NPL
National Physical Laboratory

NPL REPORT IEA 9

REVIEW OF THE EVALUATION ACTIVITIES FOR THE NMS

MIKE KING, EUGENIO RENEDO, DANNI CROUCHER

MARCH 2022

Review of Evaluation Activities for the NMS

Mike King, Eugenio Renedo, Danni Croucher
Analysis and Evaluation, Strategy Directorate

ABSTRACT

This document outlines the current state of development of the evaluation activities for the NMS. The goal of the evidence produced by these evaluation activities is to verify that the NMS is generating outcomes as anticipated, and the extent to which its impacts constitute value of money. As the NMS is ultimately funded by UK taxpayers, the costs and benefits need to be assessed from the perspective of the country as a whole, rather than just the customers of the NMS laboratories. Therefore, the evidence generated consists both of analyses that account for the direct impact on NMS users, as well as the indirect benefits that reach the rest of society.

ISSN 2633-4194

https://doi.org/10.47120/npl.IEA9

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

Approved on behalf of NPLML by
Fiona Auty, Head of Government Relations.

EXECUTIVE SUMMARY

This document outlines the current state of development of the evaluation activities for the NMS. The goal of the evidence produced by these evaluation activities is to verify that the NMS is generating outcomes as anticipated, and the extent to which its impacts constitute value of money. As the NMS is ultimately funded by UK taxpayers, the costs and benefits need to be assessed from the perspective of the country as a whole, rather than just the customers of the NMS laboratories. Therefore, the evidence generated consists both of analyses that account for the direct impact on NMS users, as well as the indirect benefits that reach the rest of society.

All the evidence summarised in this document is in line with the NMS Evaluation Framework, which sets up the general principles on designing and implementing evaluation activities. These principles align with the guidance provided by the HM Treasury's Magenta Book, as well as the overarching evaluation strategy of the Department for Business, Energy and Industrial Strategy (BEIS).

The top findings of this review are:

- The quantitative evidence that estimates the benefits of innovation and R&D activities is robust and extensive.
- Qualitative evidence needs to be expanded to reinforce the narrative of 'why' and 'how' the benefits of the NMS occur. For this reason, NPL's analysis and evaluation team plans to give special importance to this type of evaluation in the remainder of 2020 and throughout 2021.
- A fundamental piece of work that needs to be developed in the near future is the analysis of the 'fan-out' of traceability. The fan-out is a fundamental element of NMS activity that is responsible for a significant fraction of its impact on the UK's economy and society. The NMS underpins the UK's measurement technical infrastructure by supplying services to commercial laboratories – mainly calibration services, certified reference materials and proficiency testing schemes. The commercial sector then goes on to pass this information to their own customers, diffusing measurement accuracy through a chain of linked interactions. These benefits need to be modelled and estimated.

NPL has a team of professional analysts, who conduct and lead evaluation activities to support the popper functioning of the NMS programme. Moreover, NPL's analysts work closely with economists at BEIS and also submit their evidence for peer review by external academic experts who provide quality assurance as members of BEIS's evaluator network. Core pieces of econometric evidence have been rated as a four on the five-point Maryland Scale which BEIS uses to rate the robustness of a piece of evidence. Lastly, some particularly good examples of NPL's recent evaluation work are:

- 'Estimating the Price Elasticity of Demand for NPL's Services' (pending peer review)
- 'The Impact of the National Measurement System Funding on the Uptake of NPL's Services' (pending peer review)
- 'Estimating the Impact of Improved Dose Control on Clinical Outcomes in Radiotherapy'

The rest of the document is organised as follows. Section 1 provides context about the NMS and the NMS Evaluation Framework. Section 2 lists and describes all the evaluation activities carried out to date. Section 3 explains how these pieces of evidence should not be considered separately, but rather as interconnected parts of a system of evidence that provides a rationale and an estimate of the benefits generated by the NMS. Section 4 identifies gaps in the evidence portfolio and highlights the main pieces of work that need to be undertaken until March 2021. Lastly, section 5 concludes.

CONTENTS

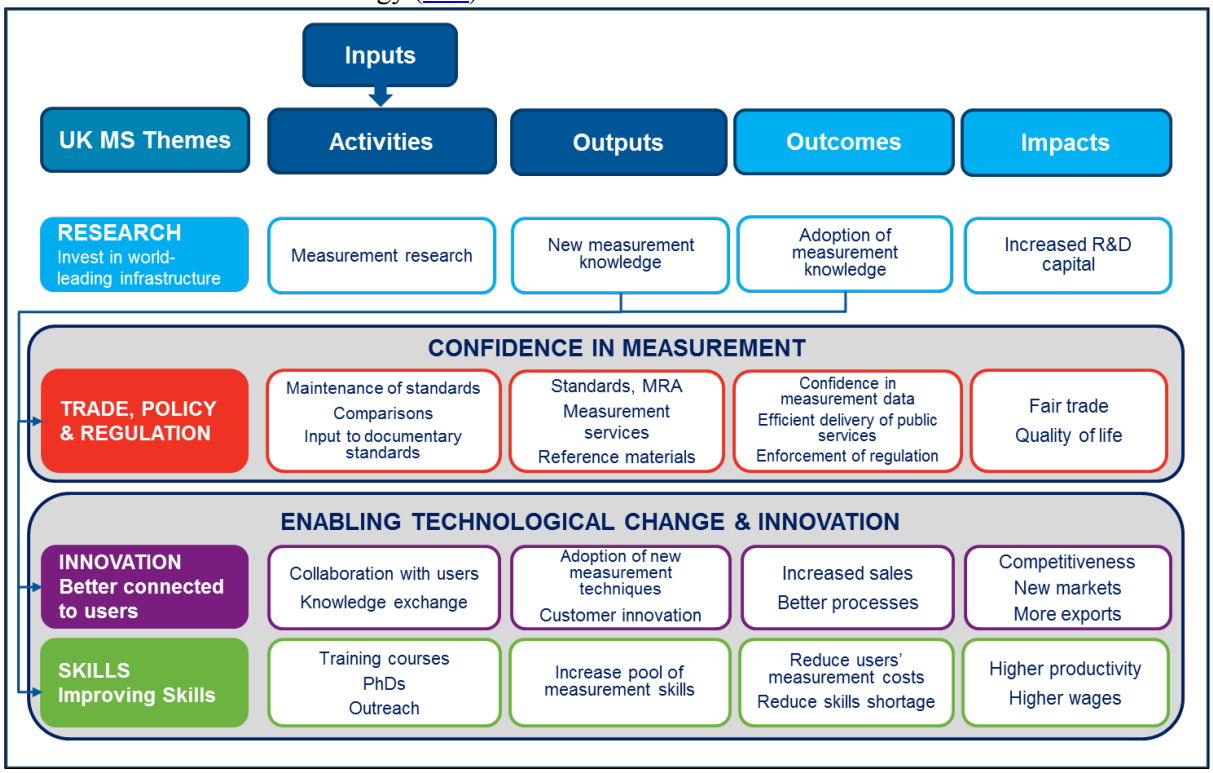
1 INTRODUCTION

1.1 ABOUT THE NMS

The NMS is the technical and organisation infrastructure which ensures a consistent and internationally recognised basis for measurement in the UK. It consists of the National Measurement Institute (NPL), and a number of laboratories which are designated to perform specific measurement functions, known as Designated Institutes (DIs); NML at LGC and NEL, among others. The NMS has two central objectives:

- To enable individuals and organisations in the UK to make measurements competently and accurately and to demonstrate the validity of such measurement
- To coordinate the UK's measurement system with what of other countries.

The following logic model summarises how the NMS creates value for the UK, based on themes in the current UK Measurement Strategy (link)[1]:



Note there is an arrow connecting the first theme to the other three. This is because value creation by the NMS starts by conducting fundamental research and performing international measurement comparisons that generate articles in peer-reviewed scientific journals. This enables the development of cutting-edge measurement capabilities that support the creation of primary standards and state-of-the-art instrumentation. This expertise is then used to deliver calibration, testing, and training services to private businesses, hospitals, and universities. In addition, the NMS works closely with Innovate UK to offer grant-funded collaborative R&D projects which involve many firms and research organisations.

1 At the time of publication, the current UK Measurement Strategy (2017) is being updated and is in the final stages before publication. Nonetheless, it is anticipated that the four key themes discussed here will largely remain consistent.

1.2 THE NMS EVALUATION FRAMEWORK

The NMS labs carry out evaluation activities to analyse the performance and impact of many aspects of the system. Specifically, the objectives of these evaluation activities are to:

- determine the causal effects of a programme – that is, which of the observable effects are attributable to the programme (i.e. the 'what'), and
- describe the purpose of a programme and the mechanisms through which the intended effect was achieved (i.e. the 'why' and the 'how').

The NMS Evaluation Framework (2017-2021) provides the over-arching approach for the evaluation of the NMS by outlining the primary evaluation methodology and the specific data collection and analysis methods that will be used to provide evaluation evidence. The Evaluation Framework builds on the guidance set by the UK Measurement Strategy (2017-2021).
The Evaluation Framework states that the impact generated by the NMS programmes should be assessed through a variety of quantitative and qualitative methods – since each method has its own strengths and limitations, by using a suite of approaches the overall assessment of impact is balanced. Quantitative techniques primarily deal with the first of the objectives (the 'what'), and therefore allow the assessment of value-for-money, whereas qualitative analyses mainly address the second of the evaluation framework's objectives (the 'why' and the 'how'), hence providing a narrative that is often missing from quantitative studies.
In terms of quantitative analysis, the framework proposes the use of:

- **Econometrics:** These studies apply statistical methods to economic data to give empirical content to economic relationships. Econometrics provide objective statistical evidence for the causal effect of the NMS laboratories on innovation and economic outcomes. The application of statistical techniques requires large samples sizes. This means that the analysis needs to be conducted at a high level, preferably by aggregating across multiple laboratories. Therefore, the main strength of econometric methods is that they are highly empirical, which gives confidence in the results obtained. Their main limitation though, is that they do not state how benefits occur. In this sense, econometrics lacks a narrative on how the impact is generated; it just estimates these positive effects.
- **Modelling:** Models are used to estimate impact in those cases where the length and complexity of the impact pathway means that it is unrealistic to pursue an empirical evaluation. That is, modelling is useful in cases where the distance between an NMS laboratory and the ultimate beneficiary is large and there is a wide range of factors that would confound an empirical evaluation. Models are based around archetypes within our archive of impact case studies and parameter estimates are typically obtained by interviews with beneficiaries. The main limitation of this approach is that it is based on strong assumptions; therefore, these must be valid for the method to be reliable. In effect, the fact that the method is not empirical is compensated by the use of well-established incontrovertible laws.
- **Indicators:** Data from the laboratories' management information systems are used to construct measures for the activities and outputs of the NMS laboratories. Indicators have been designed to meet two criteria: to align with the logic model and be based on reliable data. These indicators are reported annually to provide evidence that the NMS laboratories are conducting activities and producing outputs in line with the logic model and are therefore on the route to delivering the expected outcomes and impact. The main limitation to the use of indicators is that these are only suitable for effects and outputs that can be counted. Also, indicators cannot be used to measure the counterfactual (i.e. what would have happened without intervention). In that sense indicators must be seen as means or inputs for other analyses.
- **Surveys:** Surveys are used to obtain (1) information about the experiences and perceptions of NMS users (e.g. net promoter scores), and (2) data that is not possible to obtain from publicly available sources (e.g. what is the proportion of revenue annually spent on measurement). Such surveys enable the collection of standardised information which is then analysed to get new insight about the NMS. Although there are clear disadvantages to surveys like potential

responder positive bias and the self-reported nature of any impact figure asked to interviewees, surveys are often the only way to get certain information, and thus, remain a fundamental tool in the evaluation kit.

