National Physical Laboratory

NPL REPORT IEA 19

AN ECONOMIC MODEL FOR THE VALUE ATTRIBUTABLE TO HIGH-QUALITY CALIBRATIONS BY REDUCING MISTAKES IN CONFORMANCE TESTING

MIKE KING, SRAJAL NAYAK

MAY 2023

An Economic Model for the Value Attributable to High-Quality Calibrations by Reducing Mistakes in Conformance Testing

Mike King, Srajal Nayak
Analysis & Evaluation Team, Partnerships Directorate

ABSTRACT

Since NPL only earns income by selling services to customers wanting ‘primary’ calibrations, the value (economic benefit) created by the ‘fanout’ is missing from NPL’s revenue. The first part of this document develops an economic model for the value created when using measurement for conformity testing. This model shows how measurement information creates value by reducing mistakes (fewer “false positives” and “false negatives”) in conformity testing. The second part of this document introduces calibration as a perturbation on top of these measurement activities. Moreover, there can be value created from this reduction in uncertainty, which is where the services of NPL and the calibration laboratories play a key role.

Previous studies have explored the economic loss from making type-1 errors (“false positives”) and type-2 errors (“false negatives”) in the context of conformance testing. However, such studies do not treat the conditional probabilities of making type-1 and type-2 errors as “choice variables” that are under control of a decision maker. By treating these conditional probabilities as exogenous, they do not internalise the trade-off between type-1 and type-2 errors. In this paper, we address this gap in the literature by endogenizing the conditional probability of a type-1 error (α) and the probability of a type-2 error (β). The decision maker tries to minimise the total cost incurred due to these errors, subject to the trade-off that exists between them.

This is an economics study that borrows a little from metrology as opposed to a study by engineers that borrows a little from economics. Firstly, our analysis gives formulae for the relative uncertainty of the measurement process (σ) and a lower bound for the scrap rate (τ) as functions of the conditional probabilities for type-1 and type-2 errors (α and β). Secondly, this analysis yields a formula for the economic benefit created by using ‘primary’ calibrations to reduce the uncertainty of the measurement process. Hence, this document provides the mathematics behind a ‘calculator’ giving an estimate of the benefit from ‘primary’ calibrations based on assumptions about the parameters. Finally, this model will really come to life once we have estimates of the core parameters from a forthcoming measurement survey.

ISSN 2633-4194

DOI: https://doi.org/10.47120/npl.npl.IEA19

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW

This work was funded by the UK Government's Department for Science, Innovation & Technology through the UK's National Measurement System programmes.

Approved on behalf of NPLML by
David Skelton, Strategic Programme Leader (Partnerships Directorate)

CONTENTS

1 INTRODUCTION

The National Physical Laboratory (NPL) only supplies calibrations to a small fraction of the organisations wanting calibration services. NPL's customers are willing to pay a premium for high-quality (primary) calibrations, whereas most users are content with less accurate (secondary) calibrations supplied by calibration labs at marginal cost. These 'secondary' calibrations could not exist without the 'primary' calibrations supplied by NPL and other National Measurement Institutes. We refer to the traceable calibrations taking place outside of NPL as the 'fanout'. Since NPL only earns income from customers wanting 'primary' calibrations, the value (economic benefit) created by the fanout is missing from NPL's revenue. This document outlines a stylised economic model that will help subsequent studies to value the fanout.

There has been no serious valuation of the fanout, nor is there a clear account of the fanout derived from established economic concepts and models. So, as a first step towards understand the economics of the fanout, we need a stylised microeconomic model that can be solved (as far as possible) analytically to help us to develop the required insights and intuitions. The type of economic phenomena we wish to model is complicated, involving an interaction between a public good (NPL's realisation of the SI units) and a private good (measurement information used by firms). However, any hope of achieving the desired analytical tractability requires that we begin with a core model that is based around some strong simplifying assumptions.

The approach taken in this study is to try to value calibration services in terms of the extent to which high-quality calibrations can help firms to reduce mistakes derived from measurement errors. Hence, this study begins by setting up a stylised model for the cost of measurement errors when a measuring process is used for conformity testing. (It is important to keep in mind that some kind of conformity testing would still take place irrespective of the presence of NPL or the calibration labs.) Although maybe not all measurement activity can be characterised as a type of conformity testing, it's reasonable to suppose that a large fraction to day-to-day measurement activity does fit into this schema. Hence, the analysis in this study is hopefully relevant to a large fraction of the benefits of measurement spending even if it's not the full story.

Although, this is still very much an economics study it borrows a little from metrology. That is, the next step towards building the model is to recognise that these costs depend on the size of the standard deviation of the measurement process; and this can be split into two parts:

- The first part comes from uncertainties (errors) that aren't associated with calibration.
- The second part is a component of the uncertainty (systematic errors) that can be reduced or almost eliminated by high-quality calibration.

By this means, calibration helps to lower the standard deviation of the measurement process, which reduces the cost of measurement errors for a firm engaged in conformance testing of its products.

Previous studies have explored the economic loss from making type-1 (false positive) and type-2 errors (false negatives) in the context of conformance testing. However, such studies do not treat the probabilities of making type-1 and type-2 errors as "choice variables" that are under control of the decision maker. By treating these probabilities as exogenous, they do not internalise the trade-off between type-1 and type-2 errors.

Our study builds on the exiting literature, and its main contribution is to endogenize the probabilities of making type-1 and type-2 errors. The decision maker (firm) tries to minimize the total cost incurred due to these errors, subject to the trade-off that exists between them. Firstly, our analysis gives formulae for the relative uncertainty of the measurement process

(σ) and a lower bound for the scrap rate (τ) as functions of the probabilities for type-1 and type-2 errors (α and β). Moreover, these formulae can be inverted to give these probabilities (α and β) in terms of the model's underlying parameters (σ and τ).

Secondly, this analysis yields a formula for the economic benefit created by using 'primary' calibrations to reduce the uncertainty of the measurement process. Hence, this document provides the mathematics behind a 'calculator' giving an estimate of the benefit from 'primary' calibrations based on assumptions about the parameters.

Finally, this document considers a stylised example of a conformance testing scenario based on the following assumptions:

- In the baseline scenario, a firm uses primary calibrations provided by NPL (or other top-tier calibration labs), which eliminates any calibration-related uncertainties. Furthermore, as a generally accepted heuristic, any lab engaged in conformance testing strives for a minimum Test Accuracy Ratio (TAR)[1] of 4:1; implying that the extra uncertainty introduced through poorer calibration could be as much as 25% of the baseline uncertainty of the measurement process.
- In the baseline scenario (where the firm uses primary calibrations), the conditional likelihood of a "false positive" is 5% and the conditional likelihood of a "false negative" is 20%. (That is, the example features conventional probabilities for "false positives" and "false negatives" in the baseline scenario.)
- If a customer returns a defective item, then the firm compensates them by providing a free replacement. Consequently, a "false negative" (type-2 error) is presumed to be twice as costly as a "false positive" (type-1 error).

According to this stylised example, our formulae show that if a firm is currently benefiting from access to accurate calibrations, then a ~3% increase in the standard deviation of the measurement process – arising from losing access to accurate calibrations – would cause an 0.8 percentage point increase in the cost of mistakes (as a fraction of the output from production) during the conformance testing process. Even though this percentage seems small at first sight, it could still amount to large monetary values given that the turnover of, say, the UK's manufacturing sector is around £570 billion.

The headline results for changes in the conditional likelihoods of type-1 and type-2 errors (α and β, respectively) are as follows:

- Increased likelihood of type-1 errors: $\Delta\alpha = 0.053 - 0.050 = 0.003$, and so the percentage increase in the probability of a type-1 error becomes $0.003/0.05 \approx \mathbf{6}\%$.
- Increased likelihood of type-2 errors: $\Delta\beta = 0.214 - 0.200 = 0.014$, and so the percentage increase in the probability of a type-2 error becomes $0.014/0.20 \approx \mathbf{7}\%$.

It is helpful to conclude by giving a brief overview of how the rest of this document is structured: Section 2 discusses the relevant existing literature and the main contributions of this paper. Sections 3 and 4 discuss the general microeconomic framework and the main assumptions, respectively, underlying the model developed in this paper. Section 5 develops an economic model for the value created by using measurement for conformity testing. This model shows how measurement information creates value by reducing mistakes during conformity testing. Section 6 introduces calibration as a perturbation on top of these measurement activities. Moreover, this section explains how the value from production will be affected through this perturbation, which is where the services of NPL and other top-tier

[1] TAR is defined as the ratio of an instrument's accuracy to the accuracy of the standard that is used to calibrate the instrument.

calibration laboratories play a key role. Section 7 concludes the paper and discusses ideas for further work.

2 LITERATURE REVIEW

Before launching into the model developed in our paper it is helpful to begin with a brief review of the relevant literature.

An existing strand of the multidisciplinary literature has explored the value generated by metrology and measurements. Swann (2009) discusses the role of measurement in process control using the simple diagram presented in Figure 1, in which a system can be thought of as having a feedback loop comprised of the following four components: the Production process, the Measurement process, the Evaluation process, and the Control process.

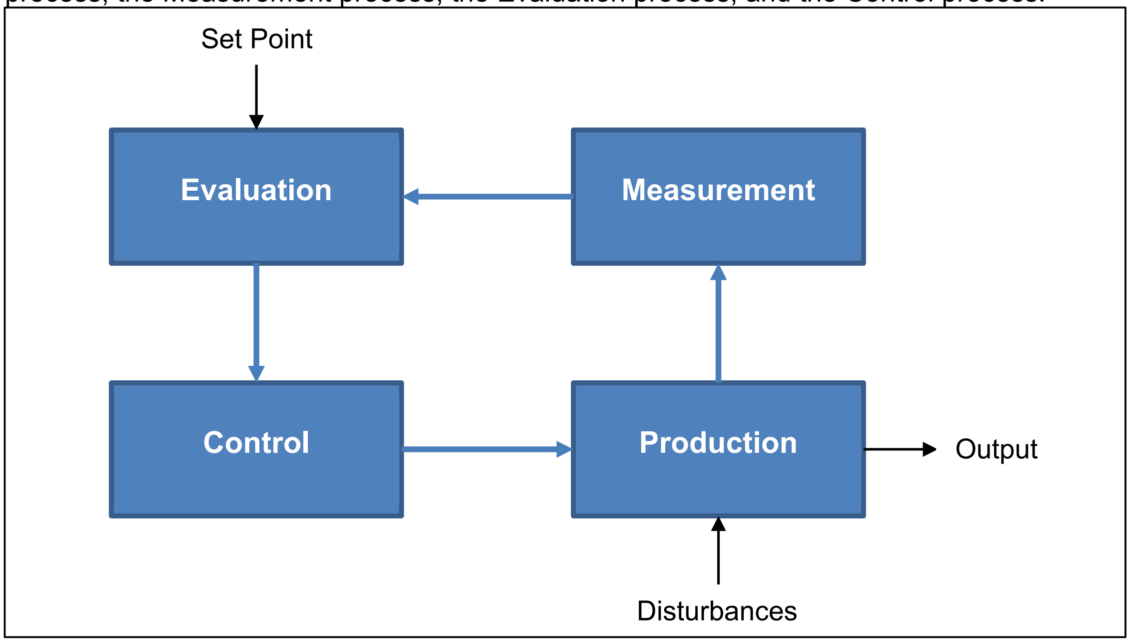

Figure 1: Role of measurement in process control.
***Source:* Based on Swann (2009)**

Production refers to efforts put into assembling the factors of production (capital, equipment, labour, etc.) to produce an output and it might be influenced by various factors, some of which are external "disturbances." *Measurement* provides information about variables that are to be controlled in the production process. Then, *evaluation* involves comparing the measurement of the variables with the desired values of those variables. Depending on the evaluation outcome, the *control* element then feeds necessary action back into the production process. Therefore, measurement plays a key role in this feedback loop and can help in optimizing the entire manufacturing system (Kunzmann et al. (2005)).

Swann (2009) also explores the idea that improvements in measurement can improve decision making and uses a simple example to show how measurement errors can lead to wrong decisions. The example, presented in Figure 2, shows the relationship between the measured value (horizontal axis) and true value (vertical axis) of a variable X that lies between 0 and 1. Let X_C denote the critical value of X, and the decision rule be: take a certain action if $X > X_C$, and take no action if $X \leq X_C$. If measurement were to be perfecty precise, then all measurements of X would lie on the $y = x$ line. However, measurements often occur with an error. And if any measured value of X is only accurate to $\pm r$, then the true value will be in the range $X \pm r$ (depicted by the grey band across the $y = x$ line). These measurement errors can lead to decision errors around the critical value X_C. When the measured value of X is above X_C but the true value is below X_C, the decision maker will take action when no action is required (type-1 error). The blue triangle indicates this situation.

Likewise, when the measured value of X is below X_C but the true value is above X_C, the decision maker will fail to take action when action is required (type-2 error). The red triangle indicates this situation. In this simple example, the distribution of points over the grey band determines the probabilities of making type-1 and type-2 errors.

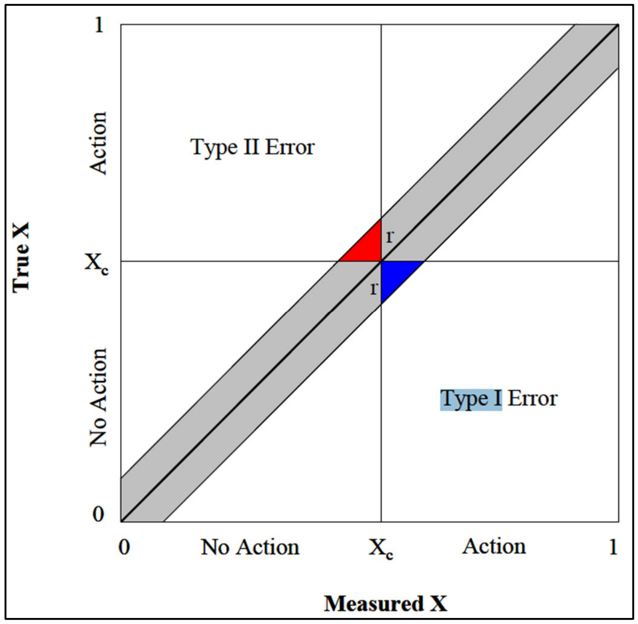

Figure 2: Measurement errors and decision errors.
***Source:* Swann (2009)**

Although Swann (2009) introduces the concept of type-1 and type-2 errors, his report does not go the full mile to set up an optimization problem that involves minimizing the cost of measurement errors subject to the trade-off between type-1 and type-2 errors. One of our contributions in this paper is to address this gap by developing a model for the cost of measurement errors that incorporates the trade-offs between the two types of errors. Understanding the cost of measurement errors is crucial for evaluating the economic value of measurement.

Another strand of engineering-style papers explores questions related to the value of measurement and metrology. Beges et al. (2010) examine the conformance testing scenario in terms of costs associated with measurement, testing, and incorrect decision-making. The study suggests the use of optimised uncertainty approach (Thompson and Fearn (1996)) to model how overall costs of maintaining calibration measurement capabilities (CMCs) in the national metrology infrastructure varies with measurement uncertainty. It defines the overall costs as a sum of testing costs (e.g., costs of providing different qualities of calibration services) and the costs incurred because of incorrect decision making; and talks about the need to strike a balance between both to determine the optimal level of CMCs. Using a numerical methods approach, it plots overall cost with respect to the value of a parameter being measured and the uncertainty surrounding its measurement. However, the paper does not include a model with a closed-form solution that can be derived in a step-by-step fashion, without which it is hard to gain general insights into the relationship between uncertainty and the cost of measurement errors that can be expressed in terms of simple formulae.

Savio (2012) presents a framework to analyse costs and benefits to evaluate the economic impact of metrology in manufacturing. The paper argues that compared to the evaluation of costs, which is usually straightforward, the nature of benefits can often be intangible and hard to quantify. It attacks the problem using the approach of modelling type-1 and type-2 errors and provides general probabilistic expressions for the two types of errors. However, it does not employ any distributional assumptions regarding the errors, and so does not obtain a closed-form expression for the relationship between the two types of error.

Usuda and Henson (2012) present a method for estimating the economic impact of the equivalence of measurement standards.[2] The method facilitates a quantitative analysis of the economic impact, based on a distribution function for the quality of the product and information about the agreement of measurement standards. That is, the quality of the product follows a certain distribution, and only products that lie within the bounds of lower and upper testing limits pass the conformity assessment (rest of the products are discarded). A monetary loss due to the absolute cost of a defective products can then be computed as the unit price multiplied by the probability that the product is defective. The setting in the paper involves two agents, manufacturer, and consumer, both of whom perform conformity testing but have deviations between their respective measurement standards. This deviation results in an additional economic loss in trade due to "false negatives" (products rejected by the manufacturer that would have been accepted by the consumer) and "false positives" (products sold by the manufacturer that are rejected by the consumer). Next, it incorporates a loss function that describes the relationship between the deviation of a measured parameter from its target (nominal value) and the economic impact that may arise due to this deviation, even when the products lie within the testing limits.[3] Finally, it merges the loss function with the prior scenario when there is a deviation between the measurement standards of the two agents (manufacturer and consumer) to estimate the economic impact of the deviation of measurement standard.

The studies discussed above (Beges et al. (2010); Savio (2012); Usuda and Henson (2012)) all pose questions similar to the one we examine in this paper: How can we quantify the cost of measurement errors and quantify the economic value of metrology? However, these studies differ from our paper in the approaches they take. Beges et al. (2010) and Savio (2012) need to employ numerical methods, as their models are not analytically tractable, which makes it harder to see the relationship between uncertainty in measurements, the costs of measurement errors, and the economic benefit that can be attained by reducing or altogether eliminating the uncertainty. We believe that Usuda and Henson (2012) is comparatively closer to the approach we take here, in the sense that their approach also involves making certain distributional assumptions to obtain a closed-form solution.

However, none of the previous studies explores the consequence of making type-1 and type-2 errors "choice variables", under control of the decision maker. By not treating these probabilities as endogenous, they forego the opportunity to internalise the trade-off between type-1 and type-2 errors. That is, given an imperfect measurement process, it is not possible to simultaneously eliminate both type-1 and type-2 errors. In this paper, we address this gap by endogenizing the probabilities of type-1 and type-2 errors. This enables the decision maker to understand the relationship between the two kinds of error and can then try to minimize the total cost of making these errors subject to the trade-off that exists between them.

3 MICROECONOMIC FRAMEWORK

This section set out a general framework for the model developed in this document, which can be summarised as follows:

- Generally, production yields products that conform to a given specification or grade. However, sometimes there is a malfunction so that the production yields defective products that don't meet these requirements. That is, organisations supply products

[2] We can think of calibrations as a way to achieve equivalence of measurement standards.

[3] In our model, we make the simplifying assumption that loss function is zero if the measurement lies within the testing limits (even if it is not exactly equal to the target value). In contrast, Usuda and Henson (2012) allow for a more generalized loss function that further studies could incorporate in our model, as well.

(goods and services) using a potentially faulty production process or technology.

- Suppose that from time to time the production stops for cleaning and maintenance, after which the process (apparatus) is set up again from scratch. Immediately after such a setup, the process always yields non-defective units of their product but, as time goes on, there is a chance that some perturbation (knock) causes the process to go awry, so that it starts to yield defective units of the product.
- Selling non-defective units of the product creates benefit (income) for the organisations. But accidentally supplying defective units of their product may harm end-users and damages the reputation of the organisation. Hence, the organisation incurs a significant cost (fine) from mistakenly selling defective units of their product.
- As part of their conformance testing activities (product verification), organisations measure certain aspects of their products and production processes. These measurements are done routinely and generate valuable information about when to intervene and stop the production process.
- The organisation incurs a cost from shutting down the production process, as it creates deadtime leading to lost production. Hence, organisations wouldn't want to shut down production unnecessarily, which is why they take regular measurements to monitor how their production process is performing.
- Measurements come with an unknown amount of error and so aren't exactly the true value. These errors can be positive or negative, and so the true value can be bigger or smaller than the measured value. The distribution of these errors is known even though the error drawn during a given measurement isn't knowable. Knowing the distribution of the errors allows the construction of confidence intervals.

