6*(2), 61-77.

Chung, Y. H., Färe, R., & Grosskopf, S. (1997). Productivity and undesirable outputs: a directional distance function approach. *Journal of Environmental Management*, *51*(3), 229-240.

Coelli, T., Lauwers, L., & Van Huylenbroeck, G. (2007). Environmental efficiency measurement and the materials balance condition. *Journal of Productivity Analysis*, *28*(1-2), 3-12.

Colombi, R., Kumbhakar, S. C., Martini, G. &and Vittadini, G. (2014). , "Closed-sSkew nNormality in sStochastic fFrontiers with iIndividual eEffects and lLong/sShort-rRun eEfficiency.," *Journal of Productivity Analysis, 42*(2), 123-1–36.

Colombi, R., Martini, G., & Vittadini, G. (2017). Determinants of transient and persistent hospital efficiency: The case of Italy. *Health economics*, *26*, 5-22.

Cornwell, C., Schmidt, P., &and Sickles, R. C. (1990)., "Production fFrontiers with cCross-sSectional and tTime-sSeries vVariation in eEfficiency lLevels.," *Journal of Econometrics, 46,* 185-–200.

De Koeijer, T. J., Wossink, G. A. A., Smit, A. B., Janssens, S. R. M., Renkema, J. A., & Struik, P. C. (2003). Assessment of the quality of farmers' environmental management and its effects on resource use efficiency: Aa Dutch case study. *Agricultural Systems*, *78*(1), 85-103.

Dubiński, J. (2013). Sustainable development of mining mineral resources. *Journal of Sustainable Mining*, *12*(1), 1-6.

Dunlop, T., & Corbera, E. (2016). Incentivizing REDD+: How developing countries are laying the groundwork for benefit-sharing. *Environmental Science & Policy*, *63*, 44-54.

Färe, R., Grosskopf, S., & Hernandez-Sancho, F. (2004). Environmental performance: Aan index number approach. *Resource and Energy Economics*, *26*(4), 343-352.

Färe, R., Grosskopf, S., Lovell, C. K., & Pasurka, C. (1989). Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *The Review of Economics and Statistics*, 90-98.

Färe, R., Grosskopf, S., Lovell, C. K., & Yaisawarng, S. (1993). Derivation of shadow prices for undesirable outputs: Aa distance function approach. *The Review of Economics and Statistics*, 374-380.

Färe, R., Grosskopf, S., Noh, D. W., & Weber, W. (2005). Characteristics of a polluting technology: Ttheory and practice. *Journal of Econometrics*, *126*(2), 469-492.

Fernandez, C., Koop, G., & Steel, M. F. J. (2002). Multiple-output production with undesirable outputs: an application to nitrogen surplus in agriculture. *Journal of the American Statistical Association*, *97*(458), 432-442.

Filippini, M., & Greene, W. (2016). Persistent and transient productive inefficiency: a maximum simulated likelihood approach. *Journal of Productivity Analysis*, *45*(2), 187-196.

Filippini, M., Geissmann, T., & Greene, W. H. (2018). Persistent and transient cost efficiency – —an application to the Swiss hydropower sector. *Journal of Productivity Analysis*, *49*(1), 65-77.

Forsund, F. R. (20089). “Good mModelling of bBad oOutputs: Pollution and mMultiple-oOutput pProduction.,” *International Review of Environmental and Resource Economics, 3,* 1-–38.

Frisch, R. (1965). A Survey of the Problems of a Dynamic Theory of Production. In *Theory of Production* (pp. 293-294). Springer, Dordrecht.

Godoy-Durán, Á., Galdeano-Gómez, E., Pérez-Mesa, J. C., & Piedra-Muñoz, L. (2017). Assessing eco-efficiency and the determinants of horticultural family-farming in southeast Spain. *Journal of Environmental Management*, *204*, 594-604.

Gorain, B. K., Kondos, P. D., & Lakshmanan, V. I. (2016). Innovations in gold and silver processing. In *Innovative Process Development in Metallurgical Industry* (pp. 393-428). Springer, Cham.

Graham, D., & Woods, N. (2006). Making corporate self-regulation effective in developing countries. *World Development*, *34*(5), 868-883.

Greene, W. (2010). A stochastic frontier model with correction for sample selection. *Journal of productivity analysis*, *34*(1), 15-24.

Greene, W. H. (2005a). , “Fixed and tRandom eEffects in sStochastic fFrontier mModels.,” *Journal of Productivity Analysis, 23,* 7-–32.