The NMS evaluation framework also recommends qualitative methods such as:

- **Structured fieldwork**: Qualitative evaluation spans from primary research, whereby individuals are interviewed in a structured way to generate new insights. The transcripts from these interactions are analysed for themes using psychological models, such as that for behaviour change or decision making. Fieldwork can be:
    - *Structured interviews:* In-depth telephone or 1-2-1 interviews with key stakeholders.
    - *Focus groups:* group interviews with a number of relevant stakeholders together.
    - *Forums:* group interviews and content generation using an online platform.

The tools used in qualitative fieldwork can be applied to projects to understand a wide range of complex issues, and the depth and nuances within these.

- **Case studies**: These are detailed examinations of a specific instance of support - project or programme. Case studies use data from interviews to show the impact that is partly or wholly attributable to an output from an NMS laboratory to understand how this impact occurred and the role played by the laboratory. They require looking at why something done in the past led to changes and impacts that are visible today. Furthermore, a case study needs to identify the elements outside the NMS lab that influence the effectiveness of its output, as well as whether the output had any unanticipated consequences. Lastly, a case study would also be helpful to find organisations or groups that did not engage directly with the NMS laboratory, but were nonetheless impacted.

- **Peer-reviewed benchmarking:** Used to assess the quality of the scientific outputs of the NMS. Scientific outputs are reviewed by external experts.

The NMS evaluation framework is designed around the following themes:

1. RESEARCH: Ensure a national approach to measurement capabilities that drives effective investment in the UK measurement laboratories in partnership with the UK science base and business to **deliver a world-leading measurement infrastructure.**
2. TRADE AND REGULATION: Champion measurement across Government to **ensure good policy, standards, and regulation.**
3. INNOVATION: Support advances across the public and private sectors through getting **better connected to end-users**; to be done by increasing awareness, access, and uptake of best practice.
4. SKILLS: **Advance the UK's capability in measurement** by upskilling measurement intensive sectors to accelerate the diffusion of new technology and fully exploit the benefits of a high-tech economy.
5. Bring together the diverse communities to understand the new capabilities in data science and develop a framework to deliver confidence in the **intelligent and effective use of data** based on measurement traceability and uncertainty analysis.

However, for simplicity, the fifth theme is subsumed into the first. That is, investment in data science is seen as a special case of investing in the UK's measurement infrastructure.
All the evaluation activities relate to one or more of these themes.

2 OVERVIEW OF THE EVIDENCE PORTFOLIO

This section contains an exhaustive summary of all the evaluation activities carried out to date or those that are currently in progress. It is divided into five sub-sections (Research; Trade and Regulation; Innovation; Skills; All) that indicate the main theme of the work: the first four sub-sections are consistent with the themes one to four defined in the Evaluation Framework and the fifth includes works that combine all themes.
Each piece of work is uses either quantitative methods (such as econometrics, surveys, etc.) or qualitative techniques (e.g. structured interviews). At the top of the description of each evaluation activity, the following information is given:

- the methodology used,
- the status of development,
- and the reach of the evaluation activity – i.e. whether the piece of work assesses impact for a specific NMS lab or for the whole NMS.

2.1 RESEARCH

2.1.1 International Science Review 2016

Method: Qualitative; review.
Status: Published (link).
Reach: NMS-wide.
The International Science Review Board was charged with the following four tasks:

- Benchmark NMS Science output against appropriate peer groups to identify areas of particular relative strength or weakness.
- Make assessments of the wider economic and social impact of NMS-supported science, measurement services and consultancy.
- Assess the appropriateness of balance in the portfolio between core metrology functions, stakeholder and 'responsive-mode' activities and long-range research.
- Assess individual NMS laboratories and collate contribution to the UK NMS as a whole.

Based primarily on written evidence prepared by senior scientists from the six Science Divisions at NPL and the teams at NML at LGC and NEL, assessments were completed by a group of 32 Expert Reviewers. The Expert Reviewers were drawn from appropriate scientific peers in NMIs, industry, academia and other research and technology organisations. Half of the Expert Reviewers were from outside the UK, ensuring that this exercise represented truly international conclusions.
The Review Board drew a collective conclusion about the overall performance and international standing of all areas of scientific activity within the NMS (including measurement services, knowledge transfer, training and consultancy) by considering the assessments from the Expert Reviewers, other documentary evidence, interviews with senior scientists and managers, personal knowledge, laboratory visits and technical expertise.
The latest International Science Review of the NMS (conducted in 2016) concluded that:

- All the NMS Science areas perform at an internationally competitive level.
- More than half of the NMS Science areas are considered to be internationally leading.

The Review Board also reported several observations and recommendations which the NMS labs have followed to improve their international leadership, get better connected to end users and further improve economic benefit.
The next International Science Review is about to be kicked off and should report by March 2021.

2.1.2 Business Case for the National Timing Centre (NTC) Programme

Method: Quantitative.
Status: Business case complete, but not made public for legal reasons. However, a full report on the economic impact of the NTC on collaborating companies has been published (link).
Reach: NPL.
This business case sought £36.27m over 5 years to deliver a new network of atomic clocks, as well as to provide training and innovation support to businesses. The NTC programme was established to address the UK's over-reliance on Global Navigation Satellite Systems (GNSS), and to put the UK in a world-leading position regarding timing and synchronisation technologies. Its objectives are:

1. Deliver a resilient UK national time infrastructure through the building and linking of a new atomic clock network that is distributed geographically in secure locations.
2. Support innovation among UK companies through funding projects in partnership with Innovate UK, based on a successful NPL and Innovate UK partnership model.
3. Respond to the shortage of people with timing and synchronisation skills through specialist apprenticeships and post graduate training opportunities.

To deliver these objectives, six options were considered based on their costs and risks. However, the preferred option was the establishment of a new resilient time network with scalable business access hosted by the National Timing Centre (NTC). The following points were presented in the economic case to help justify this decision:

- Looking at solar storms, as just one instance of GNSS interference, a fully quantitative analysis shows an expected gross benefit of £108.9m and a NPV of £34.4m over 20 years. This provides a lower bound for the net benefit of investing in the programme.
- It is predicted that, by collaborating with UK companies, support from the NTC will help to generate 800 new jobs after 2-3 years; where these jobs would have a wage premium of around £8,500.
- Using Green Book assumptions for the rate of time preference, the present value of the programme's total cost (discounted economic cost) is £74.6m; where £33.7m is for the public component and the remaining £40.9m comes from the private sector.

2.1.3 Business Case for the Quantum Test and Evaluation (QT&E) Programme

Method: Quantitative.
Status: Business case complete, but not made public for legal reasons.
Reach: NPL.
This business case sought £12.5m for 2020/2021 and, subject to Spending Review, a further £26.5m over the following three years to construct and deliver the necessary test and evaluation facilities that will accelerate the commercialisation of new technologies developed in the UK's National Quantum Technologies Programme (NQTP), by enabling industry to properly demonstrate the performance of new products. The core objectives of the (QT&E) programme are:

1. Deliver QT&E facilities to accelerate the commercialisation of new products based on quantum technologies.
2. Develop and disseminate best practice in measurement and testing to support the widespread and consistent use of quantum test and evaluation facilities across the UK by suitably qualified people.
3. Represent the UK as a leading authority within international quantum forums; particularly with regards to the international agreement of documentary standards and regulations as they emerge.
4. Provide opportunities for UK companies to make use of QT&E facilities and knowledge.
5. Respond to the shortage of people with quantum test and evaluation skills through the provision of training opportunities and specialist apprenticeships enabled by QT&E facilities, thereby supporting the broader quantum skills initiative.

The business case considered seven options against nine critical success factors. The preferred option is directly aligned with the NQTP and will aim to increase the speed of return on investment made through quantum projects in ISCF and SPF programmes, as well as the return on other industrial and government investments in quantum technologies, by enabling new products to be brought to market more quickly and more reliably.

The QT&E programme will enable NPL to build a range of knowledge assets which companies would have access to through measurement services, consultancies and so on, for a price. Access to these knowledge assets lowers barriers to innovation for these companies and enables them to undertake more innovative activities. These realized R&D projects create private knowledge assets for the firm which create greater economic opportunity for the firm and therefore leads to the firm's growth. This expansion is accompanied by an increase in employment, alongside investments in capital equipment and research activities.

The value for money (VfM) assessment of the programme provides an expected gross benefit of £256.7m and an NPV of £157.3m over 15 years, yielding an NPV-DEL ratio of 4.03:1. This means, for every £1 invested by the government, there is a net return of £4 worth to the supported firms.[2]

The QT&E programme will also make a tangible contribution to:

- achieving the 2.4% target for R&D investment, by supporting the commercialisation of quantum technologies to encourage industrial and venture capital investments;
- supporting BEIS' *Place* agenda through installing several new facilities co-located in emerging quantum technology clusters;
- delivering against the UK Measurement Strategy which identifies specific emerging and disruptive technologies as requiring measurement infrastructure development;
- addressing the widely recognised skills gap impeding the growth of the emerging quantum industry by supporting apprentices, PhD positions, and wider CPD training programmes for industry;
- ensuring UK interests are reflected in emerging standards and regulations for quantum technology.

