RESPAresearch

Benchmarking Australian electricity distribution firms: reality, simulations and robustness

1 December 2015

RESPAresearch Pty. Ltd. | www.RESPAresearch.org | info@resparesearch.org | +612 6100 4330 | ABN 73 607 564 974

Contents

1 Context **6**
1.1 The purpose of this report 6
1.2 The common partial cost function 6

2 Theoretical requirements **8**
2.1 Background neoclassical production theory 8
2.2 No uncertainty in the production set for each firm 10
2.3 The production set of each firm is feasible with costless movement 15
2.4 Physical capital use is optimal 18
2.5 All firms face the same input prices 22
2.6 All firms have the same (or similar) production sets 26

3 Statistical assumptions **28**
3.1 Estimated efficiency scores reflect actual firm efficiency 28
3.2 The distributions of error terms are correct 32
3.3 Post-modelling adjustments and the frontier 36

4 Context of the interpretation **39**
4.1 Efficiency estimates in other industries 39
4.2 Uncertainty and complexity 41
4.3 The precautionary principle 43
4.4 Interaction of model uncertainty and the precautionary principle 45

5 Recommendations of RESPAresearch **46**

Disclaimer

The content of this report is provided for information purposes only. RESPAresearch has made every attempt to ensure the accuracy and reliability of the information provided within this report, and provides this report and associated files in good faith, reflecting the knowledge and expertise of the author(s). RESPAresearch does not accept any liability or responsibility for any loss suffered by any person, stakeholder or organisation for the information, or the use of the information, provided in this report or otherwise referenced by it. Nor does RESPAresearch accept any liability or responsibility from the taking of any actions, or refraining from taking actions, as a result of reliance on this report and associated files.

The intent of this report, and others like it, is to inform stakeholders in major decisions about the complexities of the economic and statistical models used; we have ***not*** been paid by any stakeholders to produce this report. A discussion of the general principles and scientific approach used by RESPAresearch can be found at www.RESPAresearch.org/#!principles/f7izx.

Executive summary

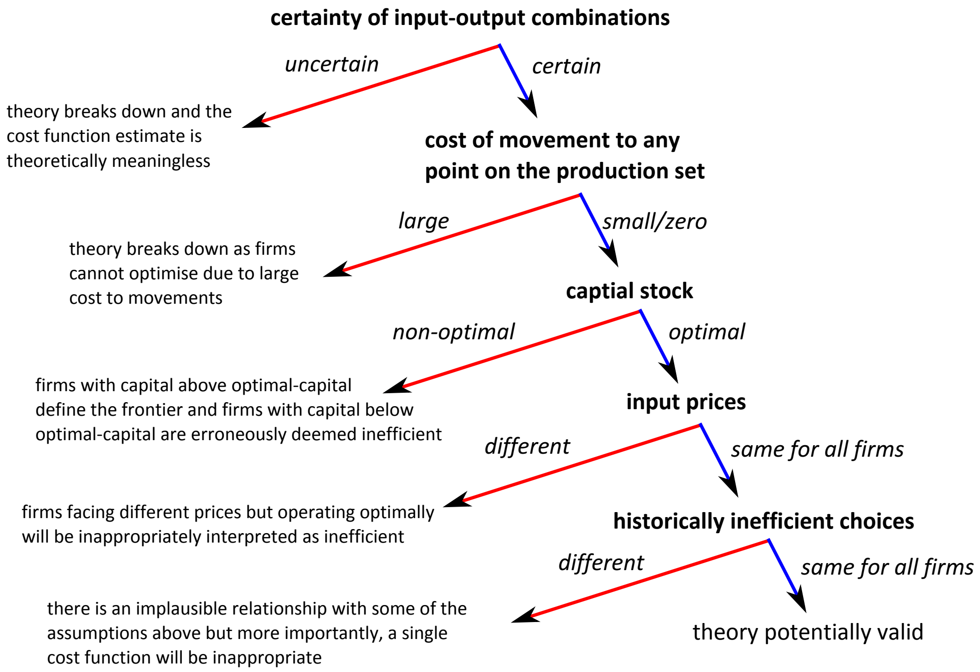
This report, undertaken by RESPAresearch independent of any stakeholder, intends to bring to the attention of all other stakeholders in the Victorian electricity distribution regulatory review process, including the AER, issues of complexity, consistency and robustness implicit in the theoretical and statistical assumptions underlying the operating expenditure benchmarking accepted by the AER.

Five implicit theoretical economic assumptions used in the operating expenditure benchmarking are explored, and the effects of deviation simulated. These assumptions are: (i) certainty of all 'technologies' and all input-output combinations; (ii) costless or low cost movement between any two input-output combinations in the production set; (iii) optimal historical capital decisions; (iv) all firms face the same input prices; and (v) all firms have the same production set.

These theoretical assumptions are discussed in light of evidence and reason, and where appropriate, subject to simulation analysis within the neoclassical framework accepted by the AER. These assumptions are found to be implausible for the following key reasons:

- (i) Firms are composed of humans, and do not have the information processing capabilities of being able to know with (even approximate) certainty, the input-output combinations from using different technologies. Knightian uncertainty, rather than risk, is the primary form of uncertainty firms face.
- (ii) The movement to different organisational structures is assumed zero in theory, but in reality, is highly costly. This inconsistency has a significant effect on the efficacy of the theory, as firms cannot implement the results of that theory.
- (iii) There is no basis for assuming that historical capital investments are, even approximately, optimal. Simulations show that firms with capital stock drawn from a (uniformly distributed) maximum deviation of $\pm 10\%$ from optimal, while having optimal operating inputs, have opex inefficiency estimates of up to 16 efficiency percentage points; capital expenditure inefficiency is erroneously interpreted as operating expenditure inefficiency.
- (iv) The highly cautious use of evidence from the ABS shows that firms in different states face different labour costs, which differ by over 5% from the mean. A simulation of firms with uniform random draws from a price range of $\pm 5\%$ from the mean, and optimal capital and operating inputs, shows that the SFA model interprets price differences of some firms as operating input inefficiencies of over 15%.
- (v) Assuming all firms have the same production set (so a single cost function can be estimated) required the acceptance of one of two highly implausible assumptions (discussed in (ii)).

These assumptions are summarised in the following figure.



Two main statistical assumptions are explored: (i) the statistically estimated efficiency scores accurately reflect the efficiencies of the firms, and (ii) the distribution of the two error terms in the SFA model are consistent with the observed empirical distributions. Simulations demonstrate that the first assumption has 9% of firms resulting in an efficiency score over 10 efficiency percentage points from their actual efficiency scores: this, and two related simulations, demonstrate the lack of robustness of the application of the statistical method in these circumstances. The second assumption fails to hold, however, the size of the effect is small; despite the small size of the effect, distributions with heavier tails retained the large deviations of actual versus estimated efficiency scores.

The theoretical and statistical issues with the model the AER has accepted demonstrate the model is not as robust or reliable as previously thought. We further provide evidence from other industries, some more competitive, which shows that a large range of efficiency estimates is the norm rather than the indication of an 'inefficient industry'. In addition to model robustness, model uncertainty and the precautionary principle need to be properly accounted for in the *interpretation* of the model results. We outline the model uncertainty present in the model accepted by the AER throughout this report by discussing evidence from heterodox economists; we also show an appropriate (qualitative) interpretation of the model results.

The implications of applying erroneous assumptions and models is not a theoretical exercise nor mere semantics: there are significant, real consequences for both firms and customers which, in the theory used, are assumed not to exist.

1 Context

1.1 The purpose of this report

The Australian Energy Regulator (**AER**) has, in this round of electricity distribution reviews, chosen to perform and rely on operating expenditure (**opex**) benchmarking – a form of efficiency analysis – to set the opex allowance for Victorian (and other Australian) electricity distribution firms. The AER has accepted the opex cost function analysis conducted by Economic Insights (**EI**) as robust, reliable and reasonable [1]. The intent of this report is to bring to the attention of all stakeholders in this regulatory process, including the AER, issues pertaining to the complexity, consistency and robustness implicit in the accepted theoretical and statistical assumptions underlying the opex cost function modelling and subsequent interpretation. This report quantifies the effects of a number of these assumptions through simulations (the value of which is discussed in Appendix 1).

Neoclassical economic theory, the theory accepted by the AER, has the underlying philosophical assumption that it accurately or approximately reflects reality. The approach taken in this report, in analysing the assumptions underlying this neoclassical theory, is not one of discussing all neoclassical and statistical assumptions and the plausibility of each: this would lead to an excessively long and complex report. Rather, we have chosen to discuss the major issues necessary for accurate estimation of the common neoclassical partial cost function accepted by the AER. The reason for this 'issues-based' approach is that the analysis accepted by the AER is the reference point; we do not have a clean slate from which to conduct our own analysis.

This report frequently uses the subjective term 'plausible'. This term is used to mean: that which a scientist would find reasonably or approximately reflects reality. The intent of this meaning is to impose a reasonable viewpoint on any analysis. This report has been read and modified in accordance with such a scientific approach.[1]

1.2 The common partial cost function

The AER has accepted and applied the result of a *common partial cost function* using *stochastic frontier analysis* (**SFA**) *estimation* [1]. In carrying out this analysis, a number of implicit assumptions have been made. Section 2 of this report analyses the required theoretical assumptions needed to accurately estimate such a common partial cost function, and Section 3 analyses the SFA model in more detail.

The common partial cost function The AER has accepted an estimate of a single industry-wide partial cost function. This cost function relates the price of a 'unit of opex' ($w_{i,t}$), four outputs (ratcheted maximum demand, customer number, line length, and underground line length[2]: $q_{i,t}$), technique ($A_{i,t}$) and country-specific effects (C_i), to opex, for each firm i at each time t. For the

---

[1] Details of the reviewing are documented in our 'laboratory book' approach to analysis, as discussed at www.RESPAresearch.org//#!constitution/rvp2t.

[2] Considering 'underground line length' as an output is equivalent to the method EI has used, but allows for the more natural interpretation as an output, rather than as an environmental variable. This can be seen by rearranging the model (when written out in full), but can more intuitively be demonstrated by running the 20150728_RESPAresearch_Replication_of_EI_results_and_distbn_analysis.R code (available on our website). The outputted figure shows EI's results alongside the results from this process: the results are numerically identical.

discussion of the theory, the country dummy variables, C_i, will be ignored: this simplifies the analysis without changing any conclusions. The common (logged) partial cost function in question is:

$$\begin{aligned}\ln(opex_{i,t}) &= \beta_0 + \beta_1 \ln(RMD_{i,t}) + \beta_2 \ln(custNum_{i,t}) + \beta_3 \ln(lineLength_{i,t}) + \\ &\quad \beta_4 \ln(UGlength_{i,t}) + \beta_5 \ln(Price_{i,t}) + \beta_6 t \qquad (1.1)\end{aligned}$$

where the only difference between this equation and that accepted by the AER [2] is the inclusion of the country dummy variables and the two error terms.

The SFA estimation of the common partial cost function The SFA estimation of Eq. (1.1) requires two error terms. To avoid any linguistic bias, these terms will be called the symmetric and asymmetric error terms, denoted $v_{i,t}$ and u_i respectively (the interpretation of these terms is discussed in Section 3). These terms are added to Eq. (1.1) to give:

$$\begin{aligned}\ln(opex_{i,t}) &= \beta_0 + \beta_1 \ln(RMD_{i,t}) + \beta_2 \ln(custNum_{i,t}) + \beta_3 \ln(lineLength_{i,t}) + \\ &\quad \beta_4 \ln(UGlength_{i,t}) + \beta_5 \ln(Price_{i,t}) + \beta_6 t + v_{i,t} + u_i \qquad (1.2)\end{aligned}$$

This is now a statistical model which, with a number of further assumptions, can be estimated.

2 Theoretical requirements

The AER has implicitly accepted the theoretical assumptions used in neoclassical production theory. This section discussed these theoretical assumptions, and so the discussion remains largely within the neoclassical realm. This neoclassical economic view of reality is the view taught in most undergraduate and postgraduate courses, but it is by no means the only view in economics. Where appropriate in this section, views and evidence from outside of neoclassical economics are provided; these views are often unknown, ignored or overlooked. As far as is possible within human biases, this report does not ignore nor discount such evidence without valid reasons.

2.1 Background neoclassical production theory

This report uses a production *set* as a primitive; this contrasts with EI, which uses a production function as a primitive [2]. Using this set, two production functions are discussed: the *ex ante* production function and the *ex post* production function. This distinction has significant implications for the theoretical plausibility of the model accepted by the AER.

Definition. (production set) The production set of a firm at a given point in time are all technologically possible combinations of *physical* inputs and *physical* outputs [3].[3]

This definition is a statement of how *physical* inputs are transformed into *physical* outputs: the cost function uses dollar costs but necessarily maintains the assumption that a production set exists in the background. The production set (more precisely, correspondence) can be depicted as a solid polygon in $\mathbb{R}_{+}^{m+n}$ space. (Figure 2.1 demonstrates and discusses this production set when one output and two inputs are assumed.) The production set is a 'black box' definition, in that there is no explanation of how the set has, for example, arisen, how it changes, how it differs between firms, how the properties it possesses have originated; it is a primitive and simply assumed to exist [3].

---

[3] If there are m inputs and n outputs, then the production set can be denoted by $Y_{i,t} \subset \mathbb{R}^{m+n}$. This set can be equivalently defined as a production correspondence, $\mathcal{Y} : \mathbb{R}_{+}^{m} \to \mathbb{R}_{+}^{n}$.

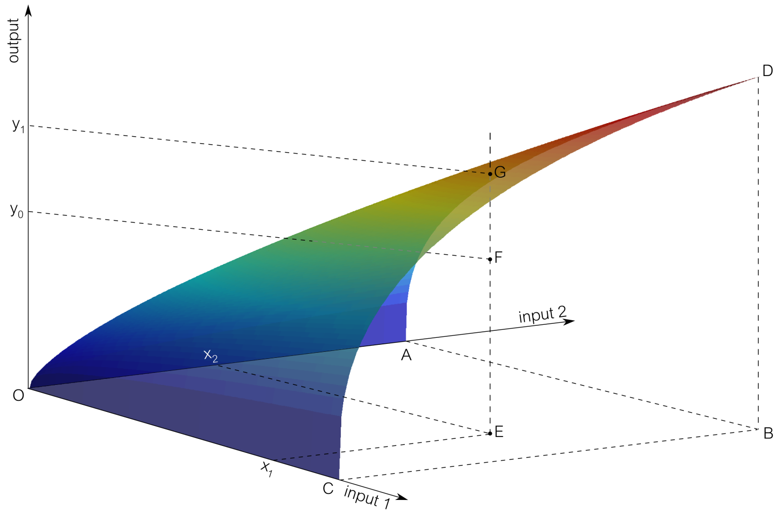


Figure 2.1:
This figure is a stylised physical input-output correspondence. (The AER's estimate has one input and four outputs, but in reality, for electricity distribution firms, there are hundreds or thousands of inputs and a handful of outputs.) The production set is the solid polygon defined by the area between the rainbow curved surface and the input 1-input 2 plane. Given the input limits of inputs 1 and 2, shown at C and A, respectively (for depiction clarity), the production set is the solid polygon defined by the four surfaces: OADC (the rainbow curved surface), ABD, BCD and OABC. Every point within this solid is associated with some combination of input 1 and input 2, some output (determined on the vertical axis), and some choice of technique used to generate that output with those given inputs. A technique is the method of combining inputs to produce outputs, and includes every idiosyncrasy of the firm which affects the possibilities available. The term 'technology' is typically used instead of 'technique', but this term has significant linguistic bias associated as its meaning in economics differs from its common meaning.
Suppose two (almost) identical firms, firm 1 and firm 2, have both chosen to use the input combination (x_1, x_2)*. Firm 1 uses a particular technique of combining inputs to produce output* y_0*, and firm 2 uses a different – and in this case optimal – technique to produce output* y_1*. Firm 2 produces the highest level of output possible from the inputs, given the current techniques available. The rainbow surface is the maximum output a firm can produce at each combination of inputs; it implicitly contains the optimal technique(s) of combining inputs. This surface is the production function. The production function is the upper envelope of all possible techniques for all levels of inputs. It does not necessarily represent a single technique: low levels of output may use one technique, and higher levels of output may use another.*

The production function is the maximum output that can be produced at any arbitrary input combination. The reason why a production function is assumed in neoclassical production theory is because no 'rational' firm would produce at any point in the interior of the production set (i.e. below maximum).

The definition of the neoclassical production set uses the phrase 'technologically possible', meaning that a firm is able to choose to operate at any point in this set. In this neoclassical realm, if a firm can costlessly choose any point in the production set and prices are given exogenously, a firm would choose the input combination that maximises profit. All firms would choose to operate on the production frontier (i.e. use the production function) as otherwise profits would not be maximised. There is result in neoclassical production theory which states that cost minimisation is a necessary condition for profit maximisation [3]; that is, to maximise profits, a firm must be minimising costs. This theorem defines a cost function which would minimise costs at every level of output; this is similar to the cost function estimated by EI. (A stylised cost function is depicted and discussed in Figure 2.2.)