Greene, W. H. (2005b). , “Reconsidering hHeterogeneity in pPanel dData eEstimators of the sStochastic fFrontier mModel.,” *Journal of Econometrics, 126,* 269-–303.

Grindle, M. S. (2004). Good enough governance: Ppoverty reduction and reform in developing countries. *Governance*, *17*(4), 525-548.

Hailu, A., and Veeman, T. S. (2001). , “Non-pParametric pProductivity aAnalysis with uUndesirable oOutputs: An aApplication to the Canadian pPulp and pPaper iIndustry.,” *American Journal of Agricultural Economics, 83*(3), 605-6–16.

Halkos, G. E., & Tzeremes, N. G. (2012). Industry performance evaluation with the use of financial ratios: An application of bootstrapped DEA. *Expert Systems with Applications*, *39*(5), 5872-5880.

Handl, G. (2012). Declaration of the United Nations conference on the human environment (Stockholm Declaration), 1972 and the Rio Declaration on Environment and Development, 1992. *United Nations Audiovisual Library of International Law*, *11*.

Heshmati, A., Kumbhakar, S. C., & Kim, J. (2018). Persistent and transient efficiency of international airlines. *European Journal of Transport and Infrastructure Research*, *18*(2).

Hua, Z., Bian, Y., & Liang, L. (2007). Eco-efficiency analysis of paper mills along the Huai River: Aan extended DEA approach. *Omega*, *35*(5), 578-587.

Hua, Z., Bian, Y., & Liang, L. (2007). Eco-efficiency analysis of paper mills along the Huai River: An extended DEA approach. *Omega*, *35*(5), 578-587.

Kumah, A. (2006). Sustainability and gold mining in the developing world. *Journal of Cleaner Production*, *14*(3-4), 315-323.

Kumbhakar, S. C. (1987). Production frontiers and panel data: an application to US class 1 railroads. *Journal of Business & Economic Statistics*, *5*(2), 249-255.

Kumbhakar, S. C. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of econometrics*, *46*(1-2), 201-211.

Kumbhakar, S. C. (1991). , "The mMeasurement and dDecomposition of cCost-iInefficiency: The tTranslog cCost sSystem.," *Oxford Economic Papers, 43,* 667-–83.

Kumbhakar, S. C., & Hjalmarsson, L. (1995). Labour-use efficiency in Swedish social insurance offices. *Journal of Applied Econometrics*, *10*(1), 33-47.

Kumbhakar, S. C., & Tsionas, E. G. (2016). The good, the bad and the technology: Endogeneity in environmental production models. *Journal of Econometrics*, *190*(2), 315-327.

Kumbhakar, S. C., and Heshmati, A. (1995), "Efficiency Measurement in Swedish Dairy Farms: An Application of Rotating Panel Data, 1976–88," *American Journal of Agricultural Economics*, **77**, 660–74.

Kumbhakar, S. C., and Heshmati, A. (1995). , "Efficiency mMeasurement in Swedish dDairy fFarms: An aApplication of rRotating pPanel dData, 1976-–88.," *American Journal of Agricultural Economics, 77,* 660-–74.

Kumbhakar, S. C., and Hjalmarsson, L. (1993). , "Technical eEfficiency and tTechnical pProgress in Swedish dDairy fFarms. In," Fried H., Schmidt S. and Lovell, C. A. K. (Eds.), *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford University Press.

Kumbhakar, S. C., and Hjalmarsson, L. (1998). , "Relative pPerformance of pPublic and pPrivate oOwnership under yYardstick cCompetition: Electricity rRetail dDistribution.," *European Economic Review, 42,* 97-–122.

Kumbhakar, S. C., and Wang, H. -J. (2005). , "Estimation of gGrowth cConvergence uUsing a sStochastic pProduction fFrontier aApproach.," *Economics Letters, 88,* 300-–5.

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2014). Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis*, *41*(2), 321-337.

Kumbhakar, S. C., Wang, H. J., & Horncastle, A. P. (2015). *A Ppractitioner's Gguide to Sstochastic Ffrontier Aanalysis Uusing Stata*. Cambridge University Press.

Kuosmanen, T., & Kortelainen, M. (2005). Measuring eco-efficiency of production with data envelopment analysis. *Journal of Industrial Ecology*, *9*(4), 59-72.

Lansink, A. O., & Bezlepkin, I. (2003). The effect of heating technologies on CO2 and energy efficiency of Dutch greenhouse firms. *Journal of Environmental Management*, *68*(1), 73-82.

Lansink, A. O., & Silva, E. (2003). CO 2 and energy efficiency of different heating technologies in the Dutch glasshouse industry. *Environmental and Resource Economics*, *24*(4), 395-407.