2.2 TRADE AND REGULATION

2.2.1 Estimating the Price Elasticity of Demand for NPL's Services

Method: Quantitative; econometrics.
Status: Published (link).
Reach: NPL.
Over half of NPL's sales are made overseas, meaning that a significant fraction of NPL's income is subject to international trade forces. This paper applies standard econometric methods to a country-level dataset to estimate the price elasticity of demand of NPL's services by exploring how variation in inflation and exchange rates between the UK and the countries affects the volume of sales that NPL makes to overseas customers.
The price elasticity of demand for a service is the percentage change in the volume of sales divided by the percentage change in the price of those services. This findings in this report are based on the following economic concepts:

1. Variation in inflation and exchange rates alters the purchasing power parity; and this exogenous source of variation can be used to estimate the price elasticity of demand.

2 Calculated using some of the top findings from both Frontier Economics and Belmana studies in accordance with the Green Book's established economic guidelines for assessing the economic impact of public programmes.

2. The price elasticity of demand provides information on customers' willingness-to-pay for NPL services.

Moreover, since NPL is a publicly funded organisation, a reliable estimate of the price elasticity of demand constitutes the cornerstone for further work aimed to estimate the value-for-money of services delivered to the supported organisations.

The analysis outlined in this paper makes use of a country-level panel dataset involving more than 100 countries over a period of 17 years (2001-2017) to estimate the price elasticity of demand for NPL's measurement services. The data used consists of NPL internal administrative information, as well as data coming from reliable and widely used publicly available databases such as the World Bank's World Development Indicators database, or the CEPII database.

The elasticity was found to be 1.24; which means that the sales of measurement services would increase by around 12.4% for a 10% drop in the price of services[3]. The analysis finds, with 95% confidence, that the price elasticity of demand for NPL's services is greater than one, meaning that these are elastic goods for which the quantity demanded will vary more than proportionally if the price changes.

Using a formula that combines NPL's income with the price elasticity of demand, the paper also explains how the benefits generated by NPL can be estimated. It tentatively suggests that £1m of funding leads to £6m of benefits; but this estimate is based on a series of assumptions that will need to be evidenced in due course[4].

2.2.2 Creatinine Impact Case Study

Method: Quantitative.
Status: Published (link).
Reach: NML at LGC.
Chronic kidney disease (CKD), a long-term condition often linked to age, high blood pressure or diabetes, affects between 7% and 10 % of the population. If is not detected early enough it can lead to an increased risk of heart attacks or, in extreme cases, even total kidney failure. In the UK over 1 % of the entire NHS budget (~£1.5bn) is spent on late-stage CKD treatments such as dialysis or kidney transplants. Although early-stage CKD exhibits minimal symptoms it can be detected by monitoring levels of the biomarker creatinine in the blood in those patients most likely to be at risk. Ensuring correct early diagnosis through accurate and reliable creatinine measurements is crucial to reducing healthcare spend for an ageing society.

In a 2003 survey run by the College of American Pathologists (CAP), over 60 % of reference material producers were shown to demonstrate significant routine biases (10 to 35 %) for creatinine measurements. At this time, 1.5m people in the UK had been assessed as having CKD. Although early diagnosis is important to prevent longer term health failure, misdiagnosis due to excessive false positive test results creates unnecessary patient distress and NHS spend. Assuming a general false positive rate of 15 % (the lower end of the bias observed) in 60 % of cases, this indicates 135,000 people may have been incorrectly diagnosed. Beyond the patient impact, these figures suggest that the NHS was spending at least £34m per year on unnecessary treatment due to lack of standardisation in this one area.

This is a strong case study for how the chemical metrology performed at NML has underpinned the standardisation of a crucial clinical biomarker for chronic kidney disease, helping save the NHS many millions of pounds each year on misdiagnosis. It uses the results of two independent studies evaluating the bias of routine measurements over a 10-year period and follows the indirect impacts of the work

3 More generally, a price elasticity greater than one means that these are elastic products for which the quantity demanded will vary more than proportionally if the price changes.

4 One of these core assumptions is developed by the evaluation activity *2.5.1*.

performed at the NML at LGC to show the benefits implementing standardisation in the clinic can have. The study finds that, over the last 15 years, this work has helped to save the NHS £290m/year and prevented over £1bn being spent globally on misdiagnosis annually.

The causal narrative explaining the route to impact is clear and very well evidenced, and the study sheds a welcome light more generally on how standardisation of diagnostic and therapeutic markers generates impacts. Nonetheless, quantification of the cost savings could benefit from further work to strengthen the mathematical and statistical modelling. This is something that analysts from NML and NPL intend to develop in the future, which, once complete, could be applied to a multitude of similar case studies where false positive/negative test results are concerned.

2.2.3 Using Stock Market Data to Measure the Effect of EASA's Amendments to Regulations Governing Components for the Aerospace Sector

Method: Quantitative.
Status: Pending publication (expected to be published imminently).
Reach: NMS-wide.
Regulations and Standards are necessary for developing products that are "safe to use", whether the product is a building, an aircraft, or a pressure vessel. NPL makes a significant contribution to standards and regulations. For example, In the field of polymer composites/fibre reinforced plastics NPL's work has contributed to 17 standards in the last two decades. To measure the effect of regulations and standards on the economy, we used the aerospace industry as a use case.

The aerospace industry has an organisation, the European Aviation Safety Agency (EASA), dedicated to providing standards and guidelines for companies operating in the sectors. The standards cover aspects of aviation from unmanned aircraft systems to the use of composite materials. Furthermore, EASA makes this information available in an open-source database. Using an Event Study methodology, stock market data and EASA's open-source data on the dates when amendments to standards took place, we calculated the effect of 64 amendments to standards regulating the composition of aerospace parts (called CS-25) on the stock prices of companies that supply aerospace parts made from composite materials.
Currently, there are two ways to prove a composite component is fit for purpose:

1. The equivalence-based approach entails proving that new components are equivalent to traditional material, such as steel.
2. The performance-based approach involves proving that the new materials can perform to the required standards in operational conditions.

A study by Southampton University argues that the performance-based approach has several advantages over the equivalence-based approach. Firstly, it provides more objective proof of performance because it relies on codified standards and guidelines rather than subjective equivalence-based assessments that may be based on loosely or poorly defined requirements; and, secondly, it increases innovation capacity as regulations can be made to adjust with technological advancements.

An economic theory called "the efficient market hypothesis" says that the share price in the stock market reflects an agent's expectations about the future performance of specific companies and sectors. Investors react to announcements by regulators and other relevant bodies concerning changes to specifications or regulations. That is, regulatory shocks are new information for investors to act upon and according to the efficient market hypothesis, the stock market can be used to measure the commercial effect of regulatory changes. This methodology has a long history of being used to assess regulatory changes in economics and finance.

Event study methodology looks at the effect of amendments to regulations for large aeroplane manufacturing on the value of European companies that use composite materials to make parts for aeroplanes. This methodology (which has a long history of being used to assess regulatory changes in economics and finance) is implemented following these steps:

1. Generate an index for the daily share-price of eight companies that make aerospace parts using composite materials and then finding the correlation between changes in the value of this index and changes in the value of an index for the whole stock market (e.g. FTSE).
2. Define the dates of the European Aviation Safety Agency's (EASA's) announcements about changes to regulation as 'events'.
3. Use the index generated in step 1 to test for abnormal returns during an 'event' by subtracting the expected return (as predicted by the market model) from the observed return.

Based on this methodology, this work estimates the contribution of these regulatory events to the growth in the share-price index after the 2009 financial crash. Considering that, since the start of 2010, the average annual return for shares in these composites companies was about 14.3%, it is estimated that, on average, the cumulative abnormal returns using a seven day 'event window' (three days either side of an event) is about 0.55%. Since 2010 there have been about 4.1 events per year; which suggests the average annual return would have been about 11.9% without the boost from EASA's amendments to regulation. Using the expected price earnings ratio for all 8 companies in the study, it is estimated that, on average, each amendment was valued at around £8m by investors in all markets. For smaller, UK companies the estimated value of each amendment was £1m. In other words, the regulatory changes are associated with about 20% of the growth in the market capitalisation of the eight composite companies.

2.2.4 The Impact of the Government Chemist Advisory Role

Method: Qualitative; survey.
Status: Underway.
Reach: NML at LGC.
This Government Chemist (GC) role is funded by BEIS and holds a statutory function as a referee analyst under several acts of Parliament which focus on public protection, value for money and consumer choice. Regulatory areas where GC advice may be important include the quality of food, animal feed, pesticides, medicines and chemicals. The GC is responsible for helping to settle technical disputes between food businesses and local authorities, carrying out research to give evidence-based advice, advising government and industry and disseminating findings with industry, academics and other organisations. The survey will offer an insight in to how the Government Chemist is helping UK industry, government and public analysts ensure the safety and authenticity of feed and food. It will evaluate the impact the GC role provides to different stakeholder groups, including industry, trade organisations, regulators and regulation enforcement organisations.

2.2.5 Estimating the Impact of Improved Dose Control on Clinical Outcomes in Radiotherapy

Method: Quantitative.
Status: Published (link).
Reach: NPL.
This report provides evidence of the importance of delivering accurate doses of radiation to cancer patients – one of the channels through which the National Physical Laboratory (NPL) contributes positively to cancer treatment in the UK.

At the most basic level, radiotherapy can be thought to have the effect of killing human body cells. This includes both cancer cells and healthy ones. Hence, the success of radiotherapy depends critically on the certainty around the treatment. If the dose is not appropriate or it is not focused on the target volume, the cancer may continue to grow, and/or healthy tissue could be damaged producing life-changing side effects. In the clinical situation, there are several sources of uncertainty. These include the initial calibration and the maintenance of the radiation machines, as well as uncertainties around organ content, hydration levels and patient positioning during the radiation treatment.