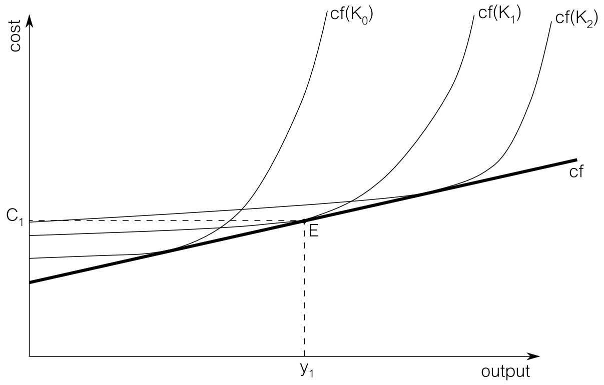


Figure 2.2:
The cost function assumed within neoclassical production theory, cf*, is the minimum cost of production at any level of output. It gives the minimum cost required to produce a certain level of output (e.g.* C_1 *is required to produce* y_1 *units of output). This cost function,* cf*, is the lower envelope of all the individual 'short run' cost functions.*

The AER has applied the results of the cost function estimate, and has stated that these results are consistent with the opex criteria in the National Electricity Rules (**NER**) [4, 5]. This acceptance implies that firms can actually change their input combination to achieve efficiency, and thus must be able to choose points on the production function. That is, all firms are able to choose the physical input-output combinations that minimise costs.

The remainder of this section discusses the plausibility of the implicit theoretical assumptions the AER has accepted by estimating and using the common partial cost function.

2.2 No uncertainty in the production set for each firm

Assumption Managers of each firm, at any given time, know their production set with certainty. This assumption can be weakened so that managers know a 'neighbourhood' of the production set around their current input-output points, including part of the 'local' frontier relevant to their firm.[4]

Context Within the neoclassical realm, under this assumption and some further technical (mathematical) assumptions, the firm is be able to globally optimise.[5] That is, the firm must be able to choose a point on the production frontier otherwise the stylised mathematisation of a firm and the subsequent optimisation breaks down. This assumption means that under all combinations and organisations of different levels of the inputs, including all different levels of operating inputs from all possibilities of the network architecture and capital choices, the outputs are known with certainty. A

---

[4] 'Neighbourhood' and 'local' are interpreted as the relevant part of the frontier for the firm around the required, exogenous level of output.

[5] The word 'optimise' in neoclassical theory is in reference to mathematical optimisation, rather than the common use of the word (such as when managers use the word). This distinction is required to avoid linguistic bias as the intention and meaning of the word are quite different.

weaker interpretation of this assumption is that managers know, or can costlessly know, with certainty, the 'local' space of input-output combinations, which must include part of the frontier at the same or similar level of output.

This assumption is an informational assumption: firm managers have perfect information regarding the levels of output that can arise from the unimaginably large number of input combinations, organised in an unimaginably large number of different ways. An equivalent interpretation is that firm managers have perfect information processing capabilities. Even in the interpretation that full information is 'local', the requirement to be able to process an unimaginably large number of choices, choices which may or may not lead to the 'local' frontier, remains.

A reasonable assumption in this context is that risk preference should be common across managers and should account for any uncertainty in firms' production sets. This requires the production sets to be the same across firms (see Section 2.6), but more fundamentally, requires the uncertainty to be classified as risk. In neoclassical production theory, all uncertainty is risk. However, uncertainty can be classified into three types [6]: (i) *risk* is an uncertainty for which the outcomes are known, and to which objective probabilities can be assigned (e.g. through historical experience); (ii) *weak Knightian uncertainty* is where outcomes are known but to which objective probabilities cannot be assigned;[6] (iii) *strong Knightian uncertainty* is where some outcomes are themselves unknown, and thus necessarily probabilities cannot be objectively assigned.[7] Risk is the simplest form of uncertainty, and the one assumed in neoclassical production theory, as it is mathematically tractable. The weak form of Knightian uncertainty requires subjective probabilities be assigned to the outcomes, and the strong form of Knightian uncertainty cannot assign even subjective probabilities since there are unknown outcomes.

Implication(s) if the assumption does not hold If there exists uncertainty in knowledge of the input-output combinations and the techniques used to produce those combinations, then neoclassical profit optimisation breaks down, and so the theoretical relationship between the production set, profit maximisation and cost minimisation breaks down. Thus the cost function estimate will bear little or no relationship to what firm managers can actually implement.

Reality Managers of firms have limited, biased human processing capabilities, and information constraints, even in the weakened version of this assumption [7]. In reality, in any decision making process, only a handful of scenarios are considered, and even these will suffer from subjective biases and uncertainties. In the context of operations management, numerous textbooks and publications advocate using the right tools for an uncertain environment, including decision analysis and scenario planning (for example, [8]): this is how managers are taught to consider uncertainty in planning.

The 'local' interpretation of this assumption still suffers from imposing full information in a 'local' area. That is, full information is still required for managers to be able to know *how* to get to the frontier (i.e. the techniques to employ). Recall that the production set is a 'black box' definition: it does not provide information of the 'techniques' required to get to any point in the set. The definition

---

[6] Humans are generally biased statisticians without previous experience of events and without computational power; this was researched extensively by Kahneman and Tversky (see, for example, [7]).

[7] [6] uses the term Knightian uncertainty for what we call 'weak Knightian uncertainty' and 'ignorance' for what we call 'strong Knightian uncertainty'.

simply assumes firms know, and have access to, the techniques to be able to produce any input-output combination in the set.

It is worth noting that a slight movement in the production set does not imply a slight change in the operation of a firm (i.e. the 'technique' chosen to combine inputs to produce outputs), as it may require significant reorganisation of inputs. This is a consequence of the production function being the upper envelope of all techniques.

Box 1 presents a contrast to neoclassical production economics: it provides a summary of four case studies of how firms *actually* operate and *actually* make decisions. Two relevant findings are that firms do not operate as if they had a neoclassical production function, and managers attempt to avoid (Knightian) uncertainty.

The risk preference of managers is not the main concern in decision making under uncertainty, as managing a firm is subject to all three forms of uncertainty. There are many instances which can plausibly be categorised as risk within a business; there are also many which are more accurately assigned Knightian uncertainty, either weak or strong:

> ...for investors, business folk, government officials, physicians, international diplomats,...risk is an intriguing subject that bears little relation to the real decisions they face. Unknown outcomes confront these players every day, and the probabilities are virtually never known or knowable. Uncertainty, not risk, is the difficulty regularly before us. ...
>
> More importantly, uncertainty characterizes much more of economic activity than does risk. ... Risk is much less important than uncertainty [6]

A survey of 1000 Austrian entrepreneurs found that 41% view future sales, a simple and intuitive variable, as most closely related to Knightian uncertainty [9]. Thus the argument that managers should have equal risk preference is irrelevant. The existence of Knightian uncertainty, in either of its forms, does not mean that, in reality, firms cannot form strategies to reign in these forms of uncertainty. However, formalising these problems remains generally intractable which is why they are not considered in neoclassical production theory. In the main, these strategies are informal and generally inconsistent with neoclassical production theory:

> A decision maker should always be aware of the factor of [strong Knightian uncertainty] and should try to draw inferences about its nature from the lessons taught by history, from experiences recounted by others, from accounts given in the media, from possibilities developed in literature, etc. [6]

This quote comes from the Handbook of the Economics of Risk and Uncertainty, written in 2014, yet the study described in Box 1, published in 1963, makes similar, empirically founded claims related to firm decision making. This evidence has been around for more than half a century; it cannot be overlooked or ignored, as it is how managers *actually* make decisions.

An issue related to uncertainty is the use of the production set; if firms do not know it, then they cannot use it. The human brain and the idiosyncratic nature of electricity distribution firms, and firms in general, means that moving the operations of a firm from one operational structure to a different one cannot be plausibly classified as risk. The outcomes, in terms of possible outputs, may be known but the probabilities assigned to each outcome cannot be assigned in an objective way: there is no historical basis from which such probabilities can be assigned. The experiences of other firms in different economic, political and geographical climates, may provide some insight but any assignment of probabilities will always remain subjective.

Even under the assumption that a firm's managers can costlessly reorganise the entire set of inputs, including how they relate to one another, it is not an inconsequential informational assumption that management knows, or, through consultants, can know, the objective probabilities of the outputs the firms will attain from the unimaginably large number of possible combinations of inputs. Through the case studies discussed in Box 1, managers in unregulated and significantly more competitive industries make decisions under uncertainty in similar ways.

The large opex reductions imposed by the AER on some firms will necessitate major restructuring [10]. Restructuring an entire business due to an application of neoclassical economic theory, an 'exogenous shock' in economic diction, implies a significant departure from 'business as usual' heuristics, and lies in the realm of decision making under Knightian uncertainty. Search for more efficient structures is likely to be undertaken but the neoclassical production function and neoclassical theory is likely to play no role in how managers can, and do, respond. Closely related to this point is the issue of application of a model under model uncertainty: this is discussed in Section 4.2.

Box 1. A summary of *A behavioural theory of the firm* [11]

This monograph is a behavioural case study and theoretical proposal of how actual firms make decisions, day-to-day and under uncommon circumstances. The motivation for this study was the dissatisfaction with two facets of neoclassical theory: (i) the motivational and cognitive assumptions made of the firm seem unrealistic, and (ii) there were few similarities of the theory which could be identified with actual businesses. This study explains why there has been little success in merging organisational theory (i.e. how firms *actually* organise themselves) and neoclassical production theory. This study takes the firm as the basic unit, and looks at: (i) how organisational goals arise and change; (ii) decisions on the timing of search and processing of information; (iii) how choices are ordered and decisions made; and (iv) the relationship between decision and implementation.

The study is an in depth look at four major decisions within three firms in the industries of manufacturing and construction: accelerated renovation of old equipment, new working quarters for a department with a doubtful future, selection of a consulting firm, and choosing a data processing system. The general findings were used to create the following sequence of events the firms followed in making decisions:

1. A firm has multiple, dynamic acceptable goals. There is not a single profit-maximising goal for all parts of the organisation and for all time.[a] The criterion for choosing goal(s) at any point in time is one which best meets all goals of the coalitions within the firm. These are goals of the individuals within the firm, and they too change depending on the environment.

2. The firm undertakes an approximately sequential consideration of alternatives, with the first satisfactory outcome chosen. When the existing policy (i.e. 'business as usual') satisfies the goal(s), little search is undertaken for alternatives; otherwise search is intensified over time.

3. Firms attempt to avoid uncertainty by following historically successful procedures, relying on a policy of feedback rather than forecasting the (uncertain) future environment.[b]

4. Firms follow standard operating procedures and heuristics to make choices, with these rules dominating the decisions made in the short run. Standard operating procedures change over time at varying rates depending on external shocks and internal motivations. The structure of the standard operating procedure arising from the 'general choice procedure' which the firms under study followed, can be summarised as: (i) minimise the need for predicting uncertain future outcomes; (ii) maintain a feasible set of decision rules that are only abandoned under duress; and (iii) maintain simple rules with some conditional flexibility.

The authors of this study also formalise this process, and provide a formal model of firm decision making and operation: the theory proposed in this monograph deviates significantly from neoclassical theory. A related monograph is [13] which discusses related aspects of how firms make decisions.

---

[a] Even if there were a profit maximising goal, [12] provides a case study using managers in the National Football League to see if optimal, profit maximising choices are made under situations more akin to risk, in a data-rich environment. The analysis of decisions made by these highly paid managers (average income of $3mil), in a highly competitive labour market (turnover of 20%), shows systematic, clear-cut, and overwhelmingly statistically significant departures from decisions which would maximise teams' profits.

[b] This uncertainty avoidance (or ambiguity aversion) is consistent with research conducted within behavioural economics [14].

Plausibility The assumption of zero or low uncertainty in the production set, if taken to reflect reality even approximately, is implausible.

2.3 The production set of each firm is feasible with costless movement

The discussion in this section assumes managers have perfect information about the production set (i.e. Section 2.2).

Assumption The production set of each firm at each point in time consists of all feasible combinations of inputs and outputs for which the movement between any two points in the set is costless. This assumption can be weakened to 'low cost' instead of costless, and to movement to the frontier rather than between any two arbitrary points.

Context This assumption can be disaggregated into two implicit assumptions: (i) a firm can actively choose a different technique with which to combine inputs to produce outputs, and (ii) the choice to change techniques is costless. These implicit assumptions are necessary in neoclassical theory as this theory is used to model how a firm *optimally* operates, meaning all points in the production set are feasible and have the same cost of implementation (typically zero). This theory is an *ex ante* theory of production, where firms choose inputs and technique *de novo* to produce outputs. For the neoclassical theory to hold, the implicit assumptions, (i) and (ii), are mathematically necessary to be able to ignore any initial input and technique 'endowment' (i.e. the fact that a firm currently exists, and has a current form of operation). These two implicit assumptions remove any path dependency in the historical choices of a firm.

Recall that a 'technique' is the way in which inputs are used to create outputs, and includes every idiosyncrasy of a firm's environment. A technique may be optimal at one point in time but be superseded by a different technique in the future. The choice and organisation of capital inputs is of significance: the choices of voltage levels, of capacities, of the physical architecture of the network, of the choice of depots, of the historical labour contracts, are all choices which, in part, determine the technique used to combine inputs to produce outputs. These historical choices are difficult and costly to reverse. If an electricity distribution network is to be on the frontier of the production set, it must have made all historical choices, including the technique used, optimally. If a firm has made inefficient historical decisions, for neoclassical theory to be retained, that firm must be able to costlessly choose different inputs and a different technique so as to be able to operate on the frontier of the production set. In reference to Figure 2.1, these assumptions means that a firm can move from point F to point G by choosing a different technique, and can do so costlessly.

It may be argued that in the economic long run, all factors of production are variable and so the efficient capital choices and organisation will be achieved by a competitive firm, and so should be achieved by a regulated firm.[8] This argument relies on the production set being stationary over time, but more importantly, on the ability of a firm to entirely redesign the interaction between inputs.

An alternative to the strict application of neoclassical theory is the view that the production set is an *ex post* production set. That is, as defining all input-output combinations which, given the

---

[8] It must be noted that empirical estimates of efficiency in significantly more competitive industries have a wide range: this is discussed in Section 4.1.

historical path of capital and technique choices, can be implemented costlessly or at low cost. This interpretation would maintain consistency with neoclassical production theory.

Implication(s) if the assumption does not hold The optimisation used in the neoclassical theory ignores any significant costs, or barriers, to moving between points in the production set. If such costs do exist and are prohibitively large, then any 'optimal' theoretical results would be inconsistent with cost-feasible options available to the firm.

Reality In each time period, firms do not make the choice of inputs, technique and outputs *de novo*: a firm does not face an *ex ante* production set, but rather faces an *ex post* production set – a production set in which history has locked in constraints, such as voltage levels, network design and architecture, etc. As [15] notes:

> [W]ith regard to the time dimension...in the study of the production process a fundamental distinction must be made between *ex-ante* analysis and *ex-post* analysis. ...
>
> [*E*]*x-ante* analysis corresponds to a 'plan' of the production process, on the basis of given hypotheses on agents' behaviour and market structure, assuming as given the availability of inputs. ...
>
> By contrast, *ex-post* analysis aims to elucidate the productive characteristics of a given process actually in operation.

Existing firms reflect the constraints of history in their production sets. Assuming that firms can change their techniques in a costless way is contrary to the historical choices of long-lived capital of electricity distribution firms, choices made decades ago via likely inefficient and *ad hoc* practices. Even labour contracts have a historical path dependence, with any changes being potentially costly. A network with such historical inefficiencies would be unlikely, at the current time, to be able to reach the frontier unless inputs could be entirely redesigned and renegotiated in an optimal way (a likely high-cost process).

To maintain the neoclassical framework, the production set needs to be interpreted as a costless (or low cost), *ex post* production set, one that takes into account the costs associated with movement within this set.[9] Figure 2.3 demonstrates and discusses a realistic interpretation of a production set while remaining within the neoclassical framework. Under this interpretation, the production function defined by this production set is one that a firm can actually achieve at low cost, and is not simply a theoretically 'possible' production function.

---

[9] A possible interpretation of 'low cost' may be the net present value of benefits to consumers of the use of a redesigned organisation, relative to the current organisation.