Lee, J. D., Park, J. B., and Kim, T. Y. (2002). , "Estimation of the shadow prices of pollutants with production/environment inefficiency taken into account: Aa nonparametric directional distance function approach. ," *Journal of Environmental Management, 64*(4), 365-–75.

Lee, Y. H., & Schmidt, P. (1993). A production frontier model with flexible temporal variation in technical efficiency. *The Mmeasurement of Pproductive Eefficiency: Techniques and Aapplications*, 237-255.

Mahdiloo, M., Saen, R. F., & Lee, K. H. (2015). Technical, environmental and eco-efficiency measurement for supplier selection: An extension and application of data envelopment analysis. *International Journal of Production Economics*, *168*, 279-289.

Meeusen, W., & van den Broeck, J. (1977). Technical efficiency and dimension of the firm: Some results on the use of frontier production functions. *Empirical economics*, *2*(2), 109-122.

Mombeuil, C. (2020). Institutional conditions, sustainable energy, and the UN sustainable development discourse: A focus on Haiti. *Journal of Cleaner Production*, *254*, 120-153.

Mudd, G. M. (2007). Global trends in gold mining: Towards quantifying environmental and resource sustainability. *Resources Policy*, *32*(1-2), 42-56.

Mundlak, Y. (1961). , "Aggregation over tTime in dDistributed lLag mModels.," *International Economic Review, 2,* 154-–63.

Murty, S., & Russell, R. R. (2002). On modeling pollution-generating technologies., Department of Economics, University of California, Riverside. *Discussion papers series*, (0(02-14)).

Murty, S., & Russell, R. R. (2018). Modeling emission-generating technologies: reconciliation of axiomatic and by-production approaches. *Empirical Economics*, *54*(1), 7-30.

Murty, S., Russell, R., and Levkoff, S. B. (2012)., "On Modeling Pollution-Generating Technologies.," *Journal of Environmental Economics and Management, 64,* 117-–35.

Neingo, P. N., & Tholana, T. (2016). Trends in productivity in the South African gold mining industry. *Journal of the Southern African Institute of Mining and Metallurgy*, *116*(3), 283-290.

Oliveira, R., Camanho, A. S., & Zanella, A. (2017). Expanded eco-efficiency assessment of large mining firms. *Journal of Cleaner Production*, *142*, 2364-2373.

Oude Lansink, A., & Van dDer Vlist, A. (2008). Non-parametric modelling of CO2 emission quota. *Journal of Agricultural Economics*, *59*(3), 487-497.

Oude Lansink, A., & Van Der Vlist, A. (2008). Non-parametric modelling of CO2 emission quota. *Journal of Agricultural Economics*, *59*(3), 487-497.

Pethig, R. (2006). Non-linear production, abatement, pollution and materials balance reconsidered. *Journal of Environmental Economics and Management*, *51*(2), 185-204.

Pimentel, B. S., Gonzalez, E. S., & Barbosa, G. N. (2016). Decision-support models for sustainable mining networks: Fundamentals and challenges. *Journal of Cleaner Production*, *112*, 2145-2157.

Piot-Lepetit, I., & Vermersch, D. (1998). Pricing organic nitrogen under the weak disposability assumption: Aan application to the French pig sector. *Journal of Agricultural Economics*, *49*(1), 85-99.

Pitt, M. M., & Lee, L. F. (1981). The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of development economics*, *9*(1), 43-64.

Rashidi, K., & Saen, R. F. (2015). Measuring eco-efficiency based on green indicators and potentials in energy saving and undesirable output abatement. *Energy Economics*, *50*, 18-26.

Reinhard, S., Lovell, C. K., & Thijssen, G. (1999). Econometric estimation of technical and environmental efficiency: Aan application to Dutch dairy farms. *American Journal of Agricultural Economics*, *81*(1), 44-60.

Reinhard, S., Lovell, C. K., & Thijssen, G. J. (2000). Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European Journal of Operational Research*, *121*(2), 287-303.

Schmidheiny, S. (1992). Guest Article: The Business of Sustainable Development. *Finance and Development*, *29*(4), 24.

Schmidt, P., &and Sickles, R. C. (1984). , “Production fFrontiers and pPanel dData.,” *Journal of Business & Economic Statistics, 2,* 367-–74.

Serra, T., Chambers, R. G., & Lansink, A. O. (2014). Measuring technical and environmental efficiency in a state-contingent technology. *European Journal of Operational Research*, *236*(2), 706-717.

Shafiee, S., & Topal, E. (2010). An overview of global gold market and gold price forecasting. *Resources Policy*, *35*(3), 178-189.