4 MAIN ASSUMPTIONS

This document offers a microeconomic account of the value of measurement information, in which particular attention is given to the benefits derived from calibration. It uses conformance testing as a scenario for modelling the cost of type-1 (false positives) and type-2 errors (false negatives). The main assumptions underlying the model are as follows:

Assumption 1: It focuses exclusively on the demand for calibrations by organisations based in the UK. We will suppose that it's impossible to buy calibrations from National Measurement Institutes located outside the UK. That is, NPL is the sole provider of primary calibrations for UK-based organisations.

Assumption 2: There is a large, fixed population of organisations that use measurement to produce their products.

- These organisations differ in terms of the importance they attach to measurement. For example, a firm that manufactures infusion pumps – a device used to administer controlled amount of chemotherapy into patients' bloodstreams – will value highly accurate measurements more than a firm that manufactures 1-litre containers for storing milk.
- Organisations also differ in terms of the extent to which accurate calibrations can have an appreciable effect on the overall uncertainty of the measurements they make. That is, some organisations might attach great importance to measurement but get little benefit from high-quality calibration because other sources of error dominate their uncertainties.

Assumption 3: In the model, time can be divided into a large number of distinct periods (days) and what happens in one time period is independent of what happens in the period before or in the period after. That is, state of the production process in a given time period is a memoryless random variable.

Assumption 4: There are two vertically differentiated varieties of calibration service: high quality (primary) calibrations, and inferior quality (secondary) calibrations.[4] The inferior calibrations are identical to one another. That is, there is no horizontal (geographical) differentiation among suppliers of the inferior calibration services.

- NPL is the monopoly supplier of high-quality calibrations, while inferior calibration services are supplied by competing calibration labs.
- One calibrated instrument is used to calibrate many other instruments. We can imagine these ‘secondary’ calibrations to be imperfect copies of the original calibration. The calibration labs come to NPL to buy a high quality (primary) calibration that they then use to supply inferior quality (secondary) calibrations.
- The market for supplying secondary calibrations will not be modelled explicitly. Rather, we will assume that there is perfect competition between many such suppliers.

Assumption 5: Organisations decide whether to pay for calibrations or rely on outside options. Outside options include shutting down production at such regular intervals that it’s unlikely to have malfunctioned during its operation. Alternatively, organisations can accept the cost of sometimes supplying defective parts and compensating adversely effected customers. If organisations decide to pay for calibrations, then they must choose between paying a premium for high quality services from NPL, or instead, rely on inferior but, nonetheless, often perfectly adequate, services from competitively priced calibration labs.

Assumption 6: The marginal cost of supplying high quality and inferior quality calibration services is the same. The monopolistic NPL incurs a fixed cost from having to maintain primary standards that are realisations of the SI units. The inferior calibrations are supplied at marginal cost, but the monopolist (NPL) applies a positive margin to the high-quality calibrations.

5 MODELLING THE COST OF MEASUREMENT ERRORS

5.1 A CONFORMANCE TESTING SCENARIO

People, organisations, and society must make decisions about what to do in certain situations. These decisions typically involve selecting from a range of options with different costs and benefits. Often there is imperfect information about the future state of the world. The agents tasked with making these decisions have an overall goal (e.g., profit maximisation in the case of a firm), and the extent to which this is achieved depends on which option they select and future state of the world. Imperfect information implies that the payoff associated with a given option may not be definite and there could be a distribution of payoffs each occurring with some likelihood.

Information is valuable because it helps people make better decisions, and so reduces the cost of bad decisions. Measurement contributes to the information that people need to make better decisions. The sectors of the economy differ in terms of the importance of measurement information. That is, some sectors of the economy, such as, manufacturing and healthcare, are more reliant on the information that comes from measurements than sectors, such as, finance and real estate.

The analysis in this section is based on a manufacturing firm who conducts conformance testing for the components it produces. However, much of the analysis could equally well be

---

[4] The label ‘inferior’ is not meant to be pejorative, rather it’s an economist’s way of saying that that all users agree that the calibrations direct from NPL are always preferred when prices are equal. There are many users for whom the calibrations from the calibration labs are perfectly adequate.

applied to a radiotherapy department of a hospital deciding whether it needs to temporarily stop treating cancer patients whilst it recalibrates its Linac. Similarly, much of the same analysis applies to a regulator deciding whether they need to suspend an economic activity (e.g., ban cars from a city centre) because the level of some pollutant (e.g. Nitrogen Oxide or NOx) may have passed a legal threshold, so that continuing the activity unabated would be dangerous to public health.

The analysis that follows is based around conformance testing because it sets up a simple binary choice: either the firm continues production for another period, or it shuts down the production process and sets it up from scratch. The first option gives the firm a positive payoff provided that the production process hasn't malfunctioned and started producing defective products, in which case the firm incurs a penalty. The second option entails a period of deadtime during which the plant ceases to produce any output, which entails an opportunity cost for the firm. However, once the production process has been set up again from scratch, the firm can be reasonably sure that it's producing non-defective units of its product.

Imagine that a firm is producing component parts for the aerospace or automotive sector, and that producing and selling these components generates a flow of profits, ν, for the firm. To be specific, let us assume that ν is the daily profit earned when everything is running correctly.

It aids the exposition to talk in terms of days, but it could be any unit of time (e.g., an hour). More generally, the unit of time used in this analysis represents the time taken to shut down production and then set up it up again from scratch. Furthermore, it's assumed that one period is equivalent to the next one, during which the plant is either active and producing outputs or the plant is inactive whilst the process is set up again from scratch.

5.2 SETTING UP THE MODEL

Let $s \in \{A, B\}$ denote the state of the production process at a given moment in time. In a certain interval of time, such as, a working day, a production process is either producing parts that conform properly to the specification, 'A', or the production process has malfunctioned and is producing defective parts, 'B'. The state of the production process on a given day does not depend on its state the day before and does not influence its state on the following day (Assumption 3). It's as if the production process starts afresh each day because it's always set up again from scratch first thing in the morning.

Let $\mathscr{v}_s$ denote the monetary value of the output that's produced when the process is in state $s \in \{A, B\}$. If the production process is in state 'A', then the goods being produced are valuable and can be sold for a profit: $\mathscr{v}_A > 0$. If the production process is in state 'B', then the output is worthless and must be scrapped: $\mathscr{v}_B = 0$.

In an ideal world, the firm would simply multiply $\mathscr{v}_A$ by the number of working days during a year to calculate its annual profit. However, the firm uses a production process that isn't entirely reliable and has a small chance of producing parts that don't meet the required specification.

Assumption 7: The firm's staff can't tell, just by eye, that the process has malfunctioned and has started to produce defective parts. When this happens, the defects always become apparent to customers because the components don't meet their tolerances, and so the defective components are returned to the firm along with a demand for compensation.

Let $\mathscr{p}_A = \Pr\{s = A\}$ denote the probability of the production process continuing to operate correctly in a certain period of time, so that the probability of a malfunction occurring during this period becomes $1 - \mathscr{p}_A = \mathscr{p}_B = \Pr\{s = B\}$. That is, the production process is either producing good parts or defective parts.

Assumption 8: $\boldsymbol{\mathscr{p}_A > \mathscr{p}_B}$. Indeed, we could also suppose that $\boldsymbol{\mathscr{p}_B}$ is very small relative to $\boldsymbol{\mathscr{p}_A}$.

Let c_A denote the cost of unnecessarily shutting down production and scrapping a day's worth of output. That is, c_A is the cost of mistaking good parts for defective parts, and so temporarily ceasing production. (Even if the materials can be recovered, the labour and machine time can't be recovered.)

Assumption 9: The cost of mistaking good parts for defective parts equals the value of the lost output that is unnecessarily scrapped: $\boldsymbol{c_A = v_A}$. (It's basically a case of mistakenly turning off the 'money machine' when you don't have to do so.)[5]

Let c_B denote the cost (penalty) incurred from the firm mistakenly failing to detect that the process has malfunctioned and that the units being produced are defective. This is the cost of recalling all the potentially defective parts sold to customers and offering compensation to them for any harm or inconvenience caused. (Note that with a service the product goes straight to the customer whereas with a good it might go to a warehouse before being sold.) In some cases, there might be reputational harm done to the company or fines issued by a regulator.

Assumption 10: The penalty exceeds the value of the lost output: $\boldsymbol{c_B > v_A}$. To simplify the exposition, we can further assume that $\boldsymbol{c_B = 2 \times v_A = 2 \times c_A}$.[6]

[5] Note that in terms of the UK's GDP, the full economic cost of unnecessarily ceasing production will exceed the lost profits of the firm, as there is also the lost output from having employees standing idle (or even sent home) whilst the production process is set up again by the firm's engineers. Whilst the plant is idle, the labour and capital employed by the firm isn't being used to produce its products. Assuming competitive factor markets for labour and capital, the lost Gross Value Added (GVA) to the economy will equal the firm's daily wage bill plus day rate for renting its plant and machinery. If we further assume that labour and capital are used in fixed proportions, determined by the optimal capital intensity of the production process, then the lost GVA will be proportional to the lost profits.

[6] The assumption that $c_B = 2 \times c_A$ is not necessary for the analysis that follows in this paper. However, it can help with simplifying the exposition. A possible way to think about $c_B = 2 \times c_A$ is through a restaurant analogy. Suppose a restaurant accidentally serves wrong food to a customer (it could be stale leftovers from the day before or a wrong dish that the customer didn't order). After the customer complains, the restaurant serves a replacement dish and does not charge the customer anything to avoid reputational damage. Thus, the cost that the restaurant incurs from serving wrong food (c_B) is twice that of the cost of a dish (c_A).

5.2.1 Firm does not perform measurements to monitor production process

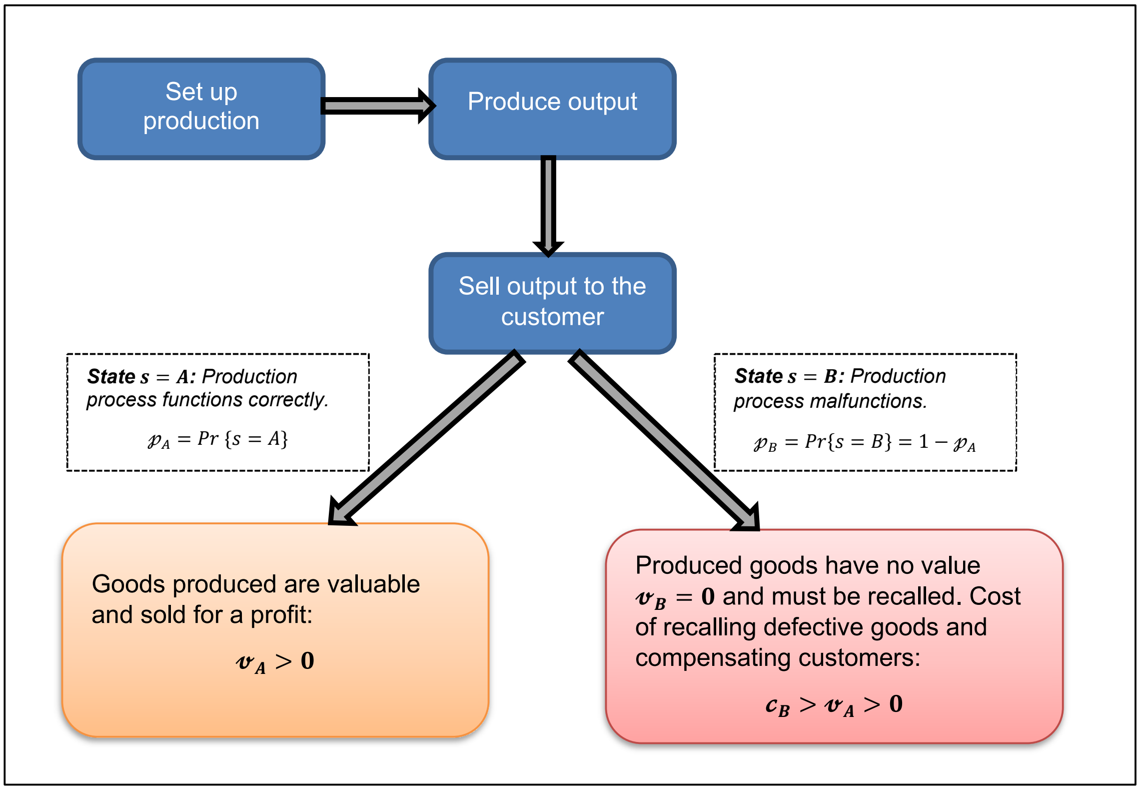

Figure 3: Scenario when the firm does not monitor production process, i.e., it does not perform conformance testing

Figure 3 represents the scenario where the firm does not perform conformance testing to monitor its production process. The situation is somewhat analogous to someone who must take the train when commuting to work but who can't really afford to buy a ticket. The flow of profits, v_A, is analogous to the price of a ticket; the likelihood of a malfunction, p_B, is analogous to the probability of a ticket inspector being on the train; and the cost incurred from dealing with defective parts, c_B, is analogous to the fixed penalty for traveling without a ticket, as well as, of course, having to buy a ticket. We will assume that the presence of inspectors on one day does not influence the likelihood of inspectors being present the following day. It's not possible for the commuter to know what will happen on an individual day, but the costs or benefits of commuting without a ticket become quite predictable over the course of a year. Lastly, it's helpful to imagine that the likelihood of meeting a ticket inspector isn't too high and receiving a fine isn't catastrophic, although, paying it off means going without any luxuries for a few weeks.

If the firm were somehow to always know when the production process had malfunctioned, then they would know when to shut down production and scrap that day's output. In this scenario, the firm would never receive a penalty and their expected (average) flow of profits would be as follows:

$$\mathbb{E}[v_S] = p_A.v_A + p_B.v_B = p_A v_A = p_A c_A.$$

... Equation 1

This is based on the assumptions that $v_B = 0$ and $c_A = v_A$.[7]

If the firm is unable to detect when the production process has malfunctioned, then their profitability will be somewhat less than the theoretical maximum given above. Even though accepting the occasional penalty is a viable option, failure to detect the intermittent malfunctioning of the production process harms the profitability of the firm. The average cost of simply ignoring such malfunctions and accepting the inevitable penalties is $p_B c_B$, so that the firm's expected flow of profits becomes:

$$\mathbb{E}[v_s] - c_B.\Pr\{s = B\} = p_A.v_A + p_B.(v_B - c_B) = p_A c_A - p_B c_B.$$

... Equation 2

Assumption 11: If $\boldsymbol{p_B c_B}$ were greater than $\boldsymbol{p_A c_A}$, then continuing production, whilst accepting the penalties, isn't economically viable. In this study we will assume that $\boldsymbol{p_B c_B} < \boldsymbol{p_A c_A}$, which is equivalent to the following condition on the cost ratio:

$$\frac{p_B c_B}{p_A c_A} < 1$$

It's helpful to use $p_A c_A$ as a baseline against which to compare other monetary values in the analysis. As a proportion of $p_A c_A$, the firm's expected payoff from simply ignoring malfunctions and accepting the penalties becomes:

$$\frac{p_A c_A - p_B c_B}{p_A c_A} = 1 - \left(\frac{p_B c_B}{p_A c_A}\right) = 1 - \tau,$$

... Equation 3

where τ denotes the relative cost of making type-2 errors, as a proportion of the maximum attainable output, so that $(p_B c_B) = \tau \times (p_A c_A)$. τ will be referred to as the 'cost ratio' and is a basic parameter of the model. In some ways, it's helpful to think of τ as a kind of tax on production that has to be paid by the firm. Assumption 11 implies that this 'tax' doesn't wipe out the value of production, i.e., $\tau < 1$.

5.2.2 Firm incorporates measurements to monitor production process

Suppose that there exists a measurement process giving the firm a reasonable likelihood of detecting defective products. For example, taking measurements might correctly detect the malfunction in 80% of the instances in which products being produced are defective.

Assumption 12: The cost of acquiring this measurement information (running the tests and performing analysis) is negligible, and so it can be ignored to a first approximation.

Assumption 13: Lastly, the overall cost of acting on potentially faulty measurement information is less than the cost of simply incurring a penalty when the production process malfunctions. That is, we suppose that the expected cost of acting on potentially faulty information is less than $\boldsymbol{p_B c_B}$, and so the firm decides to make regular measurements to monitor the production process. Figure 4 represents this scenario as a flowchart.

[7] This maximum possible level of profitability is analogous to our commuter having an uncanny prescience that enables them to buy tickets if and only if inspectors are on the train.

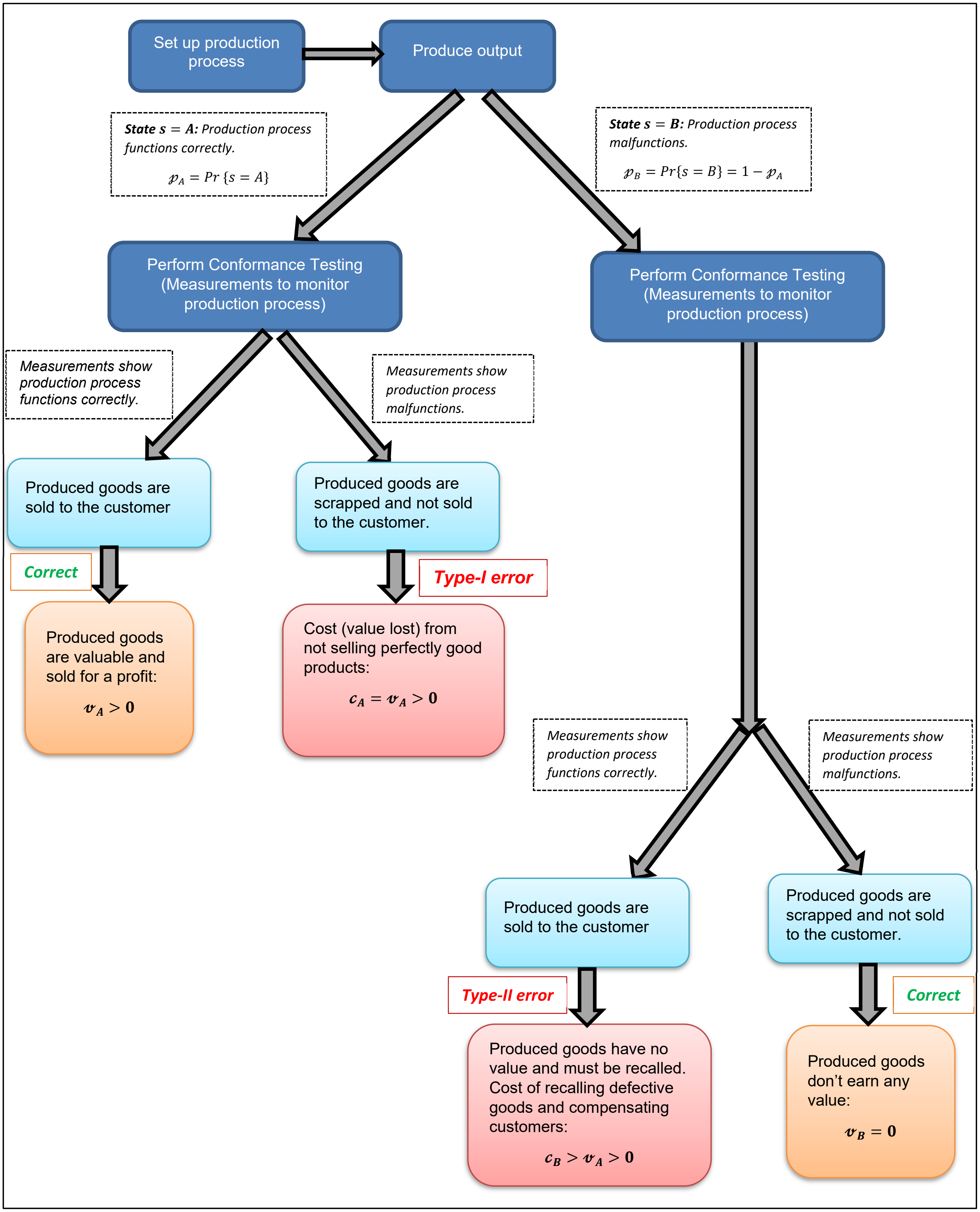

Figure 4: Scenario when the firm monitors production process, i.e., it performs conformance testing

Access to information from taking regular measurement is analogous to our commuter receiving tip-offs that there will be inspectors on a particular train. Imagine that these tip-offs turn out to be correct about 80% of the time. Moreover, suppose that these tip-offs come from a friend, and so don't cost more than a cup of coffee. That is, the cost of acquiring these tip-offs is negligible. Secondly, there is a continuous aspect to these tip-offs, such as, an estimate of the percentage of trains that day that will be containing inspectors.[8] Finally, if these tip offs were 100% accurate, then the commuter would only need to pay for a fare when they know that inspectors will be on the train. This is equivalent to the payoff received when measurement error is somehow eliminated. Notice that even in this ideal scenario, the commuter doesn't get to travel for free: there are still sometimes inspectors on the train just as the production process will sometimes malfunction.