The analysis focuses on one of the most widespread types of cancer among the UK population: prostate cancer. Although the goal of this work is to illustrate the positive impact of this reduction in the uncertainty of the delivered dose, modelling the benefits that this has had for all cancer patients

treated with radiation is extremely complex, because there are many variables that can influence each type of cancer. Hence, this work only analyses the effect on one very specific type of cancer.
In order to show the impact of a reduction in the uncertainty of the dose administered to prostate cancer patients, two commonly used radiobiological models are combined to approximate the increase in the probability of both controlling the tumour and not causing significant side effects. It applies well-established radiobiological models to show the positive effect of radiotherapy dose audit for better accuracy in radiation dose delivery. Its objective is to estimate the number of prostate cancer patients who would have received an incorrect dose because of the radiation machine not being in specification had the dosimetry programme by NPL not been carried out.

Over a 20-year period, the joint efforts of NPL and medical physicists, working in the radiotherapy community, led to a gradual increase in the accuracy of the delivered doses. As a result, this study estimated that an additional 15 prostate cancer patients are successfully treated each year. The back-of-the-envelope calculation below gives a sense of the potential size of the overall impact across all types of cancer on the basis of the strong assumption that all types of cancer are equally sensitive to the accuracy of the delivered dose.

Each year, roughly 180,000 patients in the UK receive radiotherapy as treatment for any form of cancer.[5] Approximately 2 out of 15 of these patients are receiving treatment for prostate cancer.[6] Given that it has been estimated, to a high degree of rigor, that NPL's radiation dosimetry programme, in collaboration with medical physicists, saves the lives of 15 prostate cancer patients per year, from these results we can extrapolate that the programme might save the lives of around 110 cancer patients in total.

2.3 INNOVATION

2.3.1 The Impact of Public Support for Innovation on Firm Outcomes[7]

Method: Quantitative; econometrics.
Status: Peer reviewed and published as BEIS document (link).
Reach: NMS-wide.
In 2015 Frontier Economics was commissioned by the Department for Business, Energy and Industrial Strategy (BEIS) and Innovate UK to study the economic impact of public sector support for private sector innovation. The analysis they produced focuses on policies providing direct support for business innovation through grants, loans, advice and access to specialist services operated by Innovate UK the NMS. The study examines how receipt of support affects firm-level economic performance measured in terms of survival, employment and turnover up to 5 years after support[8].
The analysis found that:

- After two years, firms supported by the NMS labs have around 20 additional employees (~15% growth in employment) compared to a matched control group of unsupported firms. This analysis suggests that support from the NMS labs boosts employment among high-tech firms, with up to 4,000 additional jobs over 2-4 years.
- Users were more likely to still be in business after one year than comparable firms that did not receive support. Among the matched control firms, the survival rates are around 95% after one

5 Cancer Research UK statistics report that there are roughly 360,000 new cases of cancer each year. Data from The Society of Radiographers suggests that 50% of cancer patients would benefit from receiving radiotherapy. Hence, 50% of 360,000 is 180,000.

6 It is estimated in the report that approximately 24,000 prostate patients per year receive radiotherapy.

7 Frontier Economics, 2016. The Impact of Public Support for Innovation on Firm Outcomes.

8 The analysis is based on a statistical technique called *propensity score matching*, which enables treated firms to be compared to a matched control group. The matching process is performed assessing the similarity between supported and unsupported businesses based on key characteristics.

year, and 85% after three years. In contrast, survival is a virtual certainty for treated firms. This survival effect is more pronounced among younger firms (2-5 years old) than older firms.

2.3.2 Public Support and Innovation Outcomes – An Econometric Analysis

Method: Quantitative; econometrics.
Status: Pending publication (ongoing discussions with Innovate UK on how to publish the study. In any case the NMS will in due course publish a summary version of it). This study has been reviewed by academic experts that are members of BEIS's evaluator network.
Reach: NMS-wide.
A forthcoming[9] piece of work by Belmana Consulting seeks to answer questions emerging from the earlier work by Frontier Economics. A key part of the new work was to look at a wider set of impacts, including pay levels in the supported businesses and the quality of the patents. From 2009 to 2015, the NMS supplied services to around 500 firms per year, and around 350 of those were regularly supported. The regularly supported firms are typically large, high-tech, and established decades ago. The support to these regularly supported firms is analysed in this study. However, it was possible to find comparators for only half of those 350 firms. For these businesses the analysis finds that:

- Around 23,000 jobs-years[10] of employment have been created by 175 businesses who sought regular support from the NMS labs from the period 2009-2017, an increase of 5.5% in economic activity. Businesses supported infrequently by the NMS labs showed a decrease in economic opportunity over the timeframe, i.e. there was a decline in their number of employees.
- Around 18,800 job years (80 %) are additional as this growth is not seen in comparable businesses. This suggests that the NMS's support helps create the conditions for these job-switchers to make better use of their capabilities, thus increasing their productivity. This translates into a wage premium. Employees moving into NMS supported businesses receive a £50 a week wage premium (£2,600 per annum), with the premium remaining significant after controlling for age and occupation of earners.
- Survival rates for regularly supported businesses are high, with only 4% of businesses closing, compared to over 35% of the wider BSD[11], and 12% of a similar comparison group. Even after controlling for business size, the NMS supported businesses are much more likely to survive.
- A second analysis has then been performed to understand the impact on job years as survival has been impacted by support. The additional impact is 4,400 jobs-years (this can be added to the previous 18,800 jobs-years).
- In general businesses supported by the NMS labs generate patents with higher spillovers, with those spillovers increasing after 2001 (knowledge spillovers are measured by patents citations). This spillovers analysis has been conducted using linked patent data prepared by Dr. Ralf Martin who has developed measures for the importance of spillovers using citations based on Google's pagerank algorithm. Note that in this analysis the importance of a patent is not just the number of citations it gets, but also the citations of any citing patent.

At first glance it might seem that the results of this study do not agree with those found in the previous analysis carried out by Frontier Economics. It should be borne in mind that this second analysis builds on the lessons learned from the latter, and therefore, makes use of more sophisticated techniques that

9 The study was submitted to BEIS' central evaluation unit for peer review, and was awarded a 4 out of 5 in the Maryland Scientific Methods Scale that rates the quality of impact evaluations.

10 A job-year is one job that lasts for a year. Two jobs-years could be one job that last two years, or two jobs that last for a year.

11 The Business Structure Database (BSD) contains a small number of variables for almost all business organisations in the UK. The BSD is derived primarily from the Inter-Departmental Business Register (IDBR), which is a live register of data collected by HM Revenue and Customs via VAT and Pay As You Earn (PAYE) records.

avoid double counting. Also, this new analysis covers a different period of time and uses job-years as opposed to jobs as a measure of impact.

2.3.3 Analysis of the Measurement for Innovators (MfI) Programme

Method: Quantitative; econometrics.
Status: Pending publication (expected to be published imminently). This study has gone through independent peer review and was awarded a 5 out of 5 on the Maryland Scientific Methods Scale.
Reach: NMS-wide.
The analysis of the MFI programme provides evidence of the key role the NMS labs play in the collaborative R&D innovation process in the UK; patent activity is a product of invention which can be thought as the first stage of innovation. The main objective of the MFI programme was to provide expanded access to the measurement knowledge, skills and facilities of the NMS laboratories. The programme was delivered via three products: consultancies, secondments, and joint industry projects. Due to data limitations the analysis only focuses on the MFI consultancies. The MFI consultancies were used to provide measurement advice to organisations to support the development of process, products or services. The consultancies were targeted at UK SMEs and trade associations with majority SME membership; however, some large companies benefited from this product.
The study finds that the MFI programme had a positive impact on companies who completed the programmes relative to a group of similar comparators:

- On average, the patenting activity of the supported firms were 11% higher than the counterfactuals during the year of support. There were no significant differences between the supported firms and the counterfactual before and after treatment.
- Based on the findings on the effect of the MFI programme on inventions, the estimated cost-benefit ratio of the programme was over 2:5[12] (this a lower bound estimate).
- After receiving treatment, the treated businesses grew more quickly than the comparators. This difference in growth rate is estimated to be about 5% annually.

Furthermore, the study shows that there was a period of stagnation for the counterfactual group 4 years after the year of treatment. However, firms who received support had a persistent growth trend even 4 years after treatment – i.e. the supported firms do not experience the same level of stagnation as the counterfactual. Growth rates this persistent show that the MFI programme did not only increase the inventive output of the firm, but it may also have changed the innovative culture of the supported firms. This means that the supported firms have adopted better ways of acquiring, assimilating and exploiting knowledge for innovating as a result of the MFI programme.
This work can also be used to extrapolate some of the impact of the Analysis for Innovators (A4I) programme, which commenced in 2017. The A4I programme has similar objectives in terms of increasing access to NMS facilities and expertise for industry to address productivity issues and improve R&D.

2.3.4 NMS Customer Survey

Method: Quantitative; survey.
Status: Published (link).
Reach: NMS-wide.
The NMS customer survey is a core requirement of the NMS evaluation plan. It is a large-scale quantitative survey designed to capture headline figures for the impact of three of the core NMS labs (NPL, NML at LGC and NEL) on their customers, understand customer's use of the labs, as well as to capture their satisfaction. The sample of users includes those working in the private sector, of which a number of these are calibration laboratories, and users and partners in the healthcare sector and academia.

12 A £1,000 of spending generates £2,500 in benefits.

The top findings of the most recent NMS Customer Survey (performed in 2019) are:

- The NMS supported around 4,500 business sites in the UK during that time. These business sites employed almost one million people and had an aggregate turnover of £188 billion.
- The supported businesses spent around 5.4% of their turnover on measurement; which equates to around £10 billion of spending per year.
- 12.5% of UK's manufacturing activity is supported by the NMS, based on employment at business sites working with an NMS lab.
- 80% of supported businesses introduced a new or improved product or process. Furthermore, 19% of those that made a change attributed the innovation to the support they received from the NMS labs.
- Users believe that, without the support of the NMS labs, their annual sales of new and improved products would have decreased by at least £539 million.

The customer survey goes beyond measuring impact to understand key metrics, such as private companies' spending on measurement and their economic importance. However, it does not currently extend to enable the undertaking of a full user segmentation, nor does it include qualitative analysis to enable the exploration of how or why a finding occurred.

2.3.5 Achieving the 2.4% GDP Target: The Role of Measurement in Increasing Investment in R&D and Innovation

Method: Quantitative (review of literature).
Status: Published (link)
Reach: NMS-wide.
Long-term economic growth is driven by technological change. Although in the short run any economy can grow by increasing the capital stock per worker, this mechanism soon runs out. This is because as the capital stock grows larger, it takes more and more investment to produce an additional unit of output, and at some point, the economy only invests to keep up with effective depreciation. Therefore, investment in R&D is key because it drives the technological change needed for productivity growth, and thus, long-term economic growth.