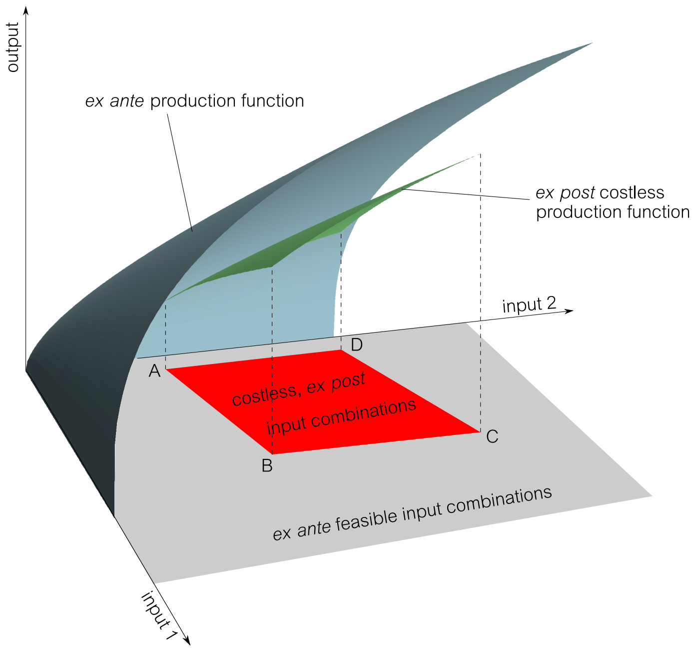


Figure 2.3:
The ex post *production set is the set of input-output combinations that a firm can costlessly (or with low cost) choose to operate within. This set takes into account costly-to-reverse historical decisions. It is implausible to assume coincidence of this* ex post *production function with the relevant part of the current* ex ante *production function, since this would require historical decisions to coincide with the decision which would be optimal at the current point in time. In this stylised figure, for any costless* ex post *input combination (i.e. any point in the red polygon, ABCD), the green surface is the maximum this firm can produce given the historical constraints. The* ex post *costless production set is the solid polygon defined by the green surface and the red polygon ABCD, which defines the feasible inputs (e.g. some inputs are required due to contracting, etc.).*

If a network has costly-to-reverse historical technical inefficiencies, as is reasonable to expect in reality, to achieve a certain level of output, relative to the *ex ante* production set, a greater use of inputs is necessary. This is a necessity arising from the neoclassical theory of production, and is discussed further in Section 2.4.[10]

The argument that the optimal design will emerge in the long run is a blind application of an economic definition to a situation unlikely to warrant its use. It is highly implausible that an electricity distribution firm would ever reach a point in time where it would choose to redesign its entire network. Decisions are always made in the short run with some factors of production fixed; decisions related to capital are always dependent on existing capital and designs (i.e. the long run is infinite). The capital in electricity distribution is long-lived in nature, and the replacement cycles of different capital

---

[10] This contrasts to the way the AER has used the theory [4]:

> If the [distribution firm] entered into a long term inefficient contract, we would be required to include the associated costs in our forecast. These decisions are not part of our role and such an approach would be contrary to the incentive basis for the regulatory regime.

The AER does not take into account any historical inefficient decisions, and therefore has assumed an *ex ante* production function.

is staggered due to different economic lives. Thus the argument that in the long run a firm will tend to the efficient design is invalid.

The AER has implicitly assumed costless movement in the production set by accepting the cost function estimate, while at the same time being unconcerned about the cost of transition to a more efficient level of production [16]. This apparent contradiction is of concern, and is why we have taken the initiative to inform stakeholders of such inconsistent assumptions.

Plausibility It is highly implausible to expect a firm to costlessly (or at low cost) redesign its network in an optimal way. It is, however, plausible to interpret the production set as a low cost, *ex post* neoclassical production set.

2.4 Physical capital use is optimal

The discussion in this section assumes managers have perfect information about a production set in which movement between any point and the frontier is costless (i.e. Sections 2.2 and 2.3).

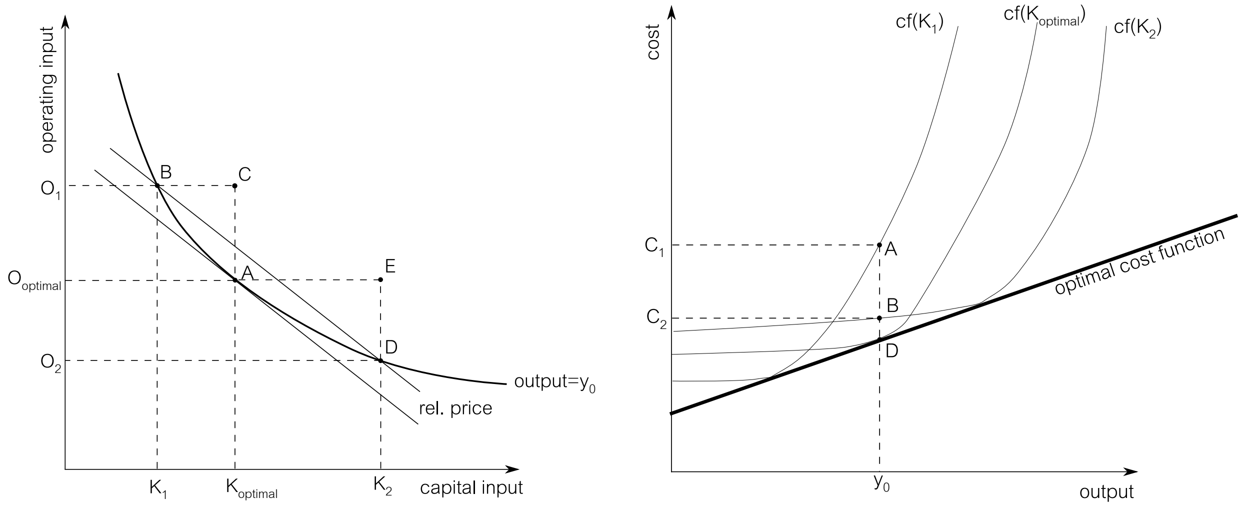
Assumption Given identical exogenous prices faced by all firms, all physical capital inputs are at a level that, if a firm were to make optimal decisions reflecting *all* inputs, that level of physical capital would be optimal.

Context The ability to accurately estimate a *partial* cost function requires an assumption on the factors of production *excluded* from the analysis. This assumption, in the context of excluding the whole range of physical capital goods, is that those capital goods are assumed to be at their optimal level. This assumption is not often used in economic analysis, and has been implicitly chosen by EI in their approach [2]. The reason for this assumption seems to be that the AER desired to pursue, and have regard to, economic benchmarking of opex, as allowed under the NER [17]. This assumption is not based on neoclassical economic theory or evidence, but rather seems to be an artificial constraint chosen by EI arising from their preference for partial cost function benchmarking.

This imposition of optimal physical capital assumes away any capital inefficiency, in terms of quantity, organisation and interaction with all other inputs, when evaluated at the optimal level of all inputs — this includes at the level of the optimal operating inputs. Figure 2.4 demonstrates and discusses the theoretical context implied by this assumption in a stylised setting, and demonstrates how historical difficult-to-reverse choices can affect the estimation and interpretation of efficiency estimates. This figure and analysis assumes substitutability between capital and operating inputs, as accepted by the AER [18].[11]

---

[11] The discussion in this figure relies on an understanding of neoclassical economics taught at the undergraduate level (see for example, [19]).



(a) (b)

Figure 2.4:

(a) This is a stylised example with only two physical inputs: capital input and operating input. The assumption of optimal capital means that each firm operates at the optimal level of physical capital, requiring the firm to be on the frontier of the production set and optimising relative to prevailing prices: this level is denoted $K_{optimal}$. At $K_{optimal}$, the optimal quantity of operating inputs is $O_{optimal}$. If the production function is accurately estimated, then this is the level to which efficiency comparisons are made. Consider two firms, 1 and 2: firm 1 has lower than optimal capital, at K_1. It chooses optimal operating inputs given prevailing prices (i.e. at B). Relative to the optimal cost function, capital inefficiency will be incorrectly viewed as operating input inefficiency: the quantity of operating inputs CA will be viewed as inefficient, despite the firm being perfectly efficient given historical capital choices. In contrast, firm 2 has higher than optimal capital, at K_2, and chooses optimal operating inputs, O_2. In this case, capital inefficiency *will be interpreted as operating efficiency: the quantity of operating input ED will be interpreted as a firm being a 'frontier' firm since it is using fewer operating inputs. Suppose these two firms were data points in an econometric estimation. Since both firms are producing the same output, but firm 1 has significantly higher (but still efficient) operating costs than the firm 2, then firm 1 will be incorrectly interpreted as being inefficient. The issue is that the data is inconsistent with the assumptions, thus the theory, and subsequent interpretation, is distorted.*

(b) This figure is a modification of Figure 5.D.7 of [3]. These cost functions each reflect the level of fixed capital inputs. Both firms in (a) seek to produce y_0 units of output, so the optimal capital would be $K_{optimal}$, however, capital is difficult to change. Instead, the firm with lower capital requires higher operating inputs, while the firm with higher capital requires lower physical operating input. The cost BD is the net additional cost to the firm with higher than optimal capital, but lower than optimal operating inputs. Similarly, AD is the net additional cost to the firm using lower than optimal capital but with higher than optimal operating inputs.
The optimal cost function is the bold line: this is the object in want of estimation. If all firms have optimal capital stock, then all efficient firms will operate on this line, and all firms with inefficient operating input levels will operate above it. Data generated in this way would be consistent with the estimation accepted by the AER. However, if the firms do not have optimal capital inputs, then the data will be above this optimal cost function and the level of operating input inefficiency will not be able to be distinguished from capital inefficiency. In fact, the case shown in this figure is plausible, where two similar firms with similar outputs have significantly different operating input costs, despite both using those operating inputs perfectly efficiently given their historical, difficult-to-reverse capital choices.
It is a simple extension of this framework to demonstrate that a firm can produce more output with less operating inputs than another firm, despite both firms being efficient conditional on their historical, costly-to-reverse choices. (This is a plausible explanation of the large dispersion of the Partial Performance Indicators and Category Analysis used by the AER.)

Implication(s) if the assumption does not hold If firms are not using the optimal level of capital, then the data used in estimating operating input efficiency will contain levels of both capital and operating input inefficiency leading to distorted and inappropriate comparisons between firms, and so comparison of efficiency scores will not reflect reality.

Reality Decisions about physical operating inputs and changes to capital inputs are generally made at the same level of management within a firm. Capital inputs are a result of multiple prior and current management teams making decisions, and are difficult to reverse, whereas operating inputs are less difficult to change but still possess, to a lesser degree, decisions made by prior management teams. Historical capital input and organisation (i.e. 'technique') decisions were made in the past by management in an environment likely containing less information than is currently available. It is thus unlikely to be efficient.

The assumption of optimal capital requires that all firms be perfectly efficient in a major area of operations (highly dependent on past decisions), while concurrently being potentially inefficient in another major area of operations. This is a highly unrealistic assumption required to ensure the mathematics of neoclassical production theory is consistent rather than being based on how firms operate, or any other evidence. It is even more unlikely that the historically chosen and organised factors of production, factors a firm cannot easily change at low cost, are perfectly efficient while the factors of production that a firm is able to change relatively easily are subject to inefficiency. It is unreasonable that, for example, optimal subtransmission voltages are different in different jurisdictions, yet this is the implicit assumption necessary to retain the validity of the theory. This is but one example where there are differences in historical capital choices across firms, all of which are implicitly assumed to be optimal for these firms.[12]

The model accepted by the AER includes circuit length, underground line length, customer numbers and ratcheted maximum demand: these are poor proxies for either (i) the different types of capital different firms use, which is the idea discussed in Figure 2.4, extended to hundreds, possibly thousands, of dimensions to account for the different observed capital choices of firms (e.g. the historical voltage level choices to which firms are locked into), and (ii) any aggregate of the capital stock, as the Cambridge capital controversies essentially showed that aggregation of capital has critical consequences to maintaining neoclassical production theory.

Simulation We have conducted simulation analysis of how capital inefficiency can be invalidly interpreted as operating inefficiency in cost function estimates. This simulation assumes 1000 firms operating using two inputs (capital input and operating input) with a common constant returns to scale Cobb-Douglas production function, each with an output requirement chosen randomly from the interval $[1, 100]$. The price of both operating and capital inputs is set to 1. Given the output of each firm and the relative price, optimal capital is determined and then set to be a random value in the interval $[0.9 \times \mathit{optimal\ capital}, 1.1 \times \mathit{optimal\ capital}]$, i.e. a maximum of $\pm 10\%$ away from

---

[12] A note is required here that the AER also made adjustments to the capital expenditure of some electricity networks in their April 2015 decisions [1]. This is somewhat inconsistent with the assumption made in this operating expenditure benchmarking analysis that firms make optimal capital choices. If the AER is to be consistent with this assumption, it requires the firm to have made, and likely to make, optimal capital choices. Consistency with this assumption has a significant effect on any capital expenditure proposed by these same firms.

the optimal capital level, uniformly distributed. Given this non-optimal capital and the relative price, the optimal level of operating input is determined (as in Figure 2.4). The operating costs are then calculated for each firm. This process gives the operating input cost and physical output data for all firms. The SFA model, similar to that accepted by the AER, is run and the corresponding efficiency analysis conducted. These efficiency results are shown and discussed in Figure 2.5.[13]

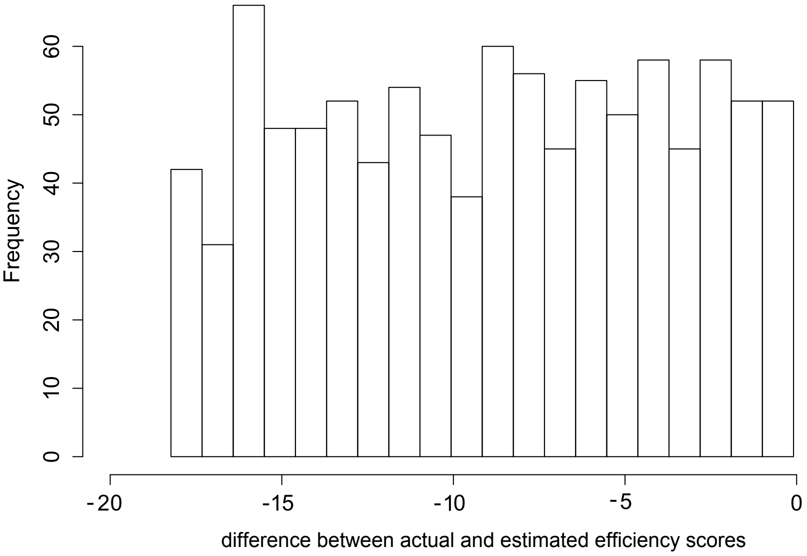


Figure 2.5:
This figure shows a typical histogram of the difference between the actual operating efficiency of the firm – all firms are perfectly efficient in operating inputs, by construction – and the estimated operating input efficiency in the presence of capital inefficiency. The median firm is typically within 0.5% of the optimal level of capital, the maximum deviation from the optimal capital of any firm is less than 10%, and all firms have perfectly efficient levels of operating inputs given their respective capital inefficiency. The difference between actual and estimated efficiency gives a non-trivial range, typically around $[-18.2, 0]$*. As expected given the discussion in Figure 2.4, all firms with capital above the optimal level have efficiency scores deviations above the median deviation, and all firms with levels of capital below the optimal amount have efficiency score deviations below the median deviation.*
In the SFA analysis, firms with higher than optimal-capital define the 'frontier' of the cost function, which is then erroneously reflected in those firms with less than optimal-capital as operating input inefficiency: that is, capital inefficiency of one firm affects the operating input efficiency estimates of other firms.
These results are robust to changes in the range of output required by the 1000 firms, and to the distribution of the output requirements of those firms (results not shown).
The code for this simulation is available online as 20150728_RESPAresearch_Non_optimal_capital_and_different_prices.R.

The lessons in application from this simple simulation are: (i) capital inefficiencies – both higher and lower than optimal capital stocks – have a significant effect on the statistical estimation of operating input efficiency levels, even if operating input choices are perfectly optimal given the capital stock; and (ii) a property of the median or mean firm (in this case, optimal capital) ignores the issue that the distribution of that property matters, and that this distribution may have a significant effect on the estimation of efficiency scores.

---

[13] Note that there is no measurement ('symmetric') error, and that all reported operating input inefficiencies are not operating inefficiencies at all; rather, they are capital inefficiencies resulting from inefficient historical choices.

Plausibility It is implausible to assume that, in reality, firms have made historically optimal capital choices. In the AER's use of the results, five out of 13 firms are over 40% inefficient in operating input costs (i.e. operating inputs can be reduced by 40% without changing output) yet these same firms are assumed to be 0% inefficient in their use of capital inputs. This result is highly unreasonable. There is no evidence for this assumption and it seems to be imposed to retain neoclassical economic theory without modification. Further, assuming perfect capital efficiency is likely to inappropriately attribute historical physical capital input inefficiencies to current physical operating input efficiency estimates.

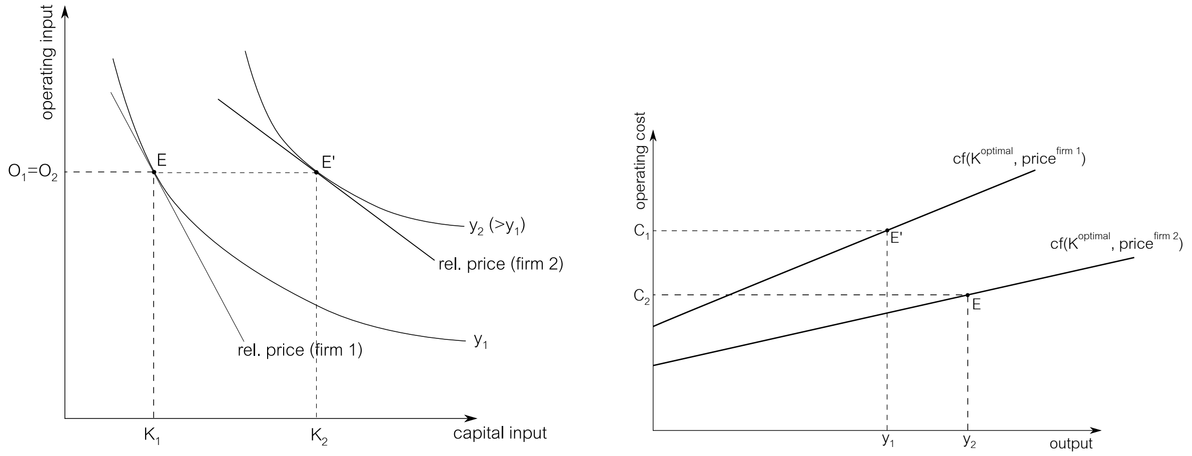
2.5 All firms face the same input prices

To remain within the neoclassical framework, the discussion in this section assumes managers have perfect information about a production set in which movements between any point and the frontier are costless, and all firms have optimal capital (i.e. Sections 2.2 to 2.4).