5.2.3 Measurement errors

Suppose there is some critical parameter of the parts being produced that is regularly measured. A practically meaningful deviation of this parameter from its desired value can be set to unity, without much loss of generality. This might represent the tolerance to which units of the product must be produced. For example, the diameter of a piston shaft might need to accurate to the nearest millimetre (mm). The idea being that if the diameter of the shaft is out by more than 1mm, then the piston-shaft combination won't work properly or maybe can't even be assembled.

Let $\boldsymbol{D}$ denote the deviation of this measured parameter from its target value; and let $\boldsymbol{\theta} = \mathbb{E}[\boldsymbol{D}]$ denote the mean of 'D'. The process is producing good parts when $\boldsymbol{\theta} < \mathbf{1}$, and is producing defective parts when $\boldsymbol{\theta} \geq \mathbf{1}$.

Assumption 14: For simplicity, we will ignore the possibility that there's some continuous random variation in the performance of production process and that all the piston shafts made on a given day either have the correct target diameter or they are all off by 1mm. That is, the components are either perfect, $\boldsymbol{\theta} = \mathbf{0}$, or they are defective, $\boldsymbol{\theta} = \mathbf{1}$.

Nonetheless, two measurements of the diameter of an individual component will typically produce two different results due to measurement error.

If there were no measurement error then the firm could find the exact value, and so know for sure whether its components meet the required specification. However, the measurements of the critical parameter are affected by a mixture of random and systematic errors. Therefore, in practice, the measured value always come with an unknown measurement error, which can be positive or negative. The effect of random errors can be largely eliminated by taking the average of multiple measurements. However, the remaining systematic errors are rather more difficult to eliminate. We can represent this as follows:

$$'measured\ value' =' true\ value' +' measurement\ error'$$

Assumption 15: Furthermore, suppose that the measurement errors follow a normal distribution.[9] Let $\boldsymbol{\sigma}$ denote the (fixed) standard deviation of 'D', which remains constant and

[8] It is rather like the Met Office telling people the proportion of a defined geographical area, or region, that will experience rain on a given day.

[9] The analysis that follows can also be performed using a more generalized distribution (having certain standard properties like symmetry around the mean, differentiability of the PDF and CDF). However, the normal distribution assumption makes the exposition easier to follow, and it is an assumption that is often made in the literature. Moreover, the headline result about the value of high-quality (primary) calibrations does not depend on what happens in the tail of the distribution. That is because we think about a change in the quality of calibration as causing a relatively small perturbation in σ, which is unlikely to be impacted by what happens in the tail of the distribution.

does not depend on the actual value of the parameter being measured. That is, the variance of 'D' is given by: $\boldsymbol{\sigma^2} = \boldsymbol{Var}[\boldsymbol{D}]$.[10]

Notice that we are expressing $\boldsymbol{\sigma}$ in terms of the size of the deviation that occurs under the hypothesis that the components are defective (i.e., off by 1mm). That is, $\boldsymbol{\sigma}$ is the relative standard deviation of the measurement process. Lastly, we suppose for the time being that a large part of the uncertainty in the measurements comes from random errors that can be eliminated by taking the average of multiple measurements.[11]

5.2.4 The null hypothesis and the alternative hypothesis

The null hypothesis $(\boldsymbol{H_0}\colon \boldsymbol{\theta} = \boldsymbol{0})$ is that the process is producing good parts, and the distribution of the deviation is given by $\boldsymbol{D} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\sigma^2})$**.** Let $\alpha \in [0,1]$ denote the probability of a type-1 error, in which the firm mistakenly rejects the null hypothesis and shuts down production, scrapping that day's output, prematurely. The probability of incorrectly rejecting the null hypothesis, $\mathbf{Pr}\{\mathbf{reject}\ \boldsymbol{H_0} | \boldsymbol{H_0}\ \mathbf{is\ true}\} = \boldsymbol{\alpha}$, is referred to as the confidence level of the test. The conventional confidence level is $\alpha = 0.05$ but $\alpha \in [0,1]$, where [0,1] represents the unit interval on the real line between '0' and '1'.

The alternative hypothesis $(\boldsymbol{H_1}\colon \boldsymbol{\theta} = \boldsymbol{1})$ is that its producing defective parts, and the distribution of the deviation is given by $\boldsymbol{D} \sim \boldsymbol{N}(\boldsymbol{1}, \boldsymbol{\sigma^2})$. Let $\beta = \Pr\{\text{accept } H_0 | H_1 \text{ is true}\} \in (0,1)$ denote the probability of a type-2 error, in which the firm mistakenly keeps the production process running even though it's producing defective parts. The probability of correctly rejecting the null hypothesis, $\mathbf{Pr}\{\mathbf{reject}\ \boldsymbol{H_0} | \boldsymbol{H_1}\ \mathbf{is\ true}\} = \boldsymbol{1} - \boldsymbol{\beta}$, usually referred to as the statistical power of the test. The conventional power of a test is 80%, implying that $\beta = 0.20$.

The following statistical analysis requires that we standardise the measured deviation from the target value by normalising it using the standard deviation. That is, the analysis that follows will focus on $\boldsymbol{T} = \boldsymbol{D}/\boldsymbol{\sigma}$, and under H_0, $T = D/\sigma \sim N(0,1)$. However, we can always rewrite *T* as follows:

$$T = D/\sigma = (D - 1 + 1)/\sigma = [(D - 1)/\sigma] + [1/\sigma]$$

... Equation 4

Under the alternative hypothesis, we must have: $(D - 1)/\sigma \sim N(0,1)$. Since $1/\sigma$ is just a constant parameter, it follows that $E[T] = 1/\sigma$ and $Var[T] = 1$. Therefore, under the alternative hypothesis, we must have: $T = D/\sigma \sim N((1/\sigma), 1)$.

The situation is illustrated by the figures below, where in both cases the horizontal axis is $T = D/\sigma$. The figure on the LHS shows the probability density of T under the null hypothesis (blue curve) and under the alternative hypothesis (red curve). The figure on the RHS shows the

---

[10] We assume that the firms have enough metrology knowledge to determine $\boldsymbol{\sigma}$. That is, $\boldsymbol{\sigma}$ is a known parameter of the model that the firms treat as given. This does not imply that the firms know the value of the measurement error that they draw out of the distribution: Although, the random error is fixed on a per firm basis, there is nonetheless a distribution from which the firm's error is drawn and the value of the error is not visible to the firm. It helps to think of measurement uncertainty as being made up of a series of components, not all of which are knowable. But the firms can gauge and characterise - through constructing 'uncertainty budgets' or participating in proficiency testing schemes - the relative uncertainty of their measurement processes. For example, different firms can participate in a proficiency testing scheme in which they all receive bits of the same sample to see how their measurements differ. By pooling the results of all the labs, they can get an estimate of $\boldsymbol{\sigma}$ based on the variation in the estimates. Similarly, we assume that firms know $\boldsymbol{p_A}$ from long experience (and hence also know $\boldsymbol{p_B} = \boldsymbol{1} - \boldsymbol{p_A}$). Similarly, we assume they know $\boldsymbol{c_A}$, and $\boldsymbol{c_B}$, so that they treat $\boldsymbol{\tau}$ as a known parameter of the model.

[11] Systematic errors can't be removed by making repeated measurements in the hope they average out. Rather, it's analogous to accidentally shaving the end of a ruler, so that it's missing part of its length, meaning that it always gives estimates of distance that are a little too high. Systematic errors, however, can be removed by calibrations and reference materials. The role of calibrations will be explored in Section 6.

cumulative probability (likelihood of T being below a given value) under these two hypotheses. The dashed green line was placed at $T = 1.64$ to represent the threshold (critical value) beyond which the null hypothesis would be rejected at the 5% confidence level. (If in fact the alternative hypothesis were true, the null hypothesis would be rejected 80% of the time.) Both these figures are based on the same numerical example in which the relative standard deviation is $\sigma = 0.40$. (The numerical values for the mean and standard deviation relate to the deviation, D.) However, these pictures would look much the same for other values of the relative standard deviation.

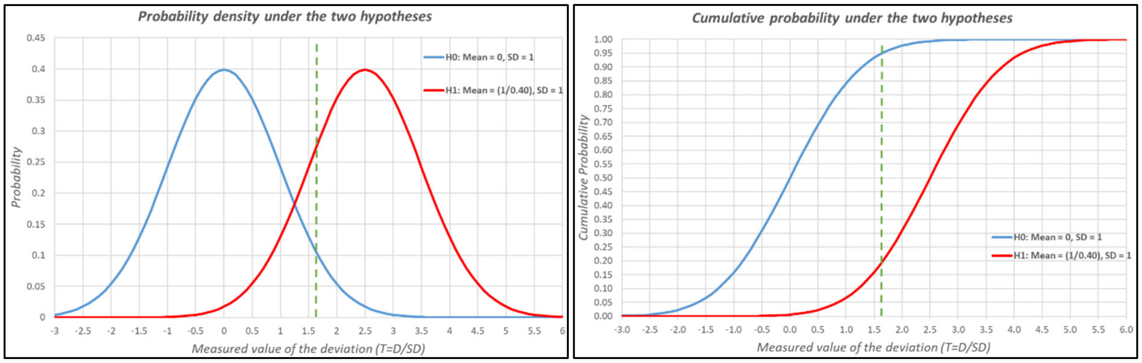

Figure 5: Probability Density Function (PDF) and Cumulative Distribution Function (CDF) for $T = D/\sigma$ under H_0 and H_1

5.2.5 Notation for CDF and PDF of Z

Let Z be a random variable drawn from the standard normal distribution: $Z \sim N(0,1)$. Then, the PDF and CDF of Z are denoted as:

$$\text{PDF: } \phi(z) = \frac{1}{\sqrt{2\pi}}. \exp\left(-\frac{1}{2}z^2\right)$$

$$\text{CDF: } \Phi(z) = \Pr\{Z \leq z\} = \int_{-\infty}^{z} \phi(x)dx$$

... Equation 5

Because $\Phi(z)$ is the integral of $\phi(z)$ and differentiation is the inverse of integration, it follows that $\phi(z)$ is the derivative of $\Phi(z)$: $d\Phi(z)/dz = \phi(z)$.[12] Moreover, Since $\Phi(z)$ is a monotonic (strictly increasing) function of z, it has an inverse, $\Phi^{-1}(.)$, which maps probabilities back on to values of z.

5.2.6 The relation between type-1 and type-2 errors

Under the null hypothesis, $T = D/\sigma \sim N(0,1)$. And, under the alternative hypothesis, $(D-1)/\sigma \sim N(0,1)$. Suppose the firm uses the conventional 5% confidence level, $\alpha = 0.05$. For a one-sided test (e.g. $D > 0$), the critical value of $T = D/\sigma$ must be 1.64, given that: $\Phi(1.64) = 0.95 = 1 - \alpha$.[13] Hence, the firm will reject the null hypothesis when $T > 1.64$, implying that the power of the test is given by:

---

[12] If a smooth curve is represented by a function, $f(x)$, then the derivative of this function, $df(x)/dx$, corresponds to its gradient and the integral of this function, $\int_a^b f(x)dx$, corresponds to the area under the curve between the vertical lines '$x = a$', '$x = b$' and the horizontal axis.

[13] The analysis that follows doesn't depend on a specific choice of confidence level, but the exposition is easier if we make it concrete.

$$1 - \beta = \Pr\{T > 1.64 | H_1 \text{ is true}\} = \Pr\{D/\sigma > 1.64 | \theta = 1\}$$

... Equation 6

Subtracting $1/\sigma$ from both sides of the inequality gives us the following expression:

$$1 - \beta = \Pr\{(D/\sigma) - (1/\sigma) > 1.64 - (1/\sigma) | \theta = 1\}$$

By slightly rearranging the LHS of this in equality the expression becomes:

$$1 - \beta = \Pr\{(D - 1)/\sigma > 1.64 - (1/\sigma) | \theta = 1\}$$

The probabilities of $(D - 1)/\sigma > 1.64 - (1/\sigma)$ and $(D - 1)/\sigma \leq 1.64 - (1/\sigma)$ must sum to unity, which implies that the previous expression can be rewritten as:

$$1 - \beta = 1 - \Pr\{(D - 1)/\sigma \leq 1.64 - (1/\sigma) | \theta = 1\}$$

Note that this probability is conditional on the alternative hypothesis being true, and under the alternative hypothesis, we know that $(D - 1)/\sigma \sim N(0,1)$. It follows that we can re express the probability in terms CDF for the standard normal distribution:

$$1 - \beta = 1 - \Phi\big(1.64 - (1/\sigma)\big)$$

This clearly simplifies to the following expression for the likelihood of a type-2 error:

$$\beta = \Phi\big(1.64 - (1/\sigma)\big)$$

... Equation 7

Since $\Phi(1.64) = 0.95 = 1 - \alpha$, and $\Phi(.)$ has an inverse we can re express this as: $\Phi^{-1}(1 - \alpha) = 1.64$. Therefore, our expression for the type-2 error can be rewritten in terms of likelihood of avoiding a type-1 error:

$$\beta = \Phi\big(\Phi^{-1}(1 - \alpha) - (1/\sigma)\big)$$

Using the inverse of the CDF, the relationship we found earlier between β and α can be rewritten as:

$$\Phi^{-1}(\beta) = \Phi^{-1}(1 - \alpha) - (1/\sigma)$$

... Equation 8

A little further rearrangement gives:

$$1/\sigma = \Phi^{-1}(1 - \alpha) - \Phi^{-1}(\beta)$$

Notice that the LHS is the deviation under the alternative hypothesis divided by the relative standard deviation (standard error) of the measurement. Taking the reciprocal of both sides gives us the relative standard deviation of the test:

$$\sigma = \frac{1}{\Phi^{-1}(1 - \alpha) - \Phi^{-1}(\beta)}$$

... Equation 9

Conventionally, an acceptable testing procedure correctly rejects the null hypothesis in 80% of instances where the alternative hypothesis is true: $1 - \beta = 0.80$. And it's also conventional to use a 5% confidence level for rejecting the null hypothesis: $\alpha = 0.05$. For this choice of α and β, the relative standard deviation becomes:

$$\sigma = \frac{1}{\Phi^{-1}(0.95) - \Phi^{-1}(0.20)} = 0.40218$$

5.2.7 The convexity of the constraint curve

It simplifies the exposition to introduce the following function:

$$g(x;\sigma) = \Phi\big(\Phi^{-1}(1-x) - (1/\sigma)\big)$$

... Equation 10

In terms of this notation, we can express the relationship between β and α as:

$$\beta = g(\alpha;\sigma)$$

... Equation 11

By choosing a particular value of σ, it is possible to plug certain values of α into this formula and find the corresponding values of β. By this means, we can plot β as a function of α. Figure 6 shows the results for σ = 0.25, σ = 0.50, and σ = 1.00. In this figure the vertical (y-axis) is the likelihood of a type-2 error, β, and the horizontal (x-axis) is the likelihood of a type-1 error, α.

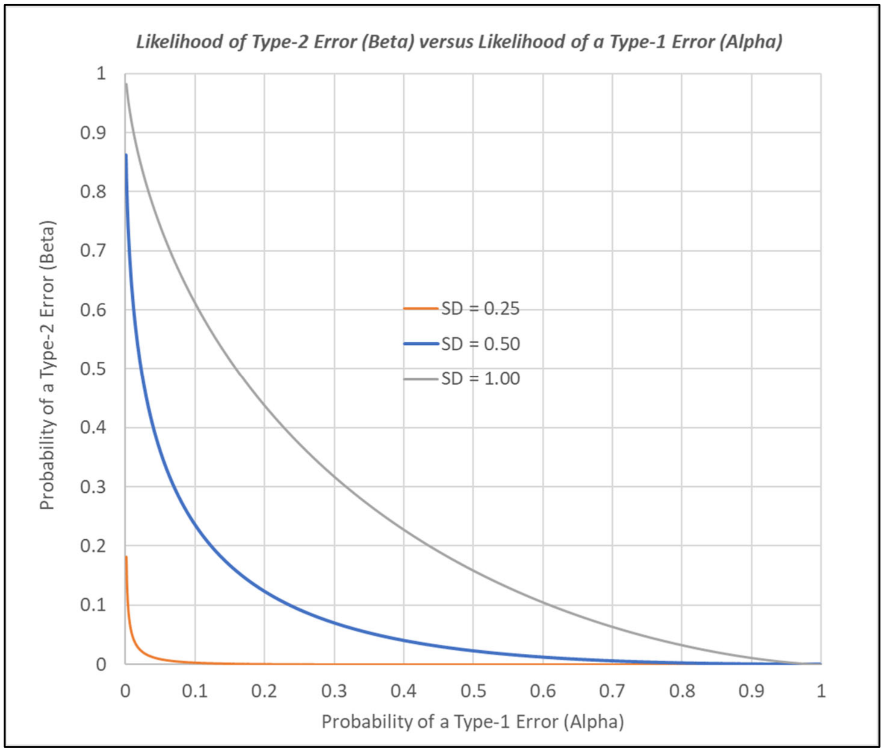

Figure 6: Relationship between Type-1 Error and Type-2 Error for different values of σ

This figure shows that there is a strong negative relationship between α and β.[14] That is, the likelihood of type-2 decreases as the likelihood of type-1 errors increases. This implies that the firm will need to trade-off one type of error against the other.[15] The figure also shows that the line moves diagonally away from the origin (0,0) as σ increases. Lastly, the curve always

---

[14] This relationship can also be seen by taking the partial derivative of the above equation with respect of α (keeping σ fixed):

$$\frac{\partial\beta}{\partial\alpha} = \frac{\partial g(\alpha;\sigma)}{\partial\alpha} = \frac{\partial\Phi\big(\Phi^{-1}(1-\alpha)-(1/\sigma)\big)}{\partial\alpha} = \underbrace{\phi\big(\Phi^{-1}(1-\alpha)-(1/\sigma)\big)}_{>0}\cdot\underbrace{\frac{d\Phi^{-1}(1-\alpha)}{d(1-\alpha)}}_{>0}\cdot\underbrace{\frac{d(1-\alpha)}{d\alpha}}_{<0} < 0$$

The first term denotes the PDF of a standard normal distribution, which is positive (by definition). The second term denotes the inverse of CDF of a standard normal distribution, which is also always positive. The third term is negative and equal to -1. Therefore, the product of the three terms is negative. That is, for a fixed value of σ, there is a negative relationship between α and β.

[15] The firm can avoid Type-I errors by never discarding a batch, but that necessarily entails a Type-II error whenever the batch is defective. Similarly, it can avoid Type-II errors by discarding almost every batch it produces but that leads to a lot of Type-I errors.

has a convex shape, but becomes less convex as σ increases. That is, although, it's first derivative is negative, $\boldsymbol{d\beta/d\alpha = g'(\alpha;\sigma) < 0}$, the slope decreases as the likelihood of a type-1 error increases, which implies that the second derivative is positive, $\boldsymbol{d^2\beta/(d\alpha)^2 = g''(\alpha;\sigma) > 0}$. The convexity of this curve has important implications for how the firm will behave when trying to minimise the cost of measurement errors.