Currently, the UK is below the OECD (Organisation for Economic Cooperation and Development) average in terms of the proportion of GDP that it invests in domestic R&D, which will negatively affect its comparative economic growth in the future. (The UK invests 1.7% of its GDP in R&D, whereas, Germany invests 2.9% of its GDP in R&D.) To remedy this the government has committed to raising the R&D spending, so that by 2027, the UK spends 2.4% of GDP on R&D. Although, some additional funding for public R&D is definitely required, this ambitious target can only be met (and sustained) by encouraging business to invest more in R&D by the government supporting interventions aimed at increasing the private returns from R&D.

This paper makes the case that although achieving the 2.4% target mainly depends on raising the private returns from R&D so that businesses invest more in R&D, the UK is close to the limit of what can be achieved through current funding mechanisms. It argues that policy makers should consider other forms of intervention aimed at fundamentally altering the productivity of R&D without just resorting to further subsidies.

Specifically, this paper makes the case for further investment in the nation's measurement infrastructure to ensure good access to research tools, techniques, and standards. Over the last 20 years, investment in such infra-technologies has not kept pace with investment in R&D. There is reason to think that underinvestment has contributed to a decline in the productivity of R&D in areas where relevant infra-technologies either don't exist or need extending into new domains. Radically updating the stock of infra-technologies requires additional public funding but would increase the efficiency by which private R&D converts generic technologies into proprietary technologies. Infra-technologies for the life sciences are used as an example but there are a host of

new scientific fields in need of support. Interventions to fundamentally increase the productivity of R&D will necessarily bring business investments in R&D closer the 2.4% target by increasing the proportion of the potential R&D projects that are privately profitable.

2.3.6 A Net Present Value Assessment of NPL's Support to UK Businesses[13]

Method: Qualitative.
Status: Published (link)
Reach: NPL.
NPL achieves its goal of generating welfare for the UK through two avenues:

1. **Direct benefits**. NPL's staff help users to either develop new products, design and implement new processes, or prevent their skills and knowledge base to become obsolete. New products enable the user to increase their market power and command a price premium, and new processes enhance their productivity and competitiveness.
2. **Indirect benefits**. Knowledge generated by NPL's activities spills over to non-supported businesses.

This document quantifies the overall impact of NPL's support to UK innovative[14] businesses[15], direct and indirect. In this paper, some of the previously described pieces of evidence are used to populate parameters in a model that connects NMS funding to employment and productivity growth among firms that received support from NPL. In particular, this business case model is a combination of:

- An adaptation of BEIS' Net Present Value (NPV) Innovation Model (see below for further details on how BEIS' model has been tailored to NPL).
- The employment growth effect of NPL's support found in "The Impact of Public Support on Firm Outcomes" (2.3.1). The results of the study show that companies who use NPL services have higher survival rates and experience an average employment growth of 20 employees within 2 to 4 years.
- The productivity gains effect (observable by a wage premium) obtained in "Public Support and Innovation Outcomes – An Econometric Analysis" (2.3.2). This study finds that employees moving into NPL supported businesses receive a £2,600 annual wage premium (which remains significant after controlling for age and occupation of earners).
- The result coming from a meta-analysis by Frontier Economics performed in 201416 that finds that existing literature estimates private rates of return to R&D investments of around 20-30%. Social returns, based on spill-over benefits from R&D, would then be typically 2 to 3 times larger than private returns. Therefore, our analysis estimates the indirect benefits by doubling the direct benefits, while applying a two-year time lag to allow time for the diffusion of new knowledge to take place.

Although this document has not been submitted to BEIS' peer-review system, the NPV assessment resulting from the analysis has been considered robust enough for publication. This is because the study combines two previous works which had already been peer reviewed, a very well-established econometric work, and the conceptual framework provided by BEIS for innovation evaluation. Thus,

13 The model is based on the Green Book's established economic guidelines for assessing the economic impact of public programmes. The Net Present Value (NPV) is calculated as the difference between the present value of streams of costs and benefits, which are obtained using a discount rate to convert all future deflated figures to values that can be compared across years.

14 Note that R&D and innovation activities inherently spill over to other firms beyond the firm that actually carries out the innovation. This occurs largely through the movement of workers involved in the innovative process to other competing firms or firms in similar/complementary industries. Thus, in economic terms, it can be said that innovation is subject to free riding. Therefore, this study not only considers the direct benefits to NPL users, but also the indirect benefits to non-users.

15 Although this analysis applies exclusively to NPL, the results could be generalised to the rest of NMS labs if we assume that they behave similarly to NPL's scientific groups. Given the much bigger size of NPL compared to other NMS labs, this assumption is not unreasonable.

16 Frontier Economics. (2014). Rates of return to investment in science and innovation. London.

there is no room for modelling choices and the parameters used have already been approved by BEIS. The only departure from BEIS' innovation model is that no second-round effects are included in our NPV model.
In this business case model, the NPV of NPL's support is assessed over a 15-year horizon. Given the current level of NMS funding received, the NPV-to-DEL ratio found was found to be 7:1.

2.3.7 Analysis for Innovators (A4I) Case Studies

Method: Qualitative.
Status: Published (link).
Reach: NMS-wide.
A4I case studies are created to showcase the impact the programme has on specific companies. Each case study is particular to a company, and details the technical issue they faced and how it affected their business, then progresses to explore how the A4I partner (such as NPL or NML at LGC) was able to offer a solution to the problem. The case studies demonstrate how projects develop and relationships are made between institutions. They cumulate in demonstrating the direct impact a project has made, on the technical output, and the wider outcomes, such as wider benefits to the business or sector, as well as indirect impacts such as improved knowledge and KT. Video case studies follow this format to present project impact in an engaging and visual way.

2.3.8 The Highlights and Effectiveness of Knowledge Transfer Activity Within the Flow Programme

Method: Qualitative.
Status: Report finished, but not made public. For dissemination to PEG members and BEIS only.
This report discusses and assess, through a series of case studies, the highlights and effectiveness of the KT activity undertaken as part of the NMS Flow Programme. The Flow Programme leads the NMS activity associated with flow measurement and gears metrology. It invests in the UK's Flow Measurement Standards based at the National Engineering Laboratory (NEL) and the Gears Metrology Standards located at Newcastle University. The programme funds research and development concerning future flow and gear metrology requirements and it contributes to the UK's international efforts on flow measurement standardisation and knowledge transfer.

This report makes the case that these KT activities, including Focus Groups, Webinars, and Lunch and Learn sessions, are found valuable by industry and academia alike. In addition, to provide even greater evidence of impact and value in the future, it identifies a number of additional activities/areas for ongoing focus.

2.3.9 NEL Case Studies

Method: Qualitative.
Status: Underway. Most of the individual case studies have already been published (see links below).
NEL produces company-level case studies that highlight the many ways in which the laboratory helps companies develop a wide range of measurement technologies. The following are prominent examples of such case studies:

- *Flow Tests of Two Thermowell Designs* (link): The UK's Designated Institute for Flow Measurement, an NMS laboratory based at NEL, was used to help WIKA, one of Europe's leading manufacturers of measurement equipment, to assess the performance of an innovative thermowell design. As a result, a new product is now in the market with respected, independent test data available to demonstrate its capabilities.
- *Optimising Check Valve Design* (link): Abacus Valves, a UK-based engineering company which designs and manufactures industry-approved check valves, required independent pressure drop flow testing on one of their prototype designs. NEL was enlisted to determine the pressure drop characteristics of a prototype check valve. This would validate the design calculations and provide independent benchmark data for future Computational Fluid Dynamics

(CFD) simulations. Abacus Valves estimate that the development of this new product range contributes an estimated 25% of the company's annual turnover.

- *Minimising Financial Exposure by Reducing Errors in Flow Measurement* (link): A global oil and gas operator enlisted NEL to investigate a damaged flow meter, present in a shared pipeline, which was found during a routine audit. The aim was to ascertain the flow measurement error and the resultant financial exposure resulting from the installation. Through testing and modelling studies carried out at NEL, it was found that the meter was overreading by up to 10% at some flow rates, which was equivalent to $1 million exposure annually. As a result, the operator was able to claim compensation from the pipeline owner.
- *Knowledge Transfer Activities are First Step in Building Relationships with Innovative UK Businesses* (pending publication; expected to be published imminently): NMS experts based at the UK's Designated Institute for Flow and Density measurement delivered a Lunch and Learn session to DASHBOARD, an innovative company based in the UK. DASHBOARD is developing cutting-edge data fusion solutions for industry. This was the beginning of an ongoing relationship with the NMS for DASHBOARD. Since winning a funding award under Innovate UK's Analysis for Innovators (A4I) Round 3 Competition, DASHBOARD's 'Limpet' measurement product has undergone development testing at the UK's National Flow Measurement Facilities at NEL. The success of the testing, according to DASHBOARD, has "super-charged" their development programme and has recently resulted in them winning another funding award to further develop their technology under A4I Round 5.

2.4 SKILLS

2.4.1 National Timing Centre Skills Survey

Method: Quantitative; survey.
Status: Underway.
Reach: NPL.
The National Timing Centre programme (NTC) is funded by the UKRI's Strategic Priorities Fund and will provide the UK's first nationally distributed time infrastructure to improve the resilience of national time and frequency distribution, paving the way for trusted time and frequency services across the country. Although the NTC is a separately funded programme, it develops off the back of the NMS knowledge and facilities. One of the main objectives of the NTC is to address the nationwide shortage of specialist skills in Time, Frequency and Synchronisation (TFS). To that end, NPL has conducted a survey assessing TFS skill requirements across a broad range of sectors. This survey was sent over the 24th March – 6th April 2020 to around 600 contacts who represented around 450 organisations. So far, around 90 responses have been submitted, with a completion rate of approximately 60%. The analysis of the results of this survey are still underway.

2.4.2 The Postgraduate Institute for Measurement Science. Five Year Review.

Method: Qualitative; review.
Status: Published (link)
Reach: NPL.
The Postgraduate Institute for Measurement Science (PGI) is a pioneering initiative to equipping the UK with an exceptionally skilled workforce in measurement science. It was founded in October 2015 by a strategic collaboration between NPL and the University of Strathclyde and the University of Surrey. This document was prepared by the PGI team within NPL to account for the progress that has occurred throughout these years.