Assumption All firms face the same, or very similar, input prices.

Context Firms are assumed to face the same input prices so that two firms with the same output requirements would choose the same inputs and so have the same costs. This assumption is effectively assuming all input markets are perfectly competitive with no imperfections, such as transport costs or geographic labour preferences, the purpose of which is to make the modelling mathematically tractable.

Within the neoclassical realm, if two identical firms faced different input prices, then their cost functions would also be different, even under optimal capital choices. Consider the simple case where a firm faces input prices 5% higher than that of another firm: the cost function would then be 5% higher for that firm, and so the resultant estimate of a single common cost function would incorrectly find that firm inefficient. This idea is extended and discussed in Figure 2.6. As the AER has accepted a single industry-wide cost function, the possibility of different firms facing different input costs and thus having different cost functions is, by assumption, excluded.



(a) (b)

Figure 2.6:
(a) This figure shows the result when different firms face different relative prices. Firm 1 has output requirement y_1, and the relative price it faces means that the optimal quantity of operating outputs is O_1. In contrast, firm 2 has output requirement y_2 which is larger than y_1, but faces a different relative price. At this relative price, the optimal choice of operating inputs for firm 2 is O_2. If only operating input was to be analysed (and prices assumed to be the same for both firms), then firm 1 would seem to be interpreted as being inefficient as it produces less output at the same level of operating input, relative to firm 2.

(b) This figure shows one possibility of different cost functions, dependent on the magnitudes of the absolute and relative price differences. In (a), firm 1 faced a lower relative price for operating inputs (and so used more of those inputs) but the absolute price may have been higher than that of firm 2. This would result in a cost function above that of firm 2 (as in this figure). The precise differences in absolute and relative prices will determine the precise nature of the differences in the cost functions the firms face: they could be as shown in this figure, with different slopes, or reversed, or intersecting.
If input prices differ (in absolute terms) across firms, then different firms will face different cost functions, and so estimating a single cost function will unfairly interpret price differences as inefficiencies.

Implication(s) if the assumption does not hold If different firms face different input costs, relative or absolute, then they will also face different cost functions; estimation of a single cost function will inappropriately interpret price differences as inefficiencies.

Reality Perfectly competitive markets are an academic economic abstraction; they are rarely, if ever, encountered in reality. This is especially true for labour markets. For any specific type of labour, wages differ significantly across geographical regions and individual skills, and depend in part on the geographic preferences of individuals, the availability of skilled labour, individual firm practices, etc. Labour markets are slow to react, with wages sticky for common sense reasons [20]: it is a strong assumption to make that such a market clears at a single wage across different states.

Under the strong aggregation assumption of a single labour market for all different labour inputs in the electricity distribution industry, the AER has in the past, and even in the latest determinations for NSW and ACT firms, set different labour cost escalators, and thus accepted differences in future labour costs [21]: this is not an unreasonable nor an unrealistic course of action taken by the AER.

In the latest cost escalation estimates, even if we implausibly assume absolute prices in the base year are the same, the AER has still accepted that there will be labour cost differences between NSW and the ACT in the range of $[0.70\%, 0.97\%]$. Figure 2.7 shows and discusses in detail differences in earnings between the states and territories in the National Electricity Market (**NEM**). Despite the publicly available data not being ideal, we have accounted for much of this in our analysis: the details are discussed in Figure 2.7.

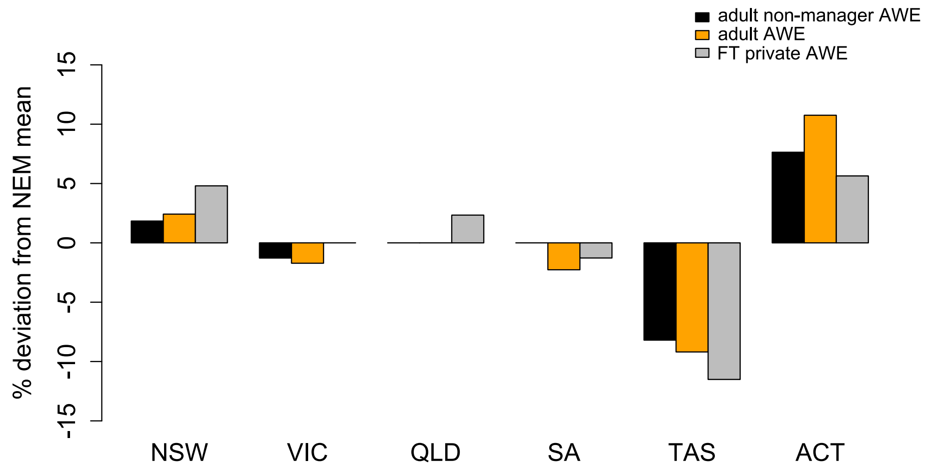
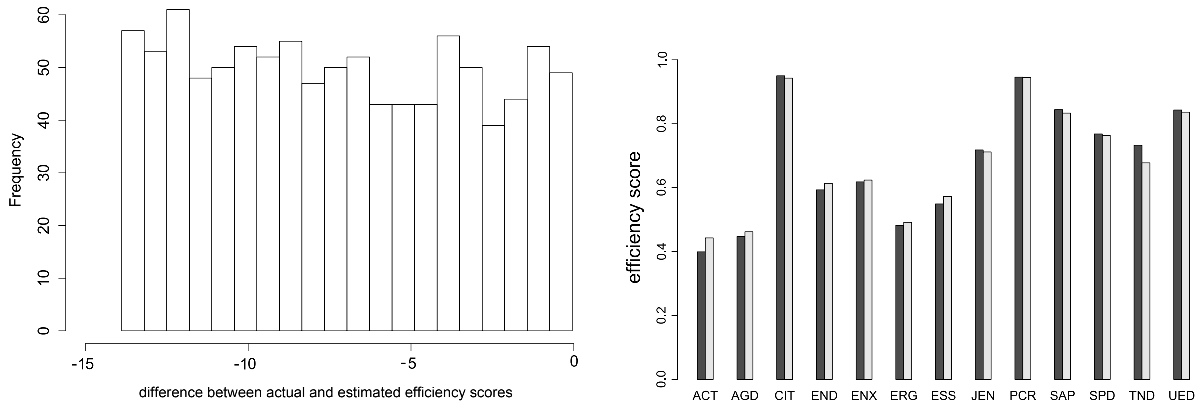


Figure 2.7:
*This analysis uses data from the ABS on average weekly earnings (**AWE**) across Australian states and territories [22]. Three related but different data sets are used to choose the one that provides a minimum range (i.e. cautionary use of the data). The three data sets used are: adult non-manager AWE, adult AWE, and full time private sector AWE. The data includes standard errors for each state/territory. A deviation of two standard errors is taken in both directions from the estimated AWE, and then a program finds the point within these intervals for each state/territory that minimises the sum of the absolute percentage deviation from the mean AWE of the relevant sample. This is done to, again, be cautious in the use of the data. States/territories not in the NEM are excluded resulting in an even smaller range, meaning an even more cautious approach. Despite all these precautions, these results show that labour costs do vary between states/territories in the NEM. The intersection of the ranges of the three data sets, in percentage deviation from the mean, is* $[-8.20, 5.64]$*; the intersection is taken so as to be even more conservative in the use of the data.*
The AWE survey data is, in this context, used to measure the average level of earnings between Australian states and territories. This data is not perfect, and has the drawback of being affected by both the level of earnings per employee and the composition of the labour force in each state/territory [23]. Despite this drawback, this data, especially when used cautiously as described above, does provide insight into differences in dollar earnings, and thus differences in dollar labour costs across states/territories. This is a deviation from neoclassical production theory; the impact of this deviation is explored through simulations below.
The code used for this analysis is available online as 20150728_RESPAresearch_Cautious_analysis_of_AWE_data.R.

Labour costs are but a single example where the costs differ across states/territories, with the price of other operating inputs in need of analysis. However, as stated by the AER, labour costs are in the range of around 70-80% of operating costs [24]. It is unrealistic to assume that the price of other operating inputs are the same across geographical regions when the presence of local market idiosyncracies, transport costs, etc. exist. These deviations can add up to produce a significant effect on this analysis.

Simulation Through simulation analysis, we explore the effects of differences in input prices, a deviation from neoclassical theory, on the efficiency estimates of firms in a controlled *in silico* environment. This simulation is identical to that in Section 2.4 with the following exceptions: capital is now assumed to be optimal and the price of the operating input is idiosyncratic for each firm at a random (uniform) value in the interval $[0.95, 1.05]$. This interval was chosen as it is a conservative choice given the analysis in Figure 2.7. The histogram of the efficiency scores are shown, along with a discussion, in Figure 2.8(a).



(a) (b)

Figure 2.8:
(a) This figure shows the histogram of the difference between actual and measured operating input cost efficiency levels when both capital inputs and operating inputs are optimal but when firms face different operating input prices. The simulated price is in the interval $[0.95, 1.05]$*, yet the difference between actual and estimated efficiency has a larger range, typically around* $[-14, 0]$*. This demonstrates that differences in relative prices between firms, even assuming all else in the neoclassical world holds perfectly, have significant effects on the estimation of firm efficiency levels.*

(b) This figure shows the effects of adjusting the opex of only *Australian firms by the conservative estimate of (private sector) AWE differences between states/territories in the NEM (i.e. from Figure 2.7). The dark bars are those accepted by the AER and the light bars are those adjusted for price differences between the states in the NEM. Issues remain with this estimation, including the fact that the costs of overseas firms have not been adjusted for internal jurisdictional differences, and that these firms largely define the frontier. Despite this and other issues, this figure can still provide some qualitative insight into the effect of price differences; the range of efficiency score differences (approximately* $[-6, 4]$ *efficiency percentage points), relative to those accepted by the AER, are not negligible.*
The code for this simulation is available with this report as 20150728_RESPAresearch_Non_optimal_capital_and_different_prices.R and 20150811_RESPAresearch_Benchmarking_data_for_R_accounting_for_price.csv.

We have also applied the prices in Figure 2.7 to the states/territories in the NEM, and then rerun the model accepted by the AER; this is shown and discussed in Figure 2.8(b). This process is not ideal for a number of reasons: labour is not the only opex cost; the 'local' price of each individual firms is unknown (i.e. how the prices differ *within* states); the interaction of the prices with other issues such as capital stock, and key to this particular case, the labour prices in New Zealand and Ontario are assumed to be the same for all firms within those jurisdictions – this is unsatisfactory as these jurisdictions form the majority of the data set, and so the frontier will be largely unchanged relative to the estimates the AER has accepted. This estimation, however, does provide a *qualitative* and *cautious* indication of the extent to which the efficiency scores are affected.

Plausibility Given the evidence of differences in labour costs, and the simulation results within the neoclassical world showing that small deviations in price have a non-trivial effect on efficiency score estimates, it is implausible that this assumption holds.

2.6 All firms have the same (or similar) production sets

For consistency with neoclassical theory, the discussion in this section assumes managers have perfect information, have made optimal capital decisions, and face the same input prices (i.e. Sections 2.2, 2.4 and 2.5).

Assumption To be able to estimate a single cost function reflective of the cost functions faced by all firms, all firms need to have the same, or very similar, production sets. This is so that the estimated cost function reflects each firm, and each firm can physically implement the results of the efficiency analysis (i.e. they can choose a point on the frontier).

Context This assumption is implicit in the AER accepting the estimation of a single, common cost function, and the setting of operating expenditure based on that cost function. There are two main reasons for this assumption: (i) there is not enough data to be able to measure the production set for each firm separately; but more importantly (ii) the intended purpose of the analysis undertaken for the AER is to benchmark the operating expenditure of Australian firms [2], so firms necessarily need to be compared to one another. This second reason is the presumed driving force for the entire analysis.

Implication(s) if the assumption does not hold If firms have different production sets resulting in different cost functions, then the estimation of a single cost function is not reflective of the possible input-output combinations firms face, and so the foundation of the analysis is undermined, meaning the efficiency results may be spurious and thus unreliable.

Reality The assumption of a common production set is closely related to the interpretation of a production set discussed in Section 2.3. Recall that the production set contains, by definition, the individual exogenous heterogeneities of a firm in a particular geographical and political jurisdiction. In reality firms face different constraints: some arising from historical decisions, some from the physical environment, some from the political or socioeconomic environment, and some even from customer preferences. These differences are, within the neoclassical world, encapsulated within a firm's production set. Assuming a single production set for all firms assumes away every difference between the firms.[14] It is these differences which makes the *ex ante* production sets of different firms different; how different is a non-trivial empirical question. It might be assumed that the *ex ante* production sets

---

[14] The argument that these differences are accounted for post-modelling is discussed in Section 3.3.

are similar for similar regions, but it needs to be recalled from Section 2.3 that it is highly implausible that firms actually have access to, and use, the *ex ante* production set.

On the other hand and as discussed in Section 2.3, the production set can be interpreted as an *ex post* costless production set; this is a plausible interpretation. However, this production set is dependent on historical choices of individual firms, for which it is implausible to assume that firms have made the same or similarly inefficient choices in the past. The assumption that firms have, on average, made the same inefficient historical choices is without basis and evidence. Even if this were true, it was demonstrated by simulation in Sections 2.4 and 2.5 that the *distribution* of inefficient choices is of relevance to efficiency analysis, not simply the mean or median choice.

Plausibility This assumption relies on one of two implausible paths – the *ex ante* production set, or all historical choices have been equally inefficient for all firms – depending on the interpretation of the production set. In either case, it is implausible.

3 Statistical assumptions

To be able to discuss the statistical SFA model accepted and used by the AER, the assumptions set out in Section 2 must be taken to be true, despite their cumulative implausibility. The SFA model accepted by the AER can be specified as the following econometric model:

$$\begin{aligned}\ln(opex_{i,t}) \quad = \quad & \beta_0 + \beta_1 \ln(RMD_{i,t}) + \beta_2 \ln(custNum_{i,t}) + \beta_3 \ln(lineLength_{i,t}) + \beta_4 \ln(UGlength_{i,t}) + \\ & \beta_5 \ln(Price_{i,t}) + \beta_6 t + \beta_7 \mathbf{1}_{NZ} + \beta_8 \mathbf{1}_{Ont} + v_{i,t} + u_i \end{aligned} \tag{3.1}$$

where $v_{i,t}$ is the symmetric error term and u_i the asymmetric error term. We have replicated the analysis undertaken for the AER from first principles.[15] The analysis in this section looks at two facets of the statistical model used: whether the efficiency estimates accurately reflect the 'actual' efficiency of firms, and whether the distributional assumptions have an effect on the efficiency estimates. That is, the two questions we are asking are: (i) given a firm has an 'actual' efficiency level (which in the following simulations, we impose by construction), what is the distribution of the difference of the actual from the estimated efficiency?; and (ii) does changing the 'actual' distributions of the two error terms affect the distribution of efficiency estimates? These questions are answered via simulations.

3.1 Estimated efficiency scores reflect actual firm efficiency

Assumption The difference between the estimated efficiency and the actual efficiency is negligible, so the estimated efficiency score accurately reflects the actual efficiency of a firm.

Context This assumption is required to be able to draw a reliable inference from the estimated efficiency scores. The AER has drawn such inference and thus has accepted this assumption.

Implication(s) if the assumption does not hold If this assumption does not hold, then any inference from this statistical model may be unacceptably biased.

Reality The difference between actual efficiency and estimated efficiency can never be known in reality. This does not mean that inference can be assumed to be appropriate, nor that it cannot be shown to be inappropriate. Instead, the size of this difference can be estimated via simulation. In fact, the simulation in Section 3.2 is how this difference was observed, but its magnitude warranted a discussion of its own.