5.2.8 The slope of the constraint curve

Before tackling the tangency condition, it's helpful to recall the following mathematical identities. Firstly, if $\gamma = \Phi(z)$ and $z = \Phi^{-1}(\gamma)$, then calculus gives us:

$$d\Phi^{-1}(\gamma)/d\gamma = [d\Phi(z)/dz]^{-1}$$

Secondly, let $\phi(z)$ denote the PDF of standard normal distribution. Since $\Phi(z)$ is the CDF, this implies that: $\Phi(z) = \int_{-\infty}^{z} \phi(x)dx$. And, because $\Phi(z)$ is the integral of $\phi(z)$ and differentiation is the inverse of integration, it follows that $\phi(z)$ is the derivative of $\Phi(z)$:

$$d\Phi(z)/dz = \phi(z) = \phi\big(\Phi^{-1}(\gamma)\big).$$

Putting these two identities together implies that:

$$\frac{d\Phi^{-1}(\gamma)}{d\beta} = \left[\phi\big(\Phi^{-1}(\gamma)\big)\right]^{-1} = \frac{1}{\phi\big(\Phi^{-1}(\gamma)\big)}$$

... Equation 12

Characterising the firm's attempt to minimise the cost of measurement errors comes down to finding an expression for the tangency condition. Hence, we need to find an expression for the gradient of the curve, for a known σ, in Figure 6. Using the inverse of the CDF, we found the relationship between β and α in ... Equation 8 as:

$$\Phi^{-1}(\beta) = \Phi^{-1}(1-\alpha) - (1/\sigma)$$

A little further rearrangement gives us:

$$1/\sigma = \Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta)$$

Using the identity for the derivative of $\Phi^{-1}(\beta)$ from ... Equation 12, the total derivative of $1/\sigma$ can be written as:

$$-\frac{d\sigma}{\sigma^2} = -\frac{d\alpha}{\phi\big(\Phi^{-1}(1-\alpha)\big)} - \frac{d\beta}{\phi\big(\Phi^{-1}(\beta)\big)}$$

Which can be written more compactly as:

$$\frac{d\alpha}{\phi\big(\Phi^{-1}(1-\alpha)\big)} + \frac{d\beta}{\phi\big(\Phi^{-1}(\beta)\big)} = \frac{d\sigma}{\sigma^2}$$

However, σ is to be regarded as a constant parameter of the model, and so it must be the case that $d\sigma = 0$, which implies the that RHS of our expression is zero. Hence, our expression for the total derivative becomes:

$$\frac{d\alpha}{\phi\big(\Phi^{-1}(1-\alpha)\big)} + \frac{d\beta}{\phi\big(\Phi^{-1}(\beta)\big)} = 0$$

A little further rearrangement yields an expression for the gradient (slope) of the dark blue curve in Figure 6 above:

$$\frac{d\beta}{d\alpha} = -\frac{\phi\big(\Phi^{-1}(\beta)\big)}{\phi\big(\Phi^{-1}(1-\alpha)\big)}$$

... Equation 13

Given that this derivative corresponds to the gradient of the of the constraint curve, $g(\alpha;\sigma)$, and $\beta = g(\alpha;\sigma)$, we can re express this as:

$$\frac{d\beta}{d\alpha} = g'(\alpha;\sigma) = -\frac{\phi\left(\Phi^{-1}(g(\alpha;\sigma))\right)}{\phi\left(\Phi^{-1}(1-\alpha)\right)}$$

... Equation 14

Notice that, holding σ constant, the RHS of ... Equation 14 becomes just a function of α. Moreover, because the constraint curve is convex, $g''(\alpha;\sigma) > 0$. That is, its first derivative, $g'(\alpha;\sigma)$, is a monotonically increasing (ever less negative) function of α, which implies the existence of an inverse function.[16] The existence of this inverse function implies a one-to-one mapping between the slope of the constraint curve, $g'(\alpha;\sigma)$, and the likelihood of type-1 errors, α. This means that a given value of the slope, $g'(\alpha;\sigma)$, corresponds to one specific value of α and vice versa. Furthermore, having found a value of α that correspond to a given value of the slope, we can find the corresponding value of β using $\beta = g(\alpha;\sigma)$.

Conventionally, an acceptable testing procedure correctly rejects the null hypothesis in 80% of instances where the alternative hypothesis is true: $1-\beta = 0.80$. And it's also conventional to use a 5% confidence level for rejecting the null hypothesis: $\alpha = 0.05$. When these values of α and β are plugged into the expression above, the calculated value of the slope becomes:

$$\frac{d\beta}{d\alpha} = -\frac{\phi\left(\Phi^{-1}(0.20)\right)}{\phi\left(\Phi^{-1}(0.95)\right)} = -2.7145$$

5.2.9 The cost of mistakes made due to measurement errors:

The cost of shutting down production and setting it up again from scratch is non-trivial. Let c_A denote the value of the lost output from the deadtime during which the firm's facilities have ceased production. In many scenarios, it's reasonable to assume that the cost of a type-1 error comes from temporarily turning off the flow of profits for a day whilst the production process is shut down and set up again: $c_A = v_A$.[17]

However, the cost of accidentally continuing to operate the production process if its malfunctioned is considerable: There is the cost of scrapping defective parts, compensating adversely effected customers, and potentially some reputational damage to the brand. Let c_B denote the cost of accidentally producing (and supplying) defective parts, and it's assumed that $c_B > c_A = v_A$. For now, it's helpful to keep c_A and c_B as distinct parameters, but there are points in the analysis where we will let $c_A = v_A = 1$, so that c_B is measured in terms of the cost of making a type-1 error.[18]

The average cost incurred due to each type of error is found by multiplying the corresponding cost of the error by the unconditional probability of making such an error. The *unconditional* likelihood of making a type-1 error is given by:

$$\Pr\{\text{reject } H_0 \cap H_0 \text{ is true}\} = \Pr\{\text{reject } H_0 | H_0 \text{ is true}\} \times \Pr\{H_0 \text{ is true}\} = \alpha \times p_A$$

... Equation 15

And the *unconditional* likelihood of making a type-2 error is given by:

$$\Pr\{\text{accept } H_0 \cap H_1 \text{ is true}\} = \Pr\{\text{accept } H_0 | H_1 \text{ is true}\} \times \Pr\{H_1 \text{ is true}\} = \beta \times p_B$$

... Equation 16

It follows that the expected cost of the mistakes made due to measurement errors is given

[16] It's not possible to write down this inverse explicitly as a function but it nonetheless exists.
[17] It's analogous to our commuter hearing a rumour that there will be ticket inspectors on the train, and so buying a ticket. If the rumour is false, then the commuter has 'wasted' they money on the fare.
[18] There is no loss of generality as we are free to set one of these cost to unity and make it the unit of account.

by:

$$\mathcal{E} = (p_A c_A).\alpha + (p_B c_B).\beta$$

... Equation 17

If there were no measurement errors, and thus no mistakes, then the expected value of the output would be $p_A v_A = p_A c_A$.[19] With measurement errors, coupled to their attendant mistakes, the firm's profitability (payoff) becomes:

$$p_A v_A - \mathcal{E} = (1-\alpha)p_A v_A - (\beta)p_B c_B$$

... Equation 18

For our purposes, perhaps the more economically relevant quantity is the additional value created by acting on the measurement information. In terms of the choices of α and β, the baseline case is where the firm foregoes the conformance testing and just accepts the penalty when the production process malfunctions. This is a kind of corner solution in which $\alpha = 0$ and $\beta = 1$. This would mean that if the production process has malfunctioned, then the firm is certain to make a type-2 error.

If the firm were to forego the information generated by conformance testing, and just accept the penalty whenever the process malfunctioned, then it's payoff would become: $p_A v_A - p_B c_B$. We have just argued that, by using the information generated by conformance testing, the firm's payoff becomes $p_A v_A - \mathcal{E}$, which implies that the value attributed to the measurement information, $\boldsymbol{\mathcal{W}}$, is the difference between these payoffs:

$$\mathcal{W} = (p_A v_A - \mathcal{E}) - (p_A v_A - p_B c_B) = p_B c_B - \mathcal{E}$$

... Equation 19

Substituting for $\mathcal{E}$, along with a little rearrangement, gives is the following expression for the value of the measurement information generated by conformance testing:

$$\mathcal{W} = p_B c_B - \mathcal{E} = (p_B c_B).(1-\beta) - (p_A c_A).\alpha$$

... Equation 20

However, we are free to use $p_A v_A = p_A c_A$ as our unit of monetary value, and so to express costs or benefit in terms of the firm's maximum potential payoff. Let $\boldsymbol{\tau}$ denote the relative cost of making type-2 errors, as a proportion of the maximum attainable output, so that $(p_B c_B) = \tau \times (p_A c_A)$. Hence, we will introduce normalised (lower case) versions of $\mathcal{E}$ and $\mathcal{W}$.

Let $\boldsymbol{e}$ denote the combined cost from both kinds of mistake, as a proportion of the maximum attainable payoff, so that $\mathcal{E} = e \times (p_A c_A)$. Dividing both sides of ... Equation 17 by the maximum attainable payoff, $p_A c_A$, gives us an expression for the combined cost from both kinds of mistake, as a proportion of the maximum attainable payoff:

$$e = \alpha + \left(\frac{p_B c_B}{p_A c_A}\right).\beta.$$

... Equation 21

This can be expressed more nicely as: $\boldsymbol{e = \alpha + \tau.\beta}$, where $\tau = (p_B c_B)/(p_A c_A)$ is referred to as the 'cost ratio' and is a basic parameter of the model.[20]

Let $\boldsymbol{w}$ denote the value of the measurement information generated by conformance testing,

---

[19] Where this identity comes from assuming that the loss from a type-1 error is dominated by the value of the output that is unnecessarily scrapped.

[20] The firm has some power over the values of α and β, but τ is a fixed parameter that isn't under the firm's control.

as a proportion of the maximum attainable payoff, so that $\mathcal{W} = w \times (p_A c_A)$. Dividing both sides of ... Equation 20 by the maximum attainable payoff, $p_A c_A$, gives us an expression for the value of measurement information, as a proportion of the maximum attainable payoff:

$$w = \left(\frac{p_B c_B}{p_A c_A}\right).(1-\beta) - \alpha.$$

... Equation 22

This can be expressed more nicely as: $\boldsymbol{w = \tau.(1-\beta) - \alpha}$, where $\tau = (p_B c_B)/(p_A c_A)$ is the 'cost ratio'. Notice that a little rearranged gives us: $w = \tau - (\alpha + \tau.\beta)$. And, because we already have $e = \alpha + \tau.\beta$, it follows that: $\boldsymbol{w = \tau - e}$.

Corollary 1: In other words, the value of measurement information is inversely related to the cost of mistakes from acting on faulty information. Lastly, since $\boldsymbol{\tau}$ is a fixed parameter of the model, it follows that maximising the value of the measurement information, $\boldsymbol{w}$, is equivalent to minimising the cost of mistakes from acting on faulty measurement information, $\boldsymbol{e}$.

5.2.10 Setting up the firm's cost minimisation problem

The firm will seek to minimise the cost of mistakes, $\mathcal{E}$, subject to the constraint imposed by the relationship we found between the likelihood of type-1 and type-2 errors in ... Equation 8. For the case where σ = 0.50, this constraint corresponds to the dark blue curve in Figure 6. More formally we can say that the firm's minimisation problem can be written as: for known values of the model's basic parameters choose $\alpha, \beta \in (0,1)$ to minimise the total cost of mistakes $\mathcal{E} = (p_A c_A).\alpha + (p_B c_B).\beta$ subject to the constraint connecting the likelihood of type-1 and type-2 errors: $\beta = g(\alpha;\sigma) = \Phi\big(\Phi^{-1}(1-\alpha) - (1/\sigma)\big)$.[21]

Corollary 2: The constraint makes $\boldsymbol{\beta}$ a function of $\boldsymbol{\alpha}$ and the objective function, that we are trying to minimise, is an increasing function of both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Together this implies that the constraint is biding, meaning that the firm is forced to trade-off a decrease in the likelihood of one kind of error against an increase in the likelihood of other kind of error.

When the constraint is substituted into the objective function, the firm's problem is equivalent to choosing the likelihood of type-1 errors, $\alpha \in (0,1)$, to minimise $\mathcal{E} = H(\alpha;\sigma,\boldsymbol{p},\boldsymbol{c})$, where $\sigma,\boldsymbol{p},\boldsymbol{c}$ are basic parameters of the model and the reduced-form of the cost function, $H(x;\sigma,\boldsymbol{p},\boldsymbol{c})$, is defined as follows:

$$H(x;\sigma,\boldsymbol{p},\boldsymbol{c}) = (p_A c_A).x + (p_B c_B).g(x;\sigma) = (p_A c_A).x + (p_B c_B).\Phi\big(\Phi^{-1}(1-x) - (1/\sigma)\big).$$

... Equation 23

In terms of our commuter, $p_A c_A$ would be their expected benefit if they only ever bought a ticket when inspectors were on the train and so never paid more than they had to but also never received a fine. If we assume that $c_A = v_A$, then by dividing through by $p_A c_A$, we can express this cost as a proportion of the maximum possible profit that would occur if the firm could eliminate the cost of measurement error. In terms of this normalised version of the cost of mistakes, the firm's problem becomes: choose the likelihood of type-1 errors, $\alpha \in (0,1)$, to minimise $e = h(\alpha;\sigma,\tau)$, where σ,τ are basic parameters of the model and the reduced-form

[21] Although the analysis in this paper sets α and β as the choice variables in the firm's cost minimization problem, we believe that firms do not necessarily solve for the optimal α and β in an explicit way. That is, they may arrive at these optimal levels through other ways, such as a mixture of repeated tatonnement (trial and error) and Darwinian selection. The first mechanism (tatonnement) is very similar to a consumer experimenting with different bundles of goods whilst seeking to maximise their utility subject to budget constraint. According to the second mechanism (Darwinian selection), firms that are far off the optimal levels will incur higher costs, thereby reducing their profitability and harming their chances of survival. Therefore, in a competitive market, the firms that will eventually survive will be the ones that use measurement processes with values of α and β close to the optimal levels. It follows that the mathematical optimisation gives us a convenient way to arrive at the outcome of what is actually a more complex evolutionary process.

of the (normalised) cost function, $h(\alpha;\sigma,\tau)$, is defined as follows:

$$h(x;\sigma,\tau) = x + \tau.\,g(x;\sigma) = x + \tau.\,\Phi\big(\Phi^{-1}(1-x) - (1/\sigma)\big).$$

... Equation 24

To solve for the firm's cost minimisation problem, it is necessary to characterise the shape of the objective function and understand how it responds to changes in the likelihood of type-1 errors, α.

5.2.11 Solving the cost minimisation problem and the tangency condition:

The figure below illustrates the (normalised) reduced-form cost function, $h(\alpha;\sigma,\tau)$, for five different combinations of the basic parameters. The five curves correspond to five different choices for the cost ratio, $\tau = (\wp_B c_B)/(\wp_A c_A)$, where $\wp_B c_B$ is the cost of type-2 errors that would be incurred if the firm were to forego the conformance testing. In each case, the relative standard deviation was set to $\sigma = 0.50$, so that we can focus on the effect of changes in the 'cost ratio'. The figure shows that, if the 'cost ratio' increases, then so does the relative harm from unmitigated type-2 errors (e.g. without conformance testing).

The figure also shows that the minimum attainable cost from measurement errors occurs at the point where the curve is momentarily parallel to the horizontal axis. That is, the minimum occurs at the point where the curve is momentarily flat: $d\mathcal{E}/d\alpha = 0$.

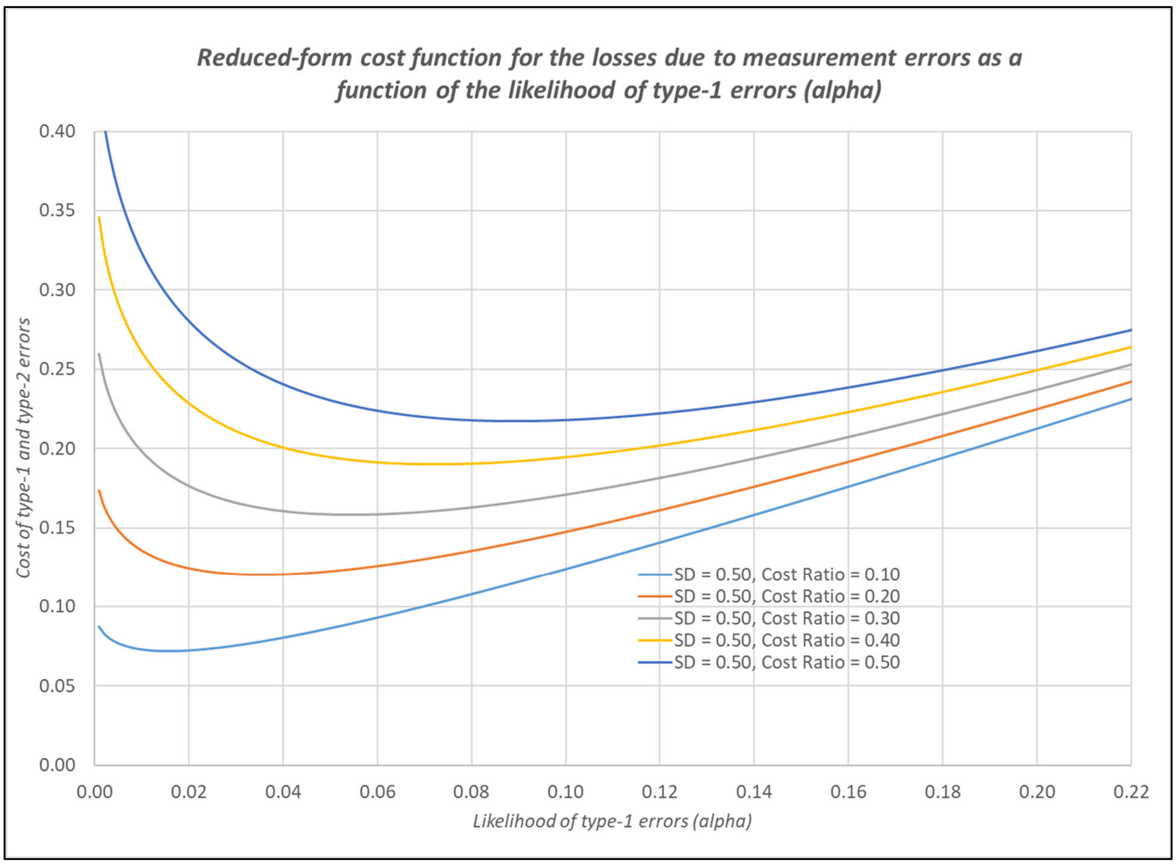

Figure 7: Relationship between the cost of Type-1 Error and the cost of Type-2 Error for different values of τ

Differentiating the objective function with respect to α gives us:

$$\frac{d\mathcal{E}}{d\alpha} = (\wp_A c_A) + (\wp_B c_B).\frac{d\beta}{d\alpha} = (\wp_A c_A) + (\wp_B c_B).\, g'(\alpha;\sigma)$$

Because the constraint makes β is decreasing convex function of α, optimisation theory (calculus) tells us that the solution to this minimisation problem is characterised by the first-order condition: $d\mathcal{E}/d\alpha = 0$. From this we get the following necessary and sufficient condition for the solution to firm's cost minimisation problem:[22]

$$(\wp_A c_A) + (\wp_B c_B).\frac{d\beta}{d\alpha} = 0$$

A little further rearrangement gives us the following tangency condition:

$$\frac{d\beta}{d\alpha} = -\frac{\wp_A c_A}{\wp_B c_B} = -\frac{1}{\tau}$$

... Equation 25

Given that this derivative corresponds to the gradient of the of the constraint curve, $g(\alpha;\sigma)$, we can re express the tangency condition as:

$$\frac{d\beta}{d\alpha} = g'(\alpha;\sigma) = -\frac{\wp_A c_A}{\wp_B c_B}$$

It remains to equate the slope of the constraint curve obtained in ... Equation 25 and the iso-cost line obtained in ... Equation 13:

$$\frac{\wp_A c_A}{\wp_B c_B} = \frac{\phi\left(\Phi^{-1}\big(g(\alpha;\sigma)\big)\right)}{\phi\big(\Phi^{-1}(1-\alpha)\big)} = \frac{\phi\big(\Phi^{-1}(\beta)\big)}{\phi\big(\Phi^{-1}(1-\alpha)\big)}$$