The review finds that, by growing the pipeline of doctoral researchers to the current cohort size of 206 (a growth of 82% in 5 years), the PGI is having a considerable impact on NMS research and its application in a range of sectors across industry and academia. Of the 140 graduates who have completed their research since 2015, the largest proportion have moved into industrial roles (43%),

indicating their industry readiness, with a number making a rapid move into leadership positions within 3 years of graduation. PGI's alumni are also established in academic and other research roles (29%) at prestigious organisations worldwide, with a further 19% becoming full-time researchers at NPL.

Furthermore, the report shows that the growth in the PGI's industrial engagement to 41 studentships with 31 industrial partners has had a significant impact on the way PhD research project are performed. This greater involvement of industrial partners, along with the development of a programme of metrology and professional training that is aligned to business requirements and tailored to student researcher needs, has provided the fundamental elements to create a pipeline of highly skilled and industry-ready postdoctoral researchers.

2.5 ALL

2.5.1 The Impact of the National Measurement System Funding on the Uptake of NPL's Services

Method: Quantitative; econometrics.
Status: Published (link)
Reach: NPL.
This econometric study analyses how NMS funding received by NPL affects the number of UK-based firms it can support.
NPL supports UK businesses by offering specialised services. This allows them to develop new products or processes. New products enable the company to increase its market power and command a price premium. On the other hand, new processes enhance productivity and competitiveness. This benefit carries over to society in the form of greater private profits of firms and better wages for their workers, as well as more value for consumers. These are direct economic benefits of public funding of NPL.

Although NPL's portfolio is quite varied, roughly speaking it can be considered to offer one unique good: the time and the expertise of highly trained scientists and engineers. In order to deliver high-accuracy measurement services, these staff must maintain the measurement capabilities required to do so. This involve carrying out a variety of research based tasks including conducting international key comparisons, participating in proficiency testing schemes, maintaining UKAS accreditation for calibration and testing services, running audits, contributing to standards and protocols, and performing research that generates articles in peer-reviewed scientific journals. Hence, the activities of NPL involve a substantial cost in preserving a knowledge stock that depreciates over time in order to meet the requirements of its users.

The economic rationale for the existence of a publicly funded organisation like NPL is that measurement R&D is subject to market failure. Indeed, the private investment needed to generate innovative measurement capabilities, will always be below the socially optimal level. This occurs because the benefits that measurement R&D generates will always spill over to firms who did not contribute, creating a strong incentive to free-ride. This paper demonstrates that more public funds allow NPL to hire more scientific staff, which leads to more UK companies receiving support
The data used consists of administrative information that NPL routinely collects to monitor the composition of the workforce, the evolution of income, as well as the number of paying customers of each of NPL's scientific groups which are specialised in different areas of metrology. The analysis finds that:

- Providing NPL with an additional £1 million per annum of resource funding, enables it to directly support around 9 more UK companies. These additional companies will either become NPL paying customers or will engage in collaborative R&D projects. This means an average cost of approximately £110,000 per company.
- Giving an additional £1 million per annum to a group leads to a 12% increase in its scientific workforce. Currently, the average group has 48 scientists and engineers, therefore each additional million would allow the group to hire 5 or 6 scientists.

The results of this study can be combined with other pieces of analysis to find the additional economic impact of NPL receiving extra public funding. To that end, it is necessary to know: (1) the effect of extra funding on the number of firms supported, and (2) the effect of such support on the growth of these companies. The scope of this report encompasses only the first of these two elements, whereas, the impact on supported companies is addressed by the aforementioned forthcoming study by Belmana Consulting.

2.5.2 Estimating a Demand Function for NPL's Competitively Won Work

Method: Quantitative; economic modelling.
Status: Report finished. Not made public.
Reach: NPL.
At first glance, one might think that NPL could double its impact by just doubling the number of staff employed on competitively won jobs. However, this is not the case because the value generated by each additional scientist employed on competitively won jobs is less than the one before – there are diminishing returns to employing more people. Thus, we expect a negative relationship between the price of competitively won work and the volume of such work, which has been steadily increasing over the recent years. This work uses NPL's financial information to model such expected decrease in the unit-price. This model has two main uses:

- In the short term, the model makes it possible to estimate the direct economic benefit to paying customers.
- In the longer term, the model will enable NPL to quantify the extent to which new strategies and business developments increase demand for NPL's services.

This analysis can also be used to estimate the price elasticity of demand for NPL services. The price elasticity found is very similar to the one obtained in the study "Estimating the Price Elasticity of Demand for NPL's Services" (see 2.2.1). This fact provides certainty in both results as they have been obtained using different techniques and using different data.

2.5.3 Quantifying Measurement Activity in the UK

Method: Quantitative; economic modelling.
Status: Published (link)
Reach: NMS-wide.
This report presents a novel method for estimating the value of measurement activity in the UK. One way to assess the value of something is to observe a buyer's willingness to pay for it. Using simple economic theory, and some reasonable assumptions combined with national statistics from credible sources, this work identifies and quantifies the value of three primary sources of measurement activity within an economy: the use of measurement tools; the manufacture of measurement tools; and metrological innovation. Each of these is then broken down into their core components, resulting in a total of four key elements of measurement activity: labour input; investment; intermediate demand; and research and development. The analysis finds that:

- **Labour input**
  - In 2017, 6.3% of total employment in the UK was comprised of occupations that involve taking measurements and the total compensation offered to employees in these measurement related occupations was £58.3bn which accounts for 2.8% of UK GDP. Of these employees, 162,000 were specifically employed to provide calibration and testing services.
  - Taking a closer look at the concentration of measurement related occupations within industries, 14.3% of all employees in the Research & Development industry and 7.9% of all employees in the Defence industry hold occupations that may require calibration or testing practices.

- Between 2007 and 2017, the number of employees in measurement related occupations rose by 13% –4 percentage points more than overall employment over the same period –to a total of 2 million. This was mostly driven by an uplift in occupations that may or may not require calibration or testing practices.

- **Investment**
  - Investment in instrumentation has increased annually and this amounted to just over 3.6bnin 2017, accounting for 4.1% of investment in productive and tangible assets and 7.2% of investment in all machinery equipment.
  - In 2017, we estimate that the gross capital stock of instrumentation was valued at £36.8bn, 0.5% of the UK's total gross capital stock, 57% of which was accounted for by instruments for Health and Life Sciences.

- **Intermediate demand**
  - Intermediate demand for instrumentation has increased, in real terms, from £6.8bn to £9bn between 2008 and 2017, while the ratio of each type of instrument produced has remained relatively constant.
  - From 2008 to 2017, the UK has become increasingly dependent on imports of instrumentation for intermediate input. This peaked at 60% of total intermediate demand in 2017.

- **Research and development**
  - In 2017, the total expenditure on measurement-related R&D was just over £2bn, 87% of which were performed by private businesses.
  - Over the last decade,7.5% of all UK originated granted patents relate to a measurement technology –almost 16k in total.
  - A similar analysis of journal articles reveals that measurement research accounts for 2.4% of all papers published in the last 10 years –almost 21k in total. Academic measurement R&D has been declining over the decade. In the UK, the proportion of total papers published that were on measurement related topics was 0.3 percentage points higher in 2007 than it was in 2017.

2.5.4 NMS Indicators

Method: Quantitative; indicators.
Status: Report finished (updated every year). Not made public.
Reach: NMS-wide.
The set of ten NMS indicators mapped across the four main themes of the UK Measurement Strategy were developed to assist with monitoring the progress of the NMS in achieving its goals – i.e. whether NMS activities are generating outputs that are on course to deliver the intended impacts. In 2018 most indicators show the NMS met or exceeded the strategic aims to maintain or grow activity. The main findings were:

- **Research:**
  - *NML at LGC:* A more strategic approach to stakeholder engagement has been implemented with targeted stakeholder activities to identify gaps and address priority areas. This has also been addressed through the development of focused stakeholder partnership programmes (e.g. A4I, NHS Knowledge Transfer Partnerships), and joint academic appointments through strategic regional partnerships.
  - *NPL:* Overall there was an increase in collaborators and reports, but particularly notable in the Digital sector, through Quantum and Graphene projects.

- **Trade and regulation:**
    - *NML at LGC:* Measurement services were maintained at 2017 levels but there was a significant leverage R&D increase.
    - *NPL:* Measurement services are generally maintained; the slight reduction showing the normal shift in demand for services. Measurement services income has continued to grow on previous years.
    - *NEL:* Income per measurement service offering has increased by more than 25% year-on-year, reflecting a strategy of sustained focus on the most industrially relevant of its flow and fluid density measurement services.
- **Innovation:**
    - *NML at LGC:* An increase in leveraged income from collaborative R&D and consultancy is observed. Also, the initial success of programmes such as A4I and its subsequent growth, and the close engagement with ISCF has led to a significant increase in leveraged funding
    - *NPL:* There has been an increased amount of engagement with SMEs; the increase in 2018 reflects the beginning of higher growth in new areas from 2016 onwards. This includes the Analysis for Innovators programme (reflected in both 2017 and 2018 metrics).
    - *NEL:* leveraged income peaked in 2017, reflecting its relative success in the 2016 EMPIR Call in Energy, which is NEL's focal sector (EMPIR Energy calls follow a 3-year cycle). Business collaborator numbers saw an increase in 2017 and 2018 for the same cyclical reason.
- **Skills:**
    - *NML at LGC:* The Community for Analytical Measurements, an industry-led skills programme co-founded by the NML at LGC was launched this year. It is dedicated to supply world-class analytical measurement science (AMS) training, and will extend the reach of our training programme.
    - *NPL:* 2018 saw a drop in the level of web-based training at NPL, which is likely to have been due to a variety of factors including a shift in demand towards more niche training, the introduction of a new software platform for e-learning with different inbuilt reporting mechanisms, and changes which mean that new NPL starters are no longer automatically enrolled on e-learning courses. NPL's Training department is expecting an increased uptake in e-learning courses during 2019, following a marketing push and changes in Business Development.
    - *NEL:* The lab has observed a steady shift in demand away from face-face toward web-based training in the last three years, facilitated by an increased range of web-based offerings. Though 2018 numbers overall were down on 2017, reflecting the general trend, web-based training numbers were still substantially greater than in 2016.