Simulation In the following statistical simulation, we suppose Eq. (3.1) is the true underlying data generating process (**DGP**) where $v_{i,t} \sim \mathcal{N}(0, \sigma_V^2)$ and $u_i \sim \left|\mathcal{N}(\mu, \sigma_U^2)\right|$ ($\mathcal{N}$ is the normal distribution, and $|\cdot|$ is the truncation at zero of the positive part of the distribution in question), and the parameter values estimated by EI to be the true values. Data is then generated by random sampling from the two distribution (assumed to be the true distributions in this simulation). Using

---

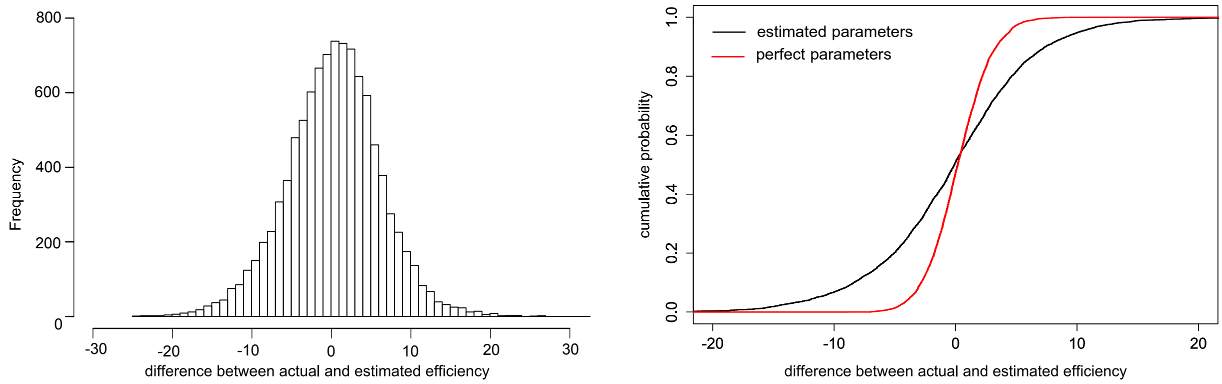
[15] Our code demonstrates that the two models – Eq. (3.1) and the model in Table 5.2 in [2] – produce identical efficiency results, as discussed in Footnote 2. This file, 20150728_RESPAresearch_Replication_of_EI_results_and_distbn_analysis.R, is available on our website.

the draws from the asymmetric distribution, actual efficiencies of the simulated firms are known with certainty. The draws from the two distributions, along with actual output data used in EI's model (i.e. the independent variables in the regression), are used to determine 'simulated opex' values (i.e. the dependent variable). At this point, we split the simulation into two parts:

(i) We use the (known) parameters to calculate the estimated efficiency scores, as per the latter part of the SFA model accepted by the AER, and compare this to the actual efficiency of each of the firms: this difference is the variable of interest in determining if this assumption holds. This proceeds without estimation of the parameters, and provides insight into the robustness of the model given perfect estimates.

(ii) We use the simulated data to estimate the parameters using the entire SFA procedure accepted by the AER, which is then used to determine the total effect on efficiency scores in a setting akin to that accepted by the AER. The difference between the actual efficiency and estimated efficiency is then calculated.

In either part, this process generates 68 data points of differences between actual and estimated efficiency scores – one for each firm in the sample. This is repeated a number of times to generate a sample of over 10,000 simulated differences. These 10,000 data points are plotted as a histogram, and discussed, in Figure 3.1(a).



(a) (b)

Figure 3.1:
(a) This histogram shows the deviation of actual efficiency from the estimated efficiency score when the assumptions of the model are known to hold (this is by construction). A wide range of differences between actual and estimated efficiency is observed; this difference may be centred at a mean of approximately zero, but a non-trivial number of firms (over 9%) have a deviation of over 10 efficiency percentage points from their true efficiency. A similar measure is the 95% empirical confidence interval: this is approximately $[-12.8, 10.7]$ *efficiency percentage points. This large deviation of actual versus estimated values is a general issue inherent in statistics; statistics is used to find general regularities and relationships rather than properties of individual data points.*

*(b) This figure shows the comparison of both the simulation with perfectly known parameters (red) and the simulation with estimated parameters (black – this is the same data as (a)) plotted as empirical cumulative distribution functions (***ecdf***). There are a couple of interesting features: first, perfectly estimated parameters do not result in the difference between actual and estimated efficiency being zero – the 95% confidence interval is approximately* $[-3.7, 4.4]$. *Secondly, estimation of the parameters of the model, using 544 data points, the same number as accepted by the AER, produces vastly increased uncertainty relative to the perfectly known parameters, as shown by the increase in the 95% empirical confidence interval (i.e.* $[-12.8, 10.7]$*). This demonstrates that the sample size is insufficient to produce approximately accurate efficiency estimates – that is, the size of the 95% empricial confidence interval has increased by over 15 efficiency percentage points due to imperfectly estimated parameters.*
The code for this simulation is available online as 20150729_RESPAresearch_Statistical_assumptions_changed_parameters.R.

In a simulated setting, where the error terms are drawn from known distributions, with a sample size identical to that accepted by the AER, the 95% confidence interval had length 23.5 efficiency percentage points. This result follows the same process as that accepted by the AER, in that the underlying parameters must be estimated, rather than are known *ex ante*.

A third simulation has been used to investigate the size of the 95% confidence interval under reasonable differences in the true underlying values of the three distributional parameters: μ, σ_V^2 and σ_U^2. This simulation provides insight into the sensitivity and relationships between these three parameters on the range of the 95% confidence interval. The effects explored are changes of each of the three parameters of: -1 standard deviation, $+0$ standard deviations, and $+1$ standard deviation to each of the three assumed (true) distributional parameters (the standard deviation is in reference to those estimated by EI [2]). This produces 27 cases to consider; the empirical cumulative distribution functions of these 27 cases are shown and discussed in Figure 3.2.

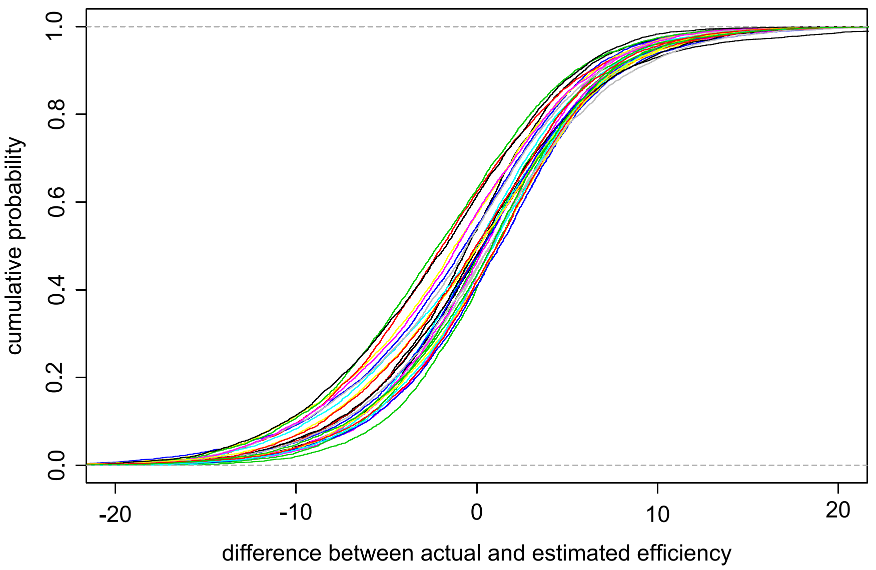


Figure 3.2:
This figure plots the empirical cumulative distribution functions of the 27 cases described above. The size *of the 95% confidence interval of these cases ranges from around 17 to around 24 efficiency percentage points; this does not consider the lesser issue of the movement away from a mean of zero (as the relative deviations are of primary concern). This result demonstrates that reasonable changes in the estimates of the distributional parameters (i.e. within one standard deviation) produces non-trivial changes in the potential to erroneously measure firm efficiency. This figure, along with similar estimations in the context of purely simulated firms (simulations not shown), suggests that the* interaction *between the three distributional variables within an SFA framework with 544 data points, is not simple, and has a non-trivial effect on the differences between actual efficiency and estimated efficiency.*

The results of these three simulations provide an analysis of statistical robustness and fairness. Given the distribution in Figure 3.1, it would be inequitable to draw deviations from true efficiency scores from such a distribution: the dispersion is simply too large. The lack of fairness is further compounded by benchmarking firms with the lowest estimated efficiency scores against a frontier determined by the firms with the highest efficiency scores. The large and uncertain range of the deviation of estimated efficiency scores from actual efficiency shows that the results are not marginal nor trivial. These three simulations demonstrate the lack of robustness of the applied use of the SFA model in this circumstance. The use of the mean difference in the above simulation does not reflect the *distribution* of the difference of actual versus estimated efficiency scores, and is another instance of the median or mean being misleading.

Confidence The simulations above show that deviations of efficiency scores of over 10 efficiency percentage points are experienced by 9% of firms. This is not a trivial nor inconsequential result: rather, it is evidence that the model accepted by the AER is not statistically robust, and has direct effects on the efficiency results of the model. It is unlikely that the model, in its current form and interpretation, produces accurate results reflective of actual efficiency levels.

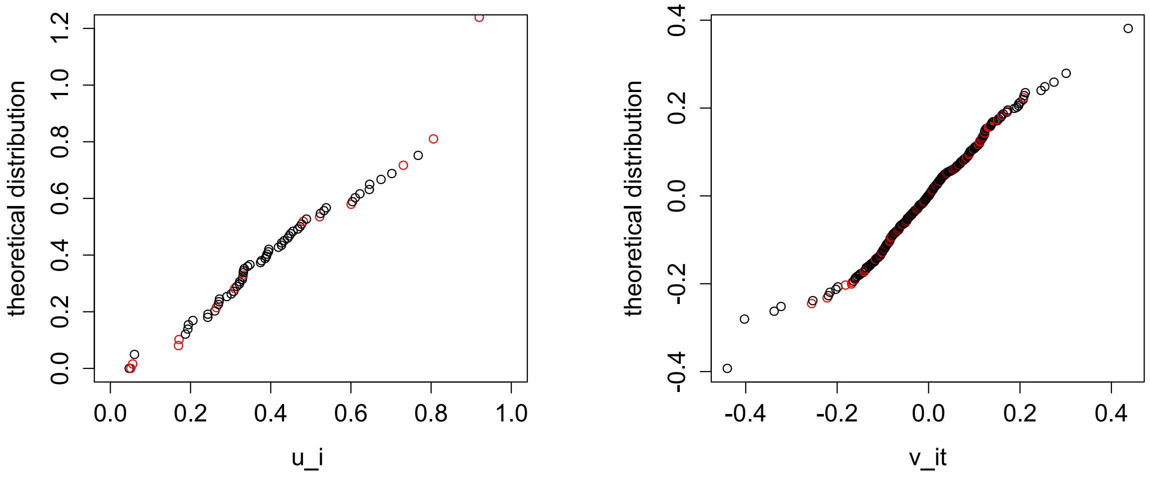
3.2 The distributions of error terms are correct

Assumption The symmetric and asymmetric error terms have the following assumed distributions, respectively: $v_{i,t} \sim \mathcal{N}(0, \sigma_V^2)$ and $u_i \sim \left|\mathcal{N}(\mu, \sigma_U^2)\right|$.

Context The intended purpose of the SFA model is to estimate the cost function along with two random terms, one a symmetric error term (typically interpreted as 'random noise') and the other an asymmetric error term (typically interpreted as 'inefficiency'). The likelihood function, used to estimate the parameters in the maximum likelihood framework, is defined by the assumptions imposed on the error terms. For both error terms to be identified within the maximum likelihood framework, distributions need to be imposed on both terms. The distributional assumptions imposed above have been used in academic papers (its first use was in [25]), mainly due to the mathematical tractability in defining the likelihood function. Mathematical simplicity does not imply the assumed distributions are appropriate in every circumstance: this needs to be examined in each individual case. This is not to say that statistical assumptions should not be made; rather, statistical assumptions are necessary to give statistical insights, but they need to be consistent with the data.

Implications if the assumptions do not hold If one or both of the distributional assumptions imposed on the error terms is misspecified, then the likelihood function is incorrectly specified and so the estimation of the parameters (i.e. the maximisation) may not be reliable or valid, and can have serious consequences for interpretation of the data [26]. It is important to check for these distributional assumptions before accepting any statistical procedure [27]. Since the parameter estimates determine both (estimated) errors, and the asymmetric errors are used to determine efficiency, then the consequences of a misspecified likelihood function flow into efficiency estimates.

Reality A number of formal and informal tests can be used to analyse distributional assumptions, as discussed in [26]. The first test is a visual test: the QQ-plot of the data against the assumed distribution for both error terms – see Appendix 2 for a discussion of the interpretation of a QQ-plot. Figure 3.3(a) shows the QQ-plot for the asymmetric error term, and part (b) of the same figure shows the QQ-plot for the symmetric error term.



(a) (b)

Figure 3.3:

(a) The QQ-plot of the asymmteric error term suggests that there might be some non-linearities in the plot, but this would be a subjective leap of faith. A more appropriate interpretation is that data seems to be approximately linear, with a significant outlier.

(b) For the symmetric error terms, there seems to be a non-linear 'S'-shaped relationship in the QQ-plot. The data seems to have fatter tails than the assumed normal distribution.
In both plots, the red points represent Australian firms.

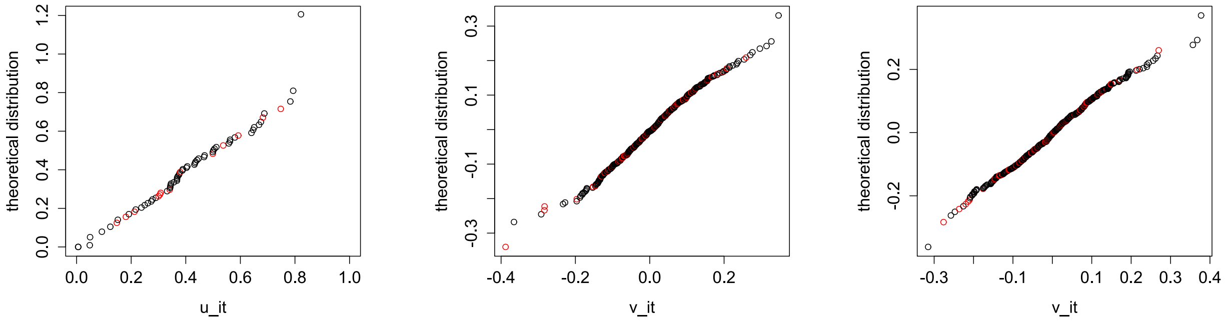
Visual inspection of the QQ-plots shows that the assumption on the symmetric error term may be inconsistent with the assumed normal distribution: the actual symmetric errors seem to have fatter tails than a normal distribution. Formal tests of normality are shown and discussed in Table 1.

test	u_i	$v_{i,t}$
Shapiro-Wilk	not reject normality $(p = 0.2053)$	reject normality $(p = 7.77 \times 10^{-8})$
Anderson-Darling	not reject normality $(p = 0.1511)$	reject normality $(p = 5.78 \times 10^{-4})$

Table 1:
This table shows tests of normality of the asymmetric (u_i) and symmetric ($v_{i,t}$) error terms at the 1% significant level. These results are consistent with the QQ-plot visual analysis: the asymmetric error term is consistent with a normal distribution (rather than a truncated normal, but since the truncation occurs at an extreme point away from the majority of the probability mass, this is quite irrelevant), whereas the symmetric error term is inconsistent with a normal distribution at the 1% level of significant. While these tests have weaknesses, they have the greatest statistical power of a number of common normality tests [27]. The statistical power of a test is the extent to which that test rejects the null hypothesis (i.e. that the data is normally distributed) when it is false (i.e. the data is not normally distributed). The weaknesses of these tests include having low power for sample sizes under approximately 200, but with a sample size of 544, these tests have been shown to have large statistical power (it needs to be noted that these results were under controlled conditions) [27].
These test are still useful elements in statistical decision making when they are used as indicators rather than certainties, as in this report.

The consequences of violating the assumption of the distribution of the symmetric error term are not isolated to that term, but rather pervade the estimation of all the parameters and the subsequent efficiency analysis. That is, it is not a valid interpretation to state that since the asymmetric error terms used in the efficiency analysis are consistent with their assumed distribution then the efficiency analysis itself is valid. The degree to which a misspecified likelihood function affects efficiency score estimates is now analysed using simulations.

Simulation This simulation is a replication of the simulation in Section 3.1, with the following differences: assume that the symmetric error term has either a $t-$distribution with 5 degrees of freedom, or an exponential power distribution with shape parameter equal to 1.5, both appropriately scaled to approximate EI's variance estimate of the symmetric error term. The reasoning behind these two choices is to be able to impose a fatter tailed distribution visually reflective of the distribution of the symmetric error term in Figure 3.3(b): the QQ-plot of a typical 'run' of the simulated data against the theoretical distributions for both error terms is shown in Figure 3.4. These simulation results are then compared to the 'control', which is the result from Figure 3.1.



(a) (b) (c)

Figure 3.4:
These QQ-plots are typical examples from the random draws described below.
(a) This asymmetric error data was created by taking $N = 68$ *(the number of firms in the sample) random draws from a truncated normal distribution with mean* $\mu = 0.385237$ *and variance* $\sigma_U^2 = 0.038796$*, as per the estimates accepted by the AER.*
(b) (Experimental condition 1) This symmetric error data was created by taking 544 random draws (the number of observations) from a $t-$*distribution with five degrees of freedom and scaled so that variance is approximately* 0.0098.
(c) (Experimental condition 2) This symmetric error data was created by taking 544 random draws from an exponential power distribution with shape parameter equal to 1.5 and scaled so that variance is approximately 0.0098.