Recalling that $\phi(.)$ denotes the PDF of a standard normal distribution, we can use this tangency condition to find expressions for α and β in terms of the basic parameters of the model. Using the definition of $\phi(.)$, the tangency condition can be re written as follows:

$$\frac{\wp_A c_A}{\wp_B c_B} = \frac{\exp\left\{-\frac{1}{2}[\Phi^{-1}(\beta)]^2\right\}}{\exp\left\{-\frac{1}{2}[\Phi^{-1}(1-\alpha)]^2\right\}} = \exp\left\{\frac{1}{2}[\Phi^{-1}(1-\alpha)]^2 - \frac{1}{2}[\Phi^{-1}(\beta)]^2\right\}$$

... Equation 26

The LHS of this expression for the tangency condition corresponds to the cost ratio, which is a basic parameter of the model. Hence, this equation helps us to connect the optimal choice of α and β to one of the basic parameters of the model.[23]

5.2.12 The cost ratio

For future analysis, it's helpful to define the '*cost ratio*' as $\tau = (\wp_B c_B)/(\wp_A c_A)$, and it's reasonable to assume that: $\tau = (\wp_B c_B)/(\wp_A c_A) < 1$. In terms of this cost ratio, we can rewrite ... Equation 26 as:

$$\tau = \frac{\exp\left\{-\frac{1}{2}[\Phi^{-1}(1-\alpha)]^2\right\}}{\exp\left\{-\frac{1}{2}[\Phi^{-1}(\beta)]^2\right\}} = \exp\left\{\frac{1}{2}[\Phi^{-1}(\beta)]^2 - \frac{1}{2}[\Phi^{-1}(1-\alpha)]^2\right\}$$

... Equation 27

Conventionally, an acceptable testing procedure correctly rejects the null hypothesis in 80% of instances where the alternative hypothesis is true: $1-\beta = 0.80$. And, it's also conventional to use a 5% confidence level for rejecting the null hypothesis: $\alpha = 0.05$. For this choice of α

---

[22] We will ignore any possibility of corner solutions.
[23] Notice that this expression does not directly involve the relative standard deviation, σ.

and β, the cost ratio becomes:

$$\tau = \exp\left\{\frac{1}{2}[\Phi^{-1}(0.20)]^2 - \frac{1}{2}[\Phi^{-1}(0.95)]^2\right\} = 0.368392$$

5.2.13 Solutions as points of intersection for two sets of curves

Notice that treating the cost ratio as a fixed parameter makes β a function of α. By taking the natural log of both sides of the equation and multiplying through by 2, it can be shown that:

$$2.\ln(\tau) = [\Phi^{-1}(\beta)]^2 - [\Phi^{-1}(1-\alpha)]^2$$

For future reference, note the parabolic form of this relationship. After a little rearrangement, we get the following expression for the likelihood of a type-2 error:

$$\beta = \Phi\left(\mp\sqrt{2.\ln(\tau) + [\Phi^{-1}(1-\alpha)]^2}\right)$$

Note that because $0 < \tau < 1$, we must have $\ln(\tau) < 0$. Because of the potential for computational confusion, it's best to rewrite this equation as:

$$\beta = \Phi\left(\mp\sqrt{[\Phi^{-1}(1-\alpha)]^2 - 2.\ln(1/\tau)}\right)$$

... Equation 28

The symmetry of the CDF of a standard normal distribution implies that $\Phi(-x) = 1 - \Phi(x)$. Using this identity, we can rewrite ... Equation 28 as:

$$\beta = \begin{cases} \Phi\left(\sqrt{[\Phi^{-1}(1-\alpha)]^2 - 2.\ln(1/\tau)}\right); \text{for the positive square root} \\ 1 - \Phi\left(\sqrt{[\Phi^{-1}(1-\alpha)]^2 - 2.\ln(1/\tau)}\right); \text{for the negative square root} \end{cases}$$

... Equation 29

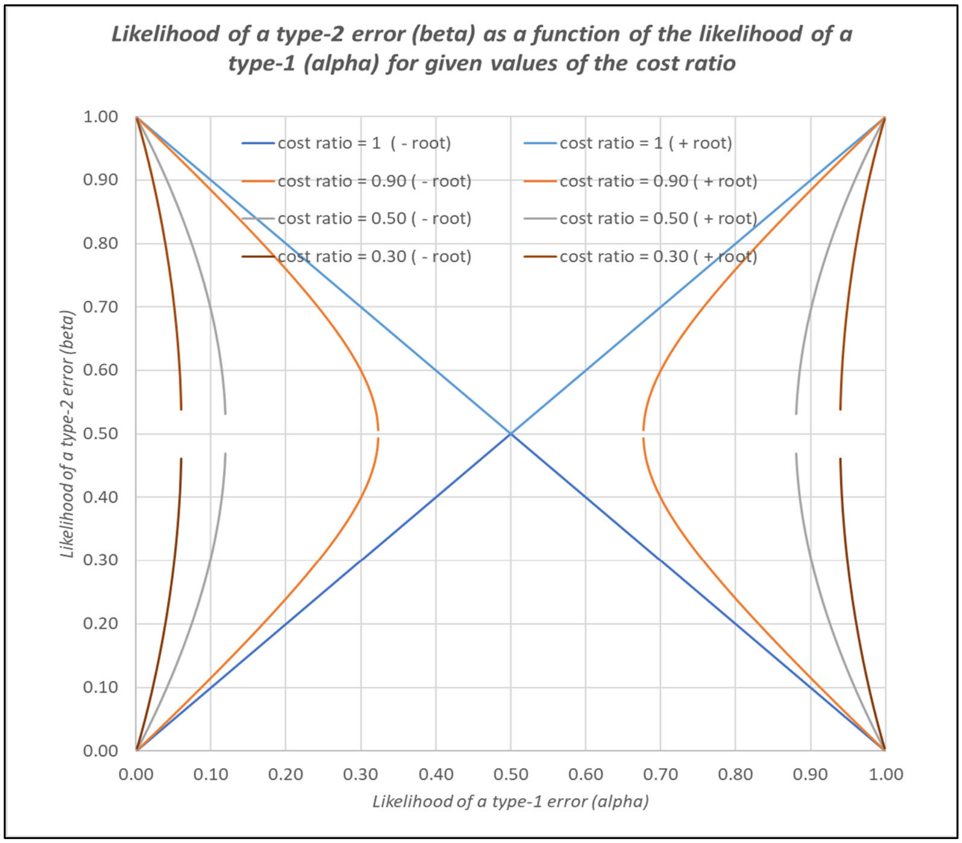

Figure 8: Relationship between the likelihood of Type-1 error and the likelihood of Type-2 Error for different values of τ

Figure 8 shows the parabolic relationship that's created between α and β when the cost ratio is treated as a fixed parameter of the model. The horizontal axis gives the likelihood of a type-1 error and the vertical axis gives the likelihood of a type-2 error. The cost ratio is given by $\tau = (\wp_B c_B)/(\wp_A c_A)$, where $\wp_B c_B$ is the cost of just accepting the penalty when the production process malfunctions and $\wp_A c_A$ is the maximum possible benefit that could be achieved if there were no measurement errors. Each parabolic curve relates to a different value of τ. The figure shows that as the cost ratio increases the curvature of the parabola increases until when the cost ratio equals unit the parabola transforms into the blue cross in the centre of the chart. Small values of cost ratio lead the parabola to almost flatten itself against the vertical axis.[24]

The parabolic nature of the relationship means that for a given value of α there are potentially two different values of β that provide a mathematically satisfactory solution. For practical purposes, it's the bottom half of the LHS segment of the figure that's most relevant to the analysis, because this is where the curve created by fixing the value of the cost ratio τ (Figure 8) will intersect the curve created by fixing the value of the relative standard deviation σ (Figure 6). That is, we have one set of curves defined by values of σ, and another separate set of curves defined by values of $\tau = (\wp_B c_B)/(\wp_A c_A)$. The points of intersection between

[24] The positive square root branch in ... Equation 29 leads to values of $\beta > 0.5$ (i.e., the upper half of the vertical axis in Figure 8. Meanwhile, the negative square root branch in ... Equation 29 leads to values of $\beta < 0.5$ (i.e., the lower half of the vertical axis in Figure 8. Therefore, it is evident that the relationship between α and β is discontinuous. That is, for a given τ, there exist values of α where β is not defined. Mathematically, it can be seen from ... Equation 29 that this would happen when the term inside the square root becomes negative. We don't yet have a complete understanding of the underlying intuition, but it is worth delving into. Lastly, the following analysis will show that for practical solutions, we need to focus on the negative square root branch in ... Equation 29, so that $\beta < 0.5$.

these families of curves show how values of the basic parameters map on to value of α and β. This is illustrated in for some specific values of the two parameters σ and τ.

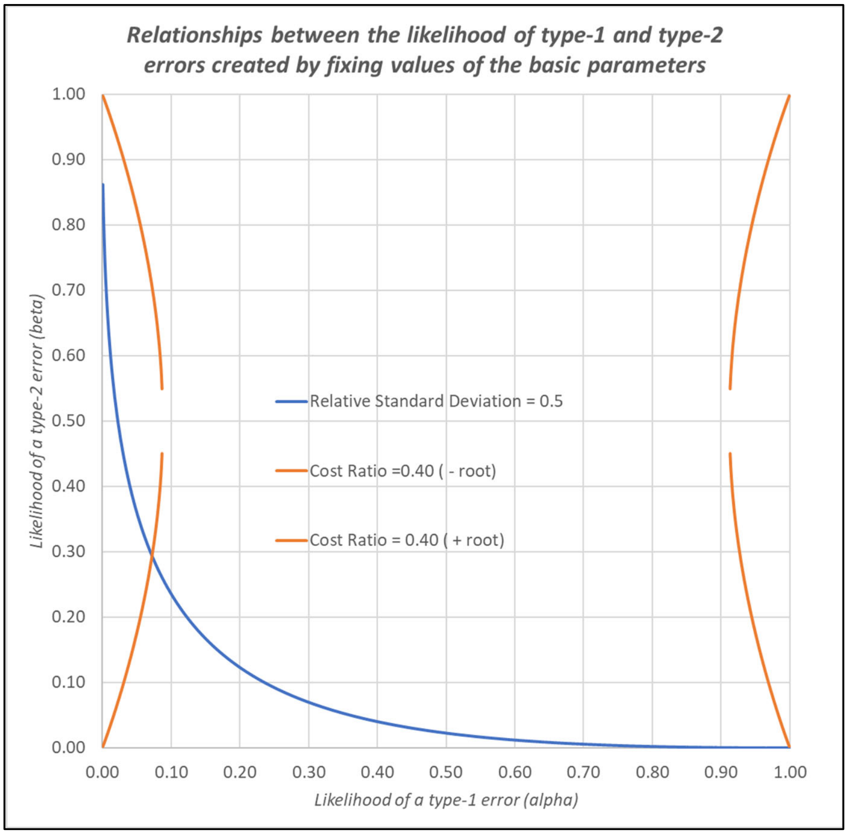

Figure 9: Relationship between the likelihood of type-1 error and the likelihood of type-2 error for fixed values of model parameters σ and τ

5.2.14 Expressions for the optimal choices of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$

Returning to our expression for the cost ratio from ... Equation 26, we found that:

$$\frac{\wp_A c_A}{\wp_B c_B} = \exp\left\{\tfrac{1}{2}[\Phi^{-1}(1-\alpha)]^2 - \tfrac{1}{2}[\Phi^{-1}(\beta)]^2\right\}$$

The term within the exponential function is the difference of two squares multiplied through by a factor of $1/2$, and so this can be re expressed as follows:

$$[\Phi^{-1}(1-\alpha)]^2 - [\Phi^{-1}(\beta)]^2 = [\Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta)].[\Phi^{-1}(1-\alpha) + \Phi^{-1}(\beta)]$$

Moreover, we saw earlier that $\Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta) = 1/\sigma$, which implies that the tangency condition becomes:

$$\frac{\wp_A c_A}{\wp_B c_B} = \exp\left(\frac{\Phi^{-1}(1-\alpha) + \Phi^{-1}(\beta)}{2\sigma}\right)$$

... Equation 30

Taking the natural log of both sides and multiplying through by 2σ, gives us the following expression:

$$\Phi^{-1}(1-\alpha) + \Phi^{-1}(\beta) = 2\sigma.\ln(\wp_A c_A/\wp_B c_B)$$

The cost ratio is given by $\tau = (\wp_B c_B)/(\wp_A c_A)$, where $\wp_B c_B$ is the cost of just accepting the penalty when the production process malfunctions and $\wp_A c_A$ is the maximum possible benefit that could be achieved if there were no measurement errors. Hence, the previous expression can be rewritten as follows:

$$\Phi^{-1}(1-\alpha) + \Phi^{-1}(\beta) = 2\sigma.\ln(1/\tau)$$

... Equation 31

Since $\Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta) = 1/\sigma$, we can substitute $\Phi^{-1}(1-\alpha) - (1/\sigma)$ for $\Phi^{-1}(\beta)$, so that the expression becomes:[25]

$$\Phi^{-1}(1-\alpha) = \sigma.\ln(1/\tau) + (1/2\sigma) = 2\sigma.\ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)$$

... Equation 32

And, because $\Phi^{-1}(.)$ is the inverse of $\Phi(.)$, this implies that:

$$\boldsymbol{\alpha = \alpha^*(\sigma,\tau) = 1 - \Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right)}$$

... Equation 33

We can now use the result for the likelihood of type-1 errors to find the likelihood of type-2 errors. Since $\Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta) = 1/\sigma$, we can substitute $\Phi^{-1}(\beta) + (1/\sigma)$ for $\Phi^{-1}(1-\alpha)$, to get the following the expression:

$$\Phi^{-1}(\beta) + (1/\sigma) = 2\sigma.\ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)$$

A little further rearrangement gives us the following result for the likelihood of type-2 errors:

$$\Phi^{-1}(\beta) = 2\sigma.\ln\left(1/\sqrt{\tau}\right) - (1/2\sigma),$$

which implies that:

$$\boldsymbol{\beta = \beta^*(\sigma,\tau) = \Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)}$$

... Equation 34

Substitution of these optimal values of α and β into the objective function, $h(.;\sigma,\tau)$, gives us a formula for the minimum possible cost, given the values of the basic parameters:

$$\tilde{h}(\sigma,\tau) = h(\alpha^*(\sigma,\tau);\sigma,\tau) = \alpha^*(\sigma,\tau) + \tau.g(\alpha^*(\sigma,\tau);\sigma)$$

It remains to substitute in for $\alpha^*(\sigma,\tau)$. After a little rearrangement and simplification, we arrive at the following expression:

$$\tilde{h}(\sigma,\tau) = 1 - \Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right) + \tau.\Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)$$

... Equation 35

5.2.15 An expression for the minimum attainable cost of mistakes due to faulty measurement information

Earlier we showed that the expected total cost of mistakes made due to the measurement errors is given by: $\mathcal{E} = (\wp_A c_A).\alpha + (\wp_B c_B).\beta$. Substituting in the formulae for optimal values of α and β into the equation for $\mathcal{E}$, gives us the following expression for the minimum attainable cost:

$$\mathcal{E} = \wp_A c_A - \wp_A c_A.\Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right) + \wp_B c_B.\Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)$$

... Equation 36

Assuming that $v_A = c_A$, the firm's payoff whilst it engages in conformance testing, and acts on the measurement information, becomes:

[25] Note that here we have used the identity $\ln(1/\tau) = 2 \times \frac{1}{2} \times \ln(1/\tau) = 2 \times \ln\left(1/\sqrt{\tau}\right)$ to rewrite the RHS in a slightly nicer way.

$$\mathcal{p}_A \mathcal{v}_A - \mathcal{E} = \mathcal{p}_A c_A . \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right) - \mathcal{p}_B c_B . \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)$$

... Equation 37

Lastly, the maximum attainable value of the measurement information generated by conformance testing becomes:

$$\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{p}_B c_B} - \boldsymbol{\mathcal{E}} = (\boldsymbol{\mathcal{p}_B c_B}).\left[\mathbf{1} - \boldsymbol{\Phi}\left(\boldsymbol{\psi}_{-}(\boldsymbol{\sigma},\boldsymbol{\tau})\right)\right] - (\boldsymbol{\mathcal{p}_A c_A}).\left[\mathbf{1} - \boldsymbol{\Phi}\left(\boldsymbol{\psi}_{+}(\boldsymbol{\sigma},\boldsymbol{\tau})\right)\right],$$

... Equation 38

where for notational convenience the expressions inside the CDF for a normal distribution is defined as: $\psi_{-}(\sigma,\tau) \equiv 2\sigma . \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)$ and $\psi_{+}(\sigma,\tau) \equiv 2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)$.

However, we are free to use $\mathcal{p}_A \mathcal{v}_A = \mathcal{p}_A c_A$ as our unit of monetary value, and so to express costs or benefit in terms of the firm's maximum potential payoff. Let τ denote the relative cost of making type-2 errors, as a proportion of the maximum attainable output, so that $(\mathcal{p}_B c_B) = \tau \times (\mathcal{p}_A c_A)$. Hence, we will introduce normalised (lower case) versions of $\mathcal{E}$ and $\mathcal{W}$.

Let e denote the combined cost from both kinds of mistake, as a proportion of the maximum attainable payoff, so that $\mathcal{E} = e \times (\mathcal{p}_A c_A)$. Dividing both sides of ... Equation 36 by the maximum attainable payoff, $\mathcal{p}_A c_A$, gives us an expression for the combined cost from both kinds of mistake, as a proportion of the maximum attainable payoff:

$$e = 1 - \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right) + \tau . \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)$$

... Equation 39

By dividing both sides of ... Equation 37 through by $\mathcal{p}_A c_A$, we can express the firm's payoff as a proportion of the maximum possible profit that would occur if the firm were to eliminate mistakes from measurement errors:

$$1 - e = \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right) - \tau . \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)$$

... Equation 40

Lastly, let $\mathcal{w}$ denote the value of the measurement information generated by conformance testing, as a proportion of the maximum attainable payoff, so that $\mathcal{W} = \mathcal{w} \times (\mathcal{p}_A c_A)$. Dividing both sides of ... Equation 38 by the maximum attainable payoff, $\mathcal{p}_A c_A$, gives us an expression for the value of measurement information, as a proportion of the maximum attainable payoff:

$$\mathcal{w} = \tau - e = \tau.\left[1 - \Phi\left(\psi_{-}(\sigma,\tau)\right)\right] - \left[1 - \Phi\left(\psi_{+}(\sigma,\tau)\right)\right],$$

... Equation 41

where for notational convenience the expressions inside the CDF for a normal distribution is defined as: $\psi_{-}(\sigma,\tau) \equiv 2\sigma . \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)$ and $\psi_{+}(\sigma,\tau) \equiv 2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)$.

5.2.16 Some further results and implications of the analysis

It's reasonable to assume that the cost of a type-1 error (false positive) can be approximated by the loss of income that comes from temporarily ceasing production whilst the process is set up again from scratch.

If we assume that $c_A = \mathcal{v}_A$, then:

$$\frac{\mathcal{p}_A \mathcal{v}_A - \mathcal{p}_B c_B}{\mathcal{p}_B c_B} = \frac{\mathcal{p}_A c_A}{\mathcal{p}_B c_B} - 1$$

The numerator of the LHS is what the firm's expected profit would be if it were to forego conformance testing and just accepted the penalty from the process malfunctioning occasionally. The denominator of the LHS is expected cost of the penalties it would incur if it were to forego conformance testing.