2.5.5 Case Study Database Meta-Analysis

Method: Quantitative (review of case studies).
Status: Publication imminent.
Reach: NPL.
The last reallocation of NMS funding occurred during 2018/19, leading to all theme budgets being reduced by 15% and the savings allocated to cross-theme projects. This move was driven by expert opinion of leading stakeholders and the judgement of NPL's management, backed up by 100 years of accrued experience running the NMS programme.

Future formulations are likely to benefit from complementing such expert opinion with greater reference to published evidence (quantitative or qualitative). This would increase the transparency of the process, further demonstrating NPL's commitment to the principles of HMT's Green Book. Moreover, previous approaches to formulation involved ranking individual projects using a common set of criteria, but this "one-size-fits-all" method can be improved upon by a more systematic consideration of where such projects are with respect to the 'technological lifecycle'.[17] The ability to locate projects with regards to their position in the lifecycle would allow for longer-term, dynamic considerations in the construction of the portfolio of funded projects.

This analysis integrates findings from multiple case studies[18]. The goal is to work out how NPL's sectors operate and how they generate impact and benefits for the UK. Since this work is contributing to build a narrative for NPL's impact, it is strongly linked to the NMS logic model. This work will in time be used as a first step to develop specific models to each of the sectors.

Case studies capture activities within each of the groups/sectors better than any other tool. In order to analyse these case studies, NPL has a database with 365 case studies that categorises all of them using a multitude of factors. Statistical analysis was conducted to gain insight from this database. By sector, the top findings were:

- **Advanced manufacturing (AM):** AM's main theme is "Trade & Regulation", with measurement services and consultancy being pushed by this sector. It is the most commercially focused out of the four sectors, as it has the lowest levels of R&D-based case studies and the high levels of consultancy and training services, and is also the most mature. This maturity is derived from the prevalence of direct impact for this sector's activities. AM's breadth of impact is the lowest of all the sectors, as many partners in this sector tend to be SMEs. This low breadth of impact is also backed up by the low level of spillover effects – as the work conducted with this sector tends to be commercial in nature, all the benefits tend to be for the firms that work with NPL. Lastly, the main impact mechanism by which Advanced Manufacturing delivers benefits to the economy is through process innovation, looking to cut costs and maximise efficiency of production.
- **Energy and environment (E&E):** For Energy & Environment, the main theme for the sector is also "Trade and Regulation". This sector shows low levels of consultancy services, yet it has higher levels of measurement services – potentially indicating that the commercial relationship is built around calibration rather than working with customers. There seems to be a prevalence of contract R&D case studies, where NPL has been tasked with producing new knowledge to solve a customer's problem. Lastly, E&E seems to have a high breadth of impact; this is backed up by the high level of spillover effects. As much of the work conducted with this sector is targeted at protecting a public good, which is liable to being underfunded by the market, the benefits often spill over outside of the direct benefactors of the work. In fact, E&E is the only sector to have a majority of its case studies exhibiting spillover effects, likely due to the major implications the work conducted within this sector has on Climate Change. Overall, the activities of this sector exist within the sweet spot on the lifecycle model, exhibiting both current and future impact.
- **Life sciences and Health (L&H):** Like E&E, the L&H sector clearly produces both current and future impact. Despite having less than 50% of its case studies within the "Trade & Regulation" theme, L&H isn't significantly different to the overall NPL average (with 51% of case studies under "Trade & Regulation"). This suggests that most of the activities conducted with this sector are through aiding industry, in this case industry would be the healthcare sector, with the NHS making up the majority of that. Collaborative and in-house R&D account

17 A conceptual model developed by an NPL sector head. The model assesses general age of NPL's sectors by judging the position of each sector's outputs on a two-dimensional scale: current impact versus future impact.

18 Similar to the REF (Research Excellence Framework) system for assessing the quality of research in UK higher education institutions.

for more than 50% of all the case studies. This emphasis on innovation through collaboration ties into the high levels of case studies that involve collaborators rather than customers, with 56% of the case studies being alongside collaborators rather than customers. L&H's breadth of impact is comparable to E&E's. The breadth of impact is also backed up by the high level of spillover effects. As much of the work conducted with this sector is targeting at boosting healthcare, which has been proven to be underfunded by the market, the benefits from work with this sector often impact 3rd parties, outside of the direct benefactors of the work.

- **Digital and Quantum technologies (D&Q):** "Research" is the only logic model theme D&Q has that is significantly different than the NPL average, in this instance being significantly higher – a large proportion of their case studies are predominantly R&D-focused and it was the only sector with a significantly larger proportion of "no impact yet" case studies, exemplifying the potential for impact in the future with the work conducted in this sector. For D&Q the emphasis is on in-house R&D, research with no outside commercial influences, either working alone or with academic collaborators. The sector shows the highest breadth of impact out of the four sectors. With the importance of much of the research conducted within this sector for Industry 4.0, there is a clear widespread reach for the work conducted within this sector. Despite this, there are few spillover impacts. This is likely due to the youthful nature of the sector, with most of it's activities emphasis being on future impact.

2.5.6 A Lifecycle Model for NPL Scientific Groups

Method: Quantitative; statistical analysis.
Status: Underway.
Reach: NPL.
This study makes use of the NMS indicators to operationalise a model of NPL's output using a statistical technique called factor analysis. The goal of this analysis is to come up with an impact assessment in two dimensions, 'impact' and 'research'. The first dimension of analysis 'impact' measures how much income is brought in by each scientific group[19], whereas the second dimension evaluates the scientific production (i.e. papers in peer-reviewed journals) by group.

The analysis identifies the types of output that each scientific group produces, whether it be strategically important research or measurement services to industry and the public sector. In general, there tends to be a trade-off with these outputs, though there are several outliers, with some groups producing both kinds of output and some doing neither. Thus, a simple two-dimensional framework can be constructed to assess the level of research output and income generated by each group. It should be noted that assessing where the group lies on this framework does not assume anything about the performance of the group. Rather, their location on the model merely informs about the maturity level of the capabilities housed by the groups. Typically, when a new group is set up, it focuses on research to develop novel measurement capabilities to fulfil a new demand in the market. In time, once these capabilities are ready, they get to the market and generate income. Lastly, other commercial measurement laboratories catch up and the capability is commoditised – at this point in time NPL might decide to shift resources into other uses. Hence, the goal of this analysis is to produce valuable insight for decision-makers to make judgements on portfolio rebalancing and long-term strategic choices.

19 The analysis is done at scientific group level – NPL is organised into a number of scientific groups that specialise in different areas of metrology such as electrochemistry or advanced engineering materials.

3 AN INTERCONNECTED PORTFOLIO OF EVIDENCE

Some of the evaluation activities described in the preceding section should not be considered separately. On the contrary, when added together, some of these works allow us to obtain a neat estimate of the impact generated. If we think in terms of the logic model, the goal for any set of evidence should be to address how inputs (e.g. NMS funding) are converted into outputs (e.g. supported businesses) through a number of activities (e.g. supply of measurement services), and then how this outputs lead to outcomes (e.g. productivity gains) and ultimately to impact for the UK (e.g. gross value added to the economy). Obviously, it is not feasible to tackle this whole chain of events with the same piece of work (i.e. model). Conversely, a fragmented approach whereby each of these links is estimated separately, and then put together to yield an estimation of impact is much more sensible.

The quantitative evidence in the innovation theme tends to be mutually supportive in that way. The first link is the evaluation activity *2.5.1*. This work is basically an input-output analysis that makes use of incontrovertible economic relationships to estimate the number of supported businesses for each additional £1m of NMS funding. The results of this study can be then combined with analysis *2.3.2* to find the additional economic impact of NPL receiving extra NMS funding. *2.3.2* provides an estimation of the effect of the NMS's support on the growth of companies. Hence combining the two we can approximate the expected growth in supported industries given an uplift in NMS funding. Finally, the connections between quantitative and qualitative evidence should be mentioned. Although these two methods involve quite different techniques, and they provide answers to slightly different questions about the impact generated, both heavily influence each other. Qualitative evidence plays a fundamental role in hypothesis formation. Hypotheses are then empirically tested by quantitative evidence. In other words, qualitative evidence can make an essential contribution to justify modelling choices (and even identify variables in a model). At the moment, most of our statistical quantitative work has been driven by models either based in NPL's understanding on how the benefits occur, or relevant precedents in the economic literature. Although we run robust specification tests to statistically check that our models are viable, nothing can replace the insight that a beneficiary can provide to identify relevant factors that affect the outcome variable of a model (e.g. a senior calibration engineer in the metrology department of a large manufacturer that regularly buys NPL's services can prove fundamental to understand what factors motivates them to seek NPL's support and/or other calibration labs). For that reason, NPL's Analysis and Evaluation team is looking to incorporate at least one social researcher that enables this *structural modelling* approach by gathering qualitative evidence of impact that feeds through the quantitative side of the analyses.

4 WHAT'S NEXT?

4.1 AN UNBALANCED EVIDENCE PORTFOLIO

The following table summarises all the evidenced developed to date:

Theme	Quantitative	Qualitative
Research	2.1.2, 2.1.3	2.1.1
Trade and Regulation	2.2.1, 2.2.2, 2.2.3, 2.2.5	2.2.4
Innovation	2.3.1, 2.3.2, 2.3.3, 2.3.4, 2.3.5	2.3.6, 2.3.7, 2.3.8, 2.3.9
Skills	2.4.1	2.4.2
All themes	2.5.1, 2.5.2, 2.5.3, 2.5.4, 2.5.5, 2.5.6	

The first impression obtained from this table of outputs is that the evidence portfolio is not balanced. Two major elements stand out: (1) there is clearly a bias towards quantitative evaluation activities, and (2) the 'research; and 'skills' themes lack evidence so far.

The 'innovation' and 'trade and regulation' themes would naturally involve more evaluation activities because the NMS places more focus on them. However, the discrepancy between these themes and those of 'research' and 'skills' is too wide. The reason for this is the lack of social researchers in the analysis and evaluation team who have the abilities to undertake qualitative analysis. In general, the 'research' and 'skills' themes require the use of qualitative techniques because there are no large and exhaustive related datasets available. To remedy this deficiency, at least one new person will be recruited throughout 2020 to professionalise the qualitative side of research. Ideally, along with these capabilities this new social researcher will also have some previous NMS work experience.
The previous table also denotes another deficiency of the current evidence portfolio. The research theme remains largely unexplored. This is a major flaw that is easily discernible in the logic model (see graph in the introductory section). Indeed, the way in which the NMS creates value is through the development of measurement capabilities that are subsequently drawn on to supply measurement and R&D services to UK businesses. This is reflected in that arrow of the logical model that connects the research theme with the rest of the themes. Thus, this should be analysed to obtain a reasoned explanation of how the benefits of NMS are produced.