Smith. J., & John. R. (2010). Globally fit leadership: four steps forward. *Journal of Global Responsibility*. 55-65.

Tamini, L.ota D., Bruno Larue, B., &and Gale West, G. (2012). "Technical and environmental efficiencies and best management practices in agriculture." *Applied Economics, 44*(.13), 1659-1672.

Thiam, A., Bravo-Ureta, B. E., & Rivas, T. E. (2001). Technical efficiency in developing country agriculture: Aa meta-analysis. *Agricultural Economics*, *25*(2-3), 235-243.

Tsionas, E. G., & Kumbhakar, S. C. (2014). Firm heterogeneity, persistent and transient technical inefficiency: A generalized true random-effects model. *Journal of Applied Econometrics*, *29*(1), 110-132.

Tsionas, EM. G., Malikov, E., & Kumbhakar, S. C. (2020). Endogenous dynamic efficiency in the intertemporal optimization models of firm behavior. *European Journal of Operational Research*, *284*(1), 313-324.

Tyteca, D. (1996). On the measurement of the environmental performance of firms – —a literature review and a productive efficiency perspective. *Journal of Environmental Management*, *46*(3), 281-308.

Tyteca, D. (1997). Linear programming models for the measurement of environmental performance of firms – —concepts and empirical results. *Journal of Productivity Analysis*, *8*(2), 183-197.

World Gold Council (2019). *The World Gold Council 2018 Annual Review*. London: World Gold Council Publisher.

Wossink, A., & Denaux, Z. S. (2006). Environmental and cost efficiency of pesticide use in transgenic and conventional cotton production. *Agricultural Systems*, *90*(1-3), 312-328.

Zhou, P., Ang, B. W., & Wang, H. (2012). Energy and CO2 emission performance in electricity generation: A non-radial directional distance function approach. *European Journal of Operational Research, 221*(3), 625-635.

Zofío, J. L., & Prieto, A. M. (2001). Environmental efficiency and regulatory standards: Tthe case of CO2 emissions from OECD industries. *Resource and Energy Economics*, *23*(1), 63-83.

Appendix 1

Appendix Table 1: Description of the mining firms used in the analysis

S/No.	Name of the mine	Country	Region
1	Buzwagi	Tanzania	East Africa
2	Bulyanhulu	Tanzania	East Africa
3	North Mara	Tanzania	East Africa
4	Geita gold mine	Tanzania	East Africa
5	Tshepong operations	South Africa	South Africa
6	Phakisa	South Africa	South Africa
7	Bambanani	South Africa	South Africa
8	Target 1	South Africa	South Africa
9	Doornkop	South Africa	South Africa
10	Joel	South Africa	South Africa
11	Kusasalethu	South Africa	South Africa
12	Masimong	South Africa	South Africa
13	Unisel	South Africa	South Africa
14	Kalgold	South Africa	South Africa
15	Phoenix	South Africa	South Africa
16	Hidden Valley	Papua New Guinea	South Africa
17	Surface dumps	South Africa	South Africa
18	South Deep	South Africa	South Africa
19	Mine Waste Solutions	South Africa	South Africa
20	Mponeng	South Africa	South Africa
21	TauTona	South Africa	South Africa
22	Kopanang	South Africa	South Africa
23	Moab Khotsong	South Africa	South Africa
24	Sadiola	Mali	West Africa
25	Morila	Mali	West Africa
26	Siguiri	Guinea	West Africa
27	Damang	Ghana	West Africa
28	Tarkwa	Ghana	West Africa
29	Iduapriem	Ghana	West Africa

30	Obuasi	Ghana	West Africa
31	Serra Grande	Brazil	Latin America
32	AGA Mineração	Brazil	Latin America
33	Cerro Corona	Peru	Latin America
34	Cerro Vanguardia	Argentina	Latin America

Appendix Table 2: Estimation for efficiencies trend

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	SMLE_TER	SMLE_TEO	SMLE_EER	SMLE_EEO	TRE_TE	TRE_EE
Year	1.42e-06	-2.97e-07	-0.000127	-0.000116	0.00278	0.00416***
	(2.03e-05)	(1.58e-05)	(0.000407)	(0.000346)	(0.00244)	(0.00122)
Constant	0.953***	0.700***	1.065	0.897	-4.780	-7.553***
	(0.0408)	(0.0331)	(0.819)	(0.697)	(4.914)	(2.452)
Observations	340	340	340	340	260	340
Number of id	34	34	34	34	26	34

Standard errors in parentheses

*** indicates significance at 1% level, ** at 5% level and * at 10% level