This simulated data is then run through the same SFA model as accepted by the AER. The difference between the true efficiency of the firms, as per the DGP for the asymmetric errors described above, and the estimated efficiency scores using this data, are shown and discussed in Figure 3.5.

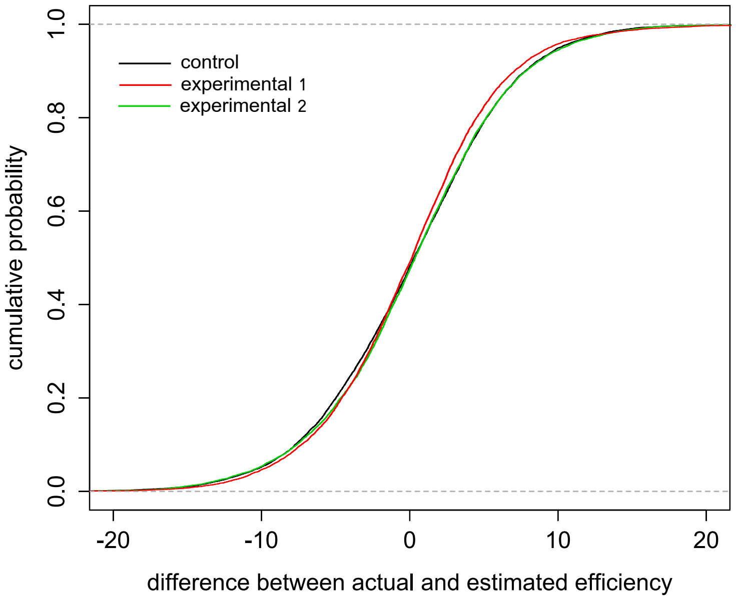


Figure 3.5:
The data used in Figure 3.1 is the 'control' condition to which the 'experimental' conditions of the heavy-tailed symmetric errors are compared with. Given the data sets generated using the process described above, this figure plots the empirical cumulative distribution functions. There is some variation away from the control condition (black) for both the heavy tailed distributions (red and green). This difference is quite small at approximately 1 efficiency percentage point difference at the maximum. However, the degree of spread of the heavy tailed distributions, relative to the control condition, remain essentially unchanged: it is still large.
The code for this simulation is available online as 20150729_RESPAresearch_Statistical_assumptions.R.

Confidence The model estimates of the symmetric error terms are inconsistent with the statistical assumptions imposed to construct the likelihood function. The simulation above demonstrates that the misspecification of the distributional assumptions of the errors seems to have a negligible effect. Despite the violation of these distributional assumptions having a small effect on the efficiency scores, the spread of the difference between the actual and estimated efficiency scores remains an issue, as discussed in Section 3.1.

3.3 Post-modelling adjustments and the frontier

Post-modelling adjustments The AER adjusted the efficiency scores by factors it, in its regulatory discretion, deemed important. These post-modelling adjustments are not founded in economics, statistics or any other science; they are subjective, and at times, inconsistent with implicit theoretical assumptions used in the modelling.

Assuming a factor is theoretically insignificant in the formal modelling but adjusting the results of that formal model for the impacts of that same factor is an arbitrary, and inconsistent, approach to analysis. For example, where post-modelling adjustments have been made to account for a specific variable, the AER has changed efficiency scores to account for that variable by between 0.5% and 8.5% for NSW and ACT distribution firms [28] (the determinations available at the time of this analysis). These adjustments are not inconsequential. Theoretically – and the results used are based on this theory – if the factor is adjusted for post-modelling, then it is significant. If it is significant, then it needs to be reflected in the production set.

Further, the size of the effect cannot be assumed to be a one-for-one adjustment, as demonstrated in different but related contexts in Sections 2.4 and 2.5. These factors, if assumed to be important, affect both the *ex ante* and low-cost *ex post* production sets. The AER assumes the size of these effects to be small, and the flow-through effects also to be small. This assumption is without evidence and, as was demonstrated with the simulation of differences in input prices, small differences in assumptions can be magnified by the theoretical framework.

The lower bound of any positive adjustment to a variable is purely arbitrary: the assumption of 0.5 percentage points change to the efficiency score assumes that the average of all such effects is this value. Leaving aside the issue with averages, there is no scientific basis for this claim, nor this general approach; this is neither a quantitative nor qualitative approach as these values are not based on any model or evidence. These factors have, in the main, not been attempted to be quantified and adjusted for prior to modelling, and as such leave the estimation open to controllable biases.

In a purely statistical sense, if a variable is assumed important but omitted from the analysis, then the parameter estimates will suffer from omitted variable bias (even if the likelihood function were of the correct form bar the omitted variable). By using post-modelling adjustments, the AER has assumed the bias in any of the estimated parameters, and thus the efficiency estimates, to be immaterial. This assumption is also without basis.

It is worth noting that we can see at least one feasible method to simulate the size of the effects of the adjustments described in this section and in Section 7 of [2], while remaining within the neoclassical realm. This has not been pursued in this report as it is expected to require discussion with stakeholders.

Frontier adjustment The frontier is the minimum cost function in neoclassical theory on which every efficient firm would operate. The frontier is an immovable benchmark in neoclassical theory, the theory on which the SFA model relies; the frontier should not change if the model has been correctly specified and estimated. The movement of the frontier accepted by the AER is an arbitrary, non-linear adjustment to the efficiency scores of the firms: after moving the frontier, any firm above the new frontier has an efficiency score of 1 while every other firm is scaled by this lower frontier.[16] The AER justifies this adjustment as:

> [I]n deciding what is materially inefficient, we consider it is appropriate to provide a margin for the effect of potential modelling and data limitations. To give effect to this

---

[16] The mathematical relationship equivalent to moving the frontier down is:

$$newEffScore = \min\left\{1, \frac{oldEffScore}{0.768}\right\}$$

Writing out this relationship helps us understand the nature of this choice. (We understand the reason for this form is a movement down of the frontier, but we are trying to analyse the technical implication of this choice.) The division by 0.768, a relatively arbitrary number, scales the SFA estimated efficiency scores and thereby increases the difference in efficiency scores between any pair of firms. However, the non-linearity introduced by the minimum function results in some pairs of firms retaining this increased efficiency score differential (i.e. those below 1 in the new efficiency score scale), other pairs of firms decreasing their efficiency score differentials (i.e. when both firms have an efficiency score of 1 in the new scale), and some pairs of firms may have an increased or decreased differential depending on the specific values. This demonstrates the inconsistent effect of this adjustment on the comparison between firms.

> consideration, we do not compare service providers to the frontier business. We consider the appropriate "benchmark comparison point" is the lowest of the efficiency scores for service providers in the top quartile of possible scores on our preferred SFA model. [29]

The AER considers the frontier adjustment accounts for theory and model uncertainty (i.e. that their model and the underlying theory may be inappropriate).[17] However, the AER still accepts (a linearly expanded version of) the size of the efficiency differences between any two firms below the frontier. This is accepting that the differences between some pairs of firms is larger than the results of the SFA model, but the differences between other pairs is smaller (as discussed in Footnote 16). This does *not* adequately account for issues with model robustness as it does not treat all firms equally. Further, this is not a cautious approach to the use of the results of the model, given the issues described in this report on which the results rely upon; this does not even consider data quality and comparability issues.

---

[17] Model uncertainty is a major source of Knightian uncertainty, one that if assumed away is inconsistent with a scientific approach. Section 4.2 discusses this in more detail.

4 Context of the interpretation

This section discusses evidence of efficiency analysis from other, some more competitive, industries, as well as the issue of model interpretation.

4.1 Efficiency estimates in other industries

The application of efficiency analysis in a regulated industry is typically justified in its use by the claim that it approximates the competitive pressures of (more) competitive industries. The underlying assumptions is that all firms in competitive industries are perfectly efficient; that is, all firms in competitive industries are on the frontier of the industry cost function (which is also assumed to exist). A number of studies related to 'X-efficiency' are relevant for this discussion: these are discussed in Table 2 below.

X-efficiency, in this report, is defined as the percent by which actual total costs could be reduced if a firm were to operate on the cost efficient frontier [30].[18] X-efficiency varies between 0% and 100%, with 0% interpreted as costs being perfectly efficient (i.e. no cost savings can be achieved) and thus on the frontier, and 100% interpreted as the limit where all costs can be reduced to get to the frontier.

Table 2 summarises a number of studies which measure X-efficiency of industries ranging from banking to community colleges. In the SFA framework, the statistical framework used by all of the studies in Table 2, the 'asymmetric error term' is used to determine the 'X-efficiency', very similar to the method accepted by the AER. Figure 4.1 plots the X-efficiency score distributions of the studies with available data discussed in Table 2, along with the Australian electricity distribution firms' (raw) efficiency scores. The efficiency results from these studies provide context for the X-efficiency scores accepted by the AER: the cost reduction potential of these industries is not that different from the Australian electricity distribution industry. The argument that all firms must be at or near the efficient cost frontier to be reflective of competition is therefore ill-founded given this evidence.

---

[18] It can be 'inversely' defined as '1 minus' the definition we use in this report.

Paper(s)	Details of study	Descriptive statistics of X-efficiency
AER/EI	This is the study accepted by the AER; it is used to give context to the results of the other studies.	Mean $= 31.6\%$, standard deviation $= 18.5$ efficiency percentage points, range $= [5, 60]$.
[30]	This study analyses the Hong Kong commercial banking sector. [31] finds that, during the period 1992 to 2002, the Hong Kong banking sector was "highly competitive"; it would be expected that most firms would be at or near the efficiency frontier. The stochastic frontier approach is used in [30] to investigate the X-efficiency of 59 commercial banks in Hong Kong over the period 1992 to 1999. The descriptive statistics of the X-efficiency estimates for Hong Kong banks are comparable to those found in other countries, including the US. The author notes that the magnitude of the measured cost inefficiency seems striking given the very high density of banks.	Mean $= 32\%$, standard deviation $= 22$ efficiency percentage points (the distribution is shown in Figure 4.1).
[32, 33]	[32] examines the properties of X-inefficiency and the relationship of X-inefficiency with risk-taking and stock returns of 254 U.S. bank holding firms during the period 1986 to 1991 using the SFA translog cost function approach. [33] examines the importance of off-balance-sheet (OBS) activities of a set of publicly listed U.S. banks on X-efficiency measures. The authors use distribution free and SFA methods to estimate the X-efficiency of a set of publicly traded U.S. commercial banks over the period 1992 to 1997. The authors note that previous studies have placed mean production cost efficiency at near 20%, with these results not too far from those estimates.	[32] (quartile of smallest firms): mean $= 19\%$, median $= 15\%$, standard deviation $= 15$ efficiency percentage points, range $= [2, 95]$; (quartile of largest firms): mean $= 8\%$, median $= 7\%$, standard deviation $= 4$ efficiency percentage points, range $= [2, 32]$; [33] (SFA, with OBS activities): mean $= 15\%$; (SFA, without OBS activities): mean $= 18\%$; (distribution free method, with OBS activities): mean $= 26\%$; (distribution free method, without OBS activities): mean $= 28\%$.
[34]	This study focuses on the causes of the differences in X-efficiency in the life insurance industry, using U.S. data from 358 life insurance firms during the 1990 to 1995 period. This study tests for a relationship between a firm's output choice and measures of X-efficiency. The findings of an older study, [35], also analysing X-efficiency in the U.S. life insurance industry, are consistent with this study.	Mean $= 58\%$, median $= 64\%$, range $= [0, 85]$.
[36]	This study examines the productive efficiency levels present in the market for residential real estate brokerage services by using the stochastic frontier approach. The data comes from 276 firms taken during a periodic, nation-wide survey of firms in the U.S. real estate brokerage industry, which included the 1990/91 survey.	Mean $= 12\%$, standard deviation $= 5$ efficiency percentage points, range $= [6, 36]$.
[37]	This (working) paper examines the X-efficiency of firms in the U.S. community college sector using SFA. Data from 950 community colleges over the period 2003 to 2010 is used.	Mean $= 41\%$ (the distribution of the 950 firms is shown in Figure 4.1 below).

Table 2:
This table is a brief summary of studies of X-efficiency in other, some more competitive, industries than the Australian electricity distribution industry.

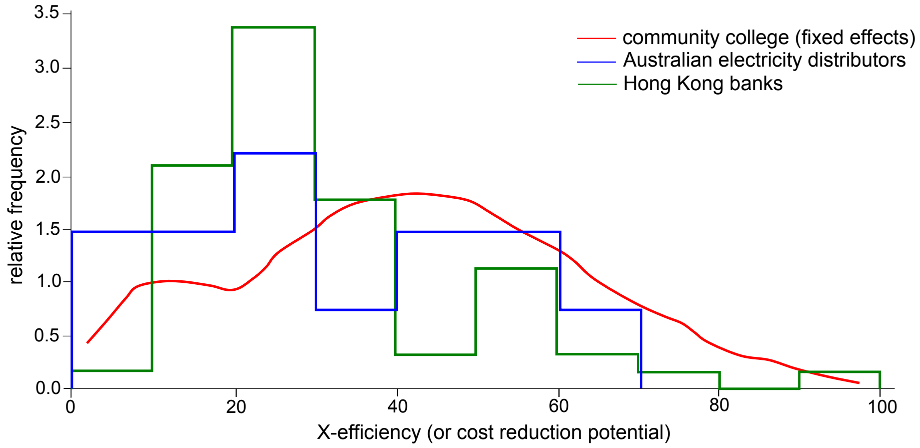


Figure 4.1:
This figure shows the distribution of the X-efficiency scores (i.e. cost reduction potential) of the studies in Table 2 that provided such detailed information. Note that for the Hong Kong banking and the Australian electricity distribution sectors, due to a small sample size (59 and 13, respectively), the data was put into bins of 10 efficiency percentage points. This figure is a simple way to demonstrate, in a general sense, that the (raw) inefficiency estimates of Australian electricity distribution firms (blue) are not that visually different to the Hong Kong banking industry (green) and the (conservative, state fixed effects) efficiency estimates of the U.S. community college sector.

4.2 Uncertainty and complexity

We at RESPAresearch recognise two major types of uncertainties when modelling, which we call model robustness and model uncertainty.[19] Model robustness is the degree to which the results of a model change when the assumptions within the model universe are altered. We analyse model robustness by altering certain assumptions in simulated environments; this allows us to understand how sensitive the results are to certain assumptions. (Model robustness is analysed in Sections 2 and 3, and forms the bulk of this report.)

Model uncertainty is the Knightian uncertainty which exists from using a model as a surrogate for a target system (e.g. the efficiency of Australian electricity distribution firms) when the accuracy and appropriateness of the model are unknown or unknowable. Model uncertainty and model robustness are two separate concepts. A model can be highly robust to modelling assumptions, but it can still be a poor reflection of the target system. Model uncertainty cannot be quantified in any single real-world application, even via simulations, since in most, if not all cases, we do not know the underlying true processes. It can, however, be shown to be small or negligible over many studies comparing model results with the phenomenon the model emulates. Unfortunately, this is not the case with efficiency analysis as we do not know the true efficiency of any firm, so cannot compare the model results with true efficiency levels.

Model uncertainty arises from the potential of inappropriate modelling of the target system, and is related to the complexity of the target system. Complexity is the scientific term used to denote systems with multiple, interacting components which produce global behaviours that cannot be easily

---

[19] We accept and *incorporate* modelling weaknesses and uncertainties into our recommendations; it is a clause in our founding constitution.

modelled or explained in terms of the interactions between individual elements.[20] The more complex the target system, the more model uncertainty is present, and the more we do not know the extent to which a model is valid, and subsequently the extent to which we can rely on the results. Rather than making detailed predictions, complexity research focuses on asking the right questions given the context, and suggesting possible strategies – regulatory strategy is the benefit that complexity science can provide.

Model uncertainty is typically ignored when applying theory to policy. In most cases, if it is acknowledged, the model results are still largely accepted. The argument that the theory used is the best theory available – a subjective assessment ignoring behavioural theories – which implies that it is the best that can be done at the current point in time, is invalid. This argument does not address the consequence of adopting a model without considering its inherent uncertainty: the "best" model may still be very poor and misleading.[21] Model uncertainty needs to be acknowledged in the *interpretation* of the modelling results, above and beyond that of model robustness. The results themselves are what need to be tempered: the more model uncertainty, the more qualitative the interpretation needs to be.

Models which give precise answers are most highly valued, even if this precision is inconsistent with evidence. Despite being a human desire, the acceptance of such precise estimates ignores both weak and strong Knightian uncertainty. The precision of the results produce a false veneer of certainty in the presence of significant model uncertainty.

Academic acceptability and the historical use of models is no guarantee that the models are appropriate and definitely does not exclude model uncertainty. "[T]he beauty of complicated mathematics is that it appears to the public as a product of scientific discourse" [48], but as discussed in [47], the weaknesses of these models needs to be understood and not underestimated.

In recent years, there has been a growing heterodoxy in economics regarding the assumptions and tools underlying neoclassical economics. Despite the myriad proximate causes of this heterodoxy, the ultimate cause is an attempt to remove part of the unrealistically strong assumptions required in neoclassical economics. The reason is that dealing with systems such as the economy or individual firms is not straightforward because these are complicated and interconnected with other non-economic systems such as social, political, biological and ecological systems [48]. Complexity economics deviates from the neoclassical framework in a discrete way, intimately affecting the framework itself.