4.2 CONTRIBUTION ANALYSIS

NPL is committed to boost the qualitative side of its evidence portfolio. To that end, the analysis and evaluation team has envisaged a contribution analysis designed to holistically understand the extent of impact attributed to NPL in a range of areas. The overall aim of this work would be to deepen the understanding of the contribution of measurement, and NPL more specifically, to outcomes and impacts amongst businesses, workers, the public sector, consumers and society more generally. Crucially, the project will allow us to explore what would have happened without NPL's influence using a hypothetical counterfactual. Note that this would be an imaginative exercise (no statistical analysis involved) of thinking through what would have happened in an alternative history where NPL had not provided support.
This study will consist of deep dive analyses into four areas of activity. The selected areas will cover the four NPL sectors and the four logic model themes. In-depth structured interviews will be conducted with NPL staff to understand the history of activity NPL has undertaken in an area, how this developed, and the impact this has had over time. Direct users will be interviewed to explore how and why outcomes have occurred for them as a result of working with NPL, and the extent to which they attribute this to NPL. Lastly, indirect impact will be captured by following the chain and exploring the impact enjoyed by the customers and beneficiaries of direct users.
As of today, the way this contribution analysis is done is yet to be decided. The preferred option is to carry out a large-scale qualitative project in collaboration with an external consulting firm (who will be selected in a competitive tender process). With this approach, the project would be complete by March 2021. However, the current Covid-19 crisis has led to a procurement freeze within NPL that would delay the project excessively. In that case, an alternative in-house approach will be developed

leading to the timeline for project completion to probably be extended beyond March 2021 because of the lack of external people to conduct the fieldwork.

4.3 FAN-OUT ANALYSIS

The other great challenge for the evaluation of the NMS is the 'fan-out' effect of calibration. The fan-out is a fundamental element of NMS activity that is responsible for a significant fraction of its impact on the UK economy and society.
The NMS underpins the UK's measurement technical infrastructure by supplying services and reference materials to commercial calibration laboratories, who then go on to pass on this information to their own customers, diffusing measurement accuracy through a chain of linked interactions. In effect, the unique knowledge that the NMS labs have is sold to commercial labs who are then able to make use of it to calibrate the instruments of their customers or validate that their customer's products meet their desired specification, thus enabling them to signal high quality goods traceable to national standards. This 'fan-out' of national measurement standards extends the reach of the NMS to thousands of organisations across the UK.
Developing evaluation activities to assess the impact of the fan-out is an absolute must, thus being one of the main priorities for the Analysis and Evaluation team in the coming months. To that end, two major streams of work need to be carried out:

- Develop a comprehensive and economically justified model of the fan-out
- Gather data to populate that model in order to come up with a figure of impact.

The first task comprises a theoretical difficulty. Given that the services provided by the NMS laboratories constitute the first link of the traceability chain, does this mean that no economic benefit that occurs further down the line would take place without the NMS? In other words, could we say that the entire impact of the calibration chain is attributable to the NMS? This certainly lacks common sense; note that we would be stating that for instance NPL is responsible for the impact coming for large primes such as Rolls Royce. Hence, although the NMS labs do play a key role in enabling industry to make precise measurements and ensuring traceability to international standards, obviously they are not responsible for all the value added created by all the companies that perform any kind of measurement. Thus, the challenge for the theoretical model to be developed is to be able to capture the incremental impact coming from the NMS labs, given their initial position in the calibration chain. In order to get valuable insight on how the benefits of accurate measurement arise in end users (i.e. private businesses operating in many markets that require precise measurement), structured interviews with users would need to be conducted. Only through speaking to people like heads of in-house calibration departments in large manufacturing primes, metrology engineers in innovative SMEs or measurement technicians of all kinds, will it be feasible to fully understand the true impact generated by the NMS labs down the calibration chain.

On the other hand, the second task involves a practical difficulty. Assuming that an acceptable model of the impact generated by the fan out is in place, probably the parameters to be used in it would be mostly information not publicly available. Specifically, the model would need to be populated with information regarding things like average prices of commercial calibration labs' measurement services, the time it takes for an instrument to cease being in specification, yield loss rates, etc. To obtain this data, a thorough survey will have to be carried out. This survey should be intelligently designed to represent accurately all segments in the end-user layer.

5 CONCLUSION

The evaluation activities for the NMS have been progressing well. In particular, it is worth highlighting the robustness of the quantitative evidence that accounts for the benefits around innovation and R&D activities. Without a doubt, the positive impact enjoyed by direct users of the NMS labs is sufficiently quantified.

The Analysis and Evaluation team at NPL will continue to expand the breadth and depth of our knowledge of the NMS through multiple publications that are expected to be underway in the forthcoming year. These include: a synopsis of the sections of the Belmana report (2.3.2) that analyse the impact of NMS labs; an econometric study for the impact of NMS support on the rate at which companies generate patentable inventions; and a life-cycle model for how each of the four NMS sectors create different forms of impact, based on findings from the Case-Study Database Meta-Analysis (2.5.5).
That said, future evaluation activities of the NMS will focus on four key areas for improvement:

1. Updating the evaluation strategy to evolve in line with the changes to the UK Measurement Strategy.
2. Developing evidence on Equality, Diversity and Inclusion (ED&I).
3. Strengthening the stock of qualitative evidence.
4. Analysis of the fan-out.

Short-to-medium-term plans to address each of these key areas are discussed in turn below.

Updating the evaluation strategy to evolve in line the changes to the UK Measurement Strategy
The development of a new UK Measurement Strategy, expected to be published imminently, presents the opportunity to rethink our approach to evaluating the NMS, and to incorporate all the lessons learned during the current evaluation period. We anticipate that the updated strategy will focus on the four core activities around which the evidence in this report is organised (research, trade and regulation, innovation, and skills), as well as three challenge areas for the UK which would benefit from the support of the NMS labs (Health and Life Sciences, Energy and Environment, Data and Digital). Going forward, evidence for each of these areas will be developed with reference to our internal SMART (Specific, Measurable, Achievable, Relevant, Timebound) objectives which will help shape the content of our evaluation activities by guiding the broad strategy into actionable tasks.

Developing evidence on Equality, Diversity and Inclusion (ED&I)
As a public sector body, NPL and the other NMS labs are duty bound by the Public Sector Equality Duty (PSED) to have due regard to the need to eliminate discrimination, advance equality of opportunity and foster good relations between different people when carrying out its duties. Currently, NPL has strived to meet this duty by publishing an annual gender pay gap report, whilst also being a Juno Practitioner, a Stonewall Diversity champion, and a member of the Business Disability Forum.[20] Monitoring and evaluating the demographics of our staff and our users is crucial to ensuring that the NMS plays its part in nurturing an equal, diverse and inclusive society. NPL has demonstrated its continued commitment to this challenge by professionalising these activities through the recent hire of a Diversity and Inclusion Human Resources Business Partner. In the forthcoming year, the Analysis and Evaluation team plans to broaden our activities in this area by piloting an NMS applicant demographic survey.

Strengthening the stock of qualitative evidence
In order to have a holistic portfolio of evidence, there is a need to reinforce the qualitative evidence in all themes. As mentioned in the document, the role of qualitative evidence is essential to having a convincing narrative of why certain projects are necessary and how the benefits are generated.

20 NPL (2020). Gender Pay Gap Report. (link).

Although existing A4I cases studies are a valuable resource in explaining how the impact of the NMS occurs, these studies revolve around the innovation activities of SMEs. Thus, a comprehensive approach to qualitative evidence that makes use of generally accepted methods in the field of social research is required. This stream of work would strengthen the case of the fundamental role that the NMS plays in the UK's economy. For this reason, NPL's Analysis and Evaluation team plans to give special importance to this type of evaluation in the remainder of 2021 and throughout 2022.

NPL will be progressively adopting the qualitative research methods that are currently being trialled for the analysis of the NTC. The plan is to produce studies that either provide analysis of user needs in markets for which there is little existing evidence, or, uncover unique insights that explore the driving forces behind the observable outcomes of NMS services. Such studies involve conducting structured or semi-structured interviews of organisations who benefit from the NMS, guided by psychological models designed to help interviewees elaborate on their company's needs. This will enable a deeper understanding of how and why companies adopt NMS services, and produce a thematic analysis of the interview responses. The subjects that these studies may focus on include the following: representation of NMS sectors; NMS collaboration with public sector organisations (e.g. NHS); and contribution of the NMS labs to tackling socio-economic objectives such as the net-zero carbon economy.

Fan-out analysis

There are also certain key impact areas which remain uncovered by our existing evidence base. Particularly, the need for an analysis of the 'fan-out' is a major piece of evidence that requires further development. The fan-out of measurement extends the reach of the NMS to thousands of organisations across the UK and, thus, it probably constitutes a significant fraction of the impact generated by the NMS labs.

One of the ways in which the Analysis and Evaluation team plans to improve the capture of evidence for this fan-out is by revising the NMS indicators to expand their scope and deepen their detail, expected to be finalised in the forthcoming year. The NMS indicators will generate metrics to help monitor the success of NMS activities against the goals of the UK Measurement Strategy, including new metrics that target data on the composition of the user-base and measurable contributions towards supporting socio-economic objectives. This work is being crafted to align closely with the activities and challenges outlined in the new UK Measurement Strategy, which will allow for more coherent evaluation activities in the future.

In addition, plans to revamp the NMS customer survey will further bolster the stock of NMS evidence on the fan-out. Based on lessons learnt from its previous outing, the survey framework will be redesigned to create different surveys for each segment of the NMS user-base (e.g. academia, healthcare, industry, etc.). It is imperative that, from the outset, a clear set of research questions are agreed upon for each survey, and the statistical methods that will be used to analyse the collected data are scrutinised and confirmed before the survey takes place. This year, to avoid picking up the effects of the Covid-19 lockdown on business operations, the NMS customer survey will be delayed until 2023. In the elapsed time, the Analysis and Evaluation team plans to run a pilot study which will serve to ensure the quality of the final survey, allowing for any necessary adjustments to be made to the methodology and survey structure based on feedback received from the pilot.