Taking the reciprocal of both sides gives us an expression for the proportion of the firm's profits that would be lost if the firm chose to forego conformance testing and just accepted the penalty from the process malfunctioning occasionally:

$$\frac{p_B c_B}{p_A v_A - p_B c_B} = \frac{1}{\left(\frac{p_A c_A}{p_B c_B}\right) - 1}$$

Using the tangency condition this becomes:

$$\frac{p_B c_B}{p_A v_A - p_B c_B} = \frac{1}{|g'(\alpha;\sigma)| - 1}$$

where

$$|g'(\alpha;\sigma)| = \frac{\phi\left(\Phi^{-1}\left(g(\alpha;\sigma)\right)\right)}{\phi\left(\Phi^{-1}(1-\alpha)\right)}$$

This implies that the relative importance of the harm done by potentially supplying defective products decreases as the steepness of the slope at the point of tangency increases. That is, if the firm were to suffer serious harm for accidentally supplying defective products, then it would select a point on the constraint curve with a relatively gentle slope, which corresponds to selecting a relatively high likelihood of type-1 errors.

If we assume that $c_A = v_A$, then the firm's payoff if it undertakes conformance testing becomes:

$$(p_A v_A) - (p_A c_A).\alpha - (p_B c_B).\beta = (p_A v_A).(1-\alpha) - (p_B c_B).\beta$$

And, if the firm had foregone conformance testing its payoff would have been: $(p_A v_A) - (p_B c_B)$. Hence, the net-benefit of undertaking conformance testing is given by: $(1-\beta).(p_B c_B) - \alpha.(p_A v_A)$. From which it follows that the proportional increase in profit attributable to conformance testing is given by:

$$\frac{(1-\beta).(p_B c_B) - \alpha.(p_A v_A)}{(p_A v_A) - (p_B c_B)} = \frac{1 - \beta - \alpha.[(p_A c_A)/(p_B c_B)]}{[(p_A c_A)/(p_B c_B)] - 1}$$

Finally, using the tangency condition, this becomes:

$$\frac{(1-\beta).(p_B c_B) - \alpha.(p_A v_A)}{(p_A v_A) - (p_B c_B)} = \frac{1 - \beta - \alpha.|g'(\alpha;\sigma)|}{|g'(\alpha;\sigma)| - 1}$$

where

$$|g'(\alpha;\sigma)| = \frac{\phi\left(\Phi^{-1}\left(g(\alpha;\sigma)\right)\right)}{\phi\left(\Phi^{-1}(1-\alpha)\right)}$$

Numerical Example:

Conventionally, an acceptable testing procedure correctly rejects the null hypothesis in 80% of instances where the alternative hypothesis is true: $1 - \beta = 0.80$. And, it's also conventional to use a 5% confidence level for rejecting the null hypothesis: $\alpha = 0.05$. Earlier we showed that this is compatible with a relative standard deviation of $\sigma = 0.40218$. When these conventional values of α and β are plugged into the expression for $|g'(\alpha;\sigma)|$, the calculated value of the slope becomes:

$$|g'(0.05; 0.40218)| = \frac{\phi\left(\Phi^{-1}(0.20)\right)}{\phi\left(\Phi^{-1}(0.95)\right)} = 2.7145$$

It follows that without conformance testing, the harm done by the occasional malfunctioning of the production process consumes over 50% of the firm's potential profit:

$$\frac{p_B c_B}{p_A v_A - p_B c_B} = \frac{1}{2.7145 - 1} = 0.58326$$

Moreover, the use of conformance testing increases the firm's profit by almost **40%**:

$$\frac{(1-\beta).(p_B c_B) - \alpha.(p_A v_A)}{(p_A v_A) - (p_B c_B)} = \frac{1 - 0.20 - 0.05 \times 2.7145}{2.7145 - 1} = 0.38745$$

5.2.17 A graphical approach using Iso-cost lines

For specific values of the basic parameters the reduced form of this problem could be solved numerically. However, it aids our understanding and intuition to attempt to solve it graphically by introducing iso-cost lines into the previous figure. Fixing the total cost, $\mathcal{E}$, so that we temporarily regard it as a kind of constant, imposes a negative relationship between α and β. In particular, β can be written as a function of α:

$$\beta = \frac{\mathcal{E} - \alpha.(p_A c_A)}{(p_B c_B)} = \frac{\mathcal{E}}{(p_B c_B)} - \alpha.\frac{(p_A c_A)}{(p_B c_B)}$$

This shows that, although, the intercepts vary with $\mathcal{E}$, the whole family of iso-cost lines must have the same negative slope:

$$\frac{d\beta}{d\alpha} = -\frac{p_A c_A}{p_B c_B} = -\left(\frac{1 - p_B}{p_B}\right)\left(\frac{c_A}{c_B}\right)$$

Moreover, if we assume that p_B is much smaller than p_A and also make c_A the unit of monetary value (set $c_A = v_A = 1$), then we get the following approximation:

$$\frac{d\beta}{d\alpha} \approx -\frac{1}{p_B c_B}$$

A little rearrangement gives the following result:

$$p_B c_B \approx -\frac{1}{d\beta/d\alpha} = \frac{1}{|d\beta/d\alpha|} = \frac{1}{slope}$$

In other words, the reciprocal of the slope of an iso-cost line is an approximation to the cost that would have been incurred by foregoing the testing process and simply accepting that occasionally the production process malfunctions and produces defective components. Moreover, when $c_A = v_A$, this gives us the cost of dealing with defective components as a proportion of the theoretically attainable profits.

Since these iso-cost lines have the same constant slope, they must run parallel to one another. Conventionally, the gradient of the iso-cost lines will be quite steep, because the probability of the production process malfunctioning is relatively small, and the penalty incurred from producing defective parts isn't too catastrophic. In the following figure, iso-cost lines (dashed lines) with a slope of -2 have been put in the same chart alongside the constraint curve for the case where $\sigma = 0.50$. Note that none of the analysis that follows depends on these specifics but including a figure for a specific case helps to explain and motive the analysis. The figure below shows that as the iso-cost lines move diagonally away from the origin (0,0) as the cost increases.

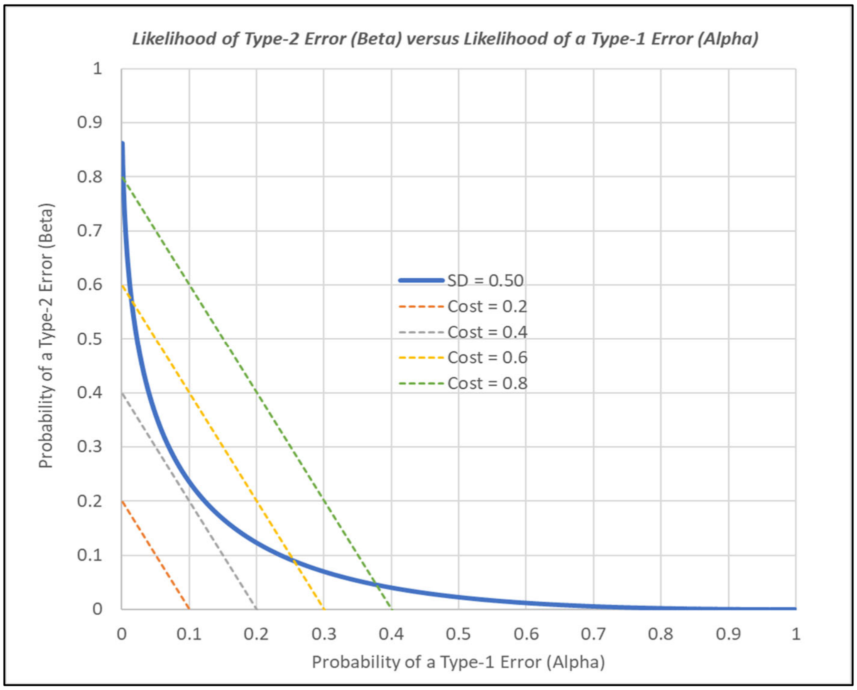

Figure 10: Relationship between the likelihood of type-1 error and the likelihood of type-2 error, plotted with the iso-cost lines

The firm is free to choose a point (a coordinate) anywhere on the dark blue curve and will act in a way that maximises its profit by minimising the cost of measurement errors. Moreover, this cost minimisation entails being on the lowest possible iso-cost line compatible with being on the dark blue curve. Because β is a decreasing convex function of α, there minimum attainable cost corresponds to the point of tangency between the dark blue line and an iso-cost line. For example, in the figure below, the dark blue line is approximately tangential to the iso-cost line for 'Cost = 0.8' at $(\alpha, \beta) = (0.10, 0.20)$. That is, being on the dark blue curve makes it impossible to reach iso-cost lines below the grey line corresponding to the iso-cost line for 'Cost = 0.8'. And all other permitted points (other than the point of tangency) will lead to higher iso-cost lines.

5.2.18 Scrap rate

So far, we've investigated the costs of measurement errors. Another indicator that is often of interest in manufacturing and relates to these costs is the scrap rate or yield loss. It is a measure of the production quality and is defined as the proportion of produced goods that are discarded or scrapped as defective. We can see from Figure 4 that a firm discards its produced goods in two cases: it can make a type-1 error and discard a perfect good product, or it can correctly discard a defective product. Thus, the scrap rate ($\mathcal{s}$) can be defined as:

$$\mathcal{s} = \mathcal{p}_A \cdot \alpha + \mathcal{p}_B \cdot (1 - \beta)$$

... Equation 42

Likewise, we can define another indicator called rebate rate $(\mathcal{r})$ that measures the rate at which defective products are returned by the customer to the firm. It can be expressed as:

$$\mathcal{r} = \mathcal{p}_B \cdot \beta$$

... Equation 43

We can incorporate the model parameter $\boldsymbol{\tau}$ into ... **Equation 42** as follows:

$$\mathcal{s} = \mathcal{p}_A \cdot \alpha + \mathcal{p}_B \cdot (1 - \beta)$$

$$= (1 - \wp_B) \cdot \alpha + \wp_B \cdot (1 - \beta)$$
$$= \left(1 - \frac{\wp_B c_B}{\wp_A c_A} \times \frac{\wp_A c_A}{c_B}\right) \cdot \alpha + \frac{\wp_B c_B}{\wp_A c_A} \times \frac{\wp_A c_A}{c_B} \cdot (1 - \beta)$$
$$= \left(1 - \tau \times \frac{\wp_A c_A}{c_B}\right) \cdot \alpha + \tau \times \frac{\wp_A c_A}{c_B} \cdot (1 - \beta)$$
$$\Rightarrow \mathcal{s} = \alpha + \tau \times \frac{\wp_A c_A}{c_B}(1 - \beta - \alpha)$$

... Equation 44

The above expression can be rearranged to express c_B/c_A in terms of $\mathcal{s}, \alpha, \beta$ and τ:

$$\frac{c_B}{c_A} = \frac{\tau \wp_A (1 - \beta - \alpha)}{\mathcal{s} - \alpha}$$
$$= \frac{\tau[(1 - \beta) \cdot \wp_A - \alpha \cdot \wp_A]}{\mathcal{s} - \alpha}$$
$$= \frac{\tau[(1 - \beta) \cdot \wp_A - \alpha \cdot \wp_A]}{\mathcal{s} - \alpha}$$
$$= \frac{\tau[(1 - \beta) \cdot (1 - \wp_B) - \alpha \cdot \wp_A]}{\mathcal{s} - \alpha}$$
$$= \frac{\tau[1 - \beta - \{(1 - \beta) \cdot \wp_B + \alpha \cdot \wp_A\}]}{\mathcal{s} - \alpha}$$
$$\Rightarrow \frac{c_B}{c_A} = \frac{\tau[1 - \beta - \mathcal{s}]}{\mathcal{s} - \alpha}$$

... Equation 45

The LHS of ... Equation 45 is a cost ratio, so it will be positive. For the RHS to be positive, one of the following two cases must be true: *1)* $1 - \beta - \mathcal{s} > 0$ & $\mathcal{s} - \alpha > 0 \Rightarrow \alpha < \mathcal{s} < 1 - \beta$; or *2)* $1 - \beta - \mathcal{s} < 0$ & $\mathcal{s} - \alpha < 0 \Rightarrow 1 - \beta < \mathcal{s} < \alpha$.

Let us look at the first case: $\alpha < \mathcal{s} < 1 - \beta$. Consider the extreme scenario where the production process never produces defective goods, i.e., $\wp_A = 1$. In this scenario, the firm will incorrectly scrap a perfectly good product whenever it makes a type-1 error. Therefore, the scrap rate will be equal to the likelihood of making a type-1 error, α. Consider the other extreme now, where the production process always malfunctions, i.e., $\wp_B = 1$. In this scenario, the firm will fail to scrap a defective product whenever it makes a type-2 error, and the likelihood of type-2 error is β. Therefore, it will correctly scrap defective goods at a rate that is equal to $1 - \beta$. In a more realistic scenario, where $\wp_A$ and $\wp_B$ both lie between 0 and 1, α and β provide the lower and upper bounds, respectively, for the scrap rate $\mathcal{s}$.[26]

Lastly, notice that we can re-insert the expression for $\mathcal{s}$ from ... Equation 42 back into the beginning step of ... Equation 45 to express $\wp_A$ and $\wp_B$ in terms of τ, α, β and c_B/c_A to obtain:

$$\frac{c_B}{c_A} = \frac{\tau \wp_A}{1 - \wp_A}$$

---

[26] The second case: $1 - \beta < \mathcal{s} < \alpha$, can similarly be understood using the more unlikely extreme scenarios where $\alpha = 0$ and $\beta = 1$.

$$\Rightarrow \wp_A = \frac{(c_B/c_A)}{\tau + \frac{c_B}{c_A}}$$

... Equation 46

Using the constraint that $\wp_A + \wp_B = 1$, we obtain:

$$\wp_B = \frac{\tau}{\tau + \frac{c_B}{c_A}}$$

... Equation 47

6 INCORPORATING CALIBRATION INTO THE MODEL

Measurement activity would be possible without the work of NPL, but it wouldn't be as effective at producing reliable information. It's reasonable to assume that the benefits of NPL mainly come through the support it provides to the provision of calibration services. Hence, the next stage of the analysis involves extending the model slightly to account for the effect of calibration services. The idea behind the following analysis is that, although, calibration services can't affect $\wp_A c_A$ or $\wp_B c_B$, calibration can reduce measurement uncertainty, and so reduce the relative standard deviation, σ.

6.1 THE IMPORTANCE OF ERRORS ASSOCIATED WITH CALIBRATION TO THE MEASUREMENT UNCERTAINTY

Suppose that the relative standard deviation, σ, derives from a variance that can split into two parts: The first part, $(\sigma_0)^2$, contains random factors that can be averaged out by repeated measurements but are not related to accurate calibration. The second part, ϵ^2, is the component of the overall measurement uncertainty that can be eliminated by accurate calibration. Hence, the overall variance is $\sigma^2 = (\sigma_0)^2 + \epsilon^2$, so that the relative standard deviation can be written as:

$$\sigma = \sqrt{(\sigma_0)^2 + \epsilon^2}$$

... Equation 48

Calibration affects the firm's cost minimisation problem by reducing the relative standard deviation. If the firm buys high quality (primary) calibrations from NPL or other top-tier calibration laboratories, then we assume that the second component of the uncertainty is eliminated: $\epsilon \to 0$ so that $\sigma \to \sigma_0$. However, it's also possible for the firm to buy secondary calibrations from non-top-tier labs. Using lower-quality calibrations reduces the second component of the uncertainty compared to not using any calibrations, but it does not eliminate it entirely: $\epsilon > 0$. If we treat σ_0 as fixed in ... Equation 48, then we can write $\sigma = \sigma(\epsilon)$. Note that an increment of $\Delta\epsilon (= \epsilon - 0)$ in the uncertainty associated with calibration errors will affect standard deviation, which can be expressed using the Taylor expansion of $\sigma(\epsilon)$ around zero:[27]

[27] If $f(x)$ is a real or complex-valued function of x that is infinitely differentiable at the point a, then it's Taylor series expansion around a can be expressed as:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

where $n!$ denotes the factorial of n, and $f^{(n)}(a)$ denotes the n^{th} derivative of f evaluated at the point a. If we assume that obtaining high-quality calibrations from NPL eliminates ϵ completely (i.e., $\epsilon = 0$), then we can write the Taylor expansion of σ around this point to determine what happens when ϵ increases.

$$\sigma(\epsilon) = \sigma(0) + \frac{\sigma'(0)}{1!}(\epsilon - 0) + \frac{\sigma''(0)}{2!}(\epsilon - 0)^2 + \frac{\sigma'''(0)}{3!}(\epsilon - 0)^3 + \cdots$$

... Equation 49

If ϵ is small relative to σ_0 (i.e., secondary calibrations increase the overall uncertainty of measurement, but the change is small relative to σ_0), higher-power terms can be ignored and ... Equation 49 can be approximated as:

$$\sigma(\epsilon) \approx \sigma(0) + \frac{\sigma'(0)}{1!}(\epsilon - 0) + \frac{\sigma''(0)}{2!}(\epsilon - 0)^2$$

... Equation 50

Substituting $\epsilon = 0$ in ... Equation 48, we get $\sigma(0) = \sigma_0$. Taking the first derivative of σ with respect to ϵ, we get: $\sigma'(\epsilon) = \frac{\epsilon}{\sqrt{(\sigma_0)^2+\epsilon^2}}$. Therefore, $\sigma'(0) = 0$. Taking the second derivative of σ with respect to ϵ, we get: $\sigma''(\epsilon) = \frac{\sqrt{(\sigma_0)^2+\epsilon^2} - \epsilon\left(\sigma'(\epsilon)\right)}{(\sigma_0)^2+\epsilon^2}$. Therefore, $\sigma''(0) = 1/\sigma_0$. Substituting these values into ... Equation 50, we get:

$$\sigma(\epsilon) \approx \sigma_0 + \frac{1}{2\sigma_0}(\epsilon)^2$$

Slightly rearranging the above expression and defining $\Delta\sigma = \sigma(\epsilon) - \sigma(0) = \sigma(\epsilon) - \sigma_0$:

$$\Delta\sigma \approx \frac{1}{2\sigma_0}(\epsilon)^2$$

... Equation 51

If the change in the standard deviation derives from an underlying change in the uncertainty associated with calibration errors, then ... Equation 51 can be re expressed as follows:

$$\frac{\Delta\sigma}{\sigma_0} \approx \frac{1}{2}\left(\frac{\epsilon}{\sigma_0}\right)^2$$

... Equation 52

These terms are easy to interpret. The LHS of ... Equation 52 gives the increase in uncertainty when a firm uses secondary calibrations relative to the scenario where it uses primary calibrations (baseline scenario). The ratio ϵ/σ_0 in the RHS of ... Equation 52 denotes the fraction of overall measurement uncertainty that comes from inaccurate, secondary calibrations.[28]

6.1.1 The sensitivity of type-1 and type-2 errors to a change in the relative standard deviation

In Section 5.2.14, we derived the expressions for the optimal choices of α and β given the model parameters σ and τ. These expressions were given by:

$$\alpha = \alpha^*(\sigma, \tau) = 1 - \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right)$$

$$\beta = \beta^*(\sigma, \tau) = \Phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)$$

[28] The fraction of overall measurement uncertainty that comes from inaccurate, second-hand calibrations is ϵ/σ. However, if ϵ is small relative to σ_0, then for practical purposes we can assume $\sigma \approx \sigma_0$ to help with the exposition (as σ_0 is fixed).