Recent publications in complexity economics are an attempt to reconcile observed behaviours and theory. For example, [48] argues that uncertainty within complex systems creates a tendency to adopt rules, in the form of habits, routines, conventions and norms that are just as rational in such conditions. However, learning plays a major role when the environment inevitably changes. There is a strong similarity of this approach with that evidenced in Box 1. [48], and to varying extents [49], discuss these simple rules and argue that optimisation is a less relevant rule the more pronounced the changes in behaviours, processes and constraints. [48] goes on to argue that firms are a good

---

[20] There is still little consensus on the definition of a complex system, however, this captures the essence of most definitions used.

[21] Suppose the 'best' model for economic efficiency analysis was precise with probability 0.05, and a random draw from the interval $[0, 1]$ with probability 0.95. Given this model is 'best' at the current time, it clearly should not be used to guide policy. This is a simple, but still not ideal, analogy to understanding model uncertainty. (It is not ideal as it assumes we know, with certainty, that the model is correct with some probability and incorrect with some other probability; the probabilities assigned to model accuracy are simply unknown in applications.)

example of a complex adaptive system (**CAS**). This is because the rules used by firms lead to the formation of network structures that have a degree of irreversibility (which further restricts the extent to which the logic of optimization can be applied). In a CAS, the production function is not a suitable analytical construct to capture how production actually takes place. It may be trivial to associate inputs with outputs over any chosen time period, however in reality, such relations alter over time. This alteration can be large and can occur in both smooth and discontinuous forms. The capital stock that currently exists within a firm is not 'optimal': it is inherited from history and only new investment can be subject to (constrained) optimization exercises. The existence of irreversible capital stock implies that constrained optimization is fraught with difficulty as a general theoretical framework for understanding production [48]. Appendix 3 discusses an evolutionary framework for understanding firms and industries which embraces dynamic adaptability rather than static optimisation.

4.3 The precautionary principle

The precautionary principle, as has been used in potential environmental damage, takes one of two forms in [38]:

> Where there are possibilities of large or irreversible serious effects, scientific uncertainty should not prevent protective actions from being taken. [Or] [w]here there are possibilities of large or irreversible serious effects, action should be taken, even if there is considerable scientific uncertainty.

The *Protection of the Environment Administration Act 1991 (NSW)* [39] adopts the precautionary principle in a similar form.[22] The judgment of *Telstra Corporation Limited v Hornsby Shire Council* [40] discussed a number of facets of the precautionary principle, including: (i) *shifting of the burden of proof*: if there is a threat of serious or irreversible damage and there is the requisite degree of scientific uncertainty, the precautionary principle will be activated, with a shift of the evidentiary burden of proof to where the decision-maker must assume that the threat of serious or irreversible damage is no longer uncertain but is a reality; (ii) *degree of precaution required*: the type and level of precautionary measures that will be appropriate will depend on the combined effect of the degree of seriousness and irreversibility of the threat and the degree of uncertainty – the more significant and the more uncertain the threat, the greater the degree of precaution required; (iii) *zero risk precautionary standard inappropriate*: the precautionary principle should not be used to try to avoid all risks; and (iv) *proportionality of response*: the precautionary principle embraces the concept of proportionality, which means that the measures should not go beyond what is appropriate and necessary in order to achieve the objectives in question.

The precautionary principle can be equally applied to potential economic policies: the potential activation of the precautionary principle depends on the individual circumstances of each potential policy. If it is activated, then the degree to which the precaution is required must be determined, including the possibility of inaction.

---

[22] Specifically: if there are threats of serious or irreversible environmental damage, lack of full scientific certainty should not be used as a reason for postponing measures to prevent environmental degradation.

In the context of electricity distribution, this report discusses the significant (theoretical) scientific uncertainty present in the analysis.[23] For the precautionary principle to be activated, the *potential* for serious and/or irreversible consequences must be demonstrated. From the outset, we would like to make clear that the following discussion focusses, as is necessary in the context of the precautionary principle, on the *potential* downside consequences.

Three main areas of concern regarding opex are network maintenance, vegetation maintenance and emergency responses. Benchmarking data from 2009 to 2013 demonstrates that Australian networks have these three categories constituting as much as 65% of opex, with the median proportion being over 45%.[24] The maximum reduction in opex, based on the (post-model adjusted) benchmarking model, is almost 33% [28]. Of consequence to the application of the precautionary principle, the size of these opex cuts may reduce the frequency of both network maintenance and vegetation clearance, and may also increase the response time to emergencies.

Estimates of the value of potential downside consequences of these opex cuts, given the uncertainty around the model used, include:[25]

1. *Increased bushfire frequency due to power-line ignitions*. The potential consequences of bushfires, rather than the causes, can be provided by looking at historical extremes. We don't have to go back far in time to see an extreme example: the 2009 'Black Saturday' Victorian bushfire. The 2009 Victorian Bushfires Royal Commission (**VBRC**) estimated the economic cost to Australia was over $4 billion, which included the loss of 173 lives [43]. The VBRC discussed the increased risks of, essentially, under-maintenance of ageing assets:

   > [Electricity] distribution businesses' capacity to respond to [an] ageing network is, however, constrained by the electricity industry's economic regulatory regime. The regime favours the status quo and makes it difficult to bring about substantial reform. As components of the distribution network age and approach the end of their engineering life, there will probably be an increase in the number of fires resulting from asset failures unless urgent preventive steps are taken.

   The VBRC further discusses how electricity-caused fires are most likely to occur when the risk of a fire getting out of control and having deadly consequences is greatest. Thus both asset maintenance and vegetation clearance are both expenses which reduce the chances of electricity bushfire ignitions.
   A less extreme example is the Tatong bushfire of 16 January 2007, which was estimated to have caused $500 million worth of economic loss to Victoria [44].

2. *Increased power outages due to degradation and non-maintenance of the network*. The economic consequences, rather than the causes, of power outages are of concern. Despite some significant differences, the U.S. economy is taken as a (cautious) comparator as data on total

---

[23] The discussion in the previous sections does not consider issues with the data or choice of variables. Further, we are not claiming that there are no highly inefficient electricity distribution firms in Australia; rather, we are stating that given the discussion and analysis in Sections 2 and 3, it is highly implausible that the assumptions of neoclassical production theory reflect the 'true' efficiencies of the firms, even approximately.

[24] This was calculated based on the values in the AER's category analysis spreadsheet for the three categories of interest [41], relative to the opex values for those same years from the benchmarking input spreadsheet [42].

[25] Note that there are also unknown consequences arising from the complexity of the system, such as whether or not the distribution network could be restored if degraded past some threshold, or whether an increase in blackouts will lead to an increased likelihood of consumers choosing to go 'off grid'.

> economic costs of electricity outages is readily available. The comparison accounts for differences in incomes and purchasing power relative to Australia: these converted annual values are estimated at between AUD$15.8 and AUD$24.8 billion [45], or at AUD$5.3 billion [46]. There are numerous differences between the U.S. and Australia in the electricity distribution infrastructure which would invalidate the use of these precise values, however, when *qualitatively* interpreted, we can say it is reasonable that the cost of power outages to the Australian economy is in the order of billions of dollars per annum.

These consequences are not definite, nor should they be in a discussion about the precautionary principle: they are *potential* downside consequences to the Australian populace. The precautionary principle states that the size of these potential events warrants robust evidence that the consequences of the policy are negligible or manageable. That is, if the consequences are large and there is scientific uncertainty regarding the policy, then the burden of proof switches to the policy proponent. The proponent must demonstrate, using robust scientific evidence, that the likelihood of such downside events resulting from the introduction of the policy, constitute an acceptable risk to the Australian community.

Given the *potential* downside economic and human consequences of underfunding electricity distribution firms, and the modelling issues surrounding the model accepted by the AER, the circumstances outlined in this report suffice to invoke the precautionary principle. However, the degree of precaution and the proportionality of the response of the precautionary principle is related to the size of the opex changes: the larger the changes, the more substantial the evidence required against the potential downside consequences. Changes to opex of up to one-third require a reversal of evidentiary burden: specifically, robust, scientific evidence that the reduction in opex will change the bushfire and blackout risks by an acceptably small or negligible amount must be provided to support such cuts.

Rather than taking the extremes of accepting the modelling (i.e. ignoring the issues) or accepting inaction (because there are so many problems with the modelling), we recommend a reasonable and pragmatic ‘middle ground’ which takes into account both forces (see Section 5).

4.4 Interaction of model uncertainty and the precautionary principle

Model uncertainty and the precautionary principle have different consequences depending on their use. An issue arises when academic models are applied to generate policy recommendations. In academia, a model may be accepted for publication based on, as is common, mathematical tractability, rather than having a strong relationship to a target system. The consequences of publishing such a model are low for the parties involved: journals run by academics have a preference for certain types of models, and other academics cater to those preferences. Economic models usually claim to ‘explain’ some particular phenomenon, without detailing the weaknesses of the model, or instances where a model cannot be applied [50]. Thus the (downside) consequences of publishing academic models are low, and so the precautionary principle is not typically required or invoked.

In contrast, the consequences of applying those same models to policy may be high. [50] discusses the philosophical approach to economics which maps model generation (i.e. ‘theory’) to cases of potential applicability of those models. This ‘case-based’ approach endorses the view that the individual applying the model must make a decision as to whether the model assumptions are approximately true (and so the model can be used as an approximation) or not.[26] Thus the burden of 'appropriate application' rests with the researcher judging whether the model reflects the applied target system sufficiently. This is a highly subjective approach, open to influences by powerful stakeholders and political inclination.

5 Recommendations of RESPAresearch

RESPAresearch has formal documented processes which mitigate overstating results as well as politicisation and misuse of science and scientific models. Given these processes, in the context of this report, we recommend a two step approach:

1. We recommend performing further research into the efficiency levels of firms, using both (moderated via simulation) neoclassical production theory and also empirically-based behavioural and psychological theories. This is likely to produce more robust models, or at least provide further insight into the degree of robustness of the models. A number of complementary, feasible research avenues are available, which include, but are not limited to:
    - ▷ Analyse capital inefficiencies, which can be done in a number of ways, including a comparison to an optimal network, which accounts for idiosyncracies of each individual firm. That is, a geographically optimal network is determined given the geographical output requirements, and an 'inefficiency' metric provides a measure of how far the actual network is from the optimal network.
    - ▷ Appropriately incorporate price differences across firms, which would require obtaining accurate input price differences across all jurisdictions in the sample, including Ontario, all areas of New Zealand, and the within-state differences in Australia (e.g. rural versus urban).
    - ▷ Conduct analysis of uncertainty aversion of customers, which would provide an indication of customer preferences for lower costs at the expense of potential unknown consequences of possible underfunding of these firms. This is a customer-led approach, and would provide an indication of the value the customers place on the current service level, in relation to an unknown comparator.
    - ▷ Accounting for the statistical errors inherent in the SFA model to some acceptable level, rather than simply assuming the mean or median firm reflects all firms. This could potentially be done by narrowing the range of estimated efficiency scores by the range of the 95% confidence interval in Figure 3.1(a).
    - ▷ Other feasible simulations include the modelling of the complexity of network structures and agent behaviours within the context of analysing comparative efficiency levels. We are currently creating the structure of an agent-based model of distribution networks to be able to dictate the data required to calibrate the model appropriately, and then perform comparative analysis in a realistic, rather than 'black box', comparative framework.
2. We recommend acceptance of firm efficiency *rankings* rather than the absolute size of the differences; this is a qualitative approach considering the complexity of the systems involved,

---

[26] Neoclassical economics is not useless; rather it is highly useful in appropriate circumstances, when uncertainties are respected.

and the significant model uncertainty present. The results from Step 1 are then used for create an interval for opex changes: the percentage opex allowance change for each firm can be determined as equally spaced points in such an interval. For example, the interval could be $[-10\%, 2\%]$, with the most inefficient (estimated) firm having imposed a 10% reduction in opex and the most efficient (estimated) firm having imposed a 2% increase in opex. We understand that this range is relatively arbitrary, but it is a qualitatively *reasonable* recommendation in the context of the precautionary principle, questionable model robustness, and model uncertainty.

This approach would provide a reasonable, ongoing incentive for cost savings and be consistent with the NER while ensuring the precautionary principle is upheld when scientific uncertainty persists.

Appendix 1. The value of simulation

Simulations are a valuable but highly underused tool in economic theory and application. [51] and [52] discuss in detail the benefits for economics and the reasons for the underuse of simulations in economics.[27]

A key reason for the underuse of simulations is the preference of most economists (most likely by training) for the use of deductive mathematical proofs. That is, neoclassical economic theory is based on mathematical assumptions and deductive reasoning. [51] notes that in these models, increased certainty is bought at the expense of reduced scope: the high degree of certainty achieved by mathematical deduction from a set of assumptions is paid for by the high degree of uncertainty in using the result outside the model. In applications like the one at hand, where millions of dollars and community safety is potentially at stake, model certainty must yield to model reality.

Simulations are vastly more flexible than mathematical models in that they require fewer assumption to produce useful results, and so can be used to ameliorate the problem of overconstraining assumptions. Simulations are significantly less overconstrained relative to neoclassical models, as fewer assumptions are made for reasons of tractability. Thus good simulations are closer to reality, and so closer to the target system: in this case, how electricity distribution firms actually operate and the inefficiencies inherent in that operation.[28] Appropriate simulations are essentially *in silico* experiments of systems significantly closer to reality.

The simulation analysis conducted in this paper would be difficult to prove mathematically. Even if a mathematical proof did exist, the question of how relevant the proof is in application still remains. This would have to be achieved by assuming a distribution with certain parameters, and would produce extremely precise results, a level of precision unwarranted in such applied work. A similar level of precision can be achieved in the few lines of code attached with this report, by increasing the simulated sample size.

Simulations are a tool which should not be ignored in applications, including within neoclassical economics, as it allows the assessment of the magnitude and implication of any deviation from the underlying assumption or assumptions. Computer simulations generate surrogate experience, and can improve learning in an experience-poor domain if they are used wisely, with clear attention to limitations [54]. In the way simulations have been used in this report, they are a highly relevant and useful tool in assessing the magnitude of the deviations from theory, while remaining within the realm of that theory.

Appendix 2. Interpreting a QQ-plot

Given a set of ordered data and a candidate distribution (i.e. a distribution which the data may have come from), a QQ-plot is a visual plot of the quantiles of the data against the quantiles of the distribution. An $a\%$ quantile is a datum in the ordered data where $a\%$ of the data is below this point.

---

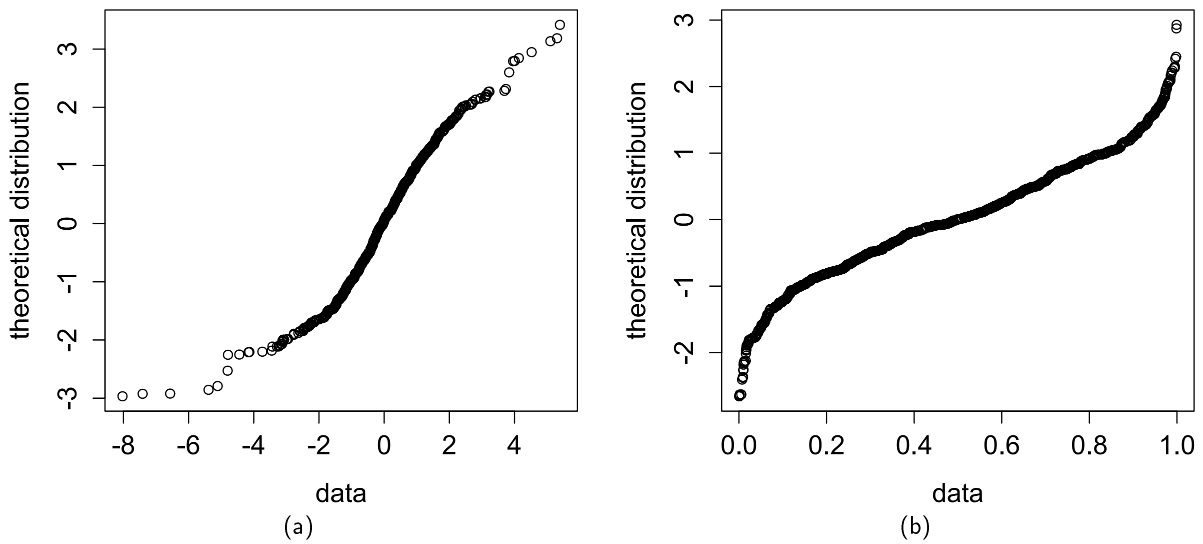
[27] These references are recommended reading (with an open, scientific mind) for anyone interested or skeptical of simulations in economics. As noted in [52], economists are "still reluctant to follow physicists in embracing computer simulation as an important tool in the search for theoretical progress."

[28] RESPAresearch has the expertise required to conduct the entirety of this efficiency analysis using simulated, realistic assumptions, unconstrained by neoclassical production theory.