To understand the effect of a change in the standard deviation, σ, on the likelihoods of making type-1 and type-2 errors, we need to conduct some comparative statics. That is, we evaluate the following:[29]

$$\frac{\partial \alpha^*(\sigma,\tau)}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left[1 - \Phi\left(2\sigma \cdot \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right)\right]$$

$$\Rightarrow \frac{\partial \alpha^*(\sigma,\tau)}{\partial \sigma} = -\phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right) \times \left[2 \cdot \ln\left(1/\sqrt{\tau}\right) - \frac{1}{2\sigma^2}\right]$$

$$= \phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right) \times \left[\frac{1}{2\sigma^2} - 2 \cdot \ln\left(1/\sqrt{\tau}\right)\right]$$

... Equation 53

Likewise,

$$\frac{\partial \beta^*(\sigma,\tau)}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left[\Phi\left(2\sigma \cdot \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right)\right]$$

$$\Rightarrow \frac{\partial \beta^*(\sigma,\tau)}{\partial \sigma} = \phi\left(2\sigma . \ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right) \times \left[2 \cdot \ln\left(1/\sqrt{\tau}\right) + \frac{1}{2\sigma^2}\right]$$

... Equation 54

From ... Equation 53, it is clear that $\frac{\partial \alpha^*(\sigma,\tau)}{\partial \sigma} > 0$ if and only if $\frac{1}{2\sigma^2} - 2 \cdot \ln\left(1/\sqrt{\tau}\right) > 0$. Slight rearrangement gives us that $\frac{\partial \alpha^*(\sigma,\tau)}{\partial \sigma} > 0$ if and only if $\exp\left(-\frac{1}{2\sigma^2}\right) < \tau$. That is, whether $\alpha^*(\sigma,\tau)$ increases or decreases with an increase in σ is determined by the value of σ.[30] On the other hand, ... Equation 54 tells us that an increase in σ always increases $\beta^*(\sigma,\tau)$.

6.1.2 The sensitivity of the cost from mistakes to a change in the relative standard deviation

In the previous section of this study, the objective function (cost function) minimised by the firm was defined as

$$\mathcal{E} = H(\alpha; \sigma, \boldsymbol{p}, \boldsymbol{c}) = (p_A c_A).\alpha + (p_B c_B).g(\alpha;\sigma)$$

where the constraint is

$$g(\alpha;\sigma) = \Phi\left(\Phi^{-1}(1-\alpha) - (1/\sigma)\right)$$

$\alpha^*(\sigma,\tau)$ denotes the value of α that minimises the cost incurred by the firm due to measurement errors, and $\beta^*(\sigma,\tau) = g(\alpha^*(\sigma,\tau);\sigma) = \Phi\left(\Phi^{-1}\left(1-\alpha^*(\sigma,\tau)\right) - (1/\sigma)\right)$ denotes the value of β that minimises the cost incurred by the firm due to measurement errors.

The minimum attainable cost then becomes $H(\alpha^*(\sigma,\tau);\sigma,\boldsymbol{p},\boldsymbol{c})$. To understand the effect of a change in the standard deviation, σ, on the outcome of firm's cost minimisation problem we conduct some comparative statics. That is, we wish to find an expression for the following derivative:

$$\frac{\partial}{\partial \sigma}\left[H(\alpha^*(\sigma,\tau);\sigma,\boldsymbol{p},\boldsymbol{c})\right]$$

---

[29] Notice that all parameters except for σ remain unchanged.

[30] $\exp\left(-\frac{1}{2\sigma^2}\right)$ is monotonically increasing in σ. Notice that $\exp\left(-\frac{1}{2\sigma^2}\right) \to 0$ as $\sigma \to 0$; and $\exp\left(-\frac{1}{2\sigma^2}\right) \to 1$ as $\sigma \to \infty$. Assumption 11 implies that $0 < \tau < 1$. Thus, for a given τ, $\exp\left(-\frac{1}{2\sigma^2}\right)$ will be less than τ for small values of σ and it will be greater than τ for large values of σ. Combining this with ... Equation 53 tells us that for a given τ, the optimal level of α will increase with measurement uncertainty if the measurement uncertainty is low. Whereas the optimal level of α will decrease with measurement uncertainty if the measurement uncertainty is high.

An ‘envelope argument’ from classic optimisation theory enables us to answer this question and find the effect of an increase in σ on the minimised cost. The first step is to find the partial derivative of the objective function with respect to σ:

$$\frac{\partial}{\partial\sigma}[H(\alpha;\sigma,\boldsymbol{p},\boldsymbol{c})] = (p_B c_B).\frac{\partial}{\partial\sigma}[g(\alpha;\sigma)] = (p_B c_B).(1/\sigma^2).\phi\big(\Phi^{-1}(1-\alpha)-(1/\sigma)\big)$$

Next, we evaluate this partial derivative at the value of α that solves the firm’s minimisation problem:

$$[(p_B c_B)/\sigma^2].\phi\big(\Phi^{-1}(1-\alpha^*(\sigma,\tau))-(1/\sigma)\big) = [(p_B c_B)/\sigma^2].\phi\big(\Phi^{-1}(\beta^*(\sigma,\tau))\big)$$

Hence, the result this comparative statics is given by:

$$\frac{\partial}{\partial\sigma}[H(\alpha^*(\sigma,\tau);\sigma,\boldsymbol{p},\boldsymbol{c})] = [(p_B c_B)/\sigma^2].\phi\big(\Phi^{-1}(\beta^*(\sigma,\tau))\big)$$

... Equation 55

Corollary 3: Notice that because the LHS of ... **Equation 55** is necessarily positive, an increase in the standard deviation always leads to an increase in the minimum attainable cost for measurement errors.

If the standard deviation were to increase by some increment, $\Delta\boldsymbol{\sigma}$, then the increase in the minimum attainable cost becomes:

$$\Delta\mathcal{E} = \Delta H(\alpha^*(\sigma,\tau);\sigma,\boldsymbol{p},\boldsymbol{c}) = (p_B c_B).\phi\big(\Phi^{-1}(\beta^*(\sigma,\tau))\big).(\Delta\sigma/\sigma^2).$$

Since $\Delta\sigma/\sigma$ is easy to interpret as the percentage change in the standard deviation, let us rewrite the above expression as follows:

$$\Delta\mathcal{E} = \Delta H(\alpha^*(\sigma,\tau);\sigma,\boldsymbol{p},\boldsymbol{c}) = \big[(1/\sigma).(p_B c_B).\phi\big(\Phi^{-1}(\beta^*(\sigma,\tau))\big)\big].(\Delta\sigma/\sigma).$$

If we consider the increase in σ happens from the baseline case where a firm uses primary calibrations (i.e., $\sigma = \sigma_0$), then using our earlier result for β^* from ... Equation 34, the above expression can be rewritten as:

$$\Delta\mathcal{E} = \Delta H(\alpha^*(\sigma_0,\tau);\sigma,\boldsymbol{p},\boldsymbol{c}) = \left[(1/\sigma_0).(p_B c_B).\phi\left(2\sigma_0.\ln(1/\sqrt{\tau})-(1/2\sigma_0)\right)\right].(\Delta\sigma/\sigma_0).$$

... Equation 56

This expression gives the *increased* cost that would come from an increase in the relative standard deviation of the measurement process from the baseline case where the firm uses primary calibrations from top-tier laboratories. Moreover, if we assume that $c_A = v_A$, then dividing this expression through by $p_A c_A$ gives us the relative increase in costs that comes from making more mistakes. The relative increase in costs created by less accurate calibration is given by:

$$\Delta\mathcal{E}/(p_A c_A) = (\Delta\sigma/\sigma_0).\left[(\tau/\sigma_0).\phi\left(2\sigma_0.\ln(1/\sqrt{\tau})-(1/2\sigma_0)\right)\right]$$

This can be re expressed as:

$$\Delta e = (\Delta\sigma/\sigma_0).\left[(\tau/\sigma_0).\phi\left(2\sigma_0.\ln(1/\sqrt{\tau})-(1/2\sigma_0)\right)\right]$$

... Equation 57

where $\tau = (p_B c_B)/(p_A c_A)$ and $\Delta\mathcal{E} = \Delta e \times (p_A c_A)$.

The term in square brackets gives us the harm caused (cost increase) from reducing the accuracy of the measurement process, and so increasing the likelihood of type-1 and type-2 errors. More specifically, this term reflects the increases in the cost of type-1 and type-2 errors that would result from increasing the relative standard deviation. In other words, it’s the partial elasticity for the cost of type-1 and type-2 errors with respect to an increase in the relative standard deviation of the measuring process. Hence, the term in square brackets gives the sensitivity of the cost from mistakes to changes in the relative standard deviation.

The sensitivity of the (normalised) cost of mistakes to changes in σ is illustrated in Figure 11.

The curves correspond to four different choices for the cost ratio $\tau = (\wp_B c_B)/(\wp_A c_A)$, where $\wp_B c_B$ is the cost of type-2 errors that would be incurred if the firm were to forego the conformance testing. Figure 11 shows that the cost is most sensitive to changes in the relative standard deviation for intermediate value of σ. That is, a 10% in σ tends to have a noticeable effect on the cost of mistakes when the relative standard deviation is in the range $0.35 < \sigma < 0.55$.

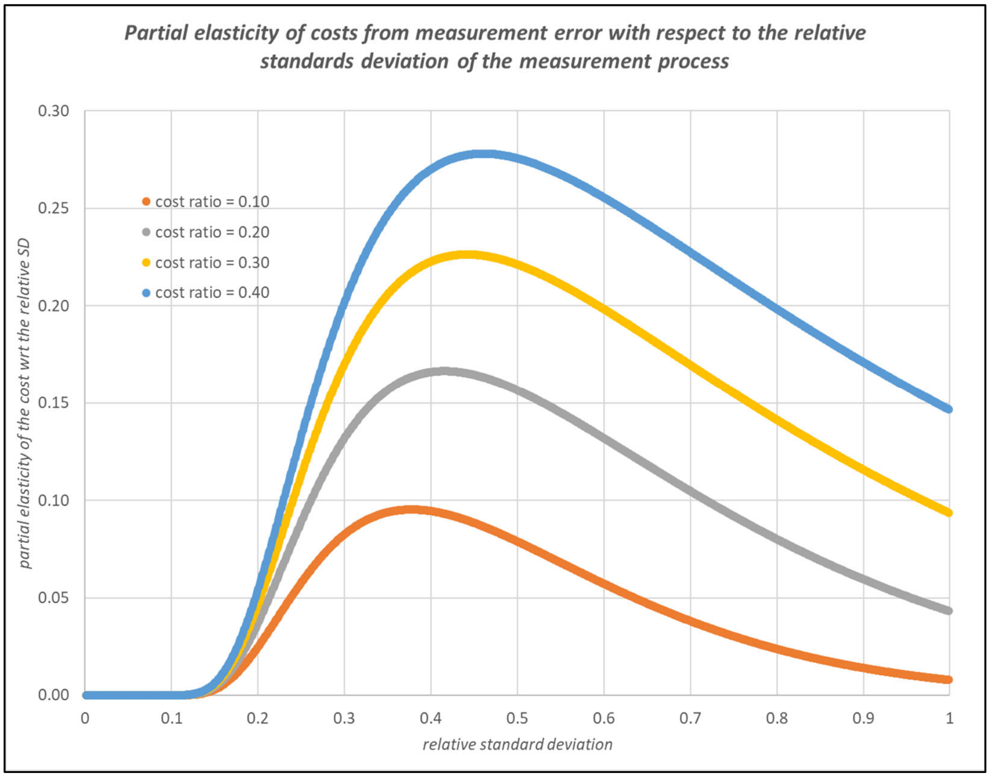

Figure 11: Partial elasticity of costs from measurement error with respect to σ

6.1.3 The benefit of high-quality calibration services

The effect of high-quality, accurate calibration on the minimum attainable cost depends on: (1) the sensitivity of the costs from mistakes to changes in the relative standard deviation; and (2) the importance of uncertainties associated with calibration to the relative standard deviation. Recall that the (normalised) value of the information generated by conformance testing is given by $w = \tau - e = \tau - \tilde{h}(\sigma, \tau)$. Hence, we let $\Delta w = -\Delta e = -\Delta \mathcal{E}/(\wp_A c_A)$ denote the relative importance of the benefits that come from high-quality, accurate calibration.

We show the potential benefits of high-quality calibration services using the following stylized numerical example. In the example, we assume values of α and β to solve for $\sigma, \tau, \wp_A$ (i.e., we solve the model discussed in Section 5.2 in a reverse order), and then explore how a change in σ resulting from higher uncertainties associated with secondary calibrations affects costs. It is important to note that in the real world, the firm has some knowledge of the model parameters, and it minimizes its costs by choosing optimal levels of α and β. However, the example below shows that it is possible to use the model to solve for α and β in terms of the parameters, and vice versa.

In the example, we first consider a baseline scenario where we assume that the firm uses primary calibrations provided by NPL or other top-tier calibration labs.

Numerical Example:

Baseline Scenario – In this scenario, we assume that the firm uses primary calibrations for the measurement process in conformance testing. Conventionally, this testing procedure correctly rejects the null hypothesis in 80% of instances where the alternative hypothesis is true: $1-\beta=0.80$ or equivalently $\beta=0.20$. Furthermore, it has a conventional 5% confidence level for rejecting the null hypothesis, i.e., $\alpha=0.05$. Plugging these values of α and β into ... Equation 9, we get the corresponding relative standard deviation:

$$\sigma=\sigma_0=\frac{1}{\Phi^{-1}(1-\alpha)-\Phi^{-1}(\beta)}=\frac{1}{\Phi^{-1}(0.95)-\Phi^{-1}(0.20)}=0.402$$

Likewise, plugging these values of α and β into ... Equation 27, we obtain cost ratio:

$$\begin{aligned}\tau=\exp\left\{\tfrac{1}{2}[\Phi^{-1}(\beta)]^2-\tfrac{1}{2}[\Phi^{-1}(1-\alpha)]^2\right\}&=\exp\left\{\tfrac{1}{2}[\Phi^{-1}(0.20)]^2-\tfrac{1}{2}[\Phi^{-1}(1-0.05)]^2\right\}\\&=\exp\left\{\tfrac{1}{2}[\Phi^{-1}(0.20)]^2-\tfrac{1}{2}[\Phi^{-1}(0.95)]^2\right\}=\exp\left\{\tfrac{1}{2}[-0.842]^2-\tfrac{1}{2}[1.645]^2\right\}\\&=\exp\{-0.999\}=0.368.\end{aligned}$$

Next, we can plug the above value of τ into ... Equation 46 and ... Equation 47, and combine them with Assumption 10 (i.e., $c_B/c_A=2$) to obtain:

$$\wp_A=\frac{\left(\frac{c_B}{c_A}\right)}{\tau+\frac{c_B}{c_A}}=\frac{2}{0.368+2}=0.845$$

$$\wp_B=\frac{\tau}{\tau+\frac{c_B}{c_A}}=\frac{0.368}{0.368+2}=0.155$$

Finally, we can solve for the scrap rate ($\mathcal{s}$) using ... Equation 42 and rebate rate ($\mathcal{r}$) using ... Equation 43:

$$\begin{aligned}\mathcal{s}&=\wp_A\cdot\alpha+\wp_B\cdot(1-\beta)\\&=0.845\times0.05+0.155\times(1-0.2)=0.166\end{aligned}$$

$$\mathcal{r}=\wp_B\cdot\beta=0.155\times0.2=0.031$$

Next, we assume a counterfactual scenario where the firm uses secondary calibration services provided by lower-tier calibration labs. Lower-tier calibration labs generally calibrate their own instruments using the services of the top-tier calibration labs, and then use these instruments to provide calibrations to a firm.

Numerical Example:

Counterfactual Scenario – It is reasonable to suppose that using these secondary calibrations leads to an increase in the uncertainty associated with the measurement process used by the firm (or end-user). This increase in uncertainty over the baseline, σ_0, is given by ... Equation 52: $\frac{\Delta\sigma}{\sigma_0}\approx\frac{1}{2}\left(\frac{\epsilon}{\sigma_0}\right)^2$. We can use the traditional concept of Test Accuracy Ratio (TAR) to roughly estimate the RHS of this expression. TAR is defined as the ratio of an instrument's accuracy to the accuracy of the standard that is used to calibrate the instrument. As a generally accepted heuristic guideline, metrology labs strive for a minimum TAR of 4:1.[31] Roughly speaking, a TAR of 4:1 implies $\frac{\epsilon}{\sigma_0} = \frac{1}{4}$ in the above expression. Thus, we get:

$$\frac{\Delta\sigma}{\sigma_0} \approx \frac{1}{2}\left(\frac{1}{4}\right)^2 = \frac{1}{32}$$

Furthermore, this implies that $\sigma = \sigma_0 + \Delta\sigma = \frac{33}{32}\sigma_0 = \frac{33}{32} \times 0.402 = 0.415.$

In contrast to the relative standard deviation, τ is solely determined by exogenous factors and is unlikely to be affected by a change in measurement uncertainty due to calibration. And, therefore, we can consider τ to be fixed at the same value as that found in the baseline Scenario: $\tau = 0.368$. By plugging these values of σ and τ into ... Equation 33 and ... Equation 34, we get:

$$\alpha^*(\sigma,\tau) = 1 - \Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) + (1/2\sigma)\right)$$
$$= 1 - \Phi\left(2 \times 0.415 \times \ln\left(1/\sqrt{0.368}\right) + (1/2 \times 0.415)\right) = 1 - \Phi(1.620)$$
$$= 1 - 0.947 = 0.053.$$
$$\beta^*(\sigma,\tau) = \Phi\left(2\sigma.\ln\left(1/\sqrt{\tau}\right) - (1/2\sigma)\right) = \Phi\left(2 \times 0.415 \times \ln\left(1/\sqrt{0.368}\right) - (1/2 \times 0.415)\right)$$
$$= \Phi(-0.791) = 0.214.$$

Note that $\mathcal{p}_A$ and $\mathcal{p}_B$ will remain unchanged from the baseline scenario as τ doesn't change. However, just like in the baseline scenario, we can solve for the scrap rate ($\mathcal{s}$) using ... Equation 42 and rebate rate ($\mathcal{r}$) using ... Equation 43:

$$\mathcal{s} = \mathcal{p}_A \cdot \alpha^*(\sigma,\tau) + \mathcal{p}_B \cdot \left(1 - \beta^*(\sigma,\tau)\right)$$
$$= 0.845 \times 0.053 + 0.155 \times (1 - 0.214) = 0.166$$

$$\mathcal{r} = \mathcal{p}_B \cdot \beta^*(\sigma,\tau) = 0.155 \times 0.214 = 0.033$$

It is evident that as the measurement uncertainty goes up, the likelihoods of making type-1 and type-2 errors also increases too. Plugging the values of $\frac{\Delta\sigma}{\sigma_0}$, σ_0, and τ into ... Equation 57, we get:

$$\Delta e = \left(\Delta\sigma/\sigma_0\right).\left[\left(\tau/\sigma_0\right) \times \phi\left(2\sigma_0.\ln\left(1/\sqrt{\tau}\right) - \left(1/2\sigma_0\right)\right)\right]$$
$$= \left(\frac{1}{32}\right).\left[\left(\frac{0.368}{0.402}\right) \times \phi\left(2 \times 0.402 \times \ln\left(1/\sqrt{0.368}\right) - (1/2 \times 0.402)\right)\right]$$
$$= \left(\frac{1}{32}\right).[0.916 \times \phi(-0.842)] = \left(\frac{1}{32}\right) \times 0.916 \times 0.280 = 0.00801 = 0.801\%$$

That is, the cost of mistakes made during the conformance testing process (as fraction of the output from production) would increase by about 0.801 percentage points if the firm were to forego high-quality, 'primary' calibrations. Since we know that $\Delta e = \Delta\mathcal{E}/(\mathcal{p}_A c_A)$, having an estimate of $\mathcal{p}_A c_A$ allows us to compute the absolute cost increase ($\Delta\mathcal{E}$). A survey of customers using the NMS labs performed by Winning Moves in 2018 (King and Tellett (2020)) revealed that the manufacturing firms that engaged with the NMS labs - representing around 12% of the UK's manufacturing workforce - have an aggregate turnover of around £82.3 billion. Remember that $\mathcal{p}_A c_A$ represents the absolute theoretical maximum value of the produced goods. Assuming this is roughly equal to the aggregate turnover of the manufacturing firms working with the NMS labs, we can infer that these firms will incur an additional cost of £660 million if they forego primary calibrations provided by top-tier labs like the NPL. (The calculation is $\Delta\mathcal{E} = \mathcal{p}_A c_A \times \Delta e =$ £82.3×0.00801 billion $\approx$ £0.660 billion.)

The Office for National Statistics (ONS) publishes yearly data on the UK's non-financial businesses as measured by the Annual Business Survey.[32] In 2019, the total turnover of the UK's manufacturing sector was about £565.3 billion, and the approximate gross value added (GVA) from manufacturing was £170.4 billion. This gives the ratio of GVA to total turnover as 0.3 ($\approx 170.4/565.3$). Using this ratio, we can say that GVA safeguarded through supplying high-quality, primary calibrations to manufacturing firms that work with the NMS labs is approximately $£197$ million ($\approx 0.3 \times £660$ million).

The main assumptions in the stylized example given above can be summarized as follows: We consider a firm that uses primary calibrations provided by NPL (or other top-tier calibration labs) and suppose that this eliminates any calibration-related uncertainties. Furthermore, the traditional concept of Test Accuracy Ratio (TAR) can be used to estimate what the percentage increase in uncertainty might be if the firm were to lose access to these primary calibrations. As a generally accepted heuristic rule, any lab engaged in conformance testing strives for a minimum TAR of 4:1, which implies the extra uncertainty introduced through poorer calibration could be as much as 25% of the baseline uncertainty (σ_0) of the measurement process.

Secondly, the baseline uncertainty - the part that cannot be removed through calibration - is such that the firm's measurement process has a 5% chance of making a type-1 error (when testing a "good" item) and a 20% chance of making a type-2 error (when testing a "bad" item). That is, our example is based on conventional levels of $\alpha = 0.05$ and $\beta = 0.20$.

Lastly, we suppose that making a type-2 error is twice as costly as making a type-1 error. (For example, the firm may offer to compensate any customers returning defective goods by supplying a replacement free of charge.)

The main findings from this stylized example can be summarized as follows: Our calculations show that if the firm loses access to such high-quality calibrations, then the overall uncertainty associated with its measurement process increases by 3.13% (the calculation is $\Delta\sigma/\sigma_0 = 1/32 \approx 3.13\%$). This increase in measurement uncertainty results in:

- $\sim 6\%$ increase in the likelihood of making a type-1 error, and $\sim 7\%$ increase in the likelihood of making a type-2 error.
    - $\Delta\alpha = 0.053 - 0.050 = 0.003$ (0.3 percentage points) and $0.003/0.05 \approx 6\%$.
    - $\Delta\beta = 0.214 - 0.200 = 0.014$ (1.4 percentage points) and $0.014/0.20 \approx 7\%$.
- $\sim 0.8\%$ of the value from production would be lost due to the extra costs incurred from making more mistakes during the conformance testing process, where these extra costs are a direct consequence of relying on a less accurate measurement process.

In the Annex at the end of this paper, we show two more worked-out numerical examples to compute the potential benefits from primary calibrations starting with different values of α and β.

7 CONCLUSION AND IDEAS FOR FURTHER WORK

7.1 CONCLUSION

The first part of this document provides an economic model for the value created when using measurement for conformity testing. This model demonstrated how measurement information creates value by reducing mistakes (fewer "false positives" and "false negatives")

[31] https://www.duncanaviation.aero/intelligence/2019/January/aircraft-tool-calibration-what-is-test-accuracy-ratio#:~:text=TAR%20is%20a%20ratio%20of,that%20the%20tool%20being%20calibrated. The minimum 4:1 TAR was a heuristic rule that has possibly been updated, however, using it here helps simplify the exposition in this stylized numerical example.

[32] The ONS dataset can be downloaded from the link below: https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/datasets/uknonfinancialbusinesseconomyannualbusinesssurveysectionsas. For the analysis that follows, we use the 2019 numbers for Manufacturing sector (SIC "C") from "Section C" sheet of the Excel file.

in conformity testing. The second part of this document then introduces calibration as a perturbation on top of these measurement activities.

Previous studies have explored the economic loss from making type-1 (false positive) and type-2 errors (false negatives) in the context of conformance testing. However, such studies do not treat the probabilities of making type-1 and type-2 errors as “choice variables” that are under control of the decision maker. Through the analysis in this document, this gap in the literature is addressed by endogenizing the probability of a type-1 error (α) and the probability of a type-2 error (β).

The specific contributions are as follows: Firstly, this document gives formulae for the relative uncertainty of the measurement process (σ) and a lower bound for the scrap rate (τ) as functions of the probabilities for type-1 and type-2 errors (α and β). Secondly, we provide a formula for the economic benefit created by using ‘primary’ calibrations to reduce the uncertainty of the measurement process. That is, this document develops the mathematics behind a ‘calculator’ for finding the benefit from ‘primary’ calibrations based on assumptions about the parameters. Lastly, we provide some worked examples that illustrate how this ‘calculator’ can be used to find the benefit of high-quality calibrations in a range of scenarios based on plausible assumptions about the parameter values that characterise each scenario.

7.2 FURTHER WORK

A follow-on study is required to quantify the value of the fan-out. Such a study would need to make assumptions about the distribution of unobservable elements of the model. For example, a distribution for the importance that end-users attach to measurement information and degree to which they are willing to pay a premium for high quality calibration services in situations where such calibrations have an appreciable effect on the uncertainty of the measurement process.

The primary goal of a follow-on study should be to solve the model and to find formulae for observable quantities, such as, the number of firms who pay for NPL’s services and its income from selling calibration services. The expressions for such observable quantities will tend to involve theoretical elements from the set-up, such as, the unobservable parameters characterising the distribution of end-users’ willingness to pay for high quality calibration services. Similarly, it may be possible to find expressions for unobservable quantities, such as, the benefit going to customers using lower tier calibration labs.

Finally, the analysis developed in this document will really come to life once we have estimates of the core parameters from a measurement survey. Hence, this model has are big implications for questions to include in the in the next NMS Customer Survey. Moreover, there is strong potential for this model to be used as the basis of some quantitative case studies. By this means it is hoped that future studies will be able to make inferences about the scale of the fanout based on what is known about observable quantities.

8 REFERENCES

Beges, G., Drnovsek, J., & Pendrill, L. R. (2010). Optimising calibration and measurement capabilities in terms of economics in conformity assessment. *Accreditation and quality assurance*, *15*, 147-154.

King, M; Tellett, G (2020). A customer survey for three of the core labs in the National Measurement System. *Technical Report.* https://eprintspublications.npl.co.uk/8655/

Kunzmann, H., Pfeifer, T., Schmitt, R., Schwenke, H., & Weckenmann, A. (2005). Productive metrology-adding value to manufacture. *CIRP annals*, *54*(2), 155-168.

Savio, E. (2012). A methodology for the quantification of value-adding by manufacturing metrology. *CIRP annals*, *61*(1), 503-506.

Swann, G. P. (2009). The Economics of Metrology and Measurement. *Report for National Measurement Office, Department of Business, Innovation and Skills (BIS)*.

Thompson, M., & Fearn, T. (1996). What exactly is fitness for purpose in analytical measurement?. *Analyst*, *121*(3), 275-278.

Usuda, T., & Henson, A. (2012). Economic impact of equivalence of measurement standards. *NCSLI Measure*, *7*(1), 62-70.

9 ANNEX

9.1 ANNEX 1 – NUMERICAL EXAMPLE WITH $\alpha = 0.01$; $\beta = 0.1$

In Section 6.1.3, we presented a stylized numerical example to compute the benefit of high-quality calibration services. In that toy example, we assumed that the measurement process for a firm that uses primary calibrations from top-tier labs like the NPL has the conventional levels of $\alpha = 0.05$ and $\beta = 0.2$.

Here, we explore what happens when we start with lower values of α and β. For simplicity, we can think of this example as representing a more mature industry where the measurement process has advanced quite a bit; hence the competition is primarily over cost and price. In comparison, the stylized example in Section 6.1.3 can be thought of as representing a relatively young industry where the measurement process hasn't become totally efficient. Thus, the competition is less over cost and price and more on the novelty or functionality of the product. That is, there is some degree of product differentiation that softens the competition.

<u>**Numerical Example**</u> ($\boldsymbol{\alpha = 0.01}$ **and** $\boldsymbol{\beta = 0.1}$):

Baseline Scenario (Firm uses primary calibrations provided by NPL or other top-tier calibration labs) – In this scenario, we assume that the firm uses primary calibrations for the measurement process in conformance testing, and the process has $\alpha = 0.01$ and $\beta = 0.1$. The α and β here are lower than the conventional levels that we assumed in Section 6.1.3. Plugging these values of α and β into ... Equation 9 and ... Equation 27, we get:

$$\sigma = \sigma_0 = \frac{1}{\Phi^{-1}(0.99) - \Phi^{-1}(0.10)} = 0.277$$

$$\tau = \exp\left\{\tfrac{1}{2}[\Phi^{-1}(0.10)]^2 - \tfrac{1}{2}[\Phi^{-1}(0.99)]^2\right\} = 0.152.$$

Next, we can plug the above value of τ into ... Equation 46 and ... Equation 47, and combine them with Assumption 10 (i.e., $c_B/c_A = 2$) to obtain:

$$\mathcal{p}_A = \frac{\left(\frac{c_B}{c_A}\right)}{\tau + \frac{c_B}{c_A}} = \frac{2}{0.152 + 2} = 0.929$$

$$\mathcal{p}_B = \frac{\tau}{\tau + \frac{c_B}{c_A}} = \frac{0.152}{0.152 + 2} = 0.071$$

Finally, we can solve for the scrap rate ($\mathcal{s}$) using ... Equation 42 and rebate rate ($\mathcal{r}$) using ... Equation 43:

$$\mathcal{s} = \mathcal{p}_A \cdot \alpha + \mathcal{p}_B \cdot (1 - \beta)$$
$$= 0.929 \times 0.01 + 0.071 \times (1 - 0.1) = 0.073$$

$$\mathcal{r} = \mathcal{p}_B \cdot \beta = 0.071 \times 0.1 = 0.007$$

Counterfactual Scenario (Firm uses secondary calibrations provided by lower-tier calibration labs) – The increase in uncertainty over the baseline, σ_0, is given by ... Equation 52: $\frac{\Delta\sigma}{\sigma_0} \approx \frac{1}{2}\left(\frac{\epsilon}{\sigma_0}\right)^2$. TAR of 4:1 implies $\frac{\epsilon}{\sigma_0} = \frac{1}{4}$. Thus, we get:

$$\frac{\Delta\sigma}{\sigma_0} \approx \frac{1}{2}\left(\frac{1}{4}\right)^2 = \frac{1}{32}$$

Furthermore, this implies that $\sigma = \sigma_0 + \Delta\sigma = \frac{33}{32}\sigma_0 = \frac{33}{32} \times 0.277 = 0.286$. And considering τ

to be fixed at the same value as that in Scenario 1: $\tau = 0.152$. By plugging these values of σ and τ into ... Equation 33 and ... Equation 34, we get:

$$\alpha^*(\sigma,\tau) = 1 - \Phi\left(2 \times 0.286 \times \ln\left(1/\sqrt{0.152}\right) + (1/2 \times 0.286)\right) = 0.011.$$

$$\beta^*(\sigma,\tau) = \Phi\left(2 \times 0.286 \times \ln\left(1/\sqrt{0.152}\right) - (1/2 \times 0.286)\right) = 0.113.$$

Plugging the values of $\frac{\Delta\sigma}{\sigma_0}$, σ_0, and τ into ... Equation 57, we get:

$$\Delta e = \left(\frac{1}{32}\right) \cdot \left[\left(\frac{0.152}{0.277}\right) \times \phi\left(2 \times 0.277 \times \ln\left(1/\sqrt{0.152}\right) - (1/2 \times 0.277)\right)\right] = 0.003 = 0.3\%$$

That is, the cost of mistakes made during the conformance testing process (as fraction of the output from production) would increase by about 0.3 percentage points if the firm were to forego high-quality, 'primary' calibrations. Assuming $\wp_A c_A$ is roughly equal to the aggregate turnover of £82.3 billion for the UK manufacturing firms that work with the NMS labs, we can infer that these organisations will incur an additional cost of $\Delta\mathcal{E} = \wp_A c_A \times \Delta e = £82.3 \times 0.003 \text{ billion} \approx £0.074 \text{ billion}$ or $£74 \text{ million}$ if they forego primary calibrations provided by top-tier labs like the NPL. Furthermore, using the ratio of GVA to total turnover as 0.3, we can say that GVA (potential benefits) from using high-quality, primary calibrations for the UK manufacturing organisations that work with the NMS labs would approximately be $£22.4 \text{ million } (\approx 0.3 \times £74 \text{ million})$. This is about a third of the GVA obtained in Section 6.1.3. The intuition behind this result is that in a mature industry with advanced measurement processes, the "outside options" (i.e., secondary calibrations) available to a firm that does not want to use calibrations from top-tier labs are also fairly accurate. Thus, the additional value from using primary calibrations will be lower than that in an industry where the "outside options" are less accurate.

The main findings from the above stylized example can be summarized as follows: Consider a firm that uses primary calibrations provided by NPL or other top-tier calibration labs, which provides it access to a measurement process with the conventional levels of $\alpha = 0.01$ and $\beta = 0.10$. Moreover, assume that if the firm loses access to such high-quality calibrations, the overall uncertainty associated with its measurement process increases by 3.13% (the calculation is $\Delta\sigma/\sigma_0 = 1/32 \approx 3.13\%$). This increase in measurement uncertainty results in:

- $\sim 10.7\%$ increase in the likelihood of making a type-1 error, and $\sim 13\%$ increase in the likelihood of making a type-2 error.
- $\sim 0.3\%$ of the value from production would be lost due to the extra costs incurred from making more mistakes during the conformance testing process, where these extra costs are a direct consequence of relying on a less accurate measurement process.

9.2 ANNEX 2 – NUMERICAL EXAMPLE WITH $\alpha = 0.1$; $\beta = 0.3$

Here, we explore what happens when we start with higher values of α and β as compared to Section 6.1.3. For simplicity, we can think of this example as representing a less mature industry producing bespoke batches of products where the measurement process has not advanced a lot. That is, there is a high degree of product differentiation that softens the competition over cost and price.

Numerical Example ($\boldsymbol{\alpha = 0.1}$ **and** $\boldsymbol{\beta = 0.3}$):

Baseline Scenario (Firm uses primary calibrations provided by NPL or other top-tier calibration labs) – In this scenario, we assume that the firm uses primary calibrations for the measurement process in conformance testing, and the process has $\alpha = 0.1$ and $\beta =$

0.3. The α and β here are higher than the conventional levels that we assumed in Section 6.1.3. Plugging these values of α and β into ... Equation 9 and ... Equation 27, we get:

$$\sigma = \sigma_0 = \frac{1}{\Phi^{-1}(0.90) - \Phi^{-1}(0.30)} = 0.554$$

$$\tau = \exp\left\{\frac{1}{2}[\Phi^{-1}(0.10)]^2 - \frac{1}{2}[\Phi^{-1}(0.99)]^2\right\} = 0.505.$$

Next, we can plug the above value of τ into ... Equation 46 and ... Equation 47, and combine them with Assumption 10 (i.e., $c_B/c_A = 2$) to obtain:

$$\mathcal{p}_A = \frac{\left(\frac{c_B}{c_A}\right)}{\tau + \frac{c_B}{c_A}} = \frac{2}{0.505 + 2} = 0.798$$

$$\mathcal{p}_B = \frac{\tau}{\tau + \frac{c_B}{c_A}} = \frac{0.505}{0.505 + 2} = 0.202$$

Finally, we can solve for the scrap rate ($\mathcal{s}$) using ... Equation 42 and rebate rate ($\mathcal{r}$) using ... Equation 43:

$$\mathcal{s} = \mathcal{p}_A \cdot \alpha + \mathcal{p}_B \cdot (1 - \beta)$$
$$= 0.798 \times 0.1 + 0.202 \times (1 - 0.3) = 0.221$$

$$\mathcal{r} = \mathcal{p}_B \cdot \beta = 0.202 \times 0.3 = 0.061$$

Counterfactual Scenario (Firm uses secondary calibrations provided by lower-tier calibration labs) – The increase in uncertainty over the baseline, σ_0, is given by ... Equation 52: $\frac{\Delta\sigma}{\sigma_0} \approx \frac{1}{2}\left(\frac{\epsilon}{\sigma_0}\right)^2$. TAR of 4:1 implies $\frac{\epsilon}{\sigma_0} = \frac{1}{4}$. Thus, we get:

$$\frac{\Delta\sigma}{\sigma_0} \approx \frac{1}{2}\left(\frac{1}{4}\right)^2 = \frac{1}{32}$$

Furthermore, this implies that $\sigma = \sigma_0 + \Delta\sigma = \frac{33}{32}\sigma_0 = \frac{33}{32} \times 0.554 = 0.571$. And considering τ to be fixed at the same value as that in Scenario 1: $\tau = 0.505$. By plugging these values of σ and τ into ... Equation 33 and ... Equation 34, we get:

$$\alpha^*(\sigma, \tau) = 1 - \Phi\left(2 \times 0.571 \times \ln\left(1/\sqrt{0.505}\right) + (1/2 \times 0.571)\right) = 0.103.$$

$$\beta^*(\sigma, \tau) = \Phi\left(2 \times 0.571 \times \ln\left(1/\sqrt{0.505}\right) - (1/2 \times 0.571)\right) = 0.314.$$

Plugging the values of $\frac{\Delta\sigma}{\sigma_0}$, σ_0, and τ into ... Equation 57, we get:

$$\Delta e = \left(\frac{1}{32}\right) \cdot \left[\left(\frac{0.505}{0.554}\right) \times \phi\left(2 \times 0.554 \times \ln\left(1/\sqrt{0.505}\right) - (1/2 \times 0.554)\right)\right] = 0.0099 \approx 1\%$$

That is, the cost of mistakes made during the conformance testing process (as fraction of the output from production) would increase by about 1 percentage point if the firm were to forego high-quality, 'primary' calibrations. Assuming $\mathcal{p}_A c_A$ is roughly equal to the aggregate turnover of £82.3 billion for the UK manufacturing firms that work with the NMS labs, we can infer that these organisations will incur an additional cost of $\Delta\mathcal{E} = \mathcal{p}_A c_A \times \Delta e = £82.3 \times 0.0099 \text{ billion} \approx £0.8147 \text{ billion}$ or $£814.7 \text{ million}$ if they forego primary calibrations provided by top-tier labs like the NPL. Furthermore, using the ratio of GVA to total turnover as 0.3, we can say that GVA (potential benefits) from using high-quality, primary calibrations for the UK manufacturing organisations that work with the NMS labs would approximately be $£244.4 \text{ million} (\approx 0.3 \times £814.7 \text{ million})$. This is higher than the GVA obtained in Section 6.1.3. The intuition behind this result is that in a less mature industry that focuses on bespoke batches of products, the "outside options" (i.e., secondary calibrations) available to a firm that does not want to use calibrations from top-tier labs have a higher measurement uncertainty. Thus, the additional value from using primary calibrations will be higher than that in an industry where the "outside options" are slightly more accurate.

The main findings from the above stylized example can be summarized as follows: Consider a firm that uses primary calibrations provided by NPL or other top-tier calibration labs, which provides it access to a measurement process with the conventional levels of $\alpha = 0.10$ and $\beta = 0.3$. Moreover, assume that if the firm loses access to such high-quality calibrations, the overall uncertainty associated with its measurement process increases by 3.13% (the calculation is $\Delta\sigma/\sigma_0 = 1/32 \approx 3.13\%$). This increase in measurement uncertainty results in:

- $\sim 2.8\%$ increase in the likelihood of making a type-1 error, and $\sim 4.6\%$ increase in the likelihood of making a type-2 error.
- $\sim 1\%$ of the value from production would be lost due to the extra costs incurred from making more mistakes during the conformance testing process, where these extra costs are a direct consequence of relying on a less accurate measurement process.