Consider the 10% quantile: the point in the data (respectively, in the distribution) where 10% of the data points are below this point (respectively, 10% of the probability mass is below this point on the candidate distribution). This will give two points: one from the data and one from the theoretical distribution; these values constitute a single point in a scatter plot. The 10% quantile is then change marginally to, say, 11%. The points in the data (respectively, the theoretical distribution) at which 11% of the data (respectively, probability mass) is below this point, constitute another point in the scatter plot. Instead of choosing each percentage, all the data is optimally used when N quantiles are considered (and thus N points on the scatter plot), where N is the number of data points in the data set. The scatter plot of these quantiles is the QQ-plot.

The purpose of the QQ-plot is that if the data came from a distribution, changing from 10% to 11% will result in approximately the same (linearly scaled) change in the quantile values for both the data and the theoretical distribution. When the N quantiles are plotted, the relationship will be visually close to linear. However, when the data differs from the candidate distribution, then the change from, for example, the 10% to 11% quantile, will result in different (linearly scaled) changes in the quantile values for the data relative to the theoretical distribution. This will appear as a non-linearity in the scatter plot of the quantiles. Thus a non-linearity is a visual indication of the data not arising from the candidate theoretical distribution.

A QQ-plot can give further information: specific patterns in a QQ-plot can indicate specific differences in the data relative to the candidate distribution. Figure 5.1 demonstrates and discusses the patterns associated with differences in the 'weights' of the tails of the distribution.



(a)

(b)

Figure 5.1:
(a) The non-linearity in this QQ-plot indicates that the distribution of the data differs from the theoretical distribution. This 'S'-shaped non-linearity means that the distribution of the data has heavier (i.e. 'fatter') tails than the theoretical distribution.
(b) Similar to (a), there is significant non-linearity in the QQ-plot. This non-linearity is an 'inverse S'-shape, which indicates that the distribution of the data has lighter (i.e. 'thinner') tails than the theoretical distribution.

Appendix 3. The firm in an evolutionary system

One specific approach to modelling and understanding economic production within the complexity framework is the view that firms can be naturally viewed as an evolutionary system. Such a system uses the evolutionary algorithm, which applies variation and selection to a given population [53, 54]; biological and cultural evolution are simply examples of this algorithm.

The unit of selection is of central importance in variation and selection within an evolutionary framework. The unit of selection in an economy (or industry) are not firms, just like organisms are not the unit of selection in biological evolution [53]. Every firm has either an explicit, or more likely, theoretical business plan, which is a comprehensive document defining the purpose of the business, its strategies, its production plans, its technology and skill requirements, its economic model, etc [53]. This business plan can in principle be written down for every firm. Each such business plans contains 'modules': a module is a component of a business plan that has provided in the past, or could provide in the future, a basis for differential selection between businesses in a competitive environment. Modules are the units of selection in economic evolution [53].[29] Modules are a malleable concept, and are analogous to genes within biological evolution or memes within cultural evolution. Given this basis, we can now discuss variation and selection.

Variation between firms in business plan modules should not be a controversial statement: different firms have different operational procedures, different cultures, different strategies, different technologies, and most importantly, different historical paths, all of which influence the modules. These variations come from current and historical choices, as well as current and historical external factors. When viewed as an evolutionary system, firm differences (i.e. variations) are not only expected but necessary [54]. This view provides a simple way to reconcile the persistent variation in 'neoclassical efficiency' distributions across time, space and industry.[30]

The fitness of a firm in economic evolution has a relationship to profitability, as firms that make consistent losses will be driven out of the industry. However, after earning survival level profit, firms can define different objectives rather than solely profitability; for example, market share [54]. Firms attempt to generate profits (or other objectives above the survival level) by combining and enacting all the modules within their business plan [53]. Profitability over some prolonged period is a necessary condition for firm survival. It, however, does not imply profit maximisation. The variation within modules generates variation between firms and thus allows for differential selection. Selection in economic evolution is determined by consumers purchasing some products and not others: consumers select for some firms and against others, with this selection determining fitness [53, 54]. Selection occurs more rapidly in those environments with more competitors, and is significantly retarded in a natural monopoly setting. These reduced pressures are, however, irrelevant to the current discussion as they simply provide a *prima facie* reason for regulation. The focus of this report is on the validity of comparing firms, which necessitates a focus on the variation component of the evolutionary algorithm, rather than the selection component.

---

[29] The definition of a module is not circular as it represents, in a simple way, intuitive concepts such as strategy, pricing plans, and so on, but that cannot be reconciled into a single context: a pricing module would look very different to a strategic takeover module.

[30] This method does not simply assume the wanted result, as it uses observations that firms are different, and tries to reconcile those differences within the framework of how firms actually operate. The reason for the differences comes from differences in modules (i.e. differences in firm practices).

Variation in modules implies differences in business plans, which gives rise to variations in economic fitness. A given firm, at a given point in time, has a current business plan (i.e. a set of modules). This produces a certain level of fitness for that firm in the current environment. This fitness is a single point on that firm's 'fitness landscape', which is a mapping from the space of 'business plans' to the fitness of each implemented business plan, for that firm, in the current environment. This fitness landscape reflects the current environment, the current business plan (which is dependent on history), the costs of movement to any other business plan, and other factors.[31] Furthermore, the fitness landscape is dynamic; it is constantly changing due to internal and external influences. At times it may be a relatively stable and linear system, and at other times it may be complex and non-linear [54]. Just like in biological and cultural evolution, this fitness landscape is a theoretical construct, and cannot be determined prior to implementation of a given business plan.

Given two firms, applying the business plan of firm 1 to firm 2 may result in vastly different fitness. This is a result of firm 2 having a different fitness landscape; significant movement away from firm 2's current business plan may require costly reversal of historical capital and contracts, or focus on different customers, or face a different physical or cultural geography. Such an imposition may force firm 2 into sustained losses and possible selection out of the population. Within the evolutionary framework, it is without basis to assume that all firms have the same fitness landscape – it is equally unreasonable to assume that two given organisms of the same species have the same fitness landscape. However, this is precisely the assumption that is required for 'efficiency' comparisons between firms to be valid.

This evolutionary model of firm behaviour is consistent with observed firm behaviour [54]: for instance, advertising and marketing affect selection; adoption of best-industry practices affects modules within business plans; historical and physical factors matter in current decision making by determining the fitness of different business plans. Furthermore, the evolutionary model is a model that acknowledges firms, industries and the economy as complex systems, which is consistent with the observed firm behaviour of using heuristics, control and learning rather than forecasting [11]. Acknowledging the view that industries and the economy are complex, evolutionary systems means trading-off certain forms of quantification and forecasting for a deeper, observation-consistent understanding of those industries. Harnessing complexity involves acting sensibly without fully understanding how the world works [54]. Understanding how firms fit within the evolutionary framework, including natural monopolies, allows intelligent intervention in designing and implementing incentives in a pragmatic way, to guide firms, under their own volition, to become fitter within their own landscape.

---

[31] Small changes in business plans will typically lead to small changes in profitability; large changes may lead to large changes in fitness, both in a positive or negative direction (for a currently profitable firm, the negative direction is more likely if the large change is purely random). This type of landscape is called a partially correlated fitness landscape.

References

[1] AER (2015) *FINAL DECISION ActewAGL distribution determination 2015—16 to 2018—19 Attachment 7 – Operating expenditure*, p.7-92; AER (2015) *FINAL DECISION Ausgrid distribution determination 2015—16 to 2018—19 Attachment 7 – Operating expenditure*, p.7-97; AER (2015) *FINAL DECISION Endeavour Energy distribution determination 2015—16 to 2018—19 Attachment 7 – Operating expenditure*, p.7-95; AER (2015) *FINAL DECISION Essential Energy distribution determination 2015—16 to 2018—19 Attachment 7 – Operating expenditure*, p.7-96.

[2] Lawrence, D., Coelli, T. and Kain, J. (2014) *Economic Benchmarking Assessment of Operating Expenditure for NSW and ACT Electricity DNSPs*, Economic Insights, Report prepared for Australian Energy Regulator.

[3] Mas-Colell, A., Whinston, M.D. and Green, J. R. (2009) *Microeconomic Theory*, Oxford University Press, Oxford, U.K., p.128, 139, 146.

[4] [1]p.7-39; p.7-42; p.7-40; and p.7-41 respectively.

[5] *National Electricity Rules Version 72*, viewed 1 August 2015, s.6.5.6(a).

[6] Zeckhauser, R. (2014) *New Frontiers Beyond Risk and Uncertainty: Ignorance, Group Decision and Unanticipated Themes*, in Machina, M. and Viscusi, K. (eds.) Handbook of the Economics of Risk and Uncertainty, Elesvier, Oxford, U.K.

[7] Tversky, A. and Kahneman, D. (1974) *Judgement under uncertainty: heuristics and biases*, Science, vol.185, no.4157, pp.1124-1131; Kahneman, D. and Tversky, A. (1979) *Prospect Theory: An Analysis of Decision under Risk*, Econometrics, vol.47, no.2, pp.263-291.

[8] van der Heijden, K. (2005) *Scenarios: The Art of Strategic Conversation*, John Wiley and Sons, West Sussex, England; Andersen, T.J. (2013) *Short Introduction to Strategic Management*, Cambridge University Press, Cambridge, U.K., p.186-187; Schwenker, B. and Wulf, T. (Eds) (2013) *Scenario-based Strategic Planning: Developing Strategies in an Uncertain World*, Springer Gabler

[9] Aiginger (1987) *Production and decision theory under uncertainty*, Basil Blackwell, Oxford, U.K., p.197.

[10] [1]p.7-9; p.7-8; NA; and p.7-8 respectively.

[11] Cyert, R.M. and March, J.G. (2013, reprint of 1963) *A behavioral theory of the firm*, Blackwell Publishers, Oxford, U.K.

[12] Romer, D. (2006) *Do firms maximize? Evidence from professional football*, Journal of Political Economy, vol. 114, no. 2, pp. 340-365.

[13] March, J. and Simon, H. (1993) Organisations (second edition), Blackwell Publishers, Oxford, U.K.

[14] Camerer, C. and Weber, M. (1992) *Recent developments in modelling preferences: uncertainty and ambiguity*, Journal of Risk and Uncertainty, vol. 5, pp. 325-370.

[15] Morroni, M. (1992) *Production process and technical change*, Cambridge University Press, Cambridge, U.K., p.36-37.

[16] [1]p.7-39; p.7-43; p.7-41; and p.7-42 respectively.

[17] [5], s.6.5.6(e)(4).

[18] [1]p.7-43; p.7-47; p.7-45; and p.7-46 respectively.

[19] Landsburg, S. (2013) *Price theory and applications* (9th edition), Cengage Learning, Stanford, CT.

[20] Bewley, T.F. (1999) *Why wages don't fall during a recession*, Harvard University Press, Cambridge, U.S.A.

[21] AER (2015) *AER Final decision Essential Energy distribution determination - Essential Energy 2015 - Opex model - April 2015*, (MS Excel file); AER (2015) *AER Final decision ActewAGL distribution determination - ActewAGL 2015 - Opex model - April 2015*, (MS Excel file).

[22] Australian Bureau of Statistics (2015), *Average Weekly Earnings, Australia, Nov 2014*, cat. no. 6302.0, ABS, Canberra.

[23] Australian Bureau of Statistics (2015), *6302.0 - Average Weekly Earnings, Australia, Nov 2014, Key Figures*, viewed 1 August 2015, http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/6302.0Main+Features1Nov%202014?OpenDocument.

[24] [1]p.7-54; p.7-58; p.7-56; and p.7-57 respectively.

[25] Battese, G.E. and Coelli, T.J. (1992) *Frontier Produciton Functions, Technical Efficiency and Panel Data: with Application to Paddy Farmers in India*, The Journal of Productivity Analysis, vol. 3, p.165.

[26] Aitkin, M., Francis, B. and Hinde, J. (2005) *Statistical Modelling in GLIM 4 (Second Edition)*, Oxford University Press, Oxford, U.K., p.115.

[27] Razali, N.M and Wah, Y.B. (2011) *Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests*, Journal of Statistical Modelling and Analytics, vol. 2, no. 1, p.21-33.

[28] Economic Insights (2015) *Response to Consultants' Reports on Economic Benchmarking of Electricity DNSPs*, Report prepared for the AER

[29] [1] p7-74; p.7-79; p.7-77; and p.7-78, respectively.

[30] Kwan, S.H. (2006) *The X-efficiency of commercial banks in Hong Kong*, Journal of Banking and Finance, vol. 30, pp. 1127-1147.

[31] Jiang, G., Wong, J., Tang, N. and Sze, A. (2004) *Banking sector competition in Hong Kong - measurement and evolution over time*, Hong Kong Monetary Authority.

[32] Kwan, S.H. and Eisenbeis, R.A. (1996) *An Analysis of Inefficiencies in Banking: A Stochastic Cost Frontier Approach*, Federal Reserve Bank of San Francisco Economic Review, no. 2, pp. 16-26.

[33] Clark, J.A. and Siems, T.F. (2002) *X-Efficiency in Banking: Looking beyond the Balance Sheet*, Journal of Money, Credit and Banking, vol. 34, no. 4, pp.987-1013.

[34] Meador, J.W., Ryan Jr., H.E. and Schellhorn, C.D. (2000) *Product Focus versus Diversification: Estimates of X-Efficiency for the U.S. Life Insurance Industry*, in Harker, P.T. and Zenios, S.A. (eds) *Performance of Financial Institutions - Efficiency, Innovation, Regulation*, Cambridge University Press, Cambridge, U.K.

[35] Yuengert, A.M. (1993) *The measurement of efficiency in life insurance: estimates of a mixed normal-gamma error model*, Journal of Banking and Finance, vol. 17, pp. 483-496.

[36] Anderson, R.I., Lewis, D. and Zumpano, L.V. (2000) *X-inefficiencies in the residential real estate market: a stochastic frontier approach*, Journal of Real Estate Research, vol. 20, no. 1, pp. 93-103.

[37] Agasisti, T. and Belfield, C. (2014) *Efficiency in the Community College Sector: Stochastic Frontier Analysis*, CCRC working paper no. 69.

[38] Ashford, N.A. (2007) *The Legacy of the Precautionary Principle in US Law: The Rise of Cost-Benefit Analysis and Risk Assessment as Undermining Factors in Health, Safety and Environmental Protection*, in de Sadeleer, N. (ed) *Implementing the Precautionary Principle: Approached from the Nordic countries, the EU and USA*, Earthscan, London, U.K.

[39] *Protection of the Environment Administration Act 1991 (NSW),* viewed 1 August 2015, http://www5.austlii.edu.au/au/legis/nsw/num_act/poteaa1991n60512.pdf.

[40] *Telstra Corporation Limited v Hornsby Shire Council (2006)*, viewed 1 August 2015, http://www.lawlink.nsw.gov.au/lecjudgments/2006nswlec.nsf/000000000000000000000000 0000000/fdf89ace6e00928bca25713800832056?opendocument.

[41] AER (2014), *AER Draft decision Essential Energy distribution determination - Essential Energy 2014 - Category Analysis Metrics - November 2014*, (MS Excel file).

[42] Economic Insights (2014), *DNSPData AusNZOnt 16Oct2014x*, AER, (MS Excel file).

[43] Teague, B., McLeod, R. and Pascoe, S. (2010) *2009 Victorian Bushfires Royal Commission Final Report Summary*, Government Printer for the State of Victoria.

[44] Victoria State Government (2007) *January Supply Interruptions - Executive Summary*, viewed 24 August 2015, http://www.energyandresources.vic.gov.au/energy/safety-and-emergencies/energy-supply-emergencies/january-supply-interruptions-executive-summary

[45] Allianz (2012) *Power trip* (Energy risks - special topic), Allianz, pp. 28-33.

[46] Other US power outage estimate

[47] Pilkey, O.H. and Pilkey-Jarvis, L. (2007) *Useless arithmetic*, Columbia Univesity Press, New York

[48] Foster, J. (2006) *Why Is Economics Not a Complex Systems Science?*, Journal of Economic Issues, vol. 40, no. 4, pp. 1069-1091.

[49] Jackson, M.O. (2008) *Social and economic networks*, Princeton University Press, N.J.; Deguchi, H. (2004) *Economics as an agent-based complex system*, Springer-Verlag, Tokyo; Tesfatsion, L. and Judd, K.L. (2006) *Handbook of computational economics, agent-based computational economics, volume 2*, North-Holland, Amsterdam; Aoyama, H., Fujiwara, Y., Ikeda, Y., Iyetomi, H. and Souma, W. (2010) *Econophysics and companies, statistical life and death in complex business networks*, Cambridge University Press, Cambridge, U.K.

[50] Gilboa, I., Postlewaite, A., Samuelson, L. and Schmeidler, D. (2014) *Economic models as analogies*, The Economic Journal, vol. 124 (August), pp. 513-533.

[51] Reiss, J. (2011) *A Plea for (Good) Simulations: Nudging Economics Towards an Experimental Science*, Simulation and Gaming, vol. 42, no. 2, pp.243-264.

[52] Lehtinen, A. and Kuorikoski, J. (2007) *Computing the Perfect Model: Why Do Economists Shun Simulation?*, Philosophy of Science, vol. 74, pp.304-329.

[53] Beinhocker, E.D. (2006) *The origin of wealth*, Harvard Business School Press, Cambridge, U.S.A.

[54] Axelrod, R. and Cohen, M.D. (2000) *Harnessing Complexity*, Basic Books